BACTERIAL NUTRIENT STRESS AND DRUG EFFLUX

NEW UNDERSTANDING FOR NEW ANTIBIOTICS: BACTERIAL NUTRIENT STRESS AND DRUG EFFLUX

By SARA S. EL ZAHED, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the

Requirements for the Degree of Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2020) McMaster University

(Biochemistry and Biomedical Sciences) Hamilton, Ontario

TITLE: New understanding for new antibiotics: bacterial nutrient stress and drug efflux

AUTHOR: Sara S. El Zahed, B.Sc.

SUPERVISOR: Eric D. Brown, Ph.D.

NUMBER OF PAGES: xiii, 137

Abstract

The dramatic decline in the discovery and development of novel antibiotics has been met with an exponential increase in antibiotic-resistant pathogenic bacteria. Calls for new chemical matter have been paradoxically answered with the withdrawal of multinational pharmaceutical companies from the field of antibiotic discovery and development. Nevertheless, scientific challenges have also been a major contributor to a lean antibiotic pipeline. Renewed efforts to gain a deeper understanding of the bacterial physiology, which governs growth and survival, are urgently needed. To this end, I have examined two important avenues with relevance to the field of new antibiotic discovery using *Escherichia coli* as a model: 1) a systems analysis of the interactions of essential functions under nutrient stress and 2) a physicochemical and structural analysis of small molecules to identify properties that influence Gram-negative activity and efflux susceptibility. I tackled the first aim by systematically combining 45 chemical probes that target essential cellular processes. I revealed a highly connected network of 186 interactions, of which 81 were synergistic and 105 were antagonistic. The network highlighted new connectivity between housekeeping functions and nutrient metabolism. I approached the second aim by screening ~314,000 diverse synthetic compounds for inhibitors of an efflux-deficient *E. coli* strain. I identified about 4,500 actives, of which approximately 84% showed high susceptibility to efflux. Using a machine learning approach, I assessed the physicochemical space occupied by these 4,500 inhibitors and determined that hydrophobic and planar small molecules with low molecular stability exhibited antibacterial activity only in efflux-compromised *E. coli*. Further, compounds

with reduced branching and compactness showed increased susceptibility to efflux. Within this dataset, I also identified some compound series highlighting structural variations that have a large impact on efflux susceptibility. In all, the work provides new insights into an emerging target in antibiotic drug discovery, namely nutrient stress, and uncovers some physicochemical properties and structural motifs that contribute to antibacterial activity and efflux susceptibility of small molecules.

Acknowledgements

I would like to thank my supervisor, Dr. Eric Brown, for providing me with the opportunity to pursue graduate work in his lab. The independent and creative environment pushed me to pursue my scientific curiosity. Thank you for your unwavering support and guidance.

Thank you to my committee members, Dr. Gerry Wright and Dr. Mike Surette, for your guidance, suggestions, and support. Gerry, thank you for sharing your love for chemistry and reminding me to put on my chemistry cap at times during my graduate work.

A special thank you to my dear and lifelong friend, Maya Farha. Thank you for all the laughter, good memories, and support. I am thankful for the bond we share. Always.

To my mom, my role model, and my best friend, Sue, you have inspired and motivated me throughout this journey. Thank you for your endless love and support. Thank you for believing in me. I would not be here without you. To my sister, Maya, and my brother, Mohammad, thank you for always being there for me and for the endless laughter that kept me going. Those are times I will cherish forever.

To my late grandfather, learning that my mom and I share some of your curiosity and love for reading has always motivated me to work hard. Although I missed the chance to share what I have learned with you, I am certain you would have been proud of your little red riding hood. This is for you and my mom.

Table of Contents

List of Figures

Chapter I

Chapter II

Chapter III

Chapter IV

List of Tables

List of Abbreviations

ABC	ATP-binding cassette
ACC	Acetyl-CoA carboxylase
ASA	Accessible surface area
ATP	Adenosine triphosphate
AUC	Area under the curve
clogD	Calculated logD
clogP	Calculated logP
DNA	Deoxyribonucleic acid
dTMP	Deoxythymidine monophosphate
EC_{50}	Half maximal effective concentration
EPI	Efflux pump inhibitor
FIC	Fractional inhibitory concentration
FICI	Fractional inhibitory concentration index
Fsp^3	Fraction of sp^3 hybridized carbon atoms
HTS	High-throughput screen
HWI	Hyper-Wiener index
IF2	Initiation factor 2
IM	Inner membrane
L	Loose
LB	Luria-Bertani
LC-MS/MS	Liquid chromatography tandem mass spectrometry
MATE	Multidrug and toxic compound extrusion
MFS	Major facilitator superfamily
MIC	Minimal inhibitory concentration
MOA	Mechanism of action
MRSA	Methicillin-resistant *S. aureus*
MW	Molecular weight
O	Open
OD	Optical density
OM	Outer membrane
PACE	Proteobacterial antimicrobial compounds efflux
PAβN	Phenylalanine-arginine-β-naphthylamine
PCA	Principal component analysis
PSA	Polar surface area
RNA	Ribonucleic acid
RND	Resistance-nodulation-division

ROC	Receiver operating characteristic
r-proteins	Ribosomal proteins
rRNA	Ribosomal ribonucleic acid
RSC	Resonant structure count
SAM	S-adenosylmethionine
SAR	Structure-activity relationship
SERF	Susceptibility to Efflux Random Forest
SMR	Small multidrug resistance
T	Tight
VRE	Vancomycin-resistant *Enterococci*

CHAPTER I – Introduction

Preface

Some parts of this chapter were adapted from a previously published book chapter:

El Zahed, S. S., Kumar, G., Tong, M., & Brown, E. D. (2018). Nutrient stress small-molecule screening platform for *Escherichia coli*. In B. Wagner (Ed.), *Phenotypic Screening: Methods and Protocols* (Vol. 1787, pp. 1-18). New York, NY: Springer New York.

For this work, I wrote the manuscript with Kumar, G., and Tong, M. provided some input. Edits were provided by Brown, E.D.

THE NEED FOR NEW APPROACHES TO ANTIBIOTIC DISCOVERY

Antibiotic discovery and the rise of antibiotic resistance

The introduction of penicillin to the clinic in the 1940's marked the beginning of the golden age of antibiotic discovery. During this period, many classes of antibiotics were discovered and developed for clinical use, revolutionizing the world of medicine, and extending life expectancy. Within two decades, however, the discovery of new classes of antibiotics dramatically declined, and ultimately the last clinically-useful antibiotic class was discovered in the 1980's. Since the beginning of the 21st century, 44 antibiotics have been introduced to the clinic, of which only five are first-in-class: linezolid (oxazolidinone, 2000), daptomycin (lipopeptide, 2003), retapamulin (pleuromutilin, 2007), fidaxomicin (tiacumicin, 2011), and bedaquiline (diarylquinoline, 2012) (Butler & Paterson, 2020). In parallel, there has been an exponential rise in antibiotic resistance where we have reached a time in which infections caused by superbugs are now common occurrences in the clinic.

Resistance to antibiotics is an evolutionary phenomenon that has existed in nature long before antibiotics were classified as therapeutic agents (D'Costa et al., 2011; Davies & Davies, 2010; Wright, 2007). For millions of years, antibiotic-producing bacteria and fungi have posed an evolutionary pressure on neighbouring bacteria to develop resistance mechanisms in order to survive (Walsh, 2003). Antibiotic resistance in the clinic developed when drug-sensitive bacteria were exposed to antibiotics during infection. The antibiotic-driven evolutionary pressure led to the rise in antibiotic resistance, where some

pathogens became resistant to our last resort antibiotics (Liu et al., 2016; Walsh et al., 2011). This has been attributed to the misuse of antibiotics in the clinic and agriculture, the paucity of new antibacterial agents, and the narrow range of bacterial pathways targeted by currently prescribed antibiotics. While campaigns continue to raise awareness about antibiotic use, the last two factors have been a major challenge in scientific research. Indeed, most of our current antibiotics are derivates of antimicrobials identified between the 1940's and 1960's from soil-derived actinomycetes (Valiquette & Laupland, 2015). This traditional antibiotic discovery platform continues to be optimized for antibiotic discovery; however, the yield for new molecules remains low (Brown & Wright, 2016). Consequently, antibiotic development has shifted to chemical modifications of known antibiotics, which target a limited number of essential bacterial processes, namely the synthesis of cell wall, DNA, RNA, and proteins. This is worrisome since resistance to one molecule in an antibiotic class tends to develop to cross resistance to other antibiotics within the same class (Munita & Arias, 2016; Walsh, 2003).

Nutrient stress and drug efflux

Seeking to expand the target base of antibacterial drug discovery, we have developed a nutrient stress screening platform that identifies growth inhibitors of *E. coli* under nutrient limitation (Zlitni et al., 2013). Under nutrient stress, bacteria require an expanded biosynthetic capacity that includes the synthesis of amino acids, vitamins, and nucleobases. Growing evidence suggests that these processes may be indispensable to certain pathogens and at particular sites of infection (El Zahed et al., 2018). Indeed, more than 100 biosynthetic enzymes become indispensable in *E. coli* grown under nutrient

stress *in vitro*. Additionally, *E. coli* requires some 300 genes that code for basic housekeeping functions, including the synthesis of cell wall, DNA, RNA, and proteins. Inhibitors targeting encoded proteins in this collection of ~400 genes can serve as probes of biology to study the interaction network of nutrient stress functions in *E. coli*. As such, I employed a chemical biology approach where I systematically combined some 45 chemical probes targeting a subset of these proteins and mapped their interactions under nutrient-limited conditions. This work revealed a highly connected network of 186 interactions and highlighted new connectivity between nutrient synthesis and housekeeping functions (El Zahed & Brown, 2018).

For more than two decades, phenotypic antibacterial screening efforts have become a modern approach to discover novel antimicrobials against Gram-negative bacteria. Large chemical collections have been screened for such molecules; however, their yield remains low (Brown & Wright, 2016; Silver, 2011; Tommasi et al., 2015). While the outer membrane barrier has been regarded as the main obstacle for this limitation, active efflux pumps are, in fact, major contributors that expel a diverse set of molecules and restrict their accumulation in bacteria. Our incomplete understanding of molecular descriptors governing efflux susceptibility has been a bottleneck in Gram-negative antibacterial drug discovery programs. In this regard, I have screened some 314,000 small molecules in efflux-compromised *E. coli* in order to identify descriptors that contribute to drug efflux. This work revealed some 4,500 actives where dose response analyses of these actives revealed that 84% were highly susceptibility to efflux. Further, using principal component analysis and a machine learning approach, this work revealed that hydrophobic and planar

small molecules with low molecular stability had antibacterial activity only in efflux-compromised *E. coli*, and compounds with reduced branching and compactness showed increased susceptibility to efflux. To examine the structural underpinnings contributing to efflux, I applied structure-based clustering of the 4,500 actives and identified some side chain moieties that caused marked changes in efflux susceptibility. In all, this novel approach identified key molecular descriptors and structural moieties that contribute to antibacterial activity and efflux.

GENETIC NETWORKS AND NUTRIENT STRESS

Gene dispensability

Since the first complete bacterial genome of *Haemophilus influenzae* was available in 1995, antibiotic discovery and development have been influenced by genomic data. With the advancements in bioinformatics, datasets of whole genome sequences are constantly mined for novel antibacterial targets (Brown & Wright, 2016; Fields et al., 2017). A critical aspect of antimicrobial compounds is that their targets need to code for functions required for bacterial growth. Methodical genome-wide single-gene deletions have been the traditional approach to explore gene dispensability, where nonviable mutants suggest that their corresponding target gene is essential. Much of this work has been inspired by the model for molecular genetics, *Saccharomyces cerevisiae*, where ~19% of its genome is essential for viability (Giaever et al., 2002; Giaever & Nislow, 2014). Construction of a genome-wide single-gene deletion collection of the model Gram-positive bacterium *Bacillus subtilis* (Kobayashi et al., 2003) and the model Gram-negative bacterium

Escherichia coli (Baba et al., 2006) soon followed suit. Compared to *S. cerevisiae*, approximately 7% of the genome in *B. subtilis* (Kobayashi et al., 2003) and *E. coli* (Baba et al., 2006) are essential. Further, transposon mutagenesis has facilitated genome-scale approaches to systematically probe gene dispensability in other organisms (Fields et al., 2017; Wetmore et al., 2015), including the pathogens *Pseudomonas aeruginosa* (Jacobs et al., 2003; Liberati et al., 2006), *Staphylococcus aureus* (Fey et al., 2013), and *Klebsiella pneumoniae* (Ramage et al., 2017).

Gene essentiality in *Escherichia coli*

A key factor in exploring gene dispensability is the environmental context of bacteria, which have evolved to survive and grow under different conditions. This aspect is important for identifying antibiotic targets since standard laboratory growth conditions may exclude some genes that are essential for bacterial growth during infection (Brown & Wright, 2016). Specifically, when the model Gram-negative bacterium *E. coli* is grown under standard nutrient-rich conditions, only 303 of ~4,400 genes in its genome are essential for growth (Baba et al., 2006). These genes code for basic housekeeping functions, such as the synthesis of cell wall, DNA, RNA, and proteins. When the growth condition is limited to a source of carbon, nitrogen, essential phosphates, and salts, *E. coli* shifts its metabolic activities to include the synthesis of essential amino acids, vitamins, nucleobases and other cofactors (El Zahed et al., 2018; Joyce et al., 2006). These functions are encoded by an additional set of 119 nutrient stress genes, which increase the total number of essential genes for viability to 422 (Figure 1). In fact, a nutrient-limited medium may provide a better proxy for the host environment. There have been many

reports of impaired growth and attenuated virulence in pathogens due to mutations in vitamin, nucleobase, and amino acid biosynthetic genes (Bange et al., 1996; Cersini et al., 1998; Chamberlain et al., 1993; Cuccui et al., 2007; Hoiseth & Stocker, 1981; Mei et al., 1997; Samant et al., 2008). From an antibiotic discovery perspective, compounds that target bacteria under nutrient-limited conditions could serve as leads for novel antibacterial drugs. As such, assessing gene essentiality in growth conditions that mimic the host environment during infection provides a more comprehensive list of potential antibacterial targets.

Genetic interaction networks

Genetic interaction networks identify functional and mechanistic relationships between genes and their corresponding pathways that can be harnessed for drug discovery. For example, using a synthetic lethality approach, Côté *et al.* have studied the interactions of some 82 nutrient stress genes (of 119) with the dispensable gene set in *E. coli* under nutrient-rich conditions. In that work, strains bearing deletions in 82 nutrient stress genes were systematically crossed with ~4,000 single-gene deletion mutants to care strains with pairs of deletions and to identify synthetic sick or lethal interactions. That work identified a dense interaction network that mapped a total of 1,881 synthetic sick interactions, including signature interactions between nutrient acquisition and biosynthesis (Côté et al., 2016). This vital connection revealed that some of the 119 nutrient stress genes that code for nutrient biosynthesis become essential in nutrient-rich conditions if the transporter of the specific nutrient is absent. Indeed, this connectivity inspired a target-specific screen for inhibitors of biotin biosynthesis (Gehrke et al., 2017), which has been validated as a

promising antimicrobial target in *Mycobacterium tuberculosis* (Woong Park et al., 2011). As such, synthetic sick interactions generated from gene-gene interaction networks can be harnessed for antimicrobial drug discovery.

However, the genetic interaction network charted by Côté *et al.* was collected in nutrient-rich conditions, where nutrient stress genes are dispensable. Therefore, the dataset lacks insight into the genetic network under nutrient-limited conditions where nutrient stress genes have an essential phenotype. The study also did not probe the collection of 303 genes that are essential under nutrient-rich conditions since these genes have an essential phenotype regardless of the growth condition. Indeed, mutagenesis has been the conventional approach to map genome-wide gene-gene interactions; however, essential genes have largely resisted these approaches. Only in *S. cerevisiae*, in which temperature sensitive alleles have been created in both essential and dispensable genes, has genetic perturbation been used to map the interactions of essential genes (Costanzo et al., 2016). That study revealed that, on average, essential genes have a higher number of interactions than those that are dispensable (Costanzo et al., 2016), suggesting that essential genes act as hubs in the molecular network of organisms. To overcome the limitations of genetic perturbation, chemical biology approaches have been employed to map uncharted interactions of functions encoded by essential genes.

Chemical biology

Chemical combinations have been used to probe biological systems in order to identify fundamental interactions and functional relationships in bacterial physiology (Lehár et al.,

2007; Yeh & Kishony, 2007). Much like genetic or epistatic interactions among genetic mutations, chemical combinations can be classified into three types of interactions: additive, synergistic, or antagonistic. An additive interaction describes one where the phenotypic effect is no greater than that expected from the combination of non-interacting chemicals. A synergistic interaction occurs when the effect of both chemicals is enhanced relative to the additive effect, while an antagonistic interaction describes one that is suppressed. Thus, compounds that act synergistically or antagonistically are thought to reveal connectivity between the functions or pathways being perturbed. Additivity, on the other hand, identifies those that are functionally non-redundant or unconnected. Classically, antibiotic combinations have been used to chart the connectivity between housekeeping functions encoded by essential genes. For instance, Yeh *et al.* combined some 21 antibiotics of diverse modes of action and generated a chemical-chemical interaction network that revealed distinct connectivity between some of these functions (Yeh et al., 2006). This study has shown that most drugs cluster based on their target class, which is consistent with functional connectivity (Yeh et al., 2006). As such, chemical interaction networks can be employed to identify connectivities among basic housekeeping processes required for bacterial growth.

Notably, the synergistic interaction of the antibiotics sulfamethoxazole and trimethoprim is a signature chemical combination highlighting functional connections that underpin bacterial metabolism. Both antibiotics inhibit key enzymes involved in folate biosynthesis. Sulfamethoxazole targets dihydropteroate synthase (FolP), an enzyme responsible for the catalysis of an early condensation step in the biosynthetic pathway and

is essential for growth only under nutrient-limited conditions (Keseler et al., 2016). Trimethoprim inhibits dihydrofolate reductase (FolA), an enzyme that is essential under nutrient-rich conditions and catalyzes the ultimate reduction step of dihydrofolate to tetrahydrofolate (Figure 2). Although the chemical combination targets sequential steps in folate biosynthesis, this inhibition cascades into a much more complex metabolic effect in bacteria. Indeed, tetrahydrofolate is required for the synthesis of glycine, methionine, purines, and thymidine triphosphate (Figure 2), which are required for the synthesis of DNA, RNA, and proteins. Interestingly, metabolomic analyses of antifolate-treated *E. coli* cells grown in different culture media revealed distinct connectivities between the synthesis of folate and some of these nutrients. Particularly, under nutrient-rich conditions, depletion of tetrahydrofolate pools in *E. coli* induces thymine starvation, which leads to DNA damage and growth inhibition (Kwon et al., 2010; Kwon et al., 2008). Under nutrient-limited growth conditions, however, folate deprivation initially results in glycine depletion, followed by reduced purine pools and, eventually, growth inhibition (Kwon et al., 2010; Minato et al., 2018). Furthermore, some chemical-chemical combination studies have also shown connectivity between folate synthesis, nucleotide homeostasis, and DNA replication in nutrient-limited growth conditions (El Zahed & Brown, 2018; Wambaugh et al., 2017; Yeh et al., 2006). In all, studies of the combination of trimethoprim and sulfamethoxazole are illustrative of the insights that are possible into the functional relationships between enzymes in nutrient metabolism and housekeeping functions.

Charting the connectivity between cellular pathways under nutrient-limited conditions reveals a complex network that underpins bacterial growth and survival. Chemical combinations can be used as tools to study functional relationships between molecular targets and their corresponding pathways. Although powerful, this approach is limited by the number of well-characterized molecules that can be used to probe housekeeping functions and nutrient metabolism. Recently, phenotypic screens have identified some probes targeting nutrient metabolism in Gram-negative bacteria (Fahnoe et al., 2012; Zlitni et al., 2013), which can be exploited to chart new connectivity in nutrient stress.

MULTIDRUG EFFLUX BY GRAM-NEGATIVE BACTERIA

An overview of the Gram-negative cell envelope

The cytoplasmic membrane of Gram-negative bacteria is enclosed by an inner membrane (IM) and an outer membrane (OM). The former is composed of a phospholipid bilayer, while the latter is an asymmetric lipid bilayer consisting of a lipopolysaccharide outer leaflet and a phospholipid inner leaflet (Bos et al., 2007). Separating these two membranes is the periplasmic space which contains a thin peptidoglycan layer. Additionally, Gram-negative bacteria are decorated with a plethora of efflux pumps, generally classified into six types: the small multidrug resistance (SMR) family, the multidrug and toxic compound extrusion (MATE) family, the proteobacterial antimicrobial compounds efflux (PACE) family, the ATP-binding cassette (ABC) transporter superfamily, the major facilitator superfamily (MFS), and the resistance-nodulation-division (RND) superfamily (Xian-Zhi Li, 2016). While the OM acts as a

robust barrier to many molecules, those that can penetrate it are subject to extrusion by efflux pumps. Most efflux pumps transport a relatively related range of substrates; however, the RND superfamily recognizes and extrudes a diverse range of substrates. Due to these pumps, Gram-negative bacteria are intrinsically resistant to many antibiotics. Indeed, overexpression of RND-type pumps has been a common resistance mechanism detected in multidrug resistant bacteria (Alcock et al., 2019; Shigemura et al., 2015; Yasufuku et al., 2011). As such, much work has been done to characterize these efflux pumps as well as to identify their range of substrates.

The AcrAB-TolC efflux pump in *Escherichia coli*

RND-type efflux pumps are multiple-component systems composed of an inner membrane transporter, a periplasmic adaptor protein, and an outer membrane channel. This tripartite composition allows for the extrusion of a broad range of compounds from the periplasmic space and outer leaflet of the inner membrane to the extracellular milieu (Du et al., 2014). Among Gram-negative bacteria, AcrAB-TolC in *E. coli* is the best characterized RND-type efflux pump. Specifically, AcrB is an inner membrane transporter driven by the proton-motive force, AcrA is a periplasmic membrane fusion protein, and TolC is an outer membrane channel (Zgurskaya et al., 2015). Structural analysis revealed a 3:6:3 compositional stoichiometry consisting of an AcrB trimer, an AcrA hexamer, and a TolC trimer (Du et al., 2014; Tikhonova et al., 2011; Xu et al., 2011). Particularly, substrates bind to AcrB and are extruded through the TolC channel with the help of AcrA (Figure 3). Asymmetric crystal structures of AcrB revealed that each protomer cycles between three conformational stages which correlate with the

different stages of substrate entry, binding, and extrusion (Figure 3) (Anes et al., 2015; Eicher et al., 2012; Li et al., 2015; Murakami et al., 2006; Yamaguchi et al., 2015). Specifically, a substrate enters either the cleft channel, vestibule channel, or central cavity channel of the monomer in a loose (L) conformation. Once one of these channels is occupied, the monomer transitions to a tight (T) conformation, moving the substrate into the distal binding pocket. Co-crystallization work of AcrB with some substrates revealed that the distal binding pocket is large and comprised of two distinct binding sites. The lower portion of the pocket, coined the hydrophobic trap, is rich in phenylalanine residues (Bohnert et al., 2008; Murakami et al., 2006; Nakashima et al., 2013), while the upper "crevice" area is rich in hydrophilic and charged residues (Li et al., 2015; Takatsuka et al., 2010; Vargiu & Nikaido, 2012). These features of the distal binding pocket in AcrB highlight the ability of RND-type pumps to accommodate and expel a broad range of substrates. Finally, substrates bound to the distal pocket induce the T protomer to transition to an open (O) conformation (Eicher et al., 2012; Li et al., 2015; Murakami et al., 2006). This collapses the binding pocket and squeezes the substrate into the central funnel of AcrB toward TolC for subsequent extrusion from the cell (Anes et al., 2015; Eicher et al., 2012). Overall, these studies show how the composition of AcrAB-TolC and its binding pocket pose a major problem in Gram-negative antibiotic discovery. As such, two approaches have been largely underway in order to overcome the intrinsic resistance conferred by RND-type efflux pumps: identifying potent efflux pump inhibitors (EPIs) and defining the physicochemical properties for drug permeation and efflux evasion.

Efflux pump inhibitors

Phenylalanine-arginine-β-naphthylamine (PAβN) was the first EPI discovered to inhibit the inner membrane transporter of RND-type efflux pumps in *E. coli* and *P. aeruginosa* (Lomovskaya et al., 2001). Since, there have been four main classes of EPIs that bind to the inner membrane transporter of such pumps and inhibit efflux: peptidomimetic compounds derived from PAβN, aryl-piperazines, pyridopyrimidines, and pyranopyridines (Li et al., 2015; Opperman & Nguyen, 2015; Xian-Zhi Li, 2016). Further, a novel class of EPIs that binds to the periplasmic adaptor protein AcrA of *E. coli* has been recently discovered; however, the exact mechanism of inhibition remains unknown (Abdali et al., 2017). Indeed, EPIs could be used in combination therapies as adjuvants that block efflux pumps and restore antibiotic activity in resistant bacteria. These inhibitors could broaden the activity spectrum of some Gram-positive antibiotics, making them effective in Gram-negative bacteria. Linezolid, for example, is a potent oxazolidinone antibiotic that inhibits the growth of Gram-positive bacteria by binding to the 50S ribosomal subunit and disrupting translation (Clemett & Markham, 2000). As previously mentioned, it is a first-in-class antibiotic introduced to the clinic in the early 2000's and has been used to treat infections caused by antibiotic-resistant pathogens such as methicillin-resistant *S. aureus* (MRSA), penicillin-resistant *Streptococcus pneumoniae* and vancomycin-resistant *Enterococci* (VRE) (Clemett & Markham, 2000; Livermore, 2003). Indeed, Gram-negative bacteria are intrinsically resistant to linezolid due to RND-type efflux pumps (Hung et al., 2013). In combination with some EPIs, on the other hand, linezolid gains antimicrobial activity in some Gram-negative bacteria (Schuster et al.,

2014). This example along with others (Li et al., 2015; Lomovskaya et al., 2001; Opperman & Nguyen, 2015) have demonstrated the potential of EPIs in blocking efflux pumps and potentiating antibiotics. To date, only the PAβN analog MP-601205 has advanced to phase I clinical trials; however, it has since been abandoned due to toxicity concerns (Lomovskaya & Bostian, 2006; Opperman & Nguyen, 2015). Nevertheless, early stages of lead optimization for the pyranopyridine EPI MBX2319 as a potential antibiotic adjuvant therapy are ongoing (Nguyen et al., 2015; Opperman et al., 2014; Sjuts et al., 2016). In all, EPIs could broaden the antibacterial activity of narrow spectrum antibiotics as well as make "old" antibiotics effective again. Cytotoxicity concerns, however, have restricted EPIs from advancing to the clinical stage of testing.

Physicochemical properties of antibiotics

Defining an "ideal" physicochemical space for molecules with Gram-negative activity has long been sought after. This stems from the fact that such molecules would have good penetration across the cell envelope and low susceptibility to efflux. Such efforts have been largely influenced by Lipinski's success at correlating the physicochemical properties of drugs with oral bioactivity. Lipinski's rule of five have outlined that orally available drugs have less than 5 hydrogen-bond donors and 10 hydrogen-bond acceptors, a molecular weight (MW) less than 500 g mol^{-1}, and a calculated logP no greater than 5 (Lipinski et al., 2001). Since these rules were established in the late 1990's, chemists have always considered them when designing compound libraries for high-throughput screens (HTS). Lipinski noted, however, that orally available antibiotics violate the rule of five due to their larger size (higher MW) and increased hydrophilicity (more polar).

Approximately a decade since the rule of five was established, O'Shea and Moser analyzed the physicochemical properties of some 147 antibacterial agents and compared them to non-antibiotic drugs (O'Shea & Moser, 2008). On average, Gram-positive antibiotics were larger (813 g mol^{-1}) than Gram-negative antibiotics (414 g mol^{-1}) and non-antibiotic drugs (338 g mol^{-1}) (O'Shea & Moser, 2008). Comparing their lipophilicity, however, Gram-negative antibiotics were more hydrophilic (clogD -2.8) than both Gram-positive antibiotics (clogD -0.2) and non-antibiotic drugs (clogD 1.6) (O'Shea & Moser, 2008). When Gram-negative antibiotics were further classified based on their target location, Moser noted that those with a periplasmic target are larger (MW 347-558 g mol^{-1}) and more hydrophilic (clogD -5.1 to -1) than those with a cytoplasmic target (MW 254-465 g mol^{-1}; -1.4 to 1.1) (Reck et al., 2019). However, the collection of antibiotics in both studies includes some that are susceptible to RND-type efflux pumps. As such, these studies propose ideal physicochemical properties for Gram-negative antimicrobials but lack insight into the properties that impact efflux susceptibility.

Physicochemical properties as guidelines for compound entry and efflux

A comprehensive analysis of some HTS campaigns at AstraZeneca revealed that molecules least susceptible to RND-type efflux pumps in *E. coli* and *P. aeruginosa* were small (MW <300 g mol^{-1}) and hydrophilic (clogD <0) or very large (MW >700 g mol^{-1}) and zwitterionic (Brown et al., 2014). Paradoxically, their hit-to-lead programs revealed that biochemical potency improved with increasing hydrophobicity, which antagonized whole-cell activity (Brown et al., 2014). This contrast in properties for efflux evasion and biochemical potency highlights the scientific challenge in Gram-negative antibiotic

discovery in the 21st century. Furthermore, a recent study that measured compound accumulation in Gram-negative bacteria challenged prior retrospective analyses (O'Shea & Moser, 2008), wherein compounds with a hydrophobic character and positive charge showed good accumulation in these bacteria (Richter et al., 2017). As such, pharmaceutical industry pundits posit that physicochemical properties offer guidelines for optimal Gram-negative penetration and efflux evasion rather than overarching rules (Tommasi et al., 2018). Most importantly, they note that these guidelines are likely to differ across chemical class as well as Gram-negative species (Silver, 2016; Tommasi et al., 2018).

Research objectives and organization of thesis

The objectives of the research in this thesis were to exploit nutrient stress in order to 1) map uncharted interactions between essential functions in *E. coli* and 2) identify the physicochemical and structural parameters contributing to antibacterial activity and efflux susceptibility. The thesis is based on the hypotheses that 1) charting functional relationships between essential processes and 2) gaining a further understanding of efflux susceptibility could be harnessed for antimicrobial discovery. To this end, Chapter 2 describes a chemical biology approach using the model bacterium *E. coli* to identify interactions between housekeeping functions and those required to respond to nutrient stress. This work suggested a highly dense and connected network of functions essential to bacteria under nutrient limitation. Further, I investigated three potent interactions between biotin and fatty acid syntheses, amino acid biosynthesis and ribosome assembly, as well as purine synthesis and translation inhibition. Chapter 3 explores empirical and

computational approaches to determine molecular descriptors that govern antibacterial activity and efflux susceptibility. Using principal component analysis and machine learning, I identified that hydrophobic and planar small molecules with low molecular stability were growth inhibitory only in efflux-compromised *E. coli*, and compounds with reduced branching and compactness showed increased susceptibility to efflux. Further, structure-based clustering and structure-activity relationship analyses revealed some side chain decorations that can render small molecules susceptible to efflux. Finally, Chapter 4 discusses some suggestions for future research, based on the work described herein.

References

Abdali, N., Parks, J. M., Haynes, K. M., Chaney, J. L., Green, A. T., Wolloscheck, D., Walker, J. K., Rybenkov, V. V., Baudry, J., Smith, J. C., & Zgurskaya, H. I. (2017). Reviving antibiotics: efflux pump inhibitors that interact with AcrA, a membrane fusion protein of the AcrAB-TolC multidrug efflux pump. *ACS Infect. Dis., 3*(1), 89-98.

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A. N., Bordeleau, E., Pawlowski, A. C., Zubyk, H. L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G. L., Beiko, R. G., Brinkman, F. S. L., Hsiao, W. W. L., Domselaar, G. V., & McArthur, A. G. (2019). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res., 48*(D1), D517-D525.

