

Methods in
Molecular Biology 1520

Springer Protocols



Peter Sass *Editor*

Antibiotics

Methods and Protocols

 Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor

John M. Walker

School of Life and Medical Sciences

University of Hertfordshire

Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
<http://www.springer.com/series/7651>

Antibiotics

Methods and Protocols

Edited by

Peter Sass

*Interfaculty Institute for Microbiology and Infection Medicine,
Microbial Bioactive Compounds, University of Tübingen, Tübingen, Germany*

 **Humana Press**

Editor

Peter Sass

Interfaculty Institute for Microbiology

and Infection Medicine

Microbial Bioactive Compounds

University of Tübingen

Tübingen, Germany

ISSN 1064-3745

Methods in Molecular Biology

ISBN 978-1-4939-6632-5

DOI 10.1007/978-1-4939-6634-9

ISSN 1940-6029 (electronic)

ISBN 978-1-4939-6634-9 (eBook)

Library of Congress Control Number: 2016956545

© Springer Science+Business Media New York 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Humana Press imprint is published by Springer Nature

The registered company is Springer Science+Business Media LLC

The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

The increasing prevalence of antibiotic-resistant microbes challenges modern medicine, posing serious threats to human and animal health. Therefore, we desperately need new antibiotics with novel mechanisms of action and resistance-breaking properties. This edition of *Antibiotic Research Protocols* in the Springer series *Methods in Molecular Biology* aims to provide state-of-the-art and novel methods on antibiotic isolation and purification, identification of antimicrobial killing mechanisms, and methods for the analysis and detection of microbial adaptation strategies. Accordingly, the chapters are organized under three major themes: Production and Design, Mode of Action, and Response and Susceptibility. In the first part on antibiotic production and design, contributions report on strategies to find new antibacterial compounds by mining bacterial genomes for antibiotic biosynthesis clusters with novel characteristics (Chapter 2) or on the production of such compounds (Chapter 3). With a new compound in hand, structure elucidation is important for compound characterization (Chapter 4) and provides the basis for further optimization, i.e., by structure- and ligand-based drug design (Chapter 5) to improve compound-target interactions in order to inhibit essential biological pathways more efficiently. To be considered as a promising new therapeutic, the selected compound should have no cytotoxic activity to eukaryotic cells. This can be monitored by methods provided in Chapter 6 and should be assessed before further steps are taken. Contributions gathered in the second part will lead the reader through methods to explore the mechanism of action of antibiotics and how to screen compound libraries for hits with a desired activity to inhibit a specific biological pathway or essential enzyme reaction. The bacterial cell envelope is a validated target for antibiotics. To detect antibiotics that interfere with cell wall integrity and synthesis, their ability to induce specific cell wall stress-responsive promoter fusions can be measured (Chapter 7). Interaction with the bacterial membrane (Chapters 8, 9, and 10) is a further means of antibiotic action that can influence membrane potential and fluidity and thus may disturb the correct function of essential membrane-associated protein machineries. Antibiotics are also frequently found to inhibit the metabolism of DNA. Here, Chapter 11 gives a detailed overview of cell-based and enzymatic assays, which can be used to screen for new inhibitors of DNA metabolism in bacteria. Moving on to RNA metabolism and translation, Chapter 12 explains a new screening method for drugs that inhibit ribonuclease P, an essential endonuclease that catalyzes the 5' end maturation of precursor tRNA, which is a necessary step prior to translation catalyzed by the ribosome. Further in this direction, Chapters 13 and 14 report on screening for translation initiation inhibitors that interfere either with regulatory elements, so-called riboswitches, or with the ribosomal initiation complex. A different target area is covered by Chapter 15, which explains a method to search for inhibitors of bacterial histidine kinases, some of which either are essential for survival or were found to be involved in the development of antibiotic resistance. Understanding the bacterial response to antibiotics as well as the established resistance mechanisms to clinically used drugs is mandatory to evaluate alternatives to commonly applied treatment strategies, which is addressed in the third part of this book. Here, global expression profiling methods to study the transcriptome (Chapter 16) or the proteome

(Chapter 17) of susceptible versus resistant strains (or untreated versus antibiotic-treated strains) have the potential to uncover the underlying antibiotic modes of action as well as resistance mechanisms. Chapter 18 goes further into this direction and characterizes alterations in the stoichiometry and composition of ribosomal and ribosome-associated proteins during antibiotic stress, which impact on protein expression profiles or hibernating ribosomes. Heading more towards antibacterial resistance, functional metagenomics emerged as a useful way to identify novel resistance genes from environmental samples (Chapter 19), which do not necessarily rely on the culturability of a specific strain in the laboratory, thus allowing to study antibiotic resistance in diverse microbial communities such as soil-, marine-, human-, wastewater-, or animal- and agriculture-associated communities. In addition to the identification of resistance factors, it is a prerequisite to track antibiotic-resistant strains by epidemiological surveillance and typing methods to understand and evaluate how resistance and its spread functions. Several typing methods are discussed in Chapter 20 including those using high-throughput sequencing technologies to identify epidemiological markers as well as antibiotic resistance and virulence determinants.

I would like to thank all contributing authors for sharing their expertise and protocols, albeit knowing to have left out other important topics that would have deserved the same attention, which again is solely assigned to the limitations that are inherent to the selection process for assembling this book. As antibiotic research is a multidisciplinary approach, this book addresses scientists from diverse fields involving microbiologists, cell biologists, molecular geneticists, pharmacists, immunologists, infectiologists, biochemists, biophysicists, bioinformaticians, and many others. We hope that the book will inspire your scientific work in the exciting field of antibiotic research, and we would be pleased to see the book more often in your lab than in the library.

Tübingen, Germany

Peter Sass

Contents

Preface	v
Contributors	ix

PART I PRODUCTION AND DESIGN

1 Antibiotics: Precious Goods in Changing Times	3
<i>Peter Sass</i>	
2 Mining Bacterial Genomes for Secondary Metabolite Gene Clusters	23
<i>Martina Adamek, Marius Spohn, Evi Stegmann, and Nadine Ziemert</i>	
3 Production of Antimicrobial Compounds by Fermentation	49
<i>Henrik Harms, Gabriele M. König, and Till F. Schäberle</i>	
4 Structure Elucidation of Antibiotics by NMR Spectroscopy	63
<i>Georgios Daletos, Elena Ancheeva, Raha S. Orfali, Victor Wray, and Peter Proksch</i>	
5 Computer-Aided Drug Design Methods	85
<i>Wenbo Yu and Alexander D. MacKerell Jr.</i>	
6 Cytotoxicity Assays as Predictors of the Safety and Efficacy of Antimicrobial Agents	107
<i>Alexander Zipperer and Dorothee Kretschmer</i>	

PART II MODE OF ACTION

7 Application of a <i>Bacillus subtilis</i> Whole-Cell Biosensor (<i>PliaI-lux</i>) for the Identification of Cell Wall Active Antibacterial Compounds	121
<i>Carolin Martina Kobras, Thorsten Mascher, and Susanne Gebhard</i>	
8 Determination of Bacterial Membrane Impairment by Antimicrobial Agents	133
<i>Miriam Wilmes and Hans-Georg Sahl</i>	
9 Mass-Sensitive Biosensor Systems to Determine the Membrane Interaction of Analytes	145
<i>Sebastian G. Hoß and Gerd Bendas</i>	
10 Measurement of Cell Membrane Fluidity by Laurdan GP: Fluorescence Spectroscopy and Microscopy	159
<i>Kathi Scheinpflug, Oxana Krylova, and Henrik Strahl</i>	
11 In Vitro Assays to Identify Antibiotics Targeting DNA Metabolism	175
<i>Allan H. Pang, Sylvie Garneau-Tsodikova, and Oleg V. Tsodikov</i>	
12 Fluorescence-Based Real-Time Activity Assays to Identify RNase P Inhibitors	201
<i>Yu Chen, Xin Liu, Nancy Wu, and Carol A. Fierke</i>	

13	Reporter Gene-Based Screening for TPP Riboswitch Activators	227
	<i>Christina E. Lünse and Günter Mayer</i>	
14	Cell-Based Fluorescent Screen to Identify Inhibitors of Bacterial Translation Initiation	237
	<i>Federica Briani</i>	
15	Bacterial Histidine Kinases: Overexpression, Purification, and Inhibitor Screen	247
	<i>Mike Gajdiss, Michael Türck, and Gabriele Bierbaum</i>	
PART III RESPONSE AND SUSCEPTIBILITY		
16	Expression Profiling of Antibiotic-Resistant Bacteria Obtained by Laboratory Evolution	263
	<i>Shingo Suzuki, Takaaki Horinouchi, and Chikara Furusawa</i>	
17	Sample Preparation for Mass-Spectrometry Based Absolute Protein Quantification in Antibiotic Stress Research	281
	<i>Florian Bonn, Sandra Maass, and Dörte Becher</i>	
18	Label-Free Quantitation of Ribosomal Proteins from <i>Bacillus subtilis</i> for Antibiotic Research	291
	<i>Sina Schäkermann, Pascal Prochnow, and Julia E. Bandow</i>	
19	Functional Metagenomics to Study Antibiotic Resistance	307
	<i>Manish Boolchandani, Sanket Patel, and Gautam Dantas</i>	
20	Epidemiological Surveillance and Typing Methods to Track Antibiotic Resistant Strains Using High Throughput Sequencing	331
	<i>Miguel Paulo Machado, Bruno Ribeiro-Gonçalves, Mickael Silva, Mário Ramirez, and João André Carrizo</i>	
	<i>Index</i>	357

Contributors

MARTINA ADAMEK • *Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbiology/Biotechnology, University of Tübingen, Tübingen, Germany; German Centre for Infection Research (DZIF), Partner Site Tübingen, Tübingen, Germany*

ELENA ANCHEEVA • *Institute of Pharmaceutical Biology and Biotechnology, Heinrich-Heine-University, Düsseldorf, Germany*

JULIA E. BANDOW • *Applied Microbiology, Ruhr-Universität Bochum, Bochum, Germany*

DÖRTE BECHER • *Department for Microbial Proteomics, Institute for Microbiology, Ernst-Moritz-Arndt University of Greifswald, Greifswald, Germany*

GERD BENDAS • *Pharmaceutical Chemistry II, Pharmaceutical Institute, University of Bonn, Bonn, Germany*

GABRIELE BIERBAUM • *Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Bonn, Germany*

FLORIAN BONN • *Department for Microbial Proteomics, Institute for Microbiology, Ernst-Moritz-Arndt University of Greifswald, Greifswald, Germany*

MANISH BOOLCHANDANI • *Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA*

FEDERICA BRIANI • *Dipartimento di Bioscienze, Università degli Studi di Milano, Milan, Italy*

JOÃO ANDRÉ CARRIÇO • *Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Alameda Da Universidade, Lisbon, Portugal*

YU CHEN • *Department of Chemistry, University of Michigan, Ann Arbor, MI, USA*

GEORGIOS DALETOS • *Institute of Pharmaceutical Biology and Biotechnology, Heinrich-Heine-University, Düsseldorf, Germany*

GAUTAM DANTAS • *Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA; Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA; Department of Biomedical Engineering, Washington University, St. Louis, MO, USA; Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA*

CAROL A. FIERKE • *Department of Chemistry, University of Michigan, Ann Arbor, MI, USA; Chemical Biology Program, University of Michigan, Ann Arbor, MI, USA; Department of Biological Chemistry, University of Michigan, Ann Arbor, MI, USA*

CHIKARA FURUSAWA • *Quantitative Biology Center, Suita, Osaka, Japan*

MIKE GAJDISS • *Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Bonn, Germany*

SYLVIE GARNEAU-TSODIKOVA • *Department of Pharmaceutical Sciences, University of Kentucky, Lexington, KY, USA*

SUSANNE GEBHARD • *Department of Biology and Biochemistry, Milner Centre for Evolution, University of Bath, Bath, UK*

- HENRIK HARMS • *Department of Pharmaceutical Biology, University of Bonn, Bonn, Germany*
- TAKAAKI HORINOUCHI • *Quantitative Biology Center, Suita, Osaka, Japan*
- SEBASTIAN G. HOß • *Pharmaceutical Chemistry II, Pharmaceutical Institute, University of Bonn, Bonn, Germany*
- CAROLIN MARTINA KOBRAS • *Department of Biology and Biochemistry, Milner Centre for Evolution, University of Bath, Bath, UK*
- GABRIELE M. KÖNIG • *Department of Pharmaceutical Biology, University of Bonn, Bonn, Germany*
- DOROTHEE KRETSCHMER • *Department of Infection Biology, Interfaculty Institute for Microbiology and Infection Medicine Tübingen (IMIT), University of Tübingen, Tübingen, Germany*
- OXANA KRYLOVA • *Department of Chemical Biology, Leibniz-Institut für Molekulare Pharmakologie (FMP), Berlin, Germany*
- XIN LIU • *Department of Chemistry, University of Michigan, Ann Arbor, MI, USA*
- CHRISTINA E. LÜNSE • *Life and Medical Sciences Institute, University of Bonn, Bonn, Germany; Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT, USA*
- SANDRA MAASS • *Department for Microbial Proteomics, Institute for Microbiology, Ernst-Moritz-Arndt University of Greifswald, Greifswald, Germany*
- MIGUEL PAULO MACHADO • *Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Alameda Da Universidade, Lisbon, Portugal*
- ALEXANDER D. MACKERELL JR. • *Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD, USA*
- THORSTEN MASCHER • *Institut für Mikrobiologie, Technische Universität Dresden, Dresden, Germany*
- GÜNTER MAYER • *Life and Medical Sciences Institute, University of Bonn, Bonn, Germany*
- RAHA S. ORFALI • *Pharmacognosy Department, Faculty of Pharmacy, King Saud University, Riyadh, Saudi Arabia*
- ALLAN H. PANG • *Department of Pharmaceutical Sciences, University of Kentucky, Lexington, KY, USA*
- SANKET PATEL • *Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA; Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA*
- PASCAL PROCHNOW • *Applied Microbiology, Ruhr-Universität Bochum, Bochum, Germany*
- PETER PROKSCH • *Institute of Pharmaceutical Biology and Biotechnology, Heinrich-Heine-University, Duesseldorf, Germany*
- MÁRIO RAMIREZ • *Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Alameda Da Univerisdade, Lisbon, Portugal*
- BRUNO RIBEIRO-GONÇALVES • *Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Alameda Da Universidade, Lisbon, Portugal*
- HANS-GEORG SAHL • *Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Bonn, Germany*
- PETER SASS • *Interfaculty Institute for Microbiology and Infection Medicine, Microbial Bioactive Compounds, University of Tübingen, Tübingen, Germany*

TILL F. SCHÄBERLE • *Department of Pharmaceutical Biology, University of Bonn, Bonn, Germany*

SINA SCHÄKERMANN • *Applied Microbiology, Ruhr-Universität Bochum, Bochum, Germany*

KATHI SCHEINPFLUG • *Department of Chemical Biology, Leibniz-Institut für Molekulare Pharmakologie (FMP), Berlin, Germany*

MICKAEL SILVA • *Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Instituto de Microbiologia, Universidade de Lisboa, Alameda Da Universidade, Lisbon, Portugal*

MARIUS SPOHN • *Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbiology/Biotechnology, University of Tübingen, Tübingen, Germany*

EVI STEGMANN • *Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbiology/Biotechnology, University of Tübingen, Tübingen, Germany; German Centre for Infection Research (DZIF), Partner Site Tübingen, Tübingen, Germany*

HENRIK STRAHL • *Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, UK*

SHINGO SUZUKI • *Quantitative Biology Center, Suita, Osaka, Japan*

OLEG V. TSODIKOV • *Department of Pharmaceutical Sciences, University of Kentucky, Lexington, KY, USA*

MICHAEL TÜRCK • *Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Bonn, Germany*

MIRIAM WILMES • *Institute of Medical Microbiology, Immunology and Parasitology, University of Bonn, Bonn, Germany*

VICTOR WRAY • *Helmholtz Centre for Infection Research, Braunschweig, Germany*

NANCY WU • *Chemical Biology Program, University of Michigan, Ann Arbor, MI, USA*

WENBO YU • *Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD, USA*

NADINE ZIEMERT • *Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Microbiology/Biotechnology, University of Tübingen, Tübingen, Germany; German Centre for Infection Research (DZIF), Partner Site Tübingen, Tübingen, Germany*

ALEXANDER ZIPPERER • *Department of Infection Biology, Interfaculty Institute for Microbiology and Infection Medicine Tübingen (IMIT), University of Tübingen, Tübingen, Germany*

Part I

Production and Design

Chapter 1

Antibiotics: Precious Goods in Changing Times

Peter Sass

Abstract

Antibiotics represent a first line of defense of diverse microorganisms, which produce and use antibiotics to counteract natural enemies or competitors for nutritional resources in their nearby environment. For antimicrobial activity, nature has invented a great variety of mechanisms of antibiotic action that involve the perturbation of essential bacterial structures or biosynthesis pathways of macromolecules such as the bacterial cell wall, DNA, RNA, or proteins, thereby threatening the specific microbial lifestyle and eventually even survival. However, along with highly inventive modes of antibiotic action, nature also developed a comparable set of resistance mechanisms that help the bacteria to circumvent antibiotic action. Microorganisms have evolved specific adaptive responses that allow appropriately reacting to the presence of antimicrobial agents, ensuring survival during antimicrobial stress. In times of rapid development and spread of antibiotic (multi-)resistance, we need to explore new, resistance-breaking strategies to counteract bacterial infections. This chapter intends to give an overview of common antibiotics and their target pathways. It will also discuss recent advances in finding new antibiotics with novel modes of action, illustrating that nature's repertoire of innovative new antimicrobial agents has not been fully exploited yet, and we still might find new drugs that help to evade established antimicrobial resistance strategies.

Key words Antimicrobial agents, Drug discovery, Mode of action, Antibiotic resistance, MRSA

1 The Times They Are Changing

Antibiotics represent one of the most important discoveries in medical history. Although the term antibiotics literally means “against life”, their use as therapeutic agents starting in the first half of the twentieth century for the first time allowed an effective treatment of even complicated, life-threatening bacterial infections (e.g., tuberculosis and septicemia), thereby saving millions of lives. In addition, antibiotics represent important tools in modern medicine to enable, e.g., cancer chemotherapy, orthopedic surgery, and transplantation. Thus, antibiotics have not just decreased mortality and morbidity caused by bacterial infections, they further revolutionized the possibilities concerning medical intervention to help increase the quality of human life.

Today, the situation has changed dramatically which is primarily due to the overuse and/or misuse of antibiotics in the clinics and to an even more serious extent in agriculture. We are now experiencing an increasing prevalence of microbes that are resistant to one, several, or almost all clinically used antibiotics. Bacterial infections that could once easily be cured are becoming harder to treat or are even untreatable, challenging modern medicine and posing one of the most serious threats to human and animal health as recognized by the World Health Organization [1]. As many pathogens (resistant and susceptible) are often characterized to sensitively respond to antibiotic stress and regulate their metabolism to survive antibacterial treatment, it can be assumed that the resistance situation will not substantially relax in the future.

In light of this alarming development and spread of antibiotic (multi-) resistance, and due to the fact that too few new antibiotics with resistance breaking properties enter the clinics (the “golden age of antibiotic discovery” has long since passed), we need fast and coordinated actions to avoid the risk of heading toward a post-antibiotic era. Such necessary actions need to be supported by physicians, pharmacists, agriculturists, scientists, politicians, and, of course, people in the community. In the community, we need a better awareness of resistance prevalence, how to avoid spread of pathogens, and when antibiotics are necessary and how they should be used. Efficient means for infection monitoring, control, and prevention are needed, accompanied by improved hygiene standards (in the community as well as in the clinic) and a far more responsible handling of antibiotic prescription and usage (in humans and livestock). In addition, a huge economic burden comes along with antimicrobial resistance with costs of approx. 55 billion USD a year only in the United States as stated by the Centers for Disease Control and Prevention (CDC) in April 2011 [2]. Whatever costs (human or economic) drive an initiative for assessing solutions to counteract antimicrobial resistance and antibiotic run out, it is an important step forward that this issue will be discussed at the UN General Assembly in 2016 and continues to rise up the agenda for the G7 and G20 groups of countries. According to The Review on Antimicrobial Resistance of March 2016 [3], which was commissioned by the UK Prime Minister and is hosted by the Wellcome Trust, without a coordinated global response to drug resistance we could be facing a threat claiming ten million lives a year by 2050, at an accumulated cost of 100 trillion USD. With the words of Professor Dame Sally Davies (UK Chief Medical Officer): “We have reached a critical point and must act now on a global scale to slow down antimicrobial resistance.”

However, decelerating development and spread of antimicrobial resistance will not be sufficient to avert this scenario, as we also need new antibiotics with extended or novel mechanisms of action that have resistance-breaking properties, i.e., improved antibiotics

of established classes and/or new compounds with novel targets. But although desperately needed, during the last decades antibiotic research suffered from questioning its economic viability and the interest of the private sector in developing new antibacterial drugs constantly decreased, which among other reasons resulted in a misguided trend that also reduced basic research on the development of new antibiotic drugs. Just recently, some academic-driven activities started to oppose this trend. In Germany, for example, a collaborative research unit of the German Research Foundation (DFG) designated “Post-Genomic Strategies for New Antibiotic Drugs and Targets” (FOR854), launched and headed by Hans-Georg Sahl (University of Bonn), started in 2009 with the goal to study promising new antibiotic compounds, molecular mechanisms, and targets using postgenome era strategies [4]. Many of the contributors to this book were involved in this research unit. In 2010, the German Federal Ministry of Education and Research initiated the German Centre for Infection Research (DZIF) that runs several task force units on infection research, one unit with a special focus on Novel Anti-infectives, which aims to bring basic research and current anti-infective development activities back together [5]. In 2011, the European Commission called for an action plan against the rising threats from antimicrobial resistance, with the goal to join research and development efforts to bring new antibiotics to the patients. As a result, the New Drugs 4 Bad Bugs (ND4BB) program was initiated in 2013, bringing industry, academia, and biotech organizations together to combat bacterial resistance in Europe. In 2006, the Gordon Research Conference on New Antibacterial Discovery & Development was launched as an excellent platform to discuss and share ideas between academia, industry, and government agencies on target discovery and validation, hit identification and chemical optimization, as well as clinical trial design and execution, thereby supporting the efforts of the above-mentioned research initiatives [6]. With a new rising commercial value due to the increasing medical need, antibiotic research may also attract more of “big pharma” back to the antibacterial drug discovery field.

2 Antibiotic Modes of Action

Microorganisms produce antibiotics for several reasons; the most prominent role is the direct counteraction of competitive or invading bacteria by inhibiting their growth. This chapter intends to highlight some of the more common antibiotics and their cellular targets, as well as examples for new antibiotics with novel modes of action identified by antibiotic researchers in recent years (Fig. 1). A large number of different antimicrobial agents have evolved which can be characterized and differentiated by their individual modes of action and cellular targets. The bacterial cell envelope represents

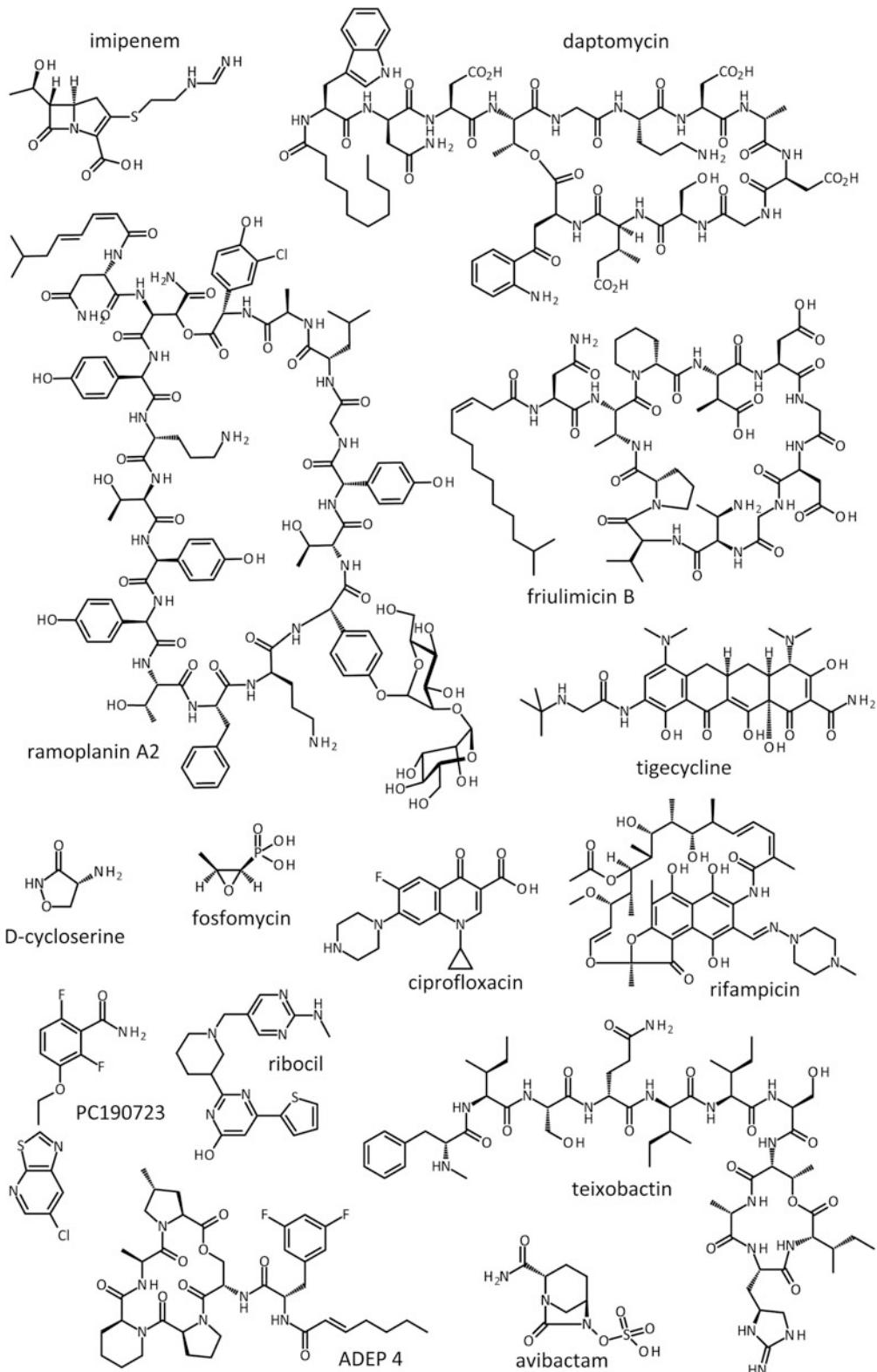


Fig. 1 Structures of exemplary antibiotics discussed in this chapter

one such prominent target pathway that is unique to bacteria and fulfills many important physiological functions. Further validated target pathways for antibiotic attack include protein biosynthesis as well as DNA and RNA metabolism, where the target sites sufficiently vary from their eukaryotic counterparts to allow selective inhibition of the bacteria. Depending on the compound and its target site, some antibiotics have a broad spectrum or a narrow spectrum of activity targeting multiple bacterial species or only a small selected group of species, respectively. Some antimicrobial agents are multi-targeted compounds that act against a broad range of microorganisms in a concentration-dependent manner. Such sanitizers possess rather unspecific modes of action, i.e., by penetrating and disturbing the integrity of the bacterial cell membrane, by producing DNA cross linkage, or by abrogating spore outgrowth [7].

However, antibacterial activity is not the only function of antibiotics as their name would suggest, and it is increasingly recognized that some members of antibiotic classes have evolved as bacterial tools for intra- and inter-domain communication [8]. For instance, antimicrobial peptides act as modulators of the innate immune response in higher organisms [9–11]. In streptomycetes, some lanthionine-containing antibiotics (*lantibiotics*) have morphogenic effects by modulating the formation of aerial hyphae [12]. Other lantibiotics induce their own production in a quorum-dependent manner, thereby showing similarities to peptide pheromones [13, 14]. Despite these exciting new functions of antibiotics, which certainly deserve and would easily fill a review on their own, the next sections will focus on the mechanisms how antibiotics exert their growth inhibitory effects or even kill bacteria.

2.1 The Bacterial Cell Envelope as Antibiotic Target

Many vital cellular functions are attributed to the bacterial cell envelope, including its role as a diffusion barrier, a shape-giving structure, and an essential communication interface between different cells in a community as well as their surrounding environment. Structurally, the cell envelope of bacteria is built up of one or two lipid membranes and the glycopeptide scaffold of the cell wall (peptidoglycan). In general, there are two different types of cell envelopes in bacteria, a Gram-positive and a Gram-negative type. The cell envelope of Gram-positive bacteria consists of a thick cell wall layer that is located outside the bacterial plasma membrane and completely covers the cell. Here, the cell wall forms a multilayered heteropolymer, which is largely composed of the sugar-peptide polymer peptidoglycan to which diverse accessory molecules like proteins, teichoic acids, teichuronic acids, carbohydrates, and polyphosphates are attached. The peptidoglycan mainly consists of linear glycan chains with the two alternating amino sugars N-acetyl-glucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc). These glycan chains are further crosslinked by

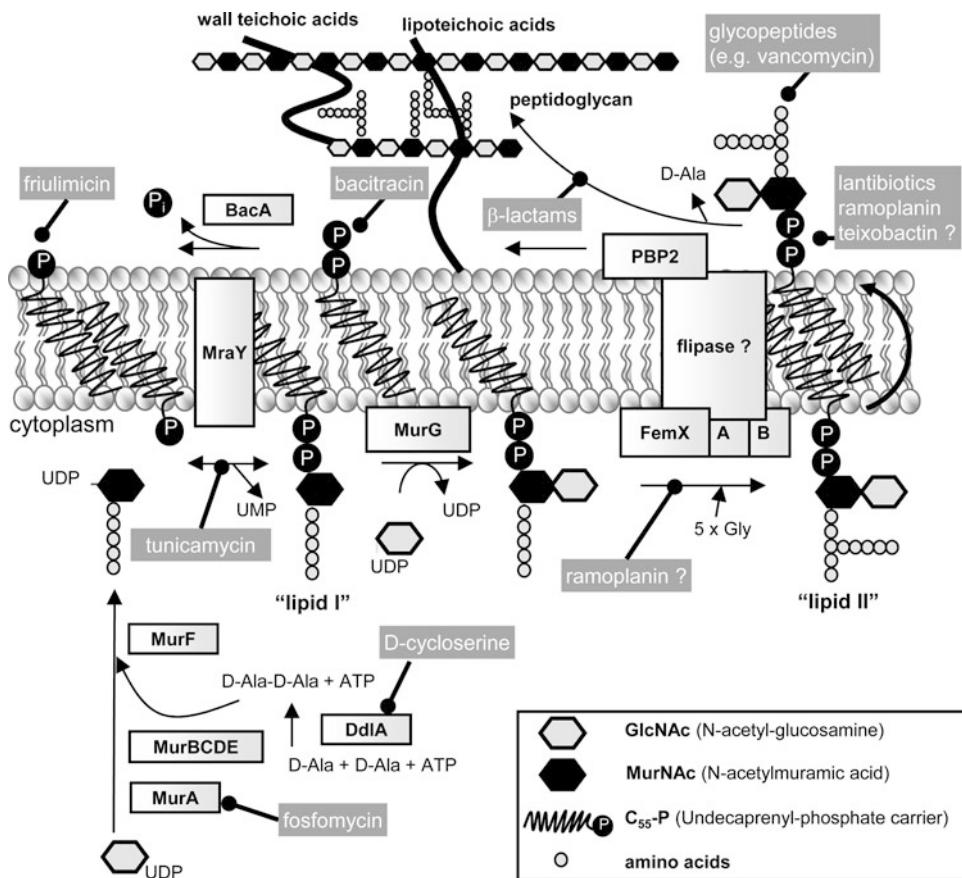


Fig. 2 Inhibition of cell wall biosynthesis of *Staphylococcus aureus* by antibiotic action

short peptides that are bound to the MurNAc moieties, and which can vary between different bacteria (Fig. 2). Depending on the bacterial species, the peptide side chains are either directly cross-linked to the peptide side chains of neighboring glycan strands, or may be connected via interpeptide bridges of characteristic additional amino acids. By that, an elastic, three-dimensional network is produced which is intimately involved in vital bacterial processes such as cell division and autolysis, and it essentially contributes to maintain bacterial cell shape and integrity by withstanding the internal osmotic pressure [15, 16]. Gram-negative bacteria possess a significantly different cell envelope architecture, as they have an additional outer lipid membrane and a much thinner peptidoglycan layer, which, however, structurally resembles the peptidoglycan of Gram-positive species. With regard to antibiotic research, the outer lipid membrane represents an intrinsic antibiotic resistance factor that acts as a diffusion barrier for a multitude of compounds including antimicrobial agents, and thus protects the bacteria from the potentially lethal effects of many antibiotics. The space between inner and outer membranes, the periplasm, harbors various enzymes that play roles in distinct physiological

pathways like peptidoglycan biosynthesis, nutrient uptake, electron transfer systems, as well as detoxification [17].

The cell wall is essential for and unique to bacteria and thus is a favored target site for antibiotics (Fig. 2). Antibiotics interfere with almost every step of peptidoglycan synthesis by either sequestering an essential substrate of the synthesis reaction or by directly inhibiting a specific enzymatic reaction, which both prevent the production of new cell wall material and finally lead to cell death of the growing bacteria [18, 19]. Penicillin G, probably the most prominent antibiotic to date, and other β -lactams interfere with enzyme reactions during peptidoglycan synthesis. To do so, β -lactams imitate the D-alanyl-D-alanine motif of the peptidoglycan precursor lipid II and bind to the active sites of transpeptidases and carboxypeptidases, which are therefore called penicillin-binding proteins (PBPs). Thereby, β -lactam binding inhibits the function of PBPs and thus prevents the cross-linking reactions of the glycan strands making the cell vulnerable to lysis during growth [20, 21]. D-cycloserine, an oxazolidinone antibiotic, inhibits both D-alanine racemase and D-alanine-D-alanine ligase and blocks the conversion of L-alanine to D-alanine and the subsequent production of D-alanyl-D-alanine dipeptide [22, 23]. Fosfomycin is a structural analog of phosphoenol pyruvate and therefore interferes with the activity of MurA, an enzyme that converts UDP-GlcNAc into UDP-MurNAc, representing the first committed step of peptidoglycan biosynthesis [24]. The enzymatic activity of MraY is inhibited by the antibiotics tunicamycin, mureidomycin, amphotomycin, muraymycin, and liposidomycin, thereby preventing the transfer of the UDP-MurNAc-pentapeptide motif onto the undecaprenyl phosphate carrier, which is the first membrane-bound step of peptidoglycan synthesis yielding lipid I [25].

Some antibiotics act by sequestering an essential substrate of cell wall biosynthesis, which leads to a significantly decreased availability of central building blocks for the synthesis of peptidoglycan. Bacitracin, a cyclic dodecylpeptide antibiotic, sequesters the lipid carriers of the essential peptidoglycan precursor molecule lipid II in its undecaprenyl-pyrophosphate state (C_{55} -PP), thereby preventing the recycling of the lipid carrier by dephosphorylation [26, 27]. The lipopeptide antibiotic friulimicin B interferes with C_{55} -P to decrease the availability of the lipid carrier and leads to an effective inhibition of the MraY reaction [28]. Antibiotics like vancomycin and teicoplanin, ramoplanin, or some lantibiotics directly bind and sequester the precursor molecule lipid II. The glycopeptide vancomycin, which is an antibiotic of last resort for the treatment of serious infections caused by methicillin-resistant *Staphylococcus aureus* (MRSA), complexes the D-alanyl-D-alanine terminus of lipid II to inhibit peptidoglycan synthesis [29]. Ramoplanin, a strongly amphipathic glycolipopeptide antibiotic with a short acyl side chain, prevents bacterial cell wall biosynthesis by interacting with the sugar phosphate head group of lipid II instead of binding to its

D-alanyl-D-alanine terminus. Ramoplanin even seems to be translocated across the membrane to inhibit intracellular lipid I- and lipid II-consuming reactions, e.g., catalyzed by MurG and FemXAB [30]. Lantibiotics are antimicrobial peptide antibiotics that undergo posttranslational modification resulting in the formation of the cyclic thioether amino acids lanthionine and 3-methyllanthionine [31, 32]. Some lantibiotics such as mersacidin, which belongs to a rather globular type of lantibiotics, also bind to the pyrophosphate group of lipid II, thereby inhibiting transglycosylation reactions and finally the incorporation of the precursor into the nascent peptidoglycan during cell wall biosynthesis [33]. Nisin, another type of lantibiotics with rather elongated structures, is commonly used as a food preservative [34]. Nisin exhibits a dual mode of action: one mode is its specific binding to the sugar-pyrophosphate moiety of lipid II to cause inhibition of cell wall biosynthesis. In addition, nisin uses lipid II as a docking molecule to insert into the bacterial membrane, which results in pore formation, ion efflux, and finally bacterial cell death [35, 36].

In contrast to lantibiotics, the mode of action of cationic antimicrobial peptides (CAMPs) appears to be mostly based on their cationic and amphiphilic nature which allows such peptides to interact with negatively charged bacterial surfaces and membranes [37]. However, mere membrane perturbation does not fully explain their antimicrobial activity. Most probably, CAMPs do not just impair the membrane but also exert unspecific, disturbing effects on multicomponent biosynthetic machineries like the peptidoglycan synthesis complex, thus acting like “sand in a gearbox” [38].

Such a “sand in a gearbox” mechanism may also account for the antimicrobial activity of the lipodepsipeptide daptomycin, one of few antibiotics that was approved for clinical use in the last two decades and is now successfully being applied to treat infections by vancomycin resistant enterococci (VRE) and MRSA. Despite the successful use of daptomycin in the clinical setting, its mode of action is incompletely understood and subject to controversial discussions hypothesizing on either a more targeted mode of action, i.e., inhibition of cell wall biosynthesis or lipoteichoic acid biosynthesis pathways, or a rather generalized killing mechanism, i.e., by membrane depolarization. It has been further discussed that Ca^{2+} -daptomycin may have a CAMP-like behavior in that it forms oligomeric structures and attaches to anionic membrane surfaces to perturb vital barrier functions of the bilayer [19]. However, the better potency of daptomycin compared to CAMPs and the fact that daptomycin specifically induces the so-called cell wall stress response in bacteria [39–41], which is commonly induced by cell wall synthesis-perturbing agents including β -lactams, glycopeptides, and mersacidin [42–44], points at a more clearly defined target in cell wall synthesis, which, however, has not yet been identified.

2.2 Inhibitors of RNA and Protein Synthesis

Bacterial transcription/translation is a further essential target pathway of various antimicrobial agents. Three different groups of antibiotics can be categorized that target different stages of the protein biosynthesis machinery (Fig. 3). One group of inhibitors interferes with DNA-dependent RNA-polymerase (RNAP) to block transcription. Another group binds to the 30S or 50S ribosomal subunits to inhibit translation initiation and elongation or affect translational accuracy. Some antibiotics interfere with tRNA synthases and elongation factors to perturb the cellular concentration of charged tRNA molecules or the delivery and release of tRNA molecules to and from the ribosome.

Bacterial DNA-dependent RNAP is an attractive antimicrobial target, since RNA synthesis is essential for bacterial growth. The semisynthetic antibiotic rifampicin (also known as rifampin) is probably the best-known representative of RNAP inhibitors, which binds to RNAP at a site adjacent to the RNAP active center and inhibits the initiation of RNA synthesis by physically blocking the formation of the phosphodiester bond in the RNA backbone by a “steric-occlusion” mechanism [45]. Other protein synthesis inhibitors directly interfere with the ribosomal complex. Aminoglycosides like streptomycin and gentamicin bind to a conserved rRNA sequence that is near the A-site of the 30S ribosomal subunit and affect the accuracy of translation. The interaction of aminoglycosides with the A-site disturbs the proofreading steps of the ribosome that are important to ensure translational fidelity and finally results in misreading of the mRNA code and the synthesis of erroneous proteins [46, 47]. Tetracyclines and glycylcyclines such as tigecycline also bind to the 30S ribosomal subunit in a region that is close to the codon-anticodon recognition site (A-site), thereby preventing the productive binding of charged aminoacyl-tRNAs to

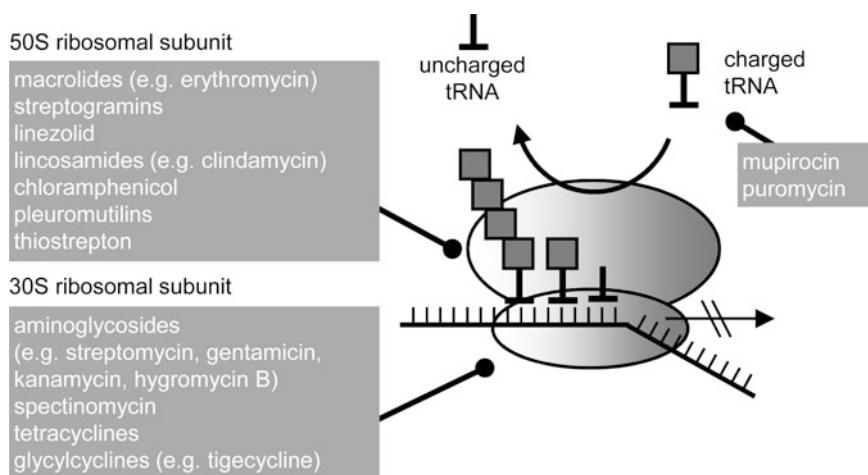


Fig. 3 Target sites of protein biosynthesis inhibitors

the ribosomal A-site [48]. Erythromycin and chloramphenicol both target the 50S ribosomal subunit and inhibit translation elongation. Erythromycin, a macrolide antibiotic, prevents the movement and release of the nascent peptide by blocking the tunnel that channels the nascent peptides away from the peptidyl transferase center [49]. Chloramphenicol prevents peptide bond formation by directly binding to the peptidyl transferase center [49–51].

Some antibiotics interfere with components in the periphery of the ribosome. For example, mupirocin, a mixture of several pseudomonic acids, blocks the activity of isoleucyl-tRNA synthetase, which prevents the charging of tRNA-Ile with the corresponding amino acid isoleucine [52, 53]. Noteworthy, the activities of aminoacyl-tRNA synthetase inhibitors like mupirocin were shown to ultimately induce the stringent response of bacteria [54, 55]. Puromycin is an aminoacyl-tRNA analog and thus incorporates into the nascent peptide chain causing premature termination and release [49]. Fusidic acid is a steroidal antibiotic that prevents the turnover of elongation factor G (EF-G) and arrests it at the ribosome after GTP hydrolysis. Although this does not lead to an inhibition of the EF-G-catalyzed translocation step of the ribosome, it will block the subsequent productive binding of new aminoacyl-tRNA to the A-site. This is because functional aminoacyl-tRNA binding requires the elongation factor Tu (EF-Tu), but EF-G and EF-Tu mutually exclude themselves regarding their presence on the same ribosome [56, 57].

2.3 Inhibitors of DNA Synthesis

Antibiotics further successfully target essential bacterial processes involved in DNA replication and turnover. Here, quinolones are probably the most famous class of DNA synthesis inhibitors, which include the antibiotics ciprofloxacin and nalidixic acid. Quinolones perturb gyrase (topoisomerase II) and topoisomerase IV activities, enzymes that play important roles in chromosome function by coiling and uncoiling of DNA. Gyrase, for example, removes knots from DNA, assists for the bending and folding of DNA, activates the chromosome for all processes involving strand separation, and even responds to environmental changes by facilitating the movement of replication and transcription complexes through DNA by adding negative supercoils in front of the complex [58]. Gyrase forms a tetramer to bind DNA in which two A- and two B-subunits wrap the DNA into a negative supercoil. In a next step, gyrase breaks and rejoins DNA strands in an ATP-dependent manner to pass one region of duplex DNA through another [59]. By introducing such negative supercoils into DNA, gyrase relieves the topological stress that occurs during the translocation of transcription and replication complexes along the DNA strand. Topoisomerase IV is a decatenating enzyme and resolves inter-linked daughter chromosomes upon DNA replication. Similar to gyrase, topoisomerase IV cleaves the phosphodiester bonds in a

DNA double strand and stabilizes the break by covalent and noncovalent interactions. Subsequently, a second double strand is passed through the open gap followed by re-ligation of the passage [60]. Here, the principles of gyrase and topoisomerase IV activities differ concerning their mechanism of DNA wrapping. While gyrase wraps DNA around itself, topoisomerase IV does not [58, 61]. Gyrase and topoisomerases play vital roles in DNA replication, transcription, repair, and recombination by ensuring an optimal level of global DNA supercoiling and removing local topological barriers, which make them essential enzymes for cell growth and division. This central role of gyrase and topoisomerase IV in DNA metabolism is one reason for the effectiveness of quinolone antibiotics. However, quinolones do not simply inhibit the enzymatic functions of these enzymes. Quinolones rather lock the enzyme complexes after cleavage of the DNA strands but before they are rejoined. By this, quinolones produce double-strand breaks that are not resealed. Thus, quinolones stall gyrase and topoisomerase IV complexes on replicative DNA strands, which eventually leads to the deleterious release of double-strand DNA breaks and fragmentation of DNA finally causing bacterial cell death.

Mitomycin C, a natural product produced by *Streptomyces caesipitosis*, is another antibiotic that efficiently interferes with DNA synthesis. Instead of inhibiting an enzyme function, mitomycin C has potent DNA cross-linking activity and catalyzes the intra- and interstrand cross-linkage of DNA strands as well as monofunctional alkyl lesions. Such cross-linking reactions inevitably lead to an arrest of DNA replication and subsequently bacterial cell death [62, 63]. However, due to its mechanism of action, which is not selectively acting against bacteria, mitomycin C is also used as an antitumor drug for the treatment of stomach, bladder, and pancreas cancer. Aside from mutations in gyrase or topoisomerase enzymes, which are a common cause of quinolone resistance, bacteria have elicited a highly effective response cascade to evade the deleterious action of DNA damaging agents including quinolones and mitomycin C, the so-called SOS response. This DNA damage-induced stress response allows the bacterial cell to minimize the lethal and mutagenic consequences of the exposure to these antibiotics and helps the cells to cope with such agents at low concentrations [58, 64–67].

2.4 Antibiotics with Extended or Novel Modes of Action

In times of an increasing spread of bacterial resistance to clinically used antibiotics, we face an urgent need to find and develop novel antibiotics with new modes of action and resistance breaking properties. This section intends to highlight some of the recent advances from academia and industry to satisfy this need, covering antibiotics that inhibit or activate novel targets to cause lethal effects, drugs that efficiently interfere with known resistance factors or even compounds that show no detectable resistance at all.

Teixobactin is a member of a new class of antibiotics that is produced by the hitherto undescribed Gram-negative soil bacterium *Eleftheria terrae* [68]. Teixobactin is antibacterially active against many pathogenic Gram-positive bacteria as well as mycobacteria including *Mycobacterium tuberculosis* but lacks activity against Gram-negatives, most probably due to ineffective penetration of the outer membrane and/or efflux. Teixobactin was also effective in reducing the bacterial load in experimental infections of MRSA and *Streptococcus pneumoniae* in mice. Teixobactin uses a dual mechanism of action that is currently not used by any clinically applied antibiotic. To kill the bacteria, teixobactin interferes with cell wall synthesis reactions at several stages by sequestering the essential precursors of peptidoglycan synthesis (lipid II) as well as of teichoic acid synthesis (lipid III). Noteworthy, it seems difficult for unrelated strains to gain resistance to teixobactin. In vitro, no teixobactin-resistant mutants of *S. aureus* or *M. tuberculosis* were isolated at four times the MIC, which may be attributed to the dual mode of action of teixobactin by targeting more than one essential bacterial macromolecule. However, bacteria have eventually always found ways to adapt to antibiotic action, and it may be just a matter of time that a resistance mechanism to teixobactin will be identified.

β -lactams are probably the most frequently used antibiotics to date and have a successful history in curing patients from infectious diseases, which is also due to their relatively small size as well as their good tolerability by the patients. However, the effectiveness of β -lactam antibiotics is severely hampered by the action of β -lactamases, which break down nearly every β -lactam by deacylation. Currently, more than thousand different β -lactamases from various structural classes and a wide range of substrate promiscuities and catalytic efficiencies are known, constantly evolving and disseminating with new β -lactam antibiotics that are introduced into clinical use. Inhibition of β -lactamases by using β -lactamase inhibitors is an effective and often practiced means to recover activity of β -lactam antibiotics. However, one disadvantage of β -lactamase inhibitors, which are also compounds with a β -lactam ring structure, is that they are also consumed by the β -lactamases, although at a much slower rate. Avibactam is a novel non- β -lactam β -lactamase inhibitor in clinical development combined with β -lactam antibiotic partners to treat infections with Gram-negative bacteria [69–71]. In contrast to other known β -lactamase inhibitors, avibactam covalently and slowly reversibly binds to various types of β -lactamases including TEM-1, i.e., deacylation of avibactam proceeds through regeneration of intact avibactam and not hydrolysis, which is a new and unique mechanism of inhibition among β -lactamase inhibitors.

Ribocil interferes with bacterial noncoding RNA (ncRNA), a new target molecule that is currently not used by any other clinically used antibiotic [72]. The researchers identified ribocil during

a phenotypical screen for inhibitors of a metabolic pathway leading to the synthesis of riboflavin, also called vitamin B2, which is a crucial precursor of essential cofactors required for various enzyme reactions. One such cofactor is flavin mononucleotide (FMN) that functions as a prosthetic group of several oxidoreductases including NADH dehydrogenase. Inside the human host, riboflavin is a rather rare metabolite that has to be produced by the bacteria to ensure their growth and vitality, rendering this pathway essential under such conditions. Ribocil resistant mutants carry mutations in a noncoding DNA region of the bacterial genome, indicating that ribocil rather acts on the level of gene regulation than direct interaction with a riboflavin biosynthesis enzyme. The involved ncRNA domain is located upstream of the translational start site of a key synthase enzyme in the riboflavin biosynthesis pathway, and constitutes a so-called riboswitch. Riboswitches are RNA regions that can change their structure upon binding of a corresponding ligand (here FMN ligand) in order to modulate the access of the transcription and translation machinery to the gene locus and thus prevent expression of this gene. This mechanism allows the bacteria to shut down the riboflavin biosynthesis pathway when sufficient riboflavin is available. Like FMN ligands, ribocil also binds to this riboswitch and shuts down riboflavin synthesis, thereby killing the bacteria by depriving them of the essential precursor metabolite. Noteworthy, ribocil is not a close structural analog of a metabolite ligand, reducing the possibility of off-target effects on other pathways that involve riboflavin and FMN in the human host, which is underlined by the observation that even high doses of the compound were not toxic in mice.

PC190723, a benzamide derivative, is a potent and selective inhibitor of the essential cell division pacemaker protein FtsZ. PC190723 shows specific antibacterial activity against staphylococci including MRSA with minimal inhibitory concentrations (MICs) in the range of 1.4–2.8 μM . Further, it is the first FtsZ inhibitor with reported in vivo efficacy as it is effective in a murine septicemia model of staphylococcal infection [73, 74]. Bacterial cell division is achieved by the divisome, a multi-protein complex that is characterized by the time-dependent assembly of specific cell division proteins [75]. At the onset of cell division, the tubulin homolog FtsZ localizes at mid cell to form the so-called FtsZ-ring or Z-ring in a GTP-dependent manner. The Z-ring functions as a scaffold for the assembly of the bacterial cytokinetic machinery. PC190723-treated rods like *Bacillus subtilis* show an elongated phenotype while staphylococci show enlarged spherical cells. Localization of FtsZ revealed the formation of multiple rings and arcs in *S. aureus* and abnormal discrete foci throughout *B. subtilis* cells, indicating an interference of PC190723 with Z-ring formation [73, 76, 77]. There are discussions about the effect of PC190723 on the GTPase activity of FtsZ. While some studies

could show a concentration-dependent inhibition of the GTPase activity of *S. aureus* and *B. subtilis* FtsZ [73, 77, 78], some more recent studies did not come to the same results but observed an increased GTPase activity of *S. aureus* FtsZ or no effect on *B. subtilis* FtsZ [79, 80]. The binding site of PC190723 maps to a cleft formed by the H7 helix, the T7-loop, and the C-terminal four-stranded β -sheet of *S. aureus* FtsZ [76, 79, 81]. Thus, the binding site is rather away from the GTP binding pocket, indicating that there seems to be at least no direct interference of PC190723 with the catalytic site of the GTPase domain. PC190723 further shows synergy with β -lactam antibiotics to kill MRSA [76]. Importantly, PC190723 resensitized MRSA to β -lactam antibiotics in vitro as well as in a mouse model of MRSA infection. This synergy is most probably achieved by the concomitant delocalization of their respective drug targets FtsZ and PBP2, since PBP2 depends on FtsZ for correct localization at the septum, where it is needed for transglycosylation of peptidoglycan in MRSA. Besides the synergistic effects, combination of imipenem with PC190723 significantly reduced the spontaneous frequency of PC190723-resistant mutants, which also showed an attenuated virulence. Thus, PC190723 represents an interesting new antibiotic that modulates the assembly/disassembly dynamics of FtsZ with promising antibacterial activity against an important human pathogen.

ADEPs belong to a new class of antibiotic acyldepsipeptides that exert prominent antibacterial activity against Gram-positive bacteria including MRSA in vitro and in vivo [82]. ADEP1, a natural product of *Streptomyces hawaiiensis* NRRL 15010, was first described in the 1980s [83]. Later, several new synthetic derivatives of ADEP1 with improved chemical and metabolic stability were obtained when researchers established a route for total ADEP synthesis and initiated a chemistry program. One of these derivatives, ADEP4, showed impressive MICs in the sub- μ g/ml range against MRSA. ADEPs demonstrate an unprecedented mode of action by targeting ClpP, the proteolytic core unit of the major bacterial protease complex (Fig. 4) [82, 84–86]. Clp proteases are important for protein turnover and homeostasis in bacteria in order to maintain vital cellular functions particularly under stress conditions. Apart from their crucial role in general protein quality control by degrading abnormally folded or otherwise aberrant or malfunctioning proteins, their temporally and spatially precise proteolysis of key regulatory proteins additionally directs developmental processes like cell motility, genetic competence, cell differentiation, sporulation, as well as important aspects of virulence. Due to their apparent relevance for many physiological processes and their conservation among diverse bacterial species including human pathogens, Clp protease emerged as a new target for antibiotic action and virulence inhibition [87, 88]. Usually, ClpP is tightly regulated by Clp-ATPases and is unable to degrade

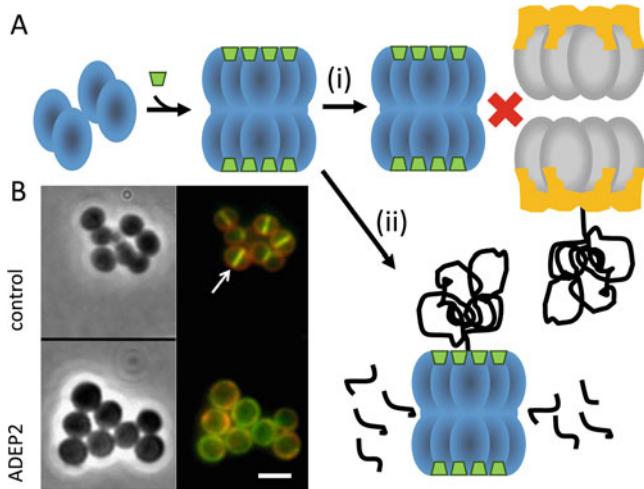


Fig. 4 ADEPs deregulate the proteolytic activity of ClpP. **(a)** Model on the ADEP mechanism of action. ADEPs perturb the activity of ClpP in a multilayered fashion: ADEPs (green) induce the oligomerization of ClpP monomers (blue) into the tetradecameric complex. Here, ADEPs share the same binding sites on ClpP with Clp-ATPases (grey), thereby abrogating the interaction of ClpP with corresponding Clp-ATPase, which leads to the inhibition of the natural functions of Clp in protein turnover (i). ADEPs bind to the outer rim of the apical and distal surfaces of ClpP in a 1:1 stoichiometry. Upon binding, ADEP induces a conformational shift in the N-terminal region of ClpP that results in the enlargement of the entrance pore to the proteolytic chamber of ClpP. Now, also nonnative protein substrates gain access to the proteolytic chamber of ClpP, releasing the degradative capacity of ClpP from strict regulation by Clp-ATPases, which leads to an untimely degradation of specific proteins or nascent polypeptides at the ribosome (ii). **(b)** ADEP-treatment of *S. aureus* leads to cell division inhibition and finally bacterial death. Fluorescence images show the bacterial membrane (red) and the divisome protein PBP2 (green). PBP2 usually localizes at mid cell of dividing bacteria (*upper panel*, arrow). Upon ADEP treatment, PBP2 delocalizes from the division site (*lower panel*), which is representative for the delocalization of several important cell division proteins under these conditions. This delocalization is a result of the degradation of the essential FtsZ protein by ADEP-activated ClpP. Scale bar, 2.5 μ m

proteins on its own. Biochemical studies demonstrated that ADEPs induce ClpP oligomerization and activate ClpP to recognize and degrade unfolded polypeptides as well as flexible proteins independently [84]. In addition, ADEPs abrogate the interaction of ClpP with cooperating Clp-ATPases, thereby preventing all natural functions of ClpP in general and regulatory proteolysis. Crystal structures and EM images of ClpP in its free form and in complex with ADEPs provided a rational for these biochemical observations. ADEPs compete with Clp-ATPases for the same binding site and finally trigger a closed- to open-gate structural transition of the ClpP N-terminal segments that opens the substrate entrance pore of ClpP, which is otherwise tightly closed [85, 89].

Using high-resolution microscopy, a significant swelling of coccoid *S. aureus* and *S. pneumoniae* cells as well as an impressive filamentation of rod-shaped *B. subtilis* cells in the presence of low inhibitory ADEP concentrations could be observed, clearly indicating stalled bacterial cell division [90]. Following the localization of fluorescently labeled cell division proteins using fluorescence microscopy revealed a mislocalization of essential members of the divisome including FtsZ. Immunodetection of FtsZ in ADEP treated cells showed a significant reduction of the concentration of FtsZ protein in a time-dependent manner, and in addition, ADEP-activated ClpP rapidly degraded purified FtsZ protein in vitro [90]. Thus, ADEPs prevent bacterial cell division by a different, yet unprecedented mechanism that is by activating a bacterial enzyme rather than inhibition of an enzymatic reaction, which destines the bacteria to death in a suicidal manner. Noteworthy, very recently, ADEPs were shown to kill mycobacteria by inhibiting the natural functions of the Clp system instead of over-activating ClpP [91], showing that depending on the microorganism, ADEPs make use of different killing mechanisms.

Antibiotics play an important role in our everyday life, since they are essential weapons in our fight against bacterial infections and additionally provide powerful tools as preservatives in the food industry. The examples in this chapter show how inventive nature is in establishing new antibiotic mechanisms of action and that there may be still more, yet unknown ways to interfere with the bacterial lifestyle. By studying such modes of action along with the coevolving resistance mechanisms, we will gain deeper insights into the bacterial way of life, which is an essential step toward the goal of developing new strategies to treat life-threatening bacterial infections while minimizing their impact on human health.

Acknowledgment

The author appreciates financial support from the German Research Foundation (DFG; SFB 766). I gratefully thank Anne Berscheid, Heike Brötz-Oesterhelt (University of Tuebingen), Gabriele Bierbaum, and Hans-Georg Sahl (University of Bonn) for their continuous support and helpful discussions.

References

1. WHO (2015). <http://www.who.int/mediacentre/factsheets/fs194/en/>. Accessed Apr 2016
2. CDC (2011) Antimicrobial resistance posing growing health threat. Centers for Disease Control and Prevention. http://www.cdc.gov/media/releases/2011/p0407_antimicrobialresistance.html. Accessed Apr 2016
3. AMR-review (2016) 22nd March 2016—Infection prevention control and surveillance: limiting the development and spread of drug resistance. <http://amr-review.org/>. Accessed Apr 2016
4. FOR854 (2016) <https://www.uni-bonn.de/forschung/startseite-forschung/dfg-geförderte-projekte/for854>. Accessed Apr 2016

5. DZIF (2016) http://www.dzif.de/en/research/novel_antifungals/. Accessed Apr 2016
6. GRC (2016) Gordon research conference: new antibacterial discovery & development. <https://www.grc.org/conferences.aspx?id=0000540>. Accessed Apr 2016
7. Russell AD (2003) Similarities and differences in the responses of microorganisms to biocides. *J Antimicrob Chemother* 52(5):750–763
8. Aminov RI (2009) The role of antibiotics and antibiotic resistance in nature. *Environ Microbiol* 11(12):2970–2988
9. Yang D, Biragyn A, Kwak LW, Oppenheim JJ (2002) Mammalian defensins in immunity: more than just microbicidal. *Trends Immunol* 23(6):291–296
10. Yang D, Biragyn A, Hoover DM, Lubkowski J, Oppenheim JJ (2004) Multiple roles of antimicrobial defensins, cathelicidins, and eosinophil-derived neurotoxin in host defense. *Annu Rev Immunol* 22:181–215
11. Bowdish DM, Davidson DJ, Hancock RE (2005) A re-evaluation of the role of host defence peptides in mammalian immunity. *Curr Protein Pept Sci* 6(1):35–51
12. Kodani S, Hudson ME, Durrant MC, Buttner MJ, Nodwell JR, Willey JM (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* 101(31):11448–11453
13. Kleerebezem M (2004) Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. *Peptides* 25(9):1405–1414
14. Schmitz S, Hoffmann A, Szekat C, Rudd B, Bierbaum G (2006) The lantibiotic mersacidin is an autoinducing peptide. *Appl Environ Microbiol* 72(11):7270–7277
15. Navarre WW, Schneewind O (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* 63(1):174–229
16. van Heijenoort J (2001) Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology* 11(3):25–36
17. Pages JM, James CE, Winterhalter M (2008) The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev Microbiol* 6(12):893–903
18. Coyette J, van der Ende A (2008) Peptidoglycan: the bacterial Achilles heel. *FEMS Microbiol Rev* 32(2):147–148
19. Schneider T, Sahl HG (2010) An oldie but a goodie—cell wall biosynthesis as antibiotic target pathway. *Int J Med Microbiol* 300(2-3):161–169
20. Yocum RR, Rasmussen JR, Strominger JL (1980) The mechanism of action of penicillin. Penicillin acylates the active site of *Bacillus stearothermophilus* D-alanine carboxypeptidase. *J Biol Chem* 255(9):3977–3986
21. Yocum RR, Waxman DJ, Rasmussen JR, Strominger JL (1979) Mechanism of penicillin action: penicillin and substrate bind covalently to the same active site serine in two bacterial D-alanine carboxypeptidases. *Proc Natl Acad Sci U S A* 76(6):2730–2734
22. Neuhaus FC, Lynch JL (1964) The enzymatic synthesis of D-alanyl-D-alanine. 3. On the inhibition of D-alanyl-D-alanine synthetase by the antibiotic D-cycloserine. *Biochemistry* 3:471–480
23. Lambert MP, Neuhaus FC (1972) Mechanism of D-cycloserine action: alanine racemase from *Escherichia coli* W. *J Bacteriol* 110(3):978–987
24. Kahan FM, Kahan JS, Cassidy PJ, Kropp H (1974) The mechanism of action of fosfomycin (phosphonomycin). *Ann N Y Acad Sci* 235:364–386
25. Bouhss A, Crouvoisier M, Blanot D, Mengin-Lecreux D (2004) Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis. *J Biol Chem* 279(29):29974–29980
26. Stone KJ, Strominger JL (1971) Mechanism of action of bacitracin: complexation with metal ion and C 55-isoprenyl pyrophosphate. *Proc Natl Acad Sci U S A* 68(12):3223–3227
27. Storm DR, Strominger JL (1973) Complex formation between bacitracin peptides and isoprenyl pyrophosphates. The specificity of lipid-peptide interactions. *J Biol Chem* 248(11):3940–3945
28. Schneider T, Gries K, Josten M, Wiedemann I, Pelzer S, Labischinski H, Sahl HG (2009) The lipopeptide antibiotic Fruilimicin B inhibits cell wall biosynthesis through complex formation with bactoprenol phosphate. *Antimicrob Agents Chemother* 53(4):1610–1618
29. Hiramatsu K (2001) Vancomycin-resistant *Staphylococcus aureus*: a new model of antibiotic resistance. *Lancet Infect Dis* 1(3):147–155
30. Cudic P, Kranz JK, Behenna DC, Kruger RG, Tadesse H, Wand AJ, Veklich YI, Weisel JW, McCafferty DG (2002) Complexation of peptidoglycan intermediates by the lipoglycopeptide antibiotic ramoplanin: minimal structural requirements for intermolecular complexation and fibril formation. *Proc Natl Acad Sci U S A* 99(11):7384–7389
31. Bierbaum G, Sahl HG (2009) Lantibiotics: mode of action, biosynthesis and bioengineering. *Curr Pharm Biotech* 10(1):2–18

32. Sahl HG, Bierbaum G (1998) Lantibiotics: biosynthesis and biological activities of uniquely modified peptides from gram-positive bacteria. *Annu Rev Microbiol* 52:41–79
33. Brötz H, Bierbaum G, Leopold K, Reynolds PE, Sahl HG (1998) The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II. *Antimicrob Agents Chemother* 42(1):154–160
34. Galvez A, Abriouel H, Lopez RL, Ben Omar N (2007) Bacteriocin-based strategies for food biopreservation. *Int J Food Microbiol* 120(1-2):51–70
35. Brötz H, Josten M, Wiedemann I, Schneider U, Gotz F, Bierbaum G, Sahl HG (1998) Role of lipid-bound peptidoglycan precursors in the formation of pores by nisin, epidermin and other lantibiotics. *Mol Microbiol* 30(2):317–327
36. Wiedemann I, Breukink E, van Kraaij C, Kuipers OP, Bierbaum G, de Kruijff B, Sahl HG (2001) Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *J Biol Chem* 276(3):1772–1779
37. Peschel A, Sahl HG (2006) The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat Rev Microbiol* 4(7):529–536
38. Sass V, Pag U, Tossi A, Bierbaum G, Sahl HG (2008) Mode of action of human beta-defensin 3 (hBD3) against *Staphylococcus aureus* and transcriptional analysis of responses to defensin challenge. *Int J Med Microbiol* 298:619–633
39. Muthaiyan A, Silverman JA, Jayaswal RK, Wilkinson BJ (2008) Transcriptional profiling reveals that daptomycin induces the *Staphylococcus aureus* cell wall stress stimulon and genes responsive to membrane depolarization. *Antimicrob Agents Chemother* 52:980–990
40. Wecke T, Zuhlke D, Mader U, Jordan S, Voigt B, Pelzer S, Labischinski H, Homuth G, Hecker M, Mascher T (2009) Daptomycin versus Friulimicin B: in-depth profiling of *Bacillus subtilis* cell envelope stress responses. *Antimicrob Agents Chemother* 53(4):1619–1623
41. Camargo IL, Neoh HM, Cui L, Hiramatsu K (2008) Serial daptomycin selection generates daptomycin-nonsusceptible *Staphylococcus aureus* strains with a heterogeneous vancomycin-intermediate phenotype. *Antimicrob Agents Chemother* 52(12):4289–4299
42. Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K (2003) Two-component system VraSR positively modulates the regulation of cell-wall biosynthesis path-way in *Staphylococcus aureus*. *Mol Microbiol* 49(3):807–821
43. Utaida S, Dunman PM, Macapagal D, Murphy E, Projan SJ, Singh VK, Jayaswal RK, Wilkinson BJ (2003) Genome-wide transcriptional profiling of the response of *Staphylococcus aureus* to cell-wall-active antibiotics reveals a cell-wall-stress stimulon. *Microbiology* 149(10):2719–2732
44. Sass P, Jansen A, Szekat C, Sass V, Sahl HG, Bierbaum G (2008) The lantibiotic mersacidin is a strong inducer of the cell wall stress response of *Staphylococcus aureus*. *BMC Microbiol* 8:186
45. Cavalleri B, Turconi M, Tamborini G, Occelli E, Cietto G, Pallanza R, Scotti R, Berti M, Romano G, Parenti F (1990) Synthesis and biological activity of some derivatives of rifamycin P. *J Med Chem* 33(5):1470–1476
46. Yoshizawa S, Fourmy D, Puglisi JD (1998) Structural origins of gentamicin antibiotic action. *EMBO J* 17(22):6437–6448
47. Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407(6802):340–348
48. Pioletti M, Schlünzen F, Harms J, Zarivach R, Gluhmann M, Avila H, Bashan A, Bartels H, Auerbach T, Jacobi C, Hartsch T, Yonath A, Franceschi F (2001) Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *EMBO J* 20(8):1829–1839
49. Schlünzen F, Zarivach R, Harms J, Bashan A, Tocilj A, Albrecht R, Yonath A, Franceschi F (2001) Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* 413(6858):814–821
50. Long KS, Porse BT (2003) A conserved chloramphenicol binding site at the entrance to the ribosomal peptide exit tunnel. *Nucleic Acids Res* 31(24):7208–7215
51. Moazed D, Noller HF (1987) Chloramphenicol, erythromycin, carbomycin and vernamycin B protect overlapping sites in the peptidyl transferase region of 23S ribosomal RNA. *Biochimie* 69(8):879–884
52. Cassels R, Oliva B, Knowles D (1995) Occurrence of the regulatory nucleotides ppGpp and pppGpp following induction of the stringent response in staphylococci. *J Bacteriol* 177(17):5161–5165
53. Crosse AM, Greenway DL, England RR (2000) Accumulation of ppGpp and ppGp in *Staphylococcus aureus* 8325-4 following nutrient starvation. *Lett Appl Microbiol* 31(4):332–337

54. Abranches J, Martinez AR, Kajfasz JK, Chavez V, Garsin DA, Lemos JA (2009) The molecular alarmone (p)ppGpp mediates stress responses, vancomycin tolerance, and virulence in *Enterococcus faecalis*. *J Bacteriol* 191(7):2248–2256
55. Reiss S, Pane-Farre J, Fuchs S, Francois P, Liebeke M, Schrenzel J, Lindequist U, Lalk M, Wolz C, Hecker M, Engelmann S (2012) Global analysis of the *Staphylococcus aureus* response to mupirocin. *Antimicrob Agents Chemother* 56(2):787–804
56. Otaka T, Kaji A (1973) Evidence that fusidic acid inhibits the binding of aminoacyl-tRNA to the donor as well as the acceptor site of the ribosomes. *Eur J Biochem* 38(1):46–53
57. Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC, Ramakrishnan V (2009) The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* 326(5953):694–699
58. Drlica K, Zhao X (1997) DNA gyrase, topoisomerase IV, and the 4-quinolones. *Microbiol Mol Biol Rev* 61(3):377–392
59. Reece RJ, Maxwell A (1991) DNA gyrase: structure and function. *Crit Rev Biochem Mol Biol* 26(3–4):335–375
60. Roca J (1995) The mechanisms of DNA topoisomerases. *Trends Biochem Sci* 20(4):156–160
61. Peng H, Marians KJ (1995) The interaction of *Escherichia coli* topoisomerase IV with DNA. *J Biol Chem* 270(42):25286–25290
62. Danshiitsoodol N, de Pinho CA, Matoba Y, Kumagai T, Sugiyama M (2006) The mitomycin C (MMC)-binding protein from MMC-producing microorganisms protects from the lethal effect of bleomycin: crystallographic analysis to elucidate the binding mode of the antibiotic to the protein. *J Mol Biol* 360(2):398–408
63. Claverys JP, Prudhomme M, Martin B (2006) Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu Rev Microbiol* 60:451–475
64. Au N, Kuester-Schoeck E, Mandava V, Bothwell LE, Canny SP, Chachu K, Colavito SA, Fuller SN, Groban ES, Hensley LA, O'Brien TC, Shah A, Tierney JT, Tomm LL, O'Gara TM, Goranov AI, Grossman AD, Lovett CM (2005) Genetic composition of the *Bacillus subtilis* SOS system. *J Bacteriol* 187(22):7655–7666
65. Friedman N, Vardi S, Ronen M, Alon U, Stavans J (2005) Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biol* 3(7), e238
66. Kelley WL (2006) Lex marks the spot: the virulent side of SOS and a closer look at the LexA regulon. *Mol Microbiol* 62(5):1228–1238
67. Michel B (2005) After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol* 3(7), e255
68. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schäberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K (2015) A new antibiotic kills pathogens without detectable resistance. *Nature* 517(7535):455–459
69. Ehmann DE, Jahic H, Ross PL, Gu RF, Hu J, Kern G, Walkup GK, Fisher SL (2012) Avibactam is a covalent, reversible, non-beta-lactam beta-lactamase inhibitor. *Proc Natl Acad Sci U S A* 109(29):11663–11668
70. Ehmann DE, Jahic H, Ross PL, Gu RF, Hu J, Durand-Reville TF, Lahiri S, Thresher J, Livchak S, Gao N, Palmer T, Walkup GK, Fisher SL (2013) Kinetics of avibactam inhibition against Class A, C, and D beta-lactamases. *J Biol Chem* 288(39):27960–27971
71. Lahiri SD, Mangani S, Jahic H, Benvenuti M, Durand-Reville TF, De Luca F, Ehmann DE, Rossolini GM, Alm RA, Docquier JD (2015) Molecular basis of selective inhibition and slow reversibility of avibactam against class D carbapenemases: a structure-guided study of OXA-24 and OXA-48. *ACS Chem Biol* 10(2):591–600
72. Howe JA, Wang H, Fischmann TO, Balibar CJ, Xiao L, Galgoci AM, Malinvern JC, Mayhood T, Villafania A, Nahvi A, Murgolo N, Barbieri CM, Mann PA, Carr D, Xia E, Zuck P, Riley D, Painter RE, Walker SS, Sherborne B, de Jesus R, Pan W, Plotkin MA, Wu J, Rindgen D, Cummings J, Garlis CG, Zhang R, Sheth PR, Gill CJ, Tang H, Roemer T (2015) Selective small-molecule inhibition of an RNA structural element. *Nature* 526(7575):672–677
73. Haydon DJ, Stokes NR, Ure R, Galbraith G, Bennett JM, Brown DR, Baker PJ, Barynin VV, Rice DW, Sedelnikova SE, Heal JR, Sheridan JM, Aiwale ST, Chauhan PK, Srivastava A, Taneja A, Collins I, Errington J, Czaplewski LG (2008) An inhibitor of FtsZ with potent and selective anti-staphylococcal activity. *Science* 321(5896):1673–1675
74. Haydon DJ, Bennett JM, Brown D, Collins I, Galbraith G, Lancett P, Macdonald R, Stokes NR, Chauhan PK, Sutariya JK, Nayal N, Srivastava A, Beanland J, Hall R, Henstock V, Noula C, Rockley C, Czaplewski L (2010) Creating an antibacterial with in vivo efficacy: synthesis and characterization of potent inhibi-

- tors of the bacterial cell division protein FtsZ with improved pharmaceutical properties. *J Med Chem* 53(10):3927–3936
75. Adams DW, Errington J (2009) Bacterial cell division: assembly, maintenance and disassembly of the Z ring. *Nat Rev Microbiol* 7(9):642–653
 76. Tan CM, Therien AG, Lu J, Lee SH, Caron A, Gill CJ, Lebeau-Jacob C, Benton-Perdomo L, Monteiro JM, Pereira PM, Elsen NL, Wu J, Deschamps K, Petcu M, Wong S, Daigneault E, Kramer S, Liang L, Maxwell E, Claveau D, Vaillancourt J, Skorey K, Tam J, Wang H, Meredith TC, Sillaots S, Wang-Jarantow L, Ramtohul Y, Langlois E, Landry F, Reid JC, Parthasarathy G, Sharma S, Baryshnikova A, Lumb KJ, Pinho MG, Soisson SM, Roemer T (2012) Restoring methicillin-resistant *Staphylococcus aureus* susceptibility to beta-lactam antibiotics. *Sci Transl Med* 4(126):126ra135
 77. Adams DW, Wu LJ, Czaplewski LG, Errington J (2011) Multiple effects of benzamide antibiotics on FtsZ function. *Mol Microbiol* 80(1):68–84
 78. Andreu JM, Schaffner-Barbero C, Huecas S, Alonso D, Lopez-Rodriguez ML, Ruiz-Avila LB, Nunez-Ramirez R, Llorca O, Martin-Galiano AJ (2010) The antibacterial cell division inhibitor PC190723 is an FtsZ polymer-stabilizing agent that induces filament assembly and condensation. *J Biol Chem* 285(19):14239–14246
 79. Elsen NL, Lu J, Parthasarathy G, Reid JC, Sharma S, Soisson SM, Lumb KJ (2012) Mechanism of action of the cell-division inhibitor PC190723: modulation of FtsZ assembly cooperativity. *J Am Chem Soc* 134(30):12342–12345
 80. Anderson DE, Kim MB, Moore JT, O'Brien TE, Sorto NA, Grove CI, Lackner LL, Ames JB, Shaw JT (2012) Comparison of small molecule inhibitors of the bacterial cell division protein FtsZ and identification of a reliable cross-species inhibitor. *ACS Chem Biol* 7(11):1918–1928
 81. Matsui T, Yamane J, Mogi N, Yamaguchi H, Takemoto H, Yao M, Tanaka I (2012) Structural reorganization of the bacterial cell-division protein FtsZ from *Staphylococcus aureus*. *Acta Crystallogr D Biol Crystallogr* 68(Pt 9):1175–1188
 82. Brötz-Oesterhelt H, Beyer D, Kroll HP, Endermann R, Ladel C, Schroeder W, Hinzen B, Raddatz S, Paulsen H, Henninger K, Bandow JE, Sahl HG, Labischinski H (2005) Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* 11(10):1082–1087
 83. Michel KH, Kastner RE (1985) A54556 antibiotics and process for production thereof. US Patent 4,492,650
 84. Kirstein J, Hoffmann A, Lilie H, Schmidt R, Rübsamen-Waigmann H, Brötz-Oesterhelt H, Mogk A, Turgay K (2009) The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Mol Med* 1(1):37–49
 85. Lee BG, Park EY, Jeon H, Sung KH, Paulsen H, Rübsamen-Schaeff H, Brötz-Oesterhelt H, Song HK (2010) Structures of ClpP in complex with a novel class of antibiotics reveal its activation mechanism. *Nat Struct Mol Biol* 17(4):471–478
 86. Gersch M, Famulla K, Dahmen M, Gobl C, Malik I, Richter K, Korotkov VS, Sass P, Rubsam-Schaeff H, Madl T, Brötz-Oesterhelt H, Sieber SA (2015) AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nat Commun* 6:6320
 87. Sass P, Brötz-Oesterhelt H (2013) Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int J Med Microbiol* 304(1):23–30
 88. Baker TA, Sauer RT (2012) ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta* 1823(1):15–28
 89. Li DH, Chung YS, Gloyd M, Joseph E, Ghirlando R, Wright GD, Cheng YQ, Maurizi MR, Guarne A, Ortega J (2010) Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: a model for the ClpX/ClpA-bound state of ClpP. *Chem Biol* 17(9):959–969
 90. Sass P, Josten M, Famulla K, Schiffer G, Sahl HG, Hamoen L, Brötz-Oesterhelt H (2011) Antibiotic acyldepsipeptides activate ClpP peptidase to degrade the cell division protein FtsZ. *Proc Natl Acad Sci U S A* 108(42):17474–17479
 91. Famulla K, Sass P, Malik I, Akopian T, Kandror O, Alber M, Hinzen B, Ruebsamen-Schaeff H, Kalscheuer R, Goldberg AL, Brötz-Oesterhelt H (2016) Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Mol Microbiol* 101:194–209

Chapter 2

Mining Bacterial Genomes for Secondary Metabolite Gene Clusters

Martina Adamek, Marius Spohn, Evi Stegmann, and Nadine Ziemert

Abstract

With the emergence of bacterial resistance against frequently used antibiotics, novel antibacterial compounds are urgently needed. Traditional bioactivity-guided drug discovery strategies involve laborious screening efforts and display high rediscovery rates. With the progress in next generation sequencing methods and the knowledge that the majority of antibiotics in clinical use are produced as secondary metabolites by bacteria, mining bacterial genomes for secondary metabolites with antimicrobial activity is a promising approach, which can guide a more time and cost-effective identification of novel compounds. However, what sounds easy to accomplish, comes with several challenges. To date, several tools for the prediction of secondary metabolite gene clusters are available, some of which are based on the detection of signature genes, while others are searching for specific patterns in gene content or regulation.