Anes, J., McCusker, M. P., Fanning, S., & Martins, M. (2015). The ins and outs of RND efflux pumps in *Escherichia coli*. *Front. Microbiol., 6*(587), 1-14.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol., 2*, 2006.0008.

Bange, F. C., Brown, A. M., & Jacobs, W. R. (1996). Leucine auxotrophy restricts growth of *Mycobacterium bovis* BCG in macrophages. *Infect. Imm., 64*(5), 1794-1799.

Bohnert, J. A., Schuster, S., Seeger, M. A., Fähnrich, E., Pos, K. M., & Kern, W. V. (2008). Site-directed mutagenesis reveals putative substrate binding residues in the *Escherichia coli* RND efflux pump AcrB. *J. Bacteriol., 190*(24), 8225-8229.

Bos, M. P., Robert, V., & Tommassen, J. (2007). Biogenesis of the Gram-negative bacterial outer membrane. *Annu. Rev. Microbiol., 61*(1), 191-214.

Brown, D. G., May-Dracka, T. L., Gagnon, M. M., & Tommasi, R. (2014). Trends and exceptions of physical properties on antibacterial activity for Gram-positive and Gram-negative pathogens. *J. Med. Chem., 57*(23), 10144-10161.

Brown, E. D., & Wright, G. D. (2016). Antibacterial drug discovery in the resistance era. *Nature, 529*(7586), 336-343.

Butler, M. S., & Paterson, D. L. (2020). Antibiotics in the clinical pipeline in October 2019. *J. Antibiot., 73*(6), 329-364.

Cersini, A., Salvia, A. M., & Bernardini, M. L. (1998). Intracellular multiplication and virulence of *Shigella flexneri* auxotrophic mutants. *Infect. Immun., 66*(2), 549-557.

Chamberlain, L. M., Strugnell, R., Dougan, G., Hormaeche, C. E., & de Hormaeche, R. D. (1993). *Neisseria gonorrhoeae* strain MS11 harbouring a mutation in gene *aroA* is attenuated and immunogenic. *Microb. Path., 15*(1), 51-63.

Clemett, D., & Markham, A. (2000). Linezolid. *Drugs, 59*(4), 815-827.

Costanzo, M., VanderSluis, B., Koch, E. N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S. D., Pelechano, V., Styles, E. B., Billmann, M., van Leeuwen, J., van Dyk, N., Lin, Z.-Y., Kuzmin, E., Nelson, J., Piotrowski, J. S., Srikumar, T., Bahr, S., Chen, Y., Deshpande, R., Kurat, C. F., Li, S. C., Li, Z., Usaj, M. M., Okada, H., Pascoe, N., San Luis, B.-J., Sharifpoor, S., Shuteriqi, E., Simpkins, S. W., Snider, J., Suresh, H. G., Tan, Y., Zhu, H., Malod-Dognin, N., Janjic, V., Przulj, N., Troyanskaya, O. G., Stagljar, I., Xia, T., Ohya, Y., Gingras, A.-C., Raught, B., Boutros, M., Steinmetz, L. M., Moore, C. L., Rosebrock, A. P., Caudy, A. A., Myers, C. L., Andrews, B., & Boone, C. (2016). A global genetic interaction network maps a wiring diagram of cellular function. *Science, 353*(6306), aaf1420-1421-1414.

Côté, J.-P., French, S., Gehrke, S. S., MacNair, C. R., Mangat, C. S., Bharat, A., & Brown, E. D. (2016). The genome-wide interaction network of nutrient stress genes in *Escherichia coli*. *mBio, 7*(6), e01714-01716.

Cuccui, J., Easton, A., Chu, K. K., Bancroft, G. J., Oyston, P. C. F., Titball, R. W., & Wren, B. W. (2007). Development of signature-tagged mutagenesis in *Burkholderia pseudomallei* to identify genes important in survival and pathogenesis. *Infect. Immun., 75*(3), 1186-1195.

D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G. B., Poinar, H. N., & Wright, G. D. (2011). Antibiotic resistance is ancient. *Nature, 477*(7365), 457-461.

Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev., 74*(3), 417-433.

Du, D., Wang, Z., James, N. R., Voss, J. E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., & Luisi, B. F. (2014). Structure of the AcrAB-TolC multidrug efflux pump. *Nature, 509*(7501), 512-515.

Eicher, T., Cha, H.-j., Seeger, M. A., Brandstätter, L., El-Delik, J., Bohnert, J. A., Kern, W. V., Verrey, F., Grütter, M. G., Diederichs, K., & Pos, K. M. (2012). Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc. Natl. Acad. Sci. U.S.A., 109*(15), 5687-5692.

El Zahed, S. S., & Brown, E. D. (2018). Chemical-chemical combinations map uncharted interactions in *Escherichia coli* under nutrient stress. *iScience, 2*, 168-181.

El Zahed, S. S., Kumar, G., Tong, M., & Brown, E. D. (2018). Nutrient stress small-molecule screening platform for *Escherichia coli*. In B. Wagner (Ed.), *Phenotypic Screening: Methods and Protocols* (Vol. 1787, pp. 1-18). New York, NY: Springer New York.

Fahnoe, K. C., Flanagan, M. E., Gibson, G., Shanmugasundaram, V., Che, Y., & Tomaras, A. P. (2012). Non-traditional antibacterial screening approaches for the identification of novel inhibitors of the glyoxylate shunt in Gram-negative pathogens. *PLoS ONE, 7*(12), e51732.

Fey, P. D., Endres, J. L., Yajjala, V. K., Widhelm, T. J., Boissy, R. J., Bose, J. L., & Bayles, K. W. (2013). A genetic resource for rapid and comprehensive phenotype screening of nonessential *Staphylococcus aureus* genes. *mBio, 4*(1), e00537-00512.

Fields, F. R., Lee, S. W., & McConnell, M. J. (2017). Using bacterial genomes and essential genes for the development of new antibiotics. *Biochem. Pharmacol., 134*, 74-86.

Gehrke, S. S., Kumar, G., Yokubynas, N. A., Côté, J.-P., Wang, W., French, S., MacNair, C. R., Wright, G. D., & Brown, E. D. (2017). Exploiting the sensitivity of nutrient transporter deletion strains in discovery of natural product antimetabolites. *ACS Infect. Dis., 3*(12), 955-965.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.-D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C.-y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., & Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature, 418*(6896), 387-391.

Giaever, G., & Nislow, C. (2014). The yeast deletion collection: a decade of functional genomics. *Genetics, 197*(2), 451-465.

Hoiseth, S. K., & Stocker, B. A. D. (1981). Aromatic-dependent *Salmonella typhimurium* are non-virulent and effective as live vaccines. *Nature, 291*(5812), 238-239.

Hung, L.-W., Kim, H.-B., Murakami, S., Gupta, G., Kim, C.-Y., & Terwilliger, T. C. (2013). Crystal structure of AcrB complexed with linezolid at 3.5 Å resolution. *J. Struct. Funct. Genomics*, *14*(2), 71-75.

Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M. V., & Manoil, C. (2003). Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A., 100*(24), 14339-14344.

Joyce, A. R., Reed, J. L., White, A., Edwards, R., Osterman, A., Baba, T., Mori, H., Lesely, S. A., Palsson, B. Ø., & Agarwalla, S. (2006). Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. *J. Bacteriol., 188*(23), 8259-8271.

Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., & Karp, P. D. (2016). The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res., 45*(D1), D543-D550.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouillé, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J. F. M. L., Sekiguchi, J., Sekowska, A., Séror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., & Ogasawara, N. (2003). Essential *Bacillus subtilis* genes. *Proc Natl. Acad. Sci. U.S.A., 100*(8), 4678-4683.

Kwon, Y. K., Higgins, M. B., & Rabinowitz, J. D. (2010). Antifolate-induced depletion of intracellular glycine and purines inhibits thymineless death in *E. coli*. *ACS Chem. Biol., 5*(8), 787-795.

Kwon, Y. K., Lu, W., Melamud, E., Khanam, N., Bognar, A., & Rabinowitz, J. D. (2008). A domino effect in antifolate drug action in *Escherichia coli*. *Nat. Chem. Biol., 4*(10), 602-608.

Lehár, J., Zimmermann, G. R., Krueger, A. S., Molnar, R. A., Ledell, J. T., Heilbut, A. M., Short, G. F., Giusti, L. C., Nolan, G. P., Magid, O. A., Lee, M. S., Borisy, A.

A., Stockwell, B. R., & Keith, C. T. (2007). Chemical combination effects predict connectivity in biological systems. *Mol. Syst. Biol., 3*, 80.

Li, X.-Z., Plésiat, P., & Nikaido, H. (2015). The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin. Microbiol. Rev., 28*(2), 337-418.

Liberati, N. T., Urbach, J. M., Miyata, S., Lee, D. G., Drenkard, E., Wu, G., Villanueva, J., Wei, T., & Ausubel, F. M. (2006). An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc. Natl. Acad. Sci. U.S.A., 103*(8), 2833-2838.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev., 46*(1-3), 3-26.

Liu, Y.-Y., Wang, Y., Walsh, T. R., Yi, L.-X., Zhang, R., Spencer, J., Doi, Y., Tian, G., Dong, B., Huang, X., Yu, L.-F., Gu, D., Ren, H., Chen, X., Lv, L., He, D., Zhou, H., Liang, Z., Liu, J.-H., & Shen, J. (2016). Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis., 16*(2), 161-168.

Livermore, D. M. (2003). Linezolid in vitro: mechanism and antibacterial spectrum. *J. Antimicrob. Chemother., 51*(Suppl 2), ii9-ii16.

Lomovskaya, O., & Bostian, K. A. (2006). Practical applications and feasibility of efflux pump inhibitors in the clinic—a vision for applied use. *Biochem. Pharmacol., 71*(7), 910-918.

Lomovskaya, O., Warren, M. S., Lee, A., Galazzo, J., Fronko, R., Lee, M., Blais, J., Cho, D., Chamberland, S., Renau, T., Leger, R., Hecker, S., Watkins, W., Hoshino, K., Ishida, H., & Lee, V. J. (2001). Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: novel agents for combination therapy. *Antimicrob. Agents Chemother., 45*(1), 105-116.

Mei, J.-M., Nourbakhsh, F., Ford, C. W., & Holden, D. W. (1997). Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol. Microbiol., 26*(2), 399-407.

Minato, Y., Dawadi, S., Kordus, S. L., Sivanandam, A., Aldrich, C. C., & Baughn, A. D. (2018). Mutual potentiation drives synergy between trimethoprim and sulfamethoxazole. *Nat. Commun., 9*(1), 1003.

Munita, J. M., & Arias, C. A. (2016). Mechanisms of antibiotic resistance. *Microbiol. Spectr., 4*(2), VMBF-0016-2015.

Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T., & Yamaguchi, A. (2006). Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature, 443*(7108), 173-179.

Nakashima, R., Sakurai, K., Yamasaki, S., Hayashi, K., Nagata, C., Hoshino, K., Onodera, Y., Nishino, K., & Yamaguchi, A. (2013). Structural basis for the inhibition of bacterial multidrug exporters. *Nature, 500*, 102-106.

Nguyen, S. T., Kwasny, S. M., Ding, X., Cardinale, S. C., McCarthy, C. T., Kim, H.-S., Nikaido, H., Peet, N. P., Williams, J. D., Bowlin, T. L., & Opperman, T. J. (2015).

Structure-activity relationships of a novel pyranopyridine series of Gram-negative bacterial efflux pump inhibitors. *Bioorg. Med. Chem., 23*(9), 2024-2034.

O'Shea, R., & Moser, H. E. (2008). Physicochemical properties of antibacterial compounds: implications for drug discovery. *J. Med. Chem., 51*(10), 2871-2878.

Opperman, T., & Nguyen, S. (2015). Recent advances toward a molecular mechanism of efflux pump inhibition. *Front. Microbiol., 6*(421), 1-16.

Opperman, T. J., Kwasny, S. M., Kim, H.-S., Nguyen, S. T., Houseweart, C., D'Souza, S., Walker, G. C., Peet, N. P., Nikaido, H., & Bowlin, T. L. (2014). Characterization of a novel pyranopyridine inhibitor of the AcrAB efflux pump of *Escherichia coli*. *Antimicrob. Agents Chemother., 58*(2), 722-733.

Ramage, B., Erolin, R., Held, K., Gasper, J., Weiss, E., Brittnacher, M., Gallagher, L., & Manoil, C. (2017). Comprehensive arrayed transposon mutant library of *Klebsiella pneumoniae* outbreak strain KPNIH1. *J. Bacteriol., 199*(20), e00352-00317.

Reck, F., Jansen, J. M., & Moser, H. E. (2019). Challenges of antibacterial drug discovery. *Arkivoc*, 227-244.

Richter, M. F., Drown, B. S., Riley, A. P., Garcia, A., Shirai, T., Svec, R. L., & Hergenrother, P. J. (2017). Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature, 545*, 299-304.

Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J. L., Mankin, A. S., & Neyfakh, A. A. (2008). Nucleotide biosynthesis is critical for growth of bacteria in human blood. *PLoS Pathog., 4*(2), e37.

Schuster, S., Kohler, S., Buck, A., Dambacher, C., König, A., Bohnert, J. A., & Kern, W. V. (2014). Random mutagenesis of the multidrug transporter AcrB from *Escherichia coli* for identification of putative target residues of efflux pump inhibitors. *Antimicrob. Agents Chemother., 58*(11), 6870.

Shigemura, K., Osawa, K., Kato, A., Tokimatsu, I., Arakawa, S., Shirakawa, T., & Fujisawa, M. (2015). Association of overexpression of efflux pump genes with antibiotic resistance in *Pseudomonas aeruginosa* strains clinically isolated from urinary tract infection patients. *J. Antibiot., 68*(9), 568-572.

Silver, L. L. (2011). Challenges of antibacterial discovery. *Clin. Microbiol. Rev., 24*(1), 71-109.

Silver, L. L. (2016). A gestalt approach to Gram-negative entry. *Bioorg. Med. Chem., 24*(24), 6379-6389.

Sjuts, H., Vargiu, A. V., Kwasny, S. M., Nguyen, S. T., Kim, H.-S., Ding, X., Ornik, A. R., Ruggerone, P., Bowlin, T. L., Nikaido, H., Pos, K. M., & Opperman, T. J. (2016). Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proc. Natl. Acad. Sci. U.S.A., 113*(13), 3509-3514.

Takatsuka, Y., Chen, C., & Nikaido, H. (2010). Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A., 107*(15), 6559-6565.

Tikhonova, Elena B., Yamada, Y., & Zgurskaya, Helen I. (2011). Sequential mechanism of assembly of multidrug efflux pump AcrAB-TolC. *Chem. Biol., 18*(4), 454-463.

Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I., & Miller, A. A. (2015). ESKAPEing the labyrinth of antibacterial discovery. *Nat. Rev. Drug Discov., 14*, 529-542.

Tommasi, R., Iyer, R., & Miller, A. A. (2018). Antibacterial drug discovery: some assembly required. *ACS Infect. Dis., 4*(5), 686-695.

Valiquette, L., & Laupland, K. B. (2015). Digging for new solutions. *Can. J. Infect. Dis. Med. Microbiol., 26*(6), 289-290.

Vargiu, A. V., & Nikaido, H. (2012). Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A., 109*(50), 20637-20642.

Walsh, C. (2003). *Antibiotics*: American Society of Microbiology.

Walsh, T. R., Weeks, J., Livermore, D. M., & Toleman, M. A. (2011). Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *Lancet Infect. Dis., 11*(5), 355-362.

Wambaugh, M. A., Shakya, V. P. S., Lewis, A. J., Mulvey, M. A., & Brown, J. C. S. (2017). High-throughput identification and rational design of synergistic small-molecule pairs for combating and bypassing antibiotic resistance. *PLoS Biol., 15*(6), e2001644.

Wetmore, K. M., Price, M. N., Waters, R. J., Lamson, J. S., He, J., Hoover, C. A., Blow, M. J., Bristow, J., Butland, G., Arkin, A. P., & Deutschbauer, A. (2015). Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *mBio, 6*(3), e00306-00315.

Woong Park, S., Klotzsche, M., Wilson, D. J., Boshoff, H. I., Eoh, H., Manjunatha, U., Blumenthal, A., Rhee, K., Barry, C. E., III, Aldrich, C. C., Ehrt, S., & Schnappinger, D. (2011). Evaluating the sensitivity of *Mycobacterium tuberculosis* to biotin deprivation using regulated gene expression. *PLoS Pathog., 7*(9), e1002264.

Wright, G. D. (2007). The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Micro., 5*(3), 175-186.

Xian-Zhi Li, C. A. E., Helen I. Zgurskaya. (2016). *Efflux-mediated antimicrobial resistance in bacteria: mechanisms, regulation and clinical implications* (1 ed.): Springer.

Xu, Y., Lee, M., Moeller, A., Song, S., Yoon, B.-Y., Kim, H.-M., Jun, S.-Y., Lee, K., & Ha, N.-C. (2011). Funnel-like hexameric assembly of the periplasmic adapter protein in the tripartite multidrug efflux pump in Gram-negative bacteria. *J. Biol. Chem., 286*(20), 17910-17920.

Yamaguchi, A., Nakashima, R., & Sakurai, K. (2015). Structural basis of RND-type multidrug exporters. *Front. Microbiol., 6*(327), 1-19.

Yasufuku, T., Shigemura, K., Shirakawa, T., Matsumoto, M., Nakano, Y., Tanaka, K., Arakawa, S., Kinoshita, S., Kawabata, M., & Fujisawa, M. (2011). Correlation of overexpression of efflux pump genes with antibiotic resistance in *Escherichia coli* strains clinically isolated from urinary tract infection patients. *J. Clin. Microbiol., 49*(1), 189-194.

Yeh, P., & Kishony, R. (2007). Networks from drug–drug surfaces. *Mol. Syst. Biol.*, *3*, 85.

Yeh, P., Tschumi, A. I., & Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. *Nature Genet.*, *38*(4), 489-494.

Zgurskaya, H. I., Löpez, C. A., & Gnanakaran, S. (2015). Permeability barrier of Gram-negative cell envelopes and approaches to bypass it. *ACS Infect. Dis.*, *1*(11), 512-522.

Zlitni, S., Ferruccio, L. F., & Brown, E. D. (2013). Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. *Nat. Chem. Biol.*, *9*(12), 796-804.

Figure Legends

Figure 1. *E. coli* gene essentiality. Shown is a schematic of gene essentiality in *E. coli* grown in nutrient-rich or nutrient-limited conditions. Dark blue represents the 4,400 genes that comprise the *E. coli* genome. Grey represents 303 genes (of 4,400) essential for viability in nutrient-rich conditions. Light blue represents 422 genes essential for growth in nutrient-limited conditions.

Figure 2. Synthesis and consumption of tetrahydrofolate in *E. coli*. Folate biosynthesis is shown on the left, starting with the condensation reaction catalyzed by FolP, followed by the addition of the glutamyl residue mediated by FolC to form 7,8-dihydrofolate, and its reduction by FolA to produce tetrahydrofolate. Tetrahydrofolate consumption is represented in dashed arrows in grey. GlyA and the gcv system (Lpd, GcvP, GcvH, GcvT) catalyze the reversible methylation of tetrahydrofolate to 5,10-methylenetetrahydrofolate. The former enzyme also produces L-glycine. 5,10-methylenetetrahydrofolate is then demethylated by ThyA to form deoxythymidine monophosphate (dTMP), reduced by MetF for the synthesis of L-methionine, or oxidized and hydrolyzed by FolD for the synthesis of purines. Shown in red are the antibiotics sulfamethoxazole and trimethoprim which inhibit FolP and FolA, respectively. Adapted from Keseler *et al.*, 2016.

Figure 3. Rotation mechanism of substrate access, binding, and extrusion mediated by AcrAB-TolC in *E. coli*. The top and bottom panels show side and horizontal views of a schematic model of AcrAB-TolC in *E. coli*, respectively. TolC is represented by black

and off-white and AcrA is represented by light blue. The three AcrB protomers are represented in green (access; loose, L, protomer), violet (binding; tight, T, protomer), and dark blue (extrusion; open, O, protomer). Substrates are represented as yellow diamonds. Arrows indicate conformational cycling of AcrB for substrate extrusion through TolC. Adapted from Yamaguchi *et al.*, 2015.

Figures

Figure 1

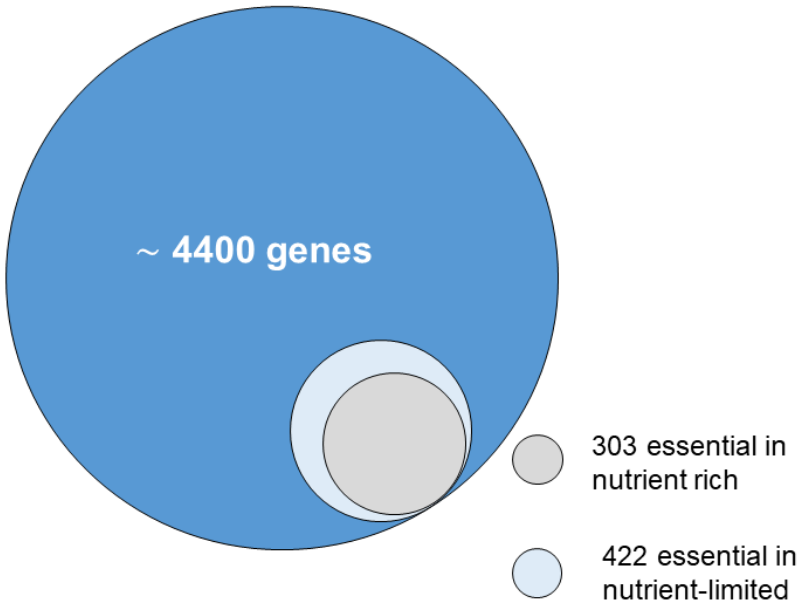


Figure 2

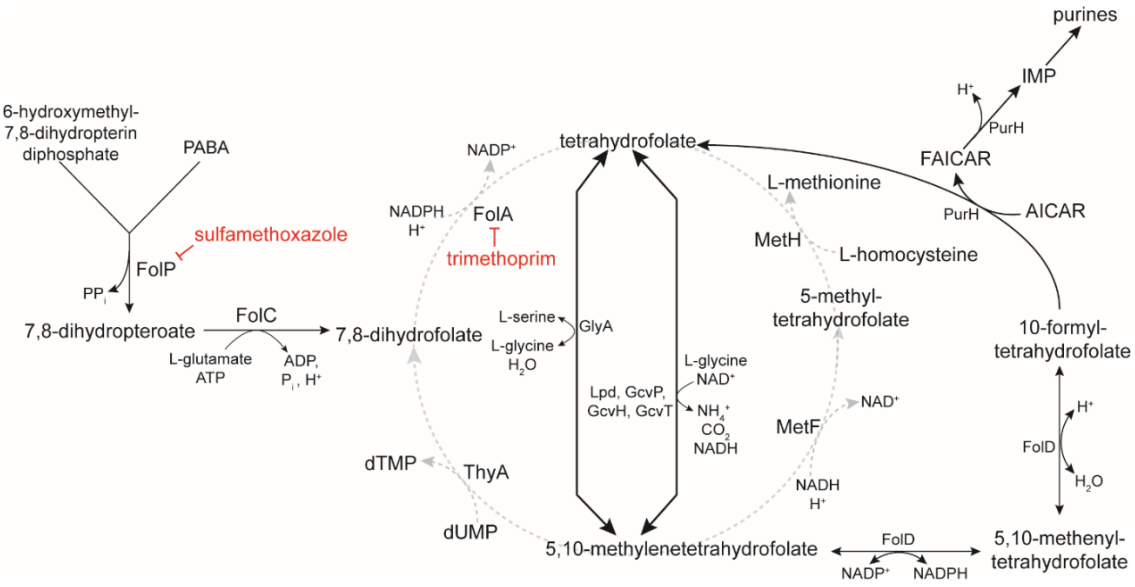
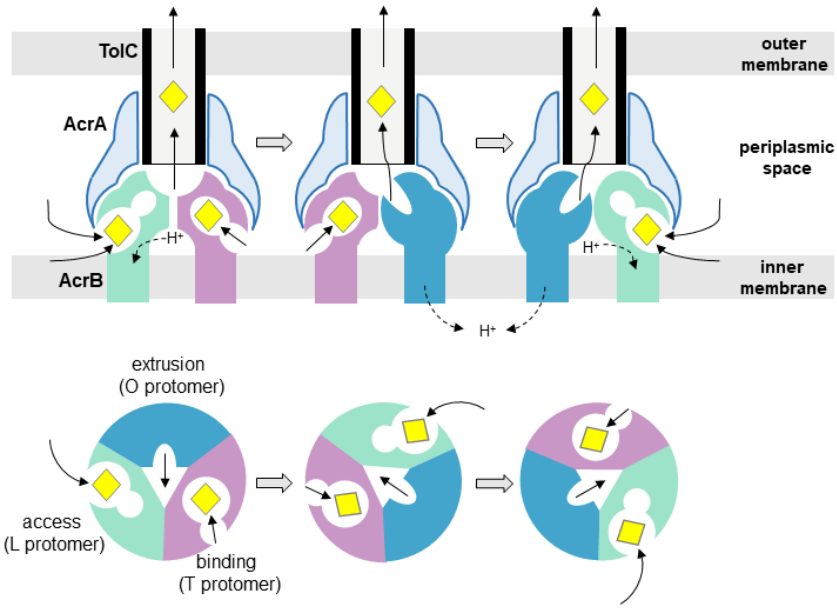

6-hydroxymethyl-
7,8-dihydropterin
diphosphate
PABA
sulfamethoxazole
FolP
PPi
7,8-dihydropteroate
FolC
L-glutamate
ATP
ADP,
Pi, H+
7,8-dihydrofolate
NADP+
NADPH
H+
FolA
trimethoprim
tetrahydrofolate
L-serine
L-glycine
H2O
GlyA
Lpd, GcvP,
GcvH, GcvT
L-glycine
NAD+
NH4+
CO2
NADH
dTMP
ThyA
dUMP
5,10-methylenetetrahydrofolate
L-methionine
MetH
L-homocysteine
5-methyl-
tetrahydrofolate
MetF
NAD+
NADH
H+
FAICAR
PurH
H+
IMP
purines
PurH
AICAR
10-formyl-
tetrahydrofolate
FolD
H+
H2O
5,10-methenyl-
tetrahydrofolate
FolD
NADP+
NADPH

Figure 3


TolC
AcrA
AcrB
outer membrane
periplasmic space
inner membrane
H+
H+
H+
extrusion
(O protomer)
access
(L protomer)
binding
(T protomer)

CHAPTER II – Chemical-chemical combinations map uncharted interactions in *Escherichia coli* under nutrient stress

Preface

The work presented in this chapter was previously published in:

El Zahed, S. S., & Brown, E. D. (2018). Chemical-chemical combinations map uncharted interactions in *Escherichia coli* under nutrient stress. *iScience, 2,* 168-181. doi:10.1016/j.isci.2018.03.018

For this work, I performed all of the experiments and wrote the manuscript, with edits provided by Brown, E.D.

Summary

Of the ~4,400 genes that constitute *Escherichia coli*'s genome, ~300 genes are indispensable for its growth in nutrient-rich conditions. These encode housekeeping functions including cell wall, DNA, RNA and protein syntheses. Under conditions where nutrients are limited to a carbon source, nitrogen source, essential phosphates, and salts, more than 100 additional genes become essential. These largely code for the synthesis of amino acids, vitamins, and nucleobases. While much is known about this collection of ~400 genes, their interactions under nutrient stress are uncharted. Using a chemical biology approach, we focused on 45 chemical probes targeting encoded proteins in this collection, and mapped their interactions under nutrient-limited conditions. Encompassing 990 unique pairwise chemical combinations, we revealed a highly-connected network of 186 interactions, where 81 were synergistic and 105 were antagonistic. The network revealed signature interactions for each probe, and highlighted new connectivity between housekeeping functions and those essential in nutrient stress.

Introduction

Systematic analyses of genome-scale gene deletion collections have characterized the basic housekeeping functions for bacterial survival. Such studies in the model bacterium *Escherichia coli*, for example, have shown that 303 genes have an essential phenotype for growth in nutrient-rich conditions while the balance, approximately 4,000 genes, are dispensable (Baba et al., 2006). In nutrient-limited conditions, where *E. coli* is grown in a medium containing glucose and ammonium chloride (carbon and nitrogen sources, respectively) as well as essential phosphates and salts, a total of 422 genes have an indispensable phenotype (Baba et al., 2006; Joyce et al., 2006). This additional set of 119 genes largely encodes proteins required for the synthesis of amino acids, vitamins, nucleobases, and other cofactors. While much is known about these housekeeping and nutrient stress functions, our understanding comes from studies that are mainly reductionist in nature, derived from one-gene-at-a-time experiments examining cell physiology or from biochemical studies of the encoded protein (Baba et al., 2006; Joyce et al., 2006; Nichols et al., 2011). Lacking, however, is an understanding of the interaction of these functions with one another as well as with the greater cell system.

We recently studied the interaction of nutrient stress genes with the dispensable gene set under nutrient-rich conditions using a synthetic lethality approach (Côté et al., 2016). Some 82 nutrient stress genes (of 119) were crossed with nearly 4,000 single gene deletion mutants to identify synthetic sick or lethal interactions. With a total of 1,881 such interactions, this study revealed a large number and density of synthetic lethal (or sick) gene combinations for the query gene set. This work revealed signature interactions

between nutrient acquisition and biosynthesis as well as pathway redundancies and the presence, in this network, of a surprising number of genes of unknown function. Of course, this gene-gene interaction dataset was collected in nutrient-rich conditions, where nutrient stress genes are dispensable. Accordingly, the dataset lacks any insight into the genetic network under nutrient-limited conditions, where nutrient stress genes have an essential phenotype. Indeed, while genetic mutation is the dominant approach to studying genetic networks, genes with indispensable phenotypes have largely resisted characterization with network mapping tools. Only in the model yeast, *Saccharomyces cerevisiae*, where temperature sensitive alleles have been created in hundreds of essential genes, has synthetic lethality been used to map the interaction network of genes with essential phenotypes (Costanzo et al., 2016; Li et al., 2011). That effort showed that essential genes had many more interactions on average than dispensable genes and established these functions as hubs in the global genetic network of *S. cerevisiae*.

In the work described here, we have taken a chemical biology approach (Parsons et al., 2004) to study the interaction network of housekeeping and nutrient stress functions in *E. coli* under nutrient-limited conditions where all of the associated genes are essential for growth. We targeted these functions with 45 chemical probes using a matrix of 990 pairwise chemical combinations. The compounds were systematically combined in 64-dose checkerboard matrices, referred to here as checkerboard assays, and assessed for the interaction phenotype of growth inhibition: synergy, antagonism, or indifference (no interaction). When growth inhibition that results from combining two compounds is simply the sum of effects of individual compounds, this is referred to as indifference. In

contrast, synergy and antagonism describe phenotypic interactions that lead to more or less growth inhibition, respectively, than predicted by the sum of the individual effects of the two compounds. Some 81 synergistic and 105 antagonistic interactions were recorded, suggesting that essential functions represent a highly connected network in bacteria. We further investigated some especially potent and paradoxical interactions between biotin and fatty acid synthesis, amino acid biosynthesis and ribosome assembly, as well as purine synthesis and protein translation inhibition. In all, the work highlighted a high density of interactions among essential functions as well as unique connectivity between nutrient biosynthesis and housekeeping functions in bacteria.

Results

Chemical-Chemical Interaction Matrix

Some 45 compounds, known to probe bacterial functions in nutrient synthesis and housekeeping functions, were selected to generate a systematic analysis of chemical-chemical combinations in order to determine the nature of each interaction, i.e., synergistic, antagonistic or indifferent (no interaction). The 45 compounds included 18 nutrient synthesis probes with growth inhibitory activities restricted to nutrient-limited conditions, and 27 housekeeping-function probes that included antibiotics that target cell wall, protein synthesis, or DNA replication, and lamotrigine, a compound that was recently identified to be an inhibitor of bacterial ribosome biogenesis (Stokes et al., 2014). A list of the probes and their targets is provided in supplementary data (Supplemental Data 1). We systematically characterized the interaction of these 45

compounds with one another using checkerboard assays that analyzed the concentration dependence of the activity of pairs of compounds. The phenotype tested for was the inhibition of *E. coli* growth in the nutrient-limited medium, M9 minimal, which consisted of only glucose as the carbon source, ammonium chloride as the nitrogen source, essential phosphates and salts. Binary combination space for n compounds is defined by the formula n*(n-1)/2 (Keith et al., 2005). Thus, 990 pairs of compounds were tested using a total of 63,360 individual data points. An analysis of the checkerboard data revealed the nature of these pairwise interactions. The fractional inhibitory concentration index (FICI) was calculated for each checkerboard assay (Transparent Methods). FICI is the sum of the fractional inhibitory concentration (FIC) of each tested compound, where the FIC for each tested compound is the ratio of its minimal inhibitory concentration (MIC) in combination divided by its MIC when used on its own (Krogstad et al., 1986). Combinations with an FICI less than or equal to 0.5 were deemed synergistic, while those with an FICI greater than 2 were antagonistic and those greater than 0.5 and less than or equal to 2 were indifferent (Krogstad et al., 1986). Of the 990 unique pairwise chemical combinations, 81 were synergistic and 105 were antagonistic, resulting in 186 interactions (Table 1 and Figure S1). Thus, on average, each compound had approximately 4 interactions (186/45) with other bioactive compounds. Furthermore, this dataset defined a unique interaction profile for each of the 45 compounds. Figure 1 presents a heat map of the entire interaction map, hierarchically clustered according to the interaction profiles of each chemical compound. All FICI values for the latter are provided in supplemental data (Supplemental Data 2).

A large fraction, some 19% (186/990) of these probe combinations, showed interactions suggesting that targeted functions represent a highly-connected network in bacteria. While these compounds were selected to probe a diverse array of bacterial physiology, we included some redundancy in our selection of housekeeping-function probes, for example, two phenicols (chloramphenicol and thiamphenicol), two macrolides (erythromycin and dirithromycin), two fluoroquinolones (norfloxacin and ciprofloxacin) and two tetracyclines (tetracycline and doxycycline). Interestingly, we noted subtle differences among the interaction profiles of closely related probes; however, even if we account for some redundancy by estimating the set of non-redundant probes to be 35 compounds, for example, we still see interactions for about 20% of our combinations. In all, this analysis represents the first comprehensive study to explore the interaction of chemical probes, targeting housekeeping functions and nutrient synthesis, when cells are grown under nutrient-limited conditions. Table 1 summarizes and categorizes the 186 interactions seen.

A total of 57 interactions occurred among the 27 probes of housekeeping functions. Previous antibiotic combination studies have charted interactions between the different classes of antibiotics (Ocampo et al., 2014; Wambaugh et al., 2017; Yeh et al., 2006), which were largely confirmed by our dataset of antibiotic combinations. An interesting pair of interactions identified in this study was the potentiation of the narrow-spectrum Gram-positive antibiotic novobiocin (Cozzarelli, 1977) against *E. coli* by two cell wall-active antibiotics, vancomycin and fosmidomycin (Figure S2). The latter interaction has not been reported by previous antibiotic combination studies. These

interactions were curious, however, we have focused herein on new chemical-chemical interactions of relevance to the study of nutrient stress in bacteria.

Interestingly, of the total 186 charted interactions, 129 involved nutrient synthesis probes. Specifically, 88 interactions were between one of the 18 nutrient synthesis probes and the remaining 27 probes targeting housekeeping functions, as well as 41 interactions that involved combinations of nutrient synthesis probes. Below, we highlight some illustrative and paradoxical interactions in these categories in addition to presenting a deeper analysis of the interactions seen with probes of biotin, S-adenosylmethionine and purine synthesis.

Hierarchical Clustering of Chemical-Chemical Interactions

Hierarchical clustering revealed that probes mapped largely based on the chemical class of each compound. This was an encouraging result from the perspective of data quality because compounds of similar structure naturally share the same mechanism of action (MOA) and thus similar chemical-chemical interaction profiles. For example, norfloxacin and ciprofloxacin belong to the fluoroquinolone antibiotic class, and inhibit DNA synthesis by targeting DNA gyrase and topoisomerase IV (Drlica et al., 2008). They show a unique fingerprint of antagonistic interactions with the tetracyclines, chloramphenicol, and the thiopurine analogues (6-thiogunaine and 6-mercaptopurine) (Figure 2A). The macrolides, erythromycin and dirithromycin, inhibit protein synthesis by targeting the 50S ribosome (Menninger and Otto, 1982). They have a similar profile of interactions including antagonistic interactions with gentamycin and D-cycloserine, and

synergistic interactions with the tetracyclines, polymyxin B, DL-3-hydroxynorvaline, lamotrigine, and *p*-fluoro-phenylalanine (Figure 2B). Similarly, analogues of nutrient synthesis probes, such as the thiopurine analogues 6-thioguanine and 6-mercaptopurine, which incorporate into DNA or RNA and inhibit purine synthesis (Nelson et al., 1975), had similar interaction profiles (Figure 2C). Interactions unique to these compounds include 6-mercaptopurine's synergistic interactions with L-norvaline and trimethoprim as well as 6-thiogunaine's antagonistic interaction with cefmetazole. These observations emphasize the concept that compounds with similar mechanism of action exhibit similar interactions when used in combination with other probes of biology.

Occasionally, interaction profiles clustered together due to a synergistic interaction between the two compounds. For instance, the interaction profiles of sulfamethoxazole and trimethoprim cluster with one another, yet only share one common antagonistic interaction with lincomycin (Figure 1). Similarly, DL-3-hydroxynorvaline, a threonine analogue (Minajigi et al., 2011), synergizes with the macrolides, erythromycin and dirithromycin, and clusters with these compounds, however, each compound's profile is largely distinct. Nonetheless, compounds with a shared MOA largely exhibit substantially similar chemical interaction profiles (Figure 1). Accordingly, the chemical-chemical interaction matrix has strong potential as a tool to elucidate the MOA of new antibacterial compounds, tested under nutrient-limited conditions.

Consistent with the goals of understanding the network that underpins nutrient stress in bacteria, we have focused herein on the interactions that impinge on targets associated with nutrient synthesis. Of the 129 interactions in this category, we prioritized

three synergistic interactions that help to elaborate the nutrient stress network in vitamin, amino acid, and nucleobase biosynthesis. Accordingly, we elaborate here on the synergistic interactions that we characterized between MAC13772 and cerulenin, L-norleucine and lamotrigine, as well as 6-mercaptopurine and aminoglycosides. Indeed, these interactions in particular revealed further insight into the complex metabolic network that underpins nutrient stress, as well as the connectivity with other biosynthetic pathways and cellular housekeeping functions.

Biotin Synthesis Interacts with Cell Wall, SAM and Fatty Acid Biosynthesis

The biotin biosynthesis inhibitor MAC13772 targets 7,8-diaminopelargonic acid synthase (BioA), an enzyme responsible for the catalysis of the antepenultimate step in biotin biosynthesis (Zlitni et al., 2013). Its chemical interaction profile identified one antagonistic and four synergistic signature interactions (Figure 3A). MAC13772 antagonized the activity of the cell wall antibiotic D-cycloserine (FICI $\geq$ 3, Figure S3A), a D-alanine analogue that competitively inhibits two essential enzymes, alanine racemase and D-alanine-D-alanine ligase (Lambert and Neuhaus, 1972; Zawadzke et al., 1991). The racemase catalyzes the conversion of L-alanine to D-alanine, and the ligase utilizes two D-alanine molecules as substrates for peptidoglycan synthesis (Figure S4). Notably, the biotin biosynthetic step preceding the antepenultimate step utilizes L-alanine as a substrate (Figure 3B). The observed antagonistic interaction suggests that the inhibition of BioA increases L-alanine availability from the preceding biotin biosynthetic step (Figure S3). Where L-alanine and D-cycloserine compete for binding to the active site of the racemase, the abundance of L-alanine enables alanine racemization and continued

peptidoglycan synthesis, suppressing D-cycloserine's activity. Recent genomic analyses of *Mycobacterium tuberculosis* have revealed that strains harboring a loss-of-function mutation in *ald*, which codes for an enzyme that catalyzes the conversion of L-alanine to pyruvate, are resistant to D-cycloserine (Desjardins et al., 2016). This Ald variant can no longer utilize L-alanine as a substrate, leading to an increase in the pool of available L-alanine and suppression of D-cycloserine activity. Accordingly, we speculate that L-alanine is redirected towards peptidoglycan synthesis on inhibition of BioA and antagonizes the effects of D-cycloserine.

In addition to its connection to cell wall synthesis, the BioA inhibitor MAC13772 also synergized with L-norleucine (FICI $\leq$ 0.5), a nutrient synthesis probe that targets S-adenosylmethionine (SAM) biosynthesis (Chattopadhyay et al., 1991). In *E. coli*, the first biotin biosynthetic step requires SAM as a methyl donor (Figure 3B). L-norleucine inhibits the methionine adenosyltransferase reaction of the MetK enzyme that catalyzes the production of SAM (Chattopadhyay et al., 1991). Thus, the first step of biotin biosynthesis is perturbed by the action of L-norleucine and synergy is seen with MAC13772, an inhibitor of the antepenultimate step of the same pathway (Figure 3B). Indeed, it has been previously shown that inhibition of SAM-dependent methyltransferases cascades to the inhibition of biotin biosynthesis (Lin et al., 2010). Accordingly, in nutrient-limited conditions, inhibition of both SAM and a late step in the biotin biosynthetic pathway exerts a synergistic inhibition of biotin biosynthesis in *E. coli*, leading to growth inhibition.

The synergistic interaction of MAC13772 and the antibiotic cerulenin (FICI ≤ 0.5) highlighted an interdependence between biotin biosynthesis and fatty acid synthesis in nutrient-limited conditions (Figures 3C and 3D). Cerulenin targets ß-ketoacyl-ACP synthase I (FabB), one of the condensing enzymes required for fatty acid biosynthesis (Price et al., 2001) and one of the four enzymes involved in the elongation of biotin's saturated chain moiety (Figure 3B). The first step in *E. coli*'s fatty acid biosynthetic pathway is a decarboxylation reaction catalyzed by the ACC complex, which uses biotin as a cofactor (Campbell and Cronan Jr, 2001). Thus, the inhibition of biotin biosynthesis appears to have a unique impact on fatty acid biosynthesis when *E. coli* is grown in nutrient-limited conditions. We assessed the effect of decreased biotin availability in nutrient-rich conditions using a Δ*bioP* mutant, harboring a deletion in the biotin transporter. In this genetic background, biotin biosynthesis becomes essential in nutrient-rich conditions. Indeed, Figure 3E revealed a synergistic interaction in the Δ*bioP* mutant strain, in which cerulenin was similarly potentiated in the presence of sub-inhibitory concentrations of MAC13772. Thus, a decrease in biotin availability impacts fatty acid biosynthesis and sensitizes this pathway to inhibition by the antibiotic cerulenin.

Perturbation of S-adenosylmethionine Biosynthesis Reveals a Growth Phenotype for the Ribosome Biogenesis Function of Initiation Factor 2

Lamotrigine is best known as an anticonvulsant drug, however, it was recently shown to have the cryptic ability to inhibit a previously unrecognized ribosome biogenesis function in *E. coli*, namely that of initiation factor 2 (IF2) (Stokes et al., 2014).

While IF2 is understood to have a key role in the initiation of protein translation, its role in ribosome assembly was revealed through an investigation of the action of lamotrigine under conditions of cold stress (15°C). Indeed, the ability of lamotrigine to inhibit the growth of *E. coli* was dependent on cold stress. Herein, under nutrient-limited conditions and at 37°C, we found that lamotrigine showed a cold-independent activity, particularly in combination with certain compounds. The chemical interaction profile of lamotrigine revealed 12 synergistic interactions (Figure 4A), including those with the antibiotics cerulenin and rifampicin that persisted in nutrient-rich conditions (Figure S5). Ribosome assembly remains a highly enigmatic process (Shajani et al., 2011). Indeed, we are at a loss to identify a precise mechanism behind many of the interactions that we observed for lamotrigine. Nevertheless, we believe the synergistic interaction between the S-adenosylmethionine (SAM) biosynthesis inhibitor L-norleucine (Chattopadhyay et al., 1991) and lamotrigine (Figures 4B and 4C) may be instructive of the network that underpins ribosome assembly. Remarkably, inhibition of SAM biosynthesis revealed a growth inhibitory phenotype for the IF2-targeted compound lamotrigine under standard temperature conditions (37°C).

Ribosome profiling analysis, an assessment of the distribution of ribosomal material among the 30S, 50S and 70S species using sucrose gradient sedimentation, is a signature phenotype that is commonly used to characterize defects in ribosome assembly. Treatment of *E. coli* growing in nutrient-rich media at 15°C with lamotrigine was previously shown to be growth inhibitory and lead to the accumulation of immature pre-30S and pre-50S ribosomal particles, consistent with its inhibition of ribosome biogenesis

(Stokes et al., 2014). Lamotrigine treatment had no such impact on growth or ribosome profiles, however, when *E. coli* was grown in nutrient-rich media at 37°C. In the work reported herein, lamotrigine treatment at 15°C in nutrient-limited media revealed the accumulation of a pre-50S species only (Figure S6). Notably, this phenotype was also seen when cells grown at 37°C in nutrient-limited conditions were treated with lamotrigine. Ribosomal subunit assembly is not thought to proceed with a linear and specific maturation pathway. Instead, emerging research suggests that there are multiple and parallel pathways for maturation. Our earlier work with lamotrigine suggested that, while the function of IF2 in late ribosomal subunit assembly was dispensable at 37°C, it became essential at cold temperatures where alternative pathways for assembly became limiting due to the effect of temperature on RNA folding. Nutrient limitation has not made this function essential for growth – Figure 4B reveals that high concentrations of lamotrigine alone are not especially growth inhibitory – but has revealed a small but detectable defect in the ribosome profile. Here again we posit that nutrient limitation may limit the assembly landscape, revealing a dependence on IF2. Notwithstanding its interaction with lamotrigine in these conditions, L-norleucine treated cells grown at 37°C in nutrient-limited conditions had a ribosome profile that was indistinguishable from untreated cells (Figure S6). Interestingly, SAM-dependent methylation of multiple sites of 30S and 50S ribosomal particles has a key role in ribosome biogenesis (Kaczanowska and Rydén-Aulin, 2007; Shajani et al., 2011). To be sure that this interaction was due to the inhibition of IF2's ribosome assembly function, we assessed the interaction with an *E. coli* IF2 mutant (mutant 1) (Stokes et al., 2014). This mutant encodes an N-terminal IF2

variant that resists the capacity of lamotrigine to inhibit the ribosome assembly activity of IF2. (We note here that lamotrigine does not affect the protein translation function of IF2.) The absence of synergy in Figure 4D confirmed that the interaction in wild-type *E. coli* was related to the role of IF2 in ribosome assembly. Although treatment of wild-type *E. coli* with L-norleucine did not produce immature ribosomal particles (Figure S6), perturbation of S-adenosylmethionine biosynthesis with L-norleucine revealed a profound growth phenotype at 37°C for the ribosome biogenesis function of initiation factor 2.

Thiopurine Analogues Showed Synergy with Aminoglycoside Antibiotics

Thiopurine antagonists are known for their incorporation into DNA and RNA and for their inhibition of purine biosynthesis (Van Scoik et al., 1985). Two anticancer drugs, 6-thioguanine and 6-mercaptopurine, are thiopurine analogues that exhibit antibacterial activity in nutrient-defined media (Coonrod and Eickhoff, 1972; Elion et al., 1954b). The mechanism of action of these drugs in both animal and bacterial cells is thought to be manifold. They are converted to purine analogues and get incorporated into DNA and RNA, but they also inhibit purine biosynthesis directly and through negative feedback mechanisms (Atkinson and Murray, 1965; Bolton and Mandel, 1957; Coggin et al., 1966; Elion et al., 1954a). These drugs showed relatively promiscuous interaction profiles in our study. For example, 10 synergistic interactions were seen for 6-mercaptopurine (Figure 2C). We posit that the central role of purine synthesis in metabolism and growth leads to pleiotropy and a high density of interactions for these probes. Nevertheless, we found it curious that, of the many protein synthesis inhibitors studied, we saw interactions with only the aminoglycoside class, namely with kanamycin and gentamycin. We

followed up on this observation, focusing on the probe 6-mercaptopurine (Figure 5) because it was more potent than 6-thioguanine (Figure S7). Aminoglycoside antibiotics are understood to corrupt ribosome function by promoting mistranslation of proteins, and these have toxic effects (Davies et al., 1965). Aminoglycosides encompass two distinct structural classes, the 4,5-disubstituted and the 4,6-disubsituted 2-deoxystreptamines, and a structurally dissimilar class that does not share a common backbone in their chemical structures (Mingeot-Leclercq et al., 1999; Recht and Puglisi, 2001). In the course of interfering with ribosome function, these compounds form sequence-specific hydrogen bonds with nucleotides in the 30S ribosome subunit (Recht and Puglisi, 2001). We reasoned that if 6-mercaptopurine was incorporated into RNA, then it may exert synergy with aminoglycosides through specific interactions. We found that not only did 6-mercaptopurine show strong synergy with the 4,6-disusbtituted aminoglycosides kanamycin and gentamycin, but also with all aminoglycosides tested regardless of structural class (Figure 6). Thus, we conclude that the synergy is not due to specific interactions but rather to the downstream consequences of inhibition of purine biosynthesis and/or incorporation into DNA and RNA. Notably, neomycin resulted in the strongest synergistic interaction profile in which 6-mercaptopurine was potentiated by more than 100-fold (Figure 6A). In all, these findings revealed an enigmatic interplay among (poly)nucleotide biosynthesis and the action of aminoglycosides on protein mistranslation.

Discussion

Systematic mutation of *E. coli*'s genome of nearly 4,400 genes has revealed that some 303 genes are essential for growth in nutrient-rich conditions. These genes code for so-called housekeeping functions in the bacteria that include cell wall, DNA, RNA and protein synthesis. Under nutrient stress, an additional set of 119 genes have an indispensable phenotype and these genes code largely for the synthesis of amino acids, vitamins, nucleobases, and other cofactors. While much is known about the function of this set of 422 genes and their gene products, relatively little is known of their interactions. Gene-gene interactions are commonly studied in model microbes with systematic genome-wide crosses of deletion mutations to detect unexpected growth defects in strains bearing deletions in two otherwise dispensable genes, resulting in a synthetic lethal or sick phenotype (Butland et al., 2008; Collins et al., 2007; Costanzo et al., 2016; Côté et al., 2016; Tong et al., 2001; Typas et al., 2008). To date, this approach has been limited, in *E. coli*, to the dispensable fraction of the genome where stable mutants can be created. Herein, we sought to probe the collection of genes involved in both housekeeping and nutrient stress functions (422) under conditions where they have an essential growth phenotype. To this end, we systematically combined 45 chemical probes targeting a subset of these functions. The 45 compounds included 18 nutrient synthesis probes with growth inhibitory activities restricted to nutrient-limited conditions, and 27 housekeeping-function probes that included antibiotics that target cell wall, protein synthesis, or DNA replication, and lamotrigine, a compound that was recently identified to be an inhibitor of bacterial ribosome biogenesis (Stokes et al., 2014).

Binary combination space for these 45 compounds encompasses 990 unique pairwise chemical-chemical combinations. Of these systematic combinations, we mapped 186 interactions, of which 81 were synergistic and 105 were antagonistic. In addition, we defined signature interaction profiles for all 45 compounds under nutrient stress conditions, averaging 4 interactions per compound. Thus 19% (186/990) of probe combinations produced an interaction in this study. This number is high relative to the 3.5% frequency recorded in the model yeast for both positive and negative gene-gene interactions seen genome-wide in that organism (Costanzo et al., 2016). Indeed, synthetic lethal/sick interactions recorded among the dispensable fraction in *E. coli* have been in a similar frequency range, for example, ~1% of the total number of double mutants created (French et al., 2017). Interestingly, interactions among the essential gene set in the model yeast *S. cerevisiae* have been probed using temperature sensitive alleles (Costanzo et al., 2016) and this work has revealed a much higher frequency of interaction (24%), closer to that seen here using chemical probes of essential functions. Together, these studies suggest that essential physiology is described by a network that is much more densely wired than that of the dispensable fraction.

Hierarchical clustering of these 45 chemical-chemical interaction profiles revealed that compounds clustered based on their chemical class. This points to a qualitative dataset since analogues share a similar mechanism of action, thus resulting in overlapping interaction profiles. Indeed, drugs belonging to the same antibiotic class clustered together, revealing the least dissimilarity in their signature interactions. Another cluster identified highly similar profiles between the nutrient synthesis probes 6-thioguanine and

6-mercaptopurine. Occasionally, some interaction profiles clustered due to a shared synergistic interaction between the two clustered compounds. This phenotype occurred for increasingly dissimilar interaction profiles, suggesting that the interaction matrix is not predictive of synergistic interactions. Rather, the chemical-chemical interaction matrix is a powerful tool that can be used to elucidate the MOA of novel compounds with antibacterial activity (unknowns). Particularly, unknowns with a growth inhibitory activity restricted to nutrient-limited conditions can have their target prioritized through nutrient suppression profiling (Zlitni et al., 2013), in which unknowns selectively targeting a unique nutrient biosynthetic pathway are suppressed in the presence of the respective nutrient. However, nutrient suppression profiling lacks the ability to determine the target of unknowns suppressed by multiple nutrients. We have previously demonstrated the utility of chemical combinations in elucidating the target and MOA of uncharted chemical probes (Farha and Brown, 2010). Therefore, chemically combining the 45 compounds with such unknowns would generate fingerprint chemical interaction profiles that are unique to each unknown. Since the chemical-chemical interaction matrix clusters compounds with a similar mechanism of action, hierarchical clustering of the unknowns' interaction profiles with that of the 45 compounds would allow the generation of testable hypotheses for the MOA of the unknown in question.

Systematic combinations of nutrient synthesis probes and other bioactive compounds charted 186 interactions in *E. coli* grown in nutrient-limited conditions. Of the 186 interactions, 57 involved combinations of housekeeping-function probes, many of which have been identified by previous antibiotic combination studies (Ocampo et al.,

2014; Wambaugh et al., 2017; Yeh et al., 2006). Herein, we report a novel synergistic interaction in *E. coli* that involves the Gram-positive antibiotic novobiocin and the cell wall active antibiotic fosmidomycin. This was an interesting observation in which a Gram-positive antibiotic was potentiated in *E. coli*, a Gram-negative bacterium. Nevertheless, we focused herein on interactions that provide further understanding of nutrient stress in *E. coli*. Of the 129 interactions, we prioritized three interactions to understand the connectivity between nutrient synthesis and other cellular housekeeping functions in nutrient-limited conditions.

Combinations of nutrient synthesis probes revealed how biosynthetic pathways are interdependent in nutrient-limited conditions. Previous *in vitro* studies have shown that inhibition of fatty acid biosynthesis by cerulenin perturbs biotin biosynthesis (Lin et al., 2010). In this work, we report a synergistic interaction between the biotin biosynthesis inhibitor, MAC13772, and cerulenin. Further, this synergy echoes a synthetic sick interaction observed between *fabH*, coding for a fatty acid biosynthetic enzyme, and *bioA*, encoding the target in biotin synthesis for the compound MAC13772 (Côté et al., 2016). Together these studies suggest a strong connectivity between these biosynthetic pathways. To further probe this connectivity, we assessed the interaction between MAC13772 and cerulenin in nutrient-rich conditions, and in an *E. coli* strain (Δ*bioP)* lacking the biotin transporter, making biotin biosynthesis essential regardless of the growth medium. Here again, we saw a strong interaction between these compounds consistent with a crucial role for the biotin biosynthetic pathway in the synthesis of fatty acids.

Other interactions that we studied in detail were those of lamotrigine, an inhibitor of the ribosome biogenesis function of initiation factor 2 (IF2) in *E. coli* (Stokes et al., 2014). Lamotrigine proved to be a promiscuous interactor with 12 novel synergistic partner compounds. The activity of lamotrigine against IF2 has formerly only been evident under cold stress, a condition that is thought to narrow the complex and redundant landscape for ribosomal subunit maturation. Thus, the activity and cold-independent phenotype of lamotrigine in combination with other several cellular probes at 37°C is a new development and may be an indication that ribosome biogenesis is a particularly important hub in the cellular network. Of the 12 synergistic interactions, those with cerulenin and rifampicin were also evident in nutrient-rich media. Such interactions of clinically used drugs are intriguing from the prospect of therapy. In this context, we note that our dataset also revealed the well-known interaction between sulfamethoxazole and trimethoprim, a highly synergistic interaction between sulfamethoxazole, an inhibitor of folate synthesis, and trimethoprim, an inhibitor of dihydrofolate reductase. Where the former is active only in nutrient-limited media, the latter is active in both nutrient-limited and rich microbiological media, and arguably inhibits a housekeeping function, namely the provision of reduced folates for a variety of cellular processes including DNA synthesis. The mechanism behind this synergistic interaction, which persists in rich microbiological media, remains elusive, but is an example of one that has had great utility in antimicrobial therapy for many decades (Masters et al., 2003).

The interaction of lamotrigine with L-norleucine, an inhibitor of S-adenosylmethionine biosynthesis was particularly interesting and consistent with the

emerging recognition of the role of a number of bacterial methyltransferases in ribosome biogenesis (Baldridge and Contreras, 2014). The exact function of these modifications has been elusive; however, it has been suggested that methylation may serve as a checkpoint in ribosomal subunit assembly. The strong synergistic interaction seen in the work reported here sheds additional light on the role of IF2 in ribosome assembly by charting a connection between subunit methylation and the role of IF2 in the late steps of ribosome assembly. Of practical note, SAM biosynthesis has recently been identified as a potential antibacterial target in *M. tuberculosis*, where strains harboring SAM biosynthetic gene deletions are impaired for growth *in vivo* (Berney et al., 2015).

Another notable synergy was that between 6-mercaptopurine and aminoglycoside antibiotics. The former is an antimetabolite used in cancer therapy but not for bacterial infection. Although the precise mechanism underpinning the synergy remains unknown, we present evidence herein that the synergy is not due to the direct interaction of aminoglycosides with thiopurines incorporated into ribosomal RNA. Where purine biosynthesis has been shown to be important for the proliferation and survival of *E. coli* in blood (Samant et al., 2008), the interaction with aminoglycosides is one worthy of additional study as is the potential of combination therapies that would exploit targets in translation and purine synthesis.

The chemical biology combinations approach reported herein has charted the first foray into pairwise interactions of functions in *E. coli* with essential growth phenotypes. The effort defined a high density of interactions among these functions and suggest that additional probes would facilitate the expansion of this effort to further understand this

aspect of the bacterial cell network as well as the importance of nutrient stress. Indeed, nutrient metabolism is emerging as a viable virulence target in pathogenic bacteria (Cersini et al., 1998; Cuccui et al., 2007; Mei et al., 1997). Full validation of nutrient metabolism as a therapeutic target will come from studies of sites of infection and specific pathogens as well as a thorough understanding of the complex network that underpins nutrient stress in bacteria.

Transparent Methods

Bacterial Strains and Culture Conditions

All antibiotics and chemicals used in the study were purchased from Sigma Aldrich. The strains used in this study were *E. coli* BW25113, Δ*bioP* (*E. coli* parent strain BW25113, (Baba et al., 2006)), and the *E. coli* IF2 mutant (mutant 1) (*E. coli* parent strain BW25113, (Stokes et al., 2014)). In all experiments, cells were prepared to a final working inoculum of 10^5 CFU/ml. Bacterial cells were grown overnight in M9 minimal medium (0.4% glucose) or LB medium and then diluted 1:50 in fresh M9 minimal medium or LB medium, respectively, and grown at 37°C with aeration at 250 rpm to an OD_{600} of 0.4 (mid-log culture). Cells grown in M9 minimal medium were then diluted 1:1,000 in fresh M9 minimal medium, while those grown in LB medium were diluted 1:10,000 in fresh LB medium, unless stated otherwise.

Determination of Minimal Inhibitory Concentrations

The minimal inhibitory concentration (MIC) for all compounds was determined to inform on the starting concentrations in the checkerboard assays. *E. coli* strains were

grown and prepared to a final working inoculum in M9 minimal or LB medium as described above, unless stated otherwise. These cells were then added to a 96-well assay plate containing 2-fold serial dilutions of one of the compounds of interest, where concentrations ranged from 256 µg/ml to 0 µg/ml. Prior to incubation, absorbance at 600 nm (OD_{600}) of the 96-well assay plates was measured using the Tecan plate reader (Infinite M1000). Assay plates were then incubated at 37°C in a stationary incubator for 18 h and OD_{600} was measured. Growth (G) at each exposed concentration was determined as follows

$$G = G_{t=18} - G_{t=0}$$

where $G_{t=18}$ corresponds to the absorbance measured after 18 h of incubation, and $G_{t=0}$ corresponds to the absorbance measured prior to incubation. From here, percent residual growth (%G) was calculated, as follows, to determine the MIC of the compound of interest

$$\%G = \frac{G_i}{G_0}$$

where G_i represents the growth in one of the 12 wells exposed to the different concentrations of the tested compound, and G_0 represents the growth in the well that was not exposed to the tested compound. The concentration that resulted in a percent residual growth ≤ 10% was deemed as the MIC of the tested compound.

Checkerboard Assays

E. coli strains were grown and prepared to a final working inoculum in M9 minimal or LB medium as described above, unless stated otherwise. The checkerboard assays were done as 8 × 8 dose-point matrices. All compounds were prepared as 2-fold serial dilutions starting at 4 × MIC to 0 µg/ml. Absorbance (OD_{600}) was measured prior to and post incubation at 37°C in a stationary incubator. To define interaction of the combination, the fractional inhibitory concentration index (FICI) for each checkerboard assay was calculated as follows

$$\mathrm{FICI} = \frac{\mathrm{MIC_{A,X}}}{\mathrm{MIC_A}} + \frac{\mathrm{MIC_{B,X}}}{\mathrm{MIC_B}}$$

where $MIC_{A,X}$ is the MIC of compound A in combination with compound B, MIC_A is the MIC of compound A on its own, $MIC_{B,X}$ is the MIC of compound B in combination with compound A, and MIC_B is the MIC of compound B on its own. FICI values ≤ 0.5 were synergistic interactions, values > 2 were antagonistic, and values > 0.5 and ≤ 2 were indifferent (Krogstad et al., 1986). All highlighted interactions were done in biological replicates. Of note, checkerboard assays constituting the chemical-chemical interaction matrix were done in the *E. coli* parent strain BW25113, which was grown and prepared to a final working inoculum in M9 minimal medium as described above.

Hierarchical Clustering of the Chemical-Chemical Interaction Profiles

The hierarchical cluster was compiled using the statistical computing and graphics programming language, R. The ward.2 clustering method, which clusters datasets based on the least variance between *n* groups (Murtagh and Legendre, 2014), was

implemented in the heatmap.2 function found in the gplots library. The dataset introduced to R comprises the FICIs of all 990 combinations.

Assessment of the Synergy of MAC13772 with cerulenin in the *E. coli* Mutant Strain Δ*bioP*

The synergy of MAC13772 with cerulenin was assessed in the *E. coli* mutant strain Δ*bioP* (Baba et al., 2006). The Δ*bioP* mutant strain was prepared to a mid-log culture in M9 minimal medium, as described above, to deplete the cells from any extracellularly available biotin. The mid-log culture was then diluted 1:10,000 in fresh LB medium. The MIC of the compounds was determined, and the checkerboard assays were prepared and analyzed as previously described.