Apart from the mere identification of gene clusters, several other factors such as determining cluster boundaries and assessing the novelty of the detected cluster are important. For this purpose, comparison of the predicted secondary metabolite genes with different cluster and compound databases is necessary. Furthermore, it is advisable to classify detected clusters into gene cluster families. So far, there is no standardized procedure for genome mining; however, different approaches to overcome all of these challenges exist and are addressed in this chapter. We give practical guidance on the workflow for secondary metabolite gene cluster identification, which includes the determination of gene cluster boundaries, addresses problems occurring with the use of draft genomes, and gives an outlook on the different methods for gene cluster classification. Based on comprehensible examples a protocol is set, which should enable the readers to mine their own genome data for interesting secondary metabolites.

Key words Genome mining, Secondary metabolite gene cluster, Antibiotics, Biosynthesis, Cluster boundaries, Prioritization, Gene cluster families, INBEKT, antiSMASH

1 Introduction

Antibiotics, immunosuppressive agents, cancer medication, the list of bacterial secondary metabolites with an immense value for healthcare is long [1, 2]. Thousands of compounds have been isolated from microbes so far, showing an immense variety in structure and bioactivities. However, with the emergence of bacterial resistances against frequently used antibiotics, the need for finding

novel bioactive compounds is urgent [3]. Traditional bioactivity guided screening methods tend to rediscover already known molecules. The increasing availability of bacterial genome sequences and the continuously improving algorithms for the computational prediction of bacterial secondary metabolites prepare the ground for the so-called genome mining [4], which aims at the identification of Secondary Metabolite Gene Clusters (SMGC) within genomic data. Understanding the composition and regulation of SMGCs can guide experiments for a more targeted isolation of molecules, to reduce time and cost for the discovery of new compounds. Additionally, genetic information can help to activate “silent” gene clusters, which are not expressed under standard laboratory culture conditions. Furthermore, knowledge of regulatory mechanisms can help with the optimization of experimental conditions for heterologous gene expression. In addition to the identification of SMGCs the prediction of their products is a challenge.

To apprehend the basic mechanisms of secondary metabolite biosynthesis, it is necessary to know certain SMGC types, typical genes, functional domains, and assembly mechanisms. A short summary, introducing the most common SMGC types, is given in the following section. However, for a deeper understanding of the various mechanisms by which secondary metabolites are produced we recommend more detailed reviews [5–9].

A major part of bacterial secondary metabolites with antibiotic activity are synthesized by large modular enzymes: polyketide synthases (PKS) and nonribosomal peptide synthetases (NRPS) [10]. Variation in number and/or structure of the different PKS or NRPS modules leads to the diversification of the encoded secondary metabolites. In general, a single PKS module is composed of three core domains: (1) an acyl-transferase (AT), which activates and binds a specific substrate (CoA activated acyl group) and transfers it onto (2) an acyl carrier protein (ACP). (3) A ketosynthase (KS) catalyzes the condensation and decarboxylation between the acyl CoA substrate and the growing polyketide chain. Further processing, which leads to chemical diversification, occurs by tailoring domains, for example, ketoreductase (KR), dehydratase (DH), or enoyl reductase (ER) domains. A thioesterase (TE) domain usually terminates the assembly [11]. An alternative mechanism is the use of trans-AT domains. Here, AT domains are not located within each module but encoded on a freestanding protein elsewhere within the cluster [12].

Likewise, NRPSs are composed of several modules, each containing particular domains. Adenylation (A) domains specifically select and activate amino acids by adenylation; the activated amino acid is subsequently transferred onto a peptidyl carrier protein (PCP). Two amino acids loaded on neighboring PCPs are condensed by the catalytic activity of a condensation (C) domain.

As in the PKS, the presence of a TE domain usually terminates the chain elongation [13, 14]. Tailoring enzymes, as for example oxygenases, glycosyltransferases, methyltransferases, racemases, or cyclases, contribute to the versatility of secondary metabolites [8]. Other domains responsible for further structure modifications can be located within the PKS or NRPS modules or on separate enzymes in the cluster. A brief summary of specific domains and their function in a secondary metabolite gene cluster is given in Table 1.

Type-I-PKSs are composed of multiple modules located on one or more proteins, successively or iteratively assembling the final molecule. Type-II-PKS clusters are composed of the same core structure domains (KS, AT, ACP), but instead of a modular structure each gene encodes a monofunctional catalytic domain. This set of proteins is iteratively used [15]. Type-III-PKSs have an iterative mechanism too, differing from type-II-PKSs by the lack of substrate binding ACP domains for the assembly [16]. Furthermore, hybrid clusters including PKS and NRPS modules are common [17].

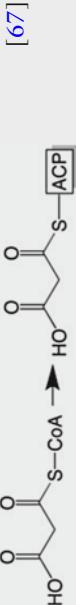
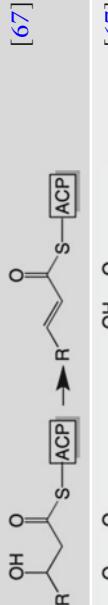
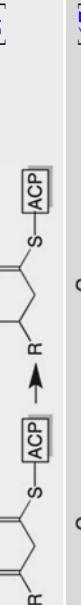
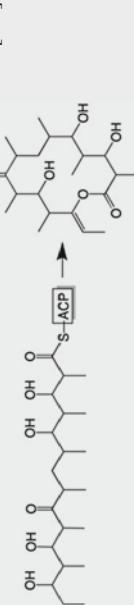
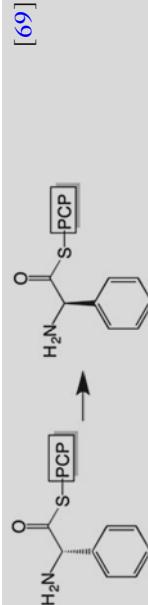
Another group of bacterial secondary metabolites that contain important antibiotic compounds are ribosomally synthesized and posttranslationally modified peptides (RiPPs). Although the final products show a high variety in length and structure, they all share a similar biosynthetic mechanism. They are synthesized as a precursor peptide comprised of a leader peptide and a core peptide section, and are further processed by modifying enzymes, such as methylases, dehydratases, or cyclases. Finally, the leader peptide is proteolytically removed, releasing the active core peptide. Important RiPP classes include lantipeptides, lassopeptides, linear azo(line) containing peptides (LAPs) and thiopeptides [5].

An additional important class of secondary metabolites are terpenes. Terpenes, or in a modified form often referred to as terpenoids, are built from isoprene units and are mainly found in plants and fungi; however, in recent years, many bacterial terpenes, including terpenes with antibiotic activity, have been discovered [18]. Terpene biosynthesis is initiated by the enzymatic ionization of one of the precursor molecules geranyl diphosphate, farnesyl diphosphate, or geranyl diphosphate, which is further processed by a cyclase. Tailoring enzymes further shape the tremendous variety of terpenoid structures. The genes encoding the precursor molecules can, but must not be necessarily, be part of the biosynthesis gene cluster [6].

Other compound classes to be mentioned within the context of potential antibiotics are: nucleosides, aminoglycosides, beta-lactams, and aminocoumarins [19–23].

Although we are not yet able to exactly predict all secondary metabolite clusters using genomic data alone, genome mining has become an important part of drug discovery. Additionally, with the computational tools available, researchers from other fields are able

Table 1
Secondary metabolites functional domains

Domain	Function	Example	References
AT	Acyltransferase	Selects and activates acyl CoA substrate by acylation, binding to an ACP	 [67]
ACP (T)	Acyl carrier protein (thiolation)	Binds the acyl CoA substrate	
KS	Ketosynthase	Catalyzes condensation	 [67]
DH	Dehydratase	Dehydrates β-hydroxy carbonyl to a C-C-double bond	 [67]
KR	Ketoreductase	Reduces the β-keto group to a hydroxyl group	 [67]
ER	Enoyl reductase	Reduces the double bond	 [67]
TE	Thioesterase	Releases the full length chain and catalyzes macro-cyclization	 [68]
E	Epimerization	Flips stereo-chemistry	 [69]

A	Adenylation	Selects and binds an amino acid to a PCP	 [7]
PCP (T)	Peptidyl carrier protein (thiolation)	Binds the amino acid	
C	Condensation	Peptide bond formation	 [7]
Cyc	Cyclase	Forms a peptide bond and catalyzes subsequent cyclization	 [7]
O/N/C MT	O/N/C Methyl-transferase	Adds a methyl group to oxygen, nitrogen, or carbon	 [7]
TD (R)	Terminal reductase (various reaction mechanisms)	Catalyzes the reductive release from PKS or NRPS assembly lines	 [70]
ECH	Enoyl CoA hydratase	Hydrates the double bond on acyl-CoA	 [71]
CAL	Coenzyme A ligase	Catalyzes the ligation of acyl groups with CoA	 [72]

(continued)

Table 1
(continued)

Domain	Function	Example	References
ACPS Acyl carrier protein synthase	Catalyzes the conversion from inactive apo-ACP to active holo-ACP	<p>Adenine-ACP + $\text{S}^{\text{-}}$ $\xrightarrow{\text{[73]}}$ Holo-ACP</p>	[73]
FkbH	Incorporates “unusual” extender units; transfers a glyceryl moiety to an ACP	<p>$\text{OH}-\text{CH}_2-\text{CH}_2-\text{OH} + \text{ACP-SH} \xrightarrow{\text{[74]}} \text{OH}-\text{CH}_2-\text{CH}(\text{ACP})-\text{OH}$</p>	[74]
X	Recruits a specific P450 oxygenase to the NRPS bound peptide for side chain cross linking in glycopeptides	<p>$\text{NRPS-Peptide} + \text{P450} \xrightarrow{\text{Oxidase}} \text{Cross-linked Peptide}$</p>	[75]

to mine their genomes for secondary metabolites and explore different aspects such as ecological function and evolution of these fascinating compounds [24].

In this book chapter, we would like to present known and validated computational tools for the discovery of secondary metabolites, but also introduce two approaches for the discovery of completely novel biosynthesis mechanisms. Finally, we address the topic of prioritization of gene clusters, and introduce three different concepts to define gene cluster families.

2 Material

In this book chapter, we provide a detailed description for some of the numerous computer programs developed for genome mining and introduce new strategies for the identification and prioritization of SMGCs (*see Note 1*). An overview of the workflow is given in Fig. 1. A list of tools and programs for the analysis of secondary metabolite gene clusters and related tasks described in this chapter is given in Table 2. For a more comprehensive list of programs and databases, please refer to some recent reviews [4, 25].

3 Methods

This section provides a guide for the computational analysis of bacterial secondary metabolite gene clusters, using the genome sequences of *Amycolatopsis orientalis* HCCB10007 and *Amycolatopsis azurea* DSM 43854 as examples. *Amycolatopsis* is an Actinomycete genus known for the production of secondary metabolites including glycopeptides and rifamycin, and therefore serves as a good example for the identification of SMGCs. *A. orientalis* HCCB10007 represents a complete genome sequence, while *A. azurea* DSM 43854 is only available as draft genome.

3.1 Obtaining the Desired Sequences

Sequence files can be downloaded for example from the National Center for Biotechnology Information (NCBI) database (<http://www.ncbi.nlm.nih.gov>) or the Joint Genome Institute (JGI) database (<http://jgi.doe.gov>) as GenBank, EMBL or FASTA files. If necessary, GenBank files can be easily converted into FASTA files using either a python script, or a web-based application (Table 2) (*see Notes 2 and 3*). If the desired sequence is not annotated, or new sequence data are used, annotation tools such as PROKKA [26] or GLIMMER3 [27] are an option. For those who are not familiar with the application of command-line tools genome annotation is offered, for example, on the GenDB annotation pipeline [28] or on the RAST webpage [29]. As an open source genome viewer Artemis (Table 2) is recommended [30].

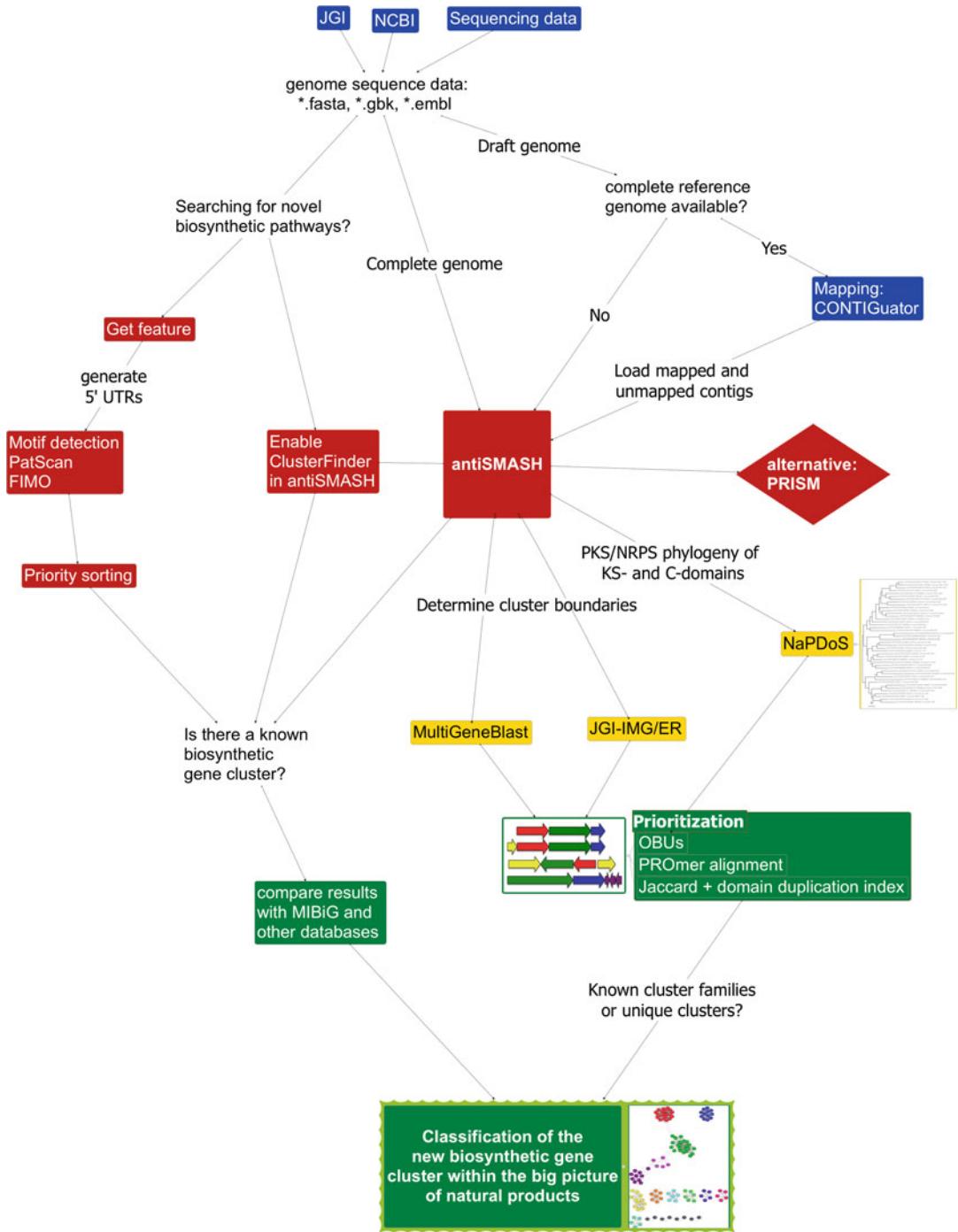


Fig. 1 Genome mining workflow overview get sequences (blue), identify SMGCs (red), refine SMGCs (yellow), compare SMGCs to known clusters, predict preliminary product (green)

Table 2
Open source programs, tools, and web applications

<i>GenBank to FASTA conversion</i>		
The Rocap lab	http://rocaplab.ocean.washington.edu/tools/genbank_to_fasta	
Molecular Organization and Assembly in Cells (Warwick)	http://www2.warwick.ac.uk/fac/sci/moac/people/students/peter_cock/python/genbank2fasta/	
<i>Secondary metabolite cluster prediction (known biosynthetic pathways)</i>		
antiSMASH	http://antismash.secondarymetabolites.org/	[33]
PRISM	http://magarveylab.ca/prism/#!/prism	[38]
<i>INBEKT progression (novel biosynthetic pathways)</i>		
GetFeature	http://webtools.secondarymetabolites.org/get_feature	
PRODORIC	http://www.prodoric.de/	[46]
CollectTF	http://www.collectf.org/home/greet/	[47]
FIMO	http://meme-suite.org/tools/fimo	[49]
PatScan	http://patscan.secondarymetabolites.org/	[50]
<i>Comparative genomics</i>		
MultiGeneBlast	http://multigeneblast.sourceforge.net/	[54]
JGI-IMG/ER	https://img.jgi.doe.gov/cgi-bin/mer/main.cgi	[53]
NaPDoS	http://napdos.ucsd.edu/	[58]
<i>Contig mapping</i>		
CONTIGuator	http://contiguator.sourceforge.net/	[57]
<i>Visualization, phylogeny, networks</i>		
Artemis (genome viewer)	http://www.sanger.ac.uk/science/tools/artemis	[30]
MEGA (phylogeny)	http://www.megasoftware.net/	[59]
Cytoscape (networks)	http://www.cytoscape.org/	[65]

The sequences used in this example setup, *Amycolatopsis orientalis* HCCB10007 (GCA_000400635.2) and *Amycolatopsis azurea* DSM 43854 (GCA_000340415.1), are available from the NCBI database and from the JGI; the GenBank files and raw nucleotide data in FASTA format can be downloaded.

3.2 Identification of Secondary Metabolite Gene Clusters

In recent years, the most commonly used platform for the detection and classification of secondary metabolite gene clusters is antiSMASH (antibiotics & Secondary Metabolite Analysis SHell) [31–33]. antiSMASH is available as a web-based application or as a standalone tool. This section focuses on the antiSMASH web version. antiSMASH uses pHMMs (profile Hidden Markov Models) of signature genes and pHMMs of PKS/NRPS signature domains for the identification of biosynthetic gene clusters. Based on the comparison of the enzymes within a cluster with an smCOG (secondary metabolite cluster of orthologous groups) database, further accessory genes are identified. Predicted clusters are subsequently automatically compared to known clusters with ClusterBlast. Furthermore, a core structure of the encoded compound is predicted based on the specificity of AT domains for type-I-PKS, A domains for NRPS, and core peptides of lantipeptides [31–33].

antiSMASH accepts GenBank, EMBL, or FASTA files as input data (*see Note 4*). For a first run, default settings are recommended. If the email address is provided, an email containing a link to a graphical display of the results is sent, after the cluster finding process is finished. In our example strain *A. orientalis* HCCB10007, 31 clusters were identified (Fig. 2a). In cases where gene clusters similar to known SMGCs were detected, the predicted cluster is linked to the MIBiG (Minimum Information about a Biosynthetic Gene cluster) database [34] of known and characterized SMGCs. Although a comparison to known clusters is often displayed, similarities can be low and therefore might not be reliable, so each cluster should be reviewed in detail (*see Note 5*). By selecting a cluster, a close-up on the specific region is displayed. Here, the biosynthetic genes, regulators, and transporters are highlighted in different color codes. Furthermore, for type-I-PKS and NRPS the different domains of a module are depicted (*see Note 6*). For a brief functional description of the different domains and some typical tailoring enzymes see Table 1. Cluster 3 of *A. orientalis* HCCB10007 was chosen as an example cluster depicted in Fig. 2b.

The “homologous gene cluster” section shows similar clusters from all NCBI database entries detected by antiSMASH (updated on a regular basis). These are not necessarily characterized clusters, but the analysis can be helpful to deduce if this type of cluster is common in any other bacterial species. The “homologous known gene cluster” section shows the most similar hits with the MIBiG database, containing only well-characterized gene clusters. The “homologous subcluster” section presents similar precursor biosynthesis subclusters. A rough prediction for the core structure of the final molecule is given for type-I-PKS, NRPS, and lantipeptides. This core structure includes the monomers predicted according to the specificity of AT or A domains but no cyclization or posttranslational modifications by tailoring enzymes. The structure prediction is based on the collinearity rule, which assumes that the

a)

Select Gene Cluster:
 Overview 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Identified secondary metabolite clusters						MiBiG BGC-ID
Cluster	Type	From	To	Most similar known cluster		
The following clusters are from record CP003410:						
Cluster 1	Lantipeptide	381927	404470	Erythreapeptin_biosynthetic_gene_cluster (75% of genes show similarity)	BGC0000513_c1	
Cluster 2	Arylpolyene	1323514	1364665	Kinamycin_biosynthetic_gene_cluster (5% of genes show similarity)	BGC0000236_c1	
Cluster 3	Nrps-T3pk	1520853	1614499	Vancomycin_biosynthetic_gene_cluster (97% of genes show similarity)	BGC0000455_c1	
Cluster 4	Terpene	1634429	1655385	Isorenieratene_biosynthetic_gene_cluster (57% of genes show similarity)	BGC0000664_c1	
Cluster 5	Bacteriocin-T1pk	1957861	2002559	Kedarcidin_biosynthetic_gene_cluster (12% of genes show similarity)	BGC0000081_c1	

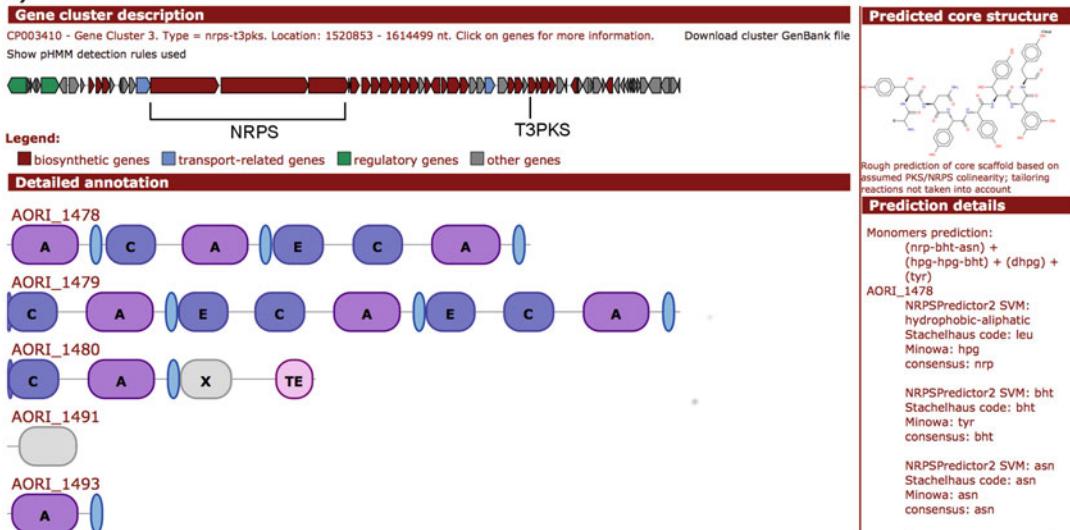
b)

Fig. 2 Example antiSMASH *.html output for *A. orientalis* HCCB10007. (a) Section from the gene cluster overview page. (b) Vancomycin biosynthesis gene cluster detected by antiSMASH. antiSMASH indicates glycopeptide biosynthesis clusters as NRPS-PKS-III hybrids. While the glycopeptide backbone is assembled via an NRPS mechanism, one of the incorporated amino acids (dihydroxyphenylglycine) is synthesized by the action of a type-III PKS

biosynthesis mechanism is linear to the order of the functional domains in a module. Since plenty of exceptions from this rule are known [35–37], the predicted core structure should be considered as only one possibility for the assembly.

3.2.1 Alternative Tools for Secondary Metabolite Gene Cluster Prediction

A recently published tool for the prediction of NRPS, type-I-, and type-II-PKS clusters is PRISM (Prediction Informatics for Secondary Metabolomes) [38]. Similar to antiSMASH, PRISM analyses open reading frames based on a large HMM library and groups them into clusters. The major difference compared to antiSMASH is that for the final structure prediction in PRISM the functions of trans-acting AT domains, deoxy sugar combinations,

tailoring, and cyclization reactions are taken into account. For the identification of known compounds, a multilocus sequence typing inspired approach, generating scaffold structures, is implemented and a database containing about 50,000 known secondary metabolites is available for comparison [38]. PRISM can be used as an alternative or in addition to antiSMASH for the prediction of PKS and NRPS clusters. It should be noted that in contrast to antiSMASH, the general cluster size determined using PRISM is usually underestimated and it is therefore recommended to take a look at the regions flanking the determined clusters. The most powerful implementation of the PRISM algorithm is provided in the GNP (Genomes to Natural Products) application, which offers the unique possibility of connecting genome sequence data with mass spectrometry data [39]. Other programs available, with more specific applications, are: NP.searcher [40], ClustScan [41] and SBSPKS [42] for PKS and NRPS clusters, or BAGEL [43] for bacteriocins.

3.2.2 Alternative Approaches for Discovery of Novel Secondary Metabolite Gene Clusters

The high confidential predictive tools are limited to the detection of already known, well-characterized gene cluster classes. Cluster Finder is an algorithm that aims to identify gene clusters of both known and unknown classes [44]. ClusterFinder is included as an optional plug-in in the most recent version of antiSMASH [33] and should be enabled if the detection of novel pathways or unknown mechanisms is desired. This tool is based on the assumption that even novel biosynthetic pathways, which are very different from known ones, utilize the same broad enzyme families for the catalysis of key reactions. ClusterFinder detects certain PFAM domains, which are located outside of a comprehensive set of known biosynthetic gene cluster classes and thereby predicts putative novel clusters. This search will increase the number of candidate secondary metabolite gene clusters (longer runtime!) to the cost of a lower confidence for some of the predicted clusters. For example, the amount of detected clusters in *A. orientalis* HCCB10007 was increased from 31 to 108 by enabling the Cluster Finder algorithm in addition to antiSMASH. If the ClusterFinder algorithm is used, the results should be evaluated critically.

A novel idea for the detection of unknown biosynthetic pathway classes is utilized by the INBEKT (Identification of Natural compound Biosynthesis pathways by Exploiting Knowledge of Transcriptional regulation) progression [45]. INBEKT has the basic concept to detect novel secondary metabolite genes by utilizing knowledge of gene regulation instead of detecting biosynthetic enzymes. The INBEKT concept follows the idea that global, environmental signal-sensing regulators control the production of certain secondary metabolites. Such regulators promote or repress gene transcription by their binding to specific DNA motifs

upstream of their target genes. These regulators may sense environmental signals such as nutrient starvation, oxidative stress, or the presence of competitive organisms, which can trigger the production of secondary metabolites. The computational screening of genome sequences for known DNA-binding sequences of such regulators will provide a number of candidate gene regions. The list of candidate gene regions can be minimized by excluding hits inside a comprehensive set of known biosynthetic gene clusters or hits that are associated with primary metabolism. The residual numbers of gene regions possibly direct synthesis of secondary metabolites by new pathway classes.

The preliminary step in the INBEKT workflow is the generation of 5' upstream regions (UTRs) of annotated open reading frames for the successive screening for regulator binding sequences. GetFeature (Table 2) is a web-based application, which can be used to generate 5'UTRs of your genome of interest. Sequence files have to be uploaded as EMBL or GenBank files with annotated open reading frames. Before submitting data, a 5'UTR and "all" locus tags need to be selected (Fig. 3). The GetFeature output data should be saved in a FASTA format (see Note 3).

Scanning these nucleotide data for the presence of provided regulator binding sequences will deliver candidate gene regions, which have to be critically sorted and evaluated. Some published consensus sequences can be accessed easily from databases like PRODORIC® (PROcaryotIC Database Of Gene Regulation) [46], a database that organizes information about gene regulation and gene expression in prokaryotes or CollecTF, a database for transcription factor binding sites in bacteria [47]. A more comprehensive list of databases is provided at the MEME-Suite web portal [48]. In general, it is advisable to use an accessible regulator binding sequence of an organism, which is highly related to the organism of interest. For example, if screening for iron-repressed genes is wanted, it has to be considered that gram-negative and low-GC gram-positive bacteria use the ferric uptake regulator (Fur) as iron responsive repressor, while high-GC gram-positive bacteria use its functional ortholog belonging to the DtxR protein family.

Scanning the nucleotide data for the presence of rationally chosen and user-provided motifs can be performed by using tools like FIMO (Find Individual Motif Occurrences) [49] or PatScanUI [50].

Here, we present an exemplary INBEKT workflow, where the *A. orientalis* HCCB10007 genome was screened for the presence of zinc uptake regulator (Zur) binding sequences. Zur is the major bacterial regulator sensing zinc concentrations. It represses the transcription of genes encoding zinc uptake and zinc mobilization functions by binding to palindromic A/T-rich sequences found in the promoters of its DNA targets [51]. All 8121 *A. orientalis* HCCB10007 5'UTRs were uploaded to PatScanUI as FASTA file.

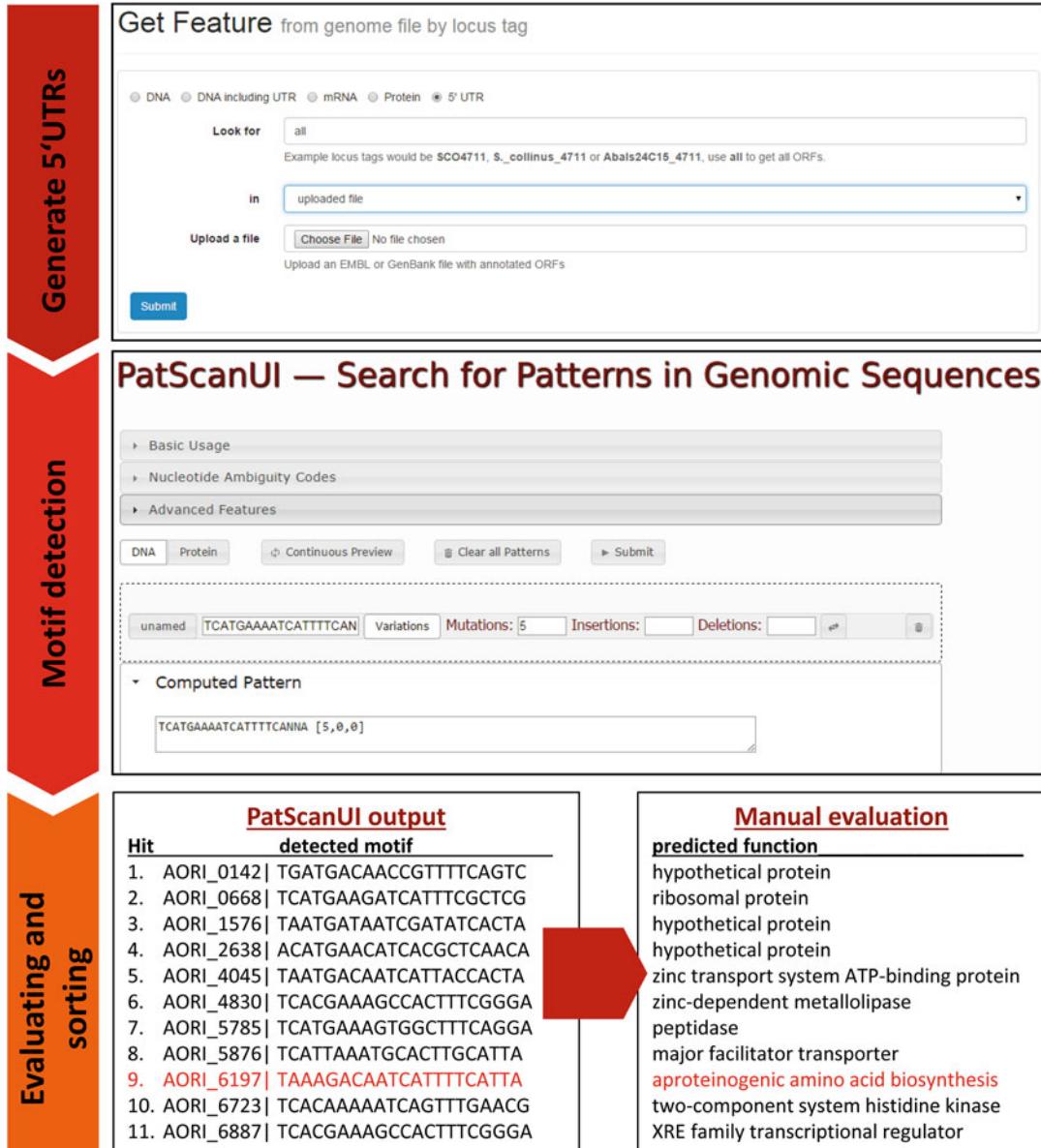


Fig. 3 Overview of the INBEKT workflow

The described Zur binding sequence (TCATGAAAATC ATTTTCANNA) of *Streptomyces coelicolor* [52] was chosen as a motif to screen for zinc repressed genes. The maximum of allowed mismatches was set to 5 (Fig. 3). The optional settings for mismatches, insertions, and deletions should be chosen empirically to find a range wide enough to detect a suitable amount of genes but exclude a lot of false positives. To estimate if the total amount of detected genes is plausible, it can be compared to the known amount of genes, which are included in corresponding regulons. Zur regulons that have been characterized so far, e.g., comprise usually 10–30 genes.

The Zur screening in *A. orientalis* HCCB10007 revealed a set of 11 genes (Fig. 3), which are putatively zinc regulated. The predicted functions of the candidate genes were assigned to known pathways when possible. Hit number 9 (AORI_6197), which is neither detected by the latest version of antiSMASH nor by the ClusterFinder algorithm, represents a protein, proposed to be involved in the synthesis of a nonproteinogenic amino acid. Such nonproteinogenic amino acids are common building blocks of various secondary metabolites. The identified AORI_6197 is highly similar to AesA of *Amycolatopsis japonica* [45] which was shown to be essential for the synthesis of an unusual zinc-responsively produced chelating agent. To date, *A. orientalis* HCCB10007 has not been described to produce such a compound.

3.3 Determining the Boundaries of a Gene Cluster

3.3.1 Comparison of antiSMASH Results with MIBiG

Clusters predicted by antiSMASH are most probably not displaying the correct boundaries, because the antiSMASH pipeline is designed to set the cluster boundaries automatically at 5, 10, or 20 kb on each side of the last signature gene, dependent on the type of gene. Without experimental validation, the real cluster boundaries cannot be exactly predicted, but they can be estimated by comparing SMGCs in different genomes of related bacterial species. In the following section, we present a set of tools to help with the estimation of gene cluster boundaries.

If the cluster of interest is similar to a known SMGC, cluster boundaries can be deduced from the additional antiSMASH output data. Thereby, the “find homologous gene clusters” and “find known homologous gene clusters” views in antiSMASH may be helpful. A comparison of *A. orientalis* HCCB10007 cluster 3 with the respective MIBiG entry for vancomycin reveals high similarity in the modular structure as well as high similarity in the set of flanking genes and therefore allows estimating the gene cluster boundaries by simple comparison. Although antiSMASH usually overestimates the gene cluster size, sometimes the known cluster is even bigger than the cluster determined by antiSMASH. In this case, the raw sequence should be carefully inspected.

Most of the SMGCs predicted by antiSMASH will not have a high similarity with clusters from the MIBiG database. In the following section, we describe the different tools that can be used to detect the SMGC boundaries by comparing the region of interest with similar regions in other, closely related bacterial strains.

3.3.2 JGI-IMG/ER: Comparative Genomics for Genomes Published in the JGI Database

An easy to use web tool is the JGI-IMG/ER (Integrated Microbial Genomes—Expert Review) genome viewer on the JGI webpage (Table 2) [53]. JGI-IMG/ER offers different genomics applications, including a genome viewer, metabolic pathway identification, annotation tools, or phylogenetic clustering programs. It is necessary to register on the JGI webpage to get access to the JGI-



Fig. 4 JGI IMG/ER “Ortholog Neighborhoods” view. Part of the vancomycin gene cluster of *A. orientalis* HCCB10007 (GCA_000400635.2) compared to the homologous regions from *A. decaplanina* DSM 44595 (GCA_000342005.1), *A. alba* DSM 44262 (GCA_000384215.1), and *A. balhimycina* DSM 44591 (GCA_000384295.1) clusters (*top down*). The cluster is highlighted in gray to indicate the boundaries. Genes conserved only in closely related strains are highlighted in orange

IMG/ER applications. One helpful tool is the gene neighborhood view, which could be difficult to find among the diverse applications. First, several genomes of interest should be loaded into the genome cart. Using the “Find Genes” function, it is possible to search for specific genes by name, function, or locus tag. For each gene an overview page with some general annotation information, such as gene families, clusters of orthologous groups (COGs), or protein family (PFAM) domains, is given. Scrolling down, the gene is shown in its direct neighborhood. Choosing the option “Show neighborhood with this gene’s bidirectional best hits” will open a view of all selected genomes from the cart that share roughly the same sized orthologs in the region of interest. By comparing the strain of interest with other strains that share the same or similar biosynthetic genes, in particular by comparing the gene cluster flanking regions, the putative boundaries of the cluster can be determined (Fig. 4).

3.3.3 Genome Comparison with MultiGeneBlast

MultiGeneBlast is an open source tool for the identification and comparison of multigene regions, such as gene clusters (Table 2) [54]. First, a database comprising the genome sequences of several closely related species must be built. This can either be done from GenBank entries on the NCBI server or from *.gbk or *.fasta files stored on a personal hard drive (see Note 7). As an example, the GenBank file of the vancomycin cluster (cluster 3) downloaded from the antiSMASH results for *A. orientalis* HCCB10007 can be used as an input file. To create an example database the “whole genome assembly” files from the NCBI server for *Amycolatopsis balhimycina* FH1894, *Amycolatopsis decaplanina* DSM 44594, *A. orientalis* B-37, *Amycolatopsis alba* DSM 44242, *Amycolatopsis mediterranei* S699, and *A. mediterranei* U32 can be downloaded by choosing “Database” → “Create from online GenBank entries.” By choosing “Make raw nucleotide database for tblastn-search,” FASTA instead of GenBank files are used. Using the option tab “File” it is possible to select the created database and open an input

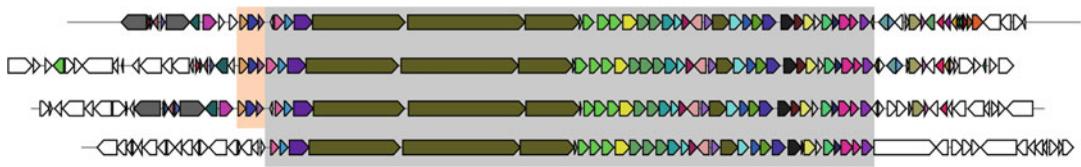


Fig. 5 MultiGeneBlast example *.xhtml output for the vancomycin biosynthesis gene cluster of *A. orientalis* HCCB10007 (GCA_000400635.2) compared to the homologous gene clusters from *A. decaplanina* DSM 44595 (GCA_000342005.1), *A. alba* DSM 44262 (GCA_000384215.1), and *A. balhimycina* DSM 44591 (GCA_000384295.1) clusters. The cluster is highlighted in gray to indicate the boundaries. Genes conserved only in closely related strains are highlighted in orange

file. For a standard approach the default settings are sufficient to determine the gene cluster boundaries. The results are displayed in a *.xhtml format. Gene clusters similar to the vancomycin cluster are present in the *A. decaplanina*, *A. alba*, and *A. balhimycina* genome. The minimum set of homologous genes shared by all strains harboring the SMGC can be seen best in *A. balhimycina*. Based on this information the boundaries can be estimated (Fig. 5). Conversely, it is sometimes possible to define a cluster by the exclusion of cluster parts that are present in all genomes, not only those harboring the SMGC of interest (see Note 8).

3.4 Working with Draft Genomes

To date, most of the published bacterial genomes are not completed, but separated on short contigs or larger scaffolds. Contigs are overlapping next-generation sequencing reads assembled into DNA sequences of high confidence, which can vary in length from a few hundred bases up to several Mb. For simple usage, a measure for the genome quality can be estimated by the number of contigs (fewer contigs = better assembly) and from the N50 value, a measure for the mean contig length, with greater weight given to longer contigs (higher N50 = better assembly). Scaffolds are assemblies of contigs, for which the relative position, but not the connecting sequence, is known. These “gaps” are expressed as larger stretches of “N”s in the genome sequence. A more sophisticated description of genome assembly methods for the interested reader is given in [55, 56].

Unfortunate for the natural product researcher, contig ends are often located in the middle of a secondary metabolite gene cluster, notably in the highly repetitive, large type-I-PKS or NRPS clusters. When working with draft genomes one could be suspicious when either the antiSMASH gene cluster ends abruptly with a type-I-PKS or NRPS like structure, or some flanking enzymes present in related SMGCs are missing. The cluster position on the contig should be taken into account, to ensure that the observed cluster is not only representing a part of the cluster, while the other part is located on a different contig.

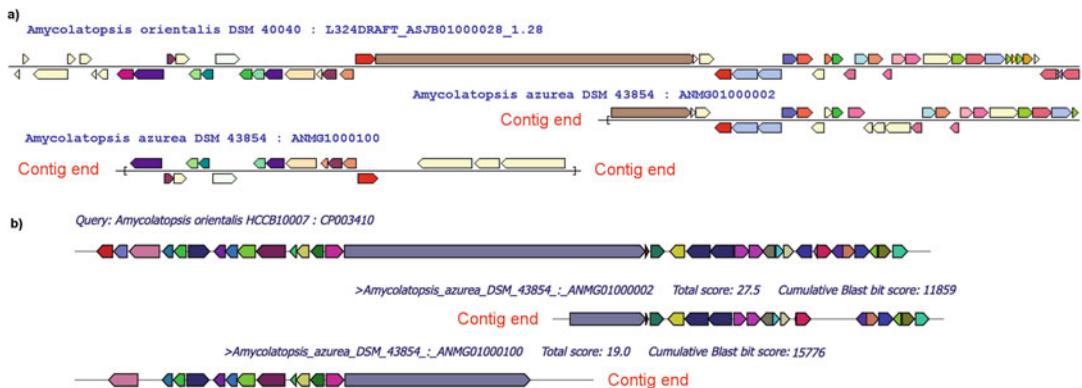


Fig. 6 (a) JGI-IMG/ER “Ortholog neighborhoods” view of *A. orientalis* DSM 40040 and the homologous regions in the *A. azurea* DSM 43854 genome which are located on two different contigs. (b) MultiGeneBlast comparison of ortholog gene clusters. *A. orientalis* HCCB10007 serves as a query. The respective region in the *A. azurea* DSM 43854 genome is located on two different contigs

The easiest method to connect gene clusters on different contigs is to map the complete draft genome to a reference genome using mapping software such as CONTIGuator (Table 2) [57]. The drawback of this approach is that it only works for highly similar SMGCs on highly similar genomes, e.g., different strains from the same bacterial species. The risk is that the mapping program misassembles the highly similar modules of type-I-PKS or NRPS. The resulting mapped cluster will be much shorter than the real cluster. Mapped clusters should therefore always be carefully validated.

The previously described JGI-IMG/ER and MultiGeneBlast applications can be of use to assign partial clusters on different contigs to one gene cluster. With the JGI-IMG/ER viewer “show neighborhood” option, contigs with sequence homology to a complete reference sequence are automatically mapped (Fig. 6a). If a draft genome is included in the MultiGeneBlast database, the separated SMGCs will be displayed as a match with a complete reference SMGC (Fig. 6b). As an example, the draft genome sequence of *A. azurea* DSM 43854 has been used. When analyzing the *A. azurea* DSM 43854 genome with antiSMASH, gene cluster 1 appeared suspicious, because cluster and contig ended right within an NRPS gene. When comparing the gene cluster with similar clusters in the JGI-IMG/ER and with MultiGeneBlast, it appeared that the SMGC is spread over two distinct contigs (Fig. 6).

Even without a complete reference gene cluster for comparison, it is possible to figure out which contigs should be merged into one cluster. Within type-I-PKS and NRPS clusters or

C-domains of the same type usually share 85–90 % homology. This information can be used to deduce which KS-/C-domains belong to the same cluster and, when comparing similar clusters on different genomes, which gene clusters are probably encoding highly similar products. A useful tool to accomplish this is NaPDoS, the *Natural Product Domain Seeker* (Table 2). NaPDoS is a web-based bioinformatic tool that can be used to identify and extract KS- and C-domains according to a BLAST search against a database of different PKS, NRPS, and hybrid clusters [58]. Furthermore, with NaPDoS it is possible to construct a phylogenetic tree based on the pairwise sequence similarity of KS- or C-domains. Another program for the construction of phylogenetic trees, which allows choosing between different tree finding algorithms and is more flexible with the choice of parameters, is MEGA 6.0 [59]. For the assignment of SMGCs according to their phylogeny, this means that KS- or C-domains belonging to the same cluster, but present on different contigs, are likely to fall in the same phylogenetic clade.

With the information gained from all three approaches it is feasible to merge the contigs. If the sequences on both contigs overlap the contigs can be assembled manually; otherwise, it is common to indicate a gap of unknown length as a stretch of “N”s in the sequence.

3.5 Prioritization of Gene Clusters

Finally, when all desired gene clusters are identified, and the boundaries have been estimated, one of the main final questions is: Which of the secondary metabolite gene clusters are worth investigating? For searching variations of already known compounds or for completely new compounds, the procedure is quite the same. It is necessary to compare the new clusters to clusters encoding already known compounds, to identify which known SMGCs show similar gene composition and therefore are likely to produce similar compounds. Based on this information, it is possible to classify gene cluster families. In the next section, the secondary metabolite databases and recent approaches for the classification of SMGC families are listed.

3.5.1 Dereplication by Comparison of Genes and Predicted Products with Secondary Metabolite Databases

To start, a comparison with the continuously growing MIBiG database [34] is a good option. A link to MIBiG is directly included in the antiSMASH output data in the “known homologous gene clusters” view. antiSMASH furthermore provides a prediction of the substrate specificity for the modular compounds of PKSs and NRPSs. A comparison of the predicted monomers can indicate if the produced compounds of two SMGCs are the same, slightly different or if they are completely different. A final similarity check can be performed by BLAST analysis.

A different approach for dereplication, which is based on the phylogeny of typical SMGC domains, is used in PRISM. Thereby,

possible variants for each product are predicted and subsequently each variant is compared to a large natural product library [38]. Further natural product databases are, ClustScan [41], DoBISCUIT [60], ClusterMine360 [61], and IMG-ABC [62].

3.5.2 Classification of Gene Cluster Families

The currently discussed approach for the prioritization of biosynthetic gene clusters is the classification of all so far discovered secondary metabolite gene clusters in gene cluster families. A SMGC classification should help to prevent replication and to predict structure and function of novel secondary metabolites by comparison with related clusters. This process is still in its early stages. Nevertheless, we would like to shortly introduce the different approaches toward solving this task.

1. A first approach was based on the definition of operational biosynthetic units (OBUs) for PKS and NRPS biosynthetic gene clusters. OBUs were classified according to a similar gene content and organization. Thereby, clusters with an amino acid sequence identity of 90% for KS-domains and 85% for C-domains were grouped [63]. However, this approach is limited to PKS and NRPS clusters and was so far only applied for *Salinispora* spp. Our recent experiences (data not published) show that for some other bacterial genera these thresholds are not applicable.
2. Another approach was based on the combination of three different similarity metrics: (a) the number of orthologous genes shared by two biosynthetic gene clusters, (b) the amount of each cluster shared in a PROmer alignment, and (c) simplified, the number of corresponding signature genes in two clusters expressed as percentage values. Creating a similarity matrix giving different weights for the different similarity indices (a: 25 %, b: 25 %, c: 50 %) allowed the implementation of these data in a distance network, clearly visualizing distinct gene cluster families [64]. The drawback of this method is the possibility that clusters producing only precursor molecules or smaller subunits can cluster in the same groups as larger and more complex secondary metabolites.
3. The third approach combined two similarity metrics: (a) the Jaccard index to measure the similarity (presence or absence) of PFAM domains from all vs. all SMGCs (weighed: 36%) and (b) the domain duplication index to measure similarities in the numbers of PFAM domains (weighed: 64%) [44]. For the graphical visualization of similarity values Cytoscape was used [65]. However, this method is not well suited for the comparison of highly repetitive multimodular clusters, such as type-1-PKS and NRPS clusters.

Finally, deciding which clusters are worth investigating is highly user dependent. If the detection of a completely new

compound is desired, clusters not belonging to any known SMGC family should be prioritized. If the objective is to find a structural and functional variant of a known compound, members of a specific SMGC family might be of interest.

For future approaches, it is necessary to define generally applicable standard rules for the classification of secondary metabolite cluster families. Furthermore, specified algorithms for the different types of gene clusters are needed.

4 Notes

1. Genome Mining is not only limited to genomic sequence data, but can also be performed for sequenced cosmid libraries and metagenomic data. However, the programs presented in this chapter are not recommended for raw sequence data or very short assembled sequences, as they need sequences of a certain length as input data. For genome mining of raw metagenomic data eSNaPD [66] is recommended.
2. When converting files, using annotation tools, etc. the same headers/identifiers should be used, so that it is always possible to identify the respective region of interest within a larger dataset. This is specifically of interest when the genome is separated on multiple contigs. Headers should be kept short, because some programs have problems with headers longer than 20 characters.
3. Filename extensions can make a difference. There are several common filename extensions for FASTA (*.fa, *.fas, *.fasta, *.fna for nucleotide, *.faa for amino acid) or GenBank (*.gbk, *.gb, *.gbf) files in use. For example, if a program does not accept FASTA data with a *.fas extension as input file, changing the extension to *.fa or *.fasta could help.
4. Poor genome annotation influences the antiSMASH results. When using antiSMASH with FASTA sequences the detected SMGCs are additionally annotated using an up-to-date annotation pipeline. Therefore, it is recommended to run antiSMASH with GenBank and FASTA sequences and to compare the results concerning missing genes or entire missing clusters.
5. antiSMASH often merges neighboring clusters into one large cluster. These false hybrid clusters can be distinguished from real hybrid clusters by comparison and detection of the cluster boundaries with MultiGeneBlast and JGI-IMG/ER.
6. Sometimes, MbtH-like structures (MbtH proteins bind to NRPS proteins to stimulate adenylation reactions) are wrongly classified as lantipeptide core peptides in antiSMASH. This is often seen in NRPS clusters. To distinguish real lantipeptide

clusters from false positives the complete cluster composition should always be taken into account. An additional blastp search against the NCBI database should be performed to validate the annotation.

7. Multi GenBank files can cause problems while running MultiGeneBlast. If such problems occur, FASTA files instead of GenBank files should be used for database creation. When a database is created from GenBank and FASTA files, the annotations in the GenBank file will be omitted and a nucleotide database (*.nal) will be created.
8. For the estimation of gene cluster boundaries, it can, in some cases, be helpful to perform a blastp search of all the genes involved. If adjacent genes show a high similarity to one genus or species and a break in this pattern is observed where a stretch of genes is most similar to a different genus or species, this could indicate the boundaries of a SMGC.

References

1. Pelaez F (2006) The historical delivery of antibiotics from microbial natural products—can history repeat? *Biochem Pharmacol* 71(7):981–990. doi:[10.1016/j.bcp.2005.10.010](https://doi.org/10.1016/j.bcp.2005.10.010)
2. Newman DJ, Cragg GM (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod* 75(3):311–335. doi:[10.1021/np200906s](https://doi.org/10.1021/np200906s)
3. Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM, Bartlett JG, Edwards Jr, Infectious Diseases Society of America (2008) The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. *Clin Infect Dis* 46(2):155–164. doi:[10.1086/524891](https://doi.org/10.1086/524891)
4. Medema MH, Fischbach MA (2015) Computational approaches to natural product discovery. *Nat Chem Biol* 11(9):639–648. doi:[10.1038/nchembio.1884](https://doi.org/10.1038/nchembio.1884)
5. Arnison PG, Bibb MJ, Bierbaum G et al (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* 30(1):108–160. doi:[10.1039/c2np20085f](https://doi.org/10.1039/c2np20085f)
6. Daum M, Herrmann S, Wilkinson B, Bechthold A (2009) Genes and enzymes involved in bacterial isoprenoid biosynthesis. *Curr Opin Chem Biol* 13(2):180–188. doi:[10.1016/j.cbpa.2009.02.029](https://doi.org/10.1016/j.cbpa.2009.02.029)
7. Hur GH, Vickery CR, Burkart MD (2012) Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat Prod Rep* 29(10):1074–1098. doi:[10.1039/c2np20025b](https://doi.org/10.1039/c2np20025b)
8. Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, Patel HM (2001) Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol* 5(5):525–534
9. Weissman KJ (2015) The structural biology of biosynthetic megaenzymes. *Nat Chem Biol* 11(9):660–670. doi:[10.1038/nchembio.1883](https://doi.org/10.1038/nchembio.1883)
10. Weber T (2010) Antibiotics: biosynthesis, generation of novel compounds. In: Flickinger, MC Encyclopedia of industrial biotechnology. John Wiley and Sons, Hoboken, NJ, USA. doi:[10.1002/9780470054581.eib040](https://doi.org/10.1002/9780470054581.eib040)
11. Khosla C, Herschlag D, Cane DE, Walsh CT (2014) Assembly line polyketide synthases: mechanistic insights and unsolved problems. *Biochemistry* 53(18):2875–2883. doi:[10.1021/bi500290t](https://doi.org/10.1021/bi500290t)
12. Piel J (2010) Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat Prod Rep* 27(7):996–1047. doi:[10.1039/b816430b](https://doi.org/10.1039/b816430b)
13. Challis GL, Naismith JH (2004) Structural aspects of non-ribosomal peptide biosynthesis. *Curr Opin Struct Biol* 14(6):748–756. doi:[10.1016/j.sbi.2004.10.005](https://doi.org/10.1016/j.sbi.2004.10.005)

14. Stricker M, Tanovic A, Marahiel MA (2010) Nonribosomal peptide synthetases: structures and dynamics. *Curr Opin Struct Biol* 20(2): 234–240. doi:[10.1016/j.sbi.2010.01.009](https://doi.org/10.1016/j.sbi.2010.01.009)
15. Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* 24(1):162–190. doi:[10.1039/b507395m](https://doi.org/10.1039/b507395m)
16. Yu D, Xu F, Zeng J, Zhan J (2012) Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* 64(4):285–295. doi:[10.1002/iub.1005](https://doi.org/10.1002/iub.1005)
17. Du L, Shen B (2001) Biosynthesis of hybrid peptide-polyketide natural products. *Curr Opin Drug Discov Devel* 4(2):215–228
18. Cane DE, Ikeda H (2012) Exploration and mining of the bacterial terpenome. *Acc Chem Res* 45(3):463–472. doi:[10.1021/ar200198d](https://doi.org/10.1021/ar200198d)
19. Chen W, Qi J, Wu P, Wan D, Liu J, Feng X, Deng Z (2016) Natural and engineered biosynthesis of nucleoside antibiotics in Actinomycetes. *J Ind Microbiol Biotechnol* 43: 401–417. doi:[10.1007/s10295-015-1636-3](https://doi.org/10.1007/s10295-015-1636-3)
20. Hamed RB, Gomez-Castellanos JR, Henry L, Ducho C, McDonough MA, Schofield CJ (2013) The enzymes of beta-lactam biosynthesis. *Nat Prod Rep* 30(1):21–107. doi:[10.1039/c2np20065a](https://doi.org/10.1039/c2np20065a)
21. Heide L (2009) The aminocoumarins: biosynthesis and biology. *Nat Prod Rep* 26(10):1241–1250. doi:[10.1039/b808333a](https://doi.org/10.1039/b808333a)
22. Kudo F, Eguchi T (2009) Biosynthetic genes for aminoglycoside antibiotics. *J Antibiot (Tokyo)* 62(9):471–481. doi:[10.1038/ja.2009.76](https://doi.org/10.1038/ja.2009.76)
23. Niu G, Tan H (2015) Nucleoside antibiotics: biosynthesis, regulation, and biotechnology. *Trends Microbiol* 23(2):110–119. doi:[10.1016/j.tim.2014.10.007](https://doi.org/10.1016/j.tim.2014.10.007)
24. Sit CS, Ruzzini AC, Van Arnam EB, Ramadhar TR, Currie CR, Clardy J (2015) Variable genetic architectures produce virtually identical molecules in bacterial symbionts of fungus-growing ants. *Proc Natl Acad Sci U S A* 112(43):13150–13154. doi:[10.1073/pnas.1515348112](https://doi.org/10.1073/pnas.1515348112)
25. Weber T, Charusanti P, Musiol-Kroll EM, Jiang X, Tong Y, Kim HU, Lee SY (2015) Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. *Trends Biotechnol* 33(1):15–26. doi:[10.1016/j.tibtech.2014.10.009](https://doi.org/10.1016/j.tibtech.2014.10.009)
26. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14): 2068–2069. doi:[10.1093/bioinformatics/btu153](https://doi.org/10.1093/bioinformatics/btu153)
27. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27(23):4636–4641
28. Meyer F, Goessmann A, McHardy AC et al (2003) GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* 31(8):2187–2195
29. Aziz RK, Bartels D, Best AA et al (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi:[10.1186/1471-2164-9-75](https://doi.org/10.1186/1471-2164-9-75)
30. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28(4):464–469. doi:[10.1093/bioinformatics/btr703](https://doi.org/10.1093/bioinformatics/btr703)
31. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T (2013) antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res* 41(Web Server issue):W204–212. doi:[10.1093/nar/gkt449](https://doi.org/10.1093/nar/gkt449)
32. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* 39(Web Server issue): W339–346. doi:[10.1093/nar/gkr466](https://doi.org/10.1093/nar/gkr466)
33. Weber T, Blin K, Duddela S et al (2015) antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* 43(W1):W237–W243. doi:[10.1093/nar/gkv437](https://doi.org/10.1093/nar/gkv437)
34. Medema MH, Kottmann R, Yilmaz P et al (2015) Minimum information about a biosynthetic gene cluster. *Nat Chem Biol* 11(9):625–631. doi:[10.1038/nchembio.1890](https://doi.org/10.1038/nchembio.1890)
35. Awakawa T, Crusemann M, Munguia J, Ziernert N, Nizet V, Fenical W, Moore BS (2015) Salinipyrone and pacificanone are biosynthetic by-products of the rosamicin polyketide synthase. *Chembiochem* 16(10): 1443–1447. doi:[10.1002/cbic.201500177](https://doi.org/10.1002/cbic.201500177)
36. Lautru S, Deeth RJ, Bailey LM, Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Biol* 1(5):265–269. doi:[10.1038/nchembio731](https://doi.org/10.1038/nchembio731)
37. Ross AC, Xu Y, Lu L, Kersten RD, Shao Z, Al-Suwailem AM, Dorrestein PC, Qian PY, Moore BS (2013) Biosynthetic multitasking facilitates thalassospiramide structural diversity

- in marine bacteria. *J Am Chem Soc* 135(3):1155–1162. doi:[10.1021/ja3119674](https://doi.org/10.1021/ja3119674)
38. Skinnider MA, DeJong CA, Rees PN, Johnston CW, Li H, Webster AL, Wyatt MA, Magarvey NA (2015) Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res* 43(20): 9645–9662. doi:[10.1093/nar/gkv1012](https://doi.org/10.1093/nar/gkv1012)
39. Johnston CW, Skinnider MA, Wyatt MA, Li X, Ranieri MR, Yang L, Zechel DL, Ma B, Magarvey NA (2015) An automated Genomes-to-Natural Products platform (GNP) for the discovery of modular natural products. *Nat Commun* 6:8421. doi:[10.1038/ncomms9421](https://doi.org/10.1038/ncomms9421)
40. Li MH, Ung PM, Zajkowski J, Garneau-Tsodikova S, Sherman DH (2009) Automated genome mining for natural products. *BMC Bioinformatics* 10:185. doi:[10.1186/1471-2105-10-185](https://doi.org/10.1186/1471-2105-10-185)
41. Starcevic A, Zucko J, Simunkovic J, Long PF, Cullum J, Hranueli D (2008) ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res* 36(21): 6882–6892. doi:[10.1093/nar/gkn685](https://doi.org/10.1093/nar/gkn685)
42. Anand S, Prasad MV, Yadav G, Kumar N, Shehara J, Ansari MZ, Mohanty D (2010) SBSPKS: structure based sequence analysis of polyketide synthases. *Nucleic Acids Res* 38 (Web Server issue):W487–496. doi:[10.1093/nar/gkq340](https://doi.org/10.1093/nar/gkq340)
43. de Jong A, van Hijum SA, Bijlsma JJ, Kok J, Kuipers OP (2006) BAGEL: a web-based bacteriocin genome mining tool. *Nucleic Acids Res* 34(Web Server issue):W273–279. doi:[10.1093/nar/gkl237](https://doi.org/10.1093/nar/gkl237)
44. Cimermancic P, Medema MH, Claesen J et al (2014) Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158(2):412–421. doi:[10.1016/j.cell.2014.06.034](https://doi.org/10.1016/j.cell.2014.06.034)
45. Spohn M, Wohlleben W, Stegmann E (2016) Elucidation of the zinc dependent regulation in *Amycolatopsis japonicum* enabled the identification of the ethylenediamine-disuccinate ([S, S]-EDDS) genes. *Environ Microbiol* 18:1249–63. doi:[10.1111/1462-2920.13159](https://doi.org/10.1111/1462-2920.13159)
46. Munch R, Hiller K, Barg H, Heldt D, Linz S, Wingender E, Jahn D (2003) PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res* 31(1):266–269
47. Kilic S, White ER, Sagitova DM, Cornish JP, Erill I (2014) CollecTF: a database of experimentally validated transcription factor-binding sites in Bacteria. *Nucleic Acids Res* 42(Database issue):D156–D160. doi:[10.1093/nar/gkt1123](https://doi.org/10.1093/nar/gkt1123)
48. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 37(Web Server issue):W202–208. doi:[10.1093/nar/gkp335](https://doi.org/10.1093/nar/gkp335)
49. Grant CE, Bailey TL, Noble WS (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27(7):1017–1018. doi:[10.1093/bioinformatics/btr064](https://doi.org/10.1093/bioinformatics/btr064)
50. Dsouza M, Larsen N, Overbeek R (1997) Searching for patterns in genomic data. *Trends Genet* 13(12):497–498
51. Fillat MF (2014) The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. *Arch Biochem Biophys* 546:41–52. doi:[10.1016/j.abb.2014.01.029](https://doi.org/10.1016/j.abb.2014.01.029)
52. Owen GA, Pascoe B, Kallifidas D, Paget MS (2007) Zinc-responsive regulation of alternative ribosomal protein genes in *Streptomyces coelicolor* involves Zur and σR. *J Bacteriol* 189 (11):4078–4086. doi:[10.1128/JB.01901-06](https://doi.org/10.1128/JB.01901-06)
53. Markowitz VM, Chen IM, Palaniappan K et al (2012) IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res* 40(Database issue):D115–D122. doi:[10.1093/nar/gkr1044](https://doi.org/10.1093/nar/gkr1044)
54. Medema MH, Takano E, Breitling R (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* 30(5):1218–1223. doi:[10.1093/molbev/mst025](https://doi.org/10.1093/molbev/mst025)
55. Bacci G (2015) Raw sequence data and quality control. *Methods Mol Biol* 1231:137–149. doi:[10.1007/978-1-4939-1720-4_9](https://doi.org/10.1007/978-1-4939-1720-4_9)
56. Orlandini V, Fondi M, Fani R (2015) Methods for assembling reads and producing contigs. *Methods Mol Biol* 1231:151–161. doi:[10.1007/978-1-4939-1720-4_10](https://doi.org/10.1007/978-1-4939-1720-4_10)
57. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A (2011) CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* 6:11. doi:[10.1186/1751-0473-6-11](https://doi.org/10.1186/1751-0473-6-11)
58. Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR (2012) The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One* 7(3), e34064. doi:[10.1371/journal.pone.0034064](https://doi.org/10.1371/journal.pone.0034064)
59. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12):2725–2729. doi:[10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197)

60. Ichikawa N, Sasagawa M, Yamamoto M, Komaki H, Yoshida Y, Yamazaki S, Fujita N (2013) DoBISCUIT: a database of secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* 41(Database issue):D408–D414. doi:[10.1093/nar/gks1177](https://doi.org/10.1093/nar/gks1177)
61. Conway KR, Boddy CN (2013) Cluster Mine360: a database of microbial PKS/NRPS biosynthesis. *Nucleic Acids Res* 41(Database issue):D402–D407. doi:[10.1093/nar/gks993](https://doi.org/10.1093/nar/gks993)
62. Hadjithomas M, Chen IM, Chu K et al (2015) IMG-ABC: a knowledge base to fuel discovery of biosynthetic gene clusters and novel secondary metabolites. *MBio* 6(4), e00932. doi:[10.1128/mBio.00932-15](https://doi.org/10.1128/mBio.00932-15)
63. Ziemert N, Lechner A, Wietz M, Millan-Aguinaga N, Chavarria KL, Jensen PR (2014) Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc Natl Acad Sci U S A* 111(12):E1130–E1139. doi:[10.1073/pnas.1324161111](https://doi.org/10.1073/pnas.1324161111)
64. Doroghazi JR, Albright JC, Goering AW, Ju KS, Haines RR, Tchálukov KA, Labeda DP, Kelleher NL, Metcalf WW (2014) A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol* 10(11):963–968. doi:[10.1038/nchembio.1659](https://doi.org/10.1038/nchembio.1659)
65. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27(3):431–432. doi:[10.1093/bioinformatics/btq675](https://doi.org/10.1093/bioinformatics/btq675)
66. Reddy BV, Milshteyn A, Charlop-Powers Z, Brady SF (2014) eSNaPD: a versatile, web-based bioinformatics platform for surveying and mining natural product biosynthetic diversity from metagenomes. *Chem Biol* 21(8): 1023–1033. doi:[10.1016/j.chembiol.2014.06.007](https://doi.org/10.1016/j.chembiol.2014.06.007)
67. Gokhale RS, Sankaranarayanan R, Mohanty D (2007) Versatility of polyketide synthases in generating metabolic diversity. *Curr Opin Struct Biol* 17(6):736–743. doi:[10.1016/j.sbi.2007.08.021](https://doi.org/10.1016/j.sbi.2007.08.021)
68. Weissman KJ (2016) Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat Prod Rep* 33:203–230. doi:[10.1039/c5np00109a](https://doi.org/10.1039/c5np00109a)
69. Stein DB, Linne U, Hahn M, Marahiel MA (2006) Impact of epimerization domains on the intermodular transfer of enzyme-bound intermediates in nonribosomal peptide synthesis. *Chembiochem* 7(11):1807–1814. doi:[10.1002/cbic.200600192](https://doi.org/10.1002/cbic.200600192)
70. Barajas JF, Phelan RM, Schaub AJ, Kliewer JT, Kelly PJ, Jackson DR, Luo R, Keasling JD, Tsai SC (2015) Comprehensive structural and biochemical analysis of the terminal myxalamid reductase domain for the engineered production of primary alcohols. *Chem Biol* 22(8):1018–1029. doi:[10.1016/j.chembiol.2015.06.022](https://doi.org/10.1016/j.chembiol.2015.06.022)
71. Agnihotri G, Liu HW (2003) Enoyl-CoA hydratase. Reaction, mechanism, and inhibition. *Bioorg Med Chem* 11(1):9–20
72. Koetsier MJ, Jekel PA, Wijma HJ, Bovenberg RA, Janssen DB (2011) Aminoacyl-coenzyme A synthesis catalyzed by a CoA ligase from *Penicillium chrysogenum*. *FEBS Lett* 585(6):893–898. doi:[10.1016/j.febslet.2011.02.018](https://doi.org/10.1016/j.febslet.2011.02.018)
73. Chan DI, Vogel HJ (2010) Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem J* 430(1):1–19. doi:[10.1042/BJ20100462](https://doi.org/10.1042/BJ20100462)
74. Chan YA, Podevels AM, Kevany BM, Thomas MG (2009) Biosynthesis of polyketide synthase extender units. *Nat Prod Rep* 26(1): 90–114
75. Haslinger K, Peschke M, Brieke C, Maximowitsch E, Cryle MJ (2015) X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* 521(7550):105–109. doi:[10.1038/nature14141](https://doi.org/10.1038/nature14141)

Chapter 3

Production of Antimicrobial Compounds by Fermentation

Henrik Harms, Gabriele M. König, and Till F. Schäberle

Abstract

The production of biologically active metabolites, e.g., antimicrobial compounds, is an essential step in the discovery and development process of medicinal drugs based on natural products. To get a hand on the compound of interest it first has to be biosynthesized by the corresponding producer, mostly a microorganism. In this chapter, a general workflow, which can easily be adapted to the lab, is described. Both fermentation on solid and in liquid medium is explained, and examples of hand on procedures are given.

Key words Microorganisms, Purification, Growth media

1 Introduction

The production of antimicrobial compounds by fermentation is an essential step throughout the whole process of identification, purification, structure elucidation, and further development of natural products. First, in the screening for biological activity, the respective compounds have to exist, meaning these metabolites of interest have to be produced. Therefore, suitable conditions for the production of the targeted compounds by the respective microorganism must be provided. Suggestions about how to achieve the requirement are given in this chapter. The following screenings are usually performed in small scale, equivalent to fermentation in volumes lower than 1 L. However, this is dependent on the test systems used. Once an interesting effect, e.g., antimicrobial activity, is detected, the next step would be to get a hand on the corresponding metabolite(s). Therefore, sufficient amount of the compound has to be produced to enable the purification and the structure elucidation of the active compound. Even though the instrumentation is getting more and more sensitive, enough crude extract must be available to start this process. Ideally, after the purification process some milligrams of the compound will be available to enable a comprehensive analysis of the natural product. This might include besides testing the chemical and physical properties

of the molecule also mode of action (MoA) studies. Also chemical synthesis of compounds of interest can be attempted. However, many biologically active natural products are structurally so complex that a time and cost-efficient synthesis becomes not suitable. Hence, the biosynthesis by microorganisms remains the method of choice for all downstream following steps. If in the long run clinical studies should be performed to develop the molecule as a medicinal drug, a lot of more aspects, e.g., good manufacturing practice (GMP, the practices required in order to meet the guidelines recommended by the authorities in charge to control that the process and the products are at high quality and do not pose any risk to the consumer or public), must be considered. However, these aspects in the development of a drug will not be considered in this article. Rather, the process to produce the targeted antimicrobial compound to perform all the experiments related to basic research will be exemplified. Of course, microorganisms are very diverse and due to that the fermentation is dependent on the respective producer strain. In some producer strains the biosynthesis can be tightly regulated, that only under highly defined conditions the compound is produced at all; or that under standard conditions the amount produced is too low for further studies. But, a general workflow can be deduced, and in the Notes section (Subheading 4) points of variation will be addressed.

2 Materials

Beside vessels for fermentation, e.g., glass flasks and petri dishes, some instrumentation is needed. An autoclave for the sterilization of the media and a laminar airflow to enable sterile inoculation must be present, as well as the opportunity for fermentation, e.g., tempered shakers or bio fermenters. Prepare all solutions using deionized water. Be aware that the pH value of deionized water can already relevantly differ from a neutral pH value. The pH value should always be checked after the preparation of media solution, but before you add any agar to the mixture (to avoid clogging of the pH meter). Thermal disinfection procedures may decrease the pH value about 0.2–0.4 units. Therefore, check and adjust the pH value afterward or take this effect in account when you prepare the media. Take into account that pH measurements with a pH meter are temperature dependent. Prepare and store all reagents at room temperature (unless indicated otherwise). Media containing vitamins should be stored in the dark (to avoid photo-induced degradation). Diligently follow all waste disposal regulations when disposing waste materials.

The detailed exemplary preparation of two fermentation media is given in the Methods part. However, general thoughts have to be considered: A cultivation medium contains some kind of

nitrogen source; either in the form of ammonium salts, selected amino acids, complex peptides (e.g., casein) or their degradation products (e.g., casitone). Furthermore, carbohydrates and energy sources are required. While many organisms can utilize amino acids (e.g., they can access the amino acids from peptone), the addition of a specific carbohydrate source either in the form of selected sugar monomers (e.g., glucose) or in a more complex form of sugar monomers and polymers (e.g., malt extract) is often beneficial for microbial growth rates.

2.1 Media Supplements

1. Trace element solution with iron [1]: Weigh 2.0 g of $\text{FeSO}_4 \times 7 \text{ H}_2\text{O}$, 70 mg $\text{ZnCl}_2 \times 7 \text{ H}_2\text{O}$, 100 mg $\text{MnCl}_2 \times 4 \text{ H}_2\text{O}$, 62 mg H_3BO_3 , 190 mg $\text{CoCl}_2 \times 6 \text{ H}_2\text{O}$, 17 mg $\text{CuCl} \times 2 \text{ H}_2\text{O}$, 24 mg $\text{NiCl}_2 \times 6 \text{ H}_2\text{O}$, and 36 mg $\text{Na}_2\text{MoO}_4 \times 2 \text{ H}_2\text{O}$. Solve each substance in 10–100 mL dem. water. Solve 5.2 g $\text{Na}_2\text{-EDTA} \times 2 \text{ H}_2\text{O}$ in 200 mL dem. water. Add metals in the given order (to avoid precipitation) to the EDTA solution to a final volume of 1 L and adjust with 2 M HCl to pH 3. The solution can be sterile filtrated (*see* Subheading 3.2), and stored at +4 °C several months. Add 1 mL of trace element solution per 1 L medium.
2. Trace element solution without iron for myxobacteria [1]: Weigh 200 mg ZnCl_2 , 1000 mg $\text{MnCl}_2 \times 4 \text{ H}_2\text{O}$, 100 mg H_3BO_3 , 100 mg $\text{CuSO}_4 \times 6 \text{ H}_2\text{O}$, 200 mg CoCl_2 , 50 mg $\text{SnCl}_2 \times 2 \text{ H}_2\text{O}$, 50 mg $\text{LiCl} \times 2 \text{ H}_2\text{O}$, 200 mg KBr, 200 mg KI, and 100 mg $\text{Na}_2\text{MoO}_4 \times 2 \text{ H}_2\text{O}$. Solve each substance in 10–100 mL dem. water. Solve 5.2 g $\text{Na}_2\text{-EDTA} \times 2 \text{ H}_2\text{O}$ in 200 mL dem. water. Add 10 mL of each metal solution in the given order (to avoid precipitation) to the EDTA solution to a final volume of 1 L. The solution can be sterile filtrated (*see* Subheading 3.2), and stored at +4 °C several months. Add 1 mL of trace element solution per 1 L medium.
3. Wolfe's vitamin solution [2]: Weigh 10.0 mg pyridoxine-HCL, 5.0 mg thiamine-HCl, 5.0 mg riboflavin, 5.0 mg nicotinic acid, 5.0 mg calcium D-(+)-pantothenate, 5.0 mg p-aminobenzoic acid, 5.0 mg thioctic acid, 2.0 mg biotin, 2.0 mg folic acid, 0.1 mg vitamin B₁₂ (*see* Note 1). Solve all substances in 1 L dem. water. The solution can be sterile filtrated (*see* Subheading 3.2), and stored at +4 °C several weeks. Keep the solution stored in darkness. Add 10 mL of vitamin solution per 1 L medium.
4. Vitamin B₁₂ stock solution: Weigh 50 mg Vitamin B₁₂. Solve in a 100 mL dem. water. The solution can be sterile filtrated (*see* Subheading 3.2), and stored in the darkness at +4 °C several months. Add 1 mL of vitamin B₁₂ stock solution per 1 L medium.

2.2 Tenellin Medium [3]

1. Mannose solution: Weigh 50 g mannitol. Solve mannitol in a suitable vessel with 200 mL dem. water (25%) (*see* Note 2), and autoclave the sugar solution (121 °C, 20 min).

2. Tenellin metal solution: Weigh 0.5 g $\text{MgSO}_4 \times 7 \text{ H}_2\text{O}$, 0.2 g CaCl_2 , and 20 mg $\text{FeSO}_4 \times 7 \text{ H}_2\text{O}$ and solve in a suitable vessel with 10 mL dem. water (see Note 3), and sterile filtrate or autoclave the solution (121 °C, 20 min). Store at +4 °C.
3. Tenellin trace metal solution: Weigh 880 mg $\text{ZnSO}_4 \times 7\text{H}_2\text{O}$, 40 mg $\text{CuSO}_4 \times 5 \text{ H}_2\text{O}$, 7.5 mg $\text{MnSO}_4 \times 4 \text{ H}_2\text{O}$, 6 mg Boric acid, 4 mg $(\text{NH}_4)_6\text{Mo}_7\text{O}_4 \times \text{H}_2\text{O}$. Solve in a suitable vessel with 1 L dem. water, and sterile filtrate the solution (see Subheading 3.2). Store at +4 °C.
4. Tenellin medium: Weigh 5.0 g KNO_3 , 1.0 g KH_2PO_4 , 0.1 g NaCl. Solve in a 1 L vessel with 780 mL dem. water, adjust pH to 7.0 with HCL and/or NaOH. Autoclave the solution (121 °C, 20 min). Allow the solution to cool down to approximately 40 °C (see Note 4). Add 200 mL mannose solution (Subheading 2.2, item 1) and 10 mL of tenellin metal (Subheading 2.2, item 2) solution. Add 10 mL of tenellin trace element solution (Subheading 2.2, item 3); the final volume should be 1 L.

2.3 MD1+G Medium [4]

1. Weigh 3.0 g casitone, 0.7 g $\text{CaCl}_2 \times 2 \text{ H}_2\text{O}$, 2.0 g $\text{MgSO}_4 \times \text{H}_2\text{O}$, 2.2 g glucose $\times \text{H}_2\text{O}$. Solve in a suitable vessel with 1 L dem. water. Adjust pH to 7.2 with HCL and/or NaOH. Autoclave solution (121 °C, 20 min).

2.4 Further Materials and Hardware

1. pH meter.
2. Disposable syringe.
3. Fitting cannula.
4. Hydrophilic membrane filter with a pore size of 0.2 µm.
5. 1 L graduated cylinder.
6. Stirring bar and a magnetic stirrer.
7. 1 L screw-capped glass bottle.
8. Water bath or drying cabinet.
9. 90 mm polystyrene petri dishes.
10. 100 mL sterilized Erlenmeyer flask.
11. 5 L Erlenmeyer flask.
12. 1.5 mL cryo tubes.
13. Separation funnel.
14. 1.8 L fernbach flasks.
15. Glycerol solution (100%).
16. Agar.
17. Adsorber resin, e.g., Amberlite XAD 16.
18. 50:50 solution of acetone/MeOH.
19. Centrifuge.

20. 750 mL centrifuge bottles.
21. Glass filter size.
22. Inoculation loop.
23. Shaker.
24. Side-arm flask with neoprene adapter.

3 Methods

3.1 Medium Selection

A sufficient media has to be selected according to the special requirements of the strain used for fermentation. Usually, a medium has to contain energy source(s) and nutrients, eventually further ingredients like trace elements and vitamins. There are compilations of cultivation media recommended for different purposes and organisms which can be found in the literature and databases, e.g., <https://www.dsmz.de>, Bergey's Manual of Systematic Bacteriology [5]. Cultivation media can inter alia be subdivided into solid or liquid media (containing gelling agents like agar or not) and in synthetic or complex cultivation media (containing only exactly defined ingredients or also complex, less defined natural products like, e.g., yeast extracts or peptone). Tenellin medium and MD1+G medium are examples of a synthetic and a complex liquid cultivation medium, respectively. (With the addition of the complex solidifying agar ingredient, the liquid synthetic tenellin medium becomes, strictly speaking, a complex medium).

1. Choose a fermentation medium for the microorganism used, containing energy source(s) and nutrients, eventually further ingredients like trace elements. Starting point should be a described medium, before variations of the medium are considered.
2. Some minerals are essential for the growth and metabolism of bacteria. In complex media they can already be provided by the utilized complex natural products (e.g., yeast extract or not purified agar). In synthetic media they have to be supplemented. Especially in such media the addition of trace metals can be beneficial for bacterial growth. (Be aware that higher metal concentrations are toxic. Thus, the concentration of those must be kept at low concentrations). Trace element solutions can be sterile filtrated to the media after autoclavation (*see Note 3*). Microorganism species have specific requirements; *see Subheading 2.1, items 1 and 2* for two examples of trace element solutions that have been proven to be useful in our working group.
3. Addition of small concentrations of NaCl (0.25–1% [w/v]) often possesses growth promoting effects. Especially marine-derived bacteria may require a certain amount of salt for optimal growth or are even not able to grow without salt (Table 1).
4. Growth factors like vitamins can be essential for some microorganisms [6]. Especially for synthetic media vitamin solutions

Table 1
Categories of halophilic microorganisms

NaCl [w/v]	
1–5 %	Weak halophilic
5–15 %	Moderate halophilic
15–30 %	Extreme halophilic

like Wolfe's Vitamin solution (*see* Subheading 2.1, item 3) can be beneficial. However, it has to be considered that complex media ingredients, e.g., yeast extracts may already contain relevant vitamin concentrations. Due to the thermolability and photo lability of some vitamins these solutions should be stored in darkness and not be added before a heat sterilization process like autoclavation. Instead, these solutions can be added to the cooled down media under a laminar airflow afterward.

3.2 Sterile Filtration

1. Choose a disposable syringe according to the filtrated volume (typical 10–50 mL).
2. Attach a fitting cannula and draw up the liquid that has to be filtered.
3. To get rid of air bubbles hold the syringe vertically so that the cannula is facing upward.
4. Flick the syringe to move any bubbles to the top.
5. Then slowly push in the plunger until you have released all air bubbles.
6. Remove the cannula. Open a sterile packaged sterile, hydrophilic membrane filter with a pore size of 0.2 µm.
7. Attach the filter with the inlet to the syringe and avoid touching the outlet.
8. Screw a sterile packaged fitting cannula to the outlet.
9. Slowly push in the plunger until the whole filter is wetted and all air has been released.
10. Pipette the desired volume in a laminar air flow system to the final container.
11. Never reclose the used cannula with their cap. Dispose them directly into the cannula waste container to avoid needle-prick injuries (*see* Note 5).

3.3 Medium Preparation for Fermentation on Solid Medium

1. Add 1 L water to a 1-L graduated cylinder and solve the media components completely, using a stirring bar and a magnetic stirrer.
2. Add around 0.5 L to a 1 L screw-capped glass bottle.

3. Add a stirring bar and start the magnetic stirrer that the water forms a maelstrom.
4. Weigh all required media ingredients except of agar and thermolabile ingredients (e.g., vitamins) and add them to the vessel.
5. Fill up to 1 L with the residual water.
6. As soon as all ingredients are solved measure the pH with a pH meter and adjust the pH value of the solution with HCl or NaOH. (Take into account that the pH value may change after the autoclavage and has eventually to be adjusted in advance).
7. Add agar to the medium (15–20 g/L, *see Note 6*). Let the medium swell for 15 min at room temperature. Heat, without boiling the medium until the agar has completely solved.
8. Sterilize the medium using an autoclave (121 °C, 20 min). Take care that the lid is not tightly closed and a steam indicator band is attached to the bottle to check the sterilization process.
9. If necessary you can add thermolabile ingredients that have been sterile filtrated (*see Subheading 3.2*), after the media has cooled down to approximately 40 °C (*see Note 4*). This operation has to be performed in a laminar airflow system.

3.4 Pour Agar Petri Dishes

1. Wait until the media has cooled down to approximately 40–60 °C. It can be helpful to store the media in such a tempered water bath or drying cabinet (*see Note 4*).
2. Label the bottom of sterile standard 90 mm polystyrene petri dishes in advance accordingly. Ideally, place the petri dishes in a laminar-airflow system.
3. Lift the lid of the plates at an angle and pour sufficient culture medium into each petri dish to cover the bottom (approximately 15–20 mL). Keep the lid above the plate and cover the plate directly afterward. You must not blot the lid with media and try to avoid the formation of air bubbles.
4. In order to avoid condensation cover the plates with a thermo-isolating material, e.g., some layers of paper.
5. If you are working in a laminar-airflow system you can also place the lid slightly diagonally on the petri dish for a while, to allow faster solidification of the plates.
6. Let the petri dishes stand until solidification (approximately 10–15 min). Once solidified, stack the dishes bottom side facing up.

3.5 Solid Preculture

1. Depending on the availability of the microbial culture either use 150 µl from a well-shaken liquid cryo culture (*see Subheading 3.8*), or use a sterilized inoculation loop to gently scratch microorganisms from the surface of an agar plate.
2. Streak out microbial material on a standard agar plate (*see Subheading 3.3*).

3. You can use dilution plating to verify the purity of the culture (*see Note 7*).
4. Incubate the plate in a tempered incubator, according to the microorganism duration of growth and temperature optimum.

3.6 Liquid Preculture

1. Prepare around 20 mL liquid medium in a 100 mL sterilized Erlenmeyer flask (flasks should only be filled to 30% of the flask volume).
2. Depending on the availability of the microbial culture, either use 150 µl from a well-shaken liquid cryo culture (*see Subheading 3.8*), or use a sterilized inoculation loop to gently scratch microorganisms from the surface of an agar plate. Transfer material into the preculture flask.
3. Incubate the flask in a tempered room/chamber according to the microorganism growth rate time and temperature optimum.
4. Put the flask on a shaker and verify that the flask is firmly attached before you start the shaker.
5. Set the shaking frequency that the liquid level in the flask rises up to two third of the flask height during shaking (Orbital shaker typically use between 140 and 350 rpm). Cultivate the microorganism until the late exponential or early stationary growth phase of the organism (*see Note 8*).

3.7 Main Culture (See Notes 9–12)

1. Prepare 1 L cultivation medium in a 5 L Erlenmeyer flask.
2. In order to avoid end product-inhibition add 2% adsorber resin, e.g., Amberlite XAD 16 (*see Note 13*).
3. Wash adsorber resin with a 50:50 solution of acetone/MeOH and afterward with 10 L water, before you add it to the Erlenmeyer flask.
4. Sterilize the 5 L Erlenmeyer flask by autoclavation.
5. After cooling add thermolabile ingredients via sterile filtration (*see Subheading 3.2*).
6. Add two precultures to the 5 L Erlenmeyer flask.
7. Incubate the flask in a tempered room/chamber according to the microorganism and temperature optimum.
8. Put the flask on a shaker and verify that the flask is firmly attached before you start the shaker.
9. Set the shaking frequency, that the liquid level in the flask rises up to two third of the flask height during shaking (Orbital shaker typically use between 140 and 350 rpm).
10. Cultivate the microorganism until the late exponential or early stationary growth phase of the organism or rather until the desired compound production has reached its maximum (*see Note 8*).

11. Without adsorber resin: Liquid-Liquid extraction using ethyl acetate with a separation funnel. Process the upper ethyl acetate phase.
12. With adsorber resin: Centrifugation for 30 min and $2666 \times g$ or through filtration with a size 2 glass filter funnel, connected to a side-arm flask by means of a neoprene adapter, with a tube leading to a vacuum pump to separate cells and adsorber resins from the medium.
13. Extract cells and adsorber resins with acetone and methanol.

3.8 Preparation of Cryo Cultures for Long-Term Preservation

1. Microorganism has to be available as pure culture, if necessary check with dilution plating (*see Note 7*).
2. Cultivate microorganism under individual optimal growth conditions. Use a 100 mL Erlenmeyer flask for 30 mL growth promoting liquid medium (usually a complex nutrient rich cultivation medium with preferably low amount of electrolytes, since high intracellular electrolyte concentrations harm cells during the freezing process).
3. Harvest cells in the late exponential growth phase.
4. Pipette 1.5 mL into a sterile 1.5 mL cryo tube.
5. Close the cryo tube and allow cells to sediment (*see Note 14*).
6. Remove the supernatant with a pipette.
7. Add 500 μ L fresh sterilized glycerol solution (100%).
8. Add 500 μ L sterilized liquid cultivation medium solution, the same that was used to prepare the cell culture (hence, the final glycerol concentration is ca. 50%).
9. Mix gently and let the glycerol react with the cell suspension for 30 min at room temperature.
10. Store the cryo culture in a -70°C or -80°C freezer (*see Note 15*).

3.9 Example for Fermentation in Synthetic Medium: *Dichotomomyces cepii* Isolation Number "293 K09"/ Strain Number 225 Fermented in Tenellin Medium [3]

1. Prepare mannose solution (*see Subheading 2.2, item 1*).
2. Prepare tenellin metal solution (*see Subheading 2.2, item 2*).
3. Prepare tenellin trace metal solution (*see Subheading 2.2, item 3*).
4. Prepare tenellin medium (*see Subheading 2.2, item 4*) and add 15 g/L agar before you start the autoclavage process. Let the solution swell for 15 min. Heat the solution without boiling until the agar has solved. Start the autoclavage process (121°C , 20 min).
5. After all solutions have cooled down to approximately 40°C (*see Note 4*) combine mannose solution, tenellin metal solution, tenellin trace metal solution, and tenellin medium.
6. Before the medium solidifies pour approx. 250 mL medium in each sterilized 1.8 L fernbach flasks.

7. Incubate the flasks either with a solid preculture (approx. 30 mm sized densely overgrown agar) or with 10 mL liquid preculture).
8. Seal each fernbach flask with a sterilized and air-permeable plug. [Steps 6–10 have to be performed in a laminar airflow system].
9. Cultivate the fungus for 40 days at 20 °C under artificial light.
10. Homogenize fungal biomass and media.
11. Extract each fernbach flask three times with approximately 40 mL ethyl acetate.
12. Combine all ethyl acetate phases in a suitable flask.
13. Remove the ethyl acetate solvent with a rotary evaporator.
14. Subsequently, the dried ethyl acetate crude extract can be further processed.

3.10 Example for Fermentation in Complex (Liquid) Medium: Corallopyronin A Producer Strain *Corallococcus coralloides* B035 Fermented in MD1 + G [4], A Standard Medium for Cultivation of Myxobacteria [7]

1. For the isolation of corallopyronin A, prepare MD1 + G medium.
2. Allow the solution to cool down to approximately 40 °C (see Note 4).
3. Add 1 mL/L vitamin B₁₂-solution. Add 1 mL/L trace-elements-solution with iron.
4. Inoculate the preculture flask with myxobacterial fruiting body fragments (yellow/orange structures that can be seen phenotypically by the pure eye) derived from an agar plate culture [7].
5. Inoculate the liquid preculture for 4 days (30 °C, shaking at 140 rpm).
6. In order to prepare a liquid preculture add 100 mL MD1 + G medium to a 300 mL flask).
7. Use a 4 days old preculture to inoculate 1.5 L medium in a 5 L Erlenmeyer flask.
8. Wash sufficient amounts of Amberlite XAD 16 with a 50:50-solution of acetone/MeOH and afterward with 10 L water.
9. Add to each flask 2 % adsorber resin Amberlite XAD 16, which binds lipophilic compounds, and is necessary to avoid end-product inhibition.
10. Incubate the main culture for 7 days (30 °C, shaking at 140 rpm).
11. After fermentation, separate cells and Amberlite XAD 16 from the medium through centrifugation for 30 min and 2666 × g or through filtration with a glass filter size 2.
12. Autoclave and discard the residual medium.
13. Cells and Amberlite XAD 16 are further processed.

4 Notes

1. Mind the minimum load of your balance and work with higher concentrated stock solutions from which you can take an aliquot when you are dealing with very small sample weights.
2. In media that contains higher contents of sugars you should autoclave the latter separately (concentration of the sugar solution max. 25 %) or rather sterile filtrate them into the media (to avoid the formation of growth inhibiting caramelization and Maillard reaction byproducts). Particular high concentrated sugar solutions (double-strength concentrates) or solutions with unsolved sugar and peptone sediments, as well as media that contain high amounts of phosphate salts as well as media with alkaline pH above 8.0, are affected.
3. Metal ions like Ca^{2+} , Mg^{2+} , $\text{Fe}^{2+/3+}$ may precipitate during autoclavage, especially in solutions with higher pH values and in the presence of higher phosphate amounts. This generally does not inhibit microbial growth. Therefore, direct addition to a medium before autoclavage can be tolerated. For more accurate condition and especially if you compare different media conditions the addition in the form of a sterile filtered solution is advised.
4. A fast check for the right temperature is to hold the autoclaved glass bottle against your cheek. If it is just warm and not painful, thermolabile ingredients can be added.
5. Simple Bubble point test to check for filter integrity: Release the filter from the syringe; Draw 10 mL of air into the syringe; Reattach the wetted filter and cannula to the syringe; Dip the tip of the prepared syringe into a water-filled vessel; Slowly push in the plunger until the first air bubbles are released; If air bubbles are released before the air is compressed to 2 mL filter integrity is doubtful and has to be repeated. This is a quick and easy but rough approximation. A more accurate alternative is the utilization of a manometer that allows determining an accurate value for the bubble point that can be compared with the sterile filter bubble point value given by the manufacturer.
6. For solid media with pH values < 5.0 higher agar concentrations are required for sufficient gel stability.
7. Dilution plating: Work in a laminar airflow system; Sterilize an inoculation loop in the external tap of a Bunsen burner; Let the inoculation loop cool down for 15 s; Check if the inoculation loop has sufficiently cooled down by tapping the inoculation loop on unused spots at the corner of the agar plate; Pick a single microbial colony from a petri dish or tap the inoculation loop in a microbial liquid suspension; Start at the left upper side of the plate and streak out serpentine like several

times to the opposite right side, slide over the agar and do not penetrate the agar surface; Sterilize the inoculation loop in the external tap of a Bunsen burner again; Turn the plate around 90° streak out to the other side, starting at the end of the last streak; Sterilize the inoculation loop in the external tap of a Bunsen burner again; Turn the plate around 90°, streak out to the other side, starting at the end of the last streak.

8. Often the desired compound is only produced under a special growth condition. Therefore, it is of advantage to test samples from different fermentation time points for the yield of the compound. In that way the ideal conditions, e.g., how long fermentation should be, can be determined.
9. For many bacteria it is known that antimicrobial compounds are not produced in complex media. Instead, a certain shortage in the fermentation broth, e.g., low phosphate availability, can trigger the biosynthesis. Thus, the use of different media for the incubation of the preculture, and of the main culture might be of advantage. A complex medium, in which the microorganism growth is fast, can be used to produce enough cell material for the preculture. This forms the inoculum for the main culture which is fermented in the optimized medium for production.
10. To get more insights into the critical parameters for production, bioreactors instead of shaking flask experiments can be used. By monitoring and/or regulating parameters in the fermentation broth, e.g., pH and O₂-saturation, the yield of the fermentation is increased.
11. The supplementation of the fermentation with substances that trigger the production might be advantageous. Although mostly the regulation mechanisms of the biosynthesis are not known, e.g., factors that trigger production positively, some compounds like hormaomycin [8] have been used to induce production.
12. It might be even the case that the production of a compound is only triggered by the presence of another microorganism. Thus, it can be that the idea of an axenic culture for fermentation is not always applicable.
13. For bacteria that particularly rely on interspecies messenger substances for their development, e.g., some marine myxobacteria, it can be beneficial not to add adsorber resins like Amberlite XAD 16 or Sephabeads directly, but rather a few days after the incubation has started.
14. Sedimentation of cells can be significantly enhanced by centrifugation of the cryo tubes for 5 min at 4000–6000×*g* for bacteria and 2000–3000×*g* for yeasts.

15. The cryo tube can be placed in a Styrofoam box (each cryo tube should be surrounded by 10–15 mm isolation material); Optimal is a low cooling rate about $-1\text{ }^{\circ}\text{C}/\text{min}$ to $-30\text{ }^{\circ}\text{C}$, followed by a fast cooling rate to $-70\text{ }^{\circ}\text{C}$, since this prevents the formation of intracellular ice crystals; For the reactivation of the cryo cultures the frozen cryo tube solution should be quickly warmed up in a tempered water bath (mostly $37\text{ }^{\circ}\text{C}$, but this has to be in accordance with the growth conditions of the microorganism), since this prevents intracellular recrystallization to larger crystals during thawing.

Acknowledgment

This work was supported by the German Federal Ministry of Education and Research (BMBF) through the German Centre for Infection Research (DZIF) initiative.

References

1. Balows A, Trüper HG, Dworkin M et al (1992) *The Prokaryotes*. Springer, New York, NY
2. Atlas RM (2010) *Handbook of microbiological media*, 4th edn. ASM Press ; CRC Press / Taylor & Francis, Washington, D.C. : Boca Raton, FL
3. Harms H, Kehraus S, Mosaferan DN et al (2015) $\text{A}\beta$ -42 lowering agents from the marine-derived fungus *Dichotomomyces cepii*. *Steroids* 104:182–188
4. Erol O, Schäberle TF, Schmitz A et al (2010) Biosynthesis of the myxobacterial antibiotic corallopyronin A. *Chembiochem* 11(9): 1253–1265
5. Garrity G (ed) (2001–2011) *Bergey's manual of systematic bacteriology*, 2nd edn, vol 1–5. Springer, New York, Berlin, Heidelberg etc
6. Koser SA (1968) Vitamin requirements of bacteria and yeasts. Charles C. Thomas Publisher, Springfield, IL
7. Behrens J, Flossdorf J, Reichenbach H (1976) Base composition of deoxyribonucleic acid from *Nannocystis exedens* (Myxobacterales). *Int J Syst Bacteriol* 26(4):561–562
8. Höfer I, Crüsemann M, Radzom M et al (2011) Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chem Biol* 18:381–391

Chapter 4

Structure Elucidation of Antibiotics by NMR Spectroscopy

**Georgios Daletos, Elena Ancheeva, Raha S. Orfali, Victor Wray,
and Peter Proksch**

Abstract

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for the structure elucidation of antibiotics in solution. Over the past 30 years there have been numerous publications describing the use of NMR to characterize naturally derived or synthetic antibiotics. A large number of one-dimensional (1D) and two-dimensional (2D) NMR methods are available today and the list continues to expand. In this chapter, we will consider the key NMR experiments that provide useful information for compound structure elucidation.

Key words NMR, Structure elucidation, Antibiotics, Natural products, Callyaerin A

1 Introduction

Natural products and their derivatives have historically been an untapped source of antibacterial leads used in the development of drugs to treat bacterial infections. Examples of naturally derived antimicrobial agents of different antibiotic classes are penicillin G originally isolated from *Penicillium notatum*, erythromycin from *Saccharopolyspora erythraea*, streptomycin from *Streptomyces griseus*, and polymyxins from *Bacillus polymixa*, among others. Most of these compounds were developed by screening soil-dwelling microorganisms, such as actinomycetes, during the golden era of antibiotic discovery in the 1940s to 1960s [1]. In recent years, however, microbial pathogens are becoming increasingly resistant to clinically applied antibiotics and pose a global threat to human health [2]. Thus, there is an overwhelming need to search for new antibiotics with new targets and novel mechanisms of action.

Nature has inspired chemists and biologists alike for many decades, providing a rich and unprecedented diversity of evolutionary preselected lead structures unparalleled even by the largest combinatorial databases [3]. Since there are still many unexplored resources in nature, the potential for finding new bioactive

compounds that could be optimized to yield therapeutic agents is also enormous. This fact coupled with advances in approaches for natural-product isolation and identification could open the door to a new era in the investigation of antibiotics [4, 5].

As part of our ongoing search for new antibiotic natural products, chemical investigation of the Indonesian sponge *Callyspongia aerizusa* revealed a group of unusual cyclic peptides, called callyaerins [6, 7]. The basic structural unit of the callyaerins comprises a cyclic peptide with a linear peptide side chain, both of variable size, linked through a nonproteinogenic (*Z*)-2,3-diaminoacrylic acid (DAA) moiety. Among the isolated peptides, callyaerin A showed potent activity against *Mycobacterium tuberculosis* with MIC₉₀ value of 2 μ M (Fig. 1). In addition, callyaerin A exhibited no cytotoxicity toward THP-1 (human acute monocytic leukemia) or MRC-5 (human fetal lung fibroblast) cells ($IC_{50} > 10 \mu$ M), which highlights the potential of this compound as a promising lead for new antitubercular agents [7].

The structure elucidation of callyaerins was a challenging task and was based mainly on extensive nuclear magnetic resonance (NMR) spectroscopic analysis. NMR spectroscopy is an extremely powerful and widely used method for the structure determination of natural products in solution, including antibiotics [5, 8]. NMR experiments are based on the magnetic properties of atomic nuclei. When placed in a powerful homogeneous magnetic field, certain nuclei (e.g., ¹H, ¹³C, ¹⁵N) undergo resonance at specific radio frequencies in the electromagnetic spectrum to produce signals that can be readily detected. Detailed analysis of these diagnostic signals in the NMR spectrum provides unique information about the molecular structure of the compound under investigation. Over the past 30 years, major improvements in spectroscopic instrumentation hardware allow structure analysis to be carried out on sub-milligram amounts of a compound [9]. Moreover, the advent of multidimensional NMR techniques has revolutionized structure elucidation of natural products so that, in the large majority of cases, unambiguous structural information is obtained, even for

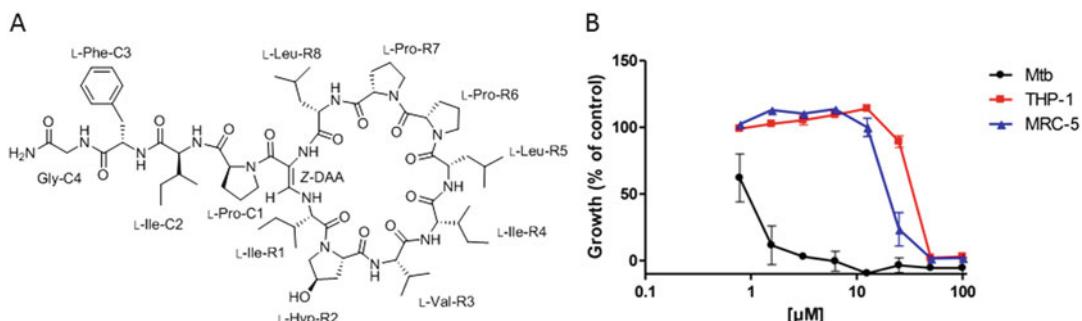


Fig. 1 (a) Structure of callyaerin A. (b) Antitubercular (*Mtb*) and cytotoxicity (THP-1 and MRC-5) profile of callyaerin A

highly complex molecular structures [10]. Herein, we offer an overview of the one-dimensional (^1H , ^{13}C , DEPT) and two-dimensional (COSY, TOCSY, HSQC, HMBC, and NOESY/ROESY) techniques that were successfully applied in the case of callyaerins and have proved most useful for the structure elucidation of other naturally derived or synthetic antibiotics.

2 Materials

1. Glass vial containing freeze-dried, purified compound.
2. NMR tube (*see Notes 1 and 2*).
3. Deuterated solvent, such as CDCl_3 , $\text{DMSO}-d_6$, or $\text{MeOH}-d_4$ (*see Notes 3 and 4*).
4. Glass Pasteur pipettes, natural rubber teats, glass wool, and Parafilm[®] (*see Note 5*).

3 Methods

3.1 Sample Preparation

1. Dry the purified compound (*see Note 6*).
2. Use a glass Pasteur pipette for adding the suitable deuterated solvent to the sample. (*see Note 7*).
3. If necessary filter your sample using glass wool (*see Note 8*).
4. Pipette the sample into an NMR tube (*see Notes 9 and 10*).
5. Wrap a 1 cm × 2 cm Parafilm[®] strip around the cap of the NMR tube (*see Note 11*).
6. Label your sample at the top of the NMR tube.

3.2 Useful Parameters from NMR Measurements

3.2.1 Chemical Shift

At this point, it is important to emphasize the most basic parameters measured in the NMR spectra, including the chemical shift, coupling constant, and integration.

1. The chemical shift (δ) is a measure of the resonant frequency of an NMR-active nucleus (e.g., ^1H , ^{13}C , or ^{15}N) and is quoted in parts per million (ppm) (*see Note 12*).
2. The chemical shift denotes the positions of the NMR peaks relative to a reference compound (usually tetramethylsilane, TMS) (*see Note 13*).
3. By convention the shielded signals of TMS are set to 0 ppm, situated on the right side of the chemical shift scale in the NMR spectrum. The resonances of common functional groups are less shielded, that is they have higher shifts, and are placed on the left side of the signal of TMS [11, 12].

3.2.2 Coupling Constant

1. Coupling constant is the absolute separation between two or more peaks (splitting) of each NMR signal, arising from coupling between nuclei, such as proton-proton or proton-carbon.
2. This intramolecular communication, caused by through-bond interactions of nuclei, is the phenomenon known as spin-spin, scalar, or J -coupling (*see Note 14*).
3. J -couplings are measured in cycles per second (Hz) and their magnitude depends on their distance apart, relative stereochemistry, and chemical environment.
4. The observation of J couplings is important, as useful structural information can be derived, including bond linkage (*see Note 15*) and molecular conformation (*see Note 16*).

3.2.3 Integration

1. In the ^1H NMR spectrum, integral values of peak areas underneath the NMR signals are proportional to the number of hydrogen atoms of related functional groups in the sample (*see Note 17*).
2. In sample mixtures the quantitative relationship between the corresponding components can be determined by the different ratio of the integrals in the spectrum.

3.3 1D NMR Methods

The 1D NMR spectrum is a plot showing amplitude along a frequency axis, which is typically the chemical shift axis. To obtain this spectrum, the nuclei are irradiated and generate a signal that is detected in the time domain and then converted mathematically into the frequency domain by employing a mathematical procedure known as Fourier transformation [11].

3.3.1 ^1H NMR

1. By measuring a ^1H NMR spectrum, we observe frequency ranges of ^1H resonances with common chemical shifts from 0 to 12 ppm. Thus, a typical spectral window for ^1H NMR is at least 12 ppm wide (*see Note 18*). An example of the ^1H NMR spectrum of callyaerin A in $\text{DMSO}-d_6$ is shown in Fig. 2. ^1H NMR analysis is performed as follows.
2. Standardize the reported chemical shifts with reference to the residual solvent peak.
3. Integrate the ^1H NMR spectrum to obtain a total hydrogen count (*see Note 19*).
4. List all ^1H NMR chemical shifts to two decimal places (*see Note 20*).
5. List the multiplicities and coupling constants (J in Hz) for all ^1H NMR signals (*see Note 21*).
6. Inspect the spectrum and identify obvious functionalities, such as aromatic protons ($\sim 7.0\text{--}8.0$ ppm), methoxyl ($\sim 3.5\text{--}4.0$ ppm) or methyl ($\sim 1.0\text{--}2.0$ ppm) groups, from their characteristic shifts, multiplicities, and integrations (*see Note 22*).

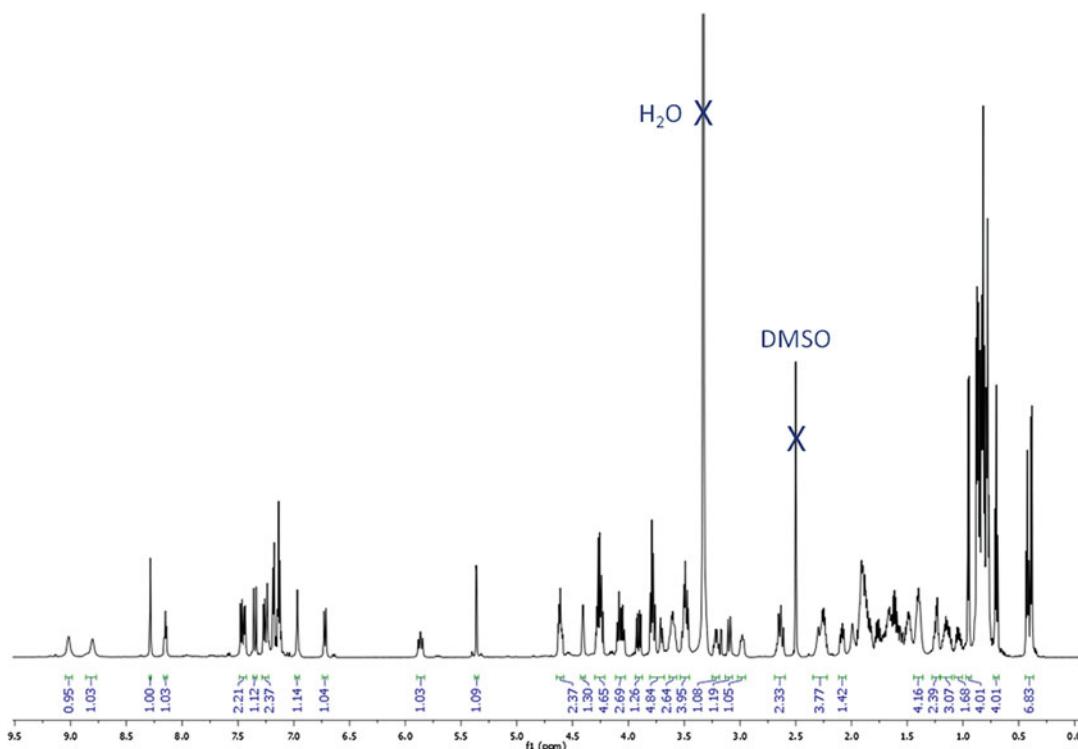


Fig. 2 ^1H NMR (600 MHz, $\text{DMSO}-d_6$) spectrum of callyaerin A

3.3.2 ^{13}C NMR

1. The ^{13}C NMR spectrum shows the chemical shifts of carbon resonances, and thus the total carbon atom number of a molecule. The frequency range for common ^{13}C shifts is from 0 to 220 ppm. Due to the low isotopic abundance (1.1%) of ^{13}C , as well as its inherent low sensitivity (~1/64 to that of ^1H), the signals are weaker than those of ^1H , and thus more time for spectra recording is acquired (*see Notes 23 and 24*). The ^{13}C NMR spectrum of callyaerin A is shown in Fig. 3. ^{13}C NMR analysis is performed as follows.
2. Reference the ^{13}C NMR spectrum to the residual deuterated solvent peak.
3. List all ^{13}C NMR chemical shifts to one decimal place (*see Note 25*).
4. As in the case of the ^1H spectrum, inspect the ^{13}C spectrum for obvious functionalities, such as carbonyl (~170–220 ppm), aromatic (~110–130 ppm), methoxy (~50–60 ppm), or methyl (~10–30 ppm) groups. [13].

3.3.3 DEPT

1. *Distortionless enhancement by polarization transfer* (DEPT) is the usual method for determining the type of carbon atoms present: tertiary (CH), secondary (CH_2), and primary (CH_3) carbons (*see Note 26*).

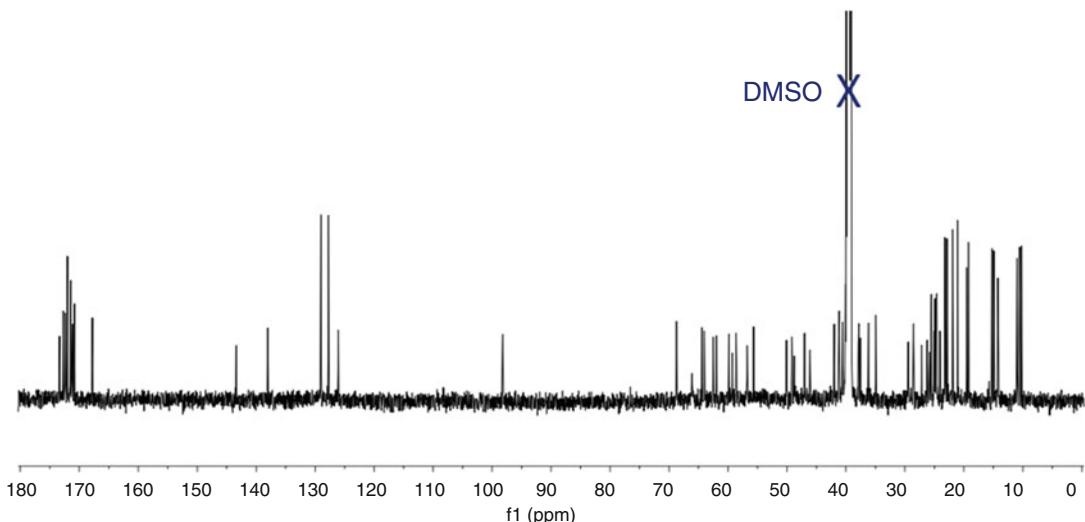


Fig. 3 ^{13}C NMR (150 MHz, DMSO- d_6) spectrum of callyaerin A

2. The DEPT-135 spectrum displays CH's and CH₃'s as positive singlet peaks and CH₂'s as negative singlet peaks, whereas the quaternary carbons are absent. The DEPT spectrum of callyaerin A is shown in Fig. 4. Interpretation of the DEPT spectrum is performed as follows.
3. Determine the carbon atom order according to the phase of the peaks and in the case of quaternary carbons by direct comparison with the ^{13}C NMR spectrum.
4. Count the number of hydrogen atoms (H's) that are directly attached to the respective carbon atoms (^{13}C 's).

3.4 2D-NMR Methods

1. A 2D NMR spectrum is obtained using multipulse experiments that correlate signals from two frequency domains (f_1 and f_2). Contour plots show cross peaks that associate information on one axis with information on the second axis [11].
2. These methods are valuable tools for structure determination of complex compounds, since either through-bond or through-space interactions are revealed between nuclei of the same (homonuclear—typically proton) or different (heteronuclear—typically proton and carbon) elements (see Note 27).
3. In addition, cross peaks observed in 2 D spectra allow the assessment of accurate chemical shift values and J -couplings that cannot be assigned directly from the 1 D spectrum due to signal overlap.

3.4.1 COSY

1. The homonuclear shift correlation spectroscopy (^1H , ^1H -COSY, or COSY) spectrum shows the through-bond coupling connectivities between groups containing hydrogen atoms,

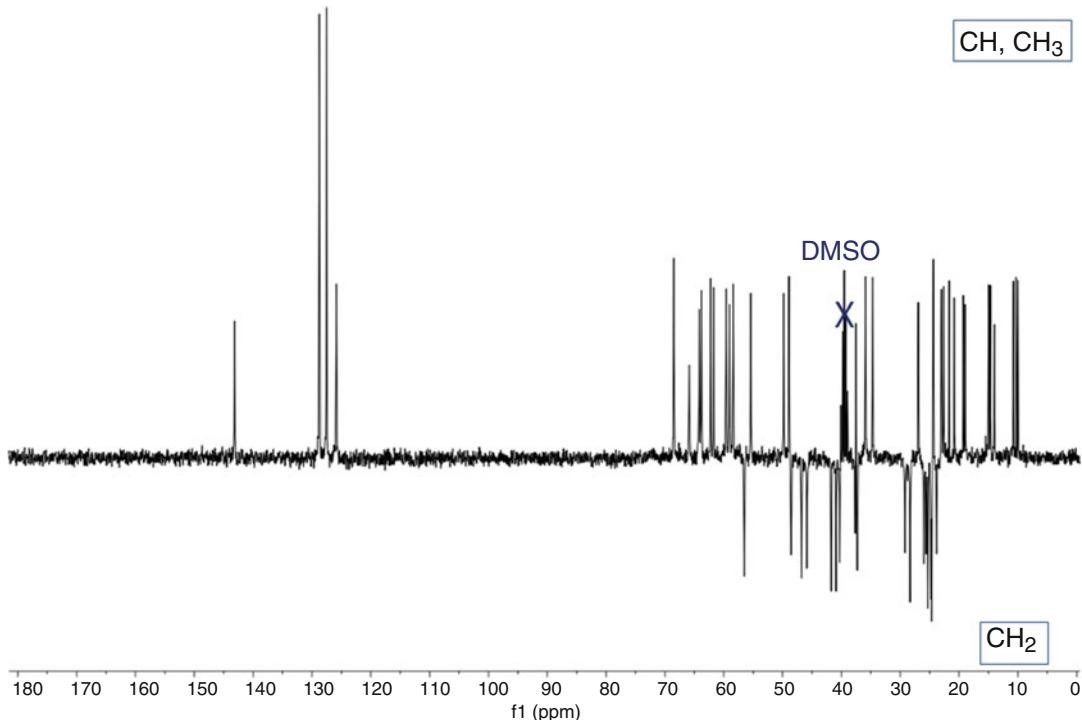


Fig. 4 DEPT-135 NMR (75 MHz, $\text{DMSO}-d_6$) spectrum of calyaerin A

based on geminal (2J) and vicinal (3J) proton couplings (see Notes 28 and 29).

2. The ^1H spectrum of the sample is set on both horizontal (f_2) and vertical (f_1) axes. Autocorrelated peaks appear on the diagonal ($\delta_1 = \delta_2$), which is the symmetrical axis of the COSY spectrum.
3. A signal situated off the diagonal is called a cross peak and appears whenever protons with resonances at δ_1 and δ_2 ($\delta_1 \neq \delta_2$) are coupled to one another.
4. A pair of coupled protons can be identified by lines through the cross peak, which is symmetrical with respect to the diagonal (see Note 30). An example of the COSY spectrum of calyaerin A is shown in Fig. 5. Analysis of the COSY spectrum is performed as follows.
5. Draw a vertical line from a known diagonal peak (H_A) until you connect with a cross peak (H_A, H_B). The horizontal line from this cross peak to the diagonal identifies the shift of the coupled proton (H_B).
6. In a similar manner, projecting from the last diagonal peak to the next cross peak and then back to the diagonal allows the assignment of the whole coupling network in the molecule.

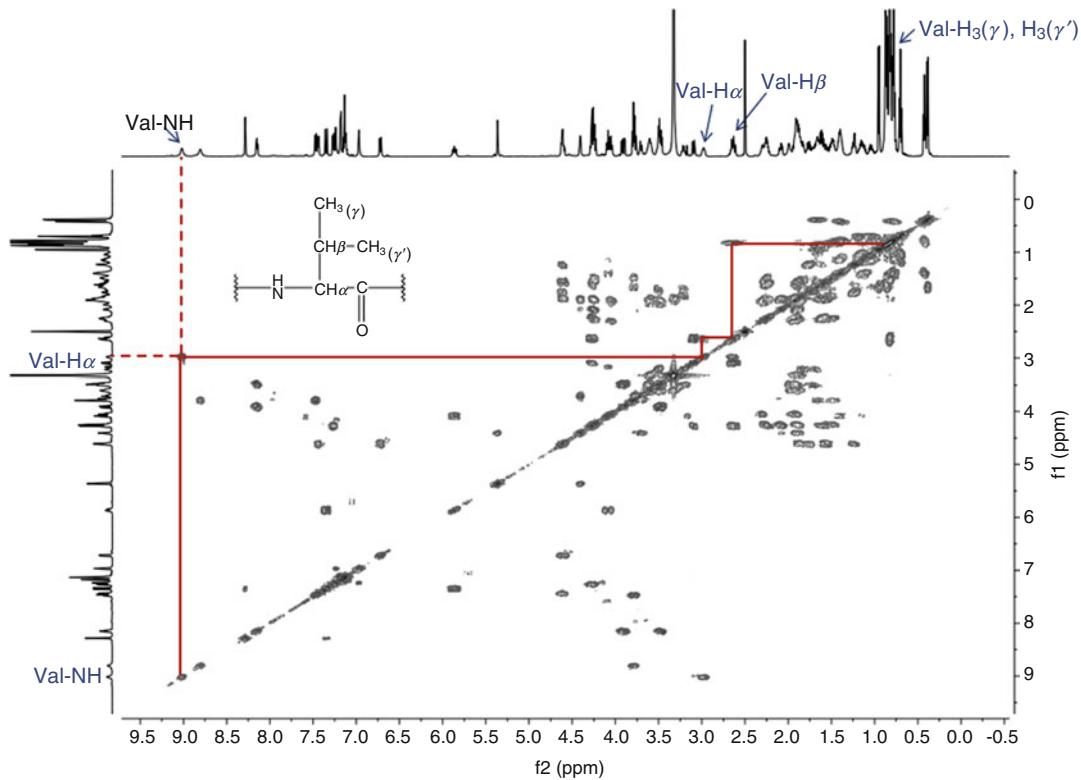


Fig. 5 COSY NMR spectrum of callyaerin A (600 MHz, $\text{DMSO}-d_6$)

The coupling network of the valine residue of callyaerin A is shown in Fig. 5.

7. Check the ^1H NMR spectrum to confirm that J couplings and integrals are in agreement with the assignments made by COSY.
- 3.4.2 TOCSY**
1. The *total correlation spectroscopy* (TOCSY) experiment allows the generation of cross peaks between virtually all protons within a given spin system (*see Notes 31 and 32*).
 2. The ^1H spectrum of the sample is set on both horizontal (f_2) and vertical (f_1) axes and cross peaks are situated off the diagonal line (*see Note 33*). The TOCSY spectrum of callyaerin A is shown in Fig. 6. Analysis of the TOCSY spectrum is performed as follows.
 3. Draw a vertical line through a selected peak at the top of the TOCSY spectrum. This line will pass through one or more cross-peaks, which are situated off the diagonal (*see Note 34*).
 4. From these cross-peaks project horizontal lines to the ^1H spectrum set along the vertical axis. The respective peaks allow the detection and assignment of all mutually coupled proton signals within a specific spin system (*see Notes 35*).

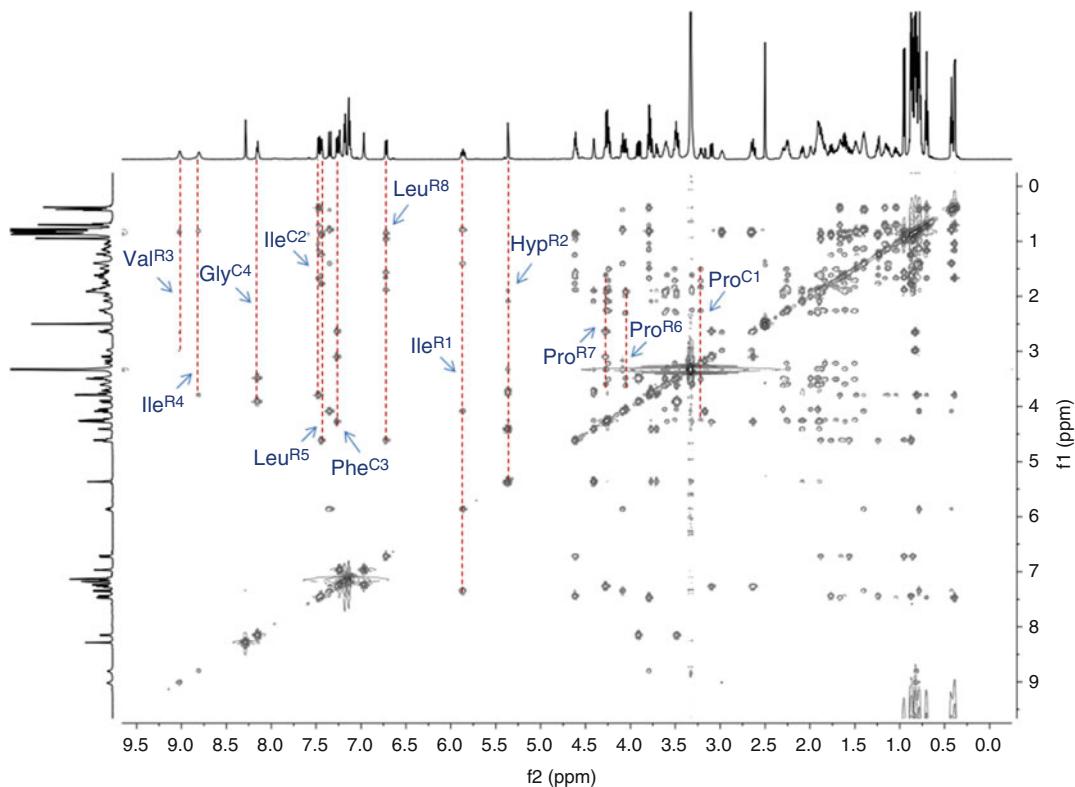


Fig. 6 TOCSY NMR spectrum of calyaerin A (600 MHz, $\text{DMSO}-d_6$)

3.4.3 NOESY/ROESY

1. The nuclear Overhauser effect spectroscopy (NOESY) and the related rotating frame Overhauser effect spectroscopy (ROESY) experiments show the relationships between protons via through-space rather than through-bond correlations (*see Notes 36 and 37*). An example of a ROESY spectrum of calyaerin A is given in Fig. 7.
2. The layout of the NOESY spectrum is analogous to that of a COSY spectrum with the ^1H NMR located along each axis (f_1 and f_2) and the diagonal (*see Note 38*). NOESY/ROESY interpretation is performed as follows.
3. Draw a horizontal and vertical line through a cross peak in the NOESY/ROESY spectrum. The lines will pass through a proton on the f_1 axis and a proton on the f_2 axis that are in close proximity ($<5 \text{ \AA}$) (*see Notes 39 and 40*).
4. Compare the NOESY/ROESY spectrum to the respective regular COSY spectrum and list all the key correlations (*see Note 41*).

3.4.4 ^1H - ^{13}C HMQC/HSQC

1. ^1H - ^{13}C HMQC (or HMQC) is an abbreviation for ^1H detected heteronuclear multiple quantum coherence and ^1H - ^{13}C HSQC (or HSQC) for ^1H detected heteronuclear single quantum

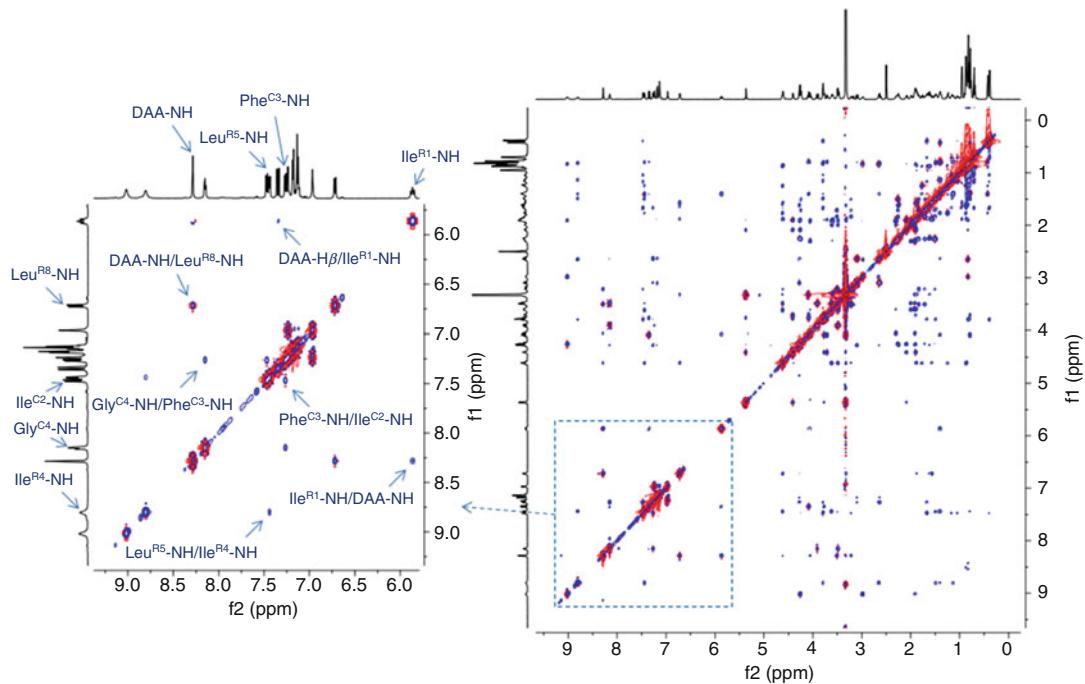


Fig. 7 ROESY NMR spectrum of callyaerin A (600 MHz, $\text{DMSO}-d_6$). Key correlations are shown in the magnified region of the spectrum

coherence. The HMQC/HSQC spectra detect one-bond couplings (ca. 140 Hz) between the protons of a molecule and the carbons to which they are directly attached (*see Note 42*).

2. The ^1H spectrum of the sample is commonly set along the horizontal (f_2) axis, whereas the ^{13}C spectrum is set along the vertical (f_1) axis. The HSQC spectrum of callyaerin A is shown in Fig. 8. HMQC/HSQC interpretation is performed as follows.
 3. Draw vertical and horizontal lines through a selected cross peak in the HSQC/HMQC spectrum. The lines will pass through a peak in the ^1H NMR spectrum and a peak in the ^{13}C NMR spectrum that indicates direct attachment of the respective nuclei (*see Notes 43 and 44*).
 4. Assign all correlations between the ^{13}C spectrum and the ^1H spectrum.
 5. Identify diastereotopic (nonequivalent) protons of methylene groups (*see Note 45*).
 6. Determine the connections of carbon atoms and assemble substructures using a combination of the HMQC/HSQC and COSY data.
- 3.4.5 ^1H - ^{13}C HMBC**
1. The ^1H detected heteronuclear multiple bond correlation (^1H - ^{13}C HMBC or HMBC) experiment shows long-range correlations between hydrogen and carbon (couplings of ca. 2–8 Hz),

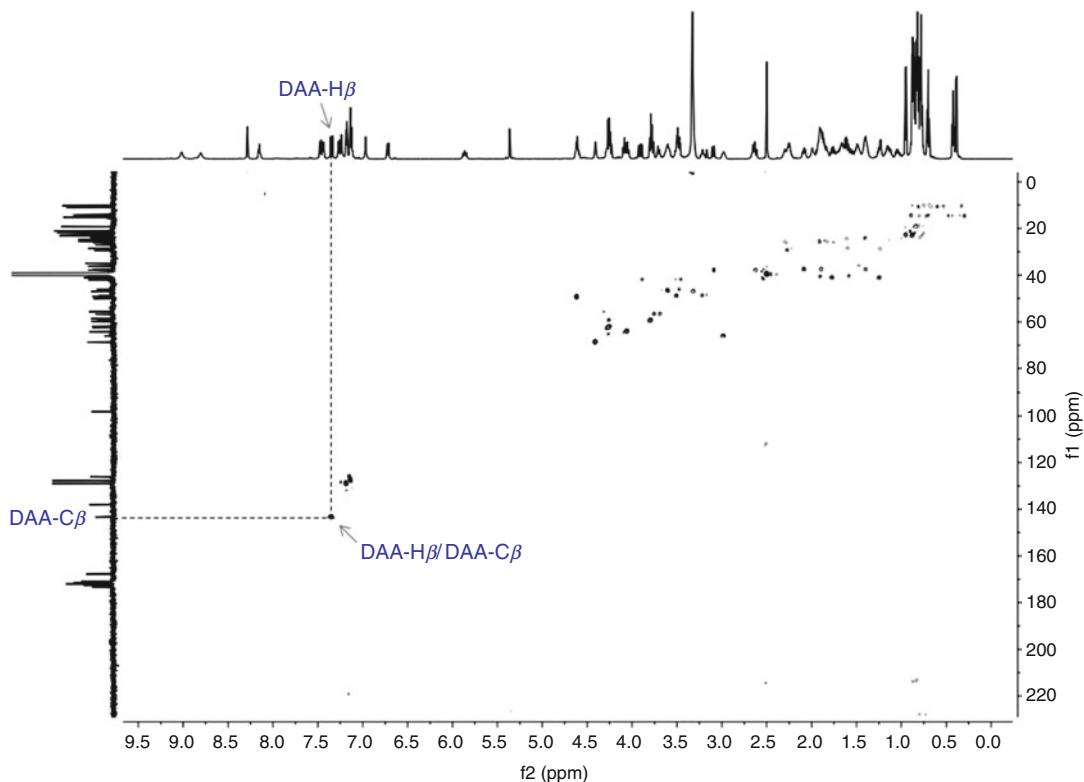


Fig. 8 HSQC NMR spectrum of callyaerin A (600 MHz, DMSO- d_6)

typically separated by two or three bonds (*see Notes 46* and *47*). The HMBC spectrum of callyaerin A is shown in Fig. 9.

2. As with the HSQC/HMQC spectra, the HMBC spectrum contains cross-peaks corresponding to proton peaks along the horizontal (f_2) axis and sets of carbon peaks along the vertical (f_1) axis (*see Notes 48* and *49*). HMBC spectrum analysis is performed as follows.
3. Draw a vertical line through a selected peak in the ^1H NMR spectrum (f_2 axis). This line will pass through one or more related cross-peaks.
4. From the respective cross-peaks project horizontal lines to the ^{13}C spectrum (f_1 axis). The lines allow the assignment of all ^{13}C peaks that are coupled through long-range couplings to the selected hydrogen atom (*see Note 50*).
5. Combine substructures into all feasible structures through long-range couplings and check consistency with the previous 1D and 2D NMR data (*see Note 51*).
6. If the structure contains chiral centers, its relative configuration may be assigned by performing NOE studies and/or analysis of its coupling constants (*see Notes 16* and *37*).

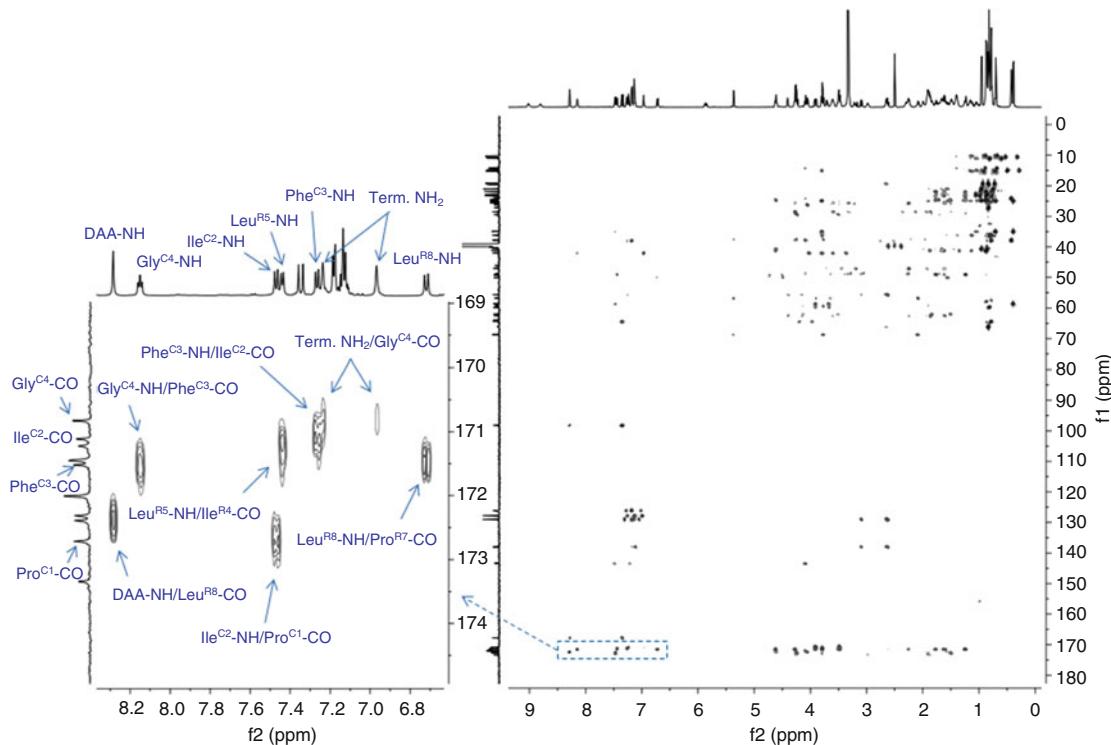


Fig. 9 HMBC NMR spectrum of callyaerin A (600 MHz, $\text{DMSO}-d_6$). Key correlations are shown in the magnified region of the spectrum

3.5 Structure Elucidation of Callyaerin A

1. The ^1H NMR spectrum of callyaerin A in $\text{DMSO}-d_6$ displays a pattern of chemical shifts typical for peptides, including a series of amide (NH) (δ_{H} 9.0–5.9) and α -amino proton ($\text{H}\alpha$) (δ_{H} 4.6–3.0) signals (see Note 52). Moreover, the presence of aromatic signals characteristic of a monosubstituted benzene ring ($\delta_{\text{H}} \sim 7$), in addition to several aliphatic methylene and methine signals (δ_{H} 1.0–3.0) and a cluster of doublet or triplet methyl signals (δ_{H} 0.4–1.0) indicates the presence of hydrophobic amino acid residues (Fig. 2).
2. The ^{13}C NMR spectrum of callyaerin A exhibits 11 amide carbonyl signals resonating between 167 and 173 ppm and 12 α -carbon signals of amino acids between 42 and 66 ppm, which further corroborated the peptidic nature of callyaerin A (Fig. 3). Additional signals include those of four aromatic signals (δ_{C} 126–138) confirming the presence of a monosubstituted benzene ring, whereas the remaining carbons are aliphatic methylenes, methines, and methyls, as supported by the DEPT-135 experiment (Fig. 4).
3. Interpretation of the COSY and TOCSY spectra of callyaerin A discloses the presence of 12 spin systems, eight of which originate from the amidic protons of one Gly, one Phe, three Ile, one Val, and two Leu residues (Fig. 5 and 6).

4. The remaining four spin systems show no correlations to any amide proton and are indicative of three proline residues, in addition to the unusual amino acid 4-hydroxyproline (Hyp) (Fig. 6).
5. The signals of one amide (δ_H 8.29; DAA-NH) and one olefinic (δ_H 7.35; DAA-H β) proton (Fig. 8) are attributed to an unusual 2,3-diaminoacrylic acid (DAA) moiety, as indicated by their HMBC correlations to DAA-C α (δ_c 98.2) and DAA-CO (δ_c 167.7) (Fig. 10).
6. Finally, two amide protons at 6.97 and 7.24 ppm show a strong COSY correlation and are assigned to a primary amide group (Fig. 5).
7. The sequential assembly of connectivities between the different amino acid residues is established by a detailed examination of the ROESY and HMBC spectra.
8. In the ROESY spectrum diagnostic correlations are observed between amide protons (NH's) of adjacent residues Ile^{C2}/Phe^{C3}/Gly^{C4} (Fig. 7), as well as between α -protons (H α) and NH's for Pro^{C1}/Ile^{C2}/Phe^{C3}/Gly^{C4}/terminal-NH₂, corresponding to the peptide fragment Pro^{C1}-Ile^{C2}-Phe^{C3}-Gly^{C4}-NH₂ (C1-C4) (Fig. 10). The through-bond heteronuclear correlations observed in the HMBC spectrum of NH signals for terminal-NH₂, Gly^{C4}, Phe^{C3}, and Ile^{C2} to the adjacent (²J) carbonyl carbons of Gly^{C4}, Phe^{C3}, Ile^{C2}, and Pro^{C1} (Fig. 9), respectively, further support the assignments made from the ROESY spectrum.
9. The observed ROESY correlations of Leu^{R5}-NH/Ile^{R4}-NH and Ile^{R4}-NH/Val^{R3}-H α suggest the partial sequence Val^{R3}-Ile^{R4}-Leu^{R5} (Figs. 7 and 10). The positions of the residues Hyp^{R2} and Pro^{R6} are apparent from the ROESY correlations

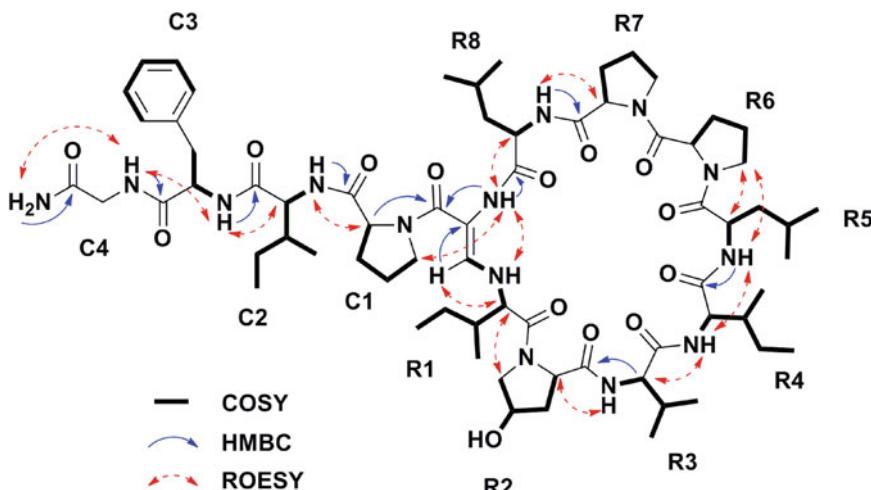


Fig. 10 COSY (***bold***), key HMBC (***plain***), and key ROESY (***dashed***) correlations of callyaerin A

observed between their δ protons with the α protons of the preceding amino acids Ile^{R1} and Leu^{R5}, respectively. Moreover, the correlation between Leu^{R8}-NH and Pro^{R7}-H α confirms the assignments of the remaining residues, leading to the substructure Ile^{R1}-Hyp^{R2}-Val^{R3}-Ile^{R4}-Leu^{R5}-Pro^{R6}-Pro^{R7}-Leu^{R8} (Fig. 10). The interpretation of the ROESY spectrum is further corroborated by the HMBC correlations of Leu^{R8}-NH/Pro^{R7}-CO, Leu^{R5}-NH/Ile^{R4}-CO, and Val^{R3}-H α /Hyp^{R2}-CO (Fig. 9).

10. The nature of the ring closure is supported by the ROESY correlations between DAA-H β /Ile^{R1}-NH, Ile^{R1}-H α and DAA-NH/Leu^{R8}-NH, Leu^{R8}-H α (Figs. 7 and 10), indicating the sequence of the cyclic part of calyaerin A as *cyclo*(DAA-Ile-Hyp^{R2}-Val^{R3}-Ile^{R4}-Leu^{R5}-Pro^{R6}-Pro^{R7}-Leu^{R8}).
11. Similarly, the linkage of the cyclic part (R1–R8) with the linear chain (C1–C4) is established through the ROESY correlation between DAA-NH and Pro^{C1}-H δ , in association with the HMBC correlation between Pro^{C1}-H α and DAA-CO (Fig. 10).
12. Finally, the Z-configuration of the DAA unit for calyaerin A is suggested by the ROESY correlation between DAA-NH and R1-NH (Fig. 10), thus leading to the overall structure of calyaerin A (*see Note 53*).

4 Notes

1. Probes for liquid NMR spectroscopy are typically designed to accommodate NMR tubes with outside diameters of 2–10 mm (and larger), with 5 mm tubes being the most common ones [11].
2. The NMR tubes should be thoroughly rinsed before use, preferably with spectroscopic grade solvents. For this purpose, the use of nonvolatile solvents, such as dimethylsulfoxide (DMSO) or dimethylformamide (DMF), which are difficult to remove completely, should be avoided [12].
3. Deuteration is used in order to suppress the signals of the solvent. However, the deuterium enrichment for each of these solvents is less than 100%, and thus they will show observable peaks in the ^1H NMR spectrum, which are detected as singlets or multiplets [13] (Fig. 2).
4. For natural products, it is preferable to use fresh solvent from an individual ampoule, rather than using solvent from a bottle. Also note that deuterated chloroform (CDCl_3) over a 6-month period may be sufficiently acidic to cause chemical modifications of acid labile compounds [11].
5. Care should be taken to avoid contamination of the purified compound. Two common examples of impurities are the

phthalates from plastic tubing and grease used to lubricate ground glass joints. Paramagnetic impurities—often originating from a spatula—should be also avoided, since these cause extensive line broadening [13].

6. The sample to be measured must be as dry as possible to prevent the residual water peak from overwhelming the solute signal. Freeze-drying is the method of choice due to the lack of elevated temperature influence on thermolabile compounds.
7. The solvent should completely dissolve the sample. Common deuterated NMR solvents are CDCl_3 , $\text{DMSO}-d_6$, $\text{MeOH}-d_4$, and benzene- d_6 , or pyridine- d_5 .
8. The solution should be clear as particles cause line broadening. These can be removed by using a Pasteur pipette packed with a small amount of glass wool to filter the sample. Addition of a small amount of solvent to a saturated solution may also minimize the line broadening caused by microscopic nucleation of particles [11].
9. Signal resolution results are optimal when there is a ca. 3 cm liquid height in the NMR tube (i.e., a volume between 0.5 and 0.7 mL in a 5 mm diameter NMR tube). Lower than optimal solution volumes cause degraded line shapes in the spectrum.
10. For natural products, where only a few mg of sample may be available, minimum solvent volume is generally preferred, as this will maximize concentration, and thus the intensity of detected signals. However, care should be taken so that the high sample concentration does not cause broadening in the NMR signals due to high viscosity or turbidity.
11. Parafilm® will limit evaporation of volatile solvents.
12. The chemical shift is characteristic of the different types of chemical groups and their chemical environment. For instance, protons next to electronegative atoms, such as oxygen or nitrogen, will have less electron density and as a result will be deshielded from the full effect of the applied field resonating at a higher frequency for a given field strength [11].
13. In the modern NMR spectrometers the standard shifts of the deuterated solvents are mostly used as internal chemical shift references, which remove the necessity of adding TMS [13].
14. J coupling is independent of the strength of the magnetic field. However, higher applied magnetic field is preferable for the analysis of structurally complex molecules with highly overlapped signals, because the latter will proportionally occupy less space on the ppm scale than they do with a lower magnetic field [11, 14].
15. J coupling is a reciprocal process and the same coupling constant is found between interacting NMR nuclei. Therefore,

according to the splitting pattern and the magnitude of the coupling constants the connection between interacting groups can be determined.

16. The dihedral angle between coupled nuclei separated by three bonds provides important information about local molecular conformation [11, 15]. For instance, in cyclohexane rings, the coupling constant of two (vicinal) protons with an axial-axial relationship ($\Phi=180^\circ$, $J=8\text{--}15$ Hz) is greater than that from two (vicinal) protons with an axial-equatorial relationship ($\Phi=60^\circ$, $J=2\text{--}5$ Hz). Likewise, in alkenes the *Z* or *E* configuration can be differentiated according to the observed J value of the respective protons (8–12 or 15–18 Hz, respectively) [13].
17. This implies that the protons of a methyl group (CH_3) produce three times the signal of a methine proton (CH). Keep in mind that the integration provides only relative intensity data, so that the experimentalist must select a resonance with a known or suspected number of protons and normalize the other integrals accordingly [14].
18. The first overview of the ^1H NMR spectrum should determine whether there are aromatics, alkenes, or methyl functionalities, the latter integrating for three protons.
19. Impurities in the sample can affect the measured integrals, especially if the signal to noise ratio of the spectrum is low. The presence of paramagnetic species, such as metal ions, or changes in the temperature and pH may also affect the intensity of the various resonances in the spectrum.
20. It is a good approach to tabulate the chemical shifts, multiplicities, and integrations of the observed resonances, so that vital information will not be overlooked. Please keep in mind that the ^1H chemical shifts are susceptible to solvent effects. When reporting ^1H chemical shifts the solvent used should always be stated.
21. In the ^1H NMR spectrum the multiplicity of the peaks is determined by the $n+1$ rule, where n is the number of the hydrogen atoms that participate in coupling (first-order spectra). Thus, the methyl group in ethanol is split from the methylene group (CH_2) into $2+1=3$ lines (triplet). In addition, the relative intensities of the peaks follow the properties of Pascal's triangle, which explains the 1:2:1 intensity of the aforementioned methyl group in the ^1H NMR spectrum. However, if we suppose that the two hydrogen atoms of the methylene group were not magnetically equivalent, that is, they have different coupling constants to the methyl group due to restricted rotation or close proximity to a stereogenic center, then the $n+1$ rule should be considered twice. As a result, the methyl group would show a peak pattern of $(1+1)(1+1)$, that is, 2×2 or doublet of doublet (dd). At this point, please keep in mind that in cases when the

chemical differences between the interacting protons ($\Delta\nu$) are not large enough compared to their coupling constants (J), i.e., $\Delta\nu/J < 5$, the lines and intensities of the peaks do not follow the aforementioned rules and their splitting pattern is more complicated (second-order spectra) [12].

22. The chemical shifts and signal dispersion are influenced by various physical conditions. Thus, when peaks coincide in the NMR spectrum, you can enhance the resolution by changing the solvent, the temperature, the pH, or even the magnetic field strength. The latter option produces more predictable results.
23. Keep in mind that the major isotope ^{12}C has no magnetic properties. On the other hand, ^{13}C has an odd mass and is NMR active. Other common NMR active nuclei are ^1H , ^{15}N , and ^{31}P .
24. The ^{13}C resonances display multiplicity due to the carbon-proton couplings, and thus signals from methyl groups (CH_3) are split into four lines (quartet), signals from methylene groups (CH_2) are split into three lines (triplet), signals from methine groups (CH) are split into two lines (doublet), and signals from quaternary carbons (C) appear as singlets. In practice, during acquisition of a $^{13}\text{CNMR}$ spectrum, the carbon-proton couplings are usually eliminated by irradiating of ^1H across a broad range of frequencies (broad-band ^1H decoupling). This technique simplifies the ^{13}C spectrum, since one singlet arises for each distinct type of carbon. However, the integration is less reliable, because the decoupling field perturbs the intensities of the peaks. Please also keep in mind that due to the low natural abundance of ^{13}C , the probability of having two adjacent ^{13}C nuclei in a single molecule is highly unlikely (0.01%), which removes complications from ^{13}C – ^{13}C couplings [12, 13].
25. When you attempt to measure ^{13}C chemical shifts, you may encounter a poor signal to noise ratio in the $^{13}\text{CNMR}$ spectrum. Moreover, quaternary carbons often exhibit low signal intensity and it is difficult to assign them. In this case, you may resort to the use of indirect detection methods, such as HSQC or HMBC to obtain the chemical shifts of missing ^{13}C atoms.
26. Different versions of this experiment have been developed for distinguishing the types of carbons within a molecule; however, the most common one is DEPT-135 [16].
27. Keep in mind that these techniques can also be applied to other common nuclei in organic molecules, such as nitrogen. Unfortunately, the very low intrinsic sensitivity and natural abundance of the NMR-active isotope ^{15}N (0.37%) has precluded the routine acquisition of ^{15}N spectra. Nevertheless, it can provide valuable structural information, especially in the case of highly substituted heterocyclic ring structures [17].