Ribosome Profiling by Sucrose Density Gradient Analysis

A single colony of *E. coli* BW25113 was inoculated in 5 ml of M9 minimal medium (0.4% glucose) and grown overnight at 37°C with aeration at 250 rpm. The overnight culture was then diluted 1:20 in 50 ml of fresh M9 minimal medium (0.4% glucose) to obtain an OD600 of ~0.05, and grown at 37°C or 15°C with aeration at 250 rpm to an OD600 of 0.2 (early-log). Consequently, the early-log culture was treated with a sub-inhibitory concentration of lamotrigine or L-norleucine, and allowed to grow for 1 h and 16 h at 37°C and 15°C, respectively, with aeration at 250 rpm. Cells were then harvested by centrifugation (4,000 rpm at 4°C for 30 min), resuspended in 5 ml chilled ribosome buffer (20 mM Tris-HCl, pH 7.0, 10.5 mM MgOAc, 100 mM NH4Cl, and 3 mM β-mercaptoethanol), and lysed using a cell disruptor set at 13 kpsi.

Cell lysates were clarified at 24,000 rpm at 4°C for 45 min, and the supernatant was collected and loaded onto 35 ml 10-45% sucrose gradients for ultracentrifugation (18,700 rpm at 4°C for 17 h). The gradients were then analyzed using an AKTA Prime FPLC equipped with a UV flow cell, which was set at an absorbance of 260 nm (Stokes et al., 2014). For each set of treated-cultures, a control culture without treatment was also harvested, lysed, clarified, and analyzed.

Supplemental References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, singlegene knockout mutants: the Keio collection. Mol Syst Biol *2*, 2006.0008-2006.0008.

Krogstad, D., Moellering, R.J., and Lorian, V. (1986). Antibiotics in laboratory medicine. Baltimore: Williams and Wilkins.

Murtagh, F., and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J Classif *31*, 274-295.

Stokes, J.M., Davis, J.H., Mangat, C.S., Williamson, J.R., and Brown, E.D. (2014). Discovery of a small molecule that inhibits bacterial ribosome biogenesis. eLife *3*, e03574.

Acknowledgements

This work was supported by a salary award to E.D.B. from the Canada Research Chairs program as well as with operating funding from the Canadian Institutes of Health Research [FDN-143215]. We thank Gerard D. Wright for providing the natural product furanomycin. We also thank Maya A. Farha for the helpful suggestions and comments on the manuscript, and Jonathan M. Stokes for the insightful discussions on ribosome biogenesis.

Author Contributions

Conceptualization, S.S.E. and E.D.B.; Methodology, S.S.E. and E.D.B.; Software, S.S.E.; Validation, S.S.E.; Formal Analysis, S.S.E.; Investigation, S.S.E. and E.D.B.; Resources, E.D.B.; Writing – Original Draft, S.S.E. and E.D.B.; Writing – Review & Editing, S.S.E. and E.D.B.; Visualization, S.S.E.; Supervision, E.D.B.; Funding Acquisition, E.D.B.

References

Atkinson, M.R., and Murray, A.W. (1965). Inhibition by 6-mercaptopurine of purine phosphoribosyltransferases from Ehrlich ascites-tumour cells that are resistant to this drug. Biochem J *94*, 71-74.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol *2*, 2006.0008-2006.0008.

Baldridge, K.C., and Contreras, L.M. (2014). Functional implications of ribosomal RNA methylation in response to environmental stress. Crit Rev Biochem Mol Biol *49*, 69-89.

Berney, M., Berney-Meyer, L., Wong, K.-W., Chen, B., Chen, M., Kim, J., Wang, J., Harris, D., Parkhill, J., Chan, J., *et al.* (2015). Essential roles of methionine and S-adenosylmethionine in the autarkic lifestyle of *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA *112*, 10008-10013.

Bolton, E.T., and Mandel, H.G. (1957). The effects of 6-mercaptopurine on biosynthesis in *Escherichia coli*. J Biol Chem *227*, 833-844.

Butland, G., Babu, M., Diaz-Mejia, J.J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A.G., Pogoutse, O., *et al.* (2008). eSGA: *E. coli* synthetic genetic array analysis. Nature Methods *5*, 789-795.

Campbell, J.W., and Cronan Jr, J.E. (2001). Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery. Annu Rev Microbiol *55*, 305-332.

Cersini, A., Salvia, A.M., and Bernardini, M.L. (1998). Intracellular multiplication and virulence of *Shigella flexneri* auxotrophic mutants. Infect Immun *66*, 549-557.

Chattopadhyay, M.K., Ghosh, A.K., and Sengupta, S. (1991). Control of methionine biosynthesis in *Escherichia coli* K-12: a closer study with analogue-resistant mutants. Microbiology *137*, 685-691.

Coggin, J.H., Loosemore, M., and Martin, W.R. (1966). Metabolism of 6-mercaptopurine by resistant *Escherichia coli* cells. J Bacteriol *92*, 446-454.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., *et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806.

Coonrod, J.D., and Eickhoff, T.C. (1972). Combined action of 6-mercaptopurine and antibiotics on gram-negative bacteria *in vitro*. Proc Soc Exp Biol Med *140*, 524-527.

Costanzo, M., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D., *et al.* (2016). A global genetic interaction network maps a wiring diagram of cellular function. Science *353*, aaf1420.

Côté, J.-P., French, S., Gehrke, S.S., MacNair, C.R., Mangat, C.S., Bharat, A., and Brown, E.D. (2016). The genome-wide interaction network of nutrient stress genes in *Escherichia coli*. mBio *7*, e01714-01716.

Cozzarelli, N.R. (1977). The mechanism of action of inhibitors of DNA synthesis. Annu Rev Biochem *46*, 641-668.

Cuccui, J., Easton, A., Chu, K.K., Bancroft, G.J., Oyston, P.C.F., Titball, R.W., and Wren, B.W. (2007). Development of signature-tagged mutagenesis in *Burkholderia pseudomallei* to identify genes important in survival and pathogenesis. Infect Immun *75*, 1186-1195.

Davies, J., Gorini, L., and Davis, B.D. (1965). Misreading of RNA codewords induced by aminoglycoside antibiotics. Mol Pharmacol *1*, 93-106.

Desjardins, C.A., Cohen, K.A., Munsamy, V., Abeel, T., Maharaj, K., Walker, B.J., Shea, T.P., Almeida, D.V., Manson, A.L., Salazar, A., *et al.* (2016). Genomic and functional analyses of *Mycobacterium tuberculosis* strains implicate *ald* in D-cycloserine resistance. Nature Genet *48*, 544-551.

Drlica, K., Malik, M., Kerns, R.J., and Zhao, X. (2008). Quinolone-mediated bacterial death. Antimicrob Agents Chemother *52*, 385-392.

Elion, G.B., Bieber, S., and Hitchings, G.H. (1954a). The fate of 6-mercaptopurine in mice. Ann N Y Acad Sci *60*, 297-303.

Elion, G.B., Singer, S., and Hitchings, G.H. (1954b). Microbiological effects of 6-mercaptopurine. Ann N Y Acad Sci *60*, 200-206.

Farha, M.A., and Brown, E.D. (2010). Chemical probes of *Escherichia coli* uncovered through chemical-chemical interaction profiling with compounds of known biological activity. Chem Biol *17*, 852-862.

French, S., Côté, J.-P., Stokes, J.M., Truant, R., and Brown, E.D. (2017). Bacteria getting into shape: genetic determinants of *E. coli* morphology. mBio *8*, e01977-01916.

Joyce, A.R., Reed, J.L., White, A., Edwards, R., Osterman, A., Baba, T., Mori, H., Lesely, S.A., Palsson, B.Ø., and Agarwalla, S. (2006). Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. J Bacteriol *188*, 8259-8271.

Kaczanowska, M., and Rydén-Aulin, M. (2007). Ribosome biogenesis and the translation process in *Escherichia coli*. Microbiol Mol Biol Rev *71*, 477-494.

Keith, C.T., Borisy, A.A., and Stockwell, B.R. (2005). Multicomponent therapeutics for networked systems. Nat Rev Drug Discov *4*, 71-78.

Krogstad, D., Moellering, R.J., and Lorian, V. (1986). Antibiotics in laboratory medicine. Baltimore: Williams and Wilkins.

Lambert, M.P., and Neuhaus, F.C. (1972). Mechanism of D-cycloserine action: alanine racemase from *Escherichia coli* W. J Bacteriol *110*, 978-987.

Li, Z., Vizeacoumar, F.J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F.S., Min, R., VanderSluis, B., Bellay, J., DeVit, M., *et al.* (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. Nat Biotechnol *29*, 361-367.

Lin, S., Hanson, R.E., and Cronan, J.E. (2010). Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nat Chem Biol *6*, 682-688.

Masters, P.A., O'Bryan, T.A., Zurlo, J., Miller, D.Q., and Joshi, N. (2003). Trimethoprim-sulfamethoxazole revisited. Arch Intern Med *163*, 402-410.

Mei, J.-M., Nourbakhsh, F., Ford, C.W., and Holden, D.W. (1997). Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. Mol Microbiol *26*, 399-407.

Menninger, J.R., and Otto, D.P. (1982). Erythromycin, carbomycin, and spiramycin inhibit protein synthesis by stimulating the dissociation of peptidyl-tRNA from ribosomes. Antimicrob Agents Chemother *21*, 811-818.

Minajigi, A., Deng, B., and Francklyn, C.S. (2011). Fidelity escape by the unnatural amino acid *β*-hydroxynorvaline: an efficient substrate for *Escherichia coli* threonyl-tRNA synthetase with toxic effects on growth. Biochemistry *50*, 1101-1109.

Mingeot-Leclercq, M.-P., Glupczynski, Y., and Tulkens, P.M. (1999). Aminoglycosides: activity and resistance. Antimicrob Agents Chemother *43*, 727-737.

Nelson, J.A., Carpenter, J.W., Rose, L.M., and Adamson, D.J. (1975). Mechanisms of action of 6-thioguanine, 6-mercaptopurine, and 8-azaguanine. Cancer Res *35*, 2872-2878.

Nichols, R.J., Sen, S., Choo, Y.J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K.M., Lee, K.J., Wong, A., *et al.* (2011). Phenotypic landscape of a bacterial cell. Cell *144*, 143-156.

Ocampo, P.S., Lázár, V., Papp, B., Arnoldini, M., Abel zur Wiesch, P., Busa-Fekete, R., Fekete, G., Pál, C., Ackermann, M., and Bonhoeffer, S. (2014). Antagonism between bacteriostatic and bactericidal antibiotics is prevalent. Antimicrob Agents Chemother *58*, 4573-4582.

Parsons, A.B., Brost, R.L., Ding, H., Li, Z., Zhang, C., Sheikh, B., Brown, G.W., Kane, P.M., Hughes, T.R., and Boone, C. (2004). Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. Nat Biotech *22*, 62-69.

Price, A.C., Choi, K.-H., Heath, R.J., Li, Z., White, S.W., and Rock, C.O. (2001). Inhibition of *β*-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin: structure and mechanism. J Biol Chem *276*, 6551-6559.

Recht, M.I., and Puglisi, J.D. (2001). Aminoglycoside resistance with homogeneous and heterogeneous populations of antibiotic-resistant ribosomes. Antimicrob Agents Chemother *45*, 2414-2419.

Samant, S., Lee, H., Ghassemi, M., Chen, J., Cook, J.L., Mankin, A.S., and Neyfakh, A.A. (2008). Nucleotide biosynthesis is critical for growth of bacteria in human blood. PLoS Pathog *4*, e37.

Shajani, Z., Sykes, M.T., and Williamson, J.R. (2011). Assembly of bacterial ribosomes. Annu Rev Biochem *80*, 501-526.

Stokes, J.M., Davis, J.H., Mangat, C.S., Williamson, J.R., and Brown, E.D. (2014). Discovery of a small molecule that inhibits bacterial ribosome biogenesis. eLife *3*, e03574.

Tong, A.H.Y., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Pagé, N., Robinson, M., Raghibizadeh, S., Hogue, C.W.V., Bussey, H., *et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364-2368.

Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L., *et al.* (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. Nature Methods *5*, 781-787.

Van Scoik, K.G., Johnson, C.A., and Porter, W.R. (1985). The pharmacology and metabolism of the thiopurine drugs 6-mercaptopurine and azathioprine. Drug Metab Rev *16*, 157-174.

Wambaugh, M.A., Shakya, V.P.S., Lewis, A.J., Mulvey, M.A., and Brown, J.C.S. (2017). High-throughput identification and rational design of synergistic small-molecule pairs for combating and bypassing antibiotic resistance. PLOS Biology *15*, e2001644.

Yeh, P., Tschumi, A.I., and Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. Nature Genet *38*, 489-494.

Zawadzke, L.E., Bugg, T.D.H., and Walsh, C.T. (1991). Existence of two D-alanine:D-alanine ligases in *Escherichia coli*: cloning and sequencing of the *ddlA* gene and purification and characterization of the DdlA and DdlB enzymes. Biochemistry *30*, 1673-1682.

Zlitni, S., Ferruccio, L.F., and Brown, E.D. (2013). Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. Nat Chem Biol *9*, 796-804.

Figure Legends

Figure 1. Chemical-Chemical Interaction Matrix under Nutrient-Limited Conditions.

A heat map of systematic chemical-chemical combinations of 45 compounds (990 unique combinations) were assessed for their growth inhibitory effects on *E. coli* BW25113 grown in M9 minimal medium. Interactions in the heat map are color-coded for the nature of the interaction as described by the Fractional Inhibitory Concentration Index (FICI) of a systematic dose response analysis of the two compounds, referred to here as

checkerboard analysis and described in the results and methods sections. Of the 990 combinations, 81 were synergistic (yellow) and 105 were antagonistic (dark blue). Hierarchical clustering was performed based on the combination profile of each compound. See also Figures S1 and S2, and Supplemental Data 1 and 2.

Figure 2. Interaction Profiles Cluster Based on Chemical Class and Mechanism.

Compounds belonging to the same chemical class resulted in similar signature interactions with the listed 45 compounds. **(A)** The two fluoroquinolone antibiotics, norfloxacin and ciprofloxacin, antagonized the tetracyclines, chloramphenicol, and the thiopurine analogues. **(B)** Similarly, the macrolide class of antibiotics, specifically erythromycin and dirithromycin, antagonized gentamycin and D-cycloserine, while they synergized with the tetracyclines, polymyxin B, DL-3-hydroxynorvaline, lamotrigine, and *p*-fluoro-phenylalanine. **(C)** The thiopurines, 6-thioguanine and 6-mercaptopurine, are anticancer drugs that revealed highly overlapping signature interactions. All interactions were done in *E. coli* BW25113 grown in M9 minimal medium, and were color-coded as in Figure 1.

Figure 3. Synergy of MAC13772 with Cerulenin Reveals the Importance of Biotin Availability for Fatty Acid Biosynthesis.

(A) The growth inhibitory interaction profile of MAC13772 with other probes. *E. coli* BW25113 was grown in M9 minimal medium. Interactions were color-coded as in Figure 1. **(B)** Biotin biosynthesis in *E. coli* requires S-adenosylmethionine (SAM) in its first committed step. FabB catalyzes one of the elongation steps. MAC13772 inhibits the

antepenultimate step in biotin biosynthesis by targeting BioA. Dashed arrows represent more than one biosynthetic step. **(C)** The synergistic interaction (FICI ≤ 0.5) of MAC13772 and cerulenin; *E. coli* BW25113 grown in M9 minimal medium. **(D)** The synergistic interaction in **(C)** was suppressed (FICI ≤ 2) in nutrient-rich conditions; *E. coli* BW25113 grown in LB medium. **(E)** The synergistic interaction (FICI ≤ 0.38) was also observed in the mutant strain *E. coli* BW25113 Δ*bioP* (Baba et al., 2006), lacking the biotin transporter, grown in LB medium. See also Figures S3 and S4.

Figure 4. The Ribosome Biogenesis Function of IF2 is Revealed when S-adenosylmethionine Biosynthesis is Perturbed.

(A) The growth inhibitory interaction profile of lamotrigine with other probes. *E. coli* BW25113 grown in M9 minimal medium at 37°C. Interactions were color-coded as in Figure 1. **(B)** The synergistic interaction of lamotrigine with L-norleucine (FICI ≤ 0.25); *E. coli* BW25113 grown in M9 minimal medium. **(C)** The synergistic interaction in **(B)** was suppressed (FICI ≤ 2) in nutrient-rich conditions; *E. coli* BW25113 grown in LB medium. **(D)** The interaction of lamotrigine with L-norleucine using the *E. coli* IF2 mutant (Stokes et al., 2014) grown in M9 minimal medium. This panel provides evidence that it is the ribosome biogenesis function of IF2 that contributes to the synergy seen between lamotrigine and L-norleucine. See also Figures S5 and S6.

Figure 5. The Antimetabolite 6-mercaptopurine Synergizes with Aminoglycoside Antibiotics.

The figure presents checkerboard assays to analyze the interaction between 6-mercatopurine and the aminoglycosides kanamycin and gentamycin in *E. coli* BW25113 grown in M9 minimal medium. The synergistic interaction between 6-mercaptopurine and **(A)** kanamycin has an FICI ≤ 0.27, while **(B)** gentamycin's interaction has an FICI ≤ 0.31. See also Figure S7.

Figure 6. Analysis of the Interaction of 6-mercaptopurine with a Panel of Structurally Diverse Aminoglycosides.

The figure presents checkerboard assays in *E. coli* BW25113 grown in M9 minimal medium to study the interaction of 6-meracptopurine with the 4,5-disubstituted 2-deoxystreptamine aminoglycosides **(A)** neomycin (FICI ≤ 0.14) and **(B)** paromomycin (FICI ≤ 0.14), as well as the structurally dissimilar class of aminoglycosides **(C)** apramycin (FICI ≤ 0.15) and **(D)** spectinomycin (FICI ≤ 0.14). See also Figure S7.

Supplemental Figure Legends

Figure S1. Summary of the Results from the Systematic Combination of 45 Compounds against *E. coli* BW25113. Related to Figure 1.

Through checkerboard assays, the chemical-chemical combination of 45 compounds against *E. coli* BW25113 grown in M9 minimal medium (0.4% glucose) resulted in more than 63,000 combination wells, constituting 990 unique chemical-chemical combinations. Of the 990 combinations, 81 were synergistic and 105 were antagonistic.

Figure S2. Novobiocin is Potentiated by Two Cell Wall Inhibitors. Related to Figure 1.

In *E. coli* BW25113 grown in M9 minimal medium (0.4% glucose), novobiocin is potentiated by both **(A)** vancomycin, FICI ≤ 0.38, and **(B)** fosmidomycin, FICI ≤ 0.5.

Figure S3. MAC13772 Antagonizes D-cycloserine in *E. coli* under Nutrient-Limited Conditions. Related to Figure 3.

(A) MAC13772's inhibition of biotin biosynthesis antagonizes D-cycloserine's growth inhibitory activity in *E. coli* under nutrient stress (FICI ≥ 3). **(B)** In nutrient-rich conditions, biotin biosynthesis is no longer required, thus suppressing MAC13772's antagonism of D-cycloserine (FICI ≤ 1).

Figure S4. Summary of L-alanine Racemization for Peptidoglycan Biosynthesis in *E. coli*. Related to Figure 3.

L-alanine is converted to D-alanine, and two molecules of D-alanine are required by D-alanine-D-alanine ligase for peptidoglycan biosynthesis. D-cycloserine inhibits both enzymes. Two consecutive arrows are used to indicate that two D-alanine molecules need to be synthesized for the ligase. Dashed arrow represents more than one biosynthetic step.

Figure S5. The Synergy of lamotrigine with Antibiotics in Nutrient-Limited and Nutrient-Rich Conditions. Related to Figure 4.

In *E. coli* BW25113 grown in M9 minimal medium (0.4% glucose) at 37°C, lamotrigine was potentiated by the antibiotics **(A)** cerulenin, FICI ≤ 0.5, and **(C)** rifampicin, FICI ≤ 0.5. Both synergistic interactions persisted in nutrient-rich conditions where lamotrigine was also potentiated at 37°C by **(B)** cerulenin, FICI < 0.28, and **(D)** rifampicin, FICI ≤ 0.16.

Figure S6. Ribosome Profile Analysis of Lamotrigine and L-norleucine-treated *E. coli* BW25113 at 15°C and 37°C. Related to Figure 4.

E. coli BW25113 at 15°C **(A-D)** and 37°C **(E-G)** in the absence or presence of sub-inhibitory lamotrigine or L-norleucine concentrations. Ribosomal accumulation was monitored using UV absorbance at 260 nm. **(A, E)** Untreated *E. coli* BW25113 cells grown in M9 minimal medium (0.4% glucose) did not accumulate immature pre-30S or pre-50S ribosomal particles. **(B)** Similar results to **(A)** were observed for untreated *E. coli* BW25113 grown in LB medium. **(C)** Lamotrigine-treated *E. coli* BW25113 cells grown in M9 minimal medium (0.4% glucose) accumulated immature pre-50S ribosomal particles. **(D)** Lamotrigine-treated *E. coli* BW25113 cells grown in LB medium resulted in immature pre-30S and pre-50S ribosomal particles. **(F)** Lamotrigine-treated *E. coli* BW25113 cells grown in M9 minimal medium (0.4% glucose) accumulated some immature pre-50S ribosomal particles. **(G)** L-norleucine-treated *E. coli* BW25113 cells grown in M9 minimal medium (0.4% glucose) did not accumulate immature ribosomal particles.

Figure S7. Growth Inhibitory Activity of Thiopurine Analogues in *E. coli* BW25113. Related to Figure 5 and Figure 6.

Potency analysis of 6-mercaptopurine (circles) and 6-thioguanine (triangles) against *E. coli* BW25113 grown in M9 minimal medium (0.4% glucose). The growth inhibitory activity of both nutrient synthesis probes was determined as described in determination of minimal inhibitory concentrations, Transparent Methods.

Figures

Figure 1

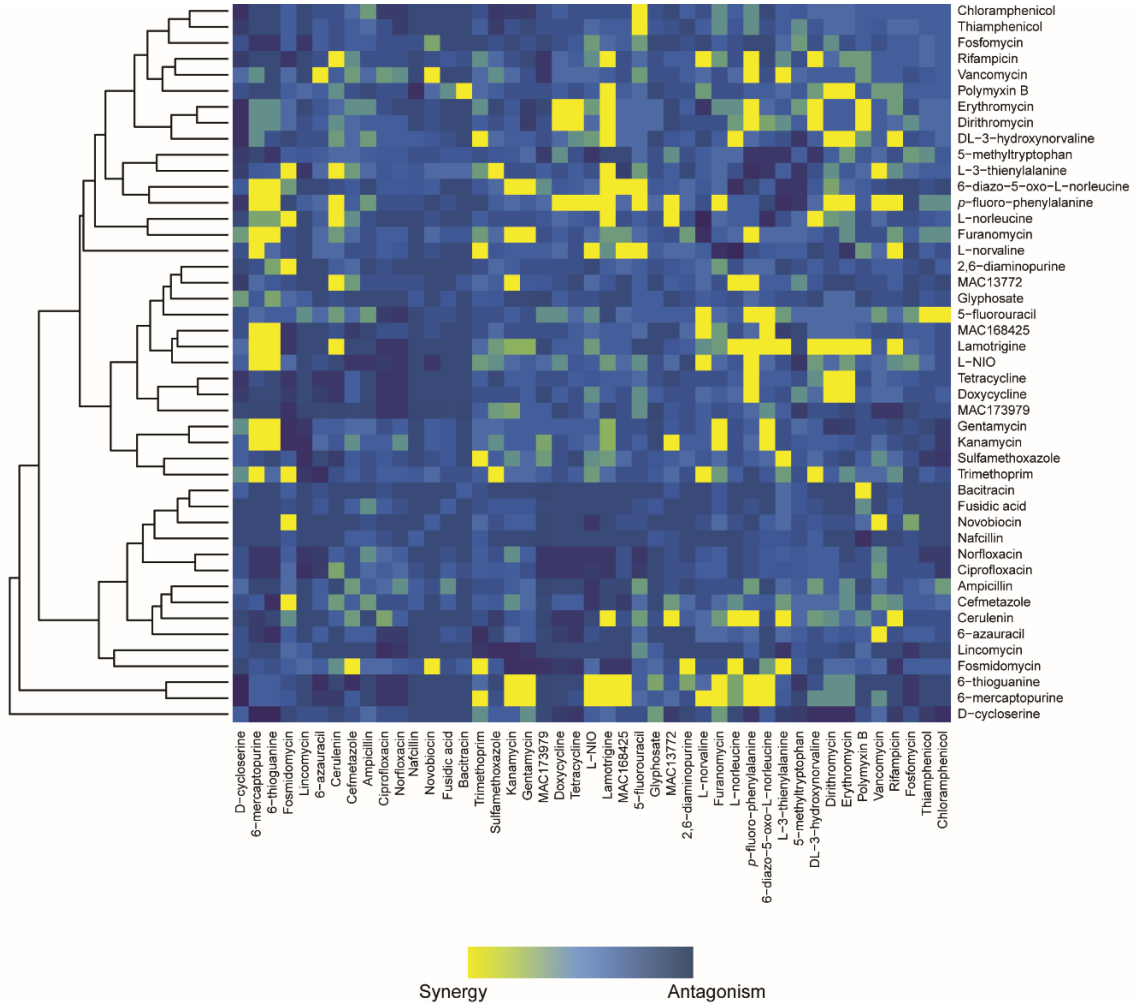


Figure 2

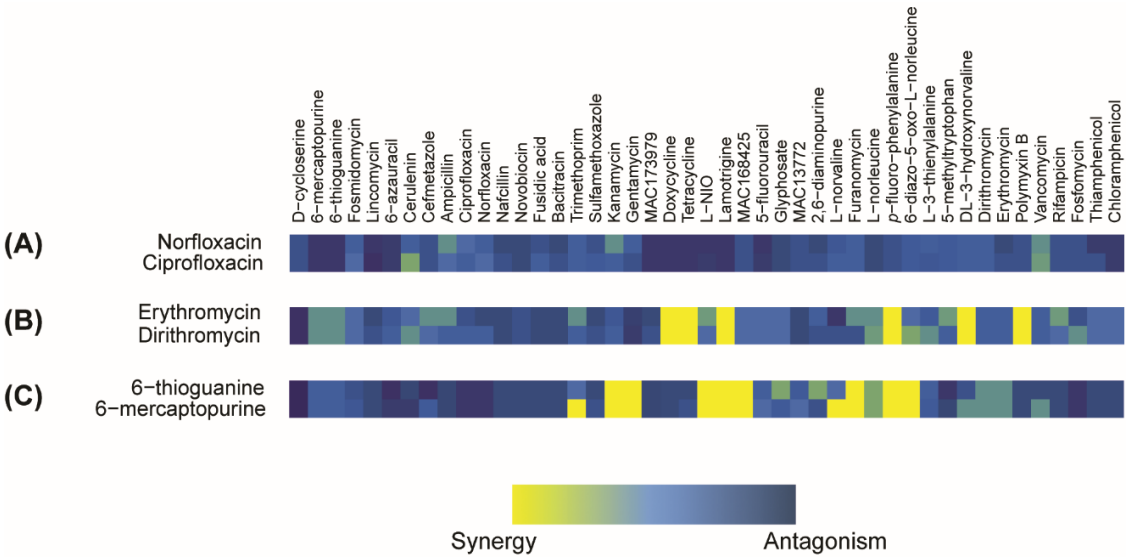

D−cycloserine
6−mercaptopurine
6−thioguanine
Fosmidomycin
Lincomycin
6−azauracil
Cerulenin
Cefmetazole
Ampicillin
Ciprofloxacin
Norfloxacin
Nafcillin
Novobiocin
Fusidic acid
Bacitracin
Trimethoprim
Sulfamethoxazole
Kanamycin
Gentamycin
MAC173979
Doxycycline
Tetracycline
L−NIO
Lamotrigine
MAC168425
5−fluorouracil
Glyphosate
MAC13772
2,6−diaminopurine
L−norvaline
Furanomycin
L−norleucine
p−fluoro−phenylalanine
6−diazo−5−oxo−L−norleucine
L−3−thienylalanine
5−methyltryptophan
DL−3−hydroxynorvaline
Dirithromycin
Erythromycin
Polymyxin B
Vancomycin
Rifampicin
Fosfomycin
Thiamphenicol
Chloramphenicol
(A)
Norfloxacin
Ciprofloxacin
(B)
Erythromycin
Dirithromycin
(C)
6−thioguanine
6−mercaptopurine
Synergy
Antagonism

Figure 3

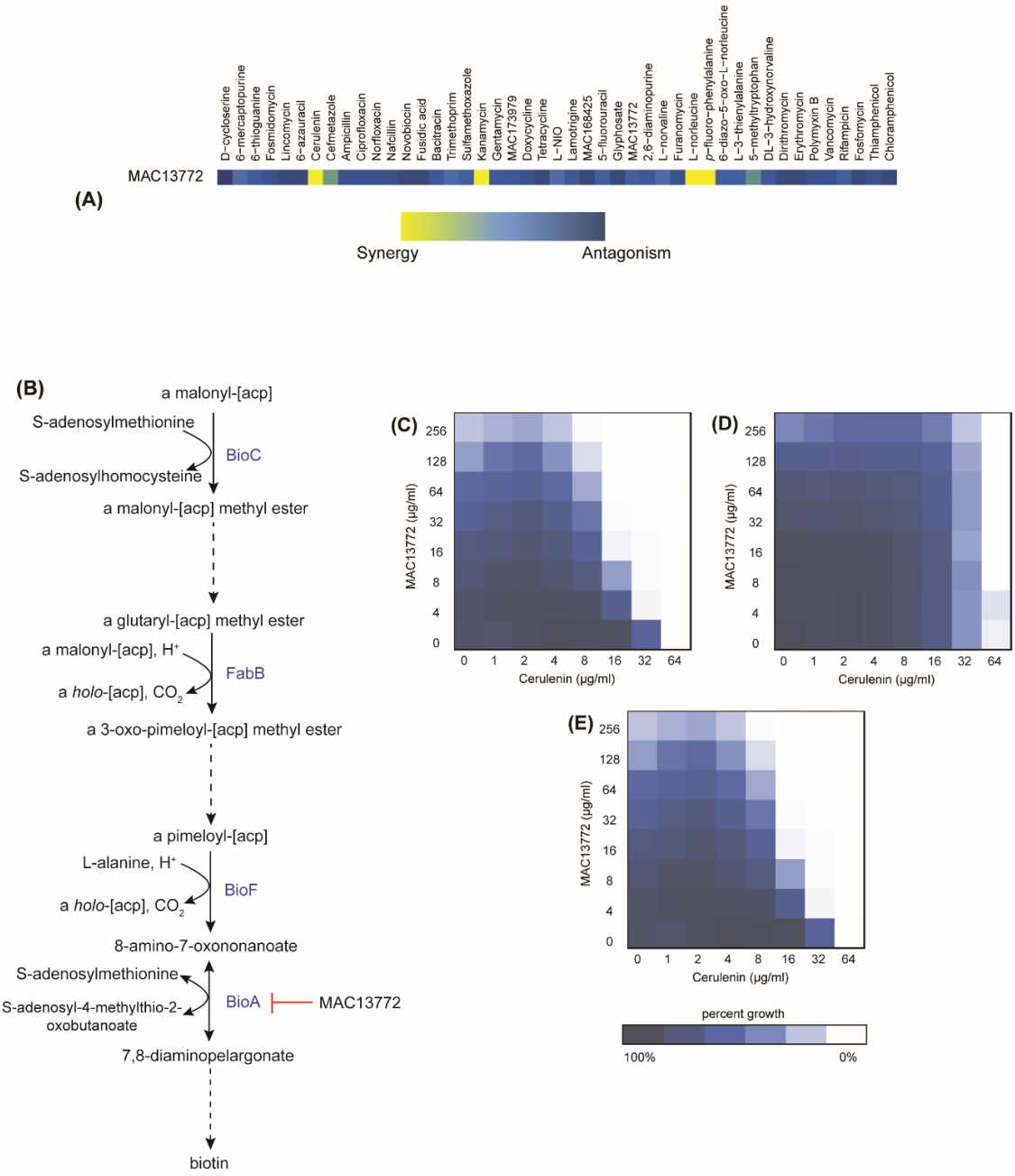

(A)
MAC13772
D-cycloserine
6-mercaptopurine
6-thioguanine
Fosmidomycin
Lincomycin
6-azauracil
Cerulenin
Cefmetazole
Ampicillin
Ciprofloxacin
Norfloxacin
Nafcillin
Novobiocin
Fusidic acid
Bacitracin
Trimethoprim
Sulfamethoxazole
Kanamycin
Gentamycin
MAC173979
Doxycycline
Tetracycline
L-NIO
Lamotrigine
MAC168425
5-fluorouracil
Glyphosate
MAC13772
2,6-diaminopurine
L-norvaline
Furanomycin
L-norleucine
p-fluoro-phenylalanine
6-diazo-5-oxo-L-norleucine
L-3-thienylalanine
5-methyltryptophan
DL-3-hydroxynorvaline
Dirithromycin
Erythromycin
Polymyxin B
Vancomycin
Rifampicin
Fosfomycin
Thiamphenicol
Chloramphenicol
Synergy
Antagonism
(B)
a malonyl-[acp]
S-adenosylmethionine
BioC
S-adenosylhomocysteine
a malonyl-[acp] methyl ester
a glutaryl-[acp] methyl ester
a malonyl-[acp], H+
FabB
a holo-[acp], CO2
a 3-oxo-pimeloyl-[acp] methyl ester
a pimeloyl-[acp]
L-alanine, H+
BioF
a holo-[acp], CO2
8-amino-7-oxononanoate
S-adenosylmethionine
BioA
MAC13772
S-adenosyl-4-methylthio-2-oxobutanoate
7,8-diaminopelargonate
biotin
(C)
MAC13772 (μg/ml)
256
128
64
32
16
8
4
0
0 1 2 4 8 16 32 64
Cerulenin (μg/ml)
(D)
MAC13772 (μg/ml)
Cerulenin (μg/ml)
(E)
MAC13772 (μg/ml)
Cerulenin (μg/ml)
percent growth
100%
0%