28. The intensity of COSY cross peaks varies in direct proportion to the magnitude of the J -coupling(s) between the correlated peaks. Please note that a coupling of zero obscures the passage of magnetization resulting in no cross-peaks. Therefore, according to the Karplus diagram, in aliphatic systems when the dihedral angle is 90° the vicinal correlation (3J) may not be observed [15].
29. Exceptions, however, can be expected for some aromatic, olefinic, and special configuration systems, since long-range couplings (4J or even 5J) with observable coupling constants can be found in a COSY spectrum.
30. This duplication can be useful as it enables us to distinguish true correlations from artifacts, since the latter are rarely symmetrical.
31. A spin system is a group of nuclei that all couple to one or more other members of the group.
32. The intensity of the cross peaks is dependent on the coupling between the protons for transmission along the pathway. A small coupling may affect the further transmission along the spin system.
33. The selective 1D TOCSY spectrum is the 1D variant of the TOCSY experiment. In this experiment, a well-resolved signal is selected for irradiation and the resulting spectrum will show only those protons that are in the same spin system as the selected proton. In addition, the coupling patterns of the ^1H NMR spectrum are visible in the subspectrum, which permits rapid assignment of all residues [12, 13].
34. In contrast to COSY, the TOCSY spectrum may assist to resolve cross-peaks obscured by the diagonal or by a very crowded region in the spectrum.
35. The TOCSY experiment is a very powerful tool for unraveling the spectra of complicated modular structures, such as saccharides and peptides. For instance, in peptides, the amide bond connecting amino acid residues is an effective barrier to J -couplings and thus, each amino acid residue constitutes a separate spin system. The process of identifying resonances within a specific amino acid is considerably simplified by this procedure and the majority of the spin systems can be delineated.
36. For intermediate molecular weight molecules (~1000 amu) only weak correlations are visible by using a standard NOESY pulse sequence. In this case, the ROESY experiment is preferred as all NOEs are displayed, affording useful NOE data. However, TOCSY artifacts may be present in the ROESY spectrum, which appear as antiphase (opposite sign) peaks [13].
37. The NOE is highly distance dependent and therefore it is a very powerful tool for drawing stereochemical assignments,

e.g., information about the relative stereochemistry around double bonds or ring junctions of cyclic compounds, and for resolving problems of positional isomerism.

38. As with TOCSY, the 1D NOE version may be applied, in which a single resonance is irradiated and its spatial coupling patterns are identified. The phenomenon that is observed is called the NOE effect. Keep in mind that the signal being irradiated should be well separated from other signals in order to avoid potential partial saturations of nearby protons. Moreover, in the case of small molecules (<600 amu) the irradiated peak phases negatively, which is opposite to the enhanced signals (positive). The 1D NOE spectrum contains fewer artifacts compared to the 2D technique. However, the 2D NOE spectrum offers the advantage of displaying all pairs of hydrogen atoms having the NOE effect simultaneously, which is the preferred choice for molecules with many stereogenic centers [11, 13].
39. The NOE magnitude is proportional to the inverse of the sixth power of the distance between the two nuclei. Therefore, the NOE effect requires close proximity between atoms in space; with distances for strong correlations of 2–3 Å and for weak correlations of 4–5 Å [13].
40. Please note that cross-peaks with solvent can be caused by chemical exchange.
41. The COSY correlated peaks can be shown in the NOESY spectrum. However, the COSY peaks have opposite phases to the genuine NOESY correlations, and this fact can be used to differentiate between them.
42. The interpretation methods for HMQC and HSQC are the same. However, the spectrum of HSQC enjoys the benefits of fewer artifacts and slightly better resolution in the ^{13}C domain.
43. If the chemical shift of a specific proton is known, the chemical shift of the coupled carbon can be determined, and vice versa, which is important for the postulation of the chemical environment of the functional group.
44. If a ^1JCH correlation is not observed then it is evident that the remaining ^{13}C NMR chemical shifts are quaternary carbons.
45. The COSY spectrum cannot differentiate between geminal and vicinal proton connectivities. However, in the HSQC spectrum, if two nonequivalent protons (diastereotopic) are attached to a common carbon, they will both correlate to the same ^{13}C peak. This information is very useful and can be used in combination with the DEPT spectrum to readily identify geminal pairs of protons. In addition, as the resolution of the HMQC (or HSQC) spectrum is higher than that of the ^1H NMR spectrum, interpretation of the HMQC/HSQC can

help to resolve the overlapped correlated peaks in the ^1H and COSY spectra [12].

46. In practice, the HMBC sequence is optimized for three-bond correlations, even though it is not obvious by inspection which ones are two- and which ones are three-bond correlations. In addition, carbons separated by more than three bonds can be detected, especially in the case of aromatic and olefinic (allylic) systems, but the intensity of the signals will not be as high.
47. In aliphatic systems three-bond couplings show a Karplus-type dihedral angle dependence and can be used for stereochemical assignments [15].
48. In some cases, one bond correlations can also be observed in the HMBC spectrum. They are characterized by a pair of signals/contours that are situated on a horizontal line that passes through the correlated peak in the ^{13}C spectrum. The signals are symmetrically displayed on either side of the proton peak that they relate to. Pay particular attention when these signals line up exactly with peaks in the ^1H NMR spectrum along the horizontal (f_2) axis, giving rise to artifacts that appear as potential cross-peaks. Therefore, it is always very useful to identify the single bond correlations before you begin interpreting the HMBC spectrum.
49. Keep in mind that the correlation peaks of broad proton resonances are usually weak or even unobservable.
50. In an analogous way, a horizontal line through a selected peak in the ^{13}C NMR spectrum (f_1 axis) allows the assignment of all ^1H peaks that are coupled through long range couplings to the selected carbon atom.
51. Note that connectivities can occur between protons and carbons separated by heteroatoms (e.g., oxygen, nitrogen or sulfur) or by quaternary carbons. Therefore, the HMBC experiment is a powerful tool for the connection of structural units within a molecule.
52. The resonances from exchangeable protons, that is, protons that exist in dynamic equilibrium with each other or with protons from the medium, such as -OH or -NH, are strongly influenced by the pH and hydrogen-bonding properties of the NMR solvent. Exchangeable protons may be readily identified by recording the ^1H spectrum before and after adding a drop of D_2O . In the latter case, the protons rapidly exchange with deuterium in D_2O , and therefore they will disappear from the spectrum. Keep in mind that in highly purified aprotic solvents, such as DMSO, the intermolecular exchange is lowered and coupling between the hydroxy and adjacent protons can be detected. The same applies if the recording temperature or

the amount of basic or acidic catalysts (such as trifluoroacetic acid—TFA) is decreased [11, 16].

53. Combination of NMR with various techniques aimed at structure determination is emphasized for an unambiguous assignment of a compound. For instance, in the case of peptides, the absolute configuration of the constituent amino acids is commonly determined by derivatization employing the Marfey's method [18]. Accordingly, acid hydrolysis of callyaerin A followed by liquid chromatography-mass spectrometry analysis of the Marfey derivatives and by comparison with standard amino acids leads to the assignment of the l-configuration for all amino acid residues [6, 7].

Acknowledgment

This work was supported by the German Federal Ministry of Education and Research (BMBF; Bundesministerium für Bildung und Forschung).

References

1. Lewis K (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* 12:371–387
2. Blair JMA, Webber MA, Baylay AJ et al (2015) Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* 13:42–51
3. Clardy J, Fischbach MA, Walsh CT (2006) New antibiotics from bacterial natural products. *Nat Biotechnol* 12:1541–1550
4. Harvey AL, Edrada-Ebel R, Quinn RJ (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov* 14:111–129
5. Koehn FE, Carter GT (2005) The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* 4:206–220
6. Ibrahim SRM, Min CC, Teuscher F et al (2010) Callyaerins A-F and H, new cytotoxic cyclic peptides from the Indonesian marine sponge *Callyspongia aerizusa*. *Bioorg Med Chem* 18:4947–4956
7. Daletos G, Kalscheuer R, Koliwer-Brandl H et al (2015) Callyaerins from the marine sponge *Callyspongia aerizusa*: cyclic peptides with antitubercular activity. *J Nat Prod* 78:1910–1925
8. Bara R, Zerfass I, Aly AH et al (2013) Atropisomeric dihydroanthracenones as inhibitors of multiresistant *Staphylococcus aureus*. *J Med Chem* 56:3257–3272
9. Molinski TF (2010) NMR of natural products at the ‘nanomole-scale’. *Nat Prod Rep* 27:321–329
10. Breton RC, Reynolds WF (2013) Using NMR to identify and characterize natural products. *Nat Prod Rep* 30:501–524
11. Simpson JH (2008) Organic structure determination using 2-D NMR spectroscopy: a problem-based approach. Elsevier, Acad Press, Amsterdam [et al]
12. Ning Y-C (2011) Interpretation of organic spectra. Wiley, Hoboken, New Jersey
13. Crews P, Rodríguez J, Jaspars M (2010) Organic structure analysis, international, 2nd edn. Oxford University Press, New York
14. Lambert JB, Gronert S, Shurvell HF, Lightner DA (2011) Organic structural spectroscopy, 2nd edn. Pearson, Prentice Hall, Boston [et al]
15. Matsumori N, Kaneno D, Murata M et al (1999) Stereochemical determination of acyclic structures based on carbon-proton spin-coupling constants. A method of configuration analysis for natural products. *J Org Chem* 64:866–876
16. Richards SA, Hollerton JC (2011) Essential practical NMR for organic chemistry. Wiley, Chichester
17. Pham C-D, Weber H, Hartmann R et al (2013) New cytotoxic 1,2,4-thiadiazole alkaloids from the ascidian *Polycarpa aurata*. *Org Lett* 15:2230–2233
18. Marfey P (1984) Determination of D-amino acids. II. Use of a bifunctional reagent, 1,5-difluoro-2,4-dinitrobenzene. *Carlsberg Res Commun* 49:591–596

Chapter 5

Computer-Aided Drug Design Methods

Wenbo Yu and Alexander D. MacKerell Jr.

Abstract

Computational approaches are useful tools to interpret and guide experiments to expedite the antibiotic drug design process. Structure-based drug design (SBDD) and ligand-based drug design (LBDD) are the two general types of computer-aided drug design (CADD) approaches in existence. SBDD methods analyze macromolecular target 3-dimensional structural information, typically of proteins or RNA, to identify key sites and interactions that are important for their respective biological functions. Such information can then be utilized to design antibiotic drugs that can compete with essential interactions involving the target and thus interrupt the biological pathways essential for survival of the microorganism(s). LBDD methods focus on known antibiotic ligands for a target to establish a relationship between their physiochemical properties and antibiotic activities, referred to as a structure-activity relationship (SAR), information that can be used for optimization of known drugs or guide the design of new drugs with improved activity. In this chapter, standard CADD protocols for both SBDD and LBDD will be presented with a special focus on methodologies and targets routinely studied in our laboratory for antibiotic drug discoveries.

Key words Computer-aided drug design, Molecular dynamics, Virtual screening, Docking, Site identification by ligand competitive saturation, SILCS, Structure-activity relationship, Pharmacophore, Force field

1 Introduction

Despite the fact that numerous antibiotic drugs are available and have been routinely used for a much longer time than most other drugs, the fight between humans and the surrounding bacteria responsible for infections is ongoing and will be so for the foreseeable future. Contributing to this is the steady rise of antibiotics drug resistance leading to the need for new antibiotics [1, 2]. Toward the design of new antibiotics, computer-aided drug design (CADD) can be combined with wet-lab techniques to elucidate the mechanism of drug resistance, to search for new antibiotic targets, and to design novel antibiotics for both known and new targets. Notably, CADD methods can produce an atomic level structure-activity relationship (SAR) used to facilitate the drug design process, thereby minimizing time and costs [3, 4].

Understanding the atomic-detailed mechanism behind the antibiotics resistance helps to reveal limitations in current antibiotics and shed light on the design of new drugs. For example, Trylska et al. studied the effects of mutations at the bacterial ribosomal A-site using molecular dynamics (MD) simulations to reveal the origins of bacterial resistance to aminoglycosidic antibiotics [5]. Our lab studied the impact of ribosomal modification on the binding of the antibiotic telithromycin using a combined Grand Canonical Monte Carlo (GCMC)/Molecular Dynamics (MD) simulation methodology [6, 7] and revealed atom-level details of how those modifications lead to resistance that will be of utility to improve the activity and spectrum of macrolide analogs thereby minimizing resistance [8].

An important alternative to solve the antibiotic resistance issue is the identification of new antibiotic targets that may represent novel mechanisms essential for bacterial survival. For example, researchers used bioinformatics approaches to screen various databases computationally and identified seven enzymes involved in bacterial metabolic pathways as well as 15 nonhomologous proteins located on membranes in the gram positive bacterium *Staphylococcus aureus* (SA), thereby indicating them as potential targets [9]. Such findings may help to overcome the resistance of this bacterium to common antibiotics such as methicillin, fluoroquinolones, and oxazolidinones. An example of a recently identified novel antibiotic target is the protein heme oxygenase, involved in the metabolism of heme by bacteria as required to access iron [10–12]. In collaborative studies with the Wilks lab, we have successfully applied CADD techniques to identify inhibitors of the bacterial heme oxygenases from *Pseudomonas aeruginosa* and *Neisseria meningitidis*, thereby confirming the potential role of heme oxygenases as novel antimicrobial targets [13, 14].

Researchers are also continuing to look for new antibiotics against existing targets and computational approaches have been successfully used in a number of studies. Using in silico database screening, Chang et al. found a new series of non-β-lactam antibiotics, the oxadiazoles, which can inhibit penicillin-binding protein 2a (PBP2a) of methicillin-resistant SA (MRSA), the cause of most infections in hospitals [15]. Using ligand-based drug design (LBDD), our lab with Andrade and coworkers investigated analogs of the third-generation ketolide antibiotic telithromycin as a possible means to address the bacterial resistance problem associated with that class of antibiotics [16–18]. In another study, based on the 3D structure of the complex of human defensin peptide HNP1 with Lipid II, which serves as a precursor for bacterial cell wall biosynthesis and is a validated target for antibiotics, our lab designed a simple pharmacophore model and used it in a database screen to search for low weight defensin mimetics [19]. From that effort, a lead compound was identified that targets Lipid II with high specificity and affinity. Notably, this is the first example of a

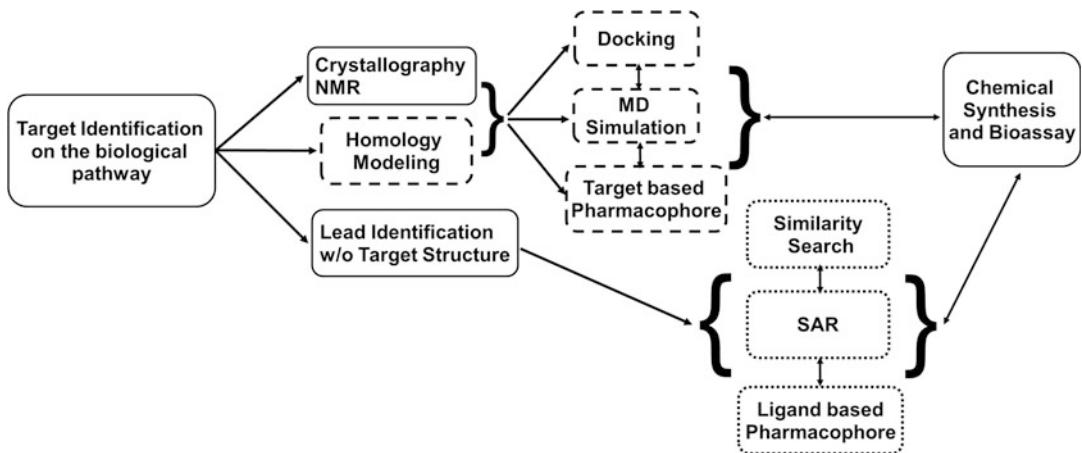


Fig. 1 Basic CADD workflow in drug discovery. Wet-lab, SBDD and LBDD CADD techniques are outlined in *solid lines*, *dashed lines*, or *dotted lines*, respectively. Double headed arrows indicate the two techniques can be used interactively in several iterative rounds of ligand design

small molecular weight compound that shows promising activity against Lipid II. Lead compound derivatives were subsequently identified again using CADD in combination with medicinal chemistry [20] and the accumulated SAR information will facilitate the development of next-generation antibiotics targeting gram positive pathogenic bacteria.

Figure 1 illustrates the basic CADD workflow that can be interactively used with experimental techniques to identify novel lead compounds as well as direct iterative ligand optimization [3, 4, 21, 22]. The process starts with the biological identification of a putative target to which ligand binding should lead to antimicrobial activity. In SDBB, the 3D structure of the target can be identified by X-ray crystallography or nuclear magnetic resonance (NMR) or using homology modeling. This lays the foundation for CADD SBDD screening using the methods described below. LBDD is used in the absence of the target 3D structure with the central theme being the development of an SAR from which information on modification of the lead compound to improve activity can be obtained. Information from the CADD methods is then used to design compounds that are subjected to chemical synthesis and biological assay, with the information from those experiments used to further develop the SAR, yielding further improvements in the compounds with respect to activity as well as absorption, disposition, metabolism, and excretion (ADME) considerations [23]. Notably, CADD methods are evolving with researchers continually updating and implementing new CADD techniques with higher levels of accuracy and speed [24–26]. In this chapter, we will present commonly used CADD approaches, including those used in our lab for the design of next-generation antibiotics.

2 Materials

CADD methods are mathematical tools to manipulate and quantify the properties of potential drug candidates as implemented in a number of programs. These include a range of publicly and commercially available software packages; the subset described below represents examples of fundamental tools for CADD with emphasis on those commonly used in our laboratory.

1. Commonly used MD simulation codes include CHARMM [27], AMBER [28], NAMD [29], GROMACS [30], and OpenMM [31]. These programs run on a variety of computer architectures including running in parallel on multicore central processing units (CPU) and, recently, optimized for graphics processing units (GPU), such as those commonly used in video games.
2. For SBDD, the 3D structure of the protein, RNA, or other macromolecule may be obtained from the Protein Data Bank (PDB) [32] if it was solved by X-ray crystallography or NMR experiments. Alternatively, a 3D structure may be constructed using homology modeling methods with a program such as MODELLER [33] or an online web server such as SWISS-MODEL [34].
3. To perform MD simulations, homology modeling, database screening, or other CADD techniques, empirical force fields for the molecules of interest are needed. These force fields are used by the respective programs to estimate the energy and forces associated with, for example, a drug-protein complex. Force fields such as those from the CHARMM [35–38] or AMBER [39, 40] families are used to describe the internal and external energetic properties of a molecular system during an energy minimization or a MD simulation. When parameters are missing in the existing force field, which is common for small drug-like molecules, automated parameter generation programs such as the CGenFF program [41, 42] or Antechamber [43] can be used to complete the force field. It is important to note that when using a force field, the parameters for different parts of the system (e.g., the protein and the ligand) need to be compatible, such that CGenFF should be used with CHARMM or Antechamber with AMBER. In addition, when parameters are estimated, it is suggested that the user check the parameters with respect to their accuracy in treating the energy as a function of conformation, as described for CGenFF [37, 44, 45]. To facilitate this process when generating parameters using the CGenFF program (*see* <https://cgenff.paramchem.org>), penalties are assigned to parameters estimated based on analogy, guiding the user with respect to parameters that require checking.

4. When no information on the binding site of a target is available, putative binding sites can be identified by various CADD methods. An example is the binding response program [46] developed in our lab. The program identifies potential binding sites by considering both the geometrical match and the binding energy of a set of diverse drug-like compounds to the sites being queried on the protein. Other programs for binding site identification include FINDSITE [47] and ConCavity [48].
5. Virtual database screening (VS) techniques are generally used to screen large *in silico* compound databases to identify potential binders for a query target. Examples of docking software commonly used for this purpose are DOCK [49] and AutoDock [50] as well as AutoDock Vina [51], all of which are well-known freeware programs. Another example is the program Pharmer [52], which uses 3D pharmacophores for database screening.
6. The *in silico* database of drug-like compounds is an essential component of CADD ligand identification based on VS. A publicly accessible database of compounds for VS is ZINC [53] which currently has about 90 million compounds that can be purchased from various chemical vendors. In-house databases can also be constructed for particular VS needs and chemical vendors such as ChemBridge and ChemDiv [54] supply their chemical catalogs in SDF format for download. However, conversion of these into 3D structures can be challenging and all physiologically accessible protonation and tautomeric states of the ligands in the database should be included.
7. Commercially available CADD software packages include Discovery Studio [55], OpenEye [56], Schrödinger [57], and MOE [58]. These programs, which can often be obtained at a discount for academic users, cover most of the capabilities required for CADD including both SBDD and LBDD methods.

3 Methods

CADD can be separated into ligand or hit identification and ligand or hit optimization, with both SBDD and LBDD methods useful in the appropriate context. Database screening methods are often used for hit identification [59] while a number of methods may be used for hit optimization [4, 24, 60]. These include the Site-identification by ligand competitive saturation (SILCS) methodology. Below we present a collection of methods that may be used for both ligand identification and optimization.

3.1 MD Simulations

MD simulations can be used to study target-ligand interactions at an atomic level of detail [61], to generate conformational ensembles for the target or for the ligand to take flexibility into account

for both SBDD and LBDD studies (*see Note 1*) and, in combination with other methods, used to estimate relative free energies of binding. Following are the steps required to perform a standard MD simulation (*see Note 2* for additional MD techniques). A convenient web-based tool to perform a number of the steps below is the CHARMM-GUI at www.charmm-gui.org [62].

1. Download the 3D structure of the bacterial target structure of interest from the PDB or use homology modeling to generate a structure.
2. Refine the target structure including adjusting the side chain orientations, add hydrogens, and determine the appropriate protonation states for titratable residues. Software such as Reduce [63] can be used for this purpose. Remove or retain cofactors, ions, and crystal waters depending on the study needs.
3. Choose a force field, such as CHARMM36 (http://mackerell.umaryland.edu/charmm_ff.shtml) to describe the system and a MD code to carry out the simulation. Prepare the input files according to the program formats. If force field parameters are missing, develop parameters using an automated program such as the CGenFF program or following a standard parametrization protocol for the chosen force field [37, 38].
4. For explicit solvent MD, solvate the system in a water box with periodic boundary conditions (PBC) [61], a process that can be performed automatically using the CHARMM-GUI mentioned above. Minimize and equilibrate the whole system step by step to allow bad atomic contacts to relax and attain relaxed geometries. Usually, harmonic restraints are first put on non-water components of the system and gradually reduced through the minimization and MD equilibration. This avoids large changes in the target structure due to bad atomic contacts in the initial model. NVT canonical ensemble MD is usually used for first step equilibration and followed by NPT ensemble MD to allow the PBC box size to adjust corresponding to the temperature and pressure, typically 298 K and 1 atm, respectively, of interest.
5. Run the MD simulation in the NPT ensemble for the time scale corresponding to the phenomena being studied. This usually involves nano- to microsecond timescales, although some phenomena can occur on shorter timescales. The user is advised to check that the event of interest (e.g., conformational change of the protein binding site) has occurred multiple times during the simulation or the phenomenon being monitored does not change significantly with increasing simulation time. However, no MD simulation is ever truly converged such that changes in the properties being monitored may occur after it appears that they are no longer changing.

6. Do a basic quality check on the MD trajectories such as analyzing the root-mean-square deviation (RMSD) of the target with respect to the starting conformation along the simulation time. Typically, there is an increase in the RMSD followed by a stable, fluctuating value. However, as stated in the preceding section, though a simulation appears stable, additional changes can occur upon additional simulation time.
7. When studying target-ligand interactions, different properties along the trajectory can be calculated for analyses such as interaction energy and hydrogen bonding profiles. In addition, structural clustering algorithms can be used to extract representative conformations from MD trajectories [64] to understand different interaction patterns between the ligand and the protein that contribute to binding.

3.2 Site Identification by Ligand Competitive Saturation (SILCS)

SILCS is a novel CADD protocol developed in our lab to facilitate ligand design [65]. It uses all-atom explicit-solvent MD simulations that include small organic solutes, such as propane, methanol, and others, to identify 3D functional-group binding patterns on the target. These patterns can be used qualitatively to direct ligand design and, when converted to free energies, termed grid free energy (GFE) FragMaps [66, 67], used to quantitatively estimate the relative binding affinities of ligands. The detailed protocol based on full MD simulations was described previously in this same book series [68]. Here, we present an updated protocol based on the use of oscillating μ_{ex} Grand Canonical Monte Carlo/MD (GCMC/MD) simulations for SILCS [69]. The GCMC/MD approach allows for the application of the SILCS method to target systems with deep or occluded pockets such as nuclear receptors and GPCRs [70].

1. Prepare the system in a similar way as described in Subheading 3.1 for MD simulations. In addition to water, add solute molecules such as benzene, propane, methanol, formamide, acetaldehyde, imidazole, methylammonium, and acetate at a concentration of about 0.25 M.
2. Place weak restraints only on the backbone C α carbon atoms with a force constant (k in $1/2 k\delta x^2$) of 0.12 kcal/mol/Å for all residues or only on core region residues in the target if additional flexibility of selected regions of the protein is desired. The use of C α restraints prevents the rotation and translation of the protein in the simulation box and prevents potential denaturation due to the presence of small solutes in the aqueous solution surrounding the target [71].
3. This system is minimized for 5000 steps with the steepest descent (SD) algorithm [72] in the presence of PBC followed by a 250 ps MD equilibration during which temperature is adjusted by velocity rescaling.

4. During GCMC, solutes and water are exchanged between their gas-phase reservoirs and the simulation system. The excess chemical potential (μ_{ex}) supplied to drive solute and water exchange is periodically oscillated over every 3 cycles for each solute or water, based on their target concentration (e.g., 0.25 M for the solutes and 55 M for water). From these calculations, which are performed over 100 or more cycles, the average μ_{ex} is close to the respective experimental hydration free energy values of the solutes and water. As described in detail elsewhere [69], there are four possible GCMC moves: insertion, deletion, translation, and rotation, with the probabilities for acceptance of these moves governed by the Metropolis criteria.
5. The configuration at the end of each GCMC cycle is used as the starting configuration for a 0.5–1 ns MD simulation during which the protein can undergo conformational changes as well as to obtain additional sampling of the water and solutes in and around the target molecule. Before the production MD, a 500 step SD minimization and a 100 ps equilibration is run. The last conformation from the production MD is used as the starting conformation of the next GCMC cycle.
6. Ten independent 100 cycle GCMC-MD runs are recommended. For each cycle, 200,000 steps of GCMC and 0.5 ns MD are conducted yielding a cumulative 200 million steps of GCMC and 500 ns of MD over all 10 independent simulations.
7. 3D probability distributions of selected atoms from the solutes, called “FragMaps,” from the GCMC/MD simulations are constructed. These are converted to GFE FragMaps based on a Boltzmann transformation, which allow for quantitative evaluation of ligand affinities, including the contribution of individual atoms. The GFE FragMaps can be used to guide ligand docking using the MC-SILCS approach [67] or for the calculation of target pharmacophore models using SILCS-Pharm [73, 74].

3.3 Database Preparation

VS against a database containing commercially available compounds is an efficient way to find potential low-molecular weight binders to the target protein [59]. While the ZINC database is available, researchers may want to prepare an in-house database for specific use.

1. Download the commercial database(s) from chemical vendors such as Chembridge, Chemdiv, Maybridge, Specs, etc. These databases are most often in 2D SDF format and need further refinement.
2. Convert 2D SDF files into 3D structure files such as MOL2 format files using a chemical data tool such as Open Babel [75]

or RDKit [76]. During the conversion, preliminary geometry optimization can be conducted to refine the 3D geometry to avoid bad contacts that may be transferred from the 2D structure. Missing hydrogens are added and appropriate protonation states are determined usually for pH 7.2 (*see Note 3*). Various tautomers can also be generated and if subsequent screening studies will use rigid ligand docking, multiple rotamers, typically 100–200, can also be generated for consideration of the conformations accessible to each molecule.

3. All 3D structures can be further optimized using a force field-based minimization to obtain more chemically accurate structures and assign atomic charges for subsequent screening studies if required. Organic molecule force fields such as CGenFF [37, 38], GAFF [40], or MMFF94 [77] can be used for this purpose.
4. When a database is prepared based on compounds from various vendors, in-house consistent identifiers are often needed to tag all the compounds for easy data management. For each compound, various entries such as physical properties and vendor information can be added for convenient use in subsequent analyses. The database, if extremely large, can be divided into several pieces for more efficient use. Finally, the database needs to be saved in the format required by the software to be used in the following studies, for example, MOE [58] uses the binary MDB format while Dock uses the readable MOL2 format.

3.4 Docking-Based VS

Docking involves posing a compound in the putative binding site on the target in an optimal way defined by a scoring function in combination with a conformational sampling method [78]. Various docking programs are available that differ based on the scoring function used to describe the interaction between small molecule and the target and the conformational sampling method used to generate the binding poses of the ligand on the protein. Here, we present a docking protocol using the DOCK program [49] to illustrate the typical docking VS workflow.

1. Prepare the target structure in the required DOCK input format. Define the desired binding pocket on the protein surface either using experimental information or by using a binding pocket prediction program as described in the Materials section. As docking typically is based on a single conformation of the target, MD simulations of the target can be used to generate multiple conformations for individual docking runs. In this scenario, each compound in the database is docked to each target conformation and the most favorable score for that compound is used for ranking as described below.
2. Choose a sampling method and scoring scheme for docking. The DOCK program adopts an incremental ligand construction

and conformational sampling scheme that divides ligands into fragments and reassembles these fragments in the binding site in a number of different conformational poses. Scoring the binding poses uses a physical force field-based scoring function that includes both van der Waals (vdW) and electrostatic terms (*see Note 4* also).

3. Dock the entire compound database using a single crystal structure of the target or multiple conformations from MD mentioned above. Compounds are then ranked based on their interactions energies and selected for further analyses. It is suggested that multiple step VS can be used to balance the efficiency and reliability of docking results [79, 80]. This approach applies a more approximate, computationally faster approach for the full database of typically >1 million compounds from which a subset of compounds are selected for a secondary, more accurate dock screen.
4. When using multiple step VS with DOCK in our laboratory, the first round of docking involves a coarse but fast optimization for each compound in the database targeting one or a few target structures. 50,000 compounds are selected from this round based on the vdW attractive energy normalized for the compound molecular weight [81]. In this way, compounds with maximal steric complementarity with the target are selected rather than compounds with very favorable electrostatic interaction that do not complement the shape of the binding pocket. The molecular weight normalization accounts for the tendency of ranking based on interaction energies to favor larger compounds.
5. The 50,000 compounds selected from the first round of VS are subject to a second round of docking using a more rigorous optimization that includes more steps of minimization and multiple protein conformations (~10) are used to take target flexibility into account. The top 1000 hits based on MW normalized total interaction energies, including both vdW and electrostatic terms, are selected for further consideration. We emphasize that each compound is docked against each target conformation with the most favorable score over all the target conformations assigned to each compound, with that score used to select the top 1000 compounds.
6. The final selection step is to obtain ~100 compounds for biological assays that are diverse as well as having properties that will likely have favorable ADME properties (*see Note 5*). Diversity is important as it will maximize the potential of selecting biologically active compounds and having diverse lead compounds will improve the probability of ultimately identifying compounds that have a high probability of success in clinical trials. The top 1000 compounds can be clustered based on chemical structure and/or physiochemical properties to maximize the chemical

diversity of the selected compounds [80]. Other descriptors such as Lipinski's rule of 5 (RO5) [82] or the 4D Bioavailability (4D-BA) ranking [83] can be used as metrics of ADME to filter the final list for testing, although using rigorous cutoffs based on these metric is not advised as there are many therapeutic agents on the market that "break the rules."

3.5 SILCS-Pharm

An alternative to docking-based VS is target-based pharmacophore VS [84]. This approach can quickly filter a database for potential binders to a specific bacterial target. A pharmacophore model is defined as spatially distributed chemical features that are essential for specific ligand-target binding. It represents a simplification of the detailed energetic information used by docking methods and so its computational requirements are much lower. While multiple methods can be used to generate pharmacophores [84], we will present a method based on information from SILCS as described

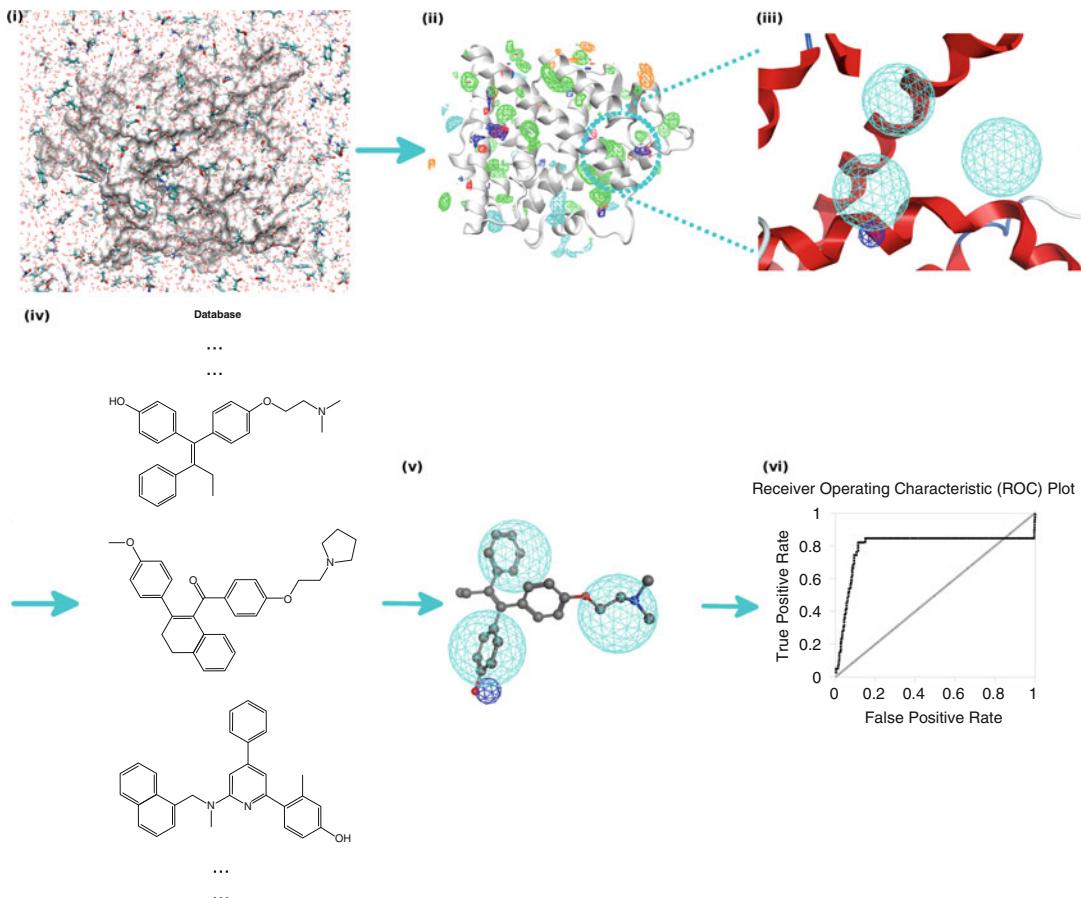


Fig. 2 SILCS-Pharm workflow for pharmacophore-based VS. The protocol starts from the SILCS simulation on the target (i), then FragMaps are generated (ii) and pharmacophore models are derived based on FragMaps (iii). The pharmacophore is then used in VS against a compound database (iv) that contains multiple conformations of each compound from which hit compounds are identified (v) and further tested in bioassays (vi)

in Subheading 3.2. The workflow for generation of a SILCS-based pharmacophore model [73, 74] is illustrated in Fig. 2.

1. Similar to docking VS, the desired binding site needs to be defined.
2. GFE FragMaps from SILCS are used as input into the SILCS-Pharm code [73, 74] to generate pharmacophore models. GFE cutoffs for FragMaps are used to define the sizes of related pharmacophore features and can be determined by visualizing FragMaps in a program such as VMD [85] and adjusting the contour value, as defined by the energy, to get well-separated, local FragMap regions. If the chosen GFE contour values are too high there will be many bulky features while contour values that are too low lead to few or no pharmacophore features for VS.
3. During generation of the pharmacophore by the SILCS-Pharm program, FragMap voxels within the defined GFE cutoffs will be clustered into intermediate SILCS features and then converted into standard pharmacophore features. The final generated pharmacophore models or hypotheses are ranked by the sum of all the feature GFEs in the model for a given number of features. More favorable GFE scores typically indicate a more effective model for use in VS as the GFE defines the strength of functional group binding obtained from the SILCS simulation. It is suggested that the most GFE favorable SILCS-Pharm model with four features can be used for VS based on tests in our lab [74].
4. Pharmacophore VS software such as Pharmer [52] or MOE [56] is then used to filter compounds in a database based on the selected SILCS-Pharm model. RMSD score, which represents the accordance between features in the pharmacophore model with related functional groups in a query compound, can be used to rank the final compound list.
5. As mentioned above, multiple, low energy conformations for each compound in the database should be pregenerated before pharmacophore VS as ligand flexibility is not included in the posing algorithm. Programs such as Open Babel [58] can be used for this purpose. 100–200 conformations for each ligand should be enough according to our in-house tests.
6. Once ligands are selected based on RMSD, alternate methods may be used to rank the ligands in a method referred to as consensus scoring [86]. For example, SILCS ligand grid free energy (LGFE) scores [67] can be used to re-rank the list to give a free energy-based ranking. The final compound list for experimental testing can be obtained by consensus scoring considering both RMSD and LGFE scores to maximize the hit potential [68].

3.6 Similarity Search

Once lead compounds are identified from experiments, LBDD methods can be utilized to start to develop an SAR or find more hit compounds. Of these, the similarity search method is the most straightforward and rapid approach [87]. It can search for compounds that are chemically or physiochemically similar to the input compound, as described later. This approach may also be used as lead validation, as a compound that has multiple analogs with biological activity from which SAR can be developed is appropriate for further studies [88].

1. Prepare the query compound in a format the program doing similarity search can recognize. The program MOE [58] has good similarity searching capabilities.
2. Choose the types of fingerprint used to define the compounds in the database. The fingerprint of a molecule refers to a collection of descriptors such as structural, physical, or chemical properties that are used to define the molecule [79]. Structural fingerprints, for example BIT MACCS [89], encode information such as the presence of specific types of atoms, bonds, or rings in the molecule and can be used to identify compounds that are structurally similar to the lead, facilitating SAR development, and may have improved binding affinity [88]. Physiochemical fingerprints such as MPMFP [90] encode properties such as the free energy of solvation, polarity, and molecular weight and can be used to identify compounds with dissimilar structures but similar physiochemical properties. This approach may help to identify novel hits that have activity but with a different chemical scaffold as compared to the lead compound, a process referred to as “lead hopping.” Such compounds could represent novel intellectual property (IP).
3. Choose a similarity comparison method and do the similarity search against an in silico database. To quantify the extent of similarity between two molecules, various similarity metrics [91] are available such as the commonly used Tanimoto coefficient [92]. Such metrics allow for giant databases to be rapidly screened. Compounds that are more similar to the query compound will have higher coefficients, such that the cutoff for the coefficient can be varied to select a desired number of similar compounds for testing. With the BIT MACCS fingerprints, a compound with a TC of 0.85 or higher (over a range of 0–1) is likely to have biological activity similar to that of the parent, query compound.

3.7 Lead Optimization Using SAR

When multiple hits for a specific bacterial target with activity data are available, structure-activity relationship (SAR) models can be developed and used to predict new compounds with improved activity [93]. LBDD SAR models use regression methods to relate a set of descriptors of the lead series of compounds to their

activities. The developed regression model can then be used to quantitatively predict the activity of the modified compounds [93]. The descriptors can be physical or chemical properties of compounds or even geometric parameters that are representative for the spatial distributions of important functional groups in the compounds, i.e., pharmacophore features. Knowledge of the relationship of these properties to activity (i.e., SAR) can be used by the medicinal chemist to qualitatively design new, synthetically accessible compounds that can be quantitatively evaluated. When developing SAR using pharmacophore descriptors, the appropriate conformations of the compounds that are responsible for the biological activity must be used. Here, we illustrate the development of SAR using our in-house developed conformationally sampled pharmacophore (CSP) protocol [94, 95].

1. Langevin dynamics-based MD simulations are conducted for all known hit compounds. Aqueous solvation effects of the simulated compounds can be included using explicit solvent or are treated using an implicit solvation model such as the generalized Born continuum solvent model [96]. Simulations should be performed for a minimum of 10 ns with the sampling of conformations of the ligand checked for convergence. If sampling is not adequate, the simulations should be extended or conducted using enhanced sampling methods, such as Temperature or Hamiltonian Replica Exchange methods [97]. Snapshots are typically saved every 0.2 ps for analysis.
2. Pharmacophore points, which are representative of well-conserved functional groups common in the hit compounds, such as aromatic ring centroid and hydrogen bond donor/acceptor atoms, are identified. Distances and angles between these pharmacophore points are measured throughout the trajectories from which probability distributions are obtained.
3. Analysis can be performed on 1- (1D) or 2-dimensional (2D) probability distributions. 1D distributions involve, for example, a distance between two important functional groups or the angle between three groups. 2D distributions can be between all possible distance or angle pairs. The 1D or 2D distributions are recorded for each hit compound. One hit compound, usually the most active compound, is selected as reference. To quantify the extent of similarity of the distributions, the overlap coefficients (OC) between the probability distributions of the reference compound and other compounds are calculated [95].
4. OCs are then used as independent variables in multiple regression analyses to fit the experimental activities. Different combinations of OCs for the various 1D and 2D pharmacophore probability distribution are regressed to identify those that yield the best correlation with the experimental data. For large

training sets of compounds, multiple SAR models can be developed [95]. The active compounds are usually divided into training and test set compounds with only the training set used for the SAR development, with the test set used to filter out the best SAR model. In studies of the opioids for a given set of compounds, CSP SAR models have been developed for both mu and delta efficacies [95, 98], allowing for identification of a compound that is both a mu agonist and a delta antagonist that may be of lower tolerance than opioids currently used in the clinic [99].

5. The regression model can be extended by the inclusion of physiochemical properties such as polar solvent accessibility, MW among others [100, 101].
6. The best CSP-SAR model can then be used to calculate predicted activities of query compounds and suggest the most potential compounds for further experimental tests. Ideally, multiple models are available for different activities allowing for both desirable and undesirable characteristics to be designed into the compounds, as done above with the opioids. In an ongoing study, as the number of compounds for which biological activity is available increases, the CSP model should be reevaluated to improve its predictability.

3.8 Single-Step Free Energy Perturbation (SSFEP)

Free energy perturbation (FEP) is a higher level, computationally demanding method with increased accuracy (*see Note 6*) that may be used to quantify the binding free energy change related to a modification in a compound [102]. To save computational time, the single step FEP (SSFEP) may be applied [103]. The approach uses a precomputed MD simulation of the hit compound-target complex from which the free energy difference due to small, single nonhydrogen atom modifications (e.g., aromatic –H to –Cl or –OH) can be rapidly evaluated [103]. This is in contrast to the need for many simulations in which the chemical modification is introduced in standard FEP methods [102]. SSFEP has the ability to give rapid predictions of binding affinity changes related to modifications and, thus, is quite useful for lead optimization [104]. The method may be applied using the following protocol with most simulations packages.

1. Run five 10 ns MD simulations of the hit compound-target complex and of the hit compound alone in solution.
2. For the chemical modification of the hit compound build in the modification onto the compounds with all other coordinates in the ligand and the remainder of the system identical to those from the original MD simulation.
3. Evaluate the interaction energy of the hit compound with the full environment for both the initial, unmodified, and modified

states for the simulations in the presence of the target and hit compound alone in solution.

4. Calculate the free energy difference, ΔG , in the presence of the protein and in aqueous solution based on the free energy perturbation formula [105] or the Bennett acceptance ratio (BAR) as described elsewhere [106]. The difference in the free energy differences in the presence of the protein and in aqueous solution yields the overall free energy difference, $\Delta\Delta G$, due to the chemical modification.

The utility of the SSFEP approach is that the $\Delta\Delta G$ values for many modifications may be rapidly evaluated as the same trajectories from the original MD simulations of the hit compound are used in each case. This approach may be of use during the fine tuning of ligand affinity or specificity for a target or as required to improve physiochemical and pharmacokinetic properties without significantly altering desirable properties such as affinity.

4 Notes

1. Conformational flexibility of molecules is a very important feature no matter if it is a small ligand or a large protein. Thus, conformational sampling of a protein or ligand that produces an ensemble of biological meaningful conformations is necessary either for SBDD or for LBDD. The CADD methods presented in the chapter such as SILCS for SBDD or CSP for LBDD take this issue into account and thus have advantages over other CADD methods that only rely on single crystal structure or limited ligand conformations.
2. MD simulation is an efficient way to generate conformational ensembles. For larger system, more advanced MD techniques can be employed to enhance the sampling efficiency such as replica exchange methods. The protocols developed in our lab such as Hamiltonian replica exchange with biasing potentials [107] and replica exchange with concurrent solute scaling and Hamiltonian biasing in one dimension [108] are efficient replica exchange methods for use to enhance the MD efficiency. However, with all MD-based methods, the user must perform careful analysis to assure that the conformational ensemble is adequately converged for effective use in CADD.
3. Protonation states of titratable residues at the targeted binding site and in the ligand being studied are quite important when setting up the CADD calculations. For example, different protonation states of histidine residues can offer different hydrogen bonding types to potential ligands. Available experimental

observations and known complex structures are useful to determine the correct protonation state of protein residue upon ligand binding. Software such as Reduce can assign the most appropriate protonation state based on environment. Constant pH MD simulation [109] where protonation state of titratable residue can change during the simulation may also be useful. With respect to ligands, many computational tools for prediction of ionization state are available, though common sense by the user is often adequate to deal with the most common ionizable groups such as carboxylates.

4. For VS, consensus scoring can be used instead of a single scoring scheme to rank hit compounds to allow more diversity of the identified compounds [86]. For example, in our SILCS-Pharm protocol, LGFE and RMSD are used together to rank compounds that pass our pharmacophore model filtering. Additional scoring metrics can include the DOCK or AUTODOCK scores [49, 50], or the average interaction energies from MD simulations, with many other variations available.
5. When constructing the final list of compound for experimental assays from VS, in addition to the binding score, drug likeness can be another criterion to further filter the list. Potential bioavailability of a compound is often judged by the Lipinski's rule of five (RO5) [82]. The 4-dimensional bioavailability (4D-BA) descriptor [83] is a scalar term derived from the four criteria in RO5 and thus facilitates the selection of potential bioavailable compounds in an automatic fashion. Pan assay interference compounds (PAINS) filter [110] can also be used to remove compounds that are likely to interfere in experimental screening techniques mainly through potential reactivity leading to false positives.
6. In the ligand optimization stage of CADD, as only a few compounds are under consideration, accuracy rather than computational efficiency is usually pursued. This means more sophisticated binding affinity evaluation methods should be used. These include the free energy methods such as SSFEP or the SILCS-based LGFE scoring discussed above.

Acknowledgments

This work was supported by NIH grants CA107331 and R43GM109635, University of Maryland Center for Biomolecular Therapeutics, Samuel Waxman Cancer Research Foundation, and the Computer-Aided Drug Design (CADD) Center at the University of Maryland, Baltimore.

Conflict of interest: A.D.M. is Co-founder and CSO of SilcsBio LLC.

References

1. Cohen ML (2000) Changing patterns of infectious disease. *Nature* 406:762–767
2. Walsh C (2003) Where will new antibiotics come from? *Nat Rev Microbiol* 1:65–70
3. Schneider G, Fechner U (2005) Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov* 4:649–663
4. Yu W, Guvench O, MacKerell AD (2013) Computational approaches for the design of protein–protein interaction inhibitors. In: Zinzalla G (ed) *Understanding and exploiting protein–protein interactions as drug targets*. Future Science Ltd., London, UK, pp 99–102
5. Panecka J, Mura C, Trylska J (2014) Interplay of the bacterial ribosomal a-site, S12 protein mutations and paromomycin binding: a molecular dynamics study. *PLoS One* 9, e111811
6. Resat H, Mezei M (1994) Grand canonical Monte Carlo simulation of water positions in crystal hydrates. *J Am Chem Soc* 116:7451–7452
7. Deng Y, Roux B (2008) Computation of binding free energy with molecular dynamics and grand canonical Monte Carlo simulations. *J Chem Phys* 128:115103
8. Small MC, Lopes P, Andrade RB, MacKerell AD Jr (2013) Impact of ribosomal modification on the binding of the antibiotic telithromycin using a combined grand canonical Monte Carlo/molecular dynamics simulation approach. *PLoS Comput Biol* 9, e1003113
9. Hossain M, Chowdhury DUS, Farhana J, Akbar MT, Chakraborty A, Islam S, Mannan A (2013) Identification of potential targets in *Staphylococcus aureus* N315 using computer aided protein data analysis. *Bioinformation* 9:187–192
10. O'Neill MJ, Wilks A (2013) The *P. aeruginosa* heme binding protein PhuS is a heme oxygenase titratable regulator of heme uptake. *ACS Chem Biol* 8:1794–1802
11. Nguyen AT, O'Neill MJ, Watts AM, Robson CL, Lamont IL, Wilks A, Oglesby-Sherrouse AG (2014) Adaptation of iron homeostasis pathways by a *Pseudomonas aeruginosa* pyoverdine mutant in the cystic fibrosis lung. *J Bacteriol* 196:2265–2276
12. Nguyen AT, Jones JW, Ruge MA, Kane MA, Oglesby-Sherrouse AG (2015) Iron depletion enhances production of antimicrobials by *Pseudomonas aeruginosa*. *J Bacteriol* 197:2265–2275. doi:[10.1128/JB.00072-15](https://doi.org/10.1128/JB.00072-15)
13. Furci LM, Lopes P, Eakanunkul S, Zhong S, MacKerell AD, Wilks A (2007) Inhibition of the bacterial heme oxygenases from *Pseudomonas aeruginosa* and *Neisseria meningitidis*: novel antimicrobial targets. *J Med Chem* 50:3804–3813
14. Hom K, Heinzl GA, Eakanunkul S, Lopes PEM, Xue F, MacKerell AD, Wilks A (2013) Small molecule antivirulents targeting the iron-regulated heme oxygenase (HemO) of *P. aeruginosa*. *J Med Chem* 56:2097–2109
15. O'Daniel PI, Peng Z, Pi H, Testero SA, Ding D, Spink E, Leemans E, Boudreau MA, Yamaguchi T, Schroeder VA, Wolter WR, Llarrull LI, Song W, Lastochkin E, Kumarasiri M, Antunes NT, Espahbodi M, Lichtenwalter K, Suckow MA, Vakulenko S, Mobashery S, Chang M (2014) Discovery of a new class of non-β-lactam inhibitors of penicillin-binding proteins with gram-positive antibacterial activity. *J Am Chem Soc* 136:3664–3672
16. Velvadapu V, Paul T, Wagh B, Klepacki D, Guvench O, MacKerell A, Andrade RB (2011) Desmethyl macrolides: synthesis and evaluation of 4,8,10-tridesmethyl telithromycin. *ACS Med Chem Lett* 2:68–72
17. Glassford I, Lee M, Wagh B, Velvadapu V, Paul T, Sandelin G, DeBrosse C, Klepacki D, Small MC, MacKerell AD, Andrade RB (2014) Desmethyl macrolides: synthesis and evaluation of 4-desmethyl telithromycin. *ACS Med Chem Lett* 5:1021–1026
18. Wagh B, Paul T, DeBrosse C, Klepacki D, Small MC, MacKerell AD, Andrade RB (2013) Desmethyl macrolides: synthesis and evaluation of 4,8,10-tridesmethyl cethromycin. *ACS Med Chem Lett* 4:1114–1118
19. Varney KM, Bonvin AMJJ, Pazgier M, Malin J, Yu W, Ateh E, Oashi T, Lu W, Huang J, Diepeveen-de Buin M, Bryant J, Breukink E, MacKerell AD Jr, de Leeuw EPH (2013) Turning defense into offense: defensin mimetics as novel antibiotics targeting lipid II. *PLoS Pathog* 9, e1003732
20. Fletcher S, Yu W, Huang J, Kwasny SM, Chauhan J, Opperman TJ, MacKerell AD Jr, de Leeuw E (2015) Structure-activity exploration of a small-molecule Lipid II inhibitor. *Drug Des Devel Ther* 9:2383–2394
21. Shijun Z, Alba TM, Alexander DM (2007) Computational identification of inhibitors of protein-protein interactions. *Curr Top Med Chem* 7:63–82
22. Shim J, MacKerell JAD (2011) Computational ligand-based rational design: role of conformational sampling and force fields in model development. *Med Chem Commun* 2:356–370

23. Ekins S, Boulanger B, Swaan P, Hupcey MZ (2002) Towards a new age of virtual ADME/TOX and multidimensional drug discovery. *J Comput Aided Mol Des* 16:381–401
24. Sliwoski G, Kothiwale S, Meiler J, Lowe EW (2014) Computational methods in drug discovery. *Pharmacol Rev* 66:334–395
25. Van Drie J (2007) Computer-aided drug design: the next 20 years. *J Comput Aided Mol Des* 21:591–601
26. Cavasotto CN (ed) (2015) In silico drug discovery and design: theory, methods, challenges, and applications. CRC Press, Boca Raton
27. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30:1545–1614
28. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688
29. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
30. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26:1701–1718
31. Eastman P, Friedrichs MS, Chodera JD, Radmer RJ, Bruns CM, Ku JP, Beauchamp KA, Lane TJ, Wang L-P, Shukla D, Tye T, Houston M, Stich T, Klein C, Shirts MR, Pande VS (2013) OpenMM 4: a reusable, extensible, hardware independent library for high performance molecular simulation. *J Chem Theory Comput* 9:461–469
32. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The protein data bank: a computer-based archival file for macromolecular structures. *J Mol Biol* 112:535–542
33. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815
34. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42:W252–W258
35. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616
36. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *J Chem Theory Comput* 8:3257–3273
37. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD (2010) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* 31:671–690
38. Yu W, He X, Vanommeslaeghe K, MacKerell AD (2012) Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J Comput Chem* 33:2451–2468
39. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117:5179–5197
40. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174
41. Vanommeslaeghe K, MacKerell AD (2012) Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* 52:3144–3154
42. Vanommeslaeghe K, Raman EP, MacKerell AD (2012) Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* 52:3155–3168
43. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25:247–260

44. Vanommeslaeghe K, Guvench O, MacKerell AD (2014) Molecular Mechanics. *Curr Pharm Des* 20:3281–3292
45. Vanommeslaeghe K, MacKerell AD Jr (2015) CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochim Biophys Acta* 1850:861–871
46. Zhong S, MacKerell AD (2007) Binding response: a descriptor for selecting ligand binding site on protein surfaces. *J Chem Inf Model* 47:2303–2315
47. Brylinski M, Skolnick J (2008) A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc Natl Acad Sci* 105:129–134
48. Capra JA, Laskowski RA, Thornton JM, Singh M, Funkhouser TA (2009) Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. *PLoS Comput Biol* 5, e1000585
49. Ewing TA, Makino S, Skillman AG, Kuntz I (2001) DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 15:411–428
50. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30:2785–2791
51. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461
52. Koes DR, Camacho CJ (2011) Pharmer: efficient and exact pharmacophore search. *J Chem Inf Model* 51:1307–1314
53. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* 52:1757–1768
54. <http://www.chembridge.com> and <http://www.chemdiv.com/>
55. Discovery studio modeling environment. Dassault Systèmes BIOVIA, San Diego. <http://accelrys.com/>. 2015
56. OEChem, OpenEye Scientific Software, Inc., Santa Fe. www.eyesopen.com. 2015
57. Schrödinger Softwares. Schrödinger, LLC, New York. <http://www.schrodinger.com>. 2015
58. Molecular Operating Environment (MOE). Chemical Computing Group Inc., Montreal. <https://www.chemcomp.com>. 2016
59. Martin YC (1992) 3D database searching in drug design. *J Med Chem* 35:2145–2154
60. Jorgensen WL (2009) Efficient drug lead discovery and optimization. *Acc Chem Res* 42:724–733
61. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Oxford University Press, Oxford, pp 1–383
62. Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 29:1859–1865
63. Word JM, Lovell SC, Richardson JS, Richardson DC (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J Mol Biol* 285:1735–1747
64. Karpen ME, Tobias DJ, Brooks CL (1993) Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV. *Biochemistry* 32:412–420
65. Guvench O, MacKerell AD Jr (2009) Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput Biol* 5, e1000435
66. Raman EP, Yu W, Guvench O, MacKerell AD (2011) Reproducing crystal binding modes of ligand functional groups using Site-Identification by Ligand Competitive Saturation (SILCS) simulations. *J Chem Inf Model* 51:877–896
67. Raman EP, Yu W, Lakkaraju SK, MacKerell AD (2013) Inclusion of multiple fragment types in the site identification by ligand competitive saturation (SILCS) approach. *J Chem Inf Model* 53:3384–3398
68. Faller C, Raman EP, MacKerell A Jr, Guvench O (2015) Site identification by ligand competitive saturation (SILCS) simulations for fragment-based drug design. In: Klon AE (ed) Fragment-based methods in drug discovery. Springer, New York, pp 75–87
69. Lakkaraju SK, Raman EP, Yu W, MacKerell AD (2014) Sampling of organic solutes in aqueous and heterogeneous environments using oscillating excess chemical potentials in grand canonical-like Monte Carlo-molecular dynamics simulations. *J Chem Theory Comput* 10:2281–2290
70. Lakkaraju SK, Yu W, Raman EP, Hershfeld AV, Fang L, Deshpande DA, MacKerell AD (2015) Mapping functional group free energy patterns at protein occluded sites: nuclear receptors and G-protein coupled receptors. *J Chem Inf Model* 55:700–708
71. Foster TJ, MacKerell AD, Guvench O (2012) Balancing target flexibility and target denaturation in computational fragment-based inhibitor discovery. *J Comput Chem* 33:1880–1891

72. Arfken G (1985) The method of steepest descents. Mathematical methods for physicists, 3rd edn. Academic, Orlando, pp 428–436
73. Yu W, Lakkaraju S, Raman EP, MacKerell A Jr (2014) Site-identification by ligand competitive saturation (SILCS) assisted pharmacophore modeling. *J Comput Aided Mol Des* 28:491–507
74. Yu W, Lakkaraju SK, Raman EP, Fang L, MacKerell AD (2015) Pharmacophore modeling using site-identification by ligand competitive saturation (SILCS) with multiple probe molecules. *J Chem Inf Model* 55:407–420
75. O’Boyle N, Banck M, James C, Morley C, Vandermeersch T, Hutchison G (2011) Open Babel: an open chemical toolbox. *J Cheminform* 3:33
76. RDKit: cheminformatics and machine learning software. <http://rdkit.org/>. 2015
77. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J Comput Chem* 17:490–519
78. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3:935–949
79. Zhong S, Oashi T, Yu W, Shapiro P, MacKerell AD Jr (2012) Prospects of modulating protein–protein interactions. Protein-ligand interactions. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp 295–329
80. Zhong S, Chen X, Zhu X, Dziegielewska B, Bachman KE, Ellenberger T, Ballin JD, Wilson GM, Tomkinson AE, MacKerell AD (2008) Identification and validation of human DNA ligase inhibitors using computer-aided drug design. *J Med Chem* 51:4553–4562
81. Pan Y, Huang N, Cho S, MacKerell AD (2003) Consideration of molecular weight during compound selection in virtual target-based database screening. *J Chem Inf Comput Sci* 43:267–272
82. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46:3–26
83. Oashi T, Ringer AL, Raman EP, MacKerell AD Jr (2011) Automated selection of compounds with physicochemical properties to maximize bioavailability and druglikeness. *J Chem Inf Model* 51:148–158
84. Koes D (2015) Pharmacophore modeling: methods and applications. Methods in pharmacology and toxicology. Humana Press, New York, pp 1–22
85. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38
86. Wang R, Wang S (2001) How does consensus scoring work for virtual library screening? An idealized computer experiment. *J Chem Inf Comput Sci* 41:1422–1426
87. Sheridan RP, Kearsley SK (2002) Why do we need so many chemical similarity search methods? *Drug Discov Today* 7:903–911
88. Macias AT, Mia MY, Xia G, Hayashi J, MacKerell AD (2005) Lead validation and SAR development via chemical similarity searching; application to compounds targeting the pY+3 site of the SH2 domain of p56lck. *J Chem Inf Model* 45:1759–1766
89. Durant JL, Leland BA, Henry DR, Nourse JG (2002) Reoptimization of MDL keys for use in drug discovery. *J Chem Inf Comput Sci* 42:1273–1280
90. Xue L, Godden JW, Stahura FL, Bajorath J (2003) Design and evaluation of a molecular fingerprint involving the transformation of property descriptor values into a binary classification scheme. *J Chem Inf Comput Sci* 43:1151–1157
91. Todeschini R, Consonni V, Xiang H, Holliday J, Buscema M, Willett P (2012) Similarity coefficients for binary chemoinformatics data: overview and extended comparison using simulated and real data sets. *J Chem Inf Model* 52:2884–2901
92. Tanimoto T (1958) An elementary mathematical theory of classification and prediction. IBM Internal Report
93. Gedeck P, Kramer C, Ertl P (2010) 4—computational analysis of structure-activity relationships. In: Witty DR, Lawton G (eds) Progress in medicinal chemistry. Elsevier, Amsterdam, pp 113–160
94. Bernard D, Coop A, MacKerell AD (2003) 2D conformationally sampled pharmacophore: a ligand-based pharmacophore to differentiate δ opioid agonists from antagonists. *J Am Chem Soc* 125:3101–3107
95. Bernard D, Coop A, MacKerell AD (2007) Quantitative conformationally sampled pharmacophore for δ opioid ligands: reevaluation of hydrophobic moieties essential for biological activity. *J Med Chem* 50:1799–1809
96. Qiu D, Shenkin PS, Hollinger FP, Still WC (1997) The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born Radii. *J Phys Chem A* 101:3005–3014

97. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151
98. Shim J, Coop A, MacKerell AD (2011) Consensus 3D model of μ -opioid receptor ligand efficacy based on a quantitative conformationally sampled pharmacophore. *J Phys Chem B* 115:7487–7496
99. Healy JR, Bezawada P, Shim J, Jones JW, Kane MA, MacKerell AD, Coop A, Matsumoto RR (2013) Synthesis, modeling, and pharmacological evaluation of UMB 425, a mixed μ agonist/ δ antagonist opioid analgesic with reduced tolerance liabilities. *ACS Chem Neurosci* 4:1256–1266
100. Rais R, Acharya C, Tririya G, MacKerell AD, Polli JE (2010) Molecular switch controlling the binding of anionic bile acid conjugates to human apical sodium-dependent bile acid transporter. *J Med Chem* 53:4749–4760
101. Chayan A, Andrew C, James EP, Alexander DM (2011) Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr Comput Aided Drug Des* 7:10–22
102. Chipot C, Pohorille A (eds) (2007) Free energy calculations: theory and applications in chemistry and biology. Springer, New York
103. Liu H, Mark AE, van Gunsteren WF (1996) Estimating the relative free energy of different molecular states with respect to a single reference state. *J Phys Chem* 100:9485–9494
104. Raman EP, Vanommeslaeghe K, MacKerell AD (2012) Site-specific fragment identification guided by single-step free energy perturbation calculations. *J Chem Theory Comput* 8:3513–3525
105. Zwanzig RW (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar gases. *J Chem Phys* 22:1420–1426
106. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:124105
107. Yang M, MacKerell AD (2015) Conformational sampling of oligosaccharides using Hamiltonian replica exchange with two-dimensional dihedral biasing potentials and the weighted histogram analysis method (WHAM). *J Chem Theory Comput* 11:788–799
108. Yang M, Huang J, MacKerell AD (2015) Enhanced conformational sampling using replica exchange with concurrent solute scaling and Hamiltonian biasing realized in one dimension. *J Chem Theory Comput* 11:2855–2867
109. Khandogin J, Brooks CL (2005) Constant pH molecular dynamics with proton tautomerism. *Biophys J* 89:141–157
110. Baell JB, Holloway GA (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 53:2719–2740

Chapter 6

Cytotoxicity Assays as Predictors of the Safety and Efficacy of Antimicrobial Agents

Alexander Zipperer and Dorothee Kretschmer

Abstract

The development of safe antimicrobial agents is important for the effective treatment of pathogens. From a multitude of discovered inhibitory compounds only few antimicrobial agents are able to enter the market. Many antimicrobials are, on the one hand, quite effective in killing pathogens but, on the other hand, cytotoxic to eukaryotic cells. Cell health can be monitored by various methods. Plasma membrane integrity, DNA synthesis, enzyme activity, and reducing conditions within the cell are known indicators of cell viability and cell death. For a comprehensive overview, methods to analyze cytotoxic and hemolytic effects, e.g., lactate dehydrogenase release, cell proliferation analysis, cell viability analysis, and hemolysis assay of antimicrobial compounds on human cells, are described in this chapter.

Key words Cytotoxicity assay, Lactate dehydrogenase release, Resazurin-based cell viability assay, Cell proliferation reagent WST-1, Hemolysis assay

1 Introduction

The emergence of multidrug-resistant organisms (MDRO) makes it essential to have access to adequate techniques to analyze drugs not only in relation to their antimicrobial effects but also regarding their cytotoxic potential. Although MDROs have evolved over decades, there are only very few antibiotics on the market, which can be used to treat such infections. The example of vancomycin shows, how important it is to develop novel antibiotics. Vancomycin is an antibiotic of last resort and of high clinical importance since it is applied to treat persistent infections. However, it also possesses nephrotoxic and ototoxic properties and can cause renal failure when injected intravenously [1]. The identification of treatment alternatives with less side effects is therefore essential to ensure patients well-being. Cytotoxicity can be analyzed by various methods, for example via measuring of the cell membrane integrity through membrane leakage assays. One example (to measure toxicity of substances) is the labeling of cells with Trypan Blue, a vital

stain (diazotized dye) that crosses only damaged membranes and will therefore stain only damaged/dead cells. The activity of the enzyme lactate dehydrogenase (LDH) is another indicator that is used very frequently to measure the integrity of cell membranes after treatment with cytotoxic compounds. LDH is present in nearly all cells and cell lines and catalyzes the conversion of lactate into pyruvate. Damage to the plasma membrane leads to rapid release of LDH into the cell culture supernatant where its activity can be determined via an enzymatic test [2–5].

Further assays that can be used to analyze cytotoxicity are based on the activity measurement of mitochondrial enzymes of healthy cells [6–14]. Resazurin functions as an indicator for cell health by using the reducing power of living cells to quantitatively measure the proliferation of various cell lines. Thereby the relative cytotoxicity of agents can be measured. Living cells maintain a reducing environment within the cytosol of the cell. Resazurin is a nontoxic, cell permeable compound that is blue in color and non-fluorescent. Upon entering the cells, resazurin is reduced to resorufin, a compound that is red in color and highly fluorescent. Viable cells continuously convert resazurin into resorufin, increasing the overall fluorescence and the color of the growth medium. Like resazurin WST-1 can be used for analysis of cytotoxic compounds based on the investigation of cell proliferation. This test is based on the cleavage of tetrazolium salts to formazan by cellular enzymes. Proliferation of cells results in an increase in the overall activity of mitochondrial dehydrogenases in the sample. This augmentation in enzyme activity leads to an increase in the amount of formazan dye and correlates directly to the number of metabolically active cells in the culture [15, 16]. In contrast to Resazurin WST-1 is a red substrate. The color of WST-1 changes into yellow, if the cells are alive. A further method to analyze toxicity of antimicrobials is to use nonmembrane permeable fluorescent dyes like 7-amino-actinomycin-D (7-AAD), which can intercalate in the DNA if cell membranes are damaged and undergoes a spectral shift upon association with DNA. Therefore, 7-AAD can be used as a marker for necrotic cells. The advantage of this method is that analyzing can occur via fluorescence-activated cell sorting (FACS) by a FACS analyzer.

A further method to analyze the toxicity of compounds is to determine the capacity of the antimicrobials to induce lysis of red blood cells. In the case of red blood cells this process is called hemolysis. This compound, an iron-containing oxygen-transport metalloprotein, is able to transport four oxygen molecules from the lung to distant tissues and organs where it is released to facilitate aerobic respiration. Simultaneously, hemoglobin returns carbon dioxide back to the lungs where the cycle is refreshed with the exchange of carbon dioxide to oxygen. Hemolysis can be easily

measured via analyzing the amount of hemoglobin in the supernatant of erythrocytes [17–19].

The dissemination of more and more MDROs results in the urgent need for potent antimicrobial compounds. To develop not only effective but also safe drugs, it is necessary to ascertain if compounds are cytotoxic or induce hemolysis of erythrocytes. To determine the potential cytotoxicity of compounds, leukocytes, which are quite sensitive cells, are commonly used. Here, we describe some basic, easy-to-perform methods that are frequently applied to analyze the cytotoxicity of various antimicrobial compounds.

2 Materials

2.1 Isolation of Human Neutrophils and Erythrocytes

1. Sodium heparin-containing tubes.
2. Pyrogen-free PBS.
3. Histopaque (density 1.119 g/ml).
4. Ficoll (density 1,077 g/ml).
5. Dulbeccos phosphor buffered saline (DPBS).
6. RPMI 1640 w/o L-glutamin.
7. Trypan Blue solution 4%.
8. Erythrocyte Lysis Buffer: to 500 ml bidest. H₂O add 4.145 g NH₄Cl, 0.5006 g KHCO₃, 0.0146 g EDTA.
9. Cell culture microscope.
10. Neubauer chamber for cell counting.

2.2 LDH Activity

1. Incubator (37 °C, 5% CO₂, 95% humidity).
2. Centrifuge with rotor for microtiter plates (MTP).
3. Microtiter plate reader with 490–492 nm filter (if a reference wavelength should be subtracted, a filter over 600 nm is recommended).
4. Cell culture microscope.
5. 96-well microplates with U-bottom for suspension cells and flat bottom for adherent cells. For color development in all assays: optically clear flat bottomed MTPs.
6. Assay medium (e.g., RPMI 1640 without phenol red containing 1% serum or 0.05% HSA).
7. Triton X-100 solution (2% Triton X-100 in assay medium). The maximum amount of releasable LDH enzyme activity is determined by lysing the cells with Triton X-100 (final concentration: 1% Triton X-100). At this concentration Triton X-100 does not affect the LDH activity.

8. Reaction mixture: consists usually of two compounds that have to be mixed shortly before use (depending on manufacturer's instructions).
9. Optional: 1 N HCl stop solution. The enzyme reaction can be stopped by the addition of 50 µl/well 1 N HCl (final concentration: 0.2 N HCl).
10. Optional: LDH standard preparation. If the released LDH activity has to be calculated in U/ml instead of relative cytotoxicity in percent of absorbance, it is recommended to use an appropriate LDH preparation as standard.

2.3 Cell Viability

Assay Based on Resazurin

1. Human leukemic monocytes (THP-1 cells).
2. Incubator (37 °C, 5% CO₂, 95% humidity).
3. Centrifuge with rotor for microtiter plates.
4. Microtiter plate reader to determine excitation at 560 nm and emission at 600 nm, optically clear flat bottomed 96-well microtiter plates.
5. Assay medium—RPMI-1640 (2 g NaHCO₃, 10% heat inactivated fetal calf serum (FCS), 1% l-glutamine and 1% penicillin-streptomycin), all contents have to be sterile and suitable for cell culture.
6. Triton X-100 as a positive control for strong cytotoxic activity (2% Triton X-100 is sufficient).
7. Alamar Blue (Invitrogen, AbD Serotec®, G-Biosciences, Sigma-Aldrich).

2.4 Cell Proliferation

Assay Using WST-1

1. Incubator (37 °C, 5% CO₂, 95% humidity).
2. Centrifuge (with rotor for MTP).
3. ELISA reader with a filter for a wavelength between 420 and 480 nm (if a reference wavelength is to be subtracted, a filter above 600 nm is recommended).
4. Cell culture microscope.
5. Hemocytometer.
6. 96-well microplates (flat bottom for adherent cells, tissue culture grade).
7. Culture medium, e.g., RPMI 1640 containing 10% heat-inactivated FCS, 2 mM l-glutamine, optionally add penicillin/streptomycin.
8. Cell Proliferation Reagent WST-1.

2.5 For Cell Death

Analysis with 7-Aminoactinomycin D (7-AAD)

1. 7-amino-actinomycin-D (7 AAD).
2. Culture medium (e.g., RPMI 1640).
3. FACS tubes (5 ml).

4. Flow cytometer with a 488 nm laser.
5. Incubator for heating cells up to 75 °C (e.g., thermomixer).
6. Flow cytometer software (Flow Jo or other) to analyze the FACS data.

2.6 Hemolysis Assay

1. Histopaque, Ficoll, heparinized tubes (*see* Subheading 2.1: isolation of leukocytes).
2. Incubator (37 °C, 5% CO₂, 95% humidity).
3. Centrifuge for reaction tubes (1.5 ml, 2 ml, 15 ml, 50 ml).
4. Spectral photometer with 540 nm absorbance measurement.
5. Centrifuge with rotor for MTP.
6. Assay medium: 1× phosphate buffered saline (PBS).
7. 2% Triton X-100 in 1× PBS as positive control.

3 Methods

3.1 Isolation of Human Neutrophils and Erythrocytes

1. Collect blood obtained from healthy human volunteers into sodium heparin-containing tubes for neutrophil isolation.
2. Dilute the heparinized blood 1:2 (v/v) with pyrogen-free PBS and layer it carefully onto a gradient of 12 ml Histopaque (density 1.119 g/ml) and 10 ml Ficoll (density 1,077 g/ml) in a 50 ml Falcon tube. A maximum of 25 ml of the diluted blood can be used per Falcon tube.
3. Centrifuge the tube at 380 × φ for 20 min at room temperature (RT) with a maximum of acceleration and without brake.
4. After centrifugation isolate the neutrophils from the Histopaque phase, transfer the cells in a new 50 ml Falcon tube, and wash them with RPMI and added 0.05% HSA (RPMI/HSA) at 250 × φ for 10 min at RT with a maximum of acceleration and with brake (Fig. 1).
5. Discard the supernatant carefully.
6. Incubate the cell pellet, which contains the neutrophils and residual erythrocytes, for 3 min at 37 °C and 5% CO₂ with 5 ml erythrocyte lysis buffer.
7. After incubation, wash neutrophils again with RPMI/HSA at 250 × φ for 10 min at RT and discard the supernatant. Solve the remaining cell pellet in 1–3 ml RPMI/HSA and determine the cell number (next step).
8. To analyze the cell number, mix 90 µl Trypan Blue with 10 µl of the cell suspension. Add 10 µl of this mixture to a Neubauer chamber. To calculate the amount of cells in the chamber, four big squares containing 16 small squares have to be enumerated,

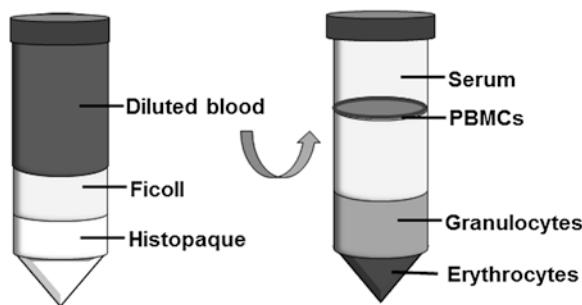


Fig. 1 Isolation of neutrophils and erythrocytes by density gradient centrifugation. Gradient before (*left side*) and after (*right side*) centrifugation

respectively. The cell number per ml can be determined with this formula:

$$\text{Cell number per ml} = \frac{\text{counted cells}}{4} \times 10 (\text{Dilution factor}) + 10^4 (\text{Chamber factor})$$

3.2 LDH Activity Assay

Colorimetric assay for the quantification of cell death and cell lysis. The assay is based on the measurement of lactate dehydrogenase (LDH) activity in cell supernatants. This allows indirect assessment of membrane integrity since LDH is released from the cytosol only if cell membranes are damaged. NAD⁺ is reduced to NADH/H⁺ by the LDH catalyzed conversion of lactate to pyruvate. In a second step the catalyst transfers H/H⁺ from NADH/H⁺ to the tetrazolium salt that is reduced to formazan resulting in a change of color.

1. Adjust cell number to 3×10^6 cells/ml (for neutrophils; for HL60 cells: 1×10^6 cells/ml) in RPMI without phenol red. For other cells first determine the optimal cell number for this assay (*see Notes 1 and 2*);
2. Use a 96-well microtiter plate (round bottom for cells grown in suspension) and set up the following reactions: a) 200 µl medium, b) 100 µl medium + 100 µl cells, c) 100 µl 2% Triton X-100 + 100 µl cells, d) 100 µl of diluted compound (dilute in RPMI 1640 without phenol red (*see Note 3*)) + 100 µl cells. Test all reactions in triplicates.
3. Incubate the microtiter plate with the samples at 37 °C, 5 % CO₂, 95 % humidity in the incubator (stimulation time depends on your sample, it is recommended to test the optimal incubation time in a first experiment, e.g., after 1 h, 3 h, 6 h, 12 h, and 24 h).
4. After incubation time centrifuge the microplate at $250 \times g$ for 10 min.