Figure 4

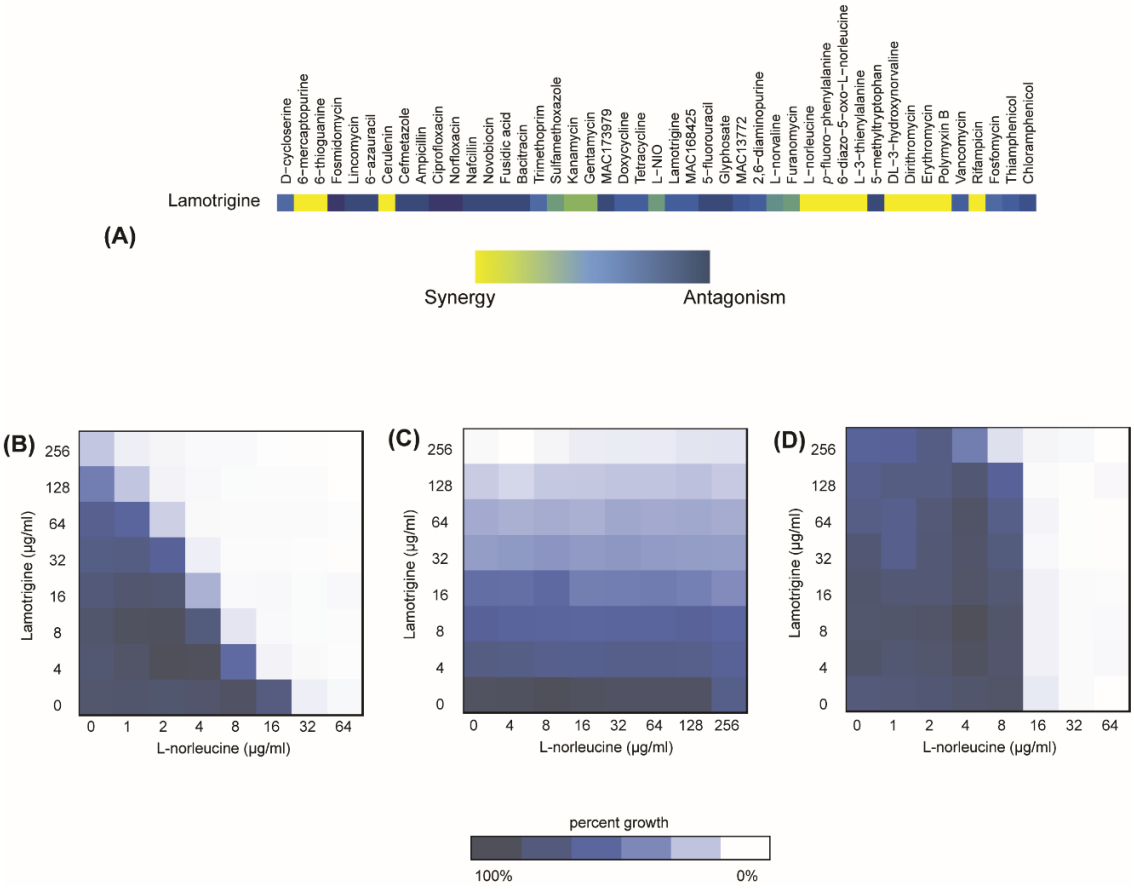

D−cycloserine
6−mercaptopurine
6−thioguanine
Fosmidomycin
Lincomycin
6−azauracil
Cerulenin
Cefmetazole
Ampicillin
Ciprofloxacin
Norfloxacin
Nafcillin
Novobiocin
Fusidic acid
Bacitracin
Trimethoprim
Sulfamethoxazole
Kanamycin
Gentamycin
MAC173979
Doxycycline
Tetracycline
L−NIO
Lamotrigine
MAC168425
5−fluorouracil
Glyphosate
MAC13772
2,6−diaminopurine
L−norvaline
Furanomycin
L−norleucine
p−fluoro−phenylalanine
6−diazo−5−oxo−L−norleucine
L−3−thienylalanine
5−methyltryptophan
DL−3−hydroxynorvaline
Dirithromycin
Erythromycin
Polymyxin B
Vancomycin
Rifampicin
Fosfomycin
Thiamphenicol
Chloramphenicol
Lamotrigine
(A)
Synergy
Antagonism
(B)
(C)
(D)
Lamotrigine (µg/ml)
L-norleucine (µg/ml)
0
4
8
16
32
64
128
256
1
2
percent growth
100%
0%

Figure 5

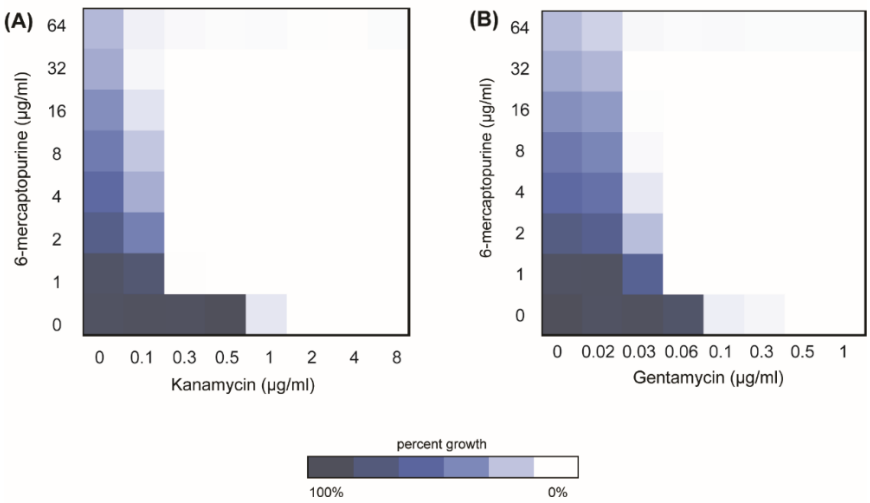
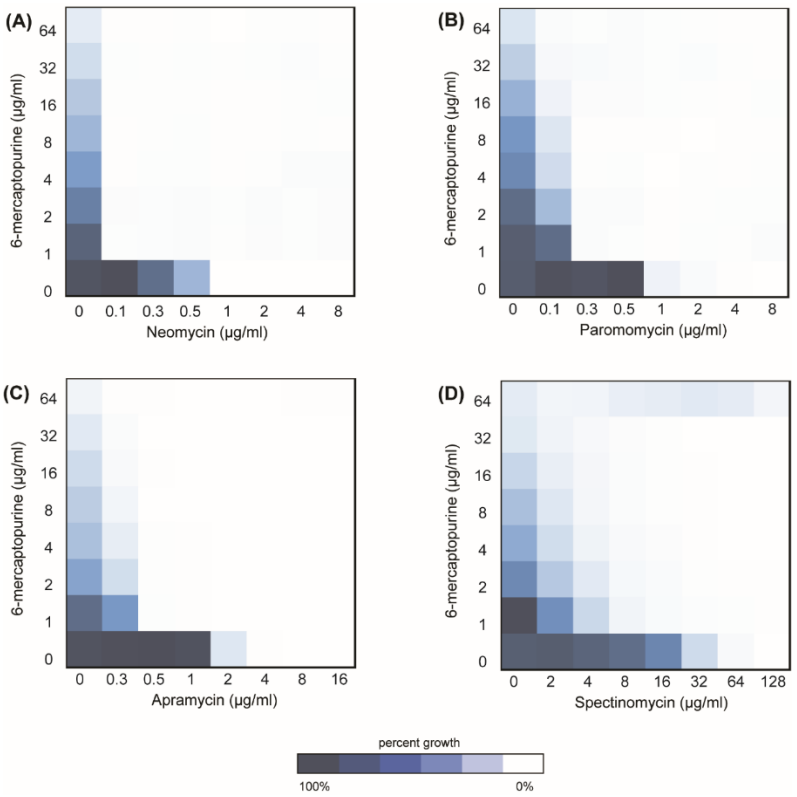


Figure 6



Tables

Table 1. Categories of Interactions among Chemical Probes of *E. coli* under Nutrient Stress.

Category	Interactions		
	Synergy	Antagonism	Total
HK probes[a]	18	39	57
NS probes[b]	22	19	41
HK probes and NS probes[c]	41	47	88
Total	81	105	186

[a]HK probes, housekeeping-function probes, refer to 27/45 compounds that target housekeeping functions (Supplemental Data 1). Interactions belonging to the HK probes category involve 49 antibiotic-antibiotic and 8 lamotrigine-antibiotic combinations.

[b]NS probes, nutrient synthesis probes, refer to 18/45 compounds that target nutrient synthesis functions (Supplemental Data 1). Interactions belonging to the NS probes category involve combinations between these 18 nutrient synthesis probes.

[c]Interactions belonging to the HK probes and NS probes category involve interactions between a housekeeping-function probe and a nutrient synthesis probe.

Supplemental Figures

Figure S1

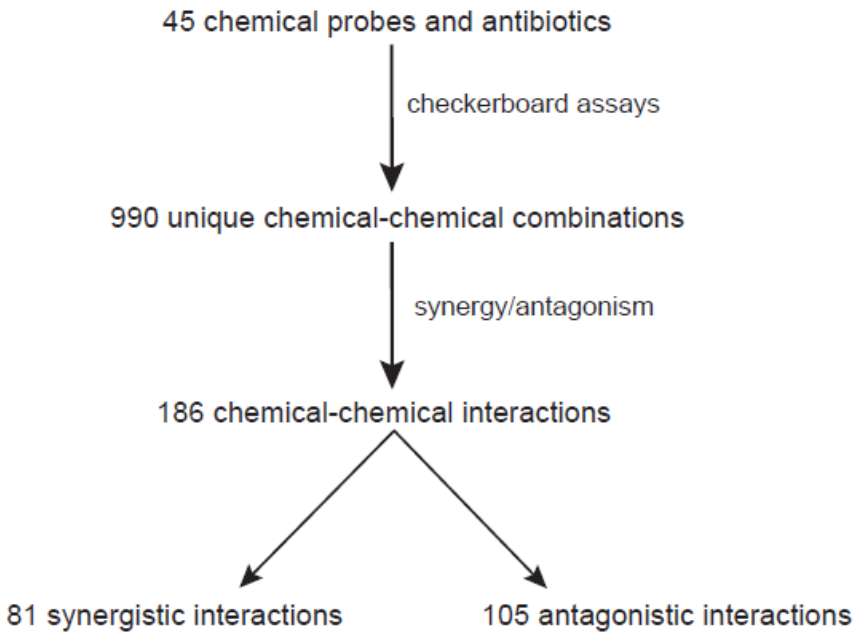


Figure S2

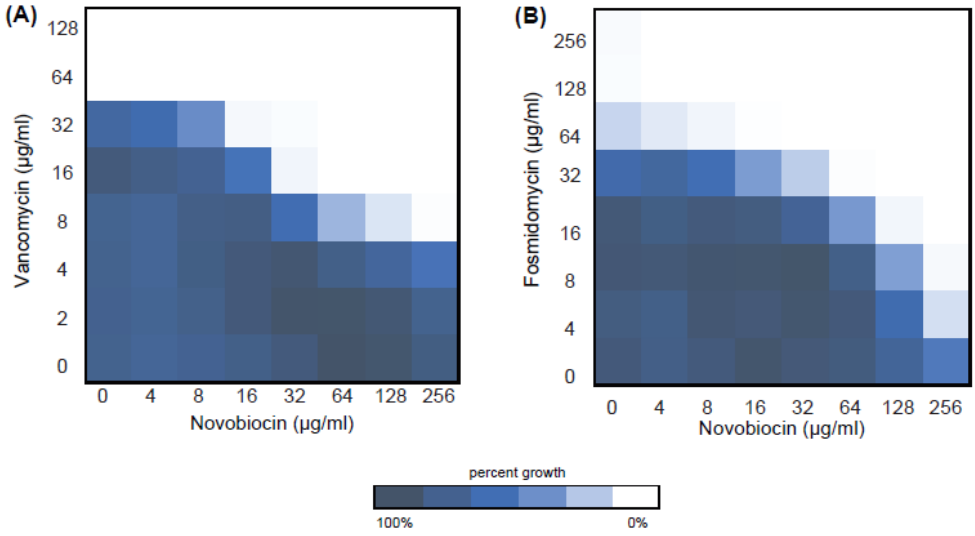


Figure S3

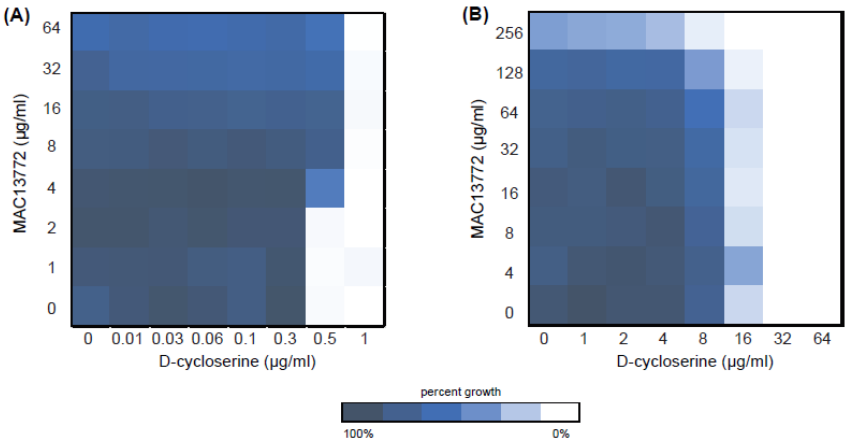


Figure S4

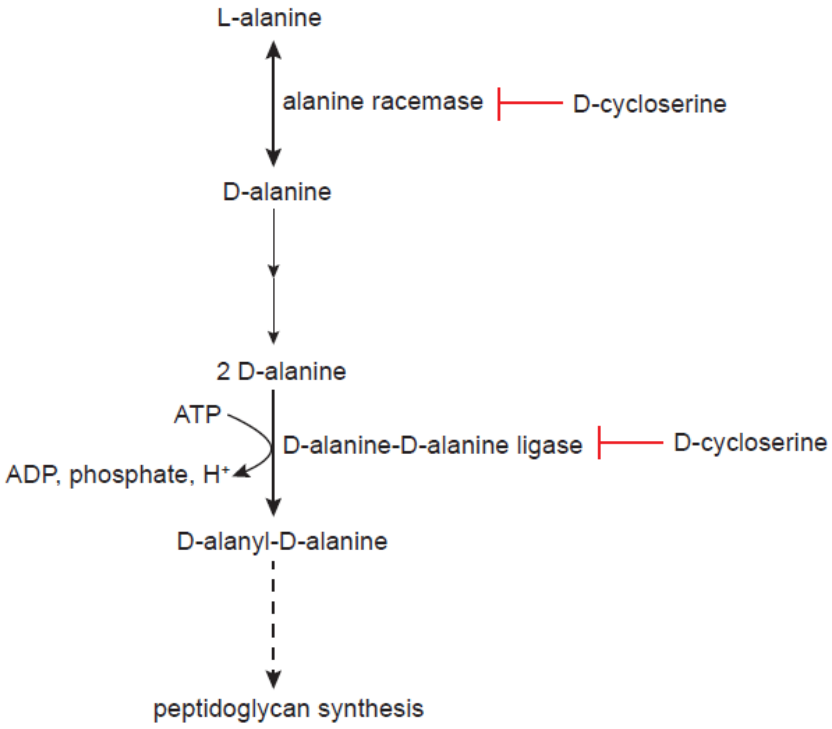


Figure S5

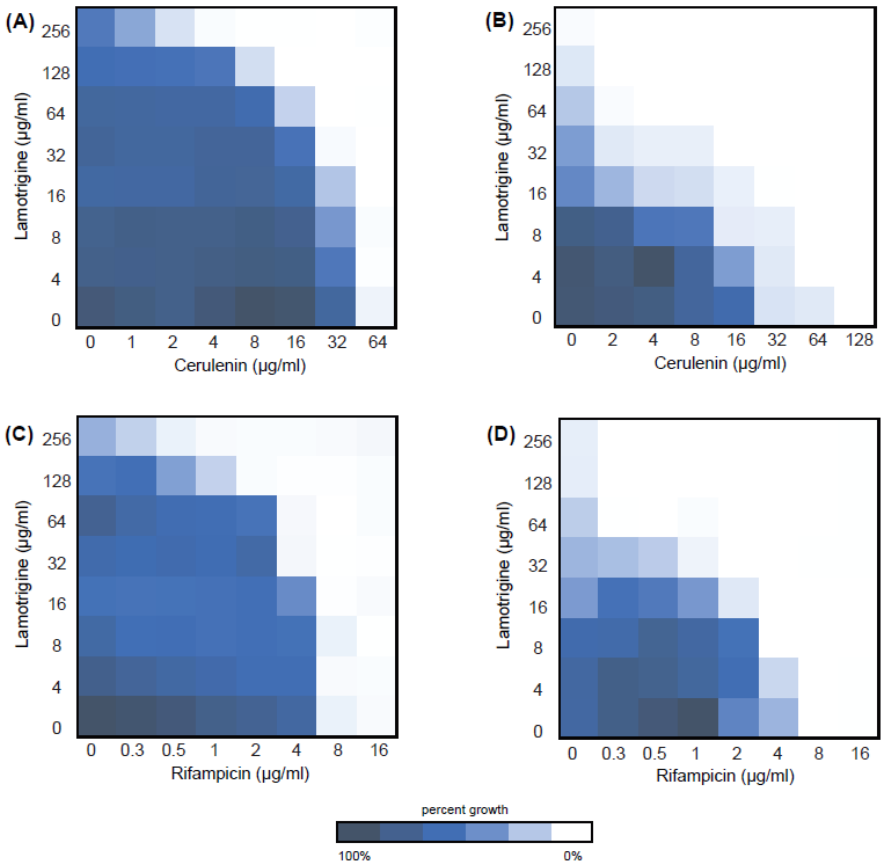

(A)
(B)
(C)
(D)
Lamotrigine (µg/ml)
Cerulenin (µg/ml)
Rifampicin (µg/ml)
256
128
64
32
16
8
4
0
0 1 2 4 8 16 32 64
0 2 4 8 16 32 64 128
0 0.3 0.5 1 2 4 8 16
percent growth
100%
0%

Figure S6

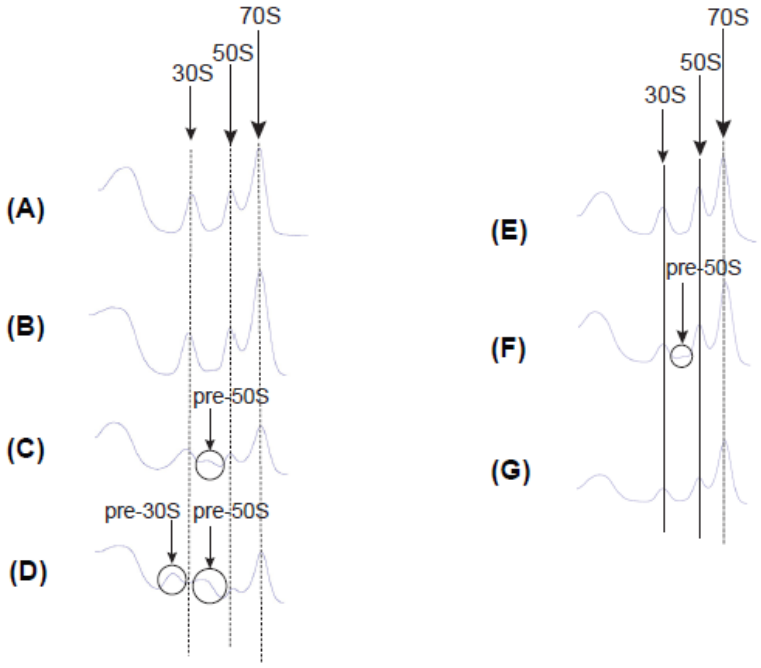
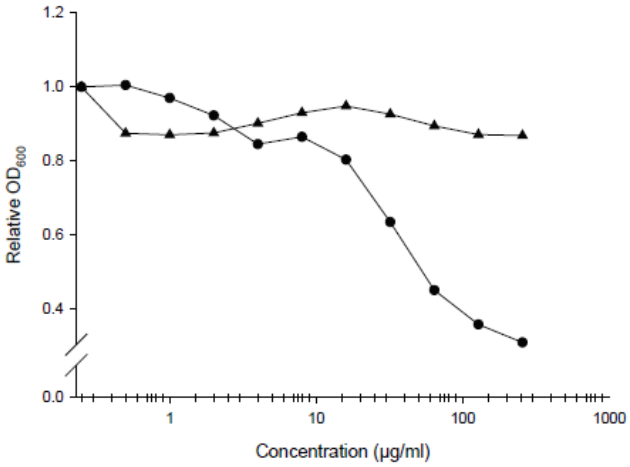


Figure S7



CHAPTER III - Physicochemical and structural parameters contributing to the antibacterial activity and efflux susceptibility of small molecule inhibitors of *Escherichia coli*

Preface

The work presented in this chapter is in preparation for submission, as of August 25, 2020:

El Zahed, S. S., French, S., Farha, M. A., Kumar, G., & Brown, E. D. Physicochemical and structural parameters contributing to the antibacterial activity and efflux susceptibility of small molecule inhibitors of *Escherichia coli*.

In preparation.

For this work, I performed all *in vitro* experiments with assistance from Farha, M.A. for the primary screen and Kumar, G. for the dose-response. I analyzed the primary screening data, dose-response curves, and structure-activity relationship. French, S. calculated molecular descriptors, developed the code for the machine-learning models, and generated the structure-activity representation. I wrote the manuscript, with significant input from French, S. Edits were provided by Brown, E.D.

Abstract

Discovering new Gram-negative antibiotics has been a challenge for decades. This has been largely attributed to a limited understanding of the molecular descriptors governing Gram-negative permeation and efflux evasion. In the work presented here, we address the contribution of efflux using a novel approach that applies multivariate analysis, machine learning, and structure-based clustering to some 4,500 actives from a small molecule screen in efflux-compromised *Escherichia coli*. We employed principal component analysis and trained two decision tree-based machine learning models to investigate descriptors contributing to the antibacterial activity and efflux susceptibility of these actives. This approach revealed that the Gram-negative activity of hydrophobic and planar small molecules with low molecular stability is limited to efflux-compromised *E. coli*. Further, molecules with reduced branching and compactness showed increased susceptibility to efflux. Given these distinct properties that govern efflux, we developed the first machine learning model, called Susceptibility to Efflux Random Forest (SERF), as a tool to analyze the molecular descriptors of small molecules and predict those that could be susceptible to efflux pumps *in silico*. Here, SERF demonstrated high accuracy and good predictive power in identifying such small molecules. Further, to examine the structural underpinnings contributing to efflux, we clustered all 4,500 actives based on their core structures. This enabled us to identify distinct clusters highlighting some side chain moieties that cause marked changes in efflux susceptibility. Such analysis provided a proof of principle for the potential of exploiting side chain modification to evade efflux pumps. In all, our work reveals a role for physicochemical and structural parameters in

governing efflux, presents a machine learning tool for rapid *in silico* analysis of efflux susceptibility, and highlights some side chain moieties that could be used to design novel antimicrobials evading efflux pumps.

Introduction

Remarkably, no truly new Gram-negative antibiotics have been discovered since the quinolones, dating back to the early 1960's[1]. In the same period, Gram-negative superbugs resistant to existing antibiotics have become a pervasive public health concern[2]. To address the growing need for new antibiotics, the World Health Organization (WHO) published a list of bacterial pathogens to guide research and development, highlighting Gram-negative pathogens as a critical priority[3]. The Centers for Disease Control and Prevention recently revealed that Gram-negative pathogens were responsible for ~50% of antibiotic-resistant microbial infections in the United States[4]. A recent analysis of the clinical pipeline revealed 50 antibiotics in development, of which only 12 are active against some of the priority Gram-negative pathogens identified by the WHO[5, 6]. Of these 12 agents, only murepavadin is considered new, with a novel pharmacophore, target, and mode of action. Most recently, and unfortunately, nephrotoxicity concerns have halted its development[5, 6]. Thus, the clinical pipeline is currently devoid of new chemical matter to treat the most troublesome infections caused by Gram-negative pathogens. Herein, we examine one of the most important impediments to identifying leads for new antibiotics targeting Gram-negative bacteria, namely, the physicochemical and structural parameters that delineate compound efflux.

It is widely accepted among pundits of antibacterial drug discovery that intrinsic resistance mechanisms, the outer membrane barrier and active efflux pumps, have presented the greatest challenge to the development of new Gram-negative antibiotics[7-9]. Indeed, there has been a growing number of calls for research to understand the

physicochemical properties that facilitate compound entry and efflux avoidance[10-12]. Inspiration comes from success correlating a drug's physicochemical properties with its pharmacokinetics, where Lipinski mapped an ideal physicochemical space for orally available drugs, establishing the rule of five: molecular weight (MW <500 g mol^{-1}), lipophilicity (clogP ≤5), and the number of hydrogen bond donors (≤10) and acceptors (≤5)[13]. Several analyses have shown, nevertheless, that antibacterial compounds are exceptional in this context. O'Shea and Moser showed that Gram-negative antibiotics, on average, are slightly larger (MW 414 g mol^{-1}) and more hydrophilic (clogD -2.8) than non-antibiotic drugs (338 g mol^{-1}; clogD 1.6)[14]. When Gram-negative antibiotics were classified by their target location, those with a cytoplasmic target were smaller (254-465 g mol^{-1}) and more hydrophobic (clogD -1.4 to 1.1) than antibiotics with a periplasmic target (347-558 g mol^{-1}; clogD -5.1 to -1)[15]. Although these studies identify an ideal physicochemical space for Gram-negative activity[14, 15], they include antibiotics that are susceptible to efflux pumps of the resistance-nodulation-division (RND) superfamily, composed of an inner membrane transporter, a periplasmic adaptor protein, and an outer membrane channel[16, 17 18]. These tripartite pumps extrude a broad range of molecules from the inner membrane and periplasmic space to the extracellular milieu[18, 19]. In an analysis of several high-throughput screening campaigns at AstraZeneca, Brown *et al.* showed that small (<300 g mol^{-1}) and hydrophilic (clogD <0) or very large (>700 g mol^{-1}) and zwitterionic compounds were least susceptible to RND efflux pumps[20]. Compared to Moser's analysis of the physicochemical space occupied by Gram-negative antibiotics[15], these results suggest that antimicrobials with a cytoplasmic target may be more

susceptible to efflux than others. In all, these initial studies highlight the complexity of designing new Gram-negative antibacterial agents that evade efflux pumps. To overcome this challenge, physicochemical and structural guidelines for Gram-negative efflux are greatly needed.

In the work described herein, we have taken a novel approach that applies multivariate analysis, machine learning, and structure-based clustering of growth inhibitors of *Escherichia coli* to identify molecular descriptors and structure-activity relationships that impact efflux susceptibility. The effort defines new physicochemical and structural parameters that contribute to the antibacterial activity and efflux susceptibility of small molecules targeting this model Gram-negative bacterium.

Results

A screen for growth inhibitors of an efflux-deficient strain of *E. coli*

Our work began with a high-throughput screen to identify compounds with growth inhibitory activity in the efflux-deficient stain *E. coli* Δ*tolC*. The screen comprised ~314,000 molecules, which were largely synthetic, and included a collection of ~3,900 previously approved drugs and bioactive molecules (bioactive collection). The data were reproducible and defined a distinct collection of growth inhibitors in the primary screen (Figure 1A). Indeed, ~4,500 actives exhibited strong growth inhibition (≤3σ below the mean, Figure 1B), of which 386 were from the bioactive collection. The potency (EC_{50}) of each active was measured in wild-type *E. coli* and its mutant strain Δ*tolC* (Figures 1C and 1D, Table S1). Of the 4,500 actives, a large fraction, some 84% (3,780/4,500),

showed little activity at the highest concentration (50 μM) tested using wild-type *E. coli* (Figure 1C), while the EC_{50} using the Δ*tolC* strain was less than 35 μM (Figure 1D). This suggested that many of our actives were susceptible to efflux, where their antibacterial activity was abolished in the presence of efflux pumps. Among these molecules were antibiotics from the bioactive collection, which include β-lactams, macrolides and other compounds, such as doxorubicin, known to have antibiotic activity and susceptibility to efflux, (Table S1). This was an encouraging result from a data quality perspective given the previous literature on their susceptibility to efflux pumps in Gram-negative bacteria[19, 21-23].

Hydrophobicity, level of saturation, and molecular stability contribute to Gram-negative activity in efflux-compromised *E. coli*

Next, we sought to identify the molecular descriptors for this set of 3,780 efflux-dependent active compounds. To this end, we calculated some 50 descriptors for all 314,000 compounds screened, which include efflux-dependent actives (Table S1), resulting in more than 15 million data points (Figure 2A). To reduce the dimensionality of this data, we initially applied a principal component analysis (PCA). The first three principal components of the PCA explained ~63% of the variances between the molecular descriptors of efflux-dependent actives and those of non-growth inhibitory (inactive) molecules in Δ*tolC* (Figure 2B). Further, the PCA identified a structural and physicochemical ‘pocket’ that defines the antibacterial activity of the 3,780 efflux-dependent actives (Figure 2C). The clustering of these actives near the origin (Figure 2C)

was largely due to molecular size, hydrophobicity, and molecular complexity, as observed by the eigenvectors in principal component 1 (PC1) and principal component 2 (PC2, Figure 2D). Notably, PC1 uniquely comprised of molecular size, which was represented by molecular weight (MW), the total atom count, and the solvent accessible surface area (ASA), as calculated by using the radius of the solvent. Particularly, ASA contained components of the surface area of partial positive charge (ASA+) and partial negative charge (ASA-). PC2, on the other hand, consisted of both hydrophobicity and molecular complexity. The former was represented by clogD (pH 7.0) and the ASA of polar atoms (ASA_P) and hydrophobic atoms (ASA_H), while the latter was described by the number of aliphatic atoms and the fraction of sp^3 hybridized carbon atoms (Fsp^3). Overall, the PCA results suggest that combinations of these nine molecular descriptors addressing molecular size, hydrophobicity, and molecular complexity contribute to Gram-negative activity in Δ*tolC*.