5. Carefully take 100 µl/well of the supernatants (without cells) and transfer to the correspondent wells of a second flat bottom microtiter plate.
6. Add 100 µl/well of freshly prepared reaction mixture to the second microtiter plate (calculate the appropriate amount for your experiment: mix compound one with compound 2, *see Note 4*).
7. Incubate the microtiter plate with the samples up to 30 min protected from light at RT.
8. Measure absorbance at 490 nm (reference wavelength should be more than 600 nm) in an ELISA reader.

3.3 Cell Viability Assay Based on Resazurin

The major component of the cell viability indicator Alamar Blue® is resazurin, a nontoxic and cell permeable compound. Viable cells convert resazurin into the red-fluorescent resorufin. The amount of fluorescence measured is proportional to the number of living cells.

1. Culture THP-1 cells in RPMI-1640 *with* penicillin-streptomycin in a cell culture flask (*see Note 5*).
2. Count cell numbers by Trypan Blue staining (*see Note 6*).
3. If your culture contains sufficient cell numbers, centrifuge (10 min, 250×*g*) and resuspend the pellet in RPMI-1640 *without* penicillin-streptomycin and *without* phenol red.
4. Prepare a 96-well microtiter plate with your test compounds, 2% Triton X-100 as positive control (final concentration per well) and corresponding solvent(s) in which the test compounds were diluted as negative control(s).
5. Add 1×10^4 THP-1 cells in 188 µl medium into each well and to your test compound (2 µl per well).
6. Incubate your samples for 48 h at 37 °C and 5% CO₂.
7. Add 10 µl Alamar Blue® to each well and incubate the samples for 24 h at 37 °C and 5% CO₂ protected from direct light (*see Note 7*).
8. Determine the relative fluorescent units (excitation 560 nm and emission 600 nm, *see Note 8*) in your microtiter plate reader with a peak emission at 585 nm (Fig. 2).
9. Results are analyzed by plotting fluorescence intensity (or absorbance) versus compound concentration.

3.4 Cell Viability Assay Based on Cell Proliferation Reagent WST-1

Analysis of cytotoxic compounds such as antimicrobials or other potential toxic compounds. This assay can be used, if the amount of dead cells in the negative control without compound is quite high.

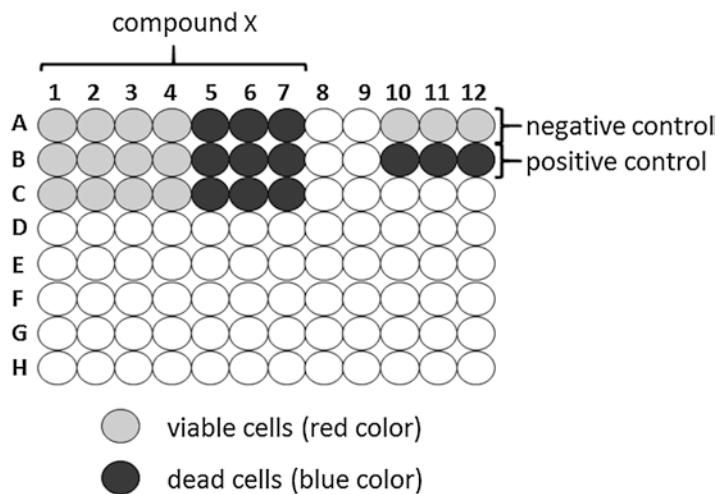


Fig. 2 Analysis of cell proliferation and cytotoxicity with resazurin. Defined cell numbers were incubated with various concentrations of compound X. After the subsequent addition of Alamar Blue® wells are stained depending on the cell condition

1. Seed cells at a concentration of 5×10^4 cells/well (depending on the cell type *see Notes 9 and 10*) in 100 µl culture medium (RPMI +10% heat inactivated FCS) and different concentrations of the test compound (depending on the compound) into 96-well microplates.
2. Incubate cell cultures for 24 h in a humidified atmosphere at 37 °C, 5% CO₂, 95% H₂O.
3. Add 10 µl Cell Proliferation Reagent WST-1 (*see Note 11*) to each well and incubate for up to 4 h (*see Note 12*) at 37 °C and 5% CO₂.
4. Shake cells thoroughly for 1 min (300 rpm) on a shaker or in the microplate reader (*see Note 13*).
5. Measure absorbance of the samples against a medium control as blank using an ELISA reader according to the filters available for the ELISA reader. The wavelength for measuring the absorbance of the formazan product is between 420 and 480 nm (max. absorption 440 nm).

3.5 Cell Death Analysis with 7-Amino-Actinomycin-D (7-AAD)

If the cell membrane is disrupted by toxic substances, 7-AAD can bind selectively to the GC-regions of the DNA. Consequently, necrotic cells are stained with 7-AAD while living cells with intact membranes are not stained. After incubation of the samples with 7-AAD the fluorescence of the samples can be measured with a flow cytometer.

1. Transfer 250 µl cells in medium (e.g., RPMI 1640) into 5 ml FACS-tubes, with a final concentration of 2.5×10^5 cells/tube. Add 250 µl of the test compound in various concentrations or controls (for the positive control heat cells up to 75 °C for 45 min, thereby cells undergo necrosis, for the negative control use medium).
2. Incubate cells with antimicrobials or controls for up to 24 h (depending on the substance and the cells) at 37 °C, 5% CO₂ in an incubator.
3. 20 min before measurement of the samples, add 7-AAD at a concentration of 5 µl/tube (0.25 µg/tube) and incubate the cells in the incubator.
4. Stop the reaction by transferring the tubes on ice and measure fluorescence by using a flow cytometer.

3.6 Hemolysis Assay

Lysis of erythrocytes leads to a release of hemoglobin into the supernatant, which can be measured spectrophotometrically [5, 20, 21]. Isolate the red blood cells from healthy human volunteers (*see Note 14*) by standard Histopaque/Ficoll centrifugation.

1. Wash the red blood cells with 1× PBS in 50 ml, 15 ml, or 2 ml reaction tubes depending on the volume of blood you use (*see Note 15*).
2. Dilute the red blood cells 1:50 with 1× PBS to prepare a 2% suspension.
3. Add your compounds to the 2% suspension of red blood cells and incubate them at 37 °C for 60 min (incubate red blood cells with Triton X-100 at a final concentration of 2% as positive control and red blood cells in 1× PBS as negative control).
4. Centrifuge samples for 10 min at 100 × g. Transfer the supernatant into a new tube and dilute it 1:10 with 1× PBS (Fig. 3).
5. Determine the hemolytic activity by measuring the optical density (hemoglobin absorbance at 540 nm) of the cell-free supernatant with a spectrophotometer.

4 Notes

1. Before you start with your experiment, determine the optimal cell concentration for the assay since different cell types contain different amounts of LDH. Therefore, the optimum cell concentration for a specific cell type should be determined in a preliminary experiment. In general, the cell concentration, in which the difference between the negative and positive control is at a maximum, should be used for the subsequent assay. The

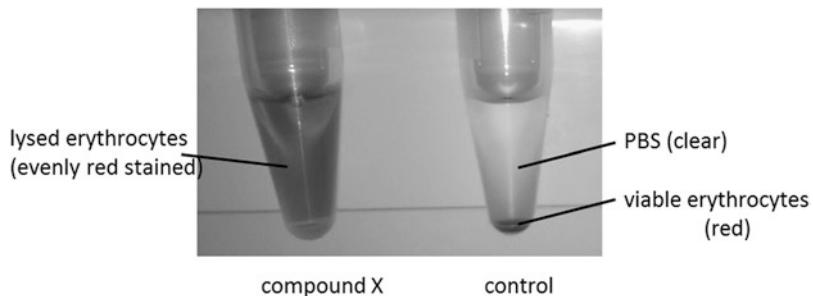


Fig. 3 Analysis of compound-dependent erythrocyte lysis. Erythrocytes were incubated with compound X. The lysis of red blood cells was subsequently monitored by OD measurement of the supernatant

optimal concentration for most cell lines is between 0.5 and 2×10^4 cells/well in 200 μ l. To this you have to adjust the cell suspensions to a concentration of 2×10^6 cell/ml and titrate the cells by twofold serial dilutions across the plates. For each concentration you need a negative control (=culture medium, spontaneous LDH release) and a positive control (100 μ l per well 2% Triton X-100 solution). The best cell concentration shows the highest difference between positive and negative control.

2. Add 100 μ l cell suspension per well to a sterile 96-well tissue culture plate (flat bottom) and incubate the cells overnight (or longer, depending on the cell type) in an incubator to allow the cells to adhere tightly. Remove the assay medium from the adherent cells and add 100 μ l fresh medium to each well. Then add your test substance to the adherent cells and incubate them (2–24 h depending on your assay). For example, seed A549 lung epithelial cells 1.14×10^4 cells/well in 1 ml in a 24-well plate and after 4 days stimulate cells for 24 h.
3. Use always media without phenol red, since phenol red can influence the absorbance.
4. The freshly prepared reaction mixture should not be stored, due to its sensitivity to light. When the two compounds for the reaction mixture are thawed they can be stored at 4 °C for up to 6 weeks. In general, all substances have to be regarded as unstable.
5. Cultivation of leukemic monocytes to obtain sufficient cell numbers: Use RMPI-1640 with 1% penicillin-streptomycin and phenol red. Renew the medium every 3–4 days by centrifugation, removal of the supernatant, and resuspension into fresh RMPI-1640.

6. Counting cells numbers: centrifuge the cell culture (10 min, $250 \times g$), remove the supernatant, and resuspend the pellet in 1 ml RPMI-1640 *without* phenol red. Mix an aliquot of this suspension in a ratio of 1:1 with Trypan Blue and count the cells (e.g., with an automated cell counter).
7. 5 % Alamar Blue (10 μ l) is less than stated in the manufacturer's protocol but completely sufficient. Optional: Add 50 μ l 3 % SDS directly to 100 μ l of cells in Alamar Blue® reagent to stop the reaction.
8. If a microplate reader is not available you can also measure your samples in a spectral photometer at an absorbance of 540–570 nm (peak excitation is 570 nm). Please note that this measurement is less sensitive! Assay plates or tubes can be wrapped in foil, stored at 4 °C, and read within 1–3 days without affecting the fluorescence or absorbance values.
9. For each cell type you should determine the optimal cell number and incubation time with WST-1, because a high cell number and proliferation leads to a detection limitation of the negative control (high absorbance).
10. After choosing a certain cell line, use only cells with low passage; otherwise, the results strongly vary.
11. The substrate is light and temperature sensitive. It is recommended to prepare aliquots of the substrate and to freeze it in adequate amounts.
12. For samples with higher cell numbers you can use shorter incubation times of 4 h or less. In general, sensitivity of detection increases with longer incubation times.
13. Before you measure absorbance, always remove bubbles, since they can falsify the result. To distribute the color equally in the well you should always shake the microtiter plate for 1 min before measuring the absorbance.
14. In case that human red blood cells are not available, the assay can be performed with rabbit or sheep red blood cells as well.
15. Blood samples can be stored at 4 °C overnight after the addition of Histopaque/Ficoll as the included heparin acts as an anticoagulant.

Acknowledgments

This work was supported by grants from the German Research Council (SFB685, SFB766, TRR156, and PE805/5-1, to A.P.; TRR34, to A.P. and D.K.) and the Fortune Program of the Medical Faculty, University of Tübingen (to D.K.).

References

1. Bruniera FR, Ferreira FM, Savioli LR, Bacci MR, Feder D, da Luz Goncalves Pedreira M, Sorgini Peterlini MA, Azzalis LA, Campos Junqueira VB, Fonseca FL (2015) The use of vancomycin with its therapeutic and adverse effects: a review. *Eur Rev Med Pharmacol Sci* 19(4):694–700
2. Decker T, Lohmann-Matthes ML (1988) A quick and simple method for the quantitation of lactate dehydrogenase release in measurements of cellular cytotoxicity and tumor necrosis factor (TNF) activity. *J Immunol Methods* 115(1):61–69
3. Korzeniewski C, Callewaert DM (1983) An enzyme-release assay for natural cytotoxicity. *J Immunol Methods* 64(3):313–320
4. Naal FD, Salzmann GM, von Knoch F, Tuebel J, Diehl P, Gradinger R, Schauwecker J (2008) The effects of clindamycin on human osteoblasts in vitro. *Arch Orthop Trauma Surg* 128(3):317–323. doi:[10.1007/s00402-007-0561-y](https://doi.org/10.1007/s00402-007-0561-y)
5. Wang R, Braughton KR, Kretschmer D, Bach TH, Queck SY, Li M, Kennedy AD, Dorward DW, Klebanoff SJ, Peschel A, DeLeo FR, Otto M (2007) Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. *Nat Med* 13(12):1510–1514. doi:[10.1038/nm1656](https://doi.org/10.1038/nm1656), nml1656 [pii]
6. Bara R, Zerfass I, Aly AH, Goldbach-Gecke H, Raghavan V, Sass P, Mandi A, Wray V, Polavarapu PL, Pretsch A, Lin W, Kurtan T, Debbab A, Brötz-Oesterhelt H, Proksch P (2013) Atropisomeric dihydroanthracenones as inhibitors of multiresistant *Staphylococcus aureus*. *J Med Chem* 56(8):3257–3272. doi:[10.1021/jm301816a](https://doi.org/10.1021/jm301816a)
7. Antczak C, Shum D, Escobar S, Bassit B, Kim E, Seshan VE, Wu N, Yang G, Ouerfelli O, Li YM, Scheinberg DA, Djabballah H (2007) High-throughput identification of inhibitors of human mitochondrial peptide deformylase. *J Biomol Screen* 12(4):521–535. doi:[10.1177/1087057107300463](https://doi.org/10.1177/1087057107300463)
8. Al-Nasiry S, Geusens N, Hanssens M, Luyten C, Pijnenborg R (2007) The use of Alamar Blue assay for quantitative analysis of viability, migration and invasion of choriocarcinoma cells. *Hum Reprod* 22(5):1304–1309. doi:[10.1093/humrep/dem011](https://doi.org/10.1093/humrep/dem011)
9. Gartlon J, Kinsner A, Bal-Price A, Coecke S, Clothier RH (2006) Evaluation of a proposed in vitro test strategy using neuronal and non-neuronal cell systems for detecting neurotoxicity. *Toxicol In Vitro* 20(8):1569–1581. doi:[10.1016/j.tiv.2006.07.009](https://doi.org/10.1016/j.tiv.2006.07.009)
10. Ovcharenko D, Jarvis R, Hunnicke-Smith S, Kelnar K, Brown D (2005) High-throughput RNAi screening in vitro: from cell lines to primary cells. *RNA* 11(6):985–993. doi:[10.1261/rna.7288405](https://doi.org/10.1261/rna.7288405)
11. Hamid R, Rotshteyn Y, Rabadi L, Parikh R, Bullock P (2004) Comparison of alamar blue and MTT assays for high through-put screening. *Toxicol In Vitro* 18(5):703–710. doi:[10.1016/j.tiv.2004.03.012](https://doi.org/10.1016/j.tiv.2004.03.012)
12. Gloeckner H, Jonuleit T, Lemke HD (2001) Monitoring of cell viability and cell growth in a hollow-fiber bioreactor by use of the dye Alamar Blue. *J Immunol Methods* 252(1-2):131–138
13. Nociari MM, Shalev A, Benias P, Russo C (1998) A novel one-step, highly sensitive fluorometric assay to evaluate cell-mediated cytotoxicity. *J Immunol Methods* 213(2):157–167
14. Nakayama GR, Caton MC, Nova MP, Parandoosh Z (1997) Assessment of the Alamar Blue assay for cellular growth and viability in vitro. *J Immunol Methods* 204(2):205–208
15. Roehm NW, Rodgers GH, Hatfield SM, Glasebrook AL (1991) An improved colorimetric assay for cell proliferation and viability utilizing the tetrazolium salt XTT. *J Immunol Methods* 142(2):257–265
16. Ishiyama M, Tominaga H, Shiga M, Sasamoto K, Ohkura Y, Ueno K (1996) A combined assay of cell viability and in vitro cytotoxicity with a highly water-soluble tetrazolium salt, neutral red and crystal violet. *Biol Pharm Bull* 19(11):1518–1520
17. Gallagher PG (2011) Hemolytic anemias: red cell membrane and metabolic defects. In: Goldman L, Schafer AI (eds) Cecil medicine, 24th edn. Saunders Elsevier, Philadelphia, PA
18. Powers A, Silberstein L (2008) Autoimmune hemolytic anemia. In: Hoffman R, Benz E, Shattil SS (eds) Hematology: basic principles and practice, 5th edn. Elsevier Churchill Livingstone, Philadelphia, PA
19. Schrier SPE (2008) Extrinsic nonimmune hemolytic anemias. In: Hoffman R, Benz E, Shattil SS (eds) Hematology: basic principles and practice, 5th edn. Elsevier Churchill Livingstone, Philadelphia, PA
20. Torres VJ, Attia AS, Masson WJ, Hood MI, Corbin BD, Beasley FC, Anderson KL, Stauff DL, McDonald WH, Zimmerman LJ, Friedman DB, Heinrichs DE, Dunman PM, Skaar EP (2010) *Staphylococcus aureus* *fur* regulates the expression of virulence factors that contribute to the pathogenesis of pneumonia. *Infect Immun* 78(4):1618–1628. doi:[10.1128/IAI.01423-09](https://doi.org/10.1128/IAI.01423-09)
21. Blevins JS, Beenken KE, Elasri MO, Hurlburt BK, Smeltzer MS (2002) Strain-dependent differences in the regulatory roles of *sarA* and *agr* in *Staphylococcus aureus*. *Infect Immun* 70(2):470–480

Part II

Mode of Action

Chapter 7

Application of a *Bacillus subtilis* Whole-Cell Biosensor (*PliaI-lux*) for the Identification of Cell Wall Active Antibacterial Compounds

Carolin Martina Kobras, Thorsten Mascher, and Susanne Gebhard

Abstract

Whole-cell biosensors, based on the visualization of a reporter strain's response to a particular stimulus, are a robust and cost-effective means to monitor defined environmental conditions or the presence of chemical compounds. One specific field in which such biosensors are frequently applied is drug discovery, i.e., the screening of large numbers of bacterial or fungal strains for the production of antimicrobial compounds. We here describe the application of a luminescence-based *Bacillus subtilis* biosensor for the discovery of cell wall active substances. The system is based on the well-characterized promoter *PliaI*, which is induced in response to a wide range of conditions that cause cell envelope stress, particularly antibiotics that interfere with the membrane-anchored steps of cell wall biosynthesis. A simple “spot-on-lawn” assay, where colonies of potential producer strains are grown directly on a lawn of the reporter strain, allows for quantitative and time-resolved detection of antimicrobial compounds. Due to the very low technical demands of this procedure, we expect it to be easily applicable to a large variety of candidate producer strains and growth conditions.

Key words Bio-assay, Reporter gene, Cell envelope stress, Cell wall, Antibiotic, Antimicrobial peptide, Stress response, Luminescence, Lipid II cycle

1 Introduction

Biosensors are “devices that use specific biochemical reactions mediated by isolated enzymes, immunosystems, tissues, organelles, or whole cells to detect chemical compounds, usually by electrical, thermal, or optical signals,” according to the IUPAC definition [1]. In recent years, whole-cell biosensors in particular have gained increasing attention in different fields of application, such as drug discovery or on-site monitoring of environmental samples, e.g., for pollutants such as heavy metal ions or xenobiotics. Compared to enzymes or the other biosensor platforms mentioned above, they offer the advantage of low costs, high stability, and ease of use [2]. Normally, whole-cell biosensors are genetically modified microorganisms that use the stimulus specificity of

signal-transducing regulatory systems to connect an input (compound or condition to be detected) with a measurable output. The latter is usually provided by a reporter gene under control of a promoter that is regulated by the signalling system.

Three different types of microbial reporter systems are most commonly employed. The β -galactosidase (encoded by *lacZ*) is the classical reporter gene and was already established in the early 1970s as a quantitative and highly reproducible measure for differential promoter activity, using the chromogenic substrate ONPG (*o*-nitrophenyl- β -D-galactopyranoside) [3]. Despite the disadvantage of having to collect cells and perform a biochemical assay for a quantitative read-out, it is still widespread, since it only requires a standard photometer for colorimetric detection of the enzymatic reaction product at a wavelength of 420 nm. Moreover, *lacZ*-based whole-cell biosensors offer the convenience of a low-cost, simple, and fast semiquantitative readout, if X-Gal (5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside) is used as a chromogenic substrate in plate-based whole-cell biosensor assays, which are suitable even for field work in the absence of any technical equipment [4].

More recently, two additional reporter systems have found widespread use in biosensors, namely fluorescent proteins, such as GFP, and bioluminescent reporters, derived from either bacterial or firefly luciferase. While both require more advanced documentation systems for quantitative measurements, their advantage over the β -galactosidase reporter is the possibility for a quantitative online monitoring of promoter activity in viable cells (e.g., growing liquid cultures in microtiter plates) without the need of harvesting cells and performing an assay. While GFP and its many derivatives have found widespread use in numerous biological applications, including whole-cell biosensors [5], the autofluorescence of the cells or media components often limits the dynamic range and hence sensitivity of fluorescence-based biosensors. In contrast, bioluminescence offers a virtually background-free reporter system, thereby enabling biosensors with a high dynamic range and hence sensitivity. Firefly luciferase requires the addition of luciferin as a substrate, which in the presence of ATP and oxygen leads to visible light emission. Bacterial luciferase, encoded by the *luxCDABE* operon of *Photinus pyralis*, catalyzes the oxidation of reduced flavin mononucleotide (FMNH₂) and fatty aldehydes to FMN and fatty acids, respectively, in the presence of molecular oxygen, resulting in blue-green light emission [2]. The substrates of this reaction are regenerated by the normal cellular metabolism and no addition of an external substrate is required, making bacterial luciferase the more convenient and cost-efficient alternative.

This chapter describes the use of a luminescence-based biosensor for the detection of antimicrobial substances that interfere with the membrane-anchored steps of cell wall biosynthesis. It is derived from the *liaI* promoter (*PliaI*) of the Gram-positive model

organism *Bacillus subtilis*, which is strictly controlled by the cell-envelope stress-responsive LiaFSR three-component system [6, 7]. It was first reported to strongly respond to the cell wall antibiotics vancomycin and bacitracin [8, 9]. Subsequently, *PliaI* was thoroughly characterized and developed into a β-galactosidase-based biosensor (*PliaI-lacZ*) for the identification and characterization of lipid-II interfering antibiotics (hence the name “lia”) [4, 10, 11]. *PliaI* possesses a very low basal promoter activity and a highly dynamic response (over 100-fold induction) to its specific inducers, making it ideally suited for screening purposes. More recently, a new *PliaI*-based whole-cell biosensor has been established by employing the bacterial luciferase system from *P. luminescence* [12]. The *luxABCDE* operon used in the resulting *PliaI-lux* biosensor has been optimized for expression in *B. subtilis* [13]. In addition to its convenience and high sensitivity, its short half-life of only 5–10 min allows an almost direct monitoring of not only the induction but also the shut-off of promoter activities, thereby providing another advantage over alternative reporter systems [12].

This chapter provides a detailed protocol of how to apply the *PliaI-lux* whole-cell biosensor for a quantitative bioluminescence detection of antimicrobial compounds interfering with cell wall biosynthesis. This “spot-on-lawn” assay is based on growing a lawn of the reporter strain on solid media and applying spots of potential antibiotic producer strains to this lawn. Production of cell wall active compounds will then induce the activity of *PliaI*, resulting in a ring-shaped luminescence signal around the producer colony that can be visualized and quantified using basic chemiluminescence detection equipment. For details on applying this biosensor strain for quantitative antibiotic induction experiments in liquid cultures, the readers are referred to the previously published procedure [12]. An alternative qualitative assay based on a *PliaI-lacZ* reporter is described in Subheading 4 (see Note 1).

2 Materials

Prepare all media and solutions using deionized water (dH₂O). All reagents can be prepared and stored at room temperature, except for agar plates (4 °C) or where stated otherwise.

All waste containing bacterial cultures should be disposed of according to local regulations. The reporter strain is a class I genetically modified organism, and all handling and disposal should follow good microbiological practice procedures.

This chapter describes standard growth conditions and media for *Bacillus subtilis*. Depending on special growth conditions that may be required by bacterial strains to be tested for antibiotic production, different media and conditions can be used, provided preliminary tests show that the *B. subtilis* reporter strain is able to grow under such conditions.

2.1 Media and Reagents

1. Lysogeny Broth (LB): 10 g tryptone, 5 g yeast extract, 10 g sodium chloride. Add dH₂O to a volume of 1 l and autoclave. Adjustment to pH 7 is not normally necessary, but can be achieved by addition of 1 M NaOH (if the pH is too low) or 1 M HCl (if the pH is too high); if required, adjust pH before autoclaving.
2. LB agar: add 15 g/L of agar (1.5 % (w/v)) to LB medium prior to autoclaving. Cool down agar to ~50 °C before adding antibiotics. Pour ca. 25 ml of agar per plate into 90 mm sterile petri dishes and let solidify.
3. LB soft agar: add 7.5 g/L of agar (0.75 %) to LB medium prior to autoclaving. For immediate use, split into aliquots after autoclaving. For this, transfer 4 ml of molten soft agar into sterile tubes and keep at 50 °C until further use. For later use, let it solidify and store at room temperature until needed.
4. Chloramphenicol stock solution: dissolve 5 mg/ml in 70% ethanol. Store at -20 °C.

2.2 Bacterial Strains

1. Reporter strain: *B. subtilis* W168 *sacA::PliaI-luxABCDE* (strain TMB1858). This strain contains the promoter *PliaI* fused to the *luxABCDE* luciferase reporter operon [12]. *PliaI* is activated by lipid II cycle interfering antibiotics, such as bacitracin, nisin, ramoplanin, and vancomycin [4]. Upon promotor induction, a chemiluminescence signal is emitted (see Notes 1 and 2). Growth media should contain 5 µg/ml chloramphenicol.
2. Positive control: *B. subtilis* ATCC6633. This strain is known as a producer of the cell envelope-active antibiotic subtilin and strongly induces the promoter *PliaI* present in the reporter strain [10].
3. Negative control: *B. subtilis* W168. This strain does not produce any compounds that induce the *PliaI-luxABCDE* reporter strain.
4. Strains to be tested for production of cell envelope-active compounds.

2.3 Special Equipment

1. For development of this assay, a FUSION-SL™ 16 bit chemiluminescence imaging system with the analysis software FUSION-CAPT™ was used (PEQLAB). Settings described in the methods section refer to this imaging system and software and may vary for other products. Other imaging systems with comparable sensitivity can be used, but exposure times may have to be optimized.
2. Chemiluminescence is quantified using the freely available software ImageJ (<http://imagej.nih.gov/ij/>). Further calculations can be carried out using Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) or any other suitable software.

3 Methods

3.1 “Spot-on-lawn” Reporter Screen

1. Preparation of overnight cultures. Set up overnight cultures of the reporter strain and all bacterial strains to be tested, including the positive and negative controls. Employ sterile technique to avoid contamination. Transfer 3 ml of LB medium into a sterile 20–50 ml test tube or universal. Add antibiotics as required, e.g., chloramphenicol from stock (5 mg/ml) to a final concentration of 5 µg/ml for the reporter strain. Inoculate with a single colony of the respective strain from a fresh agar plate. Grow cultures at 37 °C overnight (~16 h) with shaking at 180–220 rpm.
2. Preparation of plates. Warm agar plates in an incubator to 20–30 °C for at least 20 min, or leave at room temperature overnight (*see Note 3*). Melt soft agar, transfer 4 ml of soft agar to a sterile screw-cap container, and allow it to cool down to ~50 °C to prevent killing of the cells. Add 120 µl of the reporter strain overnight culture to the soft agar and mix carefully. Pour the entire mixture onto an agar plate and swirl gently. The agar plate should be evenly covered before the soft agar solidifies (*see Fig. 1a*). Avoid air bubbles. Dry plate for ~20 min or until all condensation has disappeared in order to avoid merging of culture drops in **step 3**.
3. Carefully spot 5 µl drops of the overnight cultures of the strains to be tested, including the controls, onto the soft agar (*see Note 4*; *Fig. 1a–c*). When drops are dry, incubate at 37 °C for 1–7 days (*see Note 5*).

3.2 Luminescence Detection

1. Adjust settings of the FUSION-CAPT™ software by selecting “chemiluminescence” and choosing “full resolution” mode. Open the camera’s iris to maximal aperture.
2. At each time point to be monitored during incubation, place the agar plate with the spot-on-lawn culture in the imaging system (*see Note 6*). Remove the lid of the plate.
3. With epi-white illumination switched on, bring the plate into focus of the camera using preview mode (*see Note 7*; *Fig. 1d*).
4. Switch off the epi-white light and start the chemiluminescence exposure (10 min). Do not open the door or switch on light during exposure.
5. Save the image (*see Fig. 1e*), place the lid on the plate, and return to the incubator until the next measurement is due. Repeat **steps 1–4**, if more than one plate is used for the assay (*see Note 8*).

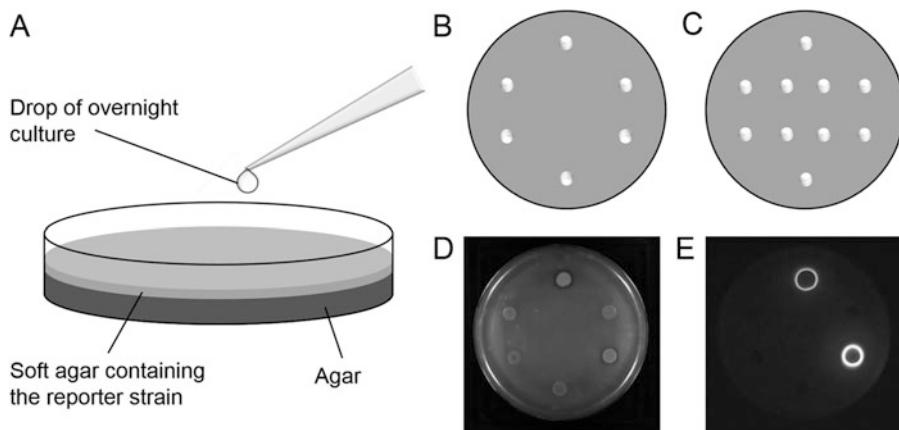


Fig. 1 Preparation of plates. **(a)** The agar plate is covered by a layer of soft agar containing the reporter strain (*light gray*). Culture drops of potential antimicrobial peptide producers are spotted carefully onto the thoroughly dried plate. **(b, c)**: Recommended distribution patterns of culture drops. **(d)** Epi-white light image of a plate with spots of potential producer strains growing on the reporter strain lawn. **(e)** Luminescence image of the same plate shown in panel **(d)**. Production of cell wall-active compounds (spots) induces the *Pial* promotor of the reporter strain (lawn). Induction results in a ring-shaped luminescence signal of the reporter strain surrounding the producer colony

3.3 Luminescence Quantification

1. After starting the ImageJ software, choose your analysis settings (“*Analyze*” → “*Set Measurements...*”). Tick the box “*Min & max gray value*.”
2. Open the image of the first time point (“*File*” → “*Open...*” or type *O* using the keyboard). Identify the precise location of each colony using the epi-white image of the same plate as a reference (see Note 9).
3. Click on the measurement tool depicting a straight line (**straight**, see Fig. 2). Set a line across the ring-shaped luminescence signal around a colony and select “*Measure*” (“*Analyze*” → “*Measure*” or type *M* using the keyboard; see Fig. 2, orange line). The measurement result will be automatically added to the “*Results*” window.
4. Repeat step 3 until 20 independent lines have been measured (see Note 10). Ensure these lines are arranged randomly to measure the different sections of the luminescence circle. Right click on “*Summarize*” to receive the mean maximum value *LS* and the standard deviation ΔLS . Repeat measurements for all other colonies on the plate. Transfer all data to Microsoft Excel® or equivalent software for further calculations.
5. To determine the brightness of luminescence intrinsically produced by the reporter strain, follow the same procedure as in step 4 to set 20 straight lines onto an area of the agar plate

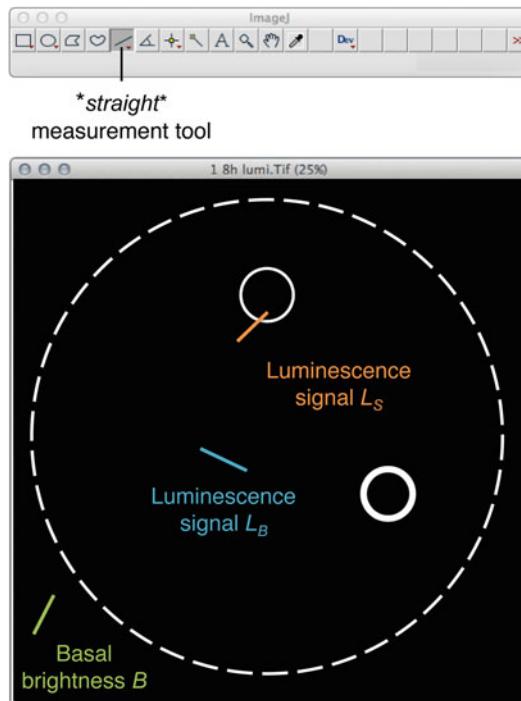


Fig. 2 Luminescence quantification using ImageJ. For determination of the luminescence signal L_S of the reporter strain around the colonies, linear measurements are taken across the luminescent ring (*orange line*). The background luminescence signal L_B , caused by autoinduction of the reporter strain [14], is determined by linear measurements in an area containing no producer colony (*blue line*). For determination of the basal image brightness B , linear measurements are taken on areas that are in the image but not on the plate (*green line*). The edge of the agar plate is indicated by a *dashed white circle*, and luminescence around producer colonies is indicated by *solid white circles*. Measurements are automatically summarized by the software in the “Results” window

image where no colonies are located. These measurements will yield the mean maximum value L_B and the standard deviation ΔL_B needed to correct the signal determined for the producer colonies (*see Note 11*; Fig. 2, blue line).

6. Similarly, determine the mean maximum value and the standard deviation of the basal brightness B of the image by analyzing areas that are still in the image but not on the plate (*see Fig. 2*, green line).
7. To calculate the luminescence signal S over the basal brightness B of the image, subtract the mean maximum value of the basal brightness B of the image from the mean maximum value of the measured luminescence L_S or the plate background L_B .

$$S = L - B$$

8. The error propagation of several error-prone values is determined according to the following formula, simplified for two error-prone values in a subtraction.

$$\Delta S = \Delta L + \Delta B$$

S: Signal over basal brightness; L : Mean maximum value of luminescence (LS or LB); B : Mean of basal brightness; Δ : Standard deviation of each parameter; ΔS : Uncertainty of S .

9. Analysis of the same plate and colonies at multiple time points allows quantification of the signal over time and thus displays the time course of production of cell wall-active compounds (see Fig. 3).

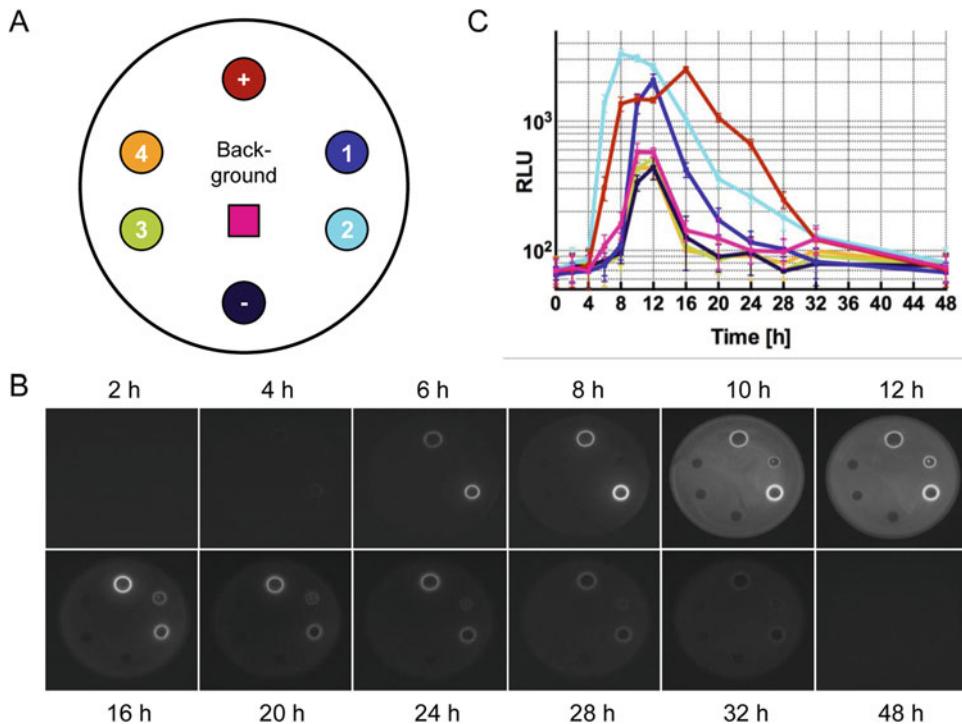


Fig. 3 Time-resolved quantification of antibiotic production. (a) Legend detailing the distribution of spots and identity of tested strains. Numbers indicate candidate producer strains; +, positive control (*B. subtilis* ATCC6633); -, negative control (*B. subtilis* W168); “Background” indicates the quantification of autoluminescence of the reporter strain. Colors match those used in panel (c). (b) Time course of antimicrobial production for example candidate strains 1–4, as well as positive and negative control strains. Spots of the producer strains are grown on a *B. subtilis* *Pial-luxABCDE* reporter lawn. (c) Quantified luminescence of potential producer strains 1–4, the controls, and the autoluminescence of the reporter strain (“Background”) caused by stationary phase induction of the reporter strain [14]. Note that tested strains 1 and 2 produce similar levels of luminescence as the positive control strain (i.e., production of a cell wall active compound), while strains 3 and 4 fail to produce a signal above the negative control or reporter strain background level of luminescence (i.e., no production of a cell wall active compound)

4 Notes

1. Should no sufficiently sensitive equipment for luminescence detection be available, the assays can be performed in a qualitative manner using a reporter strain carrying a *PliaI-lacZ* construct [4]. In this case, both the base agar and soft agar overlay should be supplemented with 100 µg/ml XGal, added after autoclaving from a stock solution of 50 mg/ml XGal dissolved in dimethylformamide. Induction of the reporter will result in the formation of a blue ring surrounding the producer colony. In theory, the intensity of the blue coloration can be quantified with an image analysis software such as ImageJ. However, in contrast to luminescence, which is a transient emission of photons and where the reporter protein has a very short half-life [12], the beta-galactosidase produced by induction of the *lacZ*-reporter is stable and the blue product of XGal cleavage will accumulate over time. It is therefore not possible to perform time-resolved analyses as described for the luciferase reporter in Fig. 3.
2. The *PliaI*-reporter will respond to cell envelope stress and is therefore a nonspecific reporter for production of substances that interfere with cell envelope integrity, including but not restricted to lipid II cycle interfering antibiotics. If a more specific screen is desired, two alternative reporters can be used. The target promoter of the BceRS-BceAB resistance system of *B. subtilis*, *PbceA*, is known to respond to bacitracin, the lantibiotics mersacidin and actagardine, and the fungal defensin plectasin [10]. The target promoter of the paralogues PsdRS-PsdAB system, *PpsdA*, responds to a broad range of cell wall active antimicrobial peptides, such as nisin, subtilin, actagardine, gallidermin, and enduracidin [10]. It is assumed that these lists of inducing substances are not exhaustive and new substrates may be identified from further screens. Fusions of these promoters to the *lux-ABCDE* reporter may be used, following the same procedures as described here, instead of the *PliaI* reporter to provide some further information on the nature of the substance produced by the strains under investigation.
3. If the agar plate is too cold, the soft agar will solidify before it is evenly spread onto the plate.
4. Culture spots should be distributed equally over the plate (see Fig. 1b), as their location on the plate and their proximity to each other could affect the luminescence signal. For an initial screening of producer strains, a higher number of cultures can be tested on a single plate (see Fig. 1c). To avoid bacterial aerosols and splashes, gently push the drop halfway out of the pipette tip before placing the drop carefully on the soft agar.

5. Plates may dry out when using a small incubator or longer incubation times. Therefore, it is helpful to wrap the plates in foil.
6. Time points depend on the growth rate of the tested strains and the desired time lapse resolution of antibiotic production and therefore may vary from 2 h intervals to daily measurements over 1–7 days. An initial screening helps to choose the optimal time points and overall duration of experiments.
7. Saving an image of the plate with the light switched on will help to remember the orientation of the plate in case of no or only low luminescence.
8. It might be necessary to enhance the brightness of the image to be able to see the luminescence on screen.
9. Do not adjust the image brightness before or during the analysis. This is essential to allow comparison of the signal strength between the time points.
10. The lines should be distributed around the luminescent ring, as the luminescence signal can be uneven around a colony. The precise length of the measurement lines is not of importance, because only the maximum value will be used for further analysis. Information of angle and length is displayed in the status bar or can be checked in the “Results” table.
11. The *PliaI* promoter was found to be autoinduced during transition of *B. subtilis* to stationary phase [14], which results in a transiently increased luciferase activity of the reporter strain in this assay. It is necessary to consider the effect of autoinduction to avoid false positive signals during luminescence quantification. Determination of the background luminescence signal, *LB*, allows appropriate correction of the actual reporter signal, *LS*, obtained for producers of cell wall-active compounds.

References

1. Nic M, Jirat J, Kosata B (2014) IUPAC compendium of chemical terminology—the gold book. <http://goldbook.iupac.org>. Accessed 15 Dec 2015
2. Park M, Tsai S-L, Chen W (2013) Microbial biosensors: engineered microorganisms as the sensing machinery. Sensors 13:5777–5795
3. Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
4. Mascher T, Zimmer SL, Smith T-A, Helmann JD (2004) Antibiotic-inducible promoter regulated by the cell envelope stress-sensing two-component system LiaRS of *Bacillus subtilis*. Antimicrob Agents Chemother 48:2888–2896
5. Kremers G-J, Gilbert SG, Cranfill PJ, Davidson MW, Piston DW (2011) Fluorescent proteins at a glance. J Cell Sci 124:157–160
6. Jordan S, Junker A, Helmann JD, Mascher T (2006) Regulation of LiaRS-dependent gene expression in *Bacillus subtilis*: identification of inhibitor proteins, regulator binding sites, and target genes of a conserved cell envelope stress-sensing two-component system. J Bacteriol 188:5153–5166
7. Schrecke K, Jordan S, Mascher T (2013) Stoichiometry and perturbation studies of the LiaFSR system of *Bacillus subtilis*. Mol Microbiol 87:769–788. doi:[10.1111/mmi.12130](https://doi.org/10.1111/mmi.12130)

8. Cao M, Wang T, Ye R, Helmann JD (2002) Antibiotics that inhibit cell wall biosynthesis induce expression of the *Bacillus subtilis* σ W and σ M regulons. Mol Microbiol 45:1267–1276
9. Mascher T, Margulis NG, Wang T, Ye RW, Helmann JD (2003) Cell wall stress responses in *Bacillus subtilis*: the regulatory network of the bacitracin stimulon. Mol Microbiol 50: 1591–1604
10. Staroń A, Finkeisen DE, Mascher T (2011) Peptide antibiotic sensing and detoxification modules of *Bacillus subtilis*. Antimicrob Agents Chemother 55:515–525
11. Helmann JD, Mascher T (2005) Compositions and methods for screening of antibacterial compounds. US Patent 7309484
12. Radeck J, Kraft K, Bartels J, Cikovic T, Dürr F, Emenegger J, Kelterborn S, Sauer C, Fritz G, Gebhard S, Mascher T (2013) The *Bacillus* BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. J Biol Eng 7:29
13. Schmalisch M, Maiques E, Nikolov L, Camp AH, Chevreux B, Muffler A, Rodriguez S, Perkins J, Losick R (2010) Small genes under sporulation control in the *Bacillus subtilis* genome. J Bacteriol 192:5402–5412
14. Jordan S, Rietkötter E, Strauch MA, Kalamorz F, Butcher BG, Helmann JD, Mascher T (2007) LiaRS-dependent gene expression is embedded in transition state regulation in *Bacillus subtilis*. Microbiology 153:2530–2540

Chapter 8

Determination of Bacterial Membrane Impairment by Antimicrobial Agents

Miriam Wilmes and Hans-Georg Sahl

Abstract

The bacterial cytoplasmic membrane separates the cell from its environment and acts as a selective permeability barrier. In addition, it functions in energy conservation, transport, and biosynthesis processes. Antimicrobial agents disrupting these functions may lead to pleiotropic effects, including leakage of low molecular weight compounds such as ions, amino acids and ATP, and subsequent membrane depolarization. This article describes two techniques to assess antibiotic-induced membrane impairment *in vivo*.

Key words Membrane permeabilization, Membrane potential, Depolarization, Tetraphenyl phosphonium bromide, Potassium efflux

1 Introduction

Many antimicrobial peptides (AMPs) display their activity by impairing the membrane barrier function via pore-formation or unspecific membrane permeabilization. For example, the lantibiotic nisin uses lipid-linked cell envelope precursors such as lipid II as docking molecule to form pores in the membrane of susceptible strains [1, 2]. Pore formation results in rapid efflux of small molecules from the cells and subsequently dissipation of the membrane potential [3].

In case of the lantibiotic Pep5 and mammalian θ -defensins the degree of membrane impairment strongly depends on the level of the membrane potential across the bacterial membrane [4, 5].

Here, we describe two methods for measuring the impact of antimicrobial substances on the bacterial membrane integrity *in vivo*: (1) determination of the bacterial membrane potential ($\Delta\psi$) using [3 H]tetraphenylphosphonium bromide and (2) potassium release from whole cells by means of a potassium-sensitive electrode.

Lipophilic cations such as tetraphenylphosphonium bromide (TPP $^+$) can easily pass through phospholipid bilayers in response to a trans-negative membrane potential ($\Delta\psi$) and were first used to investigate the mitochondrial oxidative phosphorylation [6–8].

Since then, these molecules have also been exploited to measure the membrane potential in bacteria [9, 10] and the influence of antimicrobials such as bacteriocins on $\Delta\psi$ [4, 11, 12]. For this, growing bacteria are incubated with radiolabeled TPP⁺ and the distribution of the cation inside and outside the cells is determined. The intracellular TPP⁺ concentration is calculated by the use of the internal cell volume (V_i) and by taking into account unspecific membrane binding. Intra- and extracellular TPP⁺ concentrations are then inserted into the Nernst equation for $\Delta\psi$ determination. To evaluate the effect of an antimicrobial agent on the membrane potential, the protonophore CCCP (carbonyl cyanide m-chlorophenylhydrazone) can be used as a positive control. CCCP uncouples the proton gradient across the cytoplasmic membrane leading to fast membrane depolarization and rapid TPP⁺ release from the cells.

Alternatively, ion-selective electrodes can be used to evaluate the membrane impairment by antimicrobials. Potassium is the major intracellular cation in bacteria, e.g., with a concentration of 180–200 mM in *Escherichia coli* [13] and up to 1 M in *Staphylococcus aureus* [14]. Hence, efflux of K⁺ from bacterial cells in the presence of an antibiotic indicates a membranolytic effect. Extracellular K⁺ concentrations can be calculated from the measured voltage according to Orlov et al. [15] and expressed relative to the total amount of potassium present in the cells. This method has the advantage that it is simple and broadly applicable. Moreover, it does not rely on radiolabeled substances and the measurement can be performed in real time.

2 Material

Sterilize the culture media and buffer before use. Use sterile glass ware and pipette tips.

2.1 Determination of Bacterial Membrane

Potential Using [³H]tetraphenylphosphonium Bromide (TPP⁺)

2.1.1 Medium, Strain, and Antimicrobial Substances

2.1.2 Components for Bacterial Membrane Potential Determination

1. Test strain, e.g., *Staphylococcus simulans* 22 or *S. aureus* (see Note 1).
 2. Culture medium, e.g., Mueller Hinton Broth (300 g beef infusion, 17.5 g casein hydrolysate, 1.5 g starch; pH 7.3) (see Note 2).
 3. Antibiotic or antimicrobial peptide of interest.
 4. Carbonyl cyanide 3-chlorophenylhydrazone (CCCP); prepare 10 mM stock solution in ethanol and store at –20 °C.
-
1. [³H]tetraphenylphosphonium bromide (TPP⁺).
 2. Cellulose acetate filters with a pore size of 0.2 µm (Whatman™).
 3. 50 mM Potassium phosphate buffer, pH 7. Prepare a 50 mM solution of the dibasic (K₂HPO₄) and monobasic component (KH₂PO₄) and mix both until you reach the desired pH.

- 4. n-Butanol.
 - 5. Glass vacuum filtration apparatus.
 - 6. Spectrophotometer at 600 nm.
 - 7. Shaking water bath.
 - 8. Scintillation fluid.
 - 9. Liquid scintillation counter.
- 2.1.3 Components for Protein Determination of Whole Cells**
- 1. Bacterial protein extraction reagent (B-PERTM; ThermoFisher Scientific).
 - 2. BCA Protein Assay Kit.
 - 3. Benchtop centrifuge.
- 2.2 Measurement of Potassium Release from Whole Cells**
- 2.2.1 Media, Strain, and Antimicrobial Substances**
- 1. Test strain, e.g., *Staphylococcus simulans* 22 or *Lactococcus lactis*.
 - 2. Culture medium, e.g., tryptic soy broth (17 g pancreatic digest of casein, 3 g pancreatic digest of soy bean, 5 g NaCl, 2.5 g K₂HPO₄, 2.5 g glucose; pH 7.3).
 - 3. Antibiotic or antimicrobial peptide of interest.
 - 4. Membrane-active antimicrobial as positive control, e.g., nisin.
- 2.2.2 Components for Potassium Efflux Measurement**
- 1. Potassium electrode (stored in 0.1 M KCl) (MI-442, Microelectrodes Inc, Bedford, USA).
 - 2. Reference electrode (filled with 3 M KCl saturated with AgCl) (MI-409F, Microelectrodes Inc, Bedford, USA).
 - 3. Microprocessor pH meter (HANNA[®] Instruments; Kehl am Rhein, Germany).
 - 4. Choline buffer (300 mM choline chloride, 30 mM 2-(*N*-morpholino)ethanesulfonic acid, 20 mM Tris base; pH 6.5).
 - 5. Potassium standard solutions: 0.01, 0.1, and 1 mM KCl dissolved in choline buffer.
 - 6. Centrifuge.
 - 7. Magnetic stirrer.
 - 8. 0.7% octylglucoside.

3 Method

3.1 Determination of Bacterial Membrane Potential Using [³H]tetraphenylphosphonium Bromide (TPP⁺)

3.1.1 Measurement of TPP⁺ Uptake and Distribution

- 1. Inoculate a culture of your test strain—using a 2% inoculum (v/v) from an overnight culture—and grow it to an optical density at 600 nm (OD₆₀₀) of 0.5–0.6 in a water bath with constant shaking.
- 2. To monitor the membrane potential, add 1 µCi/ml of [³H]TPP⁺ to the culture and incubate it for 30 s (see Note 2).
- 3. Transfer two samples of 100 µl directly to liquid scintillation vials for determining the total amount of radioactivity.

4. Filter a 100 µl sample through a cellulose acetate filter and wash the filter twice with 5 ml potassium phosphate buffer (*see Note 3*). Transfer the filter to a liquid scintillation vial and let it dry.
5. Split the culture into three aliquots. Treat one aliquot (i) with n-butanol for measuring the unspecific binding of [³H]TPP⁺ to the cells, treat the second aliquot (ii) with the antibiotic of interest, and run the third aliquot (iii) as a control (Fig. 1).
6. For aliquot (i), add n-butanol to the cells (10% final concentration) and mix it well by repeatedly sucking the sample into a 1 ml pipette. Take four 100 µl samples and filter them as described above (**step 4**).
7. For aliquot (ii), add the antimicrobial of interest at a given concentration (e.g., at 5× or 10× MIC, minimal inhibitory concentration) and immediately take 100 µl of the culture and filter it as described above (*see Note 4*). Take further samples at certain time points, e.g., 1, 2, 4, 5, 10, 15, and 20 min after antibiotic addition (*see Note 5*). Additionally, measure the OD₆₀₀ of the culture periodically (Fig. 1).
8. For aliquot (iii), filter 100 µl samples as described above (**step 4**) and measure the OD₆₀₀, e.g., at time points 3, 8, 13, 18, and 23 min (Fig. 1).
Optional: At the end of the experiment, 10 µM CCCP can be added as a positive control to aliquot (iii). Δψ is dissipated and the intracellular TPP⁺ concentration will decrease rapidly.
9. Add 5 ml scintillation fluid into all liquid scintillation vials.
10. Measure the radioactivity in the samples with a liquid scintillation counter for 5 min per filter.

3.1.2 Protein Determination of Whole Cells

1. Grow a culture of your test strain to an OD₆₀₀ of 1.
2. Centrifuge 2 × 1 ml of the culture (9200 × g, 5 min).
3. Wash the pellets with potassium phosphate buffer and centrifuge again.
4. Resuspend each pellet in 100 µl B-PER™ and incubate it for 10–15 min at room temperature. Optional: Freeze the cells before extraction to enhance cell lysis.
5. Determine the total protein concentration in the lysate by using the BCA Protein Assay.

3.1.3 Calculation of Bacterial Membrane Potential

1. Calculate the internal and external [³H]TPP⁺ concentration for each time point using the following formulas.
2. Correct the counts for unspecific binding of [³H]TPP⁺ by subtracting the radioactivity of the butanol-treated aliquots.

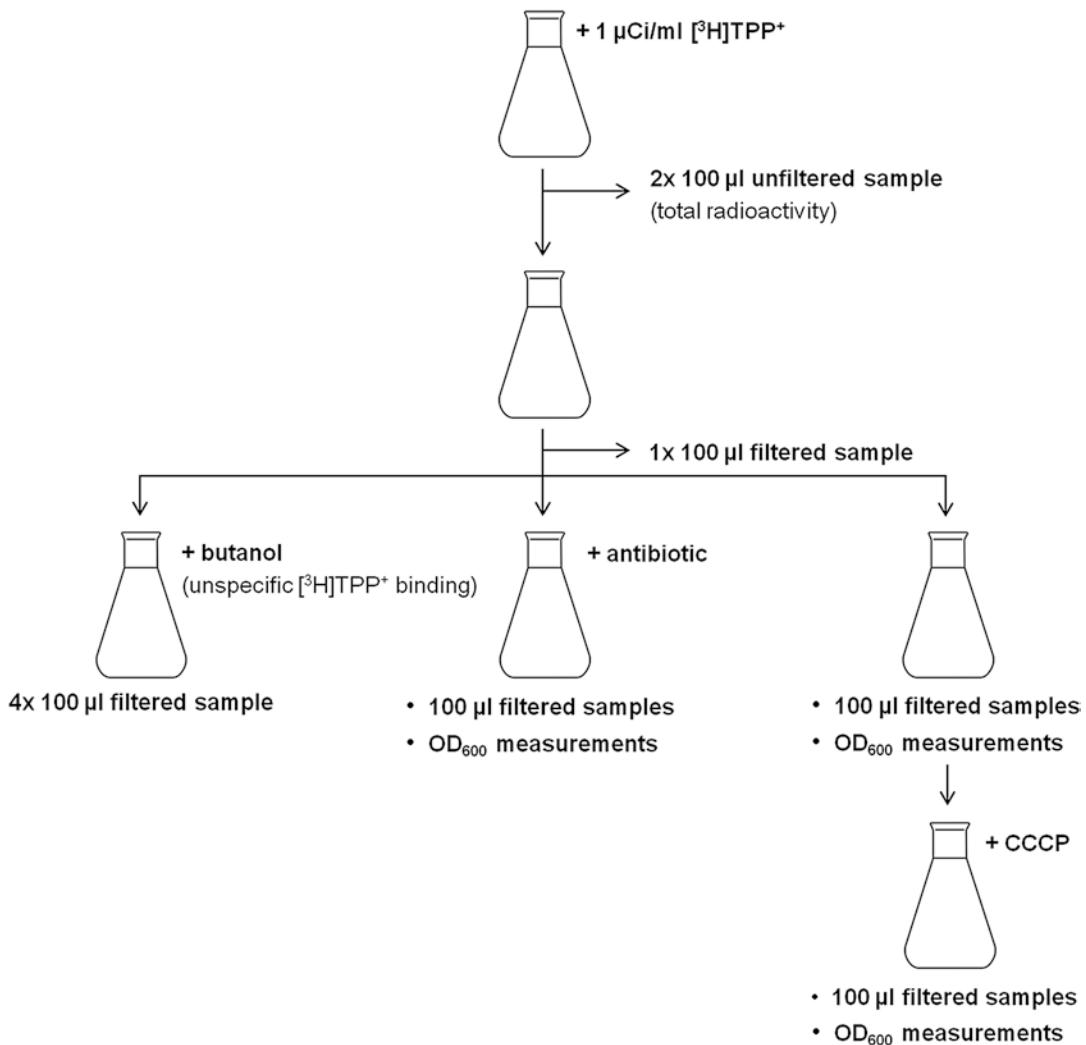


Fig. 1 Experimental scheme. $[{}^3\text{H}]TPP^+$ is added to an exponentially growing culture. After taking two samples for determining the total radioactivity, the culture is split into three aliquots. The first aliquot is treated with butanol to measure the unspecific binding of TPP^+ to the cells, the second aliquot is treated with the antibiotic of interest, and the third aliquot is run as a control. At given time points, $100 \mu\text{l}$ samples are filtered and the OD₆₀₀ is measured. CCCP can be used as a positive control

3. Calculate V_i ($\mu\text{l}/\text{ml}$ cells) by taking into account the determined protein concentration (Subheading 3.1.2) and the measured OD₆₀₀ values (Subheading 3.1.1). For example, for *S. simulans* 22 the inner volume was found to be $3.4 \mu\text{l}/\text{mg}$ cell protein [4]. Thus, V_i is $0.2 \mu\text{l}/\text{ml}$ when the determined protein concentration is $0.1 \text{ mg}/\text{ml}$ and the measured OD₆₀₀ is 0.6.

$$TPP_{\text{in}}^+ = \frac{(\text{cpm}_{\text{sample}} - \text{cpm}_{\text{BuOH}}) \times M_{TPP^+} \times 1000}{(\text{cpm}_{\text{total}} - \text{cpm}_{\text{BuOH}}) \times V_i} [\mu\text{M}]$$

$$\text{TPP}_{\text{out}}^+ = \frac{\left[(\text{cpm}_{\text{total}} - \text{cpm}_{\text{BuOH}}) - (\text{cpm}_{\text{sample}} - \text{cpm}_{\text{BuOH}}) \right] \times M_{\text{TPP}^+}}{(\text{cpm}_{\text{total}} - \text{cpm}_{\text{BuOH}})} [\text{M}]$$

TPP_{in}^+ : intracellular TPP^+ concentration; $\text{TPP}_{\text{out}}^+$: extracellular TPP^+ concentration; cpm_{BuOH} : counts per minute in the butanol control (aliquot i; mean value); $\text{cpm}_{\text{sample}}$: counts per minute in the filtered sample (aliquot ii or aliquot iii); $\text{cpm}_{\text{total}}$: counts per minute in the unfiltered sample (mean value); M_{TPP^+} : molarity of TPP^+ (μM); V : internal volume of 1 ml cells ($\mu\text{l/ml}$).

4. Insert the calculated values for the intra- and extracellular TPP^+ concentration into the Nernst equation to determine $\Delta\psi$.

$$\Delta\psi = \frac{-2.3 \times R \times T}{F} \times \log \frac{\text{TPP}_{\text{in}}^+}{\text{TPP}_{\text{out}}^+} [\text{mV}]$$

R : universal gas constant $\left(8.314 \frac{\text{J}}{\text{mol} \times \text{K}} \right)$; T : absolute temperature (K); F : Faraday constant $\left(96,485 \frac{\text{C}}{\text{mol}} \right)$

5. Plot the values of the calculated membrane potential (mV) against time (min).

Two examples of a typical experiment are shown in Fig. 2.

3.2 Measurement of Potassium Release from Whole Cells

3.2.1 Measurement of Potassium Efflux

1. Inoculate a 50 ml culture of your test strain—using a 2% inoculum (v/v) from an overnight culture—and grow it to an OD_{600} of 1–1.5 (see Note 6).
2. Harvest the bacteria by centrifugation ($2300 \times g$, 3 min, 4 °C).
3. Wash the cells with 25 ml prechilled choline buffer and centrifuge again (step 2).
4. Resuspend the cells in choline buffer to a final OD_{600} of 30 and keep them on ice until further use (see Note 7). For each measurement, dilute 200 μl cells in 1.8 ml choline buffer (final OD_{600} of 3) and gently agitate the culture by using a magnetic stirrer.
5. Calibrate the electrodes (see Note 8) with the potassium standard solutions starting with the lowest concentration. Measure five to ten values for each concentration.
6. Rinse both electrodes with distilled water and place them into the stirring culture. Monitor the potassium release for 5 min at room temperature. Collect voltage data every 10 s. Start with the untreated control to determine the K^+ concentration in the buffer ($\text{K}_{\text{initial}}^+$).

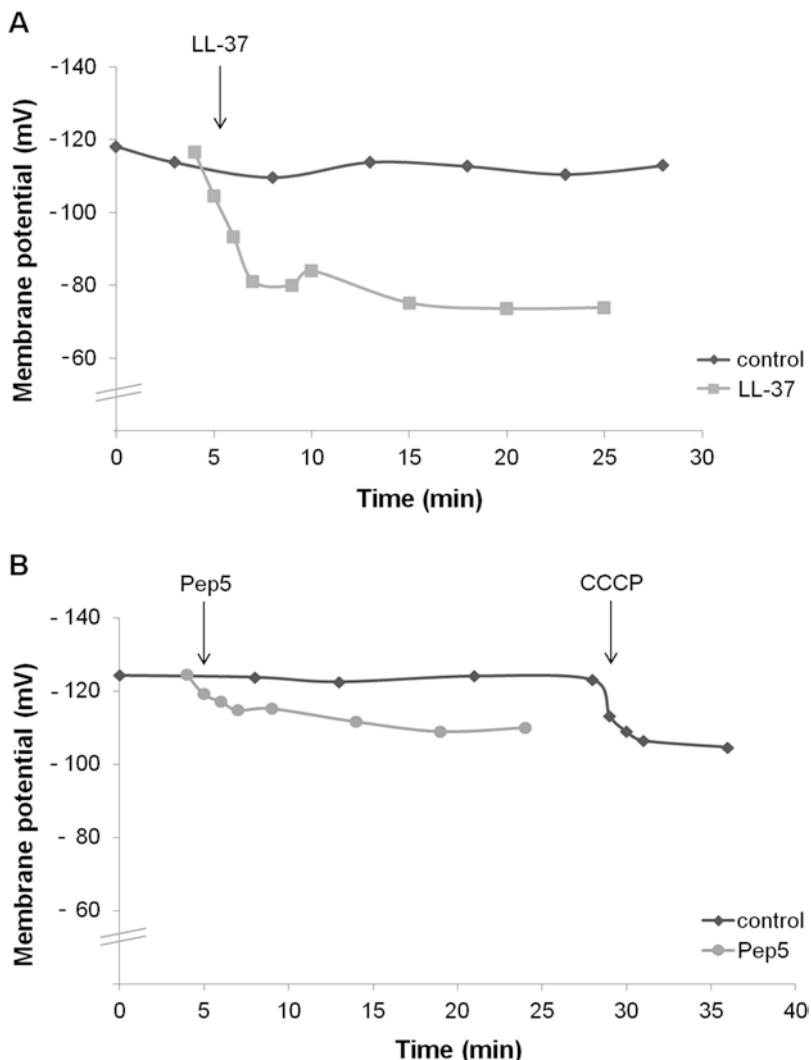


Fig. 2 Representative examples of a membrane potential measurement in the presence of an AMP. **(a)** Membrane potential of *S. aureus* SG511-Berlin in half-concentrated Mueller-Hinton broth (MHB). The human host defense peptide LL-37 was added at 5× MIC. Immediately, a rapid decrease of the membrane potential was detected. In contrast, no significant changes of the membrane potential were observed in the untreated control cells. **(b)** Membrane potential of *S. aureus* SA113 in half-concentrated MHB supplemented with 10 mM glucose. Bacteria were exposed to 10× MIC of the lantibiotic Pep5. CCCP (10 μM) was used as positive control. Both compounds induced some depolarization of the bacterial membrane

7. Start another measurement (as described in **step 6**) and induce complete potassium release (K_{total}^+) by treatment with a highly membrane-active antibiotic, e.g., 1 μM nisin (*see Note 9*).
8. Measure the membranolytic effect of your antibiotic of interest, e.g., by adding it at 5× or 10× MIC to the cells (**step 6**).
9. At the end of the experiment, wash the electrodes with distilled water and a detergent (e.g., 0.7% octylglucoside).

3.2.2 Calculation of Released Potassium Concentration

- Generate a linear standard curve of the calibration data (mean value for each concentration) to determine the slope “ m ” and the y-intercept “ z ” of the following formula, which relates the measured electrode voltage (V_{meas}) to the extracellular K^+ concentration (Fig. 3a).

$$V_{\text{meas}} = m \log_{10} [\text{K}^+] + z$$

- Calculate the initial K^+ concentration ($\text{K}_{\text{initial}}^+$) in the buffer (from your data of the untreated control) and the total K^+ concentration ($\text{K}_{\text{total}}^+$), e.g., after nisin treatment, from the measured voltages.

$$\text{K}^+ = 10^{\frac{V_{\text{meas}} - z}{m}}$$

- Finally, convert the obtained data ($\text{K}_{\text{sample}}^+$) to percent potassium release and plot the % potassium release against time (s).

$$\% \text{ release} = \frac{\text{K}_{\text{sample}}^+ - \text{K}_{\text{initial}}^+}{\text{K}_{\text{total}}^+ - \text{K}_{\text{initial}}^+} \times 100$$

An example of a typical experiment is shown in Fig. 3b.

4 Notes

- The membrane potential measurement using TPP^+ was established for some Gram-positive bacteria such as *Lactococcus lactis* [16], *Bacillus subtilis* [17], and *S. simulans* [4] but may also work with other species. However, determination of the membrane potential requires estimates of the inner aqueous volume of the cells (Subheading 3.1.3) which has to be defined for the particular strain. Additionally, in Gram-negative bacteria the permeability to TPP^+ is greatly reduced due to the presence of the outer membrane. Thus, cells have to be pretreated with EDTA [9, 12, 18] or lipophilic cation-permeable mutants have to be used as test strain [19].
- Since $\Delta\psi$ and ΔpH are two independent components of the proton motive force ($\Delta p = \Delta\psi - 59\Delta\text{pH}$), it is recommended to perform the measurement at neutral pH to keep the pH difference between the cytoplasm and the exterior of the cells low. $\Delta\psi$ may be transiently increased by addition of a suitable carbon source, e.g., 10 mM glucose. This is relevant when membrane action of a compound is dependent on a certain magnitude of $\Delta\psi$ as it has been described for AMPs such as Pep5 [4] and θ -defensins [5].
- It is recommended to add the sample and 5 ml potassium phosphate buffer simultaneously into the filtration apparatus.

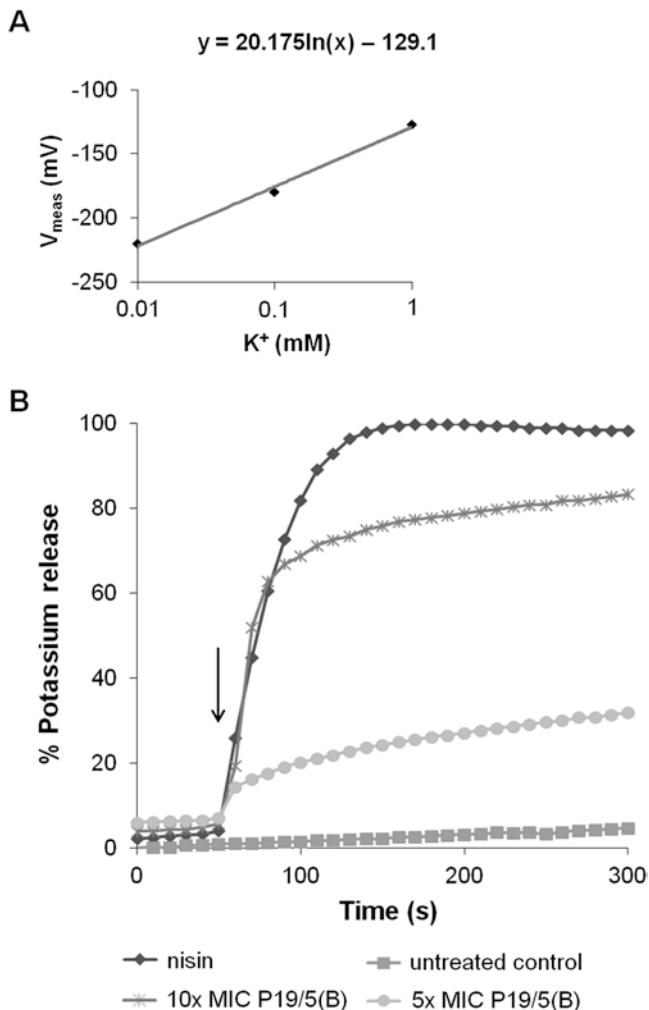


Fig. 3 Measurement of antibiotic-induced potassium efflux. **(a)** Example of a typical electrode calibration curve ($m=20.175$, $z=-129.1$). **(b)** Effect of the antimicrobial peptide P19/5(B) on K^+ release of *S. simulans* 22. Ion leakage was expressed relative to the amount of potassium released after the addition of 1 μ M of the pore-forming lantibiotic nisin (100 % efflux). The arrow indicates the moment of peptide addition

After the buffer/sample is flown through the filter, wash it again with 5 ml potassium phosphate buffer.

4. Optional: Take another 100 μ l sample before addition of the antibiotic.
5. The membrane potential decreases rapidly in the presence of a membrane-active compound (Fig. 2). Thus, it is recommended to take several samples in the first 5 min after antibiotic addition.

6. A 50 ml culture will be sufficient for measuring 6–8 samples in one experiment.
7. The bacteria dissolved in choline buffer may start lysing after a while. It is recommended to perform the experiment within 30–60 min. In addition, it may be necessary to energize the cells by addition of a suitable carbon source, e.g., 10 mM glucose.
8. It is recommended to store both electrodes in choline buffer for at least 1 h before starting the experiment.
9. Alternatively, the bacteria can be disrupted by prolonged sonication to determine the total K⁺ concentration [15].

References

1. Wiedemann I, Breukink E, van Kraaij C, Kuipers OP, Bierbaum G, de Kruijff B, Sahl HG (2001) Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *J Biol Chem* 276:1772–1779
2. Müller A, Ulm H, Reder-Christ K, Sahl HG, Schneider T (2012) Interaction of type A lantibiotics with undecaprenol-bound cell envelope precursors. *Microb Drug Resist* 18:261–270
3. Ruhr E, Sahl HG (1985) Mode of action of the peptide antibiotic nisin and influence on the membrane potential of whole cells and on cytoplasmic and artificial membrane vesicles. *Antimicrob Agents Chemother* 27:841–845
4. Sahl HG (1985) Influence of the staphylococcal-like peptide Pep 5 on membrane potential of bacterial cells and cytoplasmic membrane vesicles. *J Bacteriol* 162:833–836
5. Wilmes M, Stockem M, Bierbaum G, Schlag M, Götz F, Tran DQ, Schaal JB, Ouellette AJ, Selsted ME, Sahl HG (2014) Killing of staphylococci by theta-defensins involves membrane impairment and activation of autolytic enzymes. *Antibiotics (Basel)* 3:617–631
6. Liberman EA, Topaly VP, Tsafina LM, Jasaitis AA, Skulachev VP (1969) Mechanism of coupling of oxidative phosphorylation and the membrane potential of mitochondria. *Nature* 222:1076–1078
7. Grinius LL, Jasaitis AA, Kadziauskas YP, Liberman EA, Skulachev VP, Topali VP, Tsafina LM, Vladimirova MA (1970) Conversion of biomembrane-produced energy into electric form. I. Submitochondrial particles. *Biochim Biophys Acta* 216:1–12
8. Bakeeva LE, Grinius LL, Jasaitis AA, Kuliene VV, Levitsky DO, Liberman EA, Severina II, Skulachev VP (1970) Conversion of biomembrane-produced energy into electric form. II. Intact mitochondria. *Biochim Biophys Acta* 216:13–21
9. Schuldiner S, Kaback HR (1975) Membrane potential and active transport in membrane vesicles from *Escherichia coli*. *Biochemistry* 14:5451–5461
10. Szmelcman S, Adler J (1976) Change in membrane potential during bacterial chemotaxis. *Proc Natl Acad Sci U S A* 73:4387–4391
11. Tokuda H, Konisky J (1978) Mode of action of colicin Ia: effect of colicin on the *Escherichia coli* proton electrochemical gradient. *Proc Natl Acad Sci U S A* 75:2579–2583
12. Weiss MJ, Luria SE (1978) Reduction of membrane potential, an immediate effect of colicin K. *Proc Natl Acad Sci U S A* 75:2483–2487
13. Shabala L, Bowman J, Brown J, Ross T, McMeekin T, Shabala S (2009) Ion transport and osmotic adjustment in *Escherichia coli* in response to ionic and non-ionic osmoticities. *Environ Microbiol* 11:137–148
14. Baba T, Takeuchi F, Kuroda M, Ito T, Yuzawa H, Hiramatsu K (2004) The *Staphylococcus aureus* genome. In: Aldeen DA, Hiramatsu K (eds) *Staphylococcus aureus*: molecular and clinical aspects. Horwood Publishing, Chichester, UK, pp 66–145
15. Orlov DS, Nguyen T, Lehrer RI (2002) Potassium release, a useful tool for studying antimicrobial peptides. *J Microbiol Methods* 49:325–328
16. Kashket ER, Blanchard AG, Metzger WC (1980) Proton motive force during growth of *Streptococcus lactis* cells. *J Bacteriol* 143:128–134

17. Miller JB, Koshland DE Jr (1977) Sensory electrophysiology of bacteria: relationship of the membrane potential to motility and chemotaxis in *Bacillus subtilis*. Proc Natl Acad Sci U S A 74:4752–4756
18. Griniuviene B, Chmieliauskaite V, Grinius L (1974) Energy-linked transport of permeant ions in *Escherichia coli* cells: evidence for membrane potential generation by proton-pump. Biochem Biophys Res Commun 56:206–213
19. Hirota N, Matsuura S, Mochizuki N, Mutoh N, Imae Y (1981) Use of lipophilic cation-permeable mutants for measurement of transmembrane electrical potential in metabolizing cells of *Escherichia coli*. J Bacteriol 148:399–405

Chapter 9

Mass-Sensitive Biosensor Systems to Determine the Membrane Interaction of Analytes

Sebastian G. Hoß and Gerd Bendas

Abstract

Biosensors are devices that transform a biological interaction into a readout signal, which is evaluable for analytical purposes. The general strength of biosensor approaches is the avoidance of time-consuming and cost-intensive labeling procedures of the analytes. In this chapter, we give insight into a mass-sensitive surface-acoustic wave (SAW) biosensor, which represents an elegant and highly sensitive method to investigate binding events at a molecular level. The principle of SAW technology is based on the piezoelectric properties of the sensors, so as to binding events and their accompanied mass increase at the sensor surface are detectable by a change in the oscillation of the surface acoustic wave. In combination with model membranes, transferred to the sensor surface, the analytical value of SAW biosensors has strongly been increased and extended to different topics of biomedical investigations, including antibiotic research. The interaction with the bacterial membrane or certain target structures therein is the essential mode of action for various antibacterial compounds. Beside targeted interaction, an unspecific membrane binding or membrane insertion of drugs can contribute to the antibacterial activity by changing the lateral order of membrane constituents or by interfering with the membrane barrier function. Those pleiotropic effects are hardly to illustrate in the bacterial systems and need a detailed view at the in vitro level. Here, we illustrate the usefulness of a SAW biosensor in combination with model membranes to investigate the mode of membrane interaction of antibiotic active peptides. Using two different peptides we exemplary describe the interaction analysis in a two-step gain of information: (1) a binding intensity or affinity by analyzing the phase changes of oscillation, and (2) mode of membrane interaction, i.e., surface binding or internalization of the peptide by following the amplitude of oscillation.

Key words Biosensors, Surface acoustic wave (SAW), Model membranes

1 Introduction

Biosensors have attracted much attention during the last two decades in biosciences in light of their potential to obtain a label-free detection of biological recognition processes. Biosensors can be classified according to their principles to transform a biological event into detectable readouts, e.g., optical or electrical signals. The most established biosensor technique in biomedical research utilizes the optical phenomenon of surface plasmon resonance

(SPR). These sensors are commonly used for kinetic analysis of versatile compound-target interaction [1]. Here, we use another type of mass-sensitive biosensor, namely surface acoustic wave (SAW) sensors. SAW sensors have been developed during the last years as powerful and promising systems for detecting various biological recognition events, e.g., protein-protein, protein-nucleic acid, or cell-virus interactions [2–7]. This technique is based on piezoelectric properties of quartz sensors. Applying an electrical field to gold coated ST-cut quartz crystal slides, a Love-shear wave is generated at a thin (5 µm) guiding layer directly deposited at the sensor surface [2]. Consequently, binding events can be detected by changes of the physical properties of the shear wave in two different ways: (1) attachment of components equivalent to an increased mass leads to an angular phase shift, and (2) the viscoelastic properties of the bound analytes were reflected by changes in the oscillation amplitude and thus, give insights into the mode of analyte attachment.

Biological membranes possess not only essential barrier functions to compartmentalize cellular and subcellular components, they are also crucial elements of cellular recognition, communication, or transport [8]. In the field of antibiotic research, bacterial membranes are the most important point of attack for antibiotics, (1) either indirectly by an unspecific attachment for initial contacts and subsequently internalization, (2) by affecting the barrier properties, or (3) directly by targeting certain membrane components to interfere with essential cellular activities, such as cell wall biosynthesis [9, 10].

However, in light of the complex nature of natural cell membranes, the simplification of bacterial membranes by the use of model membranes appears as a promising strategy. To combine model membrane approaches with the above-mentioned SAW biosensor technology, we have recently developed a drying and conservation technique of model membranes at SAW sensor chips [11]. This allows, e.g., for kinetic binding investigations of different components within a simulated membrane compartment [12].

Furthermore, a well-defined model membrane at the sensor surface appears as a suitable screening tool to investigate the intrinsic capacity for membrane interactions of various compounds. Referring to antibiotic agents, the intensity and probably the mechanisms of membrane interaction should be illustrated by those investigations as a helpful contribution to interpret the mode of action. This strategy is presented here using two linear peptide structures of comparable molecular weight (~300 Da) and a net negative charge at neutral pH, both of them have been examined as experimental antibiotic active components. We made both compounds anonymous referred as “Compound A” and “Compound B” to focus the view solely on the way of membrane interaction. The SAW sensor device, used here is a sam® 5 Blue, SAW

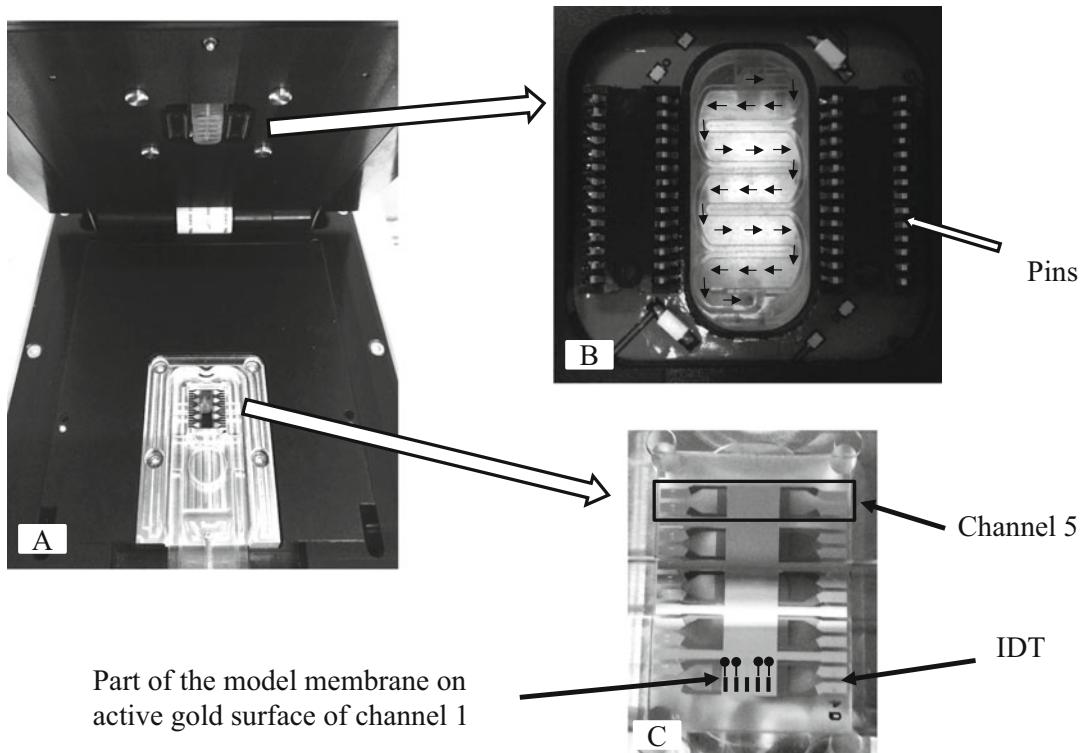


Fig. 1 (a) SAW-sensor chip mounted on the flow cell compartment of the sensor device (overview). (b) Detail of the flow cell with the electronic interface. Direction of the buffer flow is indicated by arrows. Due to the barriers of the flow cell, five compartments are formed. (c) Detail of the mounted chip and active surface in the center. During measurement, the flow cell is pressed on the sensor chip and their compartments establish five individual channels. Oscillation is initiated and the signals were collected with the help of interdigital transducers (IDT)

Instruments GmbH Bonn, now part of NanoTemper Technologies, Munich. The essential parts of the device, the sensor quartz, and the flow chamber with embedded sensor quartz are illustrated in Fig. 1.

2 Materials

2.1 Model Membrane Preparation

1. 1 mM 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-phospho-(1'-RAC-glycerol) (sodium salt) (DOPG) (Avanti Polar Lipids Inc., Alabaster, AL, USA) stock solution in chloroform.
2. 1.66 µM D-(+)-trehalose dihydrate stock solution in ultrapure water.
3. 10 mM 1-hexadecanethiol dissolved in anhydrous chloroform.
4. 10 mM 1-hexadecanoyl-2-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine (POPC) (Avanti Polar Lipids Inc., Alabaster, AL, USA) stock solution in chloroform.

5. Chloroform, anhydrous.
6. Compressed air.
7. Ethanol.
8. Langmuir-Blodgett trough (trough and damper made of Teflon®).
9. SAW-sensor quartzes (NanoTemper Technologies, Munich, Germany).
10. Teflon® tweezers.
11. Ultrapure water (“Type I”).

2.2 Running Buffer

1. 200 mM MOPS (3-morpholinopropanesulfonic acid) buffer grade, pH 7.0 (adjusted with 10 mM NaOH) (*see Note 1*).
2. Calcium chloride dihydrate.

2.3 Cleaning of Quartz Sensors

1. 35 % hydrogen peroxide.
2. Compressed air.
3. Conc. sulfuric acid.
4. Piranha solution (30% hydrogen peroxide and concentrated sulfuric acid, ratio 1–3 as described in Subheading 3.2).

2.4 Membrane Interaction of Two Peptides Detected by Biosensor Measurement

1. 1 mM stock of antibiotic peptides in DMSO.
2. Dimethyl sulfoxide (DMSO), analytical grade.

3 Methods

3.1 Model Membrane Preparation (Fig. 2)

1. A new or cleaned (*see Note 2*) sensor quartz was put in a solution of 10 mM 1-hexadecanethiol in chloroform for at least 12 h. During this period, a self-assembled monolayer (SAM) is formed on the sensor quartz by a coupling reaction of the thiol groups to the gold surface. The extended lipophilic moieties represent the “inner monolayer” (Step 1) for the subsequently formed membrane bilayer that serves as a model of the biological membrane.
2. The quartz sensor was removed from the thiol solution after the given interval, rinsed with ethanol and dried with compressed air.
3. The second part of the lipid bilayer was established by the Langmuir-Blodgett technique [11, 13]. An illustration of this procedure is given in Fig. 2 (Steps 2–4). The teflon® trough is filled with ultrapure water and the lipid mixture is carefully

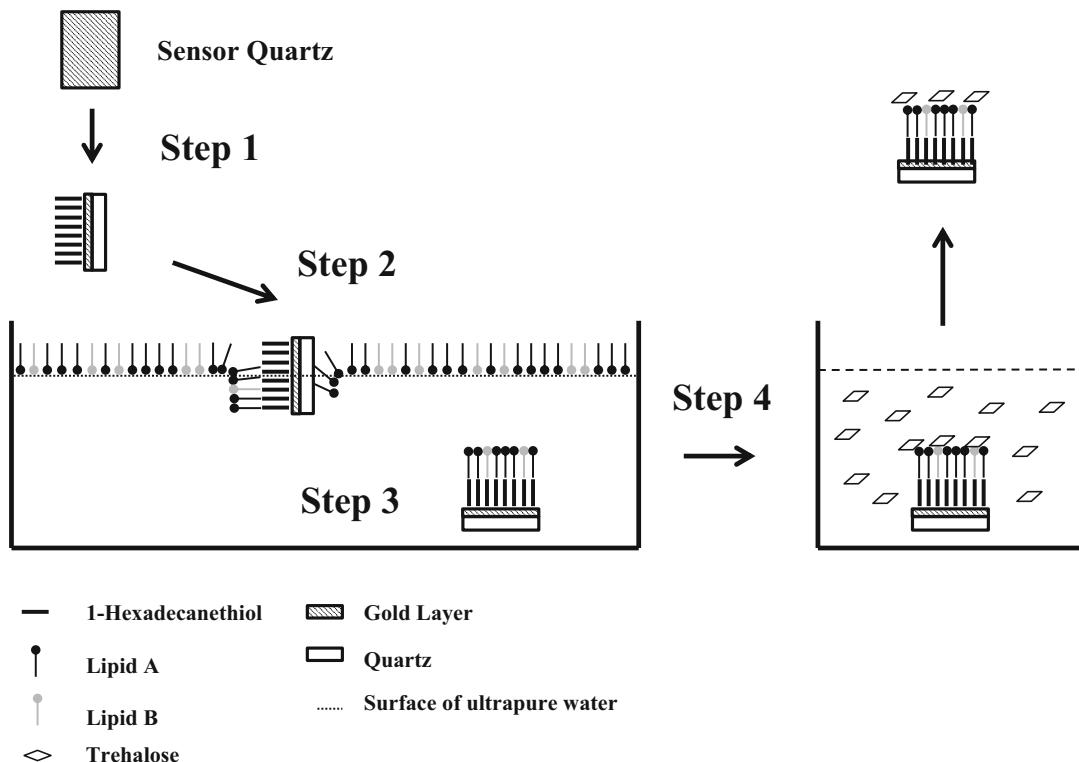


Fig. 2 Schematic illustration of the coating procedure to achieve an artificial lipid bilayer on the sensor surface. The membrane composition can be varied easily by changing the number and/or amount of individual lipids, leading to versatile available membrane functionalization

dropped (*see Note 3*) onto the surface. Depending on the composition of the lipid mixture, the characteristics of the membrane can be modified according to the analytical problem. Here, we applied a mixture of 5 µl POPC (10 mM in chloroform) and 5 µl of DOPG (1 mM in chloroform). For a randomly spreading of the lipids at the water surface they were kept for at least 10 min unaffected before the resulting lipid layer was laterally compressed by a teflon® damper until a value of 5 mN/m below the individual collapse pressure. The collapse pressure, a characteristic of the lipid mixture, has to be evaluated before performing the experiment. The procedure results in a lipid monolayer with clear orientation of lipophilic tail to the air and hydrophilic head to the aqueous compartment.

4. The sensor quartz (prepared as described in **1** and **2**) was dipped through the established monolayer (in **step 3**) resulting in completing the surface-supported lipid bilayer. The dipping procedure is done by a lifting device ensuring a very low speed not disrupting the monolayer on the water surface. The

lipid density of the surface is maintained through tracking the damper according to the difference of desired and actual pressure.

5. Figure 2 illustrates the formation of the lipid bilayer which depends in its intactness on an aqueous environment. For this reason the chip cannot easily be taken out of the water basin.
6. To avoid a disintegration of the bilayer, the sensor chip has to be kept under exclusion of air, primarily achieved by receiving the coated chip in a reservoir put into a cavity of the teflon® trough designated for this purpose before starting the coating procedure.
7. Superfluous lipids at the water surface were aspirated by a water-jet vacuum pump before the reservoir was taken out of the cavity. The reservoir contains the coated quartz and supernatant with some lipids that were collected when passing the water surface. The water is aspirated until only a small liquid portion protects the sensor surface from air.
8. To mount the sensor chip on the biosensor's contacts, one has to assure that the chip is dry and that the bilayer is not disrupted by air. Therefore, according to a protocol established by Reder-Christ et al. [11], the supernatant liquid is stepwise (to avoid dilution effects) exchanged by 1.66 µM trehalose. After 10 min equilibration in the pure trehalose solution, the sensors can be exposed to air and were wept with an additional amount of trehalose (1.66 µM) and dried overnight at 2–8 °C. Trehalose protects the bilayer from disruption when handled under exposure of air and enables a dry mounting on the instrument. Later, first amounts of running buffer will wash away the trehalose and the native model membrane solely remains and is ready for use in the binding experiment.

3.2 Cleaning of SAW-Sensor Quartzes

At the end of the experiment quartz sensors are demounted from the sensor and can be prepared for reuse by applying the following cleaning procedure. To surely detach all remnants of membranes or even proteins bound to the surface during the earlier measurement, harsh conditions have to be applied. Piranha solution was found to be a reliable agent for this purpose. Caution has to be taken while handling Piranha solution as it is very aggressive to tissues and even disrupting the gold surface (*see Note 4*) of the sensor quartz.

1. Piranha solution is always prepared fresh from 30% hydrogen peroxide and concentrated sulfuric acid in a ratio of 1–3 (caution hot!).
2. The sensor quartzes are covered (use a pasteur pipette!) with piranha solution (cooled down to room temperature) for reaction taking place 2 min.

3. The reaction is stopped by dipping (use teflon® tweezers!) the quartzes in a basin with ultrapure water. The quartzes were additionally rinsed with ultrapure water and dried with compressed air.
4. Step 3 is repeated with the variation that after drying the quartzes, they are dipped short time in an acetone containing beaker, followed by dipping in ethanol. Then the quartzes are again dried by air stream.
5. Quartzes are covered another time with piranha solution, which is washed away after 2 min. They are rinsed with ultrapure water and dried by air stream before doing the same with ethanol.
6. Quartzes prepared that way can be stored in the refrigerator until use.

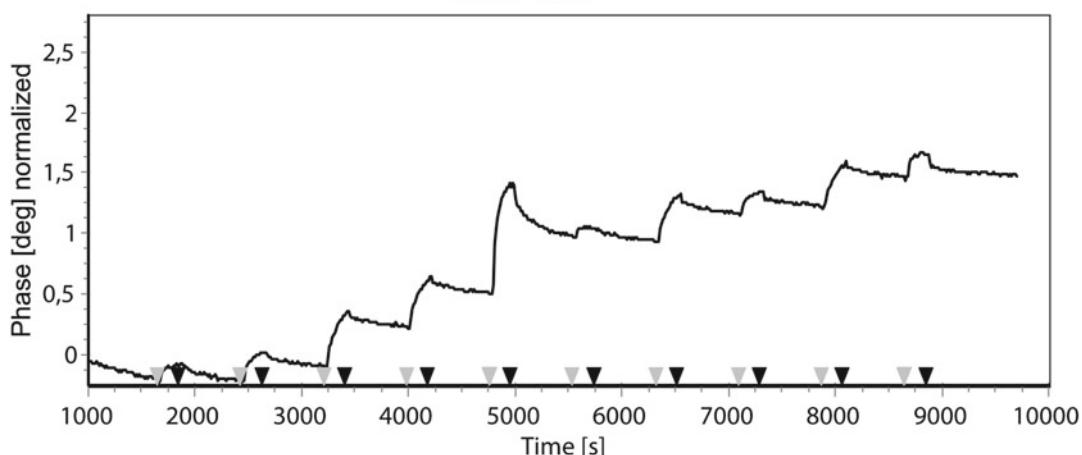
3.3 Membrane Interaction of Two Peptides Detected by Biosensor Measurement

Here, we describe the measurement of membrane interaction of two different peptides (Compounds A and B) with the biosensor. The membrane preparation according to the Langmuir-Blodgett technique [13] allows for a plenty of different assay conditions due to the major influence and easy variation of the lipid composition. In both formats we used a POPC membrane containing 10 mol% DOPG, and a 200 mM MOPS pH 7.0 flow buffer supplemented with 2.5 mM CaCl₂. During measurement, the membrane bearing sensor chip is embedded in the flow chamber compartment of the sensor device and rinsed by degassed (*see Note 5*) running buffer at a flow rate of 40 µl/min equilibrated at 22 °C (Fig. 1). The resulting sensorgrams, obtained for a single channel, are presented in Figs. 3 and 4.

1. A frequency spectrum for the individual quartz is automatically recorded for optimization of the operating frequency.
2. Before starting the measurement procedure one has to wait about 20 min until trehalose is completely washed away and the membrane is equilibrated with running buffer resulting in a stable baseline. With the start of the measurement the phase shift and the amplitude signals were recorded and can be observed in real time as a sensorgram.
3. The samples are prepared in microvials at the desired concentration (a volume of 200 µl is needed for each vial, of which 160 µl is injected). The injections were conducted by an autosampler and characteristics like injection volume and waiting time between different injections can be defined manually.
4. Here, we applied Compound A by injections with the buffer stream over the membrane in the following concentration series, beginning with the lowest: 0.075 µM, 0.1 µM, 0.25 µM, 0.5 µM, 0.75 µM, 1 µM, 2.5 µM, 5 µM, 7.5 µM, 10 µM. The same concentration series was prepared for Compound B (*see Note 6*).

a

Phase - Time

**b**

Amplitude - Time

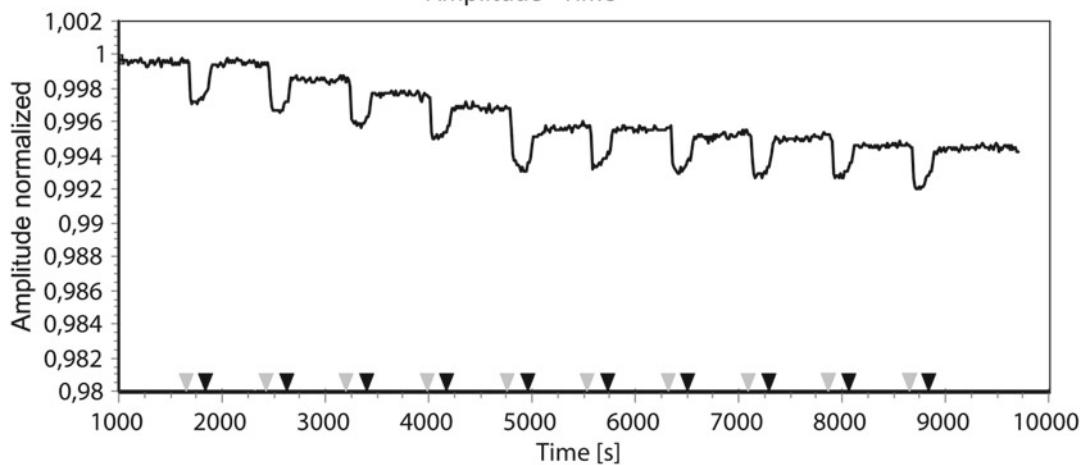
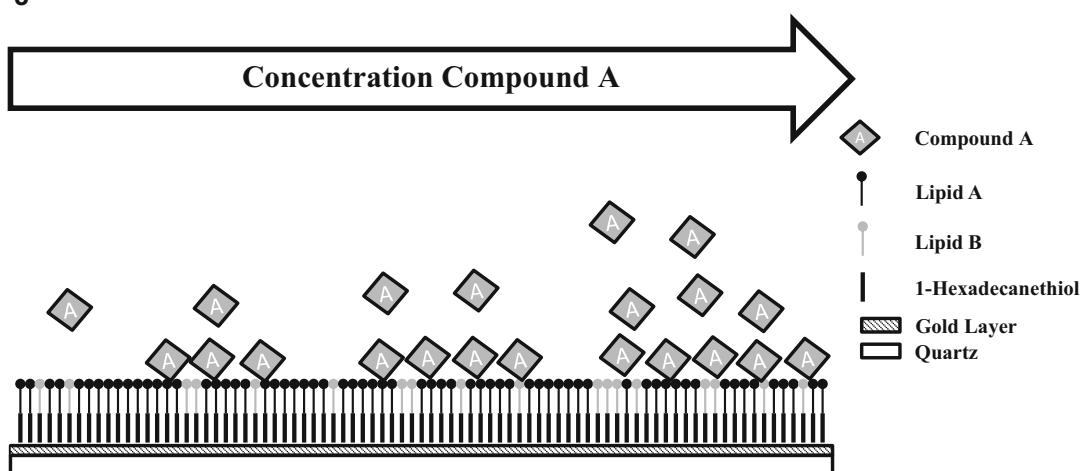
**c**

Fig. 3 (a) Sensorgram of the phase signal derived from membrane binding of Compound A. Binding to the membrane and corresponding deposition of mass on the sensor is accompanied by a shift in phase. Bound molecules were rarely washed away by flow buffer, indicated by the continuously increasing signal. (Gray triangles: Starting injection of an individual concentration of Compound A. Black triangles: End of injection.)

The above assigned concentrations were reached by dilution of 1 mM DMSO stock with running buffer instead of DMSO (*see Note 7*).

5. The injections (gray triangles in Figs. 3 and 4) start with the lowest concentration and a binding event can be identified by a phase shift establishing a higher baseline level, explained by a permanently bound mass. It has to be mentioned that this is not the case for most dynamic binding processes; where at the end of injection (black triangles in Figs. 3 and 4) dissociation of the attached analyte is the dominant process, leading to a decrease in phase shift.
6. The steeper the phase shift increases at an individual concentration, the more total mass is immobilized on the sensor surface respectively the model membrane. The reversible peaks at higher concentrations were mainly induced by viscosity changes as a result of, e.g., solvent residues and should not be considered for evaluation.
7. Considering the amplitude signals, viscosity changes occurring at the sensor surface can be monitored in real time. Binding of a flexible ligand to a pure sensor surface leads to a decrease in amplitude, while the rare event of increasing the amplitude would represent a rigidification, e.g., a membrane condensing effect.

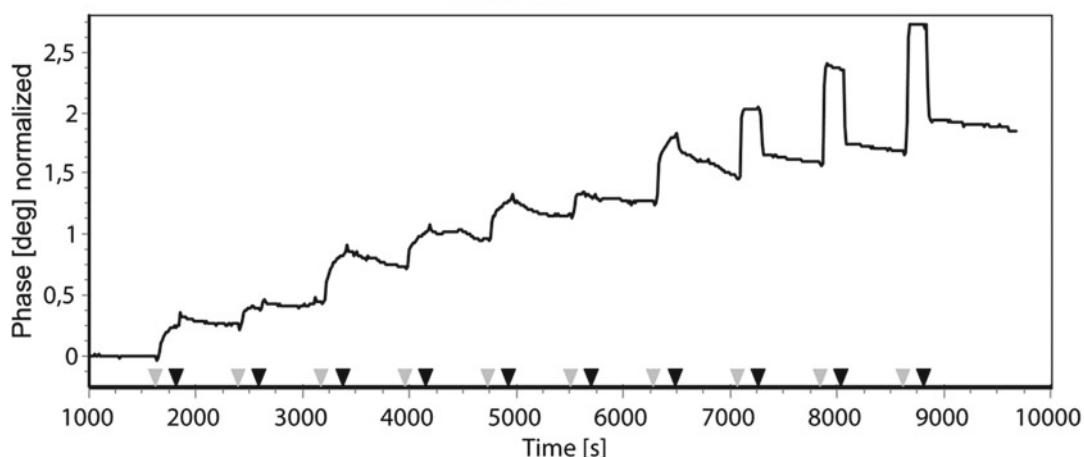
A pure POPC membrane provides an uncharged surface to an analyte at pH 7.0. Since both Compounds A and B represent peptides with a net negative charge, we simulated the physiological conditions by incorporating a negative charge into the membrane by DOPG and supplementation of divalent cations (2.5 mM Ca²⁺). This leads to an acceleration of the interaction with the model membrane at pH 7.0, driven by ionic interaction.

Comparing the sensorgrams of the phase signals obtained for Compound A (Fig. 3a) and Compound B (Fig. 4a), there is no major difference detectable. Both compounds are able to bind the membrane indicated by a step-wise increase in the phase shift, confirming a simultaneous mass deposition at the membrane. Considering the single steps, only small amounts of bound compound appear to be washed away, indicating a tight binding.

Fig. 3 (continued) (b) Sensorgram of the amplitude signal influenced by Compound A. The binding of Compound A to the membrane leads to an agglomerate that is more flexible than the pure membrane. These viscoelastic changes are recorded with the amplitude signal and the more viscous state is represented by an overall decrease. (*Gray triangles*: Starting injection of an individual concentration of Compound A. *Black triangles*: End of injection.) (c) Schematic illustration of the sensor surface during application of increasing concentrations of Compound A. The amount of Compound A deposited to the membrane increases, which is reflected in the increase in phase shift. Compound A only binds to the outer area of the model membrane and therefore the additional layer on the membrane allows for more flexibility of the whole material on the sensor surface. The more viscous properties are reflected in the decreasing amplitude

a

Phase - Time

**b**

Amplitude - Time

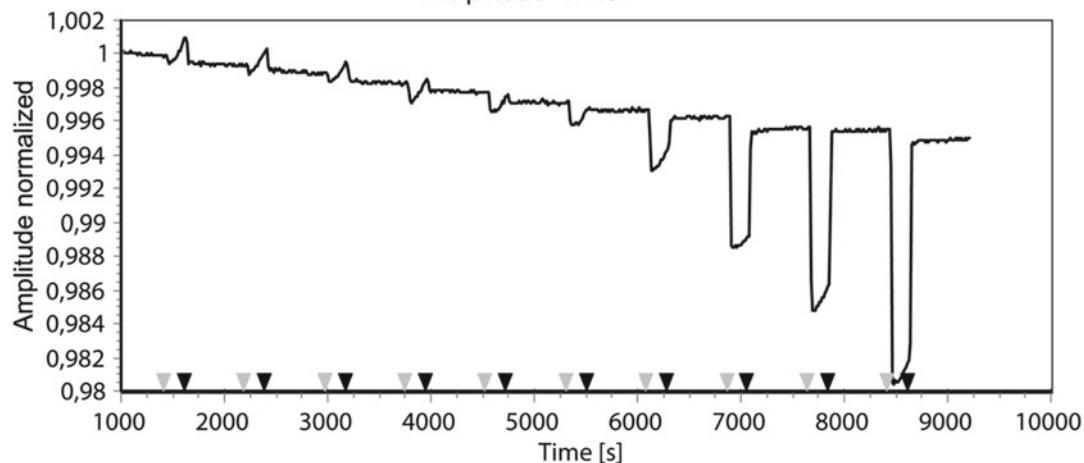
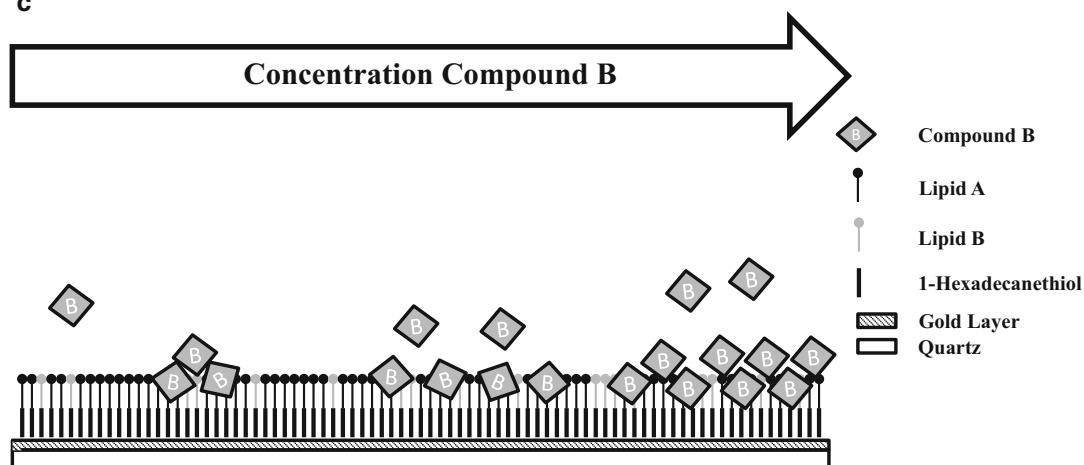
**c**

Fig. 4 (a) Sensorgram of the phase signal influenced by Compound B. In comparison to Compound A there is virtually no difference in the binding behavior of Compound B, regarding only the deposited mass. (*Gray triangles*: Starting injection of an individual concentration of Compound B. *Black triangles*: End of injection.) (b) Sensorgram of the amplitude signal influenced by Compound B. Regarding injections one to three, the amplitude

Furthermore, with increasing concentrations, the steps become smaller referring to an obvious saturation of the membrane interaction capacity (*see also Note 8*).

The amplitude signal allows deriving further information. Compound A (Fig. 3b) influences the amplitude in the expected manner as binding of the compound to the membrane surface leads to a decrease in amplitude, which results from a higher overall flexibility of the whole construct on the sensor surface. This molecular mode of action is depicted in Fig. 3c.

In case of Compound B the amplitude signal has a different appearance (Fig. 4b). During the first three injections of Compound B the amplitude increases. Increasing amplitudes can be observed if a modification turns the whole bound mass on the sensor surface more rigid. Starting with the fourth injection, a decrease in amplitude is evident which is similar to the shape resulting from Compound A. Figure 4c provides a schematic model for this behavior. The initially injected small amounts of Compound B are able to incorporate into the membrane and thus change the viscoelastic membrane properties to a more rigid nature. With the buffer flow adding more and more Compound B, the process is saturated and the molecules are thenceforward solely attached to the membrane surface.

In conclusion, these two examples of peptide interaction with model membranes illustrate that despite the obvious identical tendency for membrane binding, the SAW-biosensor approach is able to discriminate different modes of membrane interaction that justifies this technology as an appropriate tool for an initial and rapid characterization of compounds, like antibiotic drug candidates.

4 Notes

1. MOPS buffer was used instead of the common DPBS (Dulbecco's Phosphate-Buffered Saline) buffer based on optimization studies showing a better reproducibility of the binding data.

Fig. 4 (continued) signal increases upon application of Compound B to the model membrane. Afterwards the signal decreases and later shows a similar appearance as seen with Compound A. The initial increase of the amplitude indicates a more rigid nature of the construct on the sensor surface in response to compound application and is explained in Fig. 4c. (Gray triangles: Starting injection of an individual concentration of Compound B. Black triangles: End of injection.) (c) Schematic illustration of the sensor surface during application of increasing concentrations of Compound B. Compound B interacts besides an inherent binding ability and therefore overall mass deposition (Compound A) in an additional manner with the membrane. Compound B is able to integrate in the membrane and as a result changes their viscoelastic properties. At small concentrations (Injections 1–3) the integration leads to a more rigid construct and the monitored increasing amplitude signal. The ability of the membrane for integration of Compound B is limited and with increasing compound concentrations in the flow buffer, the mass attached to the surface predominates the viscoelastic properties, which leads in sum to a decrease in amplitude

2. The quality of the gold surface is directly related to the quality of the established model membrane, as it is true for the physical properties being reflected in the sensorgram. Therefore, special care has to be taken to an intact and homogenous gold film. Best results are obtained using new sensor quartzes, but the regeneration procedure with piranha solution can be executed 1–2 times without influencing data in an inappropriate manner.
3. The lipid mixture can easily be prepared by aspirating the indicated volume from the stock solution with a glass microsyringe (e.g., Hamilton[®]) and mixing with the syringe's piston. Avoid contact of the lipid containing chloroform solution with plastic parts and execute aspiration and mixing procedures with glass equipment.
4. If the quartz surface is obviously damaged (e.g., black pinholes or scratches) one should discard it.
5. Liquid serving as flow buffer should always be properly degassed because otherwise the chip surface (which is due to the membrane that in general is rather lipophilic!) tends to adsorb gas. The measurement is disturbed by these air-bubbles because the true active surface is reduced and also leads to falsified phase signal. If only one or two of the five channels of the sensor are disturbed by air, the remaining channels suffice to interpret the phase shift in the sensorgram. No significant influence on the buffer flow rate assumed as a precondition.
6. The concentration series being most suitable to be applied on the sensor has to be determined empirically in preliminary tests until binding occurs with appropriate signal intensities.
7. As the SAW technique is very sensitive to viscosity changes in the measurement medium, one has to take special care when performing dilution series. The ready diluted samples should be comprised to most possible extent of flow buffer to avoid artifacts in the sensorgram resulting from viscosity changes in the running buffer.
8. Regarding the phase changes in the original sensorgrams, indicating the bound mass, they seem not to follow a linear principle according to the concentration series. As this is not an uncommon finding, one should bear in mind that the concentration range in this kind of assay is very low and the detection principle is rather sensitive even to small deviations in sample preparation.

Acknowledgments

The authors would like to thank Hildegard Falkenstein-Paul for excellent technical competences in performing the SAW measurements and providing the data presented here.

References

1. Nguyen HH, Park J, Kang S, Kim M (2015) Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors* 15: 10481–10510
2. Gronewold TMA (2007) Surface acoustic wave sensors in the bioanalytical field: recent trends and challenges. *Anal Chim Acta* 603:119–128
3. Gronewold TMA, Schlecht U, Quandt E (2007) Analysis of proteolytic degradation of a crude protein mixture using a surface acoustic wave sensor. *Biosens Bioelectron* 22:2360–2365
4. Länge K, Rapp BE, Rapp M (2008) Surface acoustic wave biosensors: a review. *Anal Bioanal Chem* 391:1509–1519
5. Schlesinger M, Simonis D, Schmitz P, Fritzsche J, Bendas G (2009) Binding between heparin and the integrin VLA-4. *Thromb Haemost* 102:816–822
6. Baca JT, Severns V, Lovato D, Branch DW, Larson RS (2015) Rapid detection of Ebola virus with a reagent-free, point-of-care biosensor. *Sensors* 15:8605–8614
7. Slamnoiu S, Vlad C, Stumbaum M, Moise A, Lindner K, Engel N et al (2014) Identification and affinity-quantification of β -amyloid and α -synuclein polypeptides using on-line SAW-biosensor-mass spectrometry. *J Am Soc Mass Spectrom* 8:1472–1481
8. Engelman DM (2005) Membranes are more mosaic than fluid. *Nature* 438:578–580
9. Lee TH, Hall KN, Aguilar MI (2016) Antimicrobial peptide structure and mechanism of action: a focus on the role of membrane structure. *Curr Top Med Chem* 16:25–39
10. Schneider T, Sahl HG (2010) Lipid II and other bactoprenol-bound cell wall precursors as drug targets. *Curr Opin Investig Drugs* 11:157–164
11. Reder-Christ K, Schmitz P, Bota M, Gerber U, Falkenstein-Paul H, Fuss C et al (2013) A dry membrane protection technique to allow surface acoustic wave biosensor measurements of biological model membrane approaches. *Sensors* 13:12392–12405
12. Gerber U, Hoß SG, Shteingauz A, Jüngel E, Jakubzig B, Ilan N et al (2015) Latent heparanase facilitates VLA-4-mediated melanoma cell binding and emerges as a relevant target of heparin in the interference with metastatic progression. *Semin Thromb Hemost* 41:244–254
13. Girard-Egrot AP, Blum LJ (2006) Langmuir-Blodgett technique for synthesis of biomimetic lipid membranes. In: Martin D (ed) Nanotechnology of biomimetic membranes, 1st edn. Springer Science+Business Media LLC, New York, pp 23–74

Chapter 10

Measurement of Cell Membrane Fluidity by Laurdan GP: Fluorescence Spectroscopy and Microscopy

Kathi Scheinpflug, Oxana Krylova, and Henrik Strahl

Abstract

Membrane fluidity is a critical parameter of cellular membranes which cells continuously strive to maintain within a viable range. An interference with the correct membrane fluidity state can strongly inhibit cell function. Triggered changes in membrane fluidity have been postulated to contribute to the mechanism of action of membrane targeting antimicrobials, but the corresponding analyses have been hampered by the absence of readily available analytical tools. Here, we provide detailed protocols that allow straightforward measurement of antibiotic compound-triggered changes in membrane fluidity both *in vivo* and *in vitro*.

Key words Membrane fluidity, Membrane viscosity, Lipid domains, Lipid packing, Fatty acid disorder, Laurdan, Lipid adaptation, Membrane targeting antimicrobials

1 Introduction

Numerous antimicrobial compounds target bacterial cytoplasmic membranes, and disrupt the normal function of this essential cellular structure. Membrane targeting compounds frequently unfold their antimicrobial properties by interfering with the diffusion barrier function of the cytoplasmic membrane [1]. As a consequence, a comprehensive set of tools has been developed to analyze cellular parameters related to membrane permeability such as ion leakage and membrane potential. However, not all membrane targeting antimicrobials trigger permeabilization of cellular membranes. The mechanisms of action of this category of antimicrobials are considerably less understood [1].

In addition to its permeability barrier function, a correct fluidity state of the membrane is equally important in order to support the multitude of membrane associated cellular processes [2]. Interference with the fluidity state of the membrane by an antimicrobial compound, either by causing changes in the overall membrane fluidity, or by triggering formation of abnormal lipid domains, has high potential to inhibit cell growth [3]. The analysis of these important

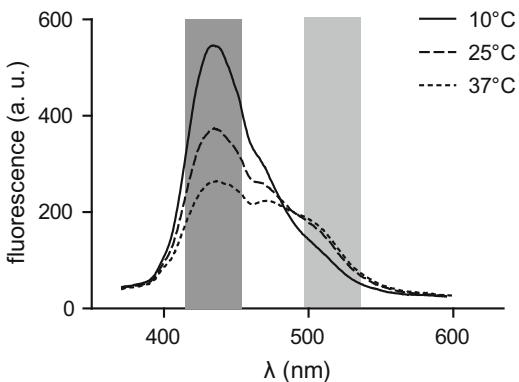


Fig. 1 Fluorescence emission spectrum of laurdan incorporated in small unilamellar vesicles (SUVs) formed of *E. coli* polar lipid extract. Note the spectral shift toward higher wavelength in higher temperatures (indicating increased membrane fluidity). The wavelength ranges used for the ratiometric measurement of laurdan fluorescence (laurdan GP) are highlighted in *light* and *dark gray*, respectively

cellular parameters has been greatly hampered by the relative absence of suitable, easy to adapt analytical tools. Here, we provide detailed protocols for the analysis of membrane fluidity of bacterial cell membranes both on a global scale, and on a single cell level with spatial resolution. The provided measurements can be carried out with widely available standard laboratory equipment such as fluorescence microplate reader and wide field epifluorescence microscope.

The protocols provided in this chapter make use of a fluorescent, fluidity-sensitive, and noninhibitory membrane dye laurdan [4, 5]. The fluorescence emission spectrum of laurdan is sensitive to the presence of H₂O close to its chromophore. The ability of H₂O to penetrate the hydrophobic membrane interior is dominated by lipid head group packing density and fatty acid disorder of lipid bilayers. As a consequence, the fluorescence emission spectrum of membrane embedded laurdan is sensitive to membrane fluidity and disorder in its surrounding (see Fig. 1) [4–8].

The provided protocols are optimized for Gram-positive model organism *Bacillus subtilis* but also offer a good starting point for measurements in other Gram-positive microorganism such as *Staphylococcus aureus*. We provide example measurements how these methods can be applied to gain insight into mechanism of action of membrane targeting antimicrobials.

2 Materials

2.1 Laurdan Fluorescence Spectroscopy In Vivo

- 1 mM Laurdan (6-Dodecanoyl-2-Dimethylaminonaphthalene; either from Molecular Probes or Sigma-Aldrich) stock solution in 100% DMF (Dimethylformamide), store in –20 °C, keep always in dark.

2. 5 M benzyl alcohol stock by dilution with DMSO (Dimethyl sulfoxide), store in -20 °C, cover stored aliquots with Argon or N₂ to prevent oxidation.
3. Fluorescence microplate reader. Both monochromator-based plate readers, and a filter-based readers equipped with 350 nm excitation filter and appropriate emission filters (ranges spanning 420–460 nm and 490–520 nm) are suitable.
4. Black, flat bottom 96-well plates; if reusable plates are used ensure proper cleaning after use.

2.2 Laurdan Fluorescence Spectroscopy In Vitro

1. Phospholipids of choice. Either natural lipid extracts, or mixtures of synthetic or purified lipids can be used. We recommend either *Escherichia coli* Polar Lipid Extract, or a mixture mimicking bacterial cytoplasmic membrane composed of a zwitterionic 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) combined either with anionic cardiolipin or with 1-palmitoyl-2-oleoyl-*sn*-glycero-3[phosphor-*rac*-(1-glycerol)] (POPG). All lipids mentioned above can be purchased from Avanti Polar Lipids.
2. Laurdan (6-Dodecanoyl-2-Dimethylaminonaphthalene, Molecular Probes or Sigma-Aldrich). Prepare a 0.2 mg/ml laurdan solution in chloroform, store in -20 °C, keep dark.
3. 10 mM sodium phosphate (NaH₂PO₄/Na₂HPO₄) buffer containing 154 mM NaCl and 0.1 mM EDTA, pH 7.4. Or other buffer of choice.
4. Chloroform and methanol of highest available purity.
5. Nitrogen or argon gas.
6. 1.5 ml and 2 ml reaction tubes and pipette tips siliconized if necessary (*see Note 1*).
7. Round-bottomed glass vials (~5 ml) with tightly sealed caps. Flat-bottomed glass vials (~2 ml) with caps.
8. Graduated glass pipettes (2 ml); Hamilton gastight syringe (100–200 µl).
9. High-vacuum pump (10⁻² to 10⁻⁴ mbar).
10. Mini-extruder and polycarbonate membranes with defined pore size (*see Note 2*). Can be purchased from Avestin Inc. or Avanti Polar Lipids.
11. Dry ice, ultrasonic bath with thermoregulation.
12. Fluorescence spectrometer (monochromator-based).
13. Disposable macro UV/VIS cuvettes (3 ml, 1 × 1 cm).
14. Magnetic stir bar (<10 mm in length).

2.3 Laurdan Fluorescence Microscopy

1. 10 mM Laurdan (6-Dodecanoyl-2-Dimethylaminonaphthalene; either from Molecular Probes or Sigma-Aldrich) stock solution in 100% DMF (Dimethylformamide), store in -20 °C, keep always in dark.
2. PBS (8.0 g/L NaCl, 0.2 g/L KCl, 1.15 g/L Na₂HPO₄, 0.2 g/L KH₂PO₄, pH 7.3) supplemented with 0.1% d-glucose.
3. Agarose (electrophoresis grade).
4. Fluorescence microscope equipped with:
 - (a) A high quality 100× objective with good chromatic correction such as Nikon Plan Apo series, Zeiss Plan Apochromat series, or equivalent.
 - (b) Appropriate filter sets (excitation at approx. 350 nm, emission at 430–460 and 500–530 nm) (*see Note 3*).
 - (c) Wide field illumination with strong light output at 350 nm. We prefer Hg-vapor or metal halide light source for this application.
 - (d) Temperature control.
 - (e) High sensitivity CCD, EM-CCD, or sCMOS camera with maximally 8 × 8 μm pixel size.
5. High quality microscope slides, coverslips, and immersion oil.
6. 0.1 μm diameter TetraSpeck™ fluorescent microspheres (Thermo Fisher Scientific).

3 Methods

3.1 Laurdan Fluorescence Spectroscopy In Vivo

3.1.1 Sample Preparation and Data Acquisition

1. Grow cells to an optical density at 600 nm (OD₆₀₀) of approx. 0.5 in suitable growth medium supplemented with 0.1% glucose at the desired temperature (*see Notes 4–6*).
2. Transfer the cell suspension to a 2 ml reaction tube and add laurdan to a final concentration of 10 μM (from a 1 mM laurdan stock solution, *see Note 7*).
3. Incubate cells with laurdan for 5 min at the desired growth temperature in a thermomixer. Cover tubes with aluminium foil to avoid light exposure.
4. Wash cells 4× in 2 ml pre-warmed PBS/glucose (centrifuge for 1 min at 16,000 × g in a table top centrifuge, carefully remove the supernatant by pipetting, resuspend in fresh PBS/glucose, repeat 4 times). After the last wash, resuspend to obtain an OD₆₀₀ of approx. 0.5 (*see Notes 8 and 9*)
5. Remove 500 μl of the cell suspension, transfer to a new reaction tube, and centrifuge as described above. Carefully harvest ~450 μl of the supernatant, which serves as laurdan background

fluorescence in subsequent measurements (background of buffer + dye not associated with cells).

6. Immediately proceed with fluorimetric measurement by transferring the stained cell suspensions, and the background sample to a pre-warmed black, flat bottom black 96-well microtiter plate (150 µl/well).
 7. Depending on the antimicrobial compound of interest, and the specific research question, three measurement options are possible:
 - (a) Preincubation of the cell culture with the antibiotic of choice, followed by staining and measurement. This measurement mode is suitable for slow acting but tightly bound antimicrobials, or for an analysis of potential adaptation to subinhibitory concentrations. In this case, we recommend a brief (2 min) shaking interval in the microplate reader before the fluorescence measurement.
 - (b) Incubation of stained cell suspension with the antibiotic of choice for a given incubation time. In this case, incubate the stained cell suspension with the compound directly in the microtiter plate for a required time under shaking, followed by fluorescence measurement. In well-energized untreated cells (PBS/glucose + shaking) laurdan GP values were found to be stable for up to 45 min.
 - (c) For a kinetic measurement, laurdan fluorescence can be measured before, and as a time series after addition of the antibiotic of interest. In order to ensure sufficient energization of the cells, we recommend either continuous shaking or a relatively low number of parallel samples. Measurement intervals of 0.5–1 min are a good starting point (*see Note 10*).
 8. As a positive control, incubate cells with 50 mM membrane fluidizer benzyl alcohol (*see Note 11*).
 9. Measure laurdan fluorescent intensities upon excitation at 350 nm at two emission wavelengths. In a monochromator-based fluorimeter, the optimal wavelengths (435 and 500 nm) should be used. In a filter-based fluorimeter, filters with wavelengths centered at 430–460 nm, and 490–520 nm are acceptable.
- 3.1.2 Data Analysis**
1. Subtract values obtained from the background sample (fluorescence of unbound dye) from the cell suspension values for each wavelength. The same background values are subtracted from both treated and untreated samples (this assumes that the compound of choice does not have fluorescent properties itself).

2. Calculate laurdan generalized polarization (GP) as follows:

$$GP = \frac{I_{435} - I_{500}}{I_{435} + I_{500}}$$

On a scale ranging from -1 to +1, high GP values correspond to low membrane fluidity and vice versa [4, 5]. Depending on the exact filters used, the absolute values vary and should be regarded as arbitrary values. Commonly, however, GP values of 0.1–0.6 are measured (see Figs. 2 and 3 for examples).

3.2 Laurdan Fluorescence Spectroscopy In Vitro

3.2.1 Lipid Handling

1. Prepare three round-bottomed glass vials with tightly sealed caps, label the vials (lipid composition, data etc.). Determine the weights of the vials (with caps on) on a microbalance taking at least five significant digits (see Note 12).
2. Pipet ~40 mg of lipid in organic solvent into one vial and add an appropriate volume of 0.2 mg/ml laurdan solution in chloroform (150 µl for 40 mg of *E. coli* polar lipid extract, see Note 13) resulting in a laurdan to lipid molar ratio of 1.5:1000. Use a graduated glass pipette for lipid stock solution, and Hamilton gastight syringe for the laurdan stock solution. If the lipid was purchased in powder form, first dissolve it in chloroform or chloroform/methanol mixture of 3:1 to yield a concentration of 20 mg/ml and then proceed as above. At last, divide the lipid-laurdan mixture into two other vials to obtain identical samples for replicate experiments.

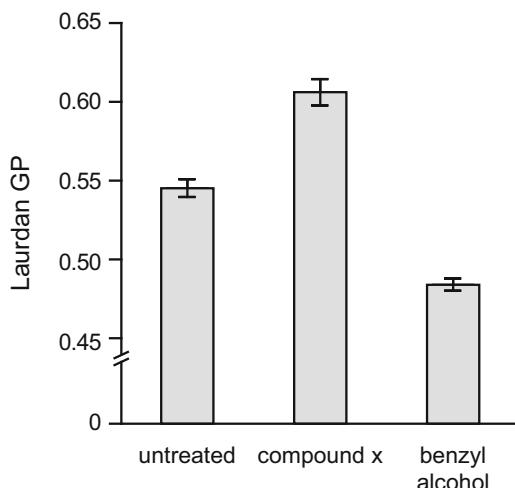


Fig. 2 Laurdan GP values measured for untreated live *B. subtilis* cells, and for cells incubated with membrane fluidizer benzyl alcohol and antimicrobial compound x, respectively. Note the reduced laurdan GP (increased fluidity) in the presence of benzyl alcohol, and the opposite effect in the presence of compound x

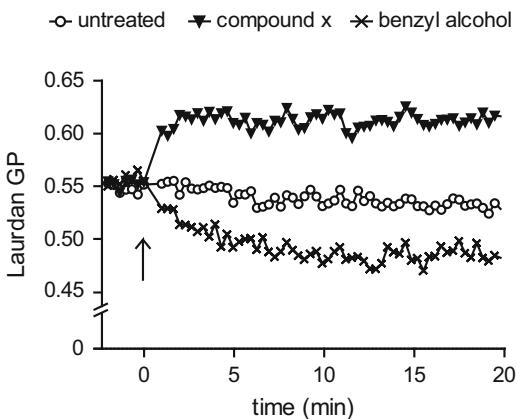


Fig. 3 Continuous measurement of laurdan GP for untreated wild-type *B. subtilis* cells, and for cells incubated with membrane fluidizer benzyl alcohol and antimicrobial compound x, respectively. The time point of benzyl alcohol and compound x addition is indicated with an arrow

3. Evaporate the organic solvent with dry nitrogen or argon by passing of a gentle gas stream over the lipid-containing solution (*see Note 14*). During the procedure, slowly rotate the vial such that a thin lipid film is deposited on the walls and bottom of vial while the solvent evaporates. Continue the procedure until the solvent has visually disappeared, followed by additional 3–4 min.
4. Use a high-vacuum pump (10^{-3} to 10^{-4} mbar) for at least 3 h to remove traces of organic solvent. Seal the vials tightly in order to prevent water absorption from the air (use parafilm sealing over the tube caps after weighing the tubes (*see step 5*)).
5. Weigh closed vials and subtract the weights of empty vials to calculate the net weight of the lipid films. Covered with argon or nitrogen, lipid-laurdan films can be stored at -20°C for 2–3 months (*see Note 15*).

3.2.2 Vesicle Preparation

1. Hydrate a lipid-laurdan film at room temperature for 30 min in a phosphate buffer pre-equilibrated to room temperature (20 mM final lipid concentration). For this aim, vortex the vial vigorously for ~5–10 min to solubilize the lipid film. Repeat additionally three to five times short (10–15 s) shaking of the suspension during the hydration step. As a result, a turbid suspension of multilamellar vesicles (MLV) will be formed (*see Note 16*).
2. Dilute the obtained lipid suspension with equal volume of buffer to obtain a 10 mM lipid stock (vortex well). Subject the stock to three cycles of freeze-thaw using dry ice and 40°C water bath, with intensive vortexing of the thawed suspension. Depending on the number of samples in the experiment, part of the 10 mM

lipid stock can be sealed with argon/nitrogen, frozen, and stored at -20 °C for replicate experiments (*see Note 17*).

3. To obtain an antimicrobial compound-free sample, dilute the 10 mM lipid stock to 1 mM with buffer. Use mini-extruder to prepare large unilamellar vesicles (LUVs). For this, fill one of two 1 ml gastight syringes with MLV suspension and connect to a mini-extruder equipped with two stacked polycarbonate filters with 100 nm pore size. Force the MLV suspension through the filter; a minimum of 21 extrusion cycles was found to be required for optimal LUV preparation (*see Notes 18 and 19*). Collect LUVs in a small flat-bottomed glass vial or a reaction tube.
4. Prepare several dilutions of the antimicrobial compound (0.5–1 ml each) and mix with equal volume of 2 mM lipid stock (diluted from 10 mM) in small flat-bottomed glass vials, shake well. This step provides samples with 1 mM lipid and different lipid-to-peptide (L/P) molar ratios (*see Note 20*).
5. If vesicles aggregate due to antibiotic adsorption, use sonication (ultrasonic bath) to dissolve the aggregates. Subject aggregated, highly turbid peptide-lipid samples to one to two cycles of sonication (5 min/cycle) in the bath in order to obtain uniformly opalescent samples.
6. At last, to obtain uniform LUVs, subject the sonicated vesicles to extrusion as described for the antimicrobial compound-free sample (**step 3**) (*see Note 21*).

3.2.3 Fluorescent Spectroscopy Parameters

Two data types can be obtained from laurdan-labeled vesicles as a characteristic for a lipid membrane state: dye emission spectra and fluorescence intensities at fixed wavelengths.

1. To obtain the emission spectra, choose the emission scan mode of the fluorescent spectrometer (the emission spectrum is scanned while the excitation wavelength is held constant). Set the excitation wavelength to 350 nm, which roughly corresponds to the adsorption maximum of laurdan [4]. Set the emission wavelength to be scanned from 370 to 600 nm. Choose appropriate values for the slit widths of both excitation and emission monochromators. Pre-equilibrate the device at the temperature of choice. (*See Figs. 1 and 4 for examples of laurdan spectra in vesicles*).
2. To obtain fluorescence intensities at fixed wavelengths, choose fixed wavelength measurement mode (use dual wavelength mode when possible). Set the excitation wavelength to 350 nm, and the emission wavelengths to 435 and/or 500 nm. When dual wavelength mode is not available, the emission wavelength needs to be changed manually in separate measurement. Choose appropriate values for the slit widths of both excitation and emission monochromators. Pre-equilibrate the device at the temperature of choice.

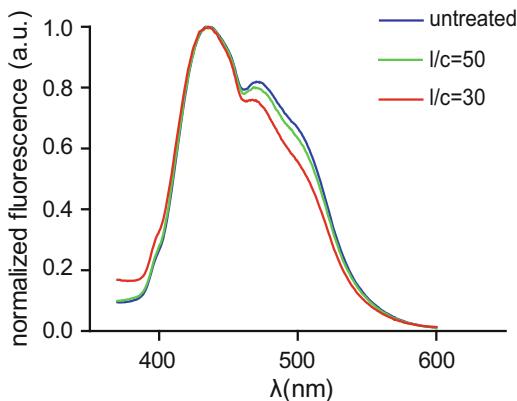


Fig. 4 Normalized fluorescence emission spectrum of laurdan incorporated in large unilamellar vesicles (LUV's) formed of *E. coli* polar lipid extract. Note the reduced intensity of the higher wavelength shoulder of the spectrum in the presence of compound x (indicating reduced membrane fluidity). I/c values represent lipid to compound molar ratios

3.2.4 Data Acquisition

1. Prepare the samples by diluting 1 mM LUV stocks with buffer (in 2 ml in disposable fluorescent cuvettes). Usually, 0.25–0.3 mM lipid is enough to obtain good fluorescence intensity signals. Equip the sample cuvette with a magnetic stir bar.
2. Place the sample in the sample compartment of the fluorescence spectrometer and equilibrate the temperature for at least 5 min (in-cell temperature sensor is recommended for accurate temperature control, *see Note 22*). Measure the sample containing only vesicles without additive (control) first. As laurdan is sensitive to photobleaching, reduce the sample's exposure time to the light source by closing the shutters whenever no data are being collected.
3. Record fluorescence emission spectrum/fluorescence intensities at fixed wavelength with the settings specified above (Subheading 3.2.3). If necessary, adjust the sensitivity of the photodetector (gain) to obtain a reasonable signal intensity and good signal-to-noise ratio, and repeat the measurement. Use identical settings for all parallel samples.
4. Repeat fluorescence recording for all parallel samples.
5. When different temperatures are scanned, start at low temperature (*see Note 23*). We found that it was more convenient from the practical point of view to scan all samples at one desired temperature, and then repeat the procedure for the next temperature points.

3.2.5 Data Analysis

1. Data of fluorescence emission spectra can be easily represented and analyzed in a commonly used spreadsheet program, such as Microsoft Office Excel, SigmaPlot, or Origin. Normalization

of the spectra to fluorescence intensity maximum gives more clear information about additive-induced changes in laurdan fluorescence (*see Fig. 4* for an example).

2. Fixed wavelength laurdan fluorescence intensities are used to calculate laurdan generalized polarization parameter, GP, which is given by the following equation [4]:

$$GP = \frac{I_{435} - I_{500}}{I_{435} + I_{500}}$$

This type of measurement is an *in vitro*-equivalent of the *in vivo* measurement shown in Fig. 2.

3.3 Laurdan Fluorescence Microscopy

3.3.1 Sample Preparation

1. Grow cells to an OD₆₀₀ of ~0.5 in a suitable growth medium supplemented with 0.1% glucose at the desired growth temperature (*see Note 6*).
2. Transfer 100 µl of the culture to a 2 ml (round bottom) reaction tube. Add laurdan from a 10 mM stock solution (in 100% DMF) to final concentrations of 100 µM laurdan and 1% DMF (*see Notes 7 and 24*).
3. Incubate the cell suspension for 5 min at the desired growth temperature under vigorous shaking (*see Note 25*) in the dark.
4. Remove unbound laurdan by adding 1.9 ml pre-warmed PBS/glucose followed by centrifugation for 1 min (at max speed on a table top centrifuge), remove the supernatant, and resuspend in 100 µl pre-warmed PBS/glucose (*see Note 26*).
5. Incubate the stained cell suspension with the compound of interested for a chosen time period under vigorous shaking. Alternatively, preincubate cells with the compound of choice before staining (**step 2**).
6. Immobilize 0.5 µl of stained cell suspension on a microscopy slide prepared in with 1.2% agarose in H₂O (*see Notes 27 and 28*).

3.3.2 Microscope Calibration

In principle, the microscopy can be carried out with any wide field fluorescence microscope but we strongly recommend the use of a microscope setup with an adjustable temperature chamber set to the growth temperature of the analyzed bacteria. It is paramount to use high quality microscopy objective with high level of correction for chromatic aberration. Even so, the residual difference in exact focal plane and magnification between the imaged wavelengths (450 and 520 nm) requires pre-calibration of the microscope setup (*see Note 29*).

1. Prepare a slide 1.2% agarose in H₂O and add an appropriate dilution of a suspension of 0.1 µm diameter TetraSpeck™ microspheres.

2. Select a field of view with some fluorescent beads located close to both the left and right edges of the image field.
3. Carry out optical sectioning (200 nm step size, 5 µm depth) with both wavelengths (450 and 520 nm).
4. Measure the distance of two pixels located on the left and right edges of the image field using both wavelengths (due to the wavelength-dependency of magnification, the distance in pixels is slightly larger for the image captured at 450 nm).
5. Note the difference in pixel-distance along the x -axis of the image between the chosen bead-pair. This will be used to correct for the wavelength-dependent difference in magnification upon image analysis.
6. Determine the difference in an exact focal plane at the different wavelengths (450 and 520 nm) using the captured optical sections.
7. Use the difference in focal planes (z-offset) to correct for the residual difference in chromatic aberration. Most image capture software allows for the use of predefined z-offset values for different wavelengths.

3.3.3 Image Capture

1. Allow the microscope to reach a stable temperature by turning on the microscope body, and the heating unit at least 2 h in advance.
2. Preprogram the microscope for capture of a fluorescence image using a 350 nm excitation filter and a 520 nm emission filter, followed by a second image using a 350 nm excitation filter and a 450 nm emission filter. Define the z-offset value to correct for the different focus.
3. Rapidly capture images on both wavelengths for a required number of cells/fields (*see Notes 30 and 31*) and save them in a raw 16-bit uncompressed tif format.

3.3.4 Image Analysis

1. Increase the size of the higher wavelength image (520 nm) to correct for the wavelength-dependent difference in magnification. For this aim, use the conversion value obtained from the measurement of fluorescent beads upon microscope calibration.
2. Carefully correct the exact position of the cells (x/y -drift) manually to ensure a perfect image overlay (*see Fig. 5*) and crop the images to contain only the selected cells.
3. Convert the images into a 32-bit format. This allows the use of decimal places in pixel intensity values which is necessary for the representation of GP values.
4. With images prepared as above, the pixel by pixel calculation of laurdan GP can now be performed using the equation:

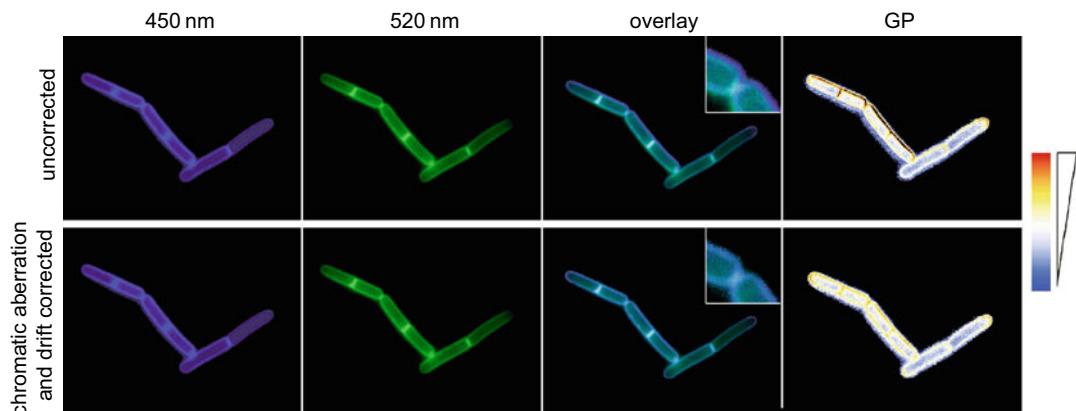


Fig. 5 Fluorescence microscopy of laurdan-stained untreated *B. subtilis* cells at two different wavelengths (450 and 520 nm, *first* and *second* panels, respectively). The *third* panel depicts an overlay of the two micrographs, and the *fourth* panel pixel-by pixel calculated laurdan GP values (*red* color representing high and *blue* low membrane fluidity, respectively). *Upper images* depict the consequence of a poor correction for chromatic aberration and x/y-drift. The *lower panels* depict the same cells with proper corrections applied

$$GP = \frac{I_{450} - I_{520}}{I_{450} + I_{520}}$$

The specific details of the calculation depend on the program of choice. The corresponding software manual should provide sufficient guidance to perform the calculations. We routinely use Wolfram Mathematica for this step but any software capable for image arithmetic can be used (*see Note 32*).

5. Due to the ratiometric nature of GP calculation, pixels with low fluorescence intensity (such as pixels outside the cell) will also be assigned a GP value. For visualization purposes, we find it useful to define an intensity cut-off below which the GP values are set to zero (and assigned the color black). An optimal cut-off value depends on the signal to background ratio but 10% is a good starting value. The pixel-GP values associated with the cell (commonly in a positive range between 0.1 and 0.4) are best represented with a blue-red color scale with blue representing low and red high membrane fluidity (*see Fig. 5* for an example).

4 Notes

1. Some hydrophobic substances strongly adsorb to plastic and/or glass surfaces. This problem can be fixed by using siliconized tubes and pipette tips for all steps handling such a substance [9].

2. Lipid vesicle size defines the scattering intensity of a solution. Light scattering can noticeably influence the signal intensities when performing fluorescence-spectroscopic experiment. Wherever possible, smaller vesicle size is preferable for the studies. Alternatively to extrusion, small lipid vesicles can be obtained by sonication [10]. Scattering effects can also be taken into account by performing control experiments where fluorescence intensities are measured in dye-free system, and then subtracted from the values obtained for the dye.
3. Commonly available filter sets used to image DAPI, FITC, and GFP can be used to carry out laurdan GP microscopy. In a microscopy setup in which excitation and emission filters, and the dichroic mirrors can be chosen independently, the combination of DAPI excitation (commonly ~350 nm), DAPI dichroic mirror (~410 nm), and either DAPI emission (~450 nm), or FITC/GFP emission (500–530 nm) is well suited. In a filter cube-based system, two separate cubes (one containing a regular DAPI filter set, and another combining DAPI excitation, DAPI dichroic mirror and GFP emission) are required.
4. The described protocol is optimized for *Bacillus subtilis* but provides a good starting point for measurements also in other Gram-positive bacteria. With Gram-negative species such as *E. coli*, laurdan staining requires permeabilization of the outer membrane with 30 µg/ml polymyxin B nonapeptide (PMBN) for 1 h prior to staining.
5. Membrane fluidity values are very sensitive to changes in temperature. Furthermore, cells rapidly adapt to changes in temperature by adapting the fatty acid composition of the membrane. As a consequence, all materials including reaction tubes, buffers, microtiter plates, plate reader, etc. should be pre-warmed to the initial growth temperature.
6. We routinely measure effects of an antibiotic on membrane fluidity for *B. subtilis* grown in LB-medium supplemented with glucose. This allows the cells to pre-adapt to utilization of glucose, and thus provides the means to keep cells energized upon wash and resuspension in PBS/glucose. As an alternative, initial growth in LB/glucose buffered with 25 mM Tris-HCl pH 7.0, followed by staining and resuspension in buffer composed of 25 mM Tris-HCl, 10 g/L NaCl, and 0.1% glucose pH 7.0 is possible. If minimal medium is preferred, cells can be stained and washed directly in the same medium.
7. A final concentration of 1% DMF was found to enhance the solubility of laurdan resulting in stronger staining of *B. subtilis*. Hence, dilution of laurdan stock solutions should not exceed 1:100. 1% DMF does not affect the viability of *B. subtilis*.

8. The signal to background ratio of laurdan-stained cells is best at cell densities of $OD_{600} > 0.5$ (for *B. subtilis*). Reliable measurement can be achieved with lower cell densities but we do not recommend values lower than OD_{600} of 0.3.
9. Laurdan staining and subsequent washing steps should be carried out in a speedy manner. In this context, it was found appropriate to limit overall number of samples per run to three reaction tubes (6 ml, sufficient for 12 samples measured in technical triplicate).
10. If changes in fluidity are followed over time, measure the laurdan fluorescence for 5–10 min under shaking to obtain a stable baseline. Pause measurement and quickly add antimicrobial compound at the desired concentration. Make sure the additional volume does not exceed 2 μl per well to minimize dilution effects.
11. Benzyl alcohol is a chemical membrane fluidizer which provides a positive control for the ability of the assay to detect changes in membrane fluidity. In order to achieve the full extent of fluidization a 5 min incubation is required.
12. Take care not to change the weight afterward by making new labels or writing on the tubes.
13. The molar mass of *E. coli* polar lipid extract was taken as 700 g/mol [11].
14. Do not use too strong gas flow and avoid spattering of the lipid-containing solution out from the tube.
15. We notice a 10–15% intensity decrease comparable to freshly prepared samples for frozen lipid films after 3 months at -20°C . Thus, a longer storage is not recommended.
16. For lipid hydration and vesicle formation the temperature of the solution should be maintained above the gel-to-liquid-crystalline phase transition temperature of the lipid. Prolonged vigorous shaking (>30 min) may be required for complete lipid hydration in some cases.
17. MLV suspensions can be stored under argon or nitrogen at -20°C for months. With laurdan-containing vesicles we usually do not exceed 2 months storage period (see also Note 15).
18. The number of extrusion cycles is always odd. The LUVs suspension is always collected in the syringe opposite to the one that initially carried the suspension.
19. It is recommended to check the size and polydispersity (particle size distribution) of the resulted LUVs by dynamic light scattering (DLS). Zetasizer Nano device (Malvern Instruments, UK) is used routinely in many laboratories for this control.
20. Here, glass vials are necessary for the following sample bath sonication. Using plastic tubes is not recommended due to a

much lower transparency for soft ultrasound waves of bath sonication.

21. Not all peptides or other additives cause vesicle aggregation. If this is the case, bath sonication step can be omitted. Nevertheless, to reach homogeneous additive redistribution between outer and inner bilayer leaflets in vesicles, MLV-additive suspension should be intensively mixed (shaking on vortex) and subjected to extrusion. Here, extrusion is preferred over bath sonication as it gives much more vesicles of uniform size.
22. The temperature of the sample must be controlled very accurately. Even a 0.5 °C temperature difference can have a measurable effect on laurdan generalized polarization parameter.
23. The fluorescent properties of laurdan are sensitive to the physical state of the lipid bilayer. Laurdan signal intensity and shape of the emission spectrum change with temperature. At low temperatures (in a more ordered bilayer) laurdan fluorescence is high with a maximum at ~440 nm. With increasing temperature, membrane order decreases and the signal intensity is reduced while the fluorescence maximum displays a red shift to 490 nm.
24. Incubation in round bottom 2 ml reaction tubes (in contrast to conical 1.5 ml tubes) ensures better aeration upon shaking, and therefore minimizes effects caused by de-energization [7, 12].
25. We use Eppendorf ThermoMixer at 850 rpm.
26. A less stringent wash to remove unbound dye is required for the microscopic assay than for the fluorimetric assay.
27. Regular agarose/H₂O microscopy slides are prepared by boiling 1.2% (w/v) in H₂O until agarose is fully dissolved, followed by cooling down to approx. 65 °C. The warm agarose solution (500 µl) is quickly spread on a microscopy slide, covered with another identical slide, and allowed to solidify at least for 10 min at room temperature. Immediately before use, the upper slide is removed, 0.5 µl of the stained cell suspension applied, gently dried until the liquid drop has evaporated, and covered with a microscopy coverslip. Prepared slides (without cells) can be stored for max. 3 h in the fridge in a container with wet tissue paper. Too dry agarose slides result in artifacts in membrane stains.
28. The cell suspension in PBS/glucose is applied on an agarose bed that contains some un-polymerized sugar which contributes toward osmolarity. During the process of sample embedding, the salts from the cell suspension increase the local osmolarity. If too high volume (more than 0.5 µl) is used, or agarose in undiluted buffer instead of H₂O is used, the resulting high local osmolarity results in visible plasmolysis. For these

reasons, the use of H₂O/agarose and 0.5 µl sample volume is strongly recommended.

29. Once defined for a given microscope setup, the values for both scaling the images and for the z-offset can be used later microscopy without need for recalibration. Recalibration is only necessary in case of a different objective, or a different fluorescent filter set.
30. Under the coverslip, the cells are not provided with an adequate supply of O₂. Aerobic microorganisms such as *B. subtilis* cells start to lose energization after approximately 10 min. For these reasons, image acquisition should be carried out in a speedy manner.
31. Laurdan fluorescence bleaches rapidly. For this reason, the illuminated sample area should be kept at minimum necessary to achieve an even illumination across the captured field. Capture of images of cells close to each other should be avoided.
32. A Wolfram Mathematica-script for the calculation of image-based laurdan-GP can be obtained from the authors upon request.

Acknowledgment

This work was supported by Wellcome Trust Institutional Strategic Support Funds (ISSF), and by Newcastle University.

References

1. Yeaman MR, Yount NY (2003) Mechanisms of antimicrobial peptide action and resistance. *Pharmacol Rev* 55:27–55
2. Zhang YM, Rock CO (2008) Membrane lipid homeostasis in bacteria. *Nat Rev Microbiol* 6:222–233
3. Epanet RM, Epanet RF (2009) Lipid domains in bacterial membranes and the action of antimicrobial agents. *Biochim Biophys Acta* 1788:289–294
4. Parasassi T, De Stasio G, Ravagnan G, Rusch RM, Gratton E (1991) Quantitation of lipid phases in phospholipid vesicles by the generalized polarization of Laurdan fluorescence. *Biophys J* 60:179–189
5. Parasassi T, De Stasio G, d'Ubaldo A, Gratton E (1990) Phase fluctuation in phospholipid membranes revealed by Laurdan fluorescence. *Biophys J* 57:1179–1186
6. Tyler KM, Wheeler G (2011) Widefield microscopy for live imaging of lipid domains and membrane dynamics. *Biochim Biophys Acta* 1808:634–641
7. Strahl H, Bürmann F, Hamoen LW (2014) The actin homologue MreB organizes the bacterial cell membrane. *Nat Commun* 5:3442
8. Bach JN, Bramkamp M (2013) Flotillins functionally organize the bacterial membrane. *Mol Microbiol* 88:1205–1217
9. Chico DE, Given RL, Miller BT (2003) Binding of cationic cell-permeable peptides to plastic and glass. *Peptides* 24:3–9
10. New RRC (1990) *Liposomes: a practical approach*. IRL Press, Oxford
11. Ingram LO (1977) Changes in lipid composition of *Escherichia coli* resulting from growth with organic solvents and with food additives. *Appl Environ Microbiol* 33:1233–1236
12. Strahl H, Hamoen LW (2010) Membrane potential is important for bacterial cell division. *Proc Natl Acad Sci U S A* 107:12281–12286

Chapter 11

In Vitro Assays to Identify Antibiotics Targeting DNA Metabolism

Allan H. Pang, Sylvie Garneau-Tsodikova, and Oleg V. Tsodikov

Abstract

DNA metabolism embodies a number of biochemical pathways, which include targets of clinically used antibiotics as well as those that are only being explored as potential targets for inhibitory compounds. We give an overview of representative cell-based and enzymatic assays suitable for high-throughput-driven search for novel DNA metabolism inhibitors of established and novel DNA metabolism targets in bacteria. The protocol for a colorimetric coupled primase-inorganic pyrophosphatase assay developed by our group is described in detail.

Key words Inhibitor discovery, High-throughput assay, DNA replication, DNA recombination, Resistance

1 Introduction

With the advent of affordable liquid handling robotics in the last 10–15 years, high-throughput screening (HTS)-driven search for inhibitors or, sometimes, activators of individual enzymes or entire pathways became possible in both industrial and academic settings. HTS assays range from traditional enzymatic assays with one purified enzyme to cell-based or even multicellular organism-based assays that have been scaled down to a microtiter plate format. The miniaturization usually involves optimization of assay conditions to achieve a reproducible signal and a sufficiently high signal-to-noise ratio, while preserving the assay sensitivity to inhibitory compounds. For inhibition of cell growth, DNA metabolism stands out as an excellent target system: Nature has already provided us with scores of natural products (and many more are to be discovered) that evolved as weapons to target DNA or DNA-interacting enzymes to win the battle of survival of the fittest. Several of such compounds (doxorubicin, bleomycin, camptothecin) have been used in clinic as anticancer therapeutics. DNA binding drugs, due to their small size, are not highly DNA sequence-specific and,

consequently, are invariably toxic to noncancerous cells as well. For these reasons, even though many of such molecules have likely evolved as antibacterial or antifungal agents, their use in clinic has been limited to life-threatening cancers. For example, doxorubicin, which functions mainly by inhibiting topoisomerase II in rapidly multiplying cells, also exhibits significant nonselective cytotoxicity due to its low-specificity DNA intercalation. HTS-enabled technologies opened a possibility to screen structurally diverse synthetic and naturally occurring compounds for inhibition of enzymes involved in DNA transactions in bacteria that differ from their counterparts in human and other eukaryotes, offering a promise of finding antibacterial agents that are not toxic to human by such an ab initio approach. Follow-up secondary assays, including cytotoxicity testing, are routinely employed to triage DNA-interacting compounds, to maximize the chances of selective antibacterial toxicity. This chapter will be dedicated to such efforts in targeting bacterium-specific DNA metabolic targets. The exposition is not meant to be comprehensive; we chose several representative assays published in the last 10 years or so. Assays included in this chapter had to satisfy two basic criteria: (1) robustness (as characterized by $Z' > 0.5$ or so) and (2) a detection method not based on radioactivity, to ensure that similar assays could be most broadly used in academia as well as in industry.

In search for selective bioactive inhibitors of enzymatic or regulatory systems (further referred to as “enzymes” for simplicity) involved in bacterial DNA metabolism, as with any targets, two main approaches have been employed: (1) cell-based assays, where the activity of the enzyme is monitored in a cell culture and (2) assays where the activity of a purified (or partially purified) enzyme is monitored, which we will further call “enzymatic assays.” Because DNA metabolic targets reside in the cytoplasm, to get to such targets, the inhibitor needs to penetrate the bacterial cell envelope and possibly withstand efflux, to be useful as a therapeutic or a biological probe. Enzymatic assays have the advantage of a relatively high probability of yielding inhibitors that are known to target the enzyme of interest among HTS hits, but these inhibitors are often not promising for pharmaceutical development, due to their poor penetration through the bacterial cell envelope. In contrast, cell-based assays are less selective and are more prone to yield false-positive hits or off-target hits, but the hit compounds are bioactive. A drawback of cell-based assays is that some inhibitors of the target of interest are missed in such assays, because they cannot go through the cell envelope, but they have a potential for bioactivity through chemical modification or formulation. With our expanding knowledge of chemical features enabling molecules to penetrate the cell envelope, enzymatic HTS assays are expected to become increasingly more attractive. With these considerations in mind, we divided the HTS assays to be discussed here into

cell-based and enzymatic, and within each category we grouped them by the method of detection (Table 1). After this overview, we present details of the HTS assay used for discovery of inhibitors of primase DnaG and inorganic pyrophosphatase that has been developed in our group.