Given that the PCA highlighted some distinct molecular descriptors governing the antibacterial activity of efflux-dependent actives, we implemented a tree-based machine learning approach to quantify descriptors that restrict this activity to efflux-compromised *E. coli*. Herein, we developed a random forest classification model for the set of 3,780 efflux-dependent actives and 3,780 inactive molecules, randomly chosen from the primary screen (Figure 3A). To eliminate bias from the PCA, all 50 molecular descriptors for each of these molecules were used to train the model, which achieved an area under the curve-receiver operating characteristic curve (AUC-ROC) of 0.808 (Figures 3B and S1). This illustrates a good measure of the model's performance in classifying molecules

with antibacterial activity in Δ*tolC* and those without antibacterial activity. The analysis revealed similar descriptors to that of the PCA (Figures 2D and 3C); indeed, the model indicated that clogD (pH 7.0), Fsp^3, and the resonant structure count of a molecule were the top three factors driving Gram-negative activity of small molecules in efflux-compromised *E. coli* (Figure 3C). Particularly, efflux-dependent actives were hydrophobic (clogD 1-5), planar (Fsp^3 0-0.5), and had low molecular stability (resonant structure count <4, Figure S2). Since these actives lost their antibacterial activity in efflux-proficient wild-type *E. coli*, these results suggest that the Gram-negative activity of molecules with such properties would be abolished due to efflux pumps. Of note, in the absence of clogD from the 50 molecular descriptors used to train this model, the model lost accuracy in classifying molecules with antibacterial activity in Δ*tolC* and those without activity (Figure 3C). This highlights the importance of hydrophobicity as a physicochemical property that affects Gram-negative activity, regardless of efflux pumps.

A random forest model predicts efflux-prone small molecules

Consistent with the goal to understand efflux susceptibility in *E. coli*, we focused our analysis on determining key molecular descriptors that contribute to efflux susceptibility. First, we categorized the 4,500 actives into efflux susceptible (pumped) molecules and unsusceptible (non-pumped) molecules based on their potency and calculated fold-change in potency (wild-type EC_{50}/Δ*tolC* EC_{50}). Actives classified as pumped molecules were potent in *E. coli* Δ*tolC* (EC_{50} <5 μM) and lacked antibacterial activity in wild-type (EC_{50} >50 μM), while those classified as non-pumped molecules showed insignificant changes in potency between both strains (fold-change ≤2, Figure 4A). The former set of molecules

consisted of ~1,070 actives, while the latter set constituted ~410 actives. Of these 1,480 actives, a random set of ~290 pumped molecules and ~290 non-pumped molecules were chosen as a training set, where each molecule's 50 descriptors were used to train a second random forest model, called Susceptibility to Efflux Random Forest (SERF, Figure 4A). This model achieved an AUC-ROC of 0.839 (Figures 4B and S3), indicating a good measure of classifying pumped molecules and non-pumped molecules. Here, we observed that along with hydrophobicity and molecular stability, molecular complexity was key to efflux susceptibility (Figure 4C). Molecular complexity was represented by topological indices, namely the hyper-Wiener index, Wiener polarity, and the Balaban index, that describe some aspects of molecular structure[24]. Notably, the three descriptors largely governing efflux susceptibility were resonant structure count, clogD (pH 7.0), and the hyper-Wiener index, which describes molecular "branching" and "compactness"[24, 25]. Molecules with a relatively low molecular stability (resonant structure count ≤6), hydrophobic nature (clogD 1-5), and reduced branching and compactness (hyper-Wiener index >6,000) showed increased susceptibility to efflux (Figure S4).

Following model development and optimization, we then assessed the predive power of SERF to identify molecules susceptible to efflux. We used a test set of ~440 actives (of 1,480) that were excluded from the training set, of which 274 were predicted by SERF to be efflux susceptible. We curated these 274 molecules and assessed their empirically measured potency in wild-type *E. coli* and Δ*tolC* (Figure 5A, Table S2). Based on the chosen cut-off for pumped molecules (EC_{50} <5 μM and wild-type EC50 >50 μM), we observed that 260 of 274 (~95%) predicted molecules were validated, illustrating a high

accuracy in the model's prediction (Figure 5A). To determine the predictive power of SERF, we further assessed the potency of the remaining 166 molecules (of 440), in wild-type and Δ*tolC*. Here, we noted 61 pumped molecules that were not reported by the model, revealing an overall 80% (260/321) predictive power in identifying efflux-susceptible molecules (Figure 5B, Table S3). In all, our analysis revealed that molecular stability, hydrophobicity as well as molecular branching and compactness contribute to efflux susceptibility. Further, we designed the first random forest model (SERF) that analyzes the molecular descriptors of small molecules and identifies those susceptible to efflux. The code for SERF can be accessed online to test other small molecules for efflux susceptibility (https://github.com/sfrench007/serf).

Structure-activity relationship analyses reveal some structural modifications that impact efflux susceptibility

To analyse the chemical space occupied by our 4,500 actives, we applied structure-based clustering using structural fingerprints. This revealed ~15-20 self-organized clusters, consisting on average of 15 or more molecules within a structural space (Figure 6A). These clusters demonstrated distinct activity cliff regions that highlight how structural modifications within a compound series of molecules led to significant changes in biological activity[26]. Some clusters were occupied by molecules known for their susceptibility to efflux, such as β-lactams and antifungals (Table 1, Table S4). Notably, the β-lactam cluster consisted of eleven penicillin-type antibiotics that further divided into five semisynthetic β-lactamase-resistant penicillins (**1a-4a**, **6a**), four aminopenicillins (**8a**,

9b, **10c**, **11a**), and two ureidopenicillins (**5a**, **7a**) (Table 1). We assessed the local structure-activity relationship (SAR) of this β-lactam cluster, where we observed, as expected, that the penam was the common core among these penicillins (Figure 6B, Table 1). Chemical modification of the acylamino side chain at the α-carbon of the penam largely contributed to this activity cliff (Table 1). Specifically, the amide group was decorated with a phenyl-isoxazolyl (**1a-4a**), naphthyl (**6a**), or benzylamine (**5a**, **7a**, **8a**, **9b**, **10c**, **11a**, **12a**) moiety. The high fold-change in the potency of penicillins **1a-4a** and **6a** suggested that both phenyl-isoxazolyl and naphthyl moieties contribute to efflux susceptibility (Table 1). Substituting these moieties with more polar benzylamine groups, on the other hand, established antibacterial activity for most of the remaining penicillins in wild-type (Table 1). This lowered the fold-change in potency and the penicillin's susceptibility to efflux (Table 1). Furthermore, compared to penicillins **1a-4a** and **6a**, the lower hydrophobic character (clogP), an overall decrease in planar, or "flat", nature (Fsp^3), and increased molecular stability (RSC) of penicillins **8a**, **9b**, **10c**, and **11a** correlated with reduced efflux susceptibility. Notably, where the resonant structure count for penicillins **7a**, **9b**, and **11a** remain constant, increased molecular compactness, as observed by lower hyper-Wiener indices (HWI, Table 1), reduces efflux susceptibility as well. In all, these results suggest that the acylamino group at the α-carbon of penams and the overall hydrophobicity and reduced molecular complexity of penicillins contribute to their efflux susceptibility in *E. coli*.

Additionally, we identified a cluster occupied by six hydroxyquinoline derivatives (Figure 6C), where the hydroxyquinoline core was decorated with different combinations

of halogen atoms on its phenol and pyridine moieties (Table 2). Although these chemical changes seem minor, addition of a chlorine atom to the ortho and para positions of the phenol in compounds **12** and **13** abolished antibacterial activity in wild-type and maintained efflux susceptibility (Table 2). Reducing the number of chlorine atoms to one or none at these positions, however, established wild-type activity and decreased efflux susceptibility of compounds **14-17** (Table 2). Here, these modifications revealed that the addition of halogen atoms with low electronegativity to the phenol moiety improved whole-cell activity and reduced susceptibility to efflux (Table 2). Furthermore, decreasing the hydrophobic character (clogP <3) reduced efflux susceptibility of compounds **15-17** (Table 2). Overall, these results highlighted modifications to the aromatic ring of hydroxyquinolines that affect their susceptibility to efflux.

The SAR and fold-changes in potency available within the clusters highlighted above revealed that some chemical modifications impact both whole-cell activity and efflux susceptibility in *E. coli*. As such, structure-based clustering has the potential to provide some guidelines for optimal antibacterial activity and efflux evasion.

Discussion

There is a need to understand compound penetration in Gram-negative bacteria, which consist of both outer membrane permeability and efflux evasion. There has been much progress in recognizing the properties that contribute to penetration of large chemical collections[14, 15, 20]; however, the exact contribution of outer membrane permeability and efflux remains unclear. Herein, we addressed the contribution of efflux by using novel

computational approaches to investigate the molecular descriptors of some 4,500 actives from a small molecule screen in efflux-compromised *E. coli*. We applied principal component analysis to visualize descriptors contributing to Gram-negative activity in the efflux-deficient strain *E. coli* Δ*tolC*, and quantitatively assessed them using a machine learning approach. In so doing, we designed a machine learning model in order to identify descriptors governing efflux susceptibility of small molecules. The latter, referred to as Susceptibility to Efflux Random Forest (SERF), was further assessed for its predictive power to analyze the molecular descriptors of a set of small molecules and identify those susceptible to efflux pumps *in silico*. Notably, the model showed ~95% accuracy in its prediction and had an overall predictive power of 80%. In all, this work presents the first large-scale study implementing PCA and machine learning in order to further resolve the contribution of efflux to compound penetration. Further, it provides the first machine learning tool (SERF) that identifies small molecules susceptible to efflux pumps *in silico*.

In this study, we identified some 4,500 actives from a phenotypic screen in the efflux-deficient strain *E. coli* Δ*tolC*. Dose-response potency analyses of these actives in wild-type *E. coli* and the mutant strain Δ*tolC* revealed that a majority (~84%) lost antibacterial activity in wild-type (efflux-dependent actives), suggesting high efflux susceptibility. Among these compounds were conventional antibiotics, including β-lactams and macrolides, and other non-antibiotic drugs known to have cryptic antibacterial activity. These findings further validated our work since these compounds have been well-recognized as efflux substrates in Gram-negative bacteria [19, 21-23].

A PCA of some 50 molecular descriptors for each efflux-dependent active and non-growth inhibitory molecules from the primary screen highlighted nine molecular descriptors that contribute to Gram-negative activity in efflux-compromised *E. coli*. These descriptors largely encompassed measures of molecular size, hydrophobicity, and molecular complexity. To quantify properties that contribute to Gram-negative activity and those that influence efflux susceptibility, we designed two random forest models. The first model addressed Gram-negative activity, highlighting that the antibacterial activity of hydrophobic and planar molecules with relatively low stability was limited to efflux-compromised *E. coli*. The second random forest model (SERF) assessed efflux susceptibility, revealing that hydrophobic molecules with reduced branching, compactness, and stability were most susceptible to efflux pumps. Overall, both random forest models suggest that hydrophobicity and molecular stability impact Gram-negative activity as well as efflux susceptibility. Further, each model identified descriptors of molecular complexity that impact each of these aspects, where the level of saturation (Fsp^3) governs antibacterial activity, potentially due to some influence from efflux pumps, while branching and compactness (hyper-Wiener index) uniquely affect efflux susceptibility.

The hydrophobic nature of efflux-susceptible Gram-negative antibacterials has been recognized by Brown *et al.*[20]. Recent analyses of the physicochemical properties of some β-lactams, fluoroquinolones, and a subset of other antibiotics have shown that lipophilicity contributes to efflux susceptibility[17]. Furthermore, O'Shea and Moser[14] have shown that Gram-negative antibiotics are more hydrophilic than other drugs. These

studies validate our analyses, where we observed that the Gram-negative activity of hydrophobic molecules was abolished in efflux-proficient *E. coli*. The impact of molecular stability and molecular complexity on efflux susceptibility, however, provides novel insight into compound penetration. Particularly, planar, or "flat", molecules with low molecular stability lost their Gram-negative activity in efflux-proficient *E. coli*, while "unbranched" and "elongated" molecules with low molecular stability showed increased susceptibility to efflux pumps. These molecular descriptors highlight ideal efflux substrates that may result in optimal binding to AcrB, for example, and extrusion. Indeed, AcrAB-TolC is a well-characterized RND efflux pump in *E. coli*[27]. This multiple-component system consists of an outer membrane channel, TolC; a periplasmic adaptor protein, AcrA; and an inner membrane transporter, AcrB[18]. Co-crystallization of AcrB with some molecules revealed that it undergoes conformational cycling in order to bind and extrude molecules through TolC[19, 22]. Hydrophilic molecules tend to bind to the upper "crevice" of the distal pocket, which is rich in hydrophilic and charged residues[27-29]. On the other hand, hydrophobic molecules interact with the phenylalanine-rich hydrophobic trap located in the lower portion of the binding pocket[22, 30, 31]. Molecular dynamics simulation studies have suggested that a tight interaction with the hydrophobic trap distorts the binding pocket and inhibits efflux, while a loose interaction facilitates binding and extrusion[27, 29, 32]. As such, our analyses suggest that hydrophobic molecules that are planar, unbranched, and elongated may be ideal candidates for loose interactions with the hydrophobic trap in AcrB. Furthermore, decreasing the molecular stability of such

molecules increases their reactivity, which may facilitate non-specific binding and extrusion.

Additionally, we used the machine learning model SERF to predict, *in silico*, which antibacterials would be susceptible to efflux pumps. For a set of some 440 actives, SERF predicted that 274 molecules would be susceptible. Following empirical validation, we found that the prediction was ~95% accurate, and the model showed good (80%) predictive power in identifying molecules susceptible to efflux. Overall, this highlights the potential of machine learning approaches to predict which compounds could be susceptible to efflux. Where novel approaches to study and overcome efflux are greatly required[10, 11, 33], SERF can be used as a tool to rapidly explore efflux susceptibility of small molecules *in silico*. As such, our machine learning model provides an opportunity to systematically develop and synthesize a novel library of Gram-negative antibacterial agents that could evade efflux pumps.

As for the structure-based clustering of the 4,500 actives, this analysis revealed ~15-20 self-organized clusters, consisting of 15 or more molecules, on average, within a self-organized structural space. We profiled the SAR within a cluster of β-lactams and hydroxyquinoline derivatives with uncharacterized modes of action. The β-lactam cluster consisted of penicillin-type antibiotics, including some five β-lactamase resistant penicillins, four aminopenicillins, and two ureidopenicillins. Modifications to the acylamino moiety of these penicillins impacted their efflux susceptibility, where the phenyl-isoxazolyl and naphthyl moieties of β-lactamase resistant penicillins showed high susceptibility to efflux. Substituting these moieties with polar benzylamine groups as well

as decreasing the hydrophobicity of the antibiotic reduced efflux susceptibility of some aminopenicillins and ureidopenicillins. Moreover, increasing the molecular complexity of these antibiotics by reducing their "flat" nature as well as increasing their molecular stability and compactness allows for better efflux evasion. Overall, this analysis highlights that chemical substitutions along the acylamino group and changes in hydrophobicity and molecular complexity may be key to efflux susceptibility for penicillins. Indeed, earlier modifications of penicillins have shown that the β-lactamase-resistant penicillins are insensitive to hydrolysis by β-lactamases and susceptible to efflux[34-37]. Further modifications largely focused on substitutions to the acylamino moiety[34, 38], which yielded the aminopenicillins and ureidopenicillins. Indeed, these penicillins have an extended activity spectrum that encompasses some *Pseudomonas* spp.[38, 39]. Although some remain susceptible to efflux, these β-lactams have shown improved Gram-negative activity in efflux-proficient strains[21, 39]. Of practical note, acylamino groups decorated with benzylamines could be used in medicinal chemistry efforts to assess for reduced efflux in other compound series.

The second cluster was composed of hydroxyquinoline derivatives exhibiting chemical modification to their aromatic rings. Substitution of chlorines (compounds **12** and **13**) with halogen atoms decreasing in electronegativity (compounds **14**-**17**) largely improved wild-type activity and decreased efflux susceptibility. A lower hydrophobic character within this compound series, also, seemed to contribute to improved activity in wild-type and reduced efflux. While hydrophobicity has been recognized for its impact on Gram-negative activity and efflux susceptibility[14, 15, 20], the hydroxyquinoline cluster revealed

that decorating aromatic rings with different halogen atoms may lead to some changes in efflux susceptibility. In previous studies, compounds containing halogenated aromatic moieties have been shown to improve antibacterial activity in Gram-positive bacteria[40] and have been used as antibiotic adjuvants to overcome colistin resistance in some ESKAPE pathogens[41]. Here, we highlight the prospect of such moieties to improve Gram-negative activity and reduce efflux susceptibility.

The empirical and computational approaches reported herein identified molecular descriptors and structural modifications that advance our understanding of efflux susceptibility. Our work here illustrates that physicochemical properties and chemical structures serve as guidelines for efflux, and combinations thereof generally define substrate quality. Leveraging these properties, we developed the first machine learning tool that can be used to assess efflux susceptibility of small molecules *in silico*. Further expansion of guidelines for efflux, however, will come from additional assessment of compound accumulation in Gram-negative bacteria, including those identified in this work.

Conclusion

Efflux pumps are major contributors to the intrinsic resistance of many antibiotics in Gram-negative bacteria. These pumps have challenged the design of candidate molecules that would be otherwise efficacious in Gram-negatives. Understanding the molecular descriptors of small molecules susceptible to efflux pumps is important to overcome this challenge. Herein, we present a novel approach to determine the physicochemical and

structural spaces occupied by such molecules. Multivariate analysis and a machine learning approach identified that hydrophobicity, the level of saturation, and molecular stability contribute to Gram-negative activity yet limit it to efflux-compromised *E. coli*. SERF, a second machine learning model, further revealed that along with hydrophobicity and molecular stability, the extent of molecular branching and compactness is a key factor that also governs efflux susceptibility. Additionally, SERF proved capable of predicting molecular efflux and is the first machine learning tool to identify small molecules susceptible to efflux pumps *in silico*. Furthermore, structure-activity relationship analyses revealed that some molecular side chains, and their associated physicochemical properties, serve as triggers for efflux. In all, the results of this work provide novel insight into the physicochemical properties and chemical structures governing efflux in Gram-negative bacteria, which can be used to guide the design of novel Gram-negative antimicrobials.

Methods

Screen for antibacterial activity in the efflux-deficient strain *E. coli* Δ*tolC*

The strains used in this study were *E. coli* BW25113 (wild-type) and Δ*tolC* (*E. coli* parent strain BW25113)[42]. In all experiments, a mid-log culture in M9 minimal medium was prepared as previously described[43] and used to prepare cells to a final working inoculum of ~10^5 CFU/mL. For the small molecule screen, compounds were added to a 384-well assay plate to a final concentration of 10 μM. Molecules, dissolved in DMSO, were sourced from Enamine, ChemDiv, Asinex, ChemBridge, Maybridge, Sigma (Lopac),

Prestwick, Biomol, and Microsource. To these plates, a final working inoculum of the *E. coli* Δ*tolC* strain was added to a final volume of 50 μL, and plates were incubated at 37°C for 18 h. All screens were performed in duplicates. Liquid handling was performed using a Beckman Coulter FX^P Laboratory Automated Workstation. After incubation, absorbance at 600 nm (OD_{600}) of the 384-well assay plates was measured using a Perkin Elmer EnVision plate reader.

Analysis of screening data

To reduce plate-to-plate variation, data from the OD_{600} measurements of each 384-well assay plate were rank-ordered and the interquartile mean of each plate was calculated. Data were first normalized on a per plate basis as previously described[44]. To account for positional effects, these data were further normalized by the interquartile mean of each well position as previously described[44]. A cut-off at three standard deviations below the mean was established to determine actives, where any wells with values below this cut-off were considered actives.

Determination of antibacterial potency in dose

Wild-type *E. coli* and Δ*tolC* strains were grown and prepared in M9 medium to final working inocula as described above. These cells were then added to a 96-well assay plate containing half-log serial dilutions of one of the actives identified from the screen described above, where concentrations ranged from 50 μM to 0 μM. Assay plates were then incubated at 37°C for 18 h and OD_{600} was measured using a Tecan plate reader (Infinite M1000). Growth at each exposed concentration was determined as follows

$$G = \frac{G_i}{G_0}$$

where G_i represents the growth in one of the wells exposed to the different concentrations of the tested compound, and the G_0 represents the growth in the well that was not exposed to the tested compound. A value of 0 represents no growth on the plate and 1 represents no growth inhibition. An EC_{50} dose-response curve was then fit to the data using the four parameter dose-response model in GraphPad Prism in order to calculate the potency of a compound, based on the ratio of wild-type EC_{50} to Δ*tolC* EC_{50} (wild-type EC_{50}/Δ*tolC* EC_{50}).

Calculation and principal component analysis of molecular properties

Structures of all compounds in this study exist as MOL file coordinates within compiled SDF files. All structures had 3D coordinates generated using cxcalc (ChemAxon), and the lowest energy conformer was chosen as the basis for molecular property calculations. All calculations were done using cxcalc in a Linux terminal, with pH set to 7.0 where relevant. An initial principal component analysis (PCA) was performed as a means of dimensionality reduction, in order to visualize the actives from the primary screen within the chemical space of the screened molecules. The R statistical programming language[45] was used here, with data appropriately scaled to eliminate bias from the different units and magnitudes of chemical descriptors. Code and dataset examples, including cxcalc commands used to generate the chemical properties, can be found on GitHub (https://github.com/sfrench007/serf).

Machine learning and prediction of efflux susceptibility

Further exploring the importance of the molecular descriptors specific to the 3,780 efflux-dependent actives, these molecules were used in a random forest machine learning approach to explore the properties key to Gram-negative activity in an efflux-deficient *E. coli* BW25113 Δ*tolC* strain. A binary (two-class) approach was used in creating a data set to train a random forest model. In this, the 3,780 efflux-dependent actives were classified as 'true' and an equal number of randomly sampled inactive molecules from the screened collection were classified as 'false'. This was performed in R, with the random forest learning done using the *caret* package[46]. Molecular descriptors that were redundant based on a similarity matrix (>85% similarity) were removed from the model training set. The model was trained with 2000 trees grown, 10 iterations, and a two-class summary function for 'true' and 'false' identifications. Validation was done through repeated random sub-sampling (70/30 split, training/test), with 10 repeats for each cross-validation. The number of properties randomly sampled when splitting at each tree node was incrementally increased and used to tune the optimal number of properties for the highest receiver operating characteristic (ROC) value. The ROC is a measure of how well the model can classify, in this case, Gram-negative active molecules. Upon training the model, molecular descriptors were ranked by their relative importance; importance in that removing that descriptor from the training set would result in incorrect classification and predictions. Frequency distributions for important descriptors were also examined, to visualize the separations, or shifts, between the two phenotypes used in the model.

To explore molecular descriptors contributing to efflux susceptibility, a second random forest model was created. This model used the same decision tree parameters as the previous, this time comparing ~290 pumped actives (EC_{50} <5 μM and wild-type EC50 >50 μM), and a random subset of ~290 non-pumped actives (fold-change in potency ≤2). This random forest model, called Susceptibility to Efflux Random Forest (SERF), was tuned and evaluated in a similar manner as the first to identify descriptors of importance to efflux. A set of ~440 actives was assessed by SERF to predict small molecules susceptible to efflux. All code and dataset examples can be found on GitHub (https://github.com/sfrench007/serf).

Structure-activity relationship analysis

Dose-response curves from primary screening actives (Δ*tolC* strain) were compared to dose-response curves of the wild-type strain. In this, EC_{50} values for each dose-response curve were calculated and compared to obtain a fold-change value. These fold-changes were tabulated along with chemical structure (as a mol representation within an SDF file), and structural fingerprints were calculated in Data Warrior[47]. The fingerprints were based on a fragment dictionary generated within Data Warrior, and a correlation matrix within the software compared all 4,507 actives in the data table. All active molecules are randomly placed within a 2D space, and force-directed clustering applied[47] based on the correlation matrix, bringing like-molecules together into groups. The result is a 2D structural representation of molecular connectivity, providing insights into the functional groups that define efflux-susceptible and efflux-evading molecules.

Clusters of actives in Figure 6 were more closely examined by breaking down the molecules to common structural cores and R-groups. This was done using Data Warrior, which calculated the common core for each chemical analog within each group of molecules and generated a list of functional groups added to that core. Using this method, several structure-activity relationship tables were generated, indicating which functional groups had an impact on how well chemical analogs in each table were susceptible to efflux.

Author contributions

S.S.E. and E.D.B. conceived and designed the research. S.S.E., M.A.F., and G.K. performed the primary screen; S.S.E. and G.K. performed the dose-response experiments; S.S.E. analyzed the primary screen, dose-response, and SAR data. S.F. calculated the molecular descriptors for all screened molecules, generated the structure-activity representations, and wrote and developed the code for the machine learning models. S.S.E., S.F., and E.D.B. wrote the manuscript.

Acknowledgements

This research was supported by a Foundation grant from the Canadian Institutes for Health Research (FRN 143215), an infrastructure funding from Canada Foundation for Innovation, a Research Excellence allocation from the Ontario Research Fund, and a Tier I Canada Research Chair award to E.D.B. In addition, S.S.E was supported by an Ontario Graduate Scholarship and a Fred and Helen Knight Enrichment Award. The high-throughput screen was performed in the Centre for Microbial Chemical Biology core

facility at McMaster University. We thank Susan McCuskar for her excellent technical assistance with the screen.

References

1. Lewis, K., The science of antibiotic discovery. *Cell* **2020,** *181* (1), 29-45.
2. Brown, E. D.; Wright, G. D., Antibacterial drug discovery in the resistance era. *Nature* **2016,** *529* (7586), 336-343.
3. *Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis*; World Health Organization: Geneva, 2017.
4. *Antibiotic resistance threats in the United States, 2019*; U.S. Department of Health and Human Services, CDC: 2019.
5. *2019 antibacterial agents in clinical development: an analysis of the antibacterial clinical development pipeline*; World Health Organization: Geneva, 2019.
6. Butler, M. S.; Paterson, D. L., Antibiotics in the clinical pipeline in October 2019. *J. Antibiot.* **2020,** *73* (6), 329-364.
7. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L., Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **2007,** *6* (1), 29-40.
8. Blair, J. M. A.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. V., Molecular mechanisms of antibiotic resistance. *Nat. Rev. Microbiol.* **2014,** *13* (1), 42-51.
9. Tommasi, R.; Brown, D. G.; Walkup, G. K.; Manchester, J. I.; Miller, A. A., ESKAPEing the labyrinth of antibacterial discovery. *Nat. Rev. Drug Discov.* **2015,** *14*, 529-542.
10. *Challenges in the discovery of Gram-negative antibacterials: the entry and efflux problem*; Pew Charitable Trusts: 2017.
11. *A scientific roadmap for antibiotic discovery*; Pew Charitable Trusts: 2016.
12. Shore, C. K.; Coukell, A., Roadmap for antibiotic discovery. *Nat. Microbiol.* **2016,** *1* (6), 16083.
13. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001,** *46* (1-3), 3-26.
14. O'Shea, R.; Moser, H. E., Physicochemical properties of antibacterial compounds: implications for drug discovery. *J. Med. Chem.* **2008,** *51* (10), 2871-2878.
15. Reck, F.; Jansen, J. M.; Moser, H. E., Challenges of antibacterial drug discovery. *Arkivoc* **2019**, 227-244.
16. Krishnamoorthy, G.; Leus, I. V.; Weeks, J. W.; Wolloscheck, D.; Rybenkov, V. V.; Zgurskaya, H. I., Synergy between active efflux and outer membrane diffusion defines rules of antibiotic permeation into Gram-negative bacteria. *mBio* **2017,** *8* (5), e01172-17.

17. Cooper, S. J.; Krishnamoorthy, G.; Wolloscheck, D.; Walker, J. K.; Rybenkov, V. V.; Parks, J. M.; Zgurskaya, H. I., Molecular properties that define the activities of antibiotics in *Escherichia coli* and *Pseudomonas aeruginosa*. *ACS Infect. Dis.* **2018,** *4* (8), 1223-1234.
18. Du, D.; Wang, Z.; James, N. R.; Voss, J. E.; Klimont, E.; Ohene-Agyei, T.; Venter, H.; Chiu, W.; Luisi, B. F., Structure of the AcrAB-TolC multidrug efflux pump. *Nature* **2014,** *509* (7501), 512-515.
19. Eicher, T.; Cha, H.-j.; Seeger, M. A.; Brandstätter, L.; El-Delik, J.; Bohnert, J. A.; Kern, W. V.; Verrey, F.; Grütter, M. G.; Diederichs, K.; Pos, K. M., Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (15), 5687-5692.
20. Brown, D. G.; May-Dracka, T. L.; Gagnon, M. M.; Tommasi, R., Trends and exceptions of physical properties on antibacterial activity for Gram-positive and Gram-negative pathogens. *J. Med. Chem.* **2014,** *57* (23), 10144-10161.
21. Li, X. Z.; Ma, D.; Livermore, D. M.; Nikaido, H., Role of efflux pump(s) in intrinsic resistance of *Pseudomonas aeruginosa*: active efflux as a contributing factor to beta-lactam resistance. *Antimicrob. Agents Chemother.* **1994,** *38* (8), 1742-1752.
22. Murakami, S.; Nakashima, R.; Yamashita, E.; Matsumoto, T.; Yamaguchi, A., Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* **2006,** *443* (7108), 173-179.
23. Nakashima, R.; Sakurai, K.; Yamasaki, S.; Nishino, K.; Yamaguchi, A., Structures of the multidrug exporter AcrB reveal a proximal multisite drug-binding pocket. *Nature* **2011,** *480*, 565-569.
24. Méndez-Lucio, O.; Medina-Franco, J. L., The many roles of molecular complexity in drug discovery. *Drug Discov. Today* **2017,** *22* (1), 120-126.
25. Sheridan, R. P.; Zorn, N.; Sherer, E. C.; Campeau, L.-C.; Chang, C.; Cumming, J.; Maddess, M. L.; Nantermet, P. G.; Sinz, C. J.; O'Shea, P. D., Modeling a crowdsourced definition of molecular complexity. *J. Chem. Inf. Model.* **2014,** *54* (6), 1604-1616.
26. Sisay, M. T.; Peltason, L.; Bajorath, J., Structural interpretation of activity cliffs revealed by systematic analysis of structure−activity relationships in analog series. *J. Chem. Inf. Model* **2009,** *49* (10), 2179-2189.
27. Li, X.-Z.; Plésiat, P.; Nikaido, H., The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin. Microbiol. Rev.* **2015,** *28* (2), 337-418.
28. Takatsuka, Y.; Chen, C.; Nikaido, H., Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (15), 6559-6565.
29. Vargiu, A. V.; Nikaido, H., Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (50), 20637-20642.
30. Bohnert, J. A.; Schuster, S.; Seeger, M. A.; Fähnrich, E.; Pos, K. M.; Kern, W. V., Site-directed mutagenesis reveals putative substrate binding residues in the *Escherichia coli* RND efflux pump AcrB. *J. Bacteriol.* **2008,** *190* (24), 8225-8229.

31. Nakashima, R.; Sakurai, K.; Yamasaki, S.; Hayashi, K.; Nagata, C.; Hoshino, K.; Onodera, Y.; Nishino, K.; Yamaguchi, A., Structural basis for the inhibition of bacterial multidrug exporters. *Nature* **2013,** *500*, 102-106.
32. Vargiu, A. V.; Ruggerone, P.; Opperman, T. J.; Nguyen, S. T.; Nikaido, H., Molecular mechanism of MBX2319 inhibition of *Escherichia coli* AcrB multidrug efflux pump and comparison with other inhibitors. *Antimicrob. Agents Chemother.* **2014,** *58* (10), 6224-6234.
33. Masi, M.; Réfregiers, M.; Pos, K. M.; Pagès, J.-M., Mechanisms of envelope permeability and antibiotic influx and efflux in Gram-negative bacteria. *Nat. Microbiol.* **2017,** *2* (3), 17001.
34. Novick, R. P., Staphylococcal penicillinase and the new penicillins. *Biochem. J.* **1962,** *83* (2), 229-235.
35. Knowles, J. R., Penicillin resistance: the chemistry of beta-lactamase inhibition. *Acc. of Chem. Res.* **1985,** *18* (4), 97-104.
36. Drawz, S. M.; Bonomo, R. A., Three decades of β-lactamase inhibitors. *Clin. Microbiol. Rev.* **2010,** *23* (1), 160-201.
37. Lim, S. P.; Nikaido, H., Kinetic parameters of efflux of penicillins by the multidrug efflux transporter AcrAB-TolC of *Escherichia coli*. *Antimicrob. Agents Chemother.* **2010,** *54* (5), 1800-1806.
38. Walsh, C., *Antibiotics*. American Society of Microbiology: 2003.
39. Verbist, L., Comparison of the activities of the new ureidopenicillins piperacillin, mezlocillin, azlocillin, and Bay k 4999 against Gram-negative organisms. *Antimicrob. Agents Chemother.* **1979,** *16* (2), 115-119.
40. Boudreau, M. A.; Ding, D.; Meisel, J. E.; Janardhanan, J.; Spink, E.; Peng, Z.; Qian, Y.; Yamaguchi, T.; Testero, S. A.; O'Daniel, P. I.; Leemans, E.; Lastochkin, E.; Song, W.; Schroeder, V. A.; Wolter, W. R.; Suckow, M. A.; Mobashery, S.; Chang, M., Structure-activity relationship for the oxadiazole class of antibacterials. *ACS Med. Chem. Lett.* **2020,** *11* (3), 322-326.
41. Barker, W. T.; Chandler, C. E.; Melander, R. J.; Ernst, R. K.; Melander, C., Tryptamine derivatives disarm colistin resistance in polymyxin-resistant Gram-negative bacteria. *Bioorg. Med. Chem.* **2019,** *27* (9), 1776-1788.
42. Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H., Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2006,** *2*, 2006.0008.
43. El Zahed, S. S.; Brown, E. D., Chemical-chemical combinations map uncharted interactions in *Escherichia coli* under nutrient stress. *iScience* **2018,** *2*, 168-181.
44. Mangat, C. S.; Bharat, A.; Gehrke, S. S.; Brown, E. D., Rank ordering plate data facilitates data visualization and normalization in high-throughput screening. *J. Biomol. Screen* **2014,** *19* (9), 1314-1320.
45. Ihaka, R.; Gentleman, R., R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **1996,** *5* (3), 299-314.
46. Kuhn, M., Building predictive models in R using the caret package. *J. Stat. Softw.* **2008,** *28* (5), 1-26.

47. Sander, T.; Freyss, J.; von Korff, M.; Rufener, C., DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J. Chem. Inf. Model* **2015,** *55* (2), 460-473.

Figure Titles and Legends

Figure 1. Primary small molecule screen and potency analysis of actives. (A) Replicate plot of the primary screen of ~314,000 small molecules in the efflux-deficient mutant strain *E. coli* BW25113 Δ*tolC*. Data were normalized as per Mangat *et al.*[44] and shows good reproducibility. Density distributions are depicted in secondary plots for each of the two replicates. (B) From this primary screen, ~4,500 actives exhibited strong growth inhibition (≤3σ below the mean; normalized mean OD value of ~0.5 or less was deemed a growth inhibitor) against the mutant strain *E. coli* BW25113 Δ*tolC*. The potency of these 4,500 actives was measured in wild-type *E. coli* BW25113 and the mutant strain *E. coli* BW25113 Δ*tolC*. A histogram of the measured potency values obtained from dose-response analyses in (C) wild-type *E. coli* BW25113 and (D) the mutant strain *E. coli* BW25113 Δ*tolC*.

Figure 2. Principal component analysis of molecular descriptors of efflux-dependent actives and non-growth inhibitory molecules from the primary screen. (A) Molecular descriptors for each efflux-dependent active and the remaining screened compounds were calculated. The former set and all non-growth inhibitory molecules from the primary screen (inactive molecules) were visualized using a principal component analysis (PCA). (B) 10 principal components (PC) explaining variances between descriptors of efflux-dependent actives and those of inactive molecules are shown. The first 3 principal

components, PC1, PC2, and PC3, explain ~63% of these variances. (C) For visualization purposes, the first two principal component scores are presented, where efflux-dependent actives occupied a distinct chemical space (blue) among the inactive molecules (grey). (D) The loading eigenvectors for this PCA plot are shown, where molecular descriptors contributing the most to the scores of PC1 and PC2 are highlighted in blue, MW: molecular weight; atom count; ASA: accessible surface area of (-) partial negative charge; (P) polar atoms; (+) partial positive charge; and (H) hydrophobic atoms; Fsp^3: fraction of sp^3 hybridized carbon atoms; aliphatic count, and descriptors with lower contributions are shown in grey.

Figure 3. A machine learning approach identifies key molecular descriptors for Gram-negative activity of efflux-dependent actives. (A) Molecular descriptors for the 3,780 efflux-dependent actives and a random set of 3,780 inactive molecules (no growth inhibition in *E. coli* BW25113 Δ*tolC*) from the primary screen were used to train a random forest model to examine descriptors contributing to Gram-negative activity in efflux-compromised *E. coli*. (B) The area under the curve-receiver operating characteristic curve (AUC-ROC) plot for the random forest model is 0.808, showing a good distinction between efflux-dependent actives and inactive molecules. Sensitivity refers to the true positive rate of the model, while specificity refers to its false positive rate. (C) The top 10 molecular descriptors that reduce the model's accuracy are shown, with clogD, Fsp^3 (fraction of sp^3 hybridized carbon atoms), and resonant structure count topping the list.

Figure 4. The Susceptibility to Efflux Random Forest model identifies molecular descriptors governing efflux. (A) The pipeline highlights the training set of compounds used to build the Susceptibility to Efflux Random Forest (SERF) model and identify key descriptors that contribute to efflux susceptibility. Based on the shown cut-offs for the 4,500 actives from the primary screen, ~1,070 actives were pumped molecules and ~410 actives were non-pumped molecules. For a random set of ~290 pumped molecules and ~290 non-pumped molecules, molecular descriptors for each of these compounds were used to train SERF and identify those contributing to efflux susceptibility in *E. coli*. (B) The area under the curve-receiver operating characteristic curve (AUC-ROC) plot for SERF is 0.839, showing a good distinction between pumped molecules and non-pumped molecules. Sensitivity refers to the true positive rate of the model, while specificity refers to its false positive rate. (C) The top 10 molecular descriptors that reduce the model's accuracy are shown, with resonant structure count, clogD, and hyper-Wiener index accounting for the greatest impact on accuracy. Polarizability tensor a(yy): principal component of polarizability along the coordinate space a(yy).

Figure 5. SERF identifies small molecules susceptible to efflux pumps. (A) Molecular descriptors of a test set of ~440 actives (~320 pumped molecules and ~120 non-pumped molecules) that were excluded from the training set were used by SERF to identify molecules susceptible to efflux. SERF predicted some 274 molecules, of which 260 were validated by dose-response potency analysis. (B) The Venn diagram represents the accuracy and predictive power of SERF to identify efflux-susceptible molecules, as determined by empirical assessment and validation; 260 molecules were correct

predictions; 14 molecules were incorrect predictions; 61 molecules were not identified by SERF to be efflux-susceptible.

Figure 6. Structure-activity representation of the 4,500 actives from the primary screen. (A) Small molecules were clustered by structural similarity with fragment-based fingerprints in Data Warrior[47], using a force-directed method of clustering within the 2D space shown. Several clusters containing >15 molecules (neighbours) are seen from the primary screening data. The points are coloured by a fold-change in EC_{50} between wild-type *E. coli* BW25113 and the mutant strain *E. coli* BW25113 Δ*tolC*. The sizes of the points indicate the structural similarity between a compound and its immediate neighbours. The location of the β-lactam cluster (black oval) and the hydroxyquinoline derivatives cluster (red oval) are highlighted. Structures of the common core for the (B) β-lactam cluster and (C) the hydroxyquinoline derivatives cluster are shown.

Supporting Information

Figure S1. The receiver operating characteristic curve for the random forest model classifying efflux-dependent actives and inactive molecules from the primary screen. The training set of molecules used to build this model was composed of 3,780 efflux-dependent actives and a random set of 3,780 inactive molecules from the primary screen. Tuning the number of molecular descriptors possible at each node in the decision trees, the number of randomly selected descriptors (features) were compared to a receiver operating characteristic curve (ROC). Here, the number of molecular descriptors at each tree node that resulted in the most accurate model was observed to be 10.

Figure S2. Physicochemical and molecular space occupied by efflux-dependent actives and inactive molecules from the primary screen. Density plots comparing some molecular descriptors of the 3,780 efflux-dependent actives (blue) and the random set of 3,780 inactive molecules (grey) from the primary screen. This set of molecules and their associated descriptors were used to build the random forest model in Figure 3. Shown are the top three descriptors contributing to this model's accuracy in classification: (A) clogD (pH 7.0), (B) Fsp^3, and (C) resonant structure count.

Figure S3. The receiver operating characteristic curve for SERF. The training set of molecules used to build this model was composed of a random set of ~290 pumped molecules and ~290 non-pumped molecules. Tuning the number of molecular descriptors possible at each node in the decision trees, the number of randomly selected descriptors (features) were compared to a receiver operating characteristic curve (ROC). Here, the number of molecular descriptors at each tree node that resulted in the most accurate model was observed to be 6.

Figure S4. Physicochemical and molecular space occupied by pumped molecules and non-pumped molecules. Density plots comparing some molecular descriptors of the random set of ~290 pumped molecules (blue) and ~290 non-pumped molecules (grey). This set of molecules and their associated descriptors were used to build SERF. Shown are the top three descriptors contributing to SERF's accuracy in classification: (A) resonant structure count, (B) clogD (pH 7.0), and (C) hyper-Wiener index.

Table S1. Measured potency and molecular descriptors of all 4,500 actives from the primary screen. All EC_{50} values reported here were assessed in duplicates against wild-type *E. coli* BW25113 and the efflux-deficient strain *E. coli* BW25113 Δ*tolC* strain grown in M9 medium, as described in methods. The fold-change in potency and molecular descriptors were calculated as described in methods.

Table S2. Measured potency of 274 molecules predicted by SERF to be susceptible to efflux pumps. All EC_{50} values reported here were assessed in duplicates in wild-type *E. coli* BW25113 and the efflux-deficient strain *E. coli* BW25113 Δ*tolC* strain grown in M9 medium, as described in methods. Molecules highlighted in blue were correct predictions by SERF while those highlighted in red were incorrect predictions.

Table S3. Measured potency of 61 molecules empirically identified to be efflux-susceptible but not predicted to be susceptible by SERF. All EC_{50} values reported here were assessed in duplicates in wild-type *E. coli* BW25113 and the efflux-deficient strain *E. coli* BW25113 Δ*tolC* strain grown in M9 medium, as described in methods.

Table S4. Structure, activity, and molecular descriptors of an antifungal compound series. Local structure-activity relationship of the structure-based cluster of some antifungals.

Figures

Figure 1

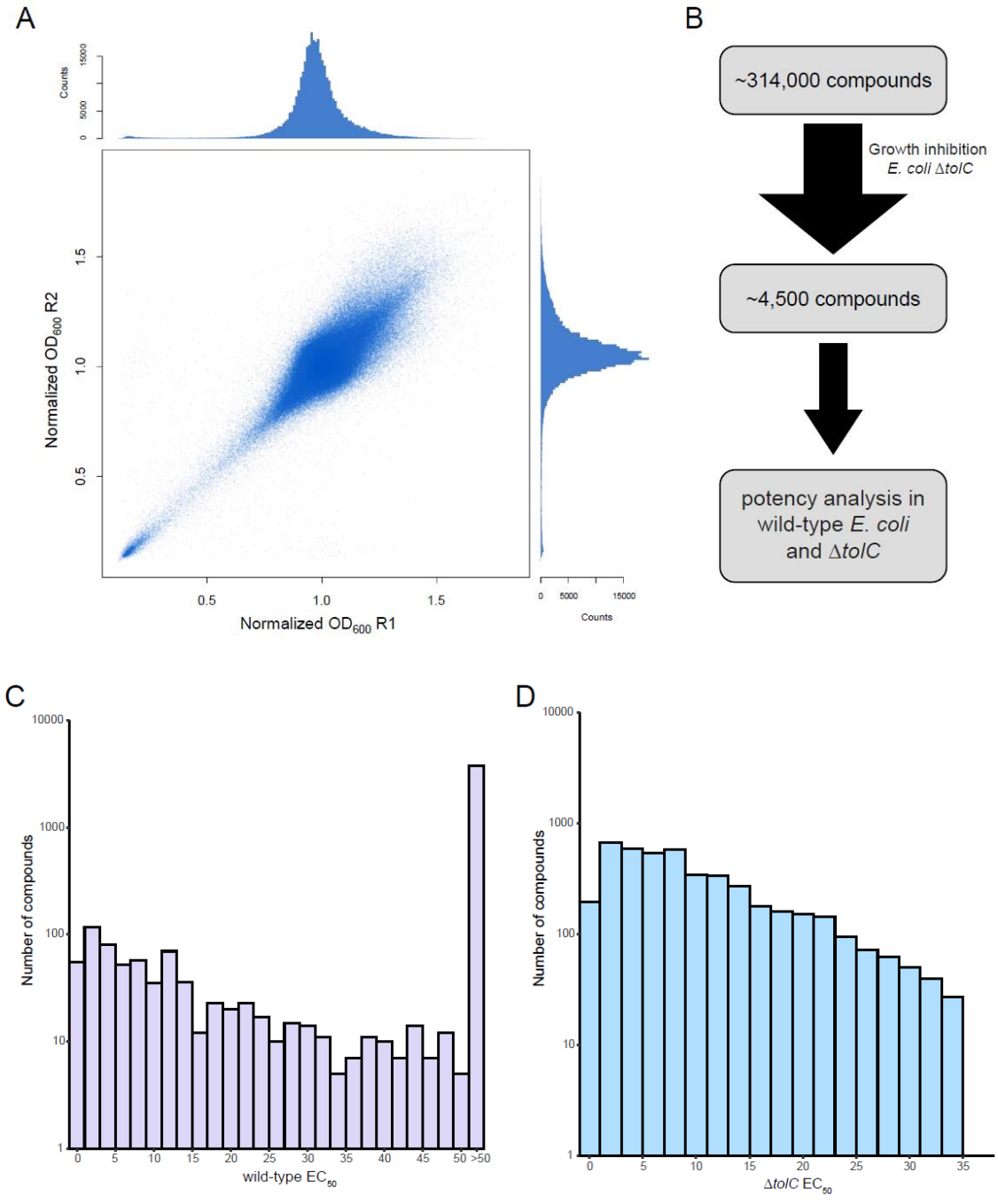


Figure 2

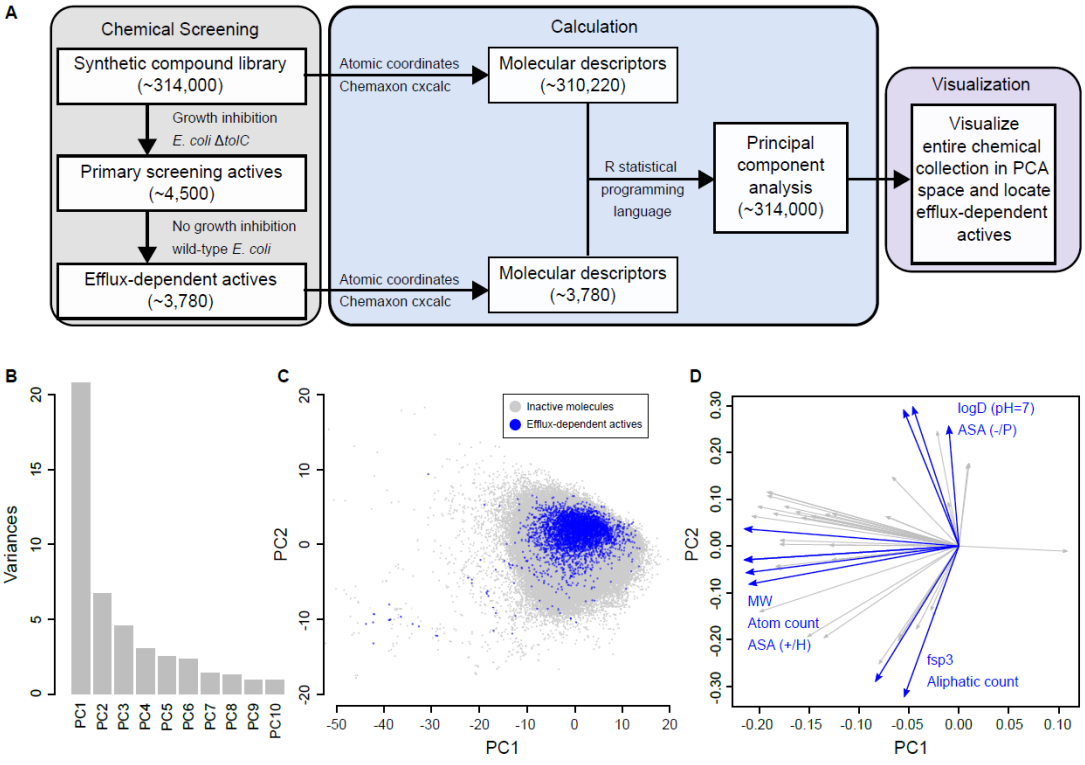

A
Chemical Screening
Synthetic compound library (~314,000)
Growth inhibition E. coli ΔtolC
Primary screening actives (~4,500)
No growth inhibition wild-type E. coli
Efflux-dependent actives (~3,780)
Calculation
Atomic coordinates Chemaxon cxcalc
Molecular descriptors (~310,220)
R statistical programming language
Principal component analysis (~314,000)
Atomic coordinates Chemaxon cxcalc
Molecular descriptors (~3,780)
Visualization
Visualize entire chemical collection in PCA space and locate efflux-dependent actives
B
Variances
PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8 PC9 PC10
C
Inactive molecules
Efflux-dependent actives
PC2
PC1
D
logD (pH=7)
ASA (-/P)
MW
Atom count
ASA (+/H)
fsp3
Aliphatic count
PC2
PC1

Figure 3

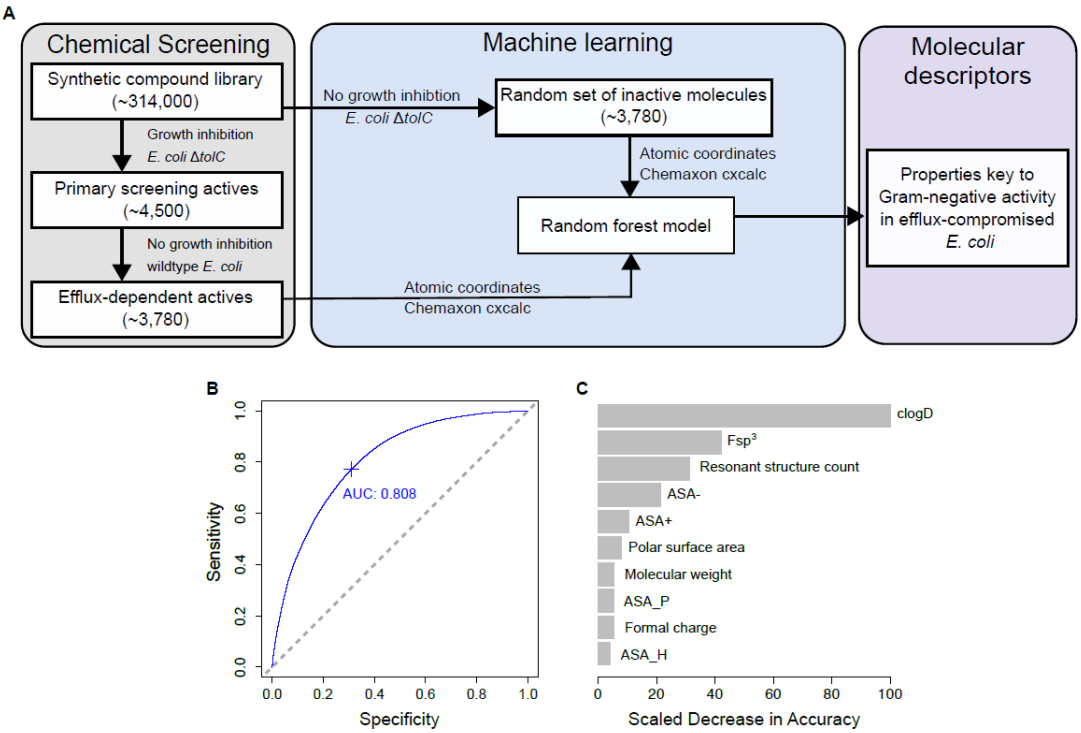

A
Chemical Screening
Synthetic compound library
(~314,000)
Growth inhibition
E. coli ΔtolC
Primary screening actives
(~4,500)
No growth inhibition
wildtype E. coli
Efflux-dependent actives
(~3,780)
Machine learning
No growth inhibtion
E. coli ΔtolC
Random set of inactive molecules
(~3,780)
Atomic coordinates
Chemaxon cxcalc
Random forest model
Atomic coordinates
Chemaxon cxcalc
Molecular descriptors
Properties key to Gram-negative activity in efflux-compromised E. coli
B
AUC: 0.808
Sensitivity
Specificity
0.0
0.2
0.4
0.6
0.8
1.0
C
clogD
Fsp3
Resonant structure count
ASA-
ASA+
Polar surface area
Molecular weight
ASA_P
Formal charge
ASA_H
0
20
40
60
80
100
Scaled Decrease in Accuracy

Figure 4

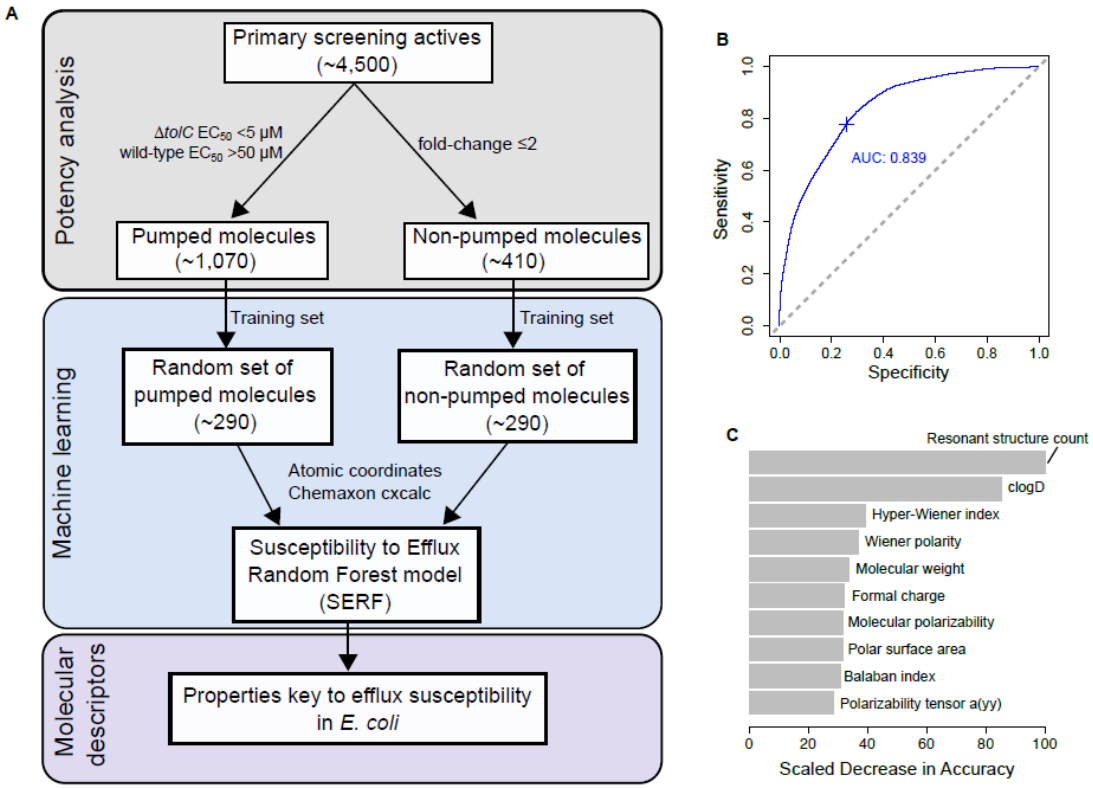

A
Potency analysis
Primary screening actives
(~4,500)
ΔtolC EC50 <5 μM
wild-type EC50 >50 μM
fold-change ≤2
Pumped molecules
(~1,070)
Non-pumped molecules
(~410)
Machine learning
Training set
Training set
Random set of
pumped molecules
(~290)
Random set of
non-pumped molecules
(~290)
Atomic coordinates
Chemaxon cxcalc
Susceptibility to Efflux
Random Forest model
(SERF)
Molecular
descriptors
Properties key to efflux susceptibility
in E. coli
B
Sensitivity
Specificity
AUC: 0.839
C
Resonant structure count
clogD
Hyper-Wiener index
Wiener polarity
Molecular weight
Formal charge
Molecular polarizability
Polar surface area
Balaban index
Polarizability tensor a(yy)
Scaled Decrease in Accuracy

Figure 5

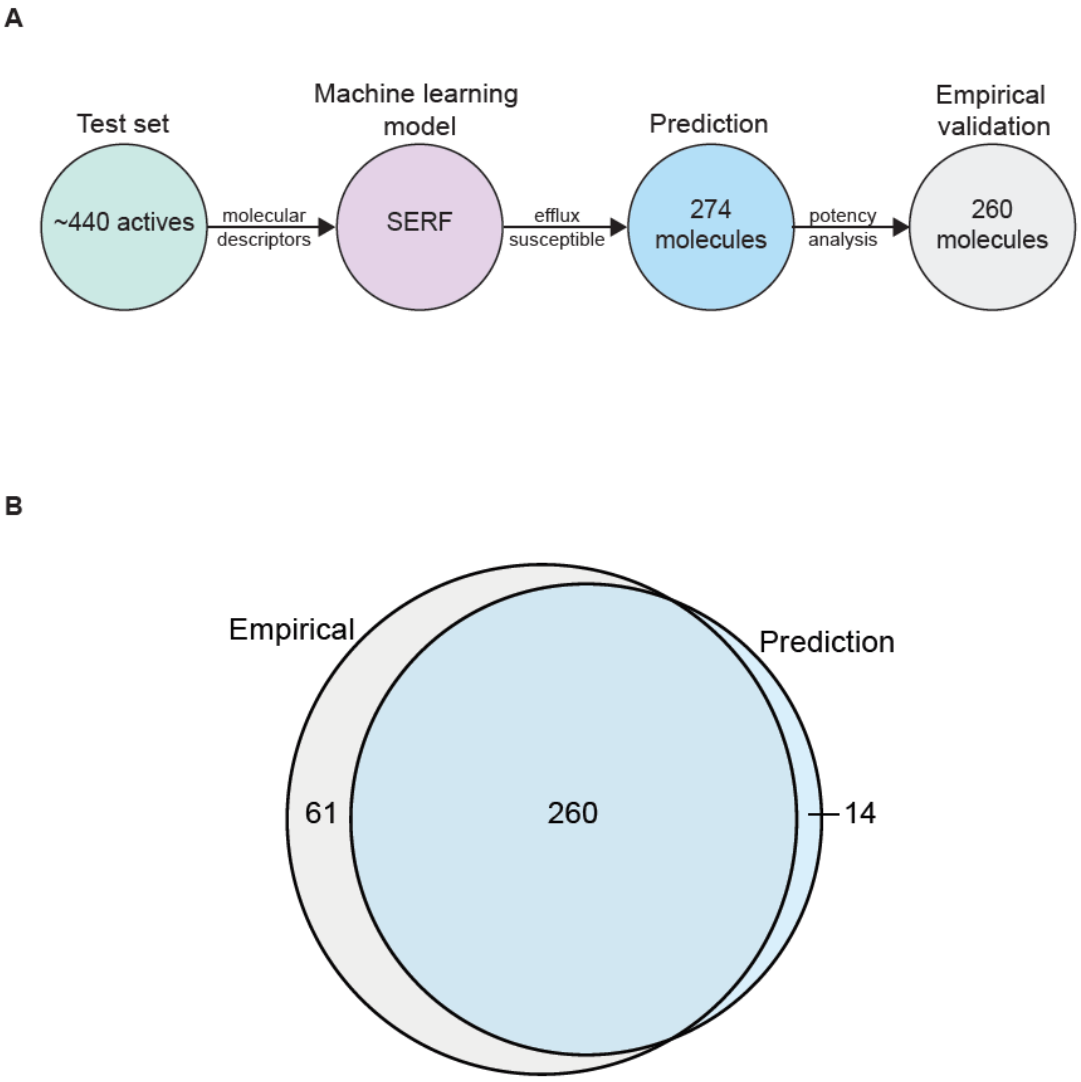
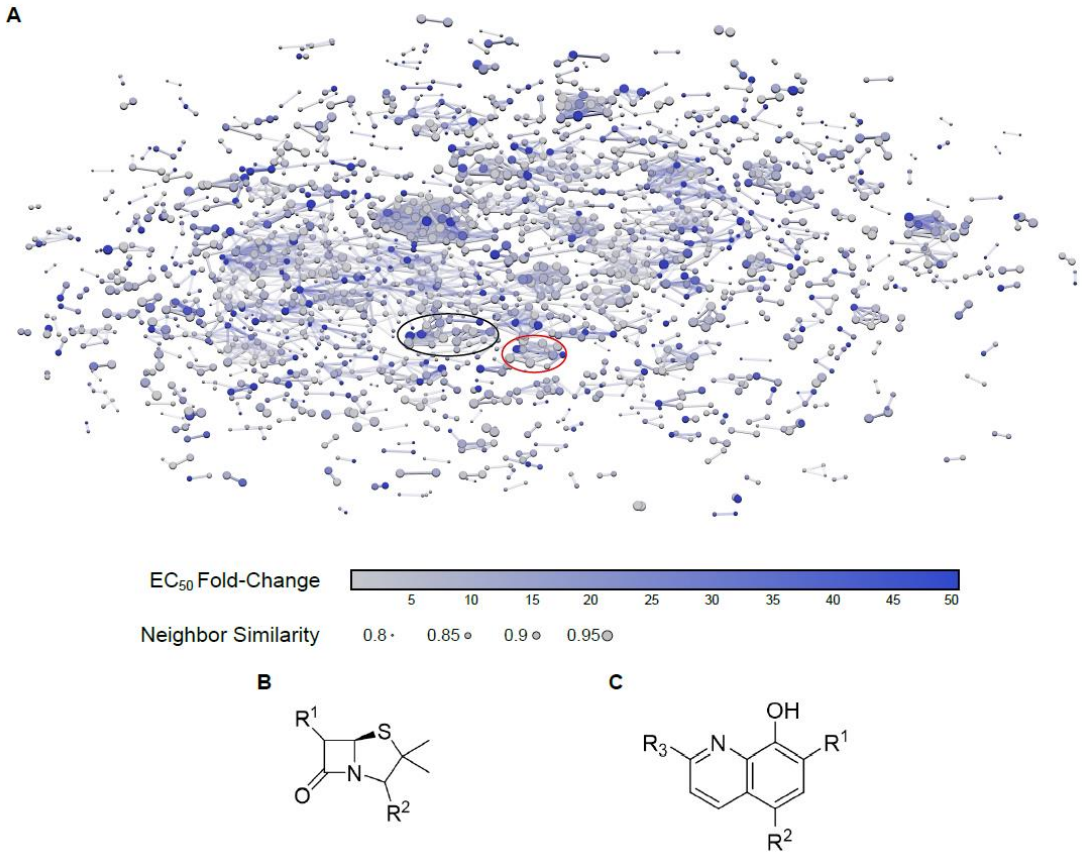
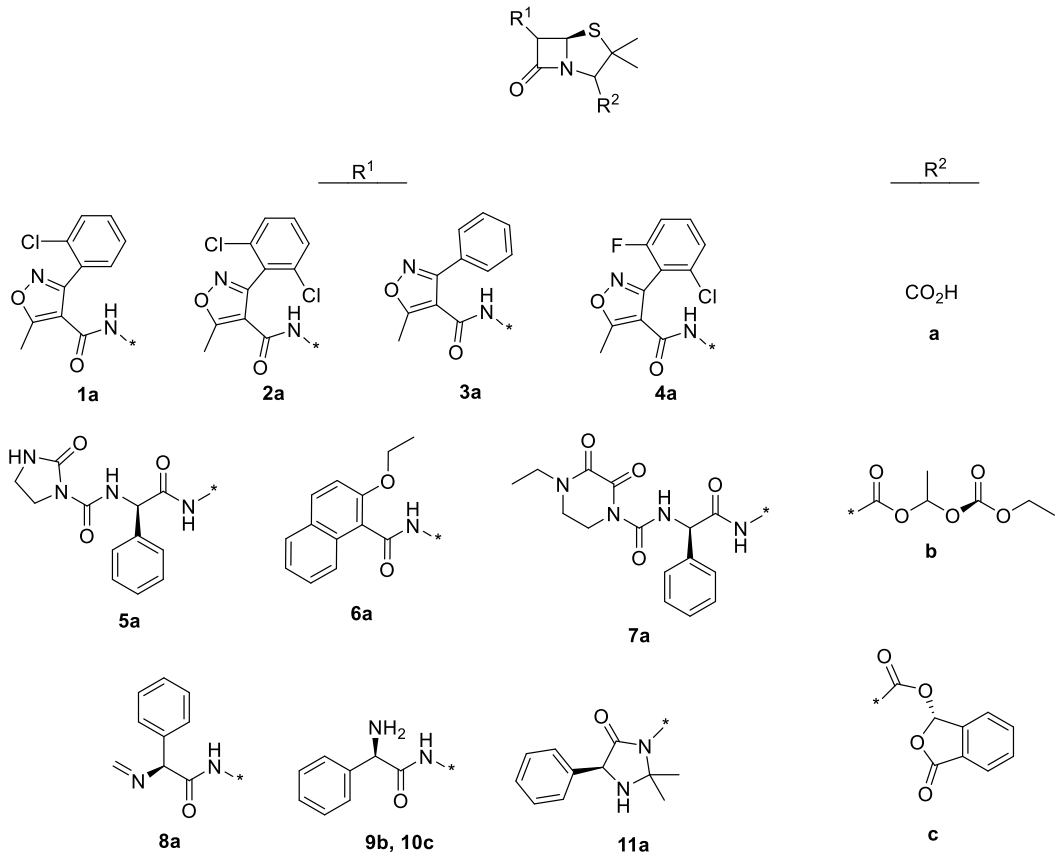