1.1 Cell-Based HTS Assays

HTS by a whole-cell target-directed assay in a bacterium can be carried out if (1) a reactant or a product of the target process is unique and readily observed or measured, or (2) a test bacterium has been mutated or otherwise conditioned to acquire a different sensitivity to the inhibition of the target process from that of the wild-type. Because DNA metabolism involves small molecules (nucleotides, DNA, or RNA oligomers) that participate in a multitude of biochemical reactions and pathways, the above option (1) is not normally available directly. For option (2), such systems can sometimes be obtained by cloning or mutagenesis involving the gene(s) encoding the target. For the genes involved in DNA metabolism, mutations of the gene of interest may sensitize the strain to the inhibitors of the respective protein or, conversely, make the strain more resistant to such inhibition. This naturally

Table 1
Assays discussed in this chapter

Assay	Detection method	Target	References
Cell-based	Fluorescence (GFP)	DNA metabolism	Fan et al. [1]
	Bioluminescence (luciferase)	Gyrase	Moir et al. [5]
	Attenuance	Replication initiator DnaA	Fossum et al. [9]
		AddAB and RecBCD	Amundsen et al. [13]
Enzymatic	Colorimetric	Gyrase	Humnabadkar et al. [17]
		Primase DnaG	Biswas et al. [18]
	Fluorescence yield	DNA adenine methyltransferase	McKelvie et al. [20]
		Ribonuclease reductase	Tholander et al. [21]
		Gyrase	Jude et al. [22]
		Gyrase	Taylor et al. [23]
		Topoisomerase I	Cheng et al. [28]
	FRET	DNA ligase	Shapiro et al. [30] Mills et al. [29]
		Gyrase	Glaser et al. [31]
	Fluorescence polarization	RecA	Peterson et al. [32]
		RecA	Lee et al. [33]

requires that a secondary HTS assay be performed with the wild-type strain with the hits from the primary HTS assay with the mutant strain. Typically, further validation is performed by other orthogonal secondary assays with simpler systems (e.g., by using the purified target enzyme) to further test the inhibition of the target of interest. Whole-cell assays usable for HTS, most commonly, are: (a) reporter-based, where one measures the level of expression of a fluorescent or luminescent protein encoded in a reporter construct that is in some way sensitive to inhibition of DNA metabolism or (b) growth rate detection-based, where one uses a strain bearing a chromosomal mutation in the target gene/promoter or contains a plasmid expressing the target (or its mutant) in such a way that this genetically modified strain is responsive differently to target inhibition than the wild-type strain. The simplest example of assay (b) is a strain in which the expression of the target is upregulated (e.g., by using an overexpression vector). The minimum inhibitory concentration (MIC) for an inhibitor of the target is expected to be greater than that for the parent strain.

1.1.1 Assays Using Fluorescent Biosensors

Fluorescence readout from reporter constructs has been routinely used in studies of transcriptional regulation, but its application to DNA metabolism is not obvious. Fan et al. developed a whole-cell assay for identification of small molecules that inhibit DNA metabolic systems [1]. This assay is based on the potential for induction of the SOS response by a DNA metabolism inhibitor (Fig. 1a). The idea is that stalling DNA replication (e.g., as a result of inhibitor action) with generation of single-stranded DNA would activate RecA [2], which would in turn relieve transcriptional repression by LexA of a number of genes, including *recA* [3]. The assay employed an *Escherichia coli* K-12 MG1655 strain, which was a knockout of a multidrug efflux transporter gene, *tolC*, and which was transformed with a reporter plasmid bearing a green fluorescent protein (GFP) gene under control of a *recA* promoter. Induction of the SOS response in this system resulted in the increase of GFP fluorescence intensity in a dose-dependent manner. After appropriate optimization, this assay in an HTS format had a *Z'* of 0.76, indicating robust performance. Out of 1.2 million compounds from the AstraZeneca compound library that were initially screened in this cell-based assay, 10,262 compounds were identified as primary hits, indicating a robust hit rate. Small molecules possessing undesirable physical and chemical properties based on clogP>6 and molecular weight >550 Da were triaged. The 7101 remaining compounds were tested in a 7-point dose-response assay; this assay also identified autofluorescent false-positive hits. The 1573 remaining compounds were tested for their growth inhibition of *Candida albicans* (an activity spectrum assay) and cytotoxicity against human lung carcinoma A549 cells and red blood cells (cytotoxicity assays). After this rigorous test cascade, 39 hit compounds (in four

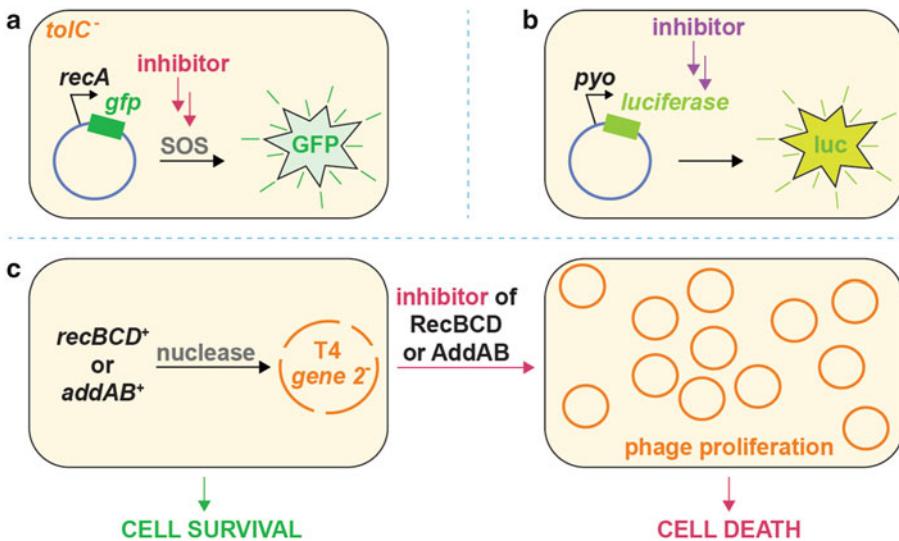


Fig. 1 Schematics of the cell-based assays. (a) Assay for inhibitors of DNA metabolism that trigger the SOS response. (b) Assay for inhibitors of gyrase using upregulation of the pyocin biosynthetic promoter as a marker of gyrase inhibition. (c) Assay for inhibitors of RecBCD and AddAB helicase-nucleases using a T4 phage with the gene 2 knocked out as a sensor of the enzyme inhibition

clusters of structurally related compounds and nine other structurally unique ones) remained as active small molecules. A series of phenylcyclohexylacetic acids was ultimately identified as DNA gyrase inhibitors by a supercoiling assay. The drawback of this assay is that DNA damaging agents, which are of little interest as antibiotics, have a high potential for induction of the SOS response. Furthermore, inhibition of some relevant targets, such as primase DnaG, would stall replication without triggering the SOS response [4]; therefore, inhibitors of such targets cannot be discovered by this assay.

While the above assay was sensitive to inhibitors of several targets (proteins and DNA), a cell-based assay designed by Moir et al. interrogated a specific DNA metabolic enzyme, DNA gyrase [5]. The assay was based on the previously reported finding that a decrease in expression of DNA gyrase (by a knockdown or chemical inhibition) upregulates gene clusters responsible for biosynthesis of pyocins in *Pseudomonas aeruginosa* [6]. For the assay, the authors created a *P. aeruginosa* strain, in which a luciferase gene was inserted into the chromosome downstream of the pyocin biosynthesis promoter (Fig. 1b). As a result, inhibition of gyrase resulted in upregulation of this luciferase gene and increase in luminescence. By using this system, the authors screened 2000 compounds from the Spectrum and Microsource Discovery library consisting of known biologically active drugs and natural products. The HTS was carried out in triplicate: the first two HTS assays were performed over three days (with Z of 0.3–0.48) and the third

one in a single day ($Z' = 0.45$). The Z' values indicated marginally acceptable performance. Out of 2000 compounds, 13 were confirmed to be positive hits. Most of the hits were quinolones (known gyrase inhibitors), but three of them were non-quinolones: nitrofuran, furazolidone, and mechlorethamine. These three compounds form DNA adducts and share the same mechanism of quinolone inhibition against gyrases.

1.1.2 Assays Monitoring Cell Growth of Specific Mutant Bacteria

Many enzymes involved in bacterial DNA replication were discovered based on pioneering studies that selected for and analyzed temperature-sensitive mutants of *E. coli* [7]. One of such mutant strains was able to grow at elevated temperature (42 °C), but ceased its growth at a lower temperature (30 °C) [8]. This phenotype was a result of overreplication, caused by disproportionately high activity of the DNA replication initiator DnaA. Partial inhibition of DnaA by a small molecule is then expected to overcome the effect of this mutation and restore normal growth at 30 °C. Fossum et al. developed an elegant assay using this idea to look for DnaA inhibitors [9]. Since DnaA is an essential gene and is absolutely required for cell growth, the investigators used an *E. coli* mutant in which *rnhA* gene was inactivated, to allow *oriC*-independent replication [10, 11] and preserve cell viability in case of full DnaA inhibition. As a proof of principle, the investigators overexpressed a truncated DnaA protein, which inhibits the assembly of the functional initiation complex and counterbalances the effect of overreplication. Indeed, the authors observed a tenfold increase in growth at 30 °C after induction of truncated DnaA. In their HTS assay, 4199 microbial extracts were tested and cell growth at 30 °C was monitored. The HTS version of the assay had a Z' of 0.55, indicating assay robustness. Despite being an excellent assay, it did not yield hits from a library of microbial fermentation extracts from Vicuron Pharmaceuticals. The investigators reasoned that the lack of hits was likely due to extract library limitations.

RecBCD and AddAB are helicase-nuclease complexes that function in genetic recombination and repair of otherwise lethal double strand breaks in DNA, which are functions commonly required for bacterial virulence. Most bacteria contain only one of these two systems [12]; therefore, they are attractive potential targets. To target specifically either of these two complexes individually, two mutant *E. coli* strains were used—a *recBCD⁺* strain and an *addAB⁺* strain, where either one or the other complex was active [13]. A wild-type T4 phage can replicate in and kill *E. coli*, in part because the T4 DNA is protected from RecBCD and AddAB nucleases by the gene 2 protein bound to the DNA ends of T4 phage virions. A T4 phage lacking gene 2 protein (T4 *gene 2-*) cannot grow in *E. coli*, as the infecting phage DNA gets digested by either RecBCD or AddAB nuclease. A small molecule that inhibits these complexes is then expected to block digestion of the

T4 gene 2 mutant DNA, which, in turn, would allow the phage to kill the host bacterium (Fig. 1c). Amundsen et al. carried out an HTS using 326,100 compounds from NIH Molecular Libraries Small Molecules Repository, with the goal of targeting AddAB. This library contained highly structurally diverse synthetic and natural products from academic and commercial sources. The assay had a Z' of 0.91, indicating robust performance. From the HTS campaign, 937 compounds were considered active hits, but only 885 compounds were tested in the secondary assay due to unavailability of 52 compounds. The secondary assay involved repetition of the primary screening assay to confirm AddAB inhibition. The 885 compounds were counter-screened for RecBCD inhibition using cell viability *recBCD*⁺ strain; cell viability of *addAB*⁺ strain without phage infection was also tested. 225 compounds showed inhibitory activity against either AddAB or RecBCD, and continued to titration assay to identify the IC₅₀ value. The 12 compounds exhibiting the best potency were selected for enzymatic assay against purified AddAB and RecBCD complex. The 12 compounds are grouped into four structurally distinct scaffolds; pyrimidopyridones, cyanothiophenes, nitrofurans, and one nitrothiazole. The compounds were tested for inhibition of double-stranded DNA exonuclease activity of AddAB and RecBCD enzymes, helicase, and Chi-cutting activities of RecBCD. Seven of these compounds were inhibitory of one or the other nuclease in 0.1–50 μM range, none significantly inhibited the helicase activity of AddAB and some compounds inhibited the helicase and the Chi-cutting activities of RecBCD. The 12 initial compounds were also tested against RecBCD or AddAB inhibition using *E. coli* Hfr-based recombination assay and phage λ recombination assay. Five of these compounds resulted in a 2- to 200-fold decrease of the Hfr recombinant frequency.

1.2 Enzymatic Assays

While cell-based assays to probe specific proteins involved in DNA metabolism require genetically modified or transformed strains, enzymatic assays rely solely on the properties of the purified enzyme. The enzyme of interest needs to be sufficiently soluble and stable in the reaction buffer over the course of the HTS assay. HTS can be protein-consuming; therefore, robust yields of protein preparations are often required. Considerable time and effort need to be dedicated to assay optimization in microtiter plates, followed by HTS workflow optimization. Normally, HTS demands that the assay be performed on a ~30-min time scale, to avoid sample evaporation and possibly other artifacts. As with any assay optimization, appropriate buffer and pH, salt, and additives are chosen. Most commonly, HTS is performed as a multiple turnover experiment. For this reason, steady-state kinetic measurements need to be carried out first to measure k_{cat} and K_m and to ascertain that the reaction times lie in the steady-state range of linear product

accumulation over time. For the most reproducible signal, the substrate concentration should be high enough, so that the rate of product formation $V \approx V_{\max}$. On the other hand, the substrate concentration should be low enough to allow competitive inhibition with concentrations of inhibitor that are commonly less than that of the substrate (usually mid- μM in primary HTS assays). We recommend using the concentration of substrate that is three- to fourfold greater than K_m . The HTS readout is most commonly absorbance, fluorescence, or luminescence. In academia, enzymatic HTS assays to identify inhibitors of bacterial DNA metabolism are by far more common than cell-based assays, as the latter not only require a specific genetic system as discussed above, but are also more time-consuming and often involve dedicated biosafety liquid handling equipment.

1.2.1 Assays Employing Colorimetric Detection of Inorganic Phosphate

Detection and quantitative measurements of a phosphate ion (P_i), a product of enzymatic reactions, at low- μM and higher concentrations in a microtiter plate format can be achieved by a number of techniques, which do not involve radioactivity [14–16]. The colorimetric detection of P_i by using the malachite green reagent is commonly employed, as it is facile, quantitative, and free from artifacts arising due to library compound fluorescence. In most cases, either P_i is a product of the reaction catalyzed by the target enzyme or inorganic pyrophosphate PP_i is a target reaction product, which is hydrolyzed into two P_i molecules by a bacterial or yeast inorganic pyrophosphatase, acting as a coupled enzyme in the assay. Increase in absorbance at 620 nm indicates formation of P_i .

The release of P_i by the hydrolytic activity of target ATPases is commonly used in HTS assay. Examples of this approach are gyrase ATPase assays for discovery of inhibitors of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* DNA gyrases [17]. The former enzyme is a rigorously validated target that is, in fact, clinically relevant, as fluoroquinolones, which target *M. tuberculosis* gyrase, are second-line antituberculosis therapeutics. The two assays performed with salmon sperm DNA as a substrate were well executed and had Z of 0.8 and 0.5 for *M. smegmatis* and *M. tuberculosis* gyrase, respectively. *M. smegmatis* DNA gyrase was eventually chosen as the target enzyme for HTS due to the higher Z value and higher enzymatic activity. The authors tested 1,000,000 compounds from the AstraZeneca library. A series of aminopyrazinamide inhibitors was identified with IC_{50} of $\sim 2 \mu\text{M}$. The HTS also yielded a series of pyrrolamides and a series of thiazolopyridine ureas from a pharmacophore-based subset of AstraZeneca library and the broad-spectrum antibacterial program AstraZeneca library, respectively. All these hits were first tested for inhibition of growth of *M. tuberculosis* bacteria in vitro and for inhibition of purified *M. tuberculosis* gyrase in the supercoiling activity assay. Further assays included inhibitor binding affinity measurements by melting temperature shift and surface plasmon resonance experiments.

A novel potential target, bacterial primase DnaG, is another enzyme required for chromosomal DNA replication. Bacterial primase is structurally distinct from the eukaryotic primase, which suggests that DnaG-binding inhibitors would not be likely to inhibit DNA replication in human. DnaG is an RNA polymerase; therefore, the released product is PP_i, not P_i, requiring coupled hydrolysis by inorganic pyrophosphatase for P_i formation and detection. Based on these principles, a coupled primase-pyrophosphatase colorimetric assay, employing *M. tuberculosis* DnaG and *M. tuberculosis* inorganic pyrophosphatase, both potential targets, was optimized and used for HTS by the authors of this chapter [18], with a Z' of 0.7 (Fig. 2a). We provide details of the protocol for this assay in the following sections. A library of 2556 chemical compounds consisting of 2000 FDA approved drugs from Spectrum Chemicals, 450 NIH Clinical collection human therapy drugs, and 106 kinase inhibitors from Enzo Life Sciences was screened. The HTS yielded nine positive hits; among them were suramin, doxorubicin, and ellagic acid. These compounds were confirmed to be inhibitors of DnaG by the primase activity assay directly monitoring synthesis of radioactively labeled RNA primers. The mode of inhibition was then established by a series of dose-response assays with varying concentrations of DNA template and NTP substrates. A similar HTS assay was carried out with *Bacillus anthracis* DnaG by the same group [19]. This HTS assay was also robust ($Z'=0.6$) and yielded nine hit compounds, including doxorubicin, epirubicin, tilorone, three fluoroquinolones (levofloxacin, enrofloxacin, and rufloxacin), and suramin. Doxorubicin and suramin were inhibitory to both *M. tuberculosis* and *B. anthracis* DnaG, likely due to the same mechanisms, which involve DNA stacking or intercalation for doxorubicin and heparin-like binding of polyanionic suramin to DnaG. In contrast, tilorone was a potent (low- μM) inhibitor specific to *B. anthracis* DnaG, whereas ellagic acid was specific to *M. tuberculosis* DnaG. Treating *B. anthracis* cell culture with these inhibitors indicated that tilorone and suramin were unable to penetrate the cell wall or were exported out by efflux, while doxorubicin displayed a bacteriostatic effect, consistent with inhibition of DNA replication.

1.2.2 Assays Using Fluorescence Intensity as a Readout

McKelvie et al. designed a double-stranded break light hairpin oligonucleotide bearing a *Dpn*I recognition site GATC, a fluorescein, and a dabcyll quencher on the two termini, and an N₆-methyladenine in one of the two complementary GATC sites (Fig. 2b) [20]. The investigators used this oligonucleotide to monitor the function of *Yersinia pestis* DNA methyltransferase (Dam). Briefly, the hemimethylated oligonucleotide becomes fully methylated by the enzymatic action of Dam. This fully methylated oligonucleotide is then cleaved by *Dpn*I, which functions only when both complementary GATC sites are methylated. The cleavage separates the fluorescein

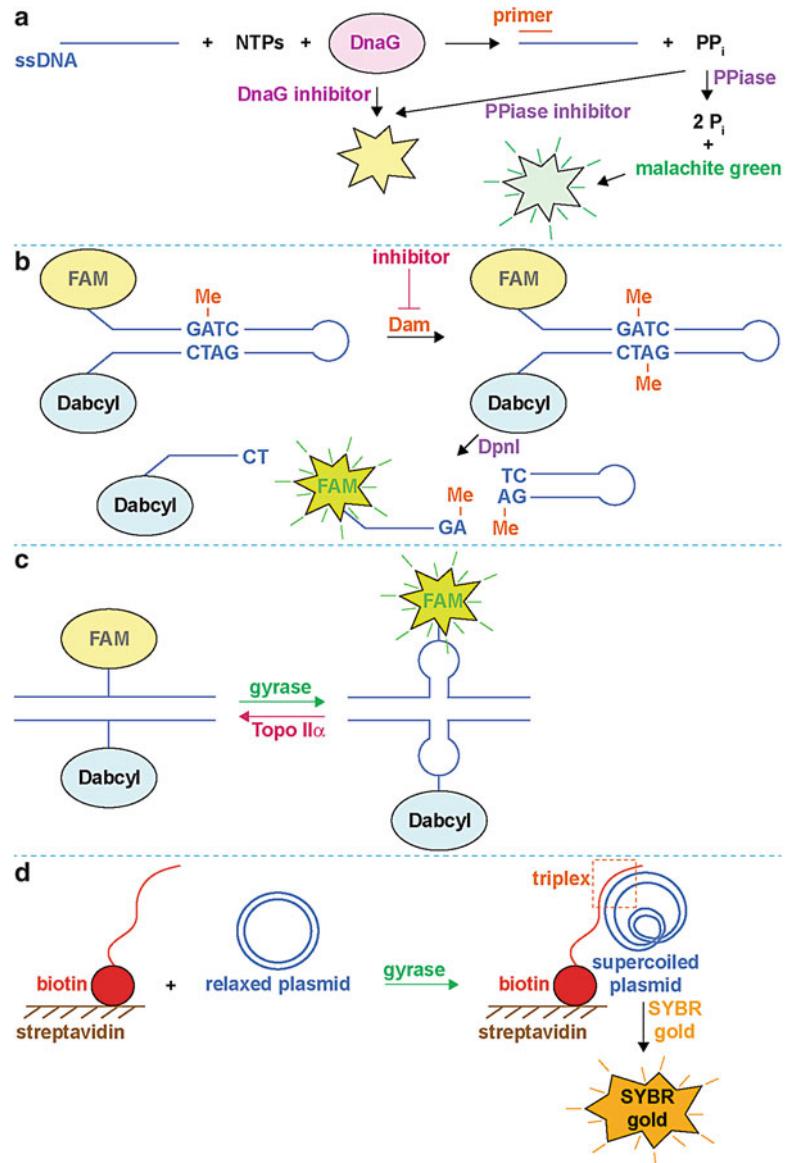


Fig. 2 Schematics of enzymatic assays. (a) Coupled primase-inorganic pyrophosphatase colorimetric assay that can be used to search for inhibitors of primase DnaG and inorganic pyrophosphatase in one assay. (b) Assay for inhibitors of Dam methyltransferase/DpnI nuclease harnessing the thermodynamic instability of a short duplex after DNA cleavage. (c) Assay for inhibitors of gyrase using cruciform extrusion as a product of gyrase activity that is amenable to detection. (d) Assay for inhibitors of gyrase using the ability of the supercoiled, but not relaxed plasmid to form DNA triplex with a specific immobilized oligonucleotide

and the dabcyt probes, thereby unquenching the fluorescein fluorescence (Fig. 2b). Comparison with the positive control (omitting SAM, the methyl donor for Dam methylation) shows an excellent screening window with a Z' of 0.71. 3082 compounds were subjected to HTS to find suitable Dam inhibitors. Six hit compounds were identified. Five of these were discarded due to their DNA intercalation, poor solubility, impure samples, or presence of inhibitory zinc. Finally, stibonic acid was subjected to further structure-activity-relationship studies. A series of stibonic acid analogues were created and tested in the Dam activity assay and in a cell culture growth inhibition assay.

Tholander et al. developed a PCR-based assay to screen for small molecule inhibitors against ribonucleotide reductase (RNR) [21]. In the assay, NDP is first reduced by RNR from NDP to dNDP, followed by its conversion to dNTP by a coupled action of NDP kinase (NDPK). The dNTP product is then used as a substrate in a PCR. The formation of the PCR product is monitored by using the SYBR green dye, whose fluorescence is enhanced upon binding DNA. A library of 1364 diverse compounds from the National Cancer Institute (NCI) was tested in a robust HTS assay with a Z' of 0.86 for their inhibition of a Class I RNR from *P. aeruginosa*. 110 compounds tested at 100 μ M were identified as initial hits, inhibiting more than 50% of RNR activity. 27 of these compounds, which exhibited over 90% inhibition, were selected for dose-response analysis by a radiolabeled CDP assay, chromatographic purification of formed dCDP, and liquid scintillation counting. Five anthraquinone-like, ten naphthoquinone-like, and four phenol-containing small molecules were noted to contain functionalities related to redox activity, known to affect or inhibit RNR activity. The remaining eight structurally diverse compounds were tested for their inhibition of cell growth of *P. aeruginosa* cultures.

By harnessing the activities of gyrase and human topoisomerase II α in introducing and removing negative supercoils, respectively, Jude et al. developed an assay that allows detection of these activities [22]. The assay employed a plasmid containing a 42-bp AT repeat, in which a fluorescein and its quencher dabcyt were incorporated near each other on the opposite strands of the DNA. Gyrase causes negative supercoiling of this plasmid, resulting in the extrusion of the repeat as a cruciform, separating the two probes, and increasing fluorescein fluorescence. The subsequent treatment with topoisomerase II α removes the supercoils and the cruciform, thereby quenching the signal (Fig. 2c). The gyrase HTS assay had a Z' of ~0.8. The investigators screened 446 compounds from the NIH Clinical Collection and obtained seven hits, which were all fluoroquinolones. Secondary assays to confirm these hits monitored product formation directly on an agarose gel. The gyrase assay yielded a large number of false positives due to intrinsic fluorescence of the compounds or their interference with probe

fluorescence. HTS on a larger scale by using this assay has not been performed yet.

A different gyrase/topoisomerase VI activity assay suitable for HTS was developed by another group [23]. This assay is based on the formation of a DNA triplex by binding of an immobilized oligonucleotide to a supercoiled, but not to a relaxed, plasmid. The supercoiling could be introduced by action of either gyrase or topoisomerase VI. In the HTS assay, a microtiter plate was coated with streptavidin, onto which a biotinylated triplex-forming oligonucleotide was immobilized. Therefore, a supercoiled plasmid would get immobilized as it binds the immobilized oligonucleotide to form a triplex region. Similarly to the previously described PCR assay, the presence of the immobilized plasmid after a wash was monitored by SYBR gold staining (Fig. 2d). The average Z 's in the assays with *E. coli* DNA gyrase and archaeal *Methanosarcina mazei* DNA topoisomerase VI were 0.64 and 0.69, respectively. A total of 960 compounds, FDA-approved drugs, and natural products were used to test for inhibition against supercoiling and relaxing activities of DNA gyrase and topoisomerase VI, respectively. 22 compounds were identified as positive hits for DNA gyrase inhibition. These were fluoroquinolones novobiocin and acriflavinium, both known DNA gyrase inhibitors. A secondary assay detecting supercoiling by DNA gyrase on a gel was carried out for the remaining inhibitors. Among those, only two, mitoxantrone and suramin, displayed inhibition. Nalidixic acid, cinoxacin, oxolinic acid, and enoxacin, which are four known gyrase inhibitors, were not identified as hit in the screen, likely due to their low-potency. A total of nine compounds were detected as inhibitors for DNA topoisomerase VI. The inhibitory action of six compounds (amsacrine, suramin, hexylresorcinol, 9-aminoacridine, purpurin, and quinacrine) was confirmed by an in-gel relaxation assay. Amsacrine was previously reported to be a topoisomerase VI [24] and a topoisomerase IIA [25] inhibitor. Suramin and quinacrine were also found to inhibit topoisomerase IIA [26, 27]. Purpurin and 9-aminoacridine are structurally related to mitoxantrone and amsacrine, respectively. Both mitoxantrone and amsacrine are known topoisomerase II inhibitors. Mitoxantrone was not identified in the HTS because the small molecule disrupts formation of DNA triplex and therefore this provides a good example of the limitation of the HTS.

Cheng et al. developed an HTS assay to identify inhibitors targeting bacterial topoisomerase I and optimized it with *Y. pestis* topoisomerase I [28]. In this assay, a stem-loop oligonucleotide bearing a 5'-CAL Fluor Red 610 fluorophore and a 3'-BHQ-2 quencher is cleaved by *E. coli* and *Y. pestis* topoisomerase I. The cleavage action of topoisomerase I separates the two probes resulting in an increase in fluorescence. Lack of the fluorescence increase would then indicate inhibition of the cleavage activity of the enzyme. The HTS assay was validated by using an *E. coli* topoisomerase I point mutant (G116S)

defective in re-ligating DNA after the cleavage step. The assay with the mutant yielded a greater than fivefold increase in fluorescence, indicating efficient separation of the two probes. The assay was robust ($Z \sim 0.7$). 7105 small molecules and 2816 natural product extracts from NERCE/NSRB were screened in the HTS assay. A total of 30 positive hits (20 small molecules and 10 natural product extracts) were identified from the HTS campaign. The ten natural product extracts were subjected to a secondary assay, to test for inhibition of relaxation activity of topoisomerase I of supercoiled plasmid. Two natural product extracts, both of which were from lichen *Hyopterachyna* sp., inhibited topoisomerase I. HPLC fractionation yielded anziaic acid as a topoisomerase inhibitor. Anziaic acid had robust antibacterial activity against *Bacillus subtilis* (MIC = 6 $\mu\text{g}/\text{mL}$), but was not active against *E. coli*, unless its cell envelope was compromised. Anziaic acid exhibited a ~10-fold preference in inhibiting bacterial topoisomerase I over human topoisomerase I in vitro, acting as a topoisomerase poison against the former. This compound was also tested for inhibition of *E. coli* DNA gyrase and human topoisomerase II α and was shown to inhibit both at IC₅₀ of ~20 $\mu\text{g}/\text{mL}$. Anziaic acid exhibited cytotoxicity against mammalian pulmonary artery endothelial cells with an MIC of 21 μM .

An assay employing fluorescence resonance energy transfer (FRET) for detection was developed to identify small molecule inhibitors of DNA ligases [29, 30]. In this assay, two short oligonucleotides were base-paired with a longer oligonucleotide followed by the nick sealing action of the ligase. The longer oligonucleotide contained a 5'-fluorescein (FAM) donor, while one of the complementary shorter oligonucleotides contained a 3'-TAMRA acceptor. The ligated product was stably annealed to the complementary strand, whereas unligated oligonucleotides were not; therefore, the ligation resulted in FRET between the two chromophores. When the ligase was inhibited, the unligated substrate oligonucleotide would not stay annealed to the FAM-labeled, and no FRET signal would ensue. The assay in the HTS format had a Z of 0.8. As a proof of principle, an AstraZeneca corporate library was tested against NAD-dependent DNA ligase (LigA) from *Haemophilus influenzae*. The assay yielded 517 hits. One structural series composed of adenosine analogues was investigated, demonstrating selective inhibition of bacterial DNA ligases from *H. influenzae*, *E. coli*, *Streptococcus pneumoniae*, *Staphylococcus aureus*, and *Mycoplasma pneumoniae* with IC₅₀ values in the range 0.01–1.27 μM . The compounds did not inhibit human ATP-dependent DNA ligase 1 or T4 bacteriophage ligase. These adenosine analogues were also tested for antibacterial activity against a number of bacterial species. It was found that some of the analogues exhibited poor permeability of the cell envelope of gram-negative bacteria. The molecules were not cytotoxic to eukaryotic cells: human cell line A549 and *Candida albicans*. To determine

the mode of action, one of the adenosine analogues was cocrystallized with *H. influenzae* LigA, identifying that the inhibitor binds in the nucleotide-binding site. The inhibitor efficacy was also evaluated in mouse *S. aureus* and *S. pneumoniae* infection models.

1.2.3 Fluorescence Polarization Assays

Fluorescence polarization (FP) is a measure of mobility of a fluorescent probe in solution. When a probe is excited with a polarized fluorescent light, emitted fluorescence stays more polarized when a probe is bound to a large ligand than when it is free in solution. In search of new inhibitors of *Francisella tularensis* gyrase, Glaser et al. developed a competition assay [31]. In this assay, the authors monitored displacement of a known inhibitor novobiocin bound to the ATP-binding site on the GyrB subunit of gyrase by another compound. This effect was observed by measuring fluorescence polarization of a Texas Red-X fluorescent dye synthetically conjugated to novobiocin. The assay ($Z = 0.8$) was used in HTS to screen 1120 FDA-approved drugs and antibiotics. From this HTS, four non-novobiocin inhibitors were found. These four inhibitors were all anthracyclines: doxorubicin, idarubicin, epirubicin, and daunorubicin. A secondary activity assay was carried out by testing for inhibition of plasmid supercoiling by *E. coli* gyrase.

Another FP-based assay used a commercial kit (Transcreener ADP FP assay kit; BellBrooks Laboratories) [32]. This assay can be used to measure ATP hydrolysis, as an FP increase results from binding of a fluorescently labeled ADP product (but not ATP) to an ADP-specific antibody. Peterson et al. used this robust assay ($Z = 0.92$) to screen 113,477 compounds (a collection from the University of North Carolina) for their inhibition of the ATPase activity of *E. coli* RecA. The 113,477 compounds are composed of 108,750 100 K diversity compound collection and 4727 kinase focus set. From the 100 K diversity compound collection, 235 compounds were considered hits, showing $> 30\%$ inhibition, while the kinase focus set yielded 14 compounds exhibiting $\geq 43\%$ inhibition. A concentration-response study was carried out for these hits with 79 inhibitors having an $IC_{50} \leq 10 \mu M$. Further evaluation identified 35 of these inhibitors to be active in biological growth inhibition assays with *E. coli* cells, 15 of which were selected for measurement of ATPase reaction rate.

Lee et al. developed an FP assay for inhibitors that target specifically active conformations of RecA [33]. RecA monomers when unbound from DNA and ATP are in an inactive state, but in the presence of DNA and ATP, RecA forms filaments, in which its conformation is active. There are two types of active RecA filaments. The first one is a RecA filament on a single-stranded DNA substrate where RecA may be bound to ATP. In this state, the active filament is capable of a signaling activity triggering the SOS response within the cell. The second active conformation state requires additional DNA strands on which the RecA monomers

form a filament assembly in the presence of ATP. In this state, the RecA filament has a motor-like activity requiring hydrolysis of ATP. Two types of inhibitors can be identified. One can look for an inhibitor that disrupts and prevents assembly of the filament, and, therefore, abolishes both signal-like and motor-like functions of RecA. Another type of inhibitor that can be identified would target specifically the second active filament conformation, preventing the motor-like activities of the RecA. To utilize fluorescence polarization to detect RecA-DNA filament formation, a fluorescein-labeled 32-bp ssDNA was used as a substrate. Formation of RecA-labeled DNA filament would produce a low fluorescence signal and a high fluorescence polarization signal. Whereas, if an inhibitor prevents the formation of active filaments, the total fluorescence signal is expected to be high while the fluorescence polarization would be low. This assay was tested in a 96-well plate with positive control reactions containing 500 μ M ADP, which acts as an inhibitor. The results showed that this assay was suitable for HTS of potential inhibitors of RecA-DNA binding, as evidenced by a Z' of 0.87.

With the ever-increasing threat of antibiotic resistance, our search for novel antibiotics continues. HTS approaches are becoming more and more routine as a discovery tool for potential lead chemical matter for antibiotic development, in both industry and academia, due to affordability of liquid handling robotic systems. The success of an HTS campaign now rests mostly on the robustness of the assay and the quality of the screening library (structural diversity and drug-like properties of the chemicals, good library maintenance). The Z' factor is typically used as a measurement of the quality of the HTS assay with realistic positive inhibition controls, and its values between 0.5 and 0.9 serves as a criterion of assay robustness [34]. One can screen a large number of structurally diverse compounds or choose a library that is enriched in certain structures (e.g., nucleotides or peptidomimetics) that are expected to yield a higher hit rate against a particular target. One can also utilize screening libraries that consist of chemical scaffolds that are known to be biologically active and/or possess drug-like properties. Such libraries are enriched in natural products, extracts, or known bioactive compounds. In considering a library for a particular assay, one should strive to avoid libraries that contain compounds that can potentially produce too many false positives. Such interference can result either from intrinsic fluorescence of the library compound or its interaction with the assay probe. Likewise, if one relies on tracking phosphate release, it is best to avoid libraries that are enriched in compounds that are prepared as phosphate salts.

HTS of inhibitors of DNA metabolic enzymes may yield non-specific inhibitors that target DNA itself or are NTP-like or poly-anionic and would bind multiple DNA-binding proteins. For instance, quinolones have been found to be potent inhibitors of

DNA gyrase [5, 22] as well as RecBCD and AddAB helicase-nucleases [13]. Anthracyclines inhibit primase [18, 19, 35] and gyrase [31]. The polyanionic suramin also inhibits primase [18, 19] and gyrase, apparently by binding the enzymes. Substrate analogues can also emerge as hits, as exemplified by the discovery of adenosine analogues as inhibitors of NAD-dependent ligases [29]. HTS campaigns have yielded novel inhibitors of DNA metabolism. For instance, a series of phenylcyclohexylacetic acids was identified as DNA gyrase inhibitors [1] and stibonic acid as a DNA methyltransferase inhibitor [20]. HTS of crude lichen extracts identified anziaic acid as an effective inhibitor of bacterial topoisomerase I [28]. Extracts contain biologically active compounds, but do not always yield hits, as in the HTS campaign targeting DnaA [9].

In summary, several robust experimental HTS assays are now available and more are being developed for identification of novel inhibitors of validated or potential targets in bacterial DNA metabolism pathways. As the assays become more and more robust and the HTS technology becomes broadly available, we can expect many useful inhibitors to emerge.

2 Materials

Below we describe a detailed protocol for the previously reported colorimetric coupled primase-inorganic pyrophosphatase HTS assay [18] for identification of inhibitors of bacterial primase DnaG and inorganic pyrophosphatase (PPiase). We expect this protocol to be applied with little to no additional optimization to DnaG and PPiase enzymes of other bacteria.

Only ultrapure water should be used for solution preparation. The reagents should be stored at room temperature, unless specified otherwise.

2.1 General Solvents

1. 4 M NaCl. Add water to 234 g of NaCl to the final volume of 1 L.
2. 1 M MgCl₂. Add water to 20.3 g of MgCl₂ hexahydrate to the final volume of 100 mL.
3. 1 M MnCl₂. Add 0.2 g of MnCl₂ tetrahydrate to 1 mL of water. Store at -20 °C.
4. 1 M Tris-HCl pH 8.0. Add water to 121.1 g of Tris base (Trizma) to the volume of 800 mL, add slowly 42 mL of concentrated HCl on a stir plate in a fume hood. Add water to the final volume of 1 L.
5. 4 M imidazole. Add water to 209.1 g of imidazole to the final volume of 500 mL.

6. 100 mM ATP Mg²⁺ salt. Add water to 0.51 g of ATP Mg²⁺ salt to the final volume of 10 mL. Store at -20 °C.
7. 100 mM phenylmethylsulfonyl fluoride (PMSF) in ethanol. Add 17.4 g of PMSF to 1 mL of ethanol. Store well sealed at -20 °C.
8. 1 M HEPES pH 7.5. Add water to 260.3 g of HEPES Na⁺ salt to 800 mL. Stir the solution to dissolve the HEPES. Add concentrated HCl to adjust pH at 7.5 upon stirring. Add water to the final volume of 1 L.
9. 0.5 M EDTA solution. Add 80 mL of water to 18.6 g of EDTA. Mix the solution thoroughly. Add NaOH dropwise until pH reaches 8.0. Adjust the final volume to 100 mL by adding water.
10. 1 M N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) pH 8.8. Dissolve 11 g of CAPS in 35 mL of water. Adjust pH with NaOH to 8.8 and add water to the final volume of 50 mL.
11. 600 mM potassium glutamate. Add water to 6 g of l-glutamic acid potassium salt monohydrate (KGlu) to the final volume of 50 mL.

2.2 Malachite Green Reagent Components

1. 0.0812% w/v malachite green, dissolved in water. Add water to 0.0812 g of malachite green to a volume of 100 mL.
2. 2.32% w/v polyvinyl alcohol, dissolved in water. Add water to 2.32 g of polyvinyl alcohol to a volume of 100 mL. To dissolve polyvinyl alcohol slowly heat up the mixture with stirring on a stir plate to 80–90 °C. Do not allow the mixture to boil. Then it should take ~2 h for polyvinyl alcohol to dissolve completely. Wrap the solution in foil and store at 4 °C. The solution should be stable for 1–2 months, until filamentous aggregates appear. This solution should be warmed up to room temperature before use.
3. 5.72% w/v ammonium molybdate, dissolved in 6 M HCl. Prepare 100 mL 6 M HCl by adding concentrated HCl (12 M) into 50 mL water (in that order) to a volume of 100 mL. Dissolve 5.72 g ammonium molybdate in this solution.
4. 10% sodium citrate, dissolved in water. Add water to 1 g of sodium citrate to a volume of 100 mL.

2.3 Protein Purification Reagents and Equipment

2.3.1 *M. tuberculosis* DnaG (*Mtb* DnaG) Purification Reagents

1. Lysogeny broth (LB) and LB agar for cell culture growth (needed to prepare both DnaG and inorganic pyrophosphatase). Autoclave before use.
2. Commercially available 15% SDS-PAGE gels, coomassie blue stain, 2× protein sample buffer, and protein molecular weight standards in the range 10–100 kDa.

3. 100 mg/mL ampicillin stock (1000×). Dilute 1 g of ampicillin in 10 mL of water. Store at -20 °C. To prepare the broth and agar for bacterial culture growth, thaw this stock and add 1 mL of this stock to 1 L of autoclaved medium or agar after they cooled down to ~40 °C and before the agar starts solidifying.
4. Basic Buffer: 40 mM Tris-HCl pH 8.0, 600 mM NaCl, 10% glycerol, 2 mM MgCl₂. Combine 40 mL of 1 M Tris-HCl pH 8.0, 150 mL of 4 M NaCl, 100 mL of glycerol, 2 mL of 1 M MgCl₂ and add water to the final volume of 1 L. Mix well.
5. DnaG Lysis Buffer: 40 mM Tris-HCl pH 8.0, 600 mM NaCl, 10% glycerol, 1 mM PMSF, 0.5 mM ATP (Mg²⁺ salt), 2 mM MgCl₂, 2 mM β-mercaptoethanol, protease inhibitors (*see Note 1*). Immediately before to use, chill the required volume of Basic Buffer (25 mL of buffer per cell pellet from 1 L of bacterial culture) on ice. Then, to 100 mL of chilled Basic buffer (scale up or down as needed) add 500 μL of 100 mM ATP, 1 mL of 100 mM PMSF, 14 μL of β-mercaptoethanol, and two tablets of complete ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche Applied Science). Mix well. Use this ice-chilled Lysis Buffer for cell pellet suspension and lysis.
6. DnaG IMAC Buffer A: 40 mM Tris-HCl pH 8.0, 600 mM NaCl, 10% glycerol, 0.5 mM ATP, 2 mM MgCl₂, 2 mM β-mercaptoethanol, and 50 mM imidazole (*see Note 2*). First, prepare 20 mL of buffer by following the DnaG Lysis buffer protocol in **item 2** above (scaled down fivefold), omitting PMSF and protease inhibitors. Add 1.25 mL of 4 M imidazole to 100 mL of this buffer. Mix well. This buffer is to be used chilled on ice immediately after its preparation.
7. DnaG IMAC Buffer B: approximately, 40 mM Tris-HCl pH 8.0, 600 mM NaCl, 10% glycerol, 0.5 mM ATP, 2 mM MgCl₂, 2 mM β-mercaptoethanol, 500 mM imidazole. First, prepare a buffer by following the Lysis buffer protocol in **item 2** above, omitting PMSF and protease inhibitors. Add 2.5 mL of 4 M imidazole to 17.5 mL of this buffer. This buffer is to be used chilled on ice immediately after its preparation.
8. DnaG Gel Filtration Buffer: 40 mM Tris-HCl pH 8.0, 600 mM NaCl, 10% glycerol, 0.25 mM ATP, 2 mM MgCl₂ and 2 mM β-mercaptoethanol. Degas and cool down 1 L of Basic Buffer at 4 °C. Immediately before use, add 140 μL of β-mercaptoethanol and 2.5 mL of 100 mM ATP, mix well.
9. DnaG HTS Buffer: 40 mM HEPES pH 7.5, 600 mM NaCl, 10% glycerol, 2 mM MgCl₂, 2 mM β-mercaptoethanol and 0.1 mM ATP (pure, Na⁺ salt). On ice, mix 4 mL of 1 M HEPES, 15 mL of 4 M NaCl, 10 mL of glycerol, 200 μL of

1 M MgCl₂, 100 µL of 100 mM ATP (Na⁺ salt) and 14 µL of β-mercaptoethanol. Add water to the final volume of 100 mL. Scale down this recipe as needed. Use this buffer immediately after preparation.

10. PreScission protease.

2.3.2 M. tuberculosis Inorganic Pyrophosphatase (Mtb PPiase) Purification Reagents

1. pJ411-*Mtb* PPiase expression vector (made available from the laboratory of Dr. Luiz Pedro Carvalho).
2. PPiase Lysis Buffer: 20 mM triethanolamine (TEA), pH 7.8. Prepare 10× lysis buffer (200 mM TEA, pH 7.8) by diluting 3.04 mL of triethanolamine in 80 mL water, adjust pH to 7.8 with HCl, add water to 100 mL. Use after tenfold dilution.
3. PPiase IMAC Buffer A: 20 mM TEA buffer pH 7.8, 50 mM imidazole, 300 mM NaCl. Combine 10 mL of 10× PPiase Lysis Buffer, 1.25 mL of 4 M imidazole and 7.5 mL of 4 M NaCl and add water to the volume of 100 mL.
4. PPiase IMAC Buffer B: 20 mM TEA buffer pH 7.8, 500 mM imidazole, 300 mM NaCl. Combine 10 mL of 10× PPiase Lysis Buffer, 12.5 mL of 4 M imidazole, and 7.5 mL of 4 M NaCl and add water to the final volume of 100 mL.
5. PPiase Gel filtration Buffer: 40 mM Tris-HCl pH 8.0, 100 mM NaCl. Add water to 40 mL of 1 M Tris-HCl pH 8.0 and 25 mL of 4 M NaCl to the final volume of 1 L. Degas this buffer and store it at 4 °C.

All the above solutions should be mixed thoroughly, to full dissolution of solid components and homogeneity. A stir plate is advisable to use for dissolving solids in volumes of 50 mL and larger.

2.3.3 Equipment for Protein Expression and Purification

1. A sonicator, a french press, or a cell homogenizer for cell disruption.
2. A refrigerated cell culture shaker.
3. A centrifuge for pelleting cells and clarifying lysate (speed up to 38,000 ×*g*).
4. A gel running apparatus for SDS-PAGE gels with a suitable power supply.
5. Fast performance liquid chromatography (FPLC) system for size-exclusion chromatography equipped with a fraction collector.
6. 0.45 µm Millex-HV PVDF filters.
7. 5 mL IMAC FPLC columns, e.g., HiTrap FF.
8. Gel filtration FPLC columns, e.g., HiPrep 26/60 Sephadryl S-200 column.

9. Amicon Ultra-15 centrifugal filter device.
10. MicroSpin G-25 columns.

2.4 Activity Assay Reagents

1. Purified recombinant homogenous proteins *Mtb* DnaG (72 μ M) and *Mtb* PPiase (750 μ M). The protocols for protein expression, purification, and storage are described below.
2. 1 μ mol of a 32-nucleotide DNA template 5'-GAAGCACCAGA CGTTTAGCATACATTCACAGA-3'. Centrifuge the DNA briefly at $5000 \times g$, dissolve it in water to a concentration of 2 mM (in strands), and store at -20 °C. The exact amount of the synthesized DNA varies; therefore, the amount of water should be adjusted accordingly.
3. 100 mM of ribonucleotide triphosphates, rCTP, rGTP, and rUTP from Promega. NTP from this company (unlike tested other suppliers) are consistently free from contamination with phosphate, which is critical for the assay. Please, note that rATP is not used in the assay, to avoid signal from potential ATPase contaminants that can be present in trace amounts in protein preparations.
4. A suitable set of test compounds as 2 mM stocks in DMSO on a 384-well plate.
5. 3x DNA-enzyme mixture: 60 mM CAPS pH 8.8, 6 mM MnCl₂, 3 mM MgCl₂, 150 mM NaCl, 450 mM KGlu, 2.1 μ M *Mtb* DnaG, and 150 nM *Mtb* PPiase. Mix 1.8 mL of 1 M CAPS pH 8.8, 180 μ L of MnCl₂, 90 μ L of 1 M MgCl₂, 1.125 mL of 4 M NaCl, and 22.5 mL of 600 mM KGlu. Add to this mixture 56.25 μ L of 2 mM 32-nucleotide DNA template, then 6 μ L of *Mtb* PPiase (750 μ M) and 875 μ L of *Mtb* DnaG (72 μ M), followed by 3.5 mL of water.
6. 3x NTP mixture: 330 μ M of CTP, GTP, and UTP. 100 μ L of each of the three 100 mM NTP (CTP, GTP, and UTP) stock was added to 30 mL of water.
7. 2 mM suramin. Dissolve 1.43 μ g of suramin in 500 mL of water.

2.5 Assay Equipment

1. 384-well flat bottom, non-binding surface, polystyrene plates for absorbance measurements.
2. A Biomek FX robot or a comparable liquid-handling instrument.
3. A 384-well microplate reader.
4. Multichannel pipette suitable for dispensing 1–20 μ L into 384-well plates.
5. A vibrating platform shaker for shaking microtiter plates.

3 Methods

3.1 Preparation of the Malachite Green Reagent

1. Prepare the malachite green reagent by mixing two parts of 0.0812% malachite green (see Note 3), two parts of water, one part of 2.32% w/v polyvinyl alcohol, and one part of 5.72% ammonium molybdate, prepared as described above. For example, prepare 90 mL of malachite green reagent by mixing 30 mL of malachite green, 30 mL of water, 15 mL of polyvinyl alcohol, and 15 mL of ammonium molybdate.
2. Shield the mixture from light (wrap it in aluminum foil) and incubate at room temperature for 15–30 min, until it turns golden-yellow.
3. Store the reagent (shielded from light) at 4 °C. The reagent is normally stable for ~1 month under these conditions.

3.2 Preparation of *Mtb* DnaG

1. Clone *Mtb dnaG* gene by standard recombinant DNA techniques. Briefly, amplify the gene from *Mtb* H37Rv genomic DNA with primers 5'-AGTTAGCACATATGTCCGGCGG ATCTCCG-3' and 5'-CCGCTCGAGTCACGCGGTGAGA TCG-3', and insert the product into a pET19b-pps vector [36] using *Nde*I and *Xho*I sites. The resulting expression vector contains a recombinant protein with an N-terminal deca-histidine tag cleavable by PreScission protease.
2. Transform the plasmid into *E. coli* BL21(DE3) cells and allow them to grow overnight on a plate with LB agar supplemented with 100 µg/mL ampicillin.
3. Pick five colonies and inoculate them in a 4 mL of LB broth supplemented with 100 µg/mL ampicillin (LB/Amp).
4. Shake this starter culture at 200 rpm at 37 °C for 4 h.
5. Inoculate 1 mL of bacterial culture into 1 L of LB/Amp broth. For this preparation, use 4 L of LB/Amp broth (in four 1 L cultures).
6. Incubate the LB broth with shaking at 220 rpm at 37 °C until an optical density of 0.2 at 600 nm is reached.
7. Move the cultures to an incubator-shaker at 18 °C and incubate similarly for 1.5 h.
8. Then induce protein production by addition of IPTG at the final concentration of 0.5 mM.
9. After 16 h of shaking at 220 rpm at 18 °C, harvest the cells by centrifugation at 5000×*g* for 10 min at 4 °C.
10. Resuspend the cells in 100 mL of DnaG Lysis buffer and disrupt the cells on ice by intermittent sonication.
11. Clarify the lysate by centrifugation at 38,000×*g* for 60 min at 4 °C.

12. Filter the supernatant through 0.45 μm PVDF filters.
13. Load the filtrate onto a 5 mL IMAC Ni^{2+} column equilibrated in DnaG Lysis Buffer.
14. Wash the column with 100 mL of DnaG IMAC Buffer A, and elute the protein with 12 mL of DnaG IMAC buffer B; collect the first 2 mL of the eluate separately and do not use this fraction in the next steps.
15. Adjust the NaCl concentration of the 10 mL of protein-containing eluate to 200 mM by adding 20 mL of DnaG IMAC buffer A that was prepared without NaCl.
16. Cleave the His tag overnight at 4 °C after adding PreScission protease.
17. Concentrate the digestion reaction by using an Amicon Ultra-15 centrifugal filter device with M_w cutoff of 10,000 Da to the volume of 5 mL.
18. Purify the primase further on an S-200 size exclusion column equilibrated and run in DnaG Gel filtration buffer.
19. Combine the fractions containing *Mtb* DnaG and concentrate the protein to 5 mg/mL. Store *Mtb* DnaG preparations on ice and use in the assay within 2 weeks of concentrating the purified protein. This protocol should yield highly pure *Mtb* DnaG (>95% purity), as assessed by Coomassie blue gel staining.
20. One hour prior to the HTS assay, exchange the buffer by using a G-25 column with DnaG HTS buffer. In this buffer, the concentration of ATP (needed for protein stability) is lower than in the previously used buffers, to avoid the presence of significant quantities of ATP in the reaction.

3.3 Preparation of *Mtb* PPiase

1. Express *Mtb* PPiase (from pJ411-*Mtb* PPiase expression vector) by the above expression protocol for *Mtb* DnaG.
2. Resuspend the cell pellet in 100 mL of PPiase Lysis buffer.
3. Disrupt the cells, clarify the lysate and filter it analogously to the *Mtb* DnaG preparation.
4. Load the filtrate onto a 5 mL IMAC Ni^{2+} column equilibrated in PPiase IMAC buffer A.
5. Pass 220 mL of IMAC buffer A through the column to wash nonspecifically bound protein.
6. Elute the protein by a linear gradient from PPiase IMAC buffers A and B in a total volume of 20 mL, collecting 2 mL fractions.
7. Pool the protein-containing fractions and pass them through an S-200 column equilibrated in PPiase gel filtration buffer.

8. Pool the fractions containing the purified protein and concentrate the protein (as with *Mtb* DnaG) to 750 μ M. The protein should be >98% pure. *Mtb* PPiase can be stored at 4 °C for up to ~1 month without significant loss of activity.

3.4 High-Throughput Screening (HTS) Assay

1. Prepare the 3× NTP and 3× DNA-enzyme mixtures prior to the assay.
2. Dispense 2 μ L of 0.5 M EDTA (the final concentration of ~30 mM) to the last two columns of wells of each 384-well plate by using a multichannel pipette, as a positive inhibition control.
3. Dispense into the rest of the wells 200 nL of each of the small molecules to be tested, each at 2 mM (the final concentration of small molecules in the assay should be 10–30 μ M, depending on how stringent the assay is) by using a Biomek FX robot or a comparable liquid-handling instrument.
4. Add 20 μ L of the 3× NTP mixture to all wells, followed by 10 μ L of the 3× DNA-enzyme mixture to initiate the reaction assay (*see Note 4*), both by using a multichannel pipette.
5. Incubate the reaction at room temperature for 30 min (*see Note 5* and comments below for design of optimum assay conditions).
6. Then add 30 μ L of malachite green reagent, followed by the addition of 10 μ L of 10% sodium citrate for color stability.
7. Shake the plates for 20–30 s on a vibrating plate shaker to ensure efficient mixing.
8. Read absorbance at 620 nm on a microplate reader.
9. Prior to testing the library compounds, perform a positive control assay, as described above with 40 μ M suramin (*see Note 6*), an established DnaG inhibitor, to ensure that the signal arises due to DnaG/PPiase activities and not due to contaminant NTPase.
10. Once the specific activity of the enzymes of interest is confirmed, perform a test HTS assay in a 384-well plate as above, but with DMSO not containing the test molecules, to make sure that the workflow allows repeating this assay for multiple plates without loss of time and to measure assay robustness (*see Note 7*).
11. Calculate the *Z'* from following formula:

$$Z' = 1 - 3 \frac{SD(pos) + SD(neg)}{|Av(pos) - Av(neg)|}$$

where SD(pos) and SD(neg) are the standard deviations of the readout from the positive control (EDTA containing) and DMSO containing wells, respectively, and AVG(pos) and Avg(neg) are the average values of the readout from these respective sets of wells (*see Note 8*).

12. Perform the assay with a double primase reaction time, to ensure that the signal increases ~2-fold (*see Note 9*).

4 Notes

1. Bacterial DnaG is prone to aggregation and proteolytic degradation. Its preparations should be stored at high salt concentration (>300 mM NaCl), to maximize protein stability.
2. In the DnaG purification protocol, 20 mM and 200 mM imidazole should be used in DnaG IMAC Buffers A and B, respectively, if a hexahistidine tag is used.
3. Handle malachite green powder with care, as even a very small amount can stain clothing, equipment, etc.
4. Because of the propensity of DnaG for aggregation at low salt concentrations, DnaG should be added to the reaction buffer (at lower salt concentration) after DNA.
5. If one uses other DnaG and PPiase homologs, the assay should be optimized to find suitable buffer, pH, reaction time, DnaG, PPiase, DNA and NTP concentrations. The HTS assay reaction time point should lie in a steady-state range, with the absorbance signal ~1.0, to ensure linear dynamic response. In order to be able to identify inhibitors of PPiase in the same assay, the concentration of PPiase needs to be chosen three- to fourfold higher (but no greater) than the minimum PPiase concentration required for full hydrolysis of the PP_i formed in the reaction. DNA templates with other sequences should be tested, if the DnaG activity is too low.
6. Because DnaG is not highly efficient or processive, one can perform a control assay by using an inactive point mutant of an active site Glu residue (Glu268Gln for *Mtb* DnaG), purified analogously to the wild-type enzyme or with wild-type DnaG and 40 μM suramin as a positive inhibition control.
7. The assay robustness is usually tested by setting up a preliminary control assay in a 384-well plate, where DMSO is dispensed in without small molecules in all but the positive (EDTA) control wells, as described above. One should perform a Z' calculation based on the readout from this plate to evaluate the assay robustness. Doxorubicin or suramin (2 mM stocks) can be used as other positive controls, although only partial inhibition is to be expected by these compounds.

8. The HTS assay can be carried out if $Z' > 0.5$. Very high Z' values of 0.9–1 may signify that the assay conditions are not steady-state.
9. If the signal stays approximately constant, kinetic experiments need to be performed at smaller concentrations of the enzyme or with higher concentrations of NTPs or template, to ensure that the reaction conditions are steady-state.

References

1. Fan J, de Jonge BL, MacCormack K, Sriram S, McLaughlin RE, Plant H, Preston M, Fleming PR, Albert R, Foulk M, Mills SD (2014) A novel high-throughput cell-based assay aimed at identifying inhibitors of DNA metabolism in bacteria. *Antimicrob Agents Chemother* 58(12):7264–7272.
2. Craig NL, Roberts JW (1981) Function of nucleoside triphosphate and polynucleotide in *Escherichia coli* recA protein-directed cleavage of phage lambda repressor. *J Biol Chem* 256(15):8039–8044
3. Brent R, Ptashne M (1981) Mechanism of action of the *lexA* gene product. *Proc Natl Acad Sci U S A* 78(7):4204–4208
4. Wang JD, Sanders GM, Grossman AD (2007) Nutritional control of elongation of DNA replication by (p)ppGpp. *Cell* 128(5):865–875.
5. Moir DT, Ming D, Opperman T, Schweizer HP, Bowlin TL (2007) A high-throughput, homogeneous, bioluminescent assay for *Pseudomonas aeruginosa* gyrase inhibitors and other DNA-damaging agents. *J Biomol Screen* 12(6):855–864.
6. Brazas MD, Hancock RE (2005) Ciprofloxacin induction of a susceptibility determinant in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 49(8):3222–3227
7. Hirota Y, Ryter A, Jacob F (1968) Thermosensitive mutants of *E. coli* affected in the processes of DNA synthesis and cellular division. *Cold Spring Harb Symp Quant Biol* 33:677–693
8. Kellenberger-Gujer G, Podhajska AJ, Caro L (1978) A cold sensitive dnaA mutant of *E. coli* which overinitiates chromosome replication at low temperature. *Mol Gen Genet* 162(1):9–16
9. Fossum S, De Pascale G, Weigel C, Messer W, Donadio S, Skarstad K (2008) A robust screen for novel antibiotics: specific knockout of the initiator of bacterial DNA replication. *FEMS Microbiol Lett* 281(2):210–214.
10. Kogoma T, von Meyenburg K (1983) The origin of replication, oriC, and the dnaA protein are dispensable in stable DNA replication (sdrA) mutants of *Escherichia coli* K-12. *EMBO J* 2(3):463–468
11. de Massy B, Fayet O, Kogoma T (1984) Multiple origin usage for DNA replication in sdrA(rnh) mutants of *Escherichia coli* K-12. Initiation in the absence of oriC. *J Mol Biol* 178(2):227–236
12. Cromie GA (2009) Phylogenetic ubiquity and shuffling of the bacterial RecBCD and AddAB recombination complexes. *J Bacteriol* 191(16):5076–5084.
13. Amundsen SK, Spicer T, Karabulut AC, Londono LM, Eberhart C, Fernandez Vega V, Bannister TD, Hodder P, Smith GR (2012) Small-molecule inhibitors of bacterial AddAB and RecBCD helicase-nuclease DNA repair enzymes. *ACS Chem Biol* 7(5):879–891.
14. Lanzetta PA, Alvarez LJ, Reinach PS, Candia OA (1979) An improved assay for nanomole amounts of inorganic phosphate. *Anal Biochem* 100(1):95–97
15. Van Veldhoven PP, Mannaerts GP (1987) Inorganic and organic phosphate measurements in the nanomolar range. *Anal Biochem* 161(1):45–48
16. de Groot H, Noll T (1985) Enzymic determination of inorganic phosphates, organic phosphates and phosphate-liberating enzymes by use of nucleoside phosphorylase-xanthine oxidase (dehydrogenase)-coupled reactions. *Biochem J* 230(1):255–260
17. Humnabadkar V, Madhvapeddi P, Basavarajappa H, Sheikh MG, Rane R, Basu R, Verma P, Sundaram A, Mukherjee K, de Sousa SM (2015) Assays, surrogates, and alternative technologies for a TB lead identification program targeting DNA gyrase ATPase. *J Biomol Screen* 20(2):265–274.
18. Biswas T, Resto-Roldan E, Sawyer SK, Artsimovitch I, Tsodikov OV (2013) A novel non-radioactive primase-pyrophosphatase activity assay and its application to the discovery of

- inhibitors of *Mycobacterium tuberculosis* primase DnaG. *Nucleic Acids Res* 41(4), e56.
- 19. Biswas T, Green KD, Garneau-Tsodikova S, Tsodikov OV (2013) Discovery of inhibitors of *Bacillus anthracis* primase DnaG. *Biochemistry* 52(39):6905–6910.
 - 20. McKelvie JC, Richards MI, Harmer JE, Milne TS, Roach PL, Oyston PC (2013) Inhibition of *Yersinia pestis* DNA adenine methyltransferase in vitro by a stibonic acid compound: identification of a potential novel class of antimicrobial agents. *Br J Pharmacol* 168(1):172–188.
 - 21. Tholander F, Sjoberg BM (2012) Discovery of antimicrobial ribonucleotide reductase inhibitors by screening in microwell format. *Proc Natl Acad Sci U S A* 109(25):9798–9803.
 - 22. Jude KM, Hartland A, Berger JM (2013) Real-time detection of DNA topological changes with a fluorescently labeled cruciform. *Nucleic Acids Res* 41(13), e133.
 - 23. Taylor JA, Mitchenall LA, Rejzek M, Field RA, Maxwell A (2013) Application of a novel microtitre plate-based assay for the discovery of new inhibitors of DNA gyrase and DNA topoisomerase VI. *PLoS One* 8(2), e58010.
 - 24. Buhler C, Gadelle D, Forterre P, Wang JC, Bergerat A (1998) Reconstitution of DNA topoisomerase VI of the thermophilic archaeon *Sulfolobus shibatae* from subunits separately overexpressed in *Escherichia coli*. *Nucleic Acids Res* 26(22):5157–5162.
 - 25. Finlay GJ, Atwell GJ, Baguley BC (1999) Inhibition of the action of the topoisomerase II poison amsacrine by simple aniline derivatives: evidence for drug-protein interactions. *Oncol Res* 11(6):249–254.
 - 26. Bojanowski K, Lelievre S, Markovits J, Couprie J, Jacquemin-Sablon A, Larsen AK (1992) Suramin is an inhibitor of DNA topoisomerase II in vitro and in Chinese hamster fibrosarcoma cells. *Proc Natl Acad Sci U S A* 89(7):3025–3029.
 - 27. Chauhan PM, Srivastava SK (2001) Present trends and future strategy in chemotherapy of malaria. *Curr Med Chem* 8(13):1535–1542.
 - 28. Cheng B, Cao S, Vasquez V, Annamalai T, Tamayo-Castillo G, Clardy J, Tse-Dinh YC (2013) Identification of anziaic acid, a lichen depside from *Hypotrachyna* sp., as a new topoisomerase poison inhibitor. *PLoS One* 8(4), e60770.
 - 29. Mills SD, Eakin AE, Buurman ET, Newman JV, Gao N, Huynh H, Johnson KD, Lahiri S, Shapiro AB, Walkup GK, Yang W, Stokes SS (2011) Novel bacterial NAD+-dependent DNA ligase inhibitors with broad-spectrum activity and antibacterial efficacy in vivo. *Antimicrob Agents Chemother* 55(3):1088–1096.
 - 30. Shapiro AB, Eakin AE, Walkup GK, Rivin O (2011) A high-throughput fluorescence resonance energy transfer-based assay for DNA ligase. *J Biomol Screen* 16(5):486–493.
 - 31. Glaser BT, Malerich JP, Duellman SJ, Fong J, Hutson C, Fine RM, Kebiansky B, Tang MJ, Madrid PB (2011) A high-throughput fluorescence polarization assay for inhibitors of gyrase B. *J Biomol Screen* 16(2):230–238.
 - 32. Peterson EJ, Janzen WP, Kireev D, Singleton SF (2012) High-throughput screening for RecA inhibitors using a transcreener adenosine 5'-O-diphosphate assay. *Assay Drug Dev Technol* 10(3):260–268.
 - 33. Lee AM, Wigle TJ, Singleton SF (2007) A complementary pair of rapid molecular screening assays for RecA activities. *Anal Biochem* 367(2):247–258.
 - 34. Zhang JH, Chung TDY, Oldenburg KR (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen* 4(2):67–73.
 - 35. Gajadeera C, Willby MJ, Green KD, Shaul P, Fridman M, Garneau-Tsodikova S, Posey JE, Tsodikov OV (2015) Antimycobacterial activity of DNA intercalator inhibitors of *Mycobacterium tuberculosis* primase DnaG. *J Antibiot (Tokyo)* 68(3):153–157.
 - 36. Biswas T, Tsodikov OV (2008) Hexameric ring structure of the N-terminal domain of *Mycobacterium tuberculosis* DnaB helicase. *FEBS J* 275(12):3064–3071.

Chapter 12

Fluorescence-Based Real-Time Activity Assays to Identify RNase P Inhibitors

Yu Chen, Xin Liu, Nancy Wu, and Carol A. Fierke

Abstract

Transfer RNA is transcribed as precursor molecules that are processed before participating in translation catalyzed by the ribosome. Ribonuclease P is the endonuclease that catalyzes the 5' end maturation of precursor tRNA and it is essential for cell survival. Bacterial RNase P has a distinct subunit composition compared to the eukaryal counterparts; therefore, it is an attractive antibacterial target. Here, we describe a real-time fluorescence-based RNase P activity assay using fluorescence polarization/anisotropy with a 5' end fluorescein-labeled pre-tRNA^{Asp} substrate. This FP/FA assay is sensitive, robust, and easy to transition to a high-throughput mode and it also detects ligands that interact with pre-tRNA. We apply this FP/FA assay to measure *Bacillus subtilis* RNase P activity under single and multiple turnover conditions in a continuous format and a high-throughput screen of inhibitors, as well as determining the dissociation constant of pre-tRNA for small molecules.

Key words RNase P, Fluorescence polarization, Fluorescence anisotropy, tRNA, High-throughput screening, Inhibitor, Antibiotics, Neomycin, Mode of inhibition

1 Introduction

Transfer RNA (tRNA) genes are transcribed as precursors with extra sequence at both the 5'- and 3'-ends, which are processed by various enzymes to generate mature and functional tRNA molecules [1–4]. Ribonuclease P (RNase P) catalyzes maturation of the 5'-end of tRNAs. In most cases, RNase P is composed of RNA and protein subunits that catalyze hydrolysis of a specific phosphodiester bond in precursor tRNAs (pre-tRNAs) to generate mature tRNAs with a phosphate at the 5'-end [5–8]. With only one exception, archaeon *Nanoarchaeum equitans* [9], this enzyme is found in all domains of life and is crucial for maintaining cell viability [10–13].

In most organisms, RNase P is a ribonucleoprotein consisting of a single catalytic RNA subunit (PRNA) and a variable number of protein subunits (P proteins) [6, 10, 14–18]. Recently, protein-only

RNase P enzymes have been identified in human mitochondria, plants, and some algae and protists [19–21]. Because of its essential role in RNA processing and the differential subunit composition from its eukaryal counterparts, bacterial RNase P could be an antimicrobial target or used for gene therapy [22–26]. However, existing RNase P activity assays require radio-labeled materials [27] or lack the sensitivity necessary for measuring steady-state activity [28–30] in real time [31].

Here, we describe a real-time fluorescence polarization/anisotropy (FP/FA) assay to analyze RNase P activity using a 5' fluorescein-labeled pre-tRNA^{Asp} substrate (Fl-pre-tRNA^{Asp}) [28, 32] (shown in Fig. 1). The degree of polarization of a fluorophore reflects the mobility of the fluorophore, which is proportional to its molecular mass [33]. When Fl-pre-tRNA^{Asp} is excited by polarized light, the emission light remains largely polarized (higher anisotropy) because the Fl-pre-tRNA^{Asp} macromolecule rotates slower than the fluorescence lifetime of fluorescein. After RNase P-catalyzed cleavage of the Fl-pre-tRNA^{Asp}, the emitted light from the fluorescein-5'-leader product is depolarized (lower anisotropy). Because the 5' leader (2- to 10-nucleotide long leaders tested in our lab) rotates more rapidly than the 82-nt Fl-pre-tRNA substrate, the anisotropy of the Fl-pre-tRNA and the 5' leader product differ by two- to threefold, providing a significant dynamic range for the activity assay. Furthermore, a time-dependent decrease in the FP/FA signal is observed upon mixing *B. subtilis* RNase P holoenzyme with the Fl-pre-tRNA^{Asp} substrate under both single turnover (STO, $[E]>[S]$) and multiple turnover

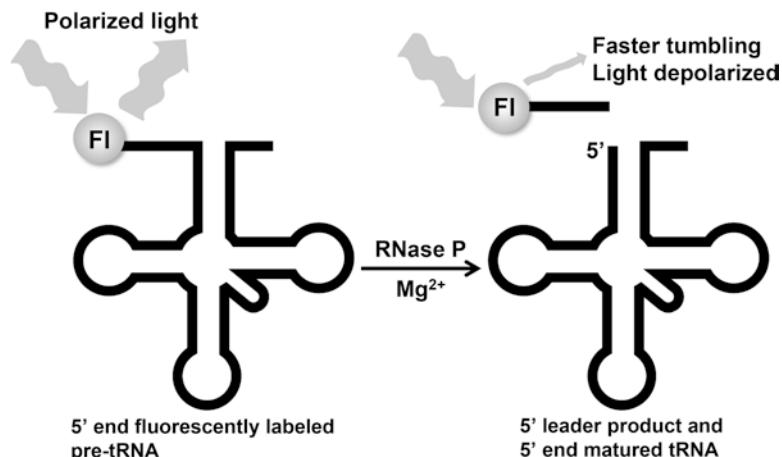


Fig. 1 Scheme of the FP/FA activity assay for RNase P-catalyzed pre-tRNA cleavage. A fluorescein dye (gray dot) is attached to the 5' end of the leader of a pre-tRNA. The fluorescein attached to Fl-pre-tRNA has high anisotropy. After RNase P catalysis, the Fl-5 nt-leader product is cleaved and produces lower anisotropy. (Reproduced from reference [32] with permission from Oxford University Press)

(MTO, $[E] < [S]$) conditions. The FP/FA signal measures the ratio of parallel and perpendicular light. This assay is ratiometric and highly accurate (error $< 5\%$ even with a twofold dynamic range) [33]. It measures RNase P activity in a continuous format, and is suitable for high-throughput screening (HTS) of RNase P activators and inhibitors, as well as detecting ligands that interact with Fl-pre-tRNA.

2 Materials

Prepare all solutions with ultrapure water (Milli-Q® treated deionized water) and chemicals in the highest purity (DNase- and RNase-free grade). Sterilize solutions using an autoclave (buffers such as Tris-HCl, MgCl₂, and KCl) or Stericup® filter units (pore size 0.22 µm, for solutions containing sodium dodecyl sulfate and MES). Measure the pH of stock buffer solutions at 25 °C.

2.1 Components for Fl-pre-tRNA^{Asp} and PRNA Preparation

1. 1 M magnesium chloride (MgCl₂) stored at room temperature.
2. 0.5 M ethylenediaminetetraacetic acid (EDTA) (pH 8.0) stored at room temperature.
3. 5 M sodium chloride (NaCl) stored at room temperature.
4. Small aliquots (200 µl) of 1 M dithiothreitol (DTT) stored at -20 °C.
5. Small aliquots of (200 µl) 1 M spermidine stored at -20 °C.
6. 10× Transcription buffer: 500 mM Tris-HCl (pH 8.0), 10 mM spermidine, and 50 mM DTT. Store at -20 °C.
7. 100 mM adenosine triphosphate (ATP), cytidine triphosphate (CTP), uridine triphosphate (UTP), and guanosine triphosphate (GTP) solutions (pH 7.0). Dissolve NTPs powder in autoclaved ultrapure water and adjusted to neutral pH using 10 M sodium hydroxide (NaOH) followed by sterile filtration. Solutions of NTPs can also be purchased. Store at -20 °C for short term or -80 °C for long term.
8. 30–100 mM guanosine 5'-monothiophosphate (GMPS) solution. Synthesize GMPS from 2', 3'-isopropylidene-guanosine and thiophosphoryl chloride as described previously [34]. Prepare GMPS solution by dissolving solid GMPS in autoclaved ultrapure water followed by sterile filtration. GMPS solution can also be purchased (Axxora). Store at -20 °C for short term or -80 °C for long term.
9. Pre-tRNA^{Asp} DNA template. Prepare the plasmid containing T7 RNA promoter and *B. subtilis* pre-tRNA^{Asp} sequence as previously described, and linearize the plasmid using *Bst*NI restriction enzyme [35]. Store at -20 °C.

10. RNase Inhibitor (20 unit/L) stored at -20 °C.
11. T7 RNA polymerase. Recombinant His₆-T7 RNA polymerase is encoded on a plasmid expressed in *E. coli*. Purify T7 RNA polymerase by Ni-NTA chromatography as described previously [36]. Store at -20 °C for short term or -80 °C for long term.
12. Pyrophosphatase (200 unit /mg) stored at -20 °C.
13. Degassed pre-tRNA labeling buffer: 10 mM Tris-HCl (pH 7.2) and 1 mM EDTA (pH 8.0). Degas a small amount of the labeling buffer using a Steriflip® filter by vacuum. Tap the tube to facilitate the release of air bubbles until no visible air bubbles can be seen. Degas the labeling buffer right before use and chill at 4 °C.
14. Denaturing polyacrylamide gel (PAGE): 6% or 10% acrylamide/bis-acrylamide (19:1), 7 M urea and 1×Tris/Borate/EDTA (TBE) buffer. Store at 4 °C and protect from light (*see Note 1*).
15. Elution buffer: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 0.1% sodium dodecyl sulfate (SDS), and 500 mM NaCl. Store at room temperature.
16. Wash buffer: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), and 500 mM NaCl. Store at 4 °C.
17. TE buffer: 10 mM Tris-HCl, (pH 8.0) and 1 mM EDTA (pH 8.0). Store at room temperature.
18. 5-Iodoacetamido-fluorescein (5-IAF): Dissolve 1 mg of 5-IAF dye into 40 µl of anhydrous dimethyl sulfoxide (DMSO) to prepare 48.5 mM solution. Store at -80 °C and protect from light (*see Note 2*).
19. 2×EDTA dye: 8 M urea, 200 mM EDTA (pH 8.0), 0.005% (w/v) bromophenol blue, and 0.005% (w/v) xylene cyanol.
20. PRNA and P protein subunits of *B. subtilis* RNase P (can be prepared as previously described [35, 37, 38]).
21. Apparatus: Amicon® Ultra-4 Centrifugal Filter (10,000 NMWL), Amicon® Ultra-15 Centrifugal Filter (10,000 NMWL), Steriflip® filter unit (pore size 0.22 µm), swing bucket centrifuge, UV-Vis spectrophotometer, UV light (medium wave UV, 312 nm), UV fluorescent TLC plate, and UV shadow (short wave UV, 254 nm).

2.2 Components for Fluorescence-Based Assays

Prepare buffers one to five fresh before use in the assays at room temperature:

1. 4× Buffer A: 200 mM Tris/MES (pH 5.5–8.8), 80 mM MgCl₂, 800 mM sodium chloride (KCl), and 80 mM DTT; Buffer A: 50 mM Tris/MES (pH 5.5–8.8), 20 mM MgCl₂, 200 mM KCl, and 20 mM DTT.