A
Test set
~440 actives
molecular descriptors
Machine learning model
SERF
efflux susceptible
Prediction
274 molecules
potency analysis
Empirical validation
260 molecules
B
Empirical
Prediction
61
260
14

Figure 6


A
EC50 Fold-Change
5
10
15
20
25
30
35
40
45
50
Neighbor Similarity
0.8
0.85
0.9
0.95
B
R1
S
N
O
R2
C
OH
R3
N
R1
R2

Tables

Table 1. Structure, activity, and molecular descriptors of a β-lactam compound series



Cmpd[a]	Cmpd Name	W.T.[b] EC_{50} (μM)	Δ*tolC* EC_{50} (μM)	F.C.[c]	M.W. (g mol^{-1})	clogP	PSA[d] ($Å^2$)	Fsp^3 [e]	RSC [f]	HWI [g]
1a	Cloxacillin	>50	1.47	>50	435.88	2.70	138	0.37	3	9,005
2a	Dicloxacillin	>50	2.88	>34.7	470.32	3.31	138	0.37	3	9,877
3a	Oxacillin	>50	3.15	>31.7	401.43	2.10	138	0.37	3	8,143
4a	Flucloxacillin	>50	3.91	>25.6	453.87	2.80	138	0.37	3	9,877
5a	Azlocillin	>50	5.72	>17.5	461.49	0.312	173	0.45	5	12,535
6a	Nafcillin	>50	10.5	>9.49	414.47	3.12	121	0.38	2	8,612
7a	Piperacillin	25.3	2.31	10.9	517.55	-0.174	182	0.48	4	19,160
8a	Metampicillin	14.9	1.82	8.19	361.41	0.542	124	0.41	5	5,716
9b	Bacampicillin	30.7	7.13	4.31	465.51	1.29	163	0.52	4	15,846
10c	Talampicillin	13.1	4.79	2.72	481.52	1.34	153	0.33	6	17,736
11a	Hetacillin	4.18	1.57	2.67	389.46	-0.234	115	0.53	4	6,164

[a]Cmpd, compound

[b]W.T., wild-type *E. coli*

[c]F.C., fold-change

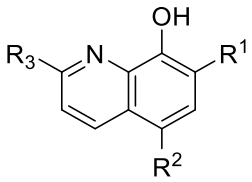
[d]PSA, polar surface area

[e]Fsp^3, ratio of sp^3 hybridized carbon atoms/total carbon atoms

[f]RSC, resonant structure count

[g]HWI, hyper-Wiener index

Table 2. Structure, activity, and molecular descriptors of a hydroxyquinoline compound series



12: R^1 = Cl, R^2 = Cl, R^3 = Cl
13: R^1 = Cl, R^2 = Cl, R^3 = H
14: R^1 = Br, R^2 = Br, R^3 = H
15: R^1 = H, R^2 = Br, R^3 = H
16: R^1 = H, R^2 = Cl, R^3 = H
17: R^1 = I, R^2 = Cl, R^3 = H

Cmpd[a]	Cmpd Name	W.T.[b] EC_{50} (μM)	Δ*tolC* EC_{50} (μM)	F.C.[c]	M.W. (g mol^{-1})	clogP	PSA[d] ($Å^2$)	Fsp^3 [e]	RSC [f]	HWI [g]
12	ID-0366	>50	0.644	>50	248.49	3.55	33.1	0.00	3	623
13	Chloroxine	>50	14.9	>6.73	214.05	2.84	33.1	0.00	4	476
14	Broxyquinoline	3.72	0.168	22.1	302.95	3.08	33.1	0.00	4	476
15	ID-0763	16.1	1.41	11.4	224.06	2.36	33.1	0.00	4	370
16	Cloxyquin	11.7	2.17	5.39	179.6	2.24	33.1	0.00	4	370
17	Clioquinol	2.64	0.560	4.72	305.5	2.67	33.1	0.00	4	476

[a]Cmpd, compound

[b]W.T., wild-type *E. coli*

[c]F.C., fold-change

[d]PSA, polar surface area

[e]Fsp^3, ratio of sp^3 hybridized carbon atoms/total carbon atoms

[f]RSC, resonant structure count

[g]HWI, hyper-Wiener index

Supplemental Figures

Figure S1

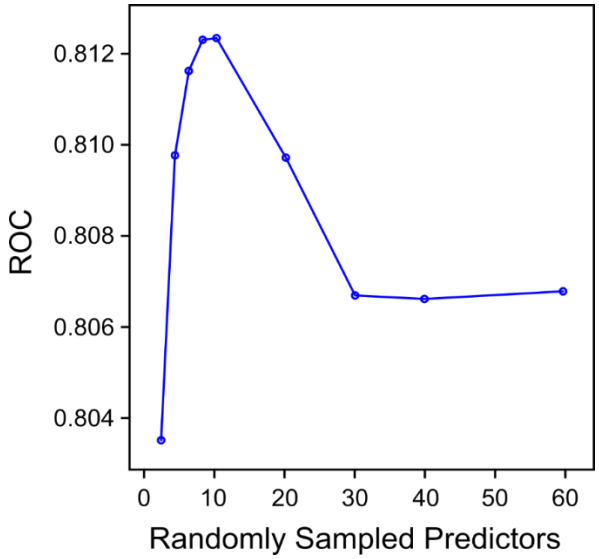


Figure S2

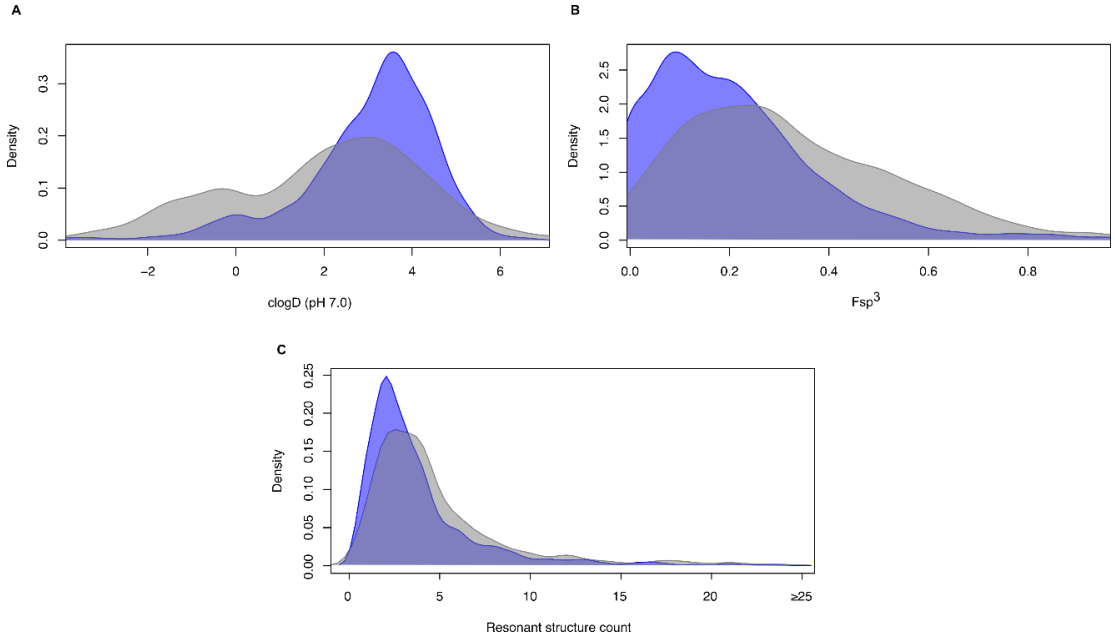


Figure S3

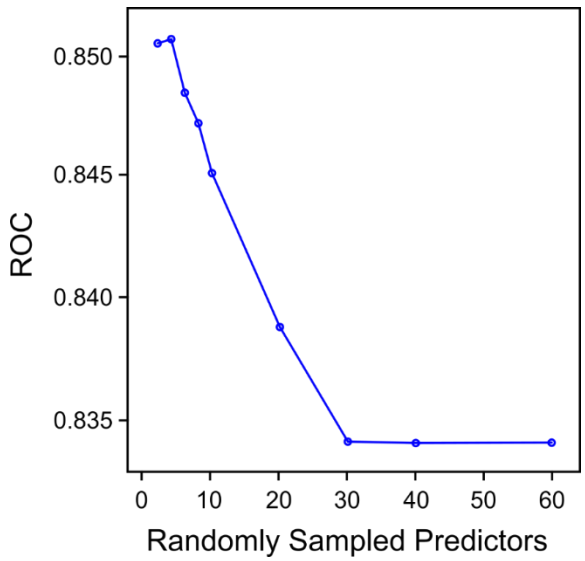
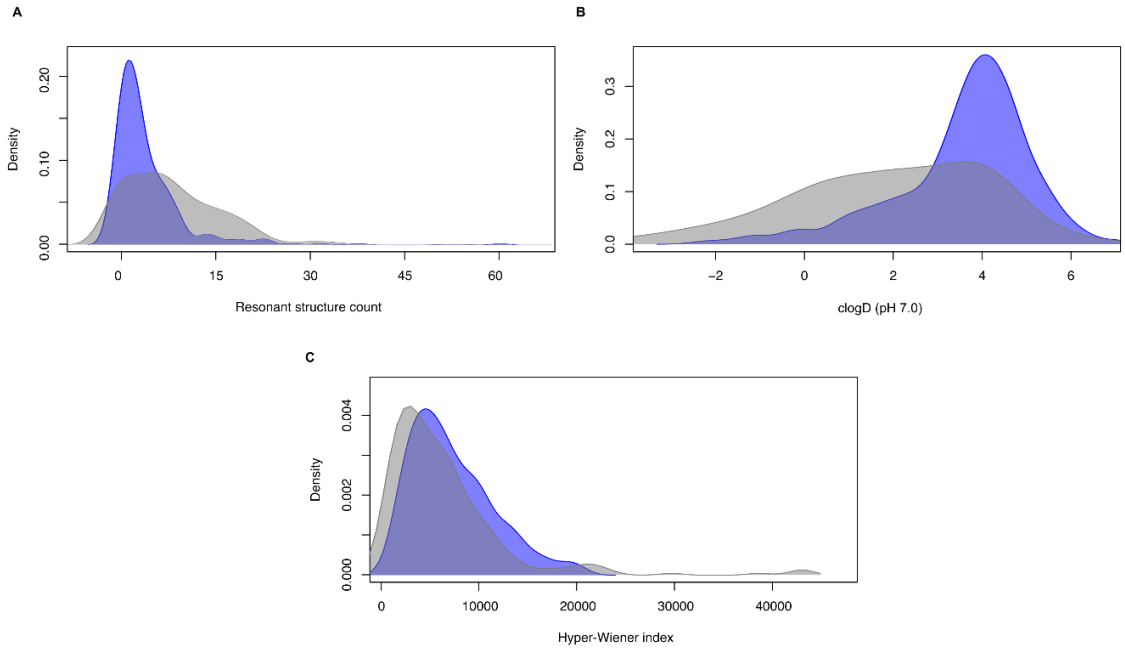
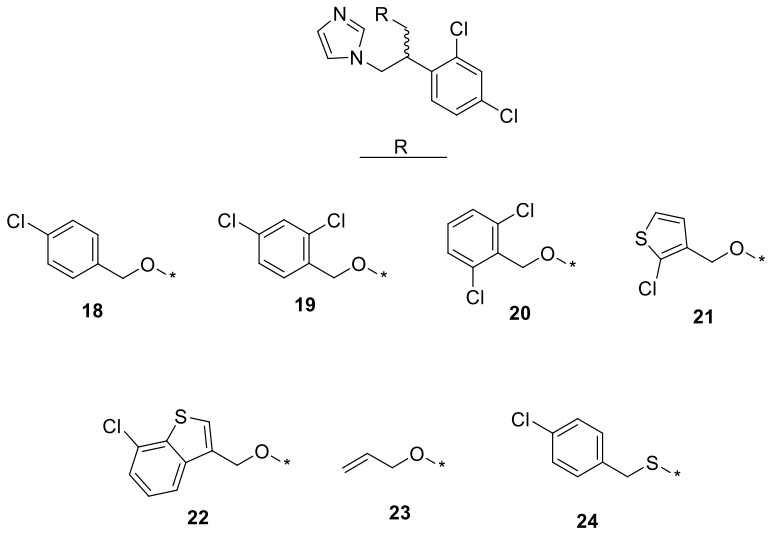


Figure S4



Supplemental Tables

Table S4. Structure, activity, and molecular descriptors of an antifungal compound series



Cmpd[a]	Cmpd Name	W.T.[b] EC_{50} (μM)	Δ*tolC* EC_{50} (μM)	F.C.[c]	M.W. (g mol^{-1})	clogP	PSA[d] ($Å^2$)	Fsp^3 [e]	RSC [f]	HWI [g]
18	Econazole	>50	4.34	>23.1	381.68	4.24	27.1	0.17	2	5,290
19	Miconazole	>50	4.80	>20.8	416.12	4.85	27.1	0.17	2	5,856
20	Isoconazole	>50	6.44	>15.5	416.12	4.85	27.1	0.17	2	5,561
21	Tioconazole	>50	6.50	>15.4	387.7	4.23	55.3	0.19	2	4,297
22	Sertaconazole	>50	7.45	>13.4	437.76	5.08	55.3	0.15	2	7,432
23	Enilconazole	>50	24.6	>4.07	297.18	2.90	27.1	0.21	2	2,273
24	Sulconazole	11.98	5.30	2.3	397.74	5.52	43.1	0.17	2	5,290

[a]Cmpd, compound

[b]W.T., wild-type *E. coli*

[c]F.C., fold-change

[d]PSA, polar surface area

[e]Fsp^3, ratio of sp^3 hybridized carbon atoms/total carbon atoms

[f]RSC, resonant structure count

[g]HWI, hyper-Wiener index

CHAPTER IV – Conclusion

Summary and future directions

The work described in this thesis, although wide in scope, highlights the utility of exploiting nutrient stress to understand the complex network that underpins bacterial growth as well as addresses efflux susceptibility of small molecules. Chapter 2 describes the first chemical biology approach to chart interactions between functions essential for viability in *E. coli* under nutrient stress. Systematic chemical combinations of antibiotics and chemical probes revealed a highly connected network in bacteria. Hierarchical clustering of the chemical-chemical interaction profiles showed that compounds clustered based on their chemical class. This simple approach generated a chemical-chemical interaction matrix that can be harnessed as a tool to elucidate the MOA of novel antimicrobial agents (unknowns). Indeed, chemical combinations and interaction matrices have proven their utility in elucidating the target and MOA of unknowns (Farha & Brown, 2010). Thus, characterizing the interaction of compounds of unknown mechanism of action with the 45 known compounds presented in Chapter 2 could be used to generate signature interaction profiles. Given that the chemical-chemical interaction matrix clusters compounds with similar MOA, hierarchical clustering of the interaction profiles of unknowns and the 45 compounds could provide testable hypotheses for the compound in question.

Chapter 2 also highlights three interactions that revealed some connectivity between biotin and fatty acid syntheses, amino acid biosynthesis and ribosome assembly, as well as purine synthesis and translation inhibition. The mechanism behind these synergies could be studied in order to further understand their functional relationships. The first

interaction shows an interdependence between biotin and fatty acids. The superpathway of fatty acid synthesis can be generally divided into two stages: fatty acid initiation and fatty acid elongation. As described in Chapter 2, cerulenin targets FabB, which is an enzyme involved in fatty acid initiation (Keseler et al., 2016). Given that biotin is used as a cofactor in the first step of fatty acid initiation (Keseler et al., 2016), I have since assessed whether the synergy between MAC13772 and cerulenin is particularly due to the inhibition of this stage. As such, I assessed the chemical combination of MAC13772 and triclosan, a fatty acid synthesis inhibitor that targets enoyl-ACP reductase (FabI), which is involved in fatty acid elongation. Figure 1 reveals additivity (FICI $\leq$0.6), suggesting that the synergistic interaction between MAC13772 and cerulenin highlights a more specific relationship between biotin and fatty acid initiation.

The remaining two synergies address functional relationships between nutrient metabolism, ribosome assembly and bacterial translation. Primarily, the interaction between lamotrigine and L-norleucine highlights a key role for the ribosome biogenesis function of IF2 in SAM-limited growth conditions. As previously mentioned, L-norleucine inhibits the synthesis of SAM, an essential methyl donor for multiple intracellular methylation sites, including ribosomal RNA (rRNA) and ribosomal proteins (r-proteins). The methylation of rRNAs and r-proteins is one of the maturation steps that occur in ribosome biogenesis. The function of the methylated sites remains unknown; however, it has been suggested that they contribute towards the stability of the ribosome (Decatur & Fournier, 2002). Consequently, the synergy suggests that the inhibition of SAM synthesis cascades to a perturbation in the overall methylation of the ribosome,

which in turn may be compensated for by the ribosome biogenesis function of IF2. As a result, lamotrigine gains growth inhibitory activity by binding to IF2 and inhibiting this compensatory role. In fact, some *in vitro* ribosome assembly work has shown that ribosome biogenesis has redundant assembly pathways that become essential under certain stress conditions, such as suboptimal growth temperatures (Williamson, 2008). Similarly, the synergy between lamotrigine and L-norleucine reveals an IF2-dependent ribosome assembly pathway in growth conditions lacking sufficient SAM pools.

The synergy between 6-mercaptopurine and aminoglycosides, on the other hand, highlights a key functional relationship between the synthesis of nucleobases and their incorporation into rRNAs. Particularly, sequence-specific hydrogen bonds between 4,5-disubstituted and 4,6-disubstituted aminoglycosides and their target rRNA have shown a purine-rich binding pocket (Recht & Puglisi, 2001). Since 6-mercaptopurine has been suggested to incorporate into RNA, these purine residues may be substituted with thiopurine analogs, which may lead to further perturbation of bacterial translation in the presence of aminoglycosides. Overall, identifying and characterizing suppressor mutants against both synergistic interactions would provide further insight into their mode of action.

With nutrient metabolism emerging as a potential antibacterial target in *M. tuberculosis* (Berney et al., 2015) and in pathogenic bacteria, such as *S. aureus* (Mei et al., 1997), *Shigella flexneri* (Cersini et al., 1998), and *Burkholderia pseudomallei* (Cuccui et al., 2007), some synergistic interactions discussed in Chapter 2 could be exploited for potential combination therapies. Notably, our lab recently validated biotin synthesis as an

in vivo target in the priority pathogens *Acinetobacter baumannii*, *K. pneumoniae*, and *P. aeruginosa* (Carfrae et al., 2020). Consequently, the synergistic interaction between the biotin inhibitor, MAC13772, and the fatty acid synthesis inhibitor, cerulenin, would be a promising candidate for combination therapy. Fatty acids have been known for their essential role in the synthesis of phospholipids (Keseler et al., 2016), which are incorporated into bacterial cell membranes. As such, enzymes catalyzing fatty acid synthesis have been recognized as potential antibacterial targets (Parker et al., 2020; Yao & Rock, 2017). Due to cytotoxicity and resistance development concerns (Wright & Reynolds, 2007; Yao & Rock, 2017), however, many inhibitors of fatty acid synthesis have yet to reach the clinical pipeline. Combination therapy would alleviate these concerns since efficacy could be achieved with dose-sparing treatments and the evolution of resistance would be reduced (Munck et al., 2014; Tyers & Wright, 2019).

Chapter 3 introduces a discussion of the molecular descriptors and side chain decorations of small molecules susceptible to efflux pumps. Here, we employ multivariate, machine learning, and structure-based clustering approaches to identify physicochemical properties and structural moieties of small molecules that render them susceptible to efflux. These analyses revealed that hydrophobic and planar molecules with low molecular stability have antibacterial activity only in efflux-compromised *E. coli*, and compounds with reduced branching and compactness showed increased susceptibility to efflux. Further, structure-based clustering revealed a β-lactam and hydroxyquinoline compound series, where structure-activity relationship analyses of these series revealed that polar acylamino side chains and some halogenated aromatic rings reduced efflux

susceptibility. Notably, trends in the hydrophobic nature, molecular stability, and molecular complexity of compounds within these series correlated with the machine learning results. Together, this work offers suggestions for medicinal chemistry efforts in hit-to-lead antibacterial programs aimed at enhancing antibacterial potency and reducing efflux susceptibility. Polar acylamino groups, for example, could be used as side chain modifications for some compounds with poor Gram-negative antibacterial activity due to their susceptibility to efflux pumps. In fact, a similar approach has been employed to modify the Gram-positive specific inhibitor deoxynybomycin and establish Gram-negative antibacterial activity (Richter et al., 2017). Further, the β-lactam and hydroxyquinoline compound series could be assessed for efflux susceptibility in some pathogenic Gram-negative bacteria, such as *Salmonella* Typhimurium, *K. pneumoniae*, and *P. aeruginosa*. This analysis will reveal whether the physicochemical properties and side chain decorations identified in Chapter 3 could serve as overarching guidelines for efflux evasion in all Gram-negative bacteria.

Of the two machine learning approaches discussed in Chapter 3, I assessed the predictive power of the Susceptibility to Efflux Random Forest (SERF) model in identifying compounds susceptible to efflux pumps. Given its high accuracy within its prediction, SERF is the first machine learning model that could be used as a tool to rapidly assess the molecular descriptors of small molecules and identify those susceptible to efflux *in silico*. As such, this tool can be used to assess a structurally diverse subset of molecules from the PubChem collection, which consists of more than 100 million compounds. Consequently, compounds strongly predicted to be susceptible to efflux pumps can be empirically tested

for efflux susceptibility by measuring their fold-change in potency and accumulation in efflux-proficient and efflux-deficient strains of *E. coli*. These results can be used to re-train SERF in order to further optimize its predictive power. Notably, measuring compound accumulation by LC-MS/MS has been recently exploited to identify physicochemical properties contributing to good accumulation and antibacterial activity (Iyer et al., 2017; Zgurskaya & Rybenkov, 2020). In fact, this method has been considered the gold standard approach for such analysis due to its broad applicability and high accuracy (Silver, 2016; Zgurskaya & Rybenkov, 2020). For example, Davis *et al.* measured the accumulation of sulfonyl adenosines in *E. coli*, *B. subtilis*, and *Mycobacterium smegmatis* by LC-MS/MS and identified that ring content and size improved accumulation (Davis et al., 2014). Similarly, Richet *et al.* revealed that molecules most likely to accumulate in *E. coli* are amphiphilic, rigid, and have low globularity (Richter et al., 2017). As such, accumulation analysis could be used as an additional parameter in SERF thus enhancing its prediction of compounds susceptible to efflux. Further, SERF provides an opportunity to systematically design a chemical collection of Gram-negative antibacterial chemical matter with low efflux susceptibility. This collection would be tested for antibacterial activity in efflux-proficient *E. coli*, where potent molecules could be investigated against Gram-negative pathogens and assessed for cytotoxicity. Compounds showing a good spectrum of growth inhibitory activity and low cytotoxicity would be prioritized for infection models in order to assess their efficacy *in vivo*.

Concluding remarks

The work presented herein aims to chart the cellular network that underpins nutrient stress in *E. coli* and identify molecular descriptors and structural moieties that render small molecules susceptible to efflux. The large fraction of interactions among characterized chemical probes suggests a dense network among essential functions. In the long term, the chemical-chemical interaction matrix can be used as a tool to elucidate the mode of action of novel antimicrobial agents. Furthermore, highlighted synergies can be assessed for their mechanism of action in order to further understand the underlying functional relationships. Additionally, some synergistic interactions within the interaction matrix offer potential combination therapies that could be further explored *in vitro* and *in vivo* across different Gram-negative bacteria. On the other hand, screening platforms continue to be optimized to discover novel antimicrobials against Gram-negative bacteria; however, efflux pumps have hampered such efforts. This stems from an incomplete understanding of efflux susceptibility in Gram-negative bacteria. I have presented some physicochemical properties and side chain decorations that contribute to the efflux susceptibility of small molecules. Further, this work highlights the first machine learning tool, SERF, that shows potential in rapidly identifying molecules susceptible to efflux *in silico*. This model allows for a systematic approach in designing novel molecules with potent Gram-negative activity and reduced efflux susceptibility.

References

Berney, M., Berney-Meyer, L., Wong, K.-W., Chen, B., Chen, M., Kim, J., Wang, J., Harris, D., Parkhill, J., Chan, J., Wang, F., & Jacobs, W. R. (2015). Essential roles

of methionine and S-adenosylmethionine in the autarkic lifestyle of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A., 112*(32), 10008-10013.

Carfrae, L. A., MacNair, C. R., Brown, C. M., Tsai, C. N., Weber, B. S., Zlitni, S., Rao, V. N., Chun, J., Junop, M. S., Coombes, B. K., & Brown, E. D. (2020). Mimicking the human environment in mice reveals that inhibiting biotin biosynthesis is effective against antibiotic-resistant pathogens. *Nat. Microbiol., 5*(1), 93-101.

Cersini, A., Salvia, A. M., & Bernardini, M. L. (1998). Intracellular multiplication and virulence of *Shigella flexneri* auxotrophic mutants. *Infect. Immun., 66*(2), 549-557.

Cuccui, J., Easton, A., Chu, K. K., Bancroft, G. J., Oyston, P. C. F., Titball, R. W., & Wren, B. W. (2007). Development of signature-tagged mutagenesis in *Burkholderia pseudomallei* to identify genes important in survival and pathogenesis. *Infect. Immun., 75*(3), 1186-1195.

Davis, T. D., Gerry, C. J., & Tan, D. S. (2014). General platform for systematic quantitative evaluation of small-molecule permeability in bacteria. *ACS Chem. Biol., 9*(11), 2535-2544.

Decatur, W. A., & Fournier, M. J. (2002). rRNA modifications and ribosome function. *Trends Biochem. Sci., 27*(7), 344-351.

Farha, M. A., & Brown, E. D. (2010). Chemical probes of *Escherichia coli* uncovered through chemical-chemical interaction profiling with compounds of known biological activity. *Chem. Biol., 17*(8), 852-862.

Iyer, R., Sylvester, M. A., Velez-Vega, C., Tommasi, R., Durand-Reville, T. F., & Miller, A. A. (2017). Whole-cell-based assay to evaluate structure permeation relationships for carbapenem passage through the *Pseudomonas aeruginosa* porin OprD. *ACS Infect. Dis., 3*(4), 310-319.

Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., Latendresse, M., Muñiz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., & Karp, P. D. (2016). The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res., 45*(D1), D543-D550.

Mei, J.-M., Nourbakhsh, F., Ford, C. W., & Holden, D. W. (1997). Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol. Microbiol., 26*(2), 399-407.

Munck, C., Gumpert, H. K., Wallin, A. I. N., Wang, H. H., & Sommer, M. O. A. (2014). Prediction of resistance development against drug combinations by collateral responses to component drugs. *Sci. Transl. Med., 6*(262), 262ra156-262ra156.

Parker, E. N., Drown, B. S., Geddes, E. J., Lee, H. Y., Ismail, N., Lau, G. W., & Hergenrother, P. J. (2020). Implementation of permeation rules leads to a FabI inhibitor with activity against Gram-negative pathogens. *Nat. Microbiol., 5*(1), 67-75.

Recht, M. I., & Puglisi, J. D. (2001). Aminoglycoside resistance with homogeneous and heterogeneous populations of antibiotic-resistant ribosomes. *Antimicrob. Agents Chemother., 45*(9), 2414-2419.

Richter, M. F., Drown, B. S., Riley, A. P., Garcia, A., Shirai, T., Svec, R. L., & Hergenrother, P. J. (2017). Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature, 545*, 299-304.

Silver, L. L. (2016). A gestalt approach to Gram-negative entry. *Bioorg. Med. Chem., 24*(24), 6379-6389.

Tyers, M., & Wright, G. D. (2019). Drug combinations: a strategy to extend the life of antibiotics in the 21st century. *Nat. Rev. Microbiol., 17*(3), 141-155.

Williamson, J. R. (2008). Biophysical studies of bacterial ribosome assembly. *Curr. Opin. Struct. Biol., 18*(3), 299-304.

Wright, H. T., & Reynolds, K. A. (2007). Antibacterial targets in fatty acid biosynthesis. *Curr. Opin. Microbiol., 10*(5), 447-453.

Yao, J., & Rock, C. O. (2017). Bacterial fatty acid metabolism in modern antibiotic discovery. *Biochim. Biophys. Acta. Mol. Cell Biol. Lipids, 1862*(11), 1300-1309.

Zgurskaya, H. I., & Rybenkov, V. V. (2020). Permeability barriers of Gram-negative pathogens. *Ann. N.Y. Acad. Sci., 1459*(1), 5-18.

Figure Legends

Figure 1. Additivity of MAC13772 with triclosan highlights importance of biotin availability for the initiation of fatty acid synthesis. The additive phenotype (FICI ≤0.6) of MAC13772 and triclosan; wild-type *E. coli* (BW25113) grown in M9 minimal medium.

Figures

Figure 1