2. 2× Buffer B: 100 mM Tris–HCl (pH 7.2), 20 mM MgCl₂, 200 mM KCl, and 40 mM DTT; Buffer B: 50 mM Tris–HCl (pH 7.2), 10 mM MgCl₂, 100 mM KCl, and 20 mM DTT.
3. 5× HTS buffer: 250 mM Tris–HCl (pH 7.2), 25 mM MgCl₂, 500 mM KCl, 100 mM DTT, 60 µg/ml yeast tRNA^{Mix}, 50 mM spermidine, and 0.05% (v/v) nonidet P-40 (NP-40); 1× HTS buffer: 50 mM Tris–HCl (pH 7.2), 5 mM MgCl₂, 100 mM KCl, 20 mM DTT, 12 µg/ml yeast tRNA^{Mix}, 10 mM spermidine, and 0.01% (v/v) nonidet P-40 (NP-40) (*see Note 3*).
4. 1× HTS buffer with 400 mM CaCl₂.
5. 1× HTS buffer with 240 mM CaCl₂.
6. 10 mM neomycin B stored at –20 °C.
7. DMSO stored dry at room temperature.
8. Apparatus: fluorescence microplate reader, 96-well half area black nonbinding surface microplate, 384-well black nonbinding surface microplate, Multidrop™ Combi reagent dispenser, Multidrop™ dispensing cassette, nanoliter pintoool, phosphorimager, GraphPad Prism software.
9. Compound library for high-throughput screening (provided by Center for Chemical Genomics (CCG) at University of Michigan).

3 Methods

Carry out all procedures at room temperature unless otherwise specified. Use low adhesion, DNase, and RNase free tubes and sterilized pipette tips.

3.1 Preparation of *B. subtilis* Fl-pre-tRNA^{Asp} [28, 30]

1. Prepare two 1 ml transcription reactions of *B. subtilis* pre-tRNA^{Asp} (*see Note 4*): Add the components according to the order specified in Table 1 into a 1.5 ml micro test tube and incubate the mixture at 37 °C overnight (up to 16 h) (*see Note 5*).
2. Stop the transcription reaction with a final concentration of 50 mM EDTA and 500 mM NaCl. For 1 ml transcription reactions, add 125 µl of 0.5 M EDTA and 125 µl of 5 M NaCl.
3. Prewash an Amicon® Ultra-4 Centrifugal Filter (10,000 NMWL) with degassed pre-tRNA labeling buffer by centrifugal filtration at 2400–2800 × g for 15 min at 4 °C (*see Note 6*). Discard the degassed labeling buffer left in the filter unit and receiving tube.
4. Add the transcription reaction into prewashed centrifugal filter and concentrate by centrifugal filtration at 2400–2800 × g for 15–20 min at 4 °C until volume is reduced to about 200 µl (*see Note 7*).

Table 1
Components and concentration for transcription of *B. subtilis* pre-tRNA^{Asp}

Component	Final concentration
Sterilized ultrapure water	
10× Transcription buffer	1× ^a
MgCl ₂ (mM)	20
ATP (mM)	4
CTP (mM)	4
UTP (mM)	4
GMPS (mM)	4–5
GTP (mM)	0.8–1
DNA template (μg/μl)	0.08–0.1
RNase inhibitor (unit/μl)	0.04
T7 RNA polymerase (μg/μl)	0.1
Pyrophosphatase (μg/ml)	2

^aAdd 100 μl of 10× Transcription buffer for 1 ml transcription reaction. Add water as needed to prepare 1 ml

5. Add 3 ml of chilled degassed labeling buffer into the centrifugal filter to resuspend the retained transcription mix and concentrate by centrifugal filtration. Repeat this buffer exchange step at least three times (*see Note 8*). The final volume is about 200 μl.
6. Transfer the sample from the centrifugal filter into a 1.5 ml amber micro test tube (*see Note 9*).
7. Dilute 1 μl of the sample into 9 μl of labeling buffer and measure A₂₆₀ of the RNA by UV-Vis spectrophotometer using the labeling buffer as a blank.
8. Calculate the concentration using Eq. 1:

$$[\text{RNA}] = \frac{A_{260}}{\varepsilon_{\text{RNA}} \times L} \times n \quad (1)$$

where ε_{260} is 685,000 (mol⁻¹ cm⁻¹) for pre-tRNA^{Asp}, L is path length (cm), and n is dilution factor that equals 10 in this experiment.

9. Calculate the yield of pre-tRNA and add 20- to 40-fold excess moles of 5-IAF.

10. Add 1–2 μ l of 20 unit/L of RNase Inhibitor and incubate the mixture at 37 °C overnight (up to 16 h) (*see Note 10*).
11. Stop the labeling reaction with a final concentration of 10 mM DTT (~1/100 of total reaction volume of 1 M DTT) and equal volume of 2 \times EDTA dye.
12. Purify the Fl-pre-tRNA^{Asp} by gel electrophoresis on 10% denaturing PAGE in the dark (*see Note 11*). Use UV light and UV shadowing to identify Fl-pre-tRNA^{Asp} bands and excise the bands (*see Note 12*). Crush the gel and transfer into a 50 ml sterilized centrifuge tube containing 45 ml of elution buffer. Gently rock the tube overnight or longer at 4 °C (less than 36 h).
13. Prewash Amicon® Ultra-15 Centrifugal Filter (10,000 NMWL) with washing buffer by centrifugal filtration at 2400–2800 $\times g$ for 15 min at 4 °C (*see Note 13*). Discard the washing buffer left in the filter unit and receiving tube.
14. Filter the eluted Fl-pre-tRNA^{Asp} using a Steriflip® filter unit (pore size 0.22 μ m). Transfer the solution into the prewashed centrifugal filter and concentrate to 0.2–1 ml by centrifugal filtration at 2400–2800 $\times g$ at 4 °C (*see Note 14*). Keep the RNA in the dark during the process. The sample should become increasing green-colored as the sample is concentrated.
15. Add 14 ml of washing buffer into centrifugal filter to dilute the retained Fl-pre-tRNA^{Asp} and concentrate by centrifugal filtration to ~500 μ l. Repeat this step at least three times to remove SDS in the RNA sample. Keep the RNA in the dark during the process.
16. After buffer exchange, transfer the Fl-pre-tRNA^{Asp} into 1.5 ml amber micro test tubes (*see Note 9*).
17. Estimate the volume and add two volumes of 100% cold ethanol (chilled at –20 °C). Gently invert the tube to mix. Store the RNA at –80 °C away from light for at least 2 h.
18. Before using the Fl-pre-tRNA^{Asp} in assays, precipitate the RNA by centrifuging at high speed (~12,000 $\times g$) for 30 min at 4 °C. Remove the supernatant carefully and wash the RNA pellet with 500 μ l of 70% ethanol. Gently invert the tube to wash the pellet. Spin the RNA at ~12,000 $\times g$ for 2 min at 4 °C and remove the supernatant. Do not disturb the RNA pellet by pipetting or vortex. Repeat this step twice. Leave the tube with lids open in the dark to dry the tubes.
19. Dissolve the Fl-pre-tRNA^{Asp} pellet in water or TE buffer. Make a tenfold dilution of the sample in water and measure A₂₆₀ ([RNA]) and A₄₉₄ ([Fluorescein]).
20. Calculate the RNA concentration using Eq. 1 and the labeling efficiency using Eqs. 2 and 3:

$$\text{Labeling efficiency} = \frac{[\text{Fluorescein}]}{[\text{RNA}]} \quad (2)$$

$$[\text{Fluorescein}] = \frac{A_{494}}{\epsilon_{\text{Fluorescein}} \times L} \times n \quad (3)$$

where $\epsilon_{\text{Fluorescein}}$ is 78,000 ($\text{mol}^{-1} \text{cm}^{-1}$), L is path length (cm), and n is the dilution factor that equals 10 in this experiment.

3.2 Preparation of *B. subtilis* PRNA and P Protein [35, 37, 38]

- After transcription of PRNA [35], follow the description from Subheading 3.1, steps 12–20 to purify the PRNA using 6% denaturing PAGE (see Note 15), buffer exchange, and ethanol precipitation. Calculate the concentration of the PRNA using Eq. 1. ϵ_{PRNA} is 3,940,000 ($\text{mol}^{-1} \text{cm}^{-1}$).
- Measure A_{280} of the P protein by UV-Vis spectrophotometer and calculate the concentration using Eq. 4.

$$[\text{P Protein}] = \frac{A_{280}}{\epsilon_{\text{P protein}} \times L} \times n \quad (4)$$

where $\epsilon_{\text{P protein}}$ is 5500 ($\text{mol}^{-1} \text{cm}^{-1}$), L is path length (cm), and n is dilution factor.

3.3 Single-Turnover (STO) Assay

- Prepare *B. subtilis* RNase P holoenzyme and Fl-pre-tRNA^{Asp} in two 1.5 ml micro test tubes (use amber tube for Fl-pre-tRNA^{Asp}, see Note 9) at the concentration listed in Table 2. The concentrations are at twice the final concentration.
- Add sterilized ultrapure water and PRNA or Fl-pre-tRNA^{Asp} into micro test tubes.
- Denature PRNA and Fl-pre-tRNA^{Asp} by heating for 3 min at 95 °C. Cool at 37 °C for 10–15 min.
- Refold PRNA and Fl-pre-tRNA^{Asp} by adding 4× Buffer A at 37 °C. Mix well and incubate the samples at 37 °C for 30 min.

Table 2
STO assay for cleavage of Fl-pre-tRNA^{Asp} catalyzed by *B. subtilis* RNase P

Component	Concentration	Component	Concentration
Sterilized ultrapure water		Sterilized ultrapure water	
PRNA (μM)	1	Fl-pre-tRNA ^{Asp} (nM)	50
4× Buffer A	1× ^a	4× Buffer A	1× ^b
P protein (μM)	1		

^aAdd 7.5 μl of 4× Buffer A for 30 μl sample. Add water as needed to prepare 30 μl

^bAdd 12.5 μl of 4× Buffer A for 50 μl sample. Add water as needed to prepare 50 μl

5. Add P protein to the PRNA sample and incubate at 37 °C for another 30 min to reconstitute RNase P holoenzyme.
6. To optimize the gain and beam position of the microplate reader, take a 96-well half area black nonbinding surface microplate, transfer 20 µl of Fl-pre-tRNA^{Asp} sample into one well, and add 20 µl of Buffer A to obtain Fl-pre-tRNA^{Asp} with a final concentration of 10 nM. Measure the FA signal of fluorescein using a microplate reader at 37 °C (see Note 16). Use an excitation filter with $\lambda_{ex} = 485 \pm 20$ nm and an emission filter with $\lambda_{em} = 535 \pm 25$ nm.
7. Transfer 40 µl of 1×Buffer A to another well and assign this well as a blank measurement (optional) (see Note 17).
8. Transfer 20 µl of Fl-pre-tRNA^{Asp} into a new well and initiate the cleavage reaction by adding 20 µl of RNase P holoenzyme (the final concentration of RNase P holoenzyme is 500 nM and Fl-pre-tRNA^{Asp} is 25 nM) (see Note 18).
9. Immediately start the measurement at 37 °C and monitor the cleavage by a decrease in the FA signal (see Note 19). Record the FA signal of the Fl-pre-tRNA^{Asp} only sample (no RNase P holoenzyme added) as well.
10. Stop the measurement when cleavage is complete (see Note 20). A typical real-time FA trace measured under STO conditions in Buffer A is shown in Fig. 2a.
11. Data analysis: Subtract the FA trace of Fl-pre-tRNA^{Asp} only from the reaction traces to adjust for temperature variations on the FA signal (see Note 21). Convert the FA values into fraction cleaved (Υ) values using Eq. 5:

$$\Upsilon = 1 - \frac{r - r_\infty}{r_0 - r_\infty + (\gamma - 1)(r_0 - r)} \quad (5)$$

The enhancement factor $\gamma = F^0/F^\infty$, defined as described previously [39], is used to correct for the total fluorescence change from the start to the end of the reaction (see Note 22). F^0 and F^∞ are the total fluorescence of RNase P holoenzyme-bound Fl-pre-tRNA^{Asp} and cleaved product, respectively; r_0 and r_∞ are the FA values of the RNase P holoenzyme-bound Fl-pre-tRNA^{Asp} and cleaved product, respectively [33, 39] under STO conditions.

The STO observed cleavage rate constant (k_{obs}) is calculated by fitting a single exponential model (Eq. 6) to the fraction cleaved as a function of time.

$$\Upsilon = \Upsilon_\infty \left(1 - e^{-k_{obs}t} \right) \quad (6)$$

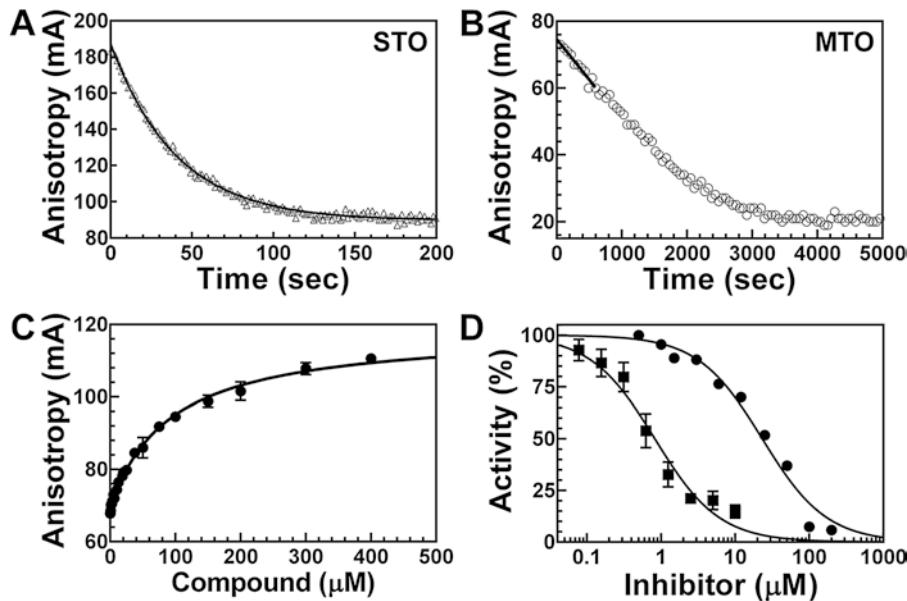


Fig. 2 Application of FA assay for RNase P. **(a)** A trace of time-dependent change in FA in Buffer A under STO condition (25 nM Fl-pre-tRNA^{Asp} and 500 nM *B. subtilis* RNase P holoenzyme) at 37 °C. The solid line is a single exponential fit to the data to obtain k_{obs} . **(b)** A trace of the time-dependent change in FA under MTO conditions (20 nM Fl-pre-tRNA^{Asp}, 0.4 nM *B. subtilis* PRNA, and 4 nM P protein) at 37 °C. The solid line indicates the linear initial rate. **(c)** The FA dependence of 50 nM of Fl-pre-tRNA^{Asp} on the concentration of neomycin B at 37 °C. The solid line is a fit of a binding isotherm to the data to obtain K_D . **(d)** Dose-response of *B. subtilis* RNase P inhibition by neomycin B (filled circle) and iriginol hexaacetate (Ir6Ac) (filled square). Measured either for neomycin B in Buffer B with 50 nM Fl-pre-tRNA^{Asp}, 0.4 nM PRNA, and 4 nM P protein at 37 °C or for Ir6Ac in HTS buffer with 1% DMSO, 50 nM Fl-pre-tRNA^{Asp}, 0.1 nM PRNA, and 2 nM P protein at 37 °C. (Reproduced from [32] with permission from Oxford University Press)

3.4 Multiple-Turnover (MTO) Assay

1. Use the concentrations listed in Table 3 and procedures from Subheading 3.3, steps 1–5 to denature, refold, and reconstitute *B. subtilis* RNase P holoenzyme and Fl-pre-tRNA^{Asp}.
2. Dilute RNase P holoenzyme to 0.8 nM (or desired concentration) by transferring 150 μl of RNase P holoenzyme sample into 1350 μl of Buffer B. The concentration of RNase P holoenzyme sample is at twice the final concentration.
3. Prepare different concentrations of Fl-pre-tRNA^{Asp} (~15 nM to 2 μM) by serial dilution. To do this, add 50 μl of 2 μM (or highest concentration in the series) Fl-pre-tRNA^{Asp} sample into 50 μl of Buffer B and mix well. This step dilutes Fl-pre-tRNA^{Asp} by twofold. Transfer 50 μl of the sample from this tube into another tube containing 50 μl of Buffer B and mix well. Repeat this dilution until the lowest Fl-pre-tRNA^{Asp} concentration in the series has been prepared (see Note 23). The concentrations of Fl-pre-tRNA^{Asp} samples are at twice the final concentration.

Table 3
MTO assay for cleavage of Fl-pre-tRNA^{Asp} catalyzed by *B. subtilis* RNase P

Component	Concentration	Component	Concentration
Sterilized ultrapure water		Sterilized ultrapure water	
PRNA (nM)	8	Fl-pre-tRNA ^{Asp} (μ M)	2
2 \times Buffer B	1 \times ^a	2 \times Buffer B	1 \times ^b
P protein (nM)	80		

^aAdd 80 μ l of 2 \times Buffer B for 160 μ l sample. Add water as needed to prepare 160 μ l

^bAdd 75 μ l of 2 \times Buffer B for 150 μ l sample. Add water as needed to prepare 150 μ l

- Take a 96-well half area black nonbinding surface microplate, transfer 20 μ l of the serial concentrations of Fl-pre-tRNA^{Asp} sample into a row of wells, and add 20 μ l of Buffer B into each well to obtain Fl-pre-tRNA^{Asp} with final concentrations from ~8 nM to 1 μ M.
- Following the procedure described in Subheading 3.3, steps 6 and 7, optimize the gain and beam position of the microplate reader at 37 °C using the well with the highest concentration of substrate (see Note 24).
- Add 20 μ l of the serial concentrations of Fl-pre-tRNA^{Asp} sample into a second row of the plate and RNase P holoenzyme sample into a third row of wells to pre-equilibrate (see Note 18).
- Initiate the cleavage reaction by transferring 20 μ l of RNase P holoenzyme sample (the final concentration of RNase P holoenzyme is 0.4 nM) into each well containing Fl-pre-tRNA^{Asp} sample (final concentrations of Fl-pre-tRNA^{Asp} are from ~8 nM to 1 μ M) using a multichannel pipette. Immediately start the measurement and monitor cleavage by a decrease in the FA signal (see Note 25). Record the signal of the Fl-pre-tRNA^{Asp} only sample (no RNase P holoenzyme added) as well.
- Stop the measurement when cleavage of at least one substrate concentration is complete (see Note 26). A typical time-dependent change in FA measured under MTO conditions in Buffer B is shown in Fig. 2b.
- Data analysis: Subtract the FA traces of the Fl-pre-tRNA^{Asp} only control from the reaction traces to adjust for fluctuations in the signal due to temperature variations (see Note 21). Convert the linear initial rates in millianisotropy per second (mA/s) from the time courses to reaction velocity (nM/s) using Eq. 7 [40]:

$$v_0 = \frac{\frac{''}{''} r_i}{\frac{''}{''} t_c [S]} \quad (7)$$

where $\Delta r_i/\Delta t$ is the initial rate, Δr_c is the total FA signal change upon complete conversion of Fl-pre-tRNA^{Asp} substrate to 5'-leader product, and $[S]$ is the initial concentration of Fl-pre-tRNA^{Asp}.

Steady-state kinetic parameters, k_{cat} , K_M , and k_{cat}/K_M can be calculated from a Michaelis–Menten fit to the dependence of the initial reaction velocity on the concentration of Fl-pre-tRNA^{Asp} (Eq. 8) [41].

$$v_0 = \frac{k_{\text{cat}} [E]}{K_M + [S]} [S] \quad (8)$$

3.5 Determination of the Dissociation Constant (K_D) of Compound Binding to Fl-pre-tRNA^{Asp}

1. Use the concentrations listed in Table 4 and procedures from Subheading 3.3, steps 1–5 to denature and refold *B. subtilis* Fl-pre-tRNA^{Asp} at twice the final concentration.
2. Prepare different concentrations of neomycin B sample (~8 μM to 1 mM) by serial dilution as described in Subheading 3.4, step 3. The concentrations of neomycin B are at twice the final concentration.
3. Take a 96-well half area black nonbinding surface microplate and add 20 μl of Fl-pre-tRNA^{Asp} sample to a row of wells.
4. Add 20 μl of Buffer B into one of the wells to obtain Fl-pre-tRNA^{Asp} with a final concentration of 50 nM. Follow the procedure described in Subheading 3.3, steps 6 and 7 to optimize the gain and beam position of the microplate reader at 37 °C.
5. Transfer 20 μl of neomycin B sample in serial concentrations to the rest of the wells and mix well (the final concentration of Fl-pre-tRNA^{Asp} is 50 nM and neomycin B are from ~4–500 μM). Incubate the plate at 37 °C for 15 min in the microplate reader.
6. Record the FA signal of all thirteen wells. An example of neomycin B-dependent binding curve for Fl-pre-tRNA^{Asp} is shown in Fig. 2c.
7. Data analysis: The total fluorescence intensity does not change significantly for neomycin B binding experiments. Thus, the

Table 4
Determine dissociation constant of Fl-pre-tRNA^{Asp} for neomycin B

Component	Concentration	Component	Concentration
Sterilized ultrapure water		Sterilized ultrapure water	
Neomycin B (mM)	1	Fl-pre-tRNA ^{Asp} (nM)	100
2× Buffer B	1× ^a	2× Buffer B	1× ^b

^aAdd 50 μl of 2× Buffer B for 100 μl sample. Add water as needed to prepare 100 μl

^bAdd 150 μl of 2× Buffer B for 300 μl sample. Add water as needed to prepare 300 μl

dissociation constant K_D can be obtained directly by fitting the data to a single binding isotherm:

$$\Upsilon = r_f + \frac{B_{\max} \times [I]}{K_D \times [I]} \quad (9)$$

where r_f is the anisotropy of free Fl-pre-tRNA^{Asp}, B_{\max} is the difference in FA between free and fully bound Fl-pre-tRNA^{Asp}-neomycin B complex, and $[I]$ is the concentration of neomycin B.

3.6 High-Throughput Screening (HTS) for Inhibitors of *B. subtilis* RNase P: Time Course to Determine

Linear Range of the Reaction

1. Follow the MTO procedures to prepare *B. subtilis* RNase P holoenzyme (0.3 nM PRNA and 3 nM P protein) and Fl-pre-tRNA^{Asp} (40 nM) in HTS buffer. The concentrations of RNase P holoenzyme and Fl-pre-tRNA^{Asp} samples are at twice the final concentration.
2. Take a 384-well black nonbinding surface microplate and add 10 μ l of RNase P holoenzyme into one row of wells using a Multidrop™ Combi reagent dispenser with Multidrop™ dispensing cassette.
3. Then add 0.2 μ l of DMSO into each well using a nanoliter pintoool and incubate the plate at 30 °C for 30 min.
4. Add 5 μ l of HTS buffer with 400 mM CaCl₂ into the first well as zero time-point.
5. Initiate the reaction by adding 10 μ l of Fl-pre-tRNA^{Asp} sample into each well and start the timer immediately (the final concentration of RNase P holoenzyme is 0.15 nM PRNA with 1.5 nM P protein and Fl-pre-tRNA^{Asp} is 20 nM).
6. Quench the reaction by adding 5 μ l of HTS buffer with 400 mM CaCl₂ (final concentration of CaCl₂ is 80 mM) at various time points (see Note 27). Conduct all reactions at 30 °C.
7. Optimize the gain and beam position of the microplate reader using the zero time-point well (see Note 16). Record the FP/FA signal of all wells using a 384-well microplate reader. For fluorescein, use an excitation filter with $\lambda_{ex} = 485$ nm and emission filter with $\lambda_{em} = 520$ nm. Determine the linear range of the reaction and choose a time point to quench the reactions in the HTS (see Note 28).

3.7 HTS for Inhibitors of *B. subtilis* RNase P Activity: Primary Screen

1. Add 5 μ l of RNase P holoenzyme sample (0.3 nM PRNA and 3 nM P protein) using a Multidrop™ Combi reagent dispenser with Multidrop™ dispensing cassette into all the wells.
2. Pinpoint 0.1 μ l of DMSO using a nanoliter pintoool into columns 1, 2, 23, and 24.
3. Pinpoint 0.1 μ l of compound (dissolved in DMSO) from HTS library into the rest of the wells (columns 3–22) in the same way.

4. Incubate the plate at 30 °C for 30 min for the compounds to incubate with RNase P.
5. Add 5 µl of HTS buffer with 240 mM CaCl₂ into columns 23 and 24. These are the positive controls (complete inhibition) in the assay.
6. Initiate the reaction by adding 5 µl of 40 nM Fl-pre-tRNA^{Asp} (the final concentration of RNase P holoenzyme is 0.15 nM PRNA with 1.5 nM P protein and Fl-pre-tRNA^{Asp} is 20 nM) into each well at 30 °C.
7. After 35 min (determined from Subheading 3.6, step 7), quench the reaction by adding 5 µl of HTS buffer with 240 mM CaCl₂ into each well except the positive control wells in columns 23 and 24 (see Subheading 3.7, step 5). Columns 1 and 2 containing RNase P holoenzyme and DMSO only (without compound) quenched by CaCl₂ are served as negative controls (no inhibition).
8. Calculate the *G* factor, gain, and beam position for each plate using a well in column 23 or 24. Set the *G* value to obtain a FP reading of 180 mP and the gain was set to obtain 80% of intensity of the parallel channel. Read the FA values for each well.
9. Evaluate the robustness of the HTS assay by calculating the *Z*-factor [42] using Eq. 10:

$$Z' = 1 - \frac{3\sigma_{c+} + 3\sigma_{c-}}{|\mu_{c+} - \mu_{c-}|} \quad (10)$$

where σ_{c+} and σ_{c-} stand for standard deviation for positive and negative controls, respectively and μ_{c+} and μ_{c-} are the average values from positive and negative controls, respectively. The *Z*-factor values range from 0 to 1. Values above 0.5 are considered to be robust assays.

10. Data analysis: Calculate the percent activity using Eq. 11:

$$\% \text{Activity} = \frac{\text{Inh} - \text{MAX}}{\text{MIN} - \text{MAX}} \times 100 \quad (11)$$

where Inh is the FA value in the presence of a given compound. MIN is the FA value of the negative control (MIN inhibition=100% activity) and MAX is the positive control (MAX inhibition=0% activity). Compounds with percent inhibition values of greater than or equal to three times the standard deviation (3SD) of the negative controls are defined as active hits. Furthermore, samples with fluorescence intensity of perpendicular channel greater or less than 3SD of negative controls are considered false positives.

3.8 HTS for Inhibitors of *B. subtilis* RNase P: Confirmation Screens

3.8.1 Eliminating Compounds That Bind Substrate

3.8.2 Orthogonal Assay to Eliminate False Positives (See Note 29)

3.9 Dose-Response Experiments

The confirmation screen follows the procedure described for primary screening with the following modifications:

1. Add 10 μ l of Fl-pre-tRNA^{Asp} (40 nM) into wells in a 384-well black nonbinding surface microplate. Add either 0.1 μ l of compound (identified as active from primary screening) or DMSO (column 1–2 and 23–24) into Fl-pre-tRNA^{Asp}.
2. Incubate the plate at 30 °C for 30 min and record the FA signal. If the difference in the Fl-pre-tRNA^{Asp} FA signal is more than 3SD of the control well (containing substrate and DMSO only) upon addition of the compound, this compound is identified as a pre-tRNA binder.
3. Add 5 μ l of HTS buffer with 400 nM CaCl₂ to columns 23 and 24.
4. Initiate the reaction by adding 10 μ l of RNase P holoenzyme (0.3 nM PRNA and 3 nM P protein) to all wells (the final concentration of RNase P holoenzyme is 0.15 nM PRNA with 1.5 nM P protein and Fl-pre-tRNA^{Asp} is 20 nM).
5. Quench the reaction at 35 min by adding 5 μ l of HTS buffer with 400 mM CaCl₂ into each well except columns 23 and 24.
6. Calculate the percent inhibition as described in primary screening. Active compounds are defined as those repeatedly show at least 30% inhibition and are not identified as pre-tRNA binders.

Perform the cleavage of Fl-pre-tRNA^{Asp} by RNase P holoenzyme in the presence of identified compound under same condition as described in primary screening. However, instead of reading the FA value with microplate reader, perform traditional gel-based assay using Fl-labeled or radioactively labeled pre-tRNA^{Asp} [35].

1. Quench the reactions with 2× EDTA dye at various time points.
2. Separate Fl-pre-tRNA^{Asp} and Fl-5 nt-leader product by 20% denaturing PAGE, and visualize using a phosphorimager with an excitation laser of 488 nm and an emission filter of 535 nm. Active inhibitors decrease product formation, leaving mainly the Fl-pre-tRNA^{Asp} band on the denaturing PAGE.

Perform the dose-response experiments with the compounds that are deemed active after the orthogonal assays either using HTS equipment or in regular lab setting. Follow the MTO procedures to prepare *B. subtilis* RNase P holoenzyme (0.2 nM PRNA and 4 nM P protein) and Fl-pre-tRNA^{Asp} (100 nM) in HTS buffer. The concentrations of RNase P holoenzyme and Fl-pre-tRNA^{Asp} samples are at twice the final concentration.

3.9.1 Perform Dose-Response Experiments Using HTS Equipment

1. Prepare varying concentrations of compound. The concentrations of compound are at twice the final concentration. Take a 384-well black nonbinding surface microplate, perform two-fold serial dilution (*see* Subheading 3.4, step 3) by adding 0.2 μ l of stock compound into 0.2 μ l of DMSO using a nanoliter pintoool (*see* Note 30).
2. Add 5 μ l of RNase P holoenzyme sample into new wells and pinpoint 0.1 μ l of compound in serial concentrations. Incubate the RNase P holoenzyme and compound mix at 37 °C for 40 min.
3. Optimize the gain and beam position of the microplate reader and perform the assay at 30 °C (the final concentration of RNase P holoenzyme is 0.1 nM PRNA with 2 nM P protein and Fl-pre-tRNA^{Asp} is 50 nM), then calculate the percent activity as described in primary screen (*see* Subheading 3.7).

3.9.2 Perform Dose-Response Experiments in Regular Lab Setting

1. Prepare varying concentrations of compound. The concentrations of compound are at twice the final concentration. Add 60 μ l of RNase P holoenzyme into one micro test tube and 30 μ l (containing 0.6 μ l of DMSO) into the rest of the tubes. Add a small volume of compound dissolved in DMSO (*see* Note 30, for example, 1.2 μ l of stock compound) into 60 μ l of RNase P holoenzyme sample and mix well. For serial dilutions, transfer 30 μ l of the sample from this tube into another 30 μ l of RNase P holoenzyme sample and mix well. This step dilutes the compound by twofold while maintaining the RNase P holoenzyme concentration (0.2 nM PRNA and 4 nM P protein). Repeat the dilution to obtain eight different concentrations. Incubate the RNase P holoenzyme and compound at 37 °C for 40 min.
2. Take a 96-well half area black nonbinding surface microplate and transfer 20 μ l of Fl-pre-tRNA^{Asp} sample into a row of wells to obtain Fl-pre-tRNA^{Asp} with a final concentration of 50 nM.
3. Add 0.4 μ l of DMSO (no compound) into columns 1 and 2 as negative control, and 0.4 μ l of DMSO (no compound) and 20 μ l of HTS buffer into columns 11 and 12 as positive control.
4. Optimize the gain and beam position of the microplate reader and perform the assay at 37 °C as described in MTO assay (*see* Subheading 3.4). The final concentration of RNase P holoenzyme is 0.1 nM PRNA with 2 nM P protein and Fl-pre-tRNA^{Asp} is 50 nM.
5. Calculate the percent inhibition of each reaction by Eq. 11 using the initial rates of the FA trace instead of FA values. Initial rates are calculated from the time-dependent FA traces as described in Subheading 3.4, step 9 and Eq. 7.

6. Plot the percent activity as a function of compound concentrations and determine IC_{50} (50% loss of the enzyme activity, Fig. 2d) using Eq. 12:

$$\% \text{ Activity} = \frac{100}{1 + \left(\frac{[I]}{IC_{50}} \right)^n} \quad (12)$$

where n is the Hill coefficient and $[I]$ is the concentration of compound. Samples that show a concentration-dependent inhibitory activity are identified as active (see Note 31).

3.10 Determining the Mode of Inhibition

Compounds that showed inhibitory activity of RNase P are further evaluated to determine the mode of inhibition as illustrated in Fig. 3 [43, 44]. If compounds exhibit high Hill coefficient values in a dose-response curve, they could be an aggregator or denaturant as described [45]. To test reversibility of the compound, time-dependence experiments are conducted at varying concentrations and varying incubation times with enzyme before reaction initiation. Inhibitors are categorized into time dependent and time independent inhibitors. Time-dependent inhibitors can be covalent modifiers or compounds that induce a conformational change in the enzyme. These inhibitors call for a different types of analysis [44]. Here, we will focus on the determination of mode of inhibition of reversible inhibitors (time-independent inhibitors). For an in-depth discussion on the specific inhibition mode of *B. subtilis* RNase P, we direct the readers to a previous publication [32].

1. Following the procedure of MTO assay and dose-response screening, perform the assays under a fixed concentration of RNase P holoenzyme with varying concentrations of Fl-pre-tRNA^{Asp} and inhibitor (see Subheadings 3.4 and 3.9). Determine

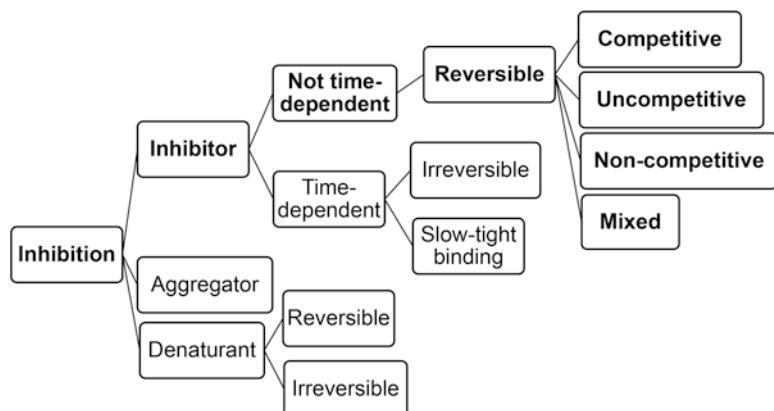


Fig. 3 A scheme showing the steps to determine the mode of enzyme inhibition in vitro [41, 42]

the initial cleavage rate under varying inhibitor and substrate concentrations and convert the values to activity (%). Obtain apparent steady-state kinetic parameters, $k_{\text{cat,app}}$ and $(k_{\text{cat}}/K_M)_{\text{app}}$, by fitting the Michaelis–Menten equation to the dependence of the RNase P cleavage activity on the Fl-pre-tRNA^{Asp} concentration [41] (Eq. 8) at various concentrations of the inhibitor. Fit Eq. 13 to the dependence of $k_{\text{cat,app}}$ and $(k_{\text{cat}}/K_M)_{\text{app}}$ on the inhibitor concentrations (Fig. 4a, b) to determine inhibition constants (K) and the Hill coefficient of cooperativity (n).

$$k_{\text{app}} = \frac{k}{1 + \left(\frac{[I]}{K}\right)^n} \quad (13)$$

where $k = k_{\text{cat,app}}$ or $(k_{\text{cat}}/K_M)_{\text{app}}$ and K is the inhibition constant of the inhibitor for RNase P.

2. Perform a global fit to all of the data for the inhibition of RNase P in GraphPad Prism software (Fig. 4c, d). Allow the fitting program to fit parameters (k_{cat} , K_M , K_i , K_{is} , and n) as “shared value for all data sets” and indicate the value of the inhibitor concentration by specifying the numbers in the titles of the data table. Use different models of inhibition (noncompetitive, competitive, uncompetitive, and mixed inhibition model), and compare the goodness of fit based on R^2 values to determine the best model for the type of inhibition. Generally, a competitive inhibitor, which mainly binds to the enzyme, has no effect on k_{cat} and decreases k_{cat}/K_M with increasing inhibitor concentration. An uncompetitive inhibitor, which mainly binds to the enzyme-substrate complex, decreases k_{cat} while k_{cat}/K_M does not change, with increasing inhibitor concentration. A noncompetitive inhibitor decreases k_{cat} and k_{cat}/K_M with increasing inhibitor concentration. Lastly, a mixed inhibitor decreases k_{cat} and varies in its effect on K_M and k_{cat}/K_M [44].

4 Notes

1. Light and alkali catalyze the conversion of acrylamide and bis-acrylamide to acrylic acid and bis-acrylic acid over time.
2. Equilibrate vial to room temperature before dissolving in DMSO to avoid moisture condensation. To dissolve 5-IAF in anhydrous DMSO, dry a needle in the oven and use it to draw DMSO under argon or nitrogen gas. 5-IAF is unstable when exposed to light and the functional group hydrolyzes in aqueous solution. Store 5-IAF powder in an original container at –80 °C and protect from light.

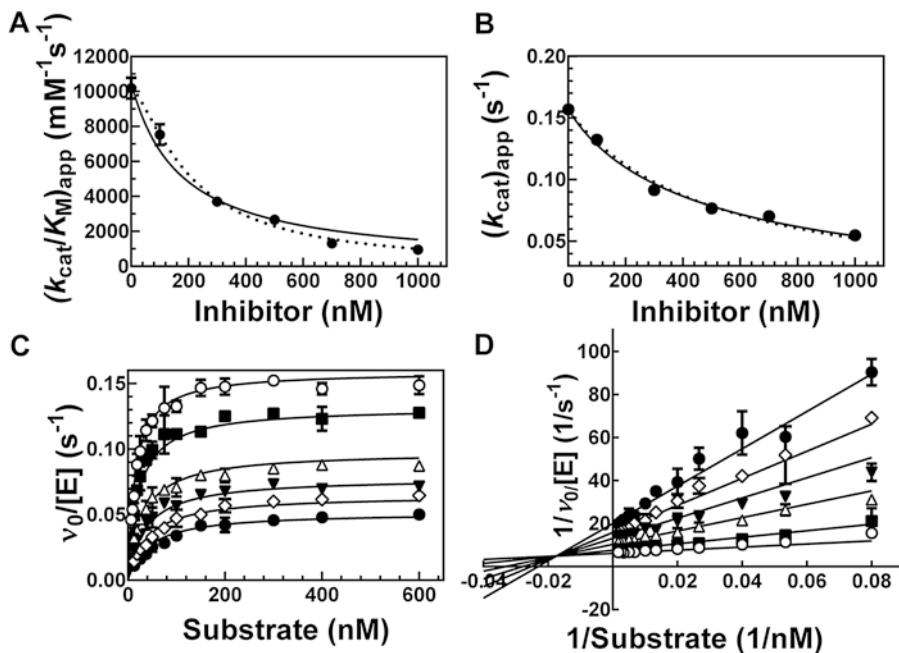


Fig. 4 Determination of inhibition mechanism of *B. subtilis* RNase P by Ir6Ac. The assays were carried out at a fixed RNase P concentration of 0.4 nM PRNA (2 nM P protein) with varying concentrations of Ir6Ac (100–1000 nM) and Fl-pre-tRNA^{Asp} (6–600 nM) in HTS buffer at 37 °C. Equation 13 is fit to the apparent k_{cat} (a) and k_{cat}/K_M (b) values as a function of the concentration of Ir6Ac. The solid line is a fit with $n=1$ and the dotted line is a fit where n is a variable: for k_{cat}/K_M , $n=1.4 \pm 0.1$ ($R^2=0.9969$) and for k_{cat} , $n=0.9 \pm 0.1$ ($R^2=0.9919$). (c) Best global fit for inhibition of RNase P in the presence of varying concentrations of Fl-pre-tRNA^{Asp} and Ir6Ac. (d) Lineweaver-Burk plot for the dependence of RNase P activity on the concentrations of Ir6Ac and Fl-pre-tRNA^{Asp}. A noncooperative mixed inhibition model is fit to the data ($R^2=0.9670$ for global fit). Symbols represent means \pm SD determined from two to three independent experiments at each concentration. (Reproduced from [32] with permission from Oxford University Press)

3. Precipitation can form when spermidine and MgCl₂ are mixed on ice.
4. Two tubes of 1 ml transcription are recommended for one batch of labeling. If a larger scale is desired, do multiple 1 ml transcription reactions instead of increasing the volume of the transcription reaction.
5. For a control reaction, remove 10 µl reaction mix immediately after the addition of T7 RNA polymerase but prior to adding pyrophosphatase. White precipitate (formation of Mg₂P₂O₇) will appear in 2–4 h after incubating at 37 °C indicating positive transcription reactivity.
6. Prewashing the filter decreases the loss of RNA caused by adhesion to the membrane. If doing a larger scale transcription reaction (more than 2 ml), Amicon® Ultra-15 Centrifugal Filter can be used, but the larger membrane leads to greater loss of RNA. A swing bucket centrifuge is recommended for

better filtration. Stop the centrifugation when ~1 ml of labeling buffer is left in the filter. Do not let the membrane dry.

7. If there is more than 200 μ l of the sample left after one round of centrifugation, resuspend the sample by gently inverting the tube to remix and centrifuge again. Keep each round of centrifugation between 15 and 20 min. Do not spin for more than 20 min without resuspending the RNA. This will avoid applying too much pressure on the sample at the bottom of the V-shaped filter.
8. This step removes excess nucleotide triphosphate, DTT, and GMPS from the transcribed 5'-GMPS-pre-tRNA^{Asp} and exchanges the RNA into the labeling buffer (pH 7.2).
9. Use amber or other dark micro test tubes. Alternatively, use a clear tube covered with aluminum foil to avoid light.
10. Normally, 2 ml of transcription yields ~200 μ l of ~150–200 μ M RNA, which requires ~1 mg of 5-IAF (40 μ l of 48.5 mM 5-IAF) for a 40-fold excess. The color of the 5-IAF solution should be slightly orange. Some precipitation might be observed if the solution is old. After overnight reaction the reaction mix clears if efficient labeling happens. If precipitation still exists, it suggests inefficient labeling, likely because the iodoacetamido moiety of the dye has been hydrolyzed.
11. Purify Fl-pre-tRNA^{Asp} from 2 ml of transcription using one 16.5 cm wide gel (2–3 mm in comb size) and run the gel at 20 W for 4 h. Use additional gels if the reaction is scaled up.
12. For 10% denaturing PAGE, Fl-pre-tRNA^{Asp} runs slower than xylene cyanol (upper dye on the gel). Detect fluorescence band on the gel by UV light (medium wave UV, 312 nm) then place the gel on a UV fluorescent TLC plate and detect the Fl-pre-tRNA^{Asp} by UV shadow (short wave UV, 254 nm). The dark RNA band should overlap with the fluorescence band. If there is DNA template band above the Fl-pre-tRNA^{Asp}, try not to cut the DNA band or add a DNase treatment step prior to PAGE. Free fluorescein also migrates throughout the gel. Therefore, an ethanol precipitation step prior to the adding the 2×EDTA dye and PAGE purification is beneficial when working on a pre-tRNAs of different lengths.
13. Prewashing the filter decreases the loss of RNA caused by adhesion to the membrane. A swing bucket centrifuge is recommended for better filtration. Stop the centrifugation when ~2 ml of washing buffer left in the filter. Do not let the membrane dry.
14. Keep adding eluted Fl-pre-tRNA^{Asp} into the centrifugal filter if the volume is so large that the concentration cannot be achieved in one round centrifugation. Keep each round of centrifugation to 15–20 min.

15. Purify PRNA from 1 ml of transcription using one 16.5 cm wide gel (comb thickness of 3 mm) and run the gel at 23 W for 3 h. Use additional gels if larger scale transcription reaction is desired. For 6% denaturing PAGE, PRNA runs above the xylene cyanol (upper dye on the gel).
16. To obtain a reasonable gain reading, calculate the gain for 50–80% intensity to avoid oversaturating the signal beyond the range of detection upon addition of enzyme.
17. When setting the assay buffer as “Blank,” the microplate reader corrects the background signal from the assay buffer. With G-factor = 0.926 in TECAN Infinity F500 microplate reader, the experimental FA reading for Fl-pre-tRNA^{Asp} is ~80 mA and FP reading is ~120 mA.
18. To minimize FA signal fluctuations due to temperature fluctuations, utilize the empty wells on the plate to pre-incubate enzyme or substrate. By pipetting the reagent from the plate using a multichannel and multidispense pipette, multiple reactions can be performed at the same time and the amounts of reagents can be minimized compared to using a reagent reservoir.
19. Under STO conditions, at the beginning of the reaction, the FA signal is higher compared to Fl-pre-tRNA^{Asp} alone. This is a result of RNase P binding to Fl-pre-tRNA^{Asp}. Upon cleavage of the 5' end leader catalyzed by RNase P, the Fl-5 nt-leader product leads to the decrease of FA signal.
20. The FA signal of Fl-5 nt-leader product is 20–35 mP (depending on the batch of substrate and the instrument used), indicating cleavage completion.
21. When measuring Fl-pre-tRNA^{Asp} at 37 °C, there is a nonlinear decrease in FA for the first few minutes after initiating reactions due to a change in temperature after the plate is moved into the microplate reader chamber (although this phenomenon is less prominent once we used the method described in **Note 19**). This is not detected when the experiment is performed at room temperature. The substrate Fl-pre-tRNA^{Asp} only control is used to adjust for this effect.
22. For STO reaction, the total fluorescence intensity decreases from the initial Fl-pre-tRNA^{Asp}-RNase P complex (F^0) to cleaved product (F^∞), and the extent of this decrease varies with pH. Therefore, an enhancement factor γ is used to correct for the total fluorescence change.
23. For twofold serial dilutions, make two series to cover a wider range of concentrations. For example, prepare two concentrations of Fl-pre-tRNA^{Asp}, 1.5 μM and 2 μM, and then make two individual series starting with 1.5 μM and 2 μM, respectively. Alternatively, a 1.4-fold serial dilution can be used by mixing

100 μ l of first concentration of substrate with 40 μ l of buffer B. Then taking out 100 μ l of the mixture to the next tube of 40 μ l of Buffer B for serial dilution.

24. Theoretically, the FA reading for all concentrations of Fl-pre-tRNA^{Asp} should be the same. However, in our experience there are variations in certain instruments.
25. Under MTO conditions, there is no significant increase in FA signal upon the addition of RNase P since Fl-pre-tRNA^{Asp} is in excess compared with RNase P. Cleavage of the 5' end leader, catalyzed by RNase P, to form the free Fl-5 nt-leader product leads to the decrease in the FA signal.
26. The FA signal of the Fl-5 nt-leader product is ~35 mP indicating cleavage completion. Collect at least one full time course to calculate the FA signal change upon complete conversion of Fl-pre-tRNA^{Asp} substrate to 5'-leader product. For the remainder of the reactions, the measurement can be stopped after collecting data in the linear range.
27. The cleavage of Fl-pre-tRNA^{Asp} catalyzed by RNase P is very slow in the presence of Ca²⁺. Therefore, the addition of excess Ca²⁺ can quench the reaction because Ca²⁺ rapidly competes with Mg²⁺. Using CaCl₂ instead of EDTA to quench the reaction enhances the dynamic range of the FA signal because the high FA signal of Fl-pre-tRNA^{Asp} is dependent on the pre-tRNA structure. EDTA chelates metal ions, which are important for stabilizing the pre-tRNA structure.
28. The target linear range should be between 30 and 60 min. If the linear range is less than 30 min, decrease the concentration of RNase P holoenzyme to increase the linear range. Ensure that optimal activity can still be achieved with a lower enzyme concentration. Choose a time to quench the reaction that is within the linear range of the reaction.
29. The primary assay relies on the fluorescence from Fl-pre-tRNA^{Asp}. If the compound has fluorescence, it affects the FA values from Fl-pre-tRNA^{Asp} by either enhancing or quenching the fluorophore and results in false positives in the primary screen.
30. Perform serial dilution in DMSO to ensure a consistent DMSO concentration in each reaction. Keep the concentration of DMSO low is crucial since RNase P is inhibited by DMSO. Furthermore, addition of a smaller volume of the compound has a smaller effect on the concentration of RNase P holoenzyme and Fl-pre-tRNA^{Asp}.
31. A value of $n=1$ means no cooperativity, while n values that are larger or smaller than 1 indicate positive or negative cooperativity, respectively.

Acknowledgment

This work was supported by grants from National Institute of Health [R01 GM55387 to C.A.F.], Pilot Screen Grant from the Center for Chemical Genomics at the University of Michigan [to C.A.F.], Rackham Graduate Student Research Grant [to X.L.]. Thanks go to Drs. Elaina Zverina, Lyra Chang, John Hsieh and Daina Zeng for helpful discussions on the development of the HTS assay and Professors Jason Gestwicki and Anna Mapp for sharing plate-reader instruments. We thank Martha Larsen, Steven Swaney, and Paul Kirchhoff at the Center for Chemical Genomics (CCG) at University of Michigan for their help with the compound library screen and data mining.

References

1. Hartmann RK, Gossringer M, Spath B, Fischer S, Marchfelder A (2009) The making of tRNAs and more—RNase P and tRNase Z. *Prog Nucleic Acid Res Mol Biol* 85:319–368
2. Iwata-Reuy D (2008) An embarrassment of riches: the enzymology of RNA modification. *Curr Opin Chem Biol* 12(2):126–133
3. Xiong Y, Steitz TA (2006) A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr Opin Struct Biol* 16(1):12–17
4. Phizicky EM, Hopper AK (2010) tRNA biology charges to the front. *Genes Dev* 24(17):1832–1860
5. Kazantsev AV, Pace NR (2006) Bacterial RNase P: a new view of an ancient enzyme. *Nat Rev Microbiol* 4(10):729–740
6. Smith JK, Hsieh J, Fierke CA (2007) Importance of RNA-protein interactions in bacterial ribonuclease P structure and catalysis. *Biopolymers* 87(5-6):329–338
7. Kirsebom LA (2007) RNase P RNA mediated cleavage: substrate recognition and catalysis. *Biochimie* 89(10):1183–1194
8. Harris ME, Christian EL (2003) Recent insights into the structure and function of the ribonucleoprotein enzyme ribonuclease P. *Curr Opin Struct Biol* 13(3):325–333
9. Randau L, Schroder I, Soll D (2008) Life without RNase P. *Nature* 453(7191):120–123
10. Walker SC, Engelke DR (2006) Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol* 41(2):77–102
11. Apirion D (1980) Genetic mapping and some characterization of the rnpA49 mutation of *Escherichia coli* that affects the RNA-processing enzyme ribonuclease P. *Genetics* 94(2):291–299
12. Eder PS, Hatfield C, Vioque A, Gopalan V (2003) Bacterial RNase P as a potential target for novel anti-infectives. *Curr Opin Investig Drugs* 4(8):937–943
13. Frank DN, Pace NR (1998) Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* 67:153–180
14. Hartmann E, Hartmann RK (2003) The enigma of ribonuclease P evolution. *Trends Genet* 19(10):561–569
15. Hall TA, Brown JW (2001) The ribonuclease P family. *Methods Enzymol* 341:56–77
16. Hsieh J, Andrews AJ, Fierke CA (2004) Roles of protein subunits in RNA-protein complexes: lessons from ribonuclease P. *Biopolymers* 73(1):79–89
17. Lai LB, Vioque A, Kirsebom LA, Gopalan V (2010) Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects. *FEBS Lett* 584(2):287–296
18. Esakova O, Krasilnikov AS (2010) Of proteins and RNA: the RNase P/MRP family. *RNA* 16(9):1725–1747
19. Thomas BC, Gao L, Stomp D, Li X, Gegenheimer PA (1995) Spinach chloroplast RNase P: a putative protein enzyme. *Nucleic Acids Symp Ser* 33:95–98
20. Gobert A, Buttmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P (2010) A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol* 17(6):740–744
21. Holzmann J, Frank P, Loffler E, Bennett KL, Gerner C, Rossmanith W (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial

- tRNA processing enzyme. *Cell* 135(3): 462–474
22. Trang P, Kim K, Liu F (2004) Developing RNase P ribozymes for gene-targeting and antiviral therapy. *Cell Microbiol* 6(6): 499–508
 23. Willkomm DK, Gruegelsiepe H, Goudinakis O, Kretschmer-Kazemi Far R, Bald R, Erdmann VA, Hartmann RK (2003) Evaluation of bacterial RNase P RNA as a drug target. *Chembiochem* 4(10):1041–1048
 24. Gossringer M, Kretschmer-Kazemi Far R, Hartmann RK (2006) Analysis of RNase P protein (*rnpA*) expression in *Bacillus subtilis* utilizing strains with suppressible *rnpA* expression. *J Bacteriol* 188(19):6816–6823
 25. Waugh DS, Pace NR (1990) Complementation of an RNase P RNA (*rnpB*) gene deletion in *Escherichia coli* by homologous genes from distantly related eubacteria. *J Bacteriol* 172(11):6316–6322
 26. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, Asai K, Ashikaga S, Aymerich S, Bessieres P, Boland F, Brignell SC, Bron S, Bunai K, Chapuis J, Christiansen LC, Danchin A, Debarbouille M, Dervyn E, Deuerling E, Devine K, Devine SK, Dreesen O, Errington J, Fillinger S, Foster SJ, Fujita Y, Galizzi A, Gardan R, Eschevins C, Fukushima T, Haga K, Harwood CR, Hecker M, Hosoya D, Hullo MF, 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, Mauel C, Meima R, Mellado RP, Moir A, Moriya S, Nagakawa E, Nanamiya H, Nakai S, Nygaard P, Ogura M, Ohanen T, O'Reilly M, O'Rourke M, Pragai Z, Pooley HM, Rapoport G, Rawlins JP, Rivas LA, Rivolta C, Sadaie A, Sadaie Y, Sarvas M, Sato T, Saxild HH, Scanlan E, Schumann W, Seegers JF, Sekiguchi J, Sekowska A, Seror SJ, Simon M, Stragier P, Studer R, Takamatsu H, Tanaka T, Takeuchi M, Thomaides HB, Vagner V, van Dijl JM, 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
 27. Mikkelsen NE, Brannvall M, Virtanen A, Kirsebom LA (1999) Inhibition of RNase P RNA cleavage by aminoglycosides. *Proc Natl Acad Sci U S A* 96(11):6155–6160
 28. Rueda D, Hsieh J, Day-Storms JJ, Fierke CA, Walter NG (2005) The 5' leader of precursor tRNA(Asp) bound to the *Bacillus subtilis* RNase P holoenzyme has an extended conformation. *Biochemistry* 44(49):16130–16139
 29. Hsieh J, Fierke CA (2009) Conformational change in the *Bacillus subtilis* RNase P holoenzyme—pre-tRNA complex enhances substrate affinity and limits cleavage rate. *RNA* 15(8): 1565–1577
 30. Hsieh J, Koutmou KS, Rueda D, Koutmos M, Walter NG, Fierke CA (2010) A divalent cation stabilizes the active conformation of the *B. subtilis* RNase P x pre-tRNA complex: a role for an inner-sphere metal ion in RNase P. *J Mol Biol* 400(1):38–51
 31. Giordano T, Sturgess MA, Rao SJ (2006) Inhibitors of RNase P proteins as antibacterial compounds. Google Patents
 32. Liu X, Chen Y, Fierke CA (2014) A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucleic Acids Res* 42(20), e159
 33. Jameson DM, Ross JA (2010) Fluorescence polarization/anisotropy in diagnostics and imaging. *Chem Rev* 110(5):2685–2708
 34. Behrman EJ (2000) An improved synthesis of guanosine 5'-monothiophosphate. *J Chem Res* 2000(9):446–447
 35. Beebe JA, Fierke CA (1994) A kinetic mechanism for cleavage of precursor tRNA(Asp) catalyzed by the RNA component of *Bacillus subtilis* ribonuclease P. *Biochemistry* 33(34): 10294–10304
 36. He B, Rong M, Lyakhov D, Gartenstein H, Diaz G, Castagna R, McAllister WT, Durbin RK (1997) Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expr Purif* 9(1):142–151
 37. Nirajanakumari S, Kurz JC, Fierke CA (1998) Expression, purification and characterization of the recombinant ribonuclease P protein component from *Bacillus subtilis*. *Nucleic Acids Res* 26(13):3090–3096
 38. Nirajanakumari S, Stams T, Crary SM, Christianson DW, Fierke CA (1998) Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. *Proc Natl Acad Sci U S A* 95(26):15212–15217
 39. Mocz G, Helms MK, Jameson DM, Gibbons IR (1998) Probing the nucleotide binding sites of axonemal dynein with the fluorescent nucleotide analogue 2'(3')-O-(N-Methylanthraniloyl)-adenosine 5'-triphosphate. *Biochemistry* 37(27):9862–9869
 40. Sem DS, McNeeley PA (1999) Application of fluorescence polarization to the steady-state enzyme kinetic analysis of calpain II. *FEBS Lett* 443(1):17–19
 41. Michaelis L, Menten ML, Johnson KA, Goody RS (2011) The original Michaelis constant:

- translation of the 1913 Michaelis-Menten paper. *Biochemistry* 50(39):8264–8269
42. Zhang JH, Chung TD, Oldenburg KR (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen* 4(2): 67–73
43. Seidler J, McGovern SL, Doman TN, Shoichet BK (2003) Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J Med Chem* 46(21): 4477–4486
44. Copeland RA (2000). Enzymes: a practical introduction to structure, mechanism, and data analysis. Wiley-VCH Inc.: 266–349
45. Prinz H (2010) Hill coefficients, dose-response curves and allosteric mechanisms. *J Chem Biol* 3(1):37–44

Chapter 13

Reporter Gene-Based Screening for TPP Riboswitch Activators

Christina E. Lünse and Günter Mayer

Abstract

With the rise of multidrug resistant bacteria and a growing number of nosocomial infections, there has been an increased interest in finding new antibacterial drugs and drug targets. Riboswitches represent attractive new antibacterial drug targets, because they not only inherently recognize a specific metabolite or ion with their RNA aptamer domain, but also often regulate essential metabolic pathways. Here, we describe a reporter gene-based screen to identify compounds that activate the thiamine pyrophosphate (TPP) riboswitch in bacteria. This assay can be easily adapted for different riboswitch classes and thus has the potential to target many essential metabolic pathways and a broad spectrum of bacterial pathogens.

Key words Translational fusion, Miller assay, Thiamine pyrophosphate, *lacZ* reporter gene, Antibiotics

1 Introduction

Riboswitches are structured RNA elements, mostly found in the 5' UTR of bacterial mRNAs. They regulate gene expression either on the transcriptional or translational level in response to binding a specific metabolite or ion [1]. Since their discovery, riboswitches have been explored as new antibacterial drug targets [2, 3]. Natural and synthetic ligand analogues have been identified for several riboswitch classes to date [4–9]. Recently, the first riboswitch-targeting compound was reported that is not a metabolite analogue, but structurally distinct from the natural ligand [10].

Here, we present a detailed description of a reporter gene-based screening for thiamine pyrophosphate (TPP) riboswitch-activating compounds. A TPP riboswitch found in the 5' UTR of the *Escherichia coli thiM* gene (Fig. 1a) is used to construct a translational fusion plasmid (Fig. 1b) [6] that is placed in a strain of *E. coli* (BW25113). These cells are grown with and without thiamine or candidate compounds to screen for riboswitch activation (Fig. 1c, d). In this case, the riboswitch is “activated,” when it binds its cognate metabolite and changes its secondary structure so

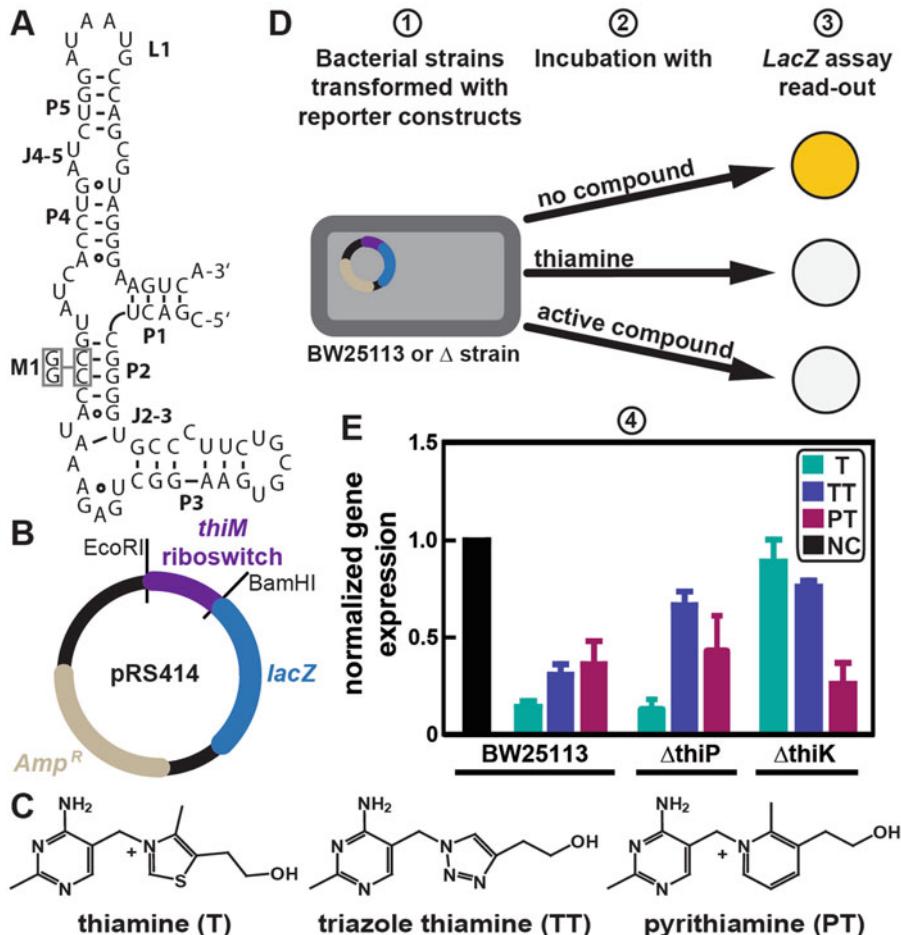


Fig. 1 Reporter gene-based screening for thiamine pyrophosphate riboswitch-activating compounds. (a) Secondary structure of the *thim* riboswitch aptamer domain as determined by structural probing [18]. CC to GG nucleotide changes of inactive mutant M1 are highlighted by gray boxes. Watson-Crick base pairing is indicated by dashes, open circles designate non-Watson-Crick interactions. Stems and junctions in the RNA secondary structure are labeled P1 through P5 as well as J2-3 and J4-5, black lines indicate 0 nucleotide connections to aid secondary structure representation. (b) Schematic for the construction of a *thim* riboswitch translational fusion reporter plasmid, which uses restriction sites *EcoRI* and *BamHI* to directionally insert the target sequence in frame with the *lacZ* coding sequence. (c) Chemical structures of thiamine (T) and its analogs triazolethiamine (TT) and pyritthiamine (PT). (d) Main steps of reporter gene-based screening for TPP riboswitch activators: Step 1. *E. coli* strain BW25113 or variants thereof in which different genes of thiamine biosynthesis or transport are deleted (e.g., Δ thiP, or Δ thiK) were transformed with the translational fusion plasmid. Step 2. Overnight cultures are diluted and grown to exponential phase, when they are incubated with thiamine or compounds, e.g., TT or PT. Control experiments are carried out using a translational fusion of the inactive mutant *thim* M1 whose reporter gene expression remains uninfluenced by thiamine or compounds (not shown). Step 3. Cell lysis and determination of β -galactosidase expression by reading the colorimetric change induced by β -galactosidase mediated breakdown of ONPG (yellow color, samples represented by circles). When reporter gene expression is decreased in the presence of thiamine or compound, less yellow color is generated (represented by white circles). (e) Step 4. Plot of normalized reporter gene expression (black bar, NC stands for “no compound”) in different *E. coli* strains (BW25113, Δ thiP, Δ thiK) and in the presence of thiamine (T), triazolethiamine (TT), and pyritthiamine (PT). Use of the different deletion strains illustrates that each compound assayed requires different degrees of phosphorylation mediated by thiamine kinase (ThiK) or prefers different routes of cell entry, such as active transport (ThiP) versus diffusion

that the ribosome binding site is sequestered and translation cannot occur (OFF-switch). Therefore, *thiM* riboswitch reporter gene fusions exhibit thiamine-dependent suppression of β -galactosidase expression (Fig. 1e). This repression is abolished when a *thiM* riboswitch mutant (M1, Fig. 1a) is used that disrupts essential base pairing in stem P2 [11] and thus prohibits ligand binding. This control allows the exclusion of compounds with riboswitch-unspecific effects from further analysis. After cell lysis, β -galactosidase activity is measured using the colorimetric change induced by the enzymatic cleavage of *o*-nitrophenyl- β -galactopyranoside (ONPG). Lastly, the β -galactosidase expression relative to its expression in the absence of thiamine is plotted (Fig. 1e).

After initial screening with a wild-type *E. coli* strain, deletion strains, for example taken from the Keio collection, can be used to investigate how the absence of certain metabolic enzymes influences intracellular hit compound conversion. As an example we show here how phosphorylation or transport of thiamine effects TPP riboswitch-dependent reporter gene expression in strains in which genes of thiamine biosynthesis or transport are deleted (Table 1, Fig. 1e).

The reporter gene-based screening approach described here is based on the thiamine pyrophosphate riboswitch, but can be easily adapted to feature other riboswitch classes or further modified for HT-compatible screenings [13]. This assay is not only inexpensive, but also easy to set up as it does not require extraordinary equipment or materials. Furthermore, a positive assay outcome immediately reflects compound take-up and metabolic conversions, some of which can be followed using the appropriate deletion strain.

2 Materials

Prepare all solutions using deionized water. Store all materials at room temperature unless indicated otherwise, bacterial glycerol

Table 1
Examples of bacterial strains used to screen for TPP riboswitch-targeting compounds [6]

Bacterial strain	Biological role	Genotype	ECSC number
BW25113	Wild-type strain	F-, Δ (araD-araB)567, λ -, <i>rph</i> -1, Δ lacZ4787(:rrnB-3), Δ (rhaD-rhaB)568, <i>hsdR514</i>	7636
Δ <i>thiP</i>	Transporter	F-, Δ (araD-araB)567, Δ <i>thiP774::kan</i> , Δ lacZ4787(:rrnB-3), λ -, <i>rph</i> -1, Δ (rhaD-rhaB)568, <i>hsdR514</i>	8368
Δ <i>thiK</i>	Kinase	F-, Δ (araD-araB)567, Δ <i>thiK768::kan</i> , <i>rph</i> -1, Δ lacZ4787(:rrnB-3), λ -, Δ (rhaD-rhaB)568, <i>hsdR514</i>	9034

Specific deletion strains from the Keio collection [14] were used to investigate how compound efficiency depends on key metabolic enzymes such as transporter proteins or kinases

stocks are stored at -80 °C. Cell culture components need to be sterilized, either by autoclaving at 120 °C for 15 min or filtration.

2.1 Cell Culture

1. 5× M9 salts stock: 15 g/l KH₂PO₄, 5 g/l NH₄Cl, 2.5 g/l NaCl, 30 g/l Na₂HPO₄, add deionized water to 1 l, sterilize by autoclaving.
2. 1 M magnesium sulfate (MgSO₄) solution.
3. 20% glucose (w/v): filter-sterilize and store at 4 °C.
4. 1 M calcium chloride (CaCl₂) solution.
5. M9 medium: For 1 l mix 200 ml 5× M9 minimal medium salt stock, add MgSO₄ (5 mM) and glucose (0.2% w/v), casamino acids (0.2 µg/µl) to final concentration indicated in parenthesis. Add water to 1 l (see Notes 1 and 2).
6. Lysogeny broth (LB) Lennox standard medium: 10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl in water, sterilize by autoclaving (see Note 2).
7. Glycerol (100%): sterilize by autoclaving.

2.2 Make Competent Cells

1. Solution of 50 mM magnesium chloride (MgCl₂) and 80 mM CaCl₂.
2. 100 mM CaCl₂ solution.
3. Glycerol (50% in ultrapure water, v/v): sterilize by autoclaving.
4. 50 mg/ml kanamycin (final concentration 50 µg/ml) (see Note 3).
5. Spectrophotometer.
6. Dewar and liquid nitrogen (see Note 4).
7. Bacterial strains: BW25113 and Keio collection strains with deletions of genes of interest [14], in this case thiamine-related genes involved in transport (such as $\Delta thiP$, Table 1), precursor phosphorylation (such as $\Delta thiK$, Table 1), as well as biosynthesis and rescue [6] (not shown) can be used for compound characterization.

2.3 Transformation

1. translational fusion plasmid pRS414 [15].
2. *thiM* riboswitch insert containing the natural promoter sequence (see Note 5) and flanked by *EcoRI* and *BamHI* restriction sites generated from genomic DNA (*E. coli* strain DH10b) by PCR using primers EcoRI.thiM.F2 (5'-TCTACAAGTGAAT TCCTGCCGTTTCCTCGTTACAA-3') and Rev thiM-RS (5'-TTGCGCTGGATCCAGCAGGTCGA-3') [16].
3. SOC medium: 20 g tryptone, 5 g yeast extract, 10 mM NaCl, 2.5 mM KCl, add 800 ml of water. Autoclave. Add sterile filtered MgCl₂ (10 mM), MgSO₄ (10 mM) and glucose (20 mM) to the final concentrations given in parenthesis. Add sterile water to 1 l (see Note 2).

4. 50 mg/ml kanamycin stock solution (final concentration 50 µg/ml) (*see Note 3*).
5. 100 mg/ml ampicillin stock solution (final concentration 100 µg/ml).

2.4 LacZ-Based Reporter Gene Assay

1. Kanamycin: 50 mg/ml (final concentration 50 µg/ml) (*see Note 3*).
2. Ampicillin: 100 mg/ml (final concentration 100 µg/ml).
3. Cells transformed with pRS414 with *thiM* riboswitch insert or with *thiM M1* riboswitch insert.
4. Phosphate buffered saline, pH 7.4 (1× PBS): 8 g NaCl, 0.2 g KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, add 800 ml of water to dissolve salts, adjust pH to 7.4 using HCl, add water to a total volume of 1 l.
5. Lysis buffer: 5× reporter lysis buffer diluted to 1× with water (*see Note 6*).
6. 2× Assay buffer: 200 mM sodium phosphate buffer (pH 7.3, *see Note 7*), 2 mM MgCl₂, 100 mM β-mercaptoethanol, 1.33 mg/ml *o*-nitrophenyl-β-galactopyranosid (ONPG).
7. 1 M sodium carbonate (Na₂CO₃) solution.
8. Thiamine and library compounds in DMSO (500 µM final concentration in assay).
9. Microplate reader, such as Nanoquant Infinite 200 (Tecan).

3 Methods

3.1 Plasmid Design and Cloning

1. The *thiM* riboswitch sequence including its natural promoter (*see Note 5*) is cloned into the translational fusion plasmid pRS414 [15] upstream of, and in-frame with the β-galactosidase reporter gene using *EcoRI* and *BamHI* restriction sites.
2. Transform ligated plasmid into competent *E. coli* cloning host cells (commercially available or self-made, *see Subheading 3.2*), such as *E. coli* DH5α, and select for transformants with ampicillin resistance by streaking cells on agar plates containing 100 µg/ml ampicillin.
3. Confirm the integrity of the resulting plasmid by DNA sequencing.
4. Transform confirmed translational fusion plasmids into Keio Collection strains (*see Subheading 3.3*).

3.2 Make Competent Cells

1. Grow pre-cultures overnight of strains to be made competent, e.g., *E. coli* BW25113 (no antibiotic resistance) or Keio knock-out strains (kanamycin resistance).

2. Dilute overnight culture to an optical density at 600 nm (OD_{600}) of 0.05–0.1 in 5 ml and grow cells at 37 °C under vigorous shaking.
3. Monitor growth by measuring OD_{600} at different time points, at OD_{600} of ~0.96 harvest cells [17] by centrifugation ($4500 \times g$ for 5 min).
4. Resuspend the cell pellet in 500 µl ice cold 80 mM CaCl₂ and 50 mM MgCl₂ solution and incubate on ice for 10 min.
5. Repeat step 4 twice.
6. Resuspend pellet in 200 µl ice-cold 0.1 M CaCl₂ solution to about 5×10^9 cells/ml
7. Mix suspension with an equal volume of 50% glycerol. Immediately flash freeze competent cell aliquots (80 µl) using liquid nitrogen (see Note 4) and store at –80 °C until used for transformation.

3.3 Transformation

1. Thaw 80 µl aliquot of competent cells on ice for 10 min.
2. Mix cells with 50 pg of plasmid (pRS414 *thiM* or pRS414 *thiM* M1) and incubate for 30 min on ice.
3. Heat shock cells at 42 °C for 40 s in a water bath and rapidly return them to ice for 3 min.
4. Add 400 µl SOC medium and incubate at 37 °C for 1 h.
5. Plate cells on LB agar plates with respective antibiotic and incubate overnight at 37 °C (see Note 8).

3.4 *lacZ* Assay and Analysis

1. Streak control and knock out cells containing the appropriate reporter plasmids from glycerol stocks on LB agar plates with required antibiotic (see Note 8). Incubate plates overnight at 37 °C.
2. Pick a single colony to inoculate pre-cultures (3 ml LB with antibiotic in a 15 ml culture tube (see Note 8)). Incubate pre-cultures overnight at 37 °C under vigorous shaking (155 rpm).
3. Dilute cells to an OD_{600} of 0.5.
4. Use this dilution to inoculate β-galactosidase expression cultures at a ratio of 1:500 (2 ml final volume) in M9 medium containing the appropriate antibiotics (see Note 8).
5. Add thiamine (20 µM) or compounds to be screened (500 µM) to each culture tube (see Note 9).
6. Incubate cultures for 24 h at 37 °C at 155 rpm.
7. Measure and record OD_{600} (see Note 10). Sediment remaining cells by centrifugation at $4500 \times g$ for 5 min, discard supernatant, and wash cell pellets in 400 µl 1× PBS. Repeat centrifugation and washing step.

8. Resuspend pellets in 200 μ l 1 \times lysis buffer and incubate for 15 min at room temperature. Pellet cells by centrifugation (4500 \times g, 3 min, *see Note 11*).
9. Mix 75 μ l cleared cell lysate (supernatant) to 75 μ l 2 \times assay buffer (*see Note 12*). Incubate for 5–15 min at 37 °C.
10. Stop reaction by adding 250 μ l Na₂CO₃. Add 200 μ l of stopped reaction to a clear, flat bottom 96-well plate. Measure and record absorbance at 420 nm using a microplate reader to monitor β -galactosidase activity in duplicate.
11. Calculate Miller units using the following equation: Miller units = (1000 \times OD₄₂₀) / (OD₆₀₀ \times incubation time [min] \times culture volume [l]). Miller units of control without compound or thiamine added were set to 1 (*see Note 13*).

4 Notes

1. Even though casamino acids that are used to make minimal medium are vitamin deprived, thiamine auxotrophic strains can grow without the addition of thiamine. This is due to small but constant amounts of thiamine present in the casamino acid stock.
2. Before autoclaving the pH of media can be adjusted to 7.5 using sodium hydroxide (NaOH).
3. Keio collection strains with a deletion of a nonessential gene contain a kanamycin resistance cassette. Therefore, any time such a knock-out strain is chosen, kanamycin has to be added to the culture medium.
4. When handling liquid nitrogen wear personal protective gear such as lab coat, gloves, and lab goggles at all times. Be aware that contact of liquid nitrogen with the skin or eyes may cause serious freezing (frostbite) injuries.
5. Here, the natural promoter was used; however, the use of a constitutively active promoter such as *E. coli lysC* can be indicated for some riboswitches, especially when the expression provided by the natural promoter is too weak or interferes with standard culturing conditions.
6. We used 5 \times reporter lysis buffer obtained from Promega (E3971). Other lysis buffer products could be used as long as they lead to the disruption of cells and do not interfere with reporter gene read-out.
7. To make sodium phosphate buffer prepare a 1 M dibasic sodium phosphate solution (Na₂HPO₄) and a 1 M monobasic sodium phosphate solution (NaH₂PO₄). To prepare the stock

solutions, dissolve 138 g of NaH₂PO₄·H₂O (monobasic; m.w. = 138) in sufficient deionized water to make a final volume of 1 l and dissolve 142 g of Na₂HPO₄ (dibasic; m.w. = 142) in sufficient deionized water to make a final volume of 1 l. To obtain a 0.2 M sodium phosphate solution with a pH of 7.3 mix 145.8 ml of 1 M Na₂HPO₄ and 54.2 ml of 1 M NaH₂PO₄ with 800 ml deionized water.

8. This means BW25113 cells containing the reporter plasmid are grown with 100 µg/ml ampicillin and Keio collection knock-out strains containing the reporter plasmid are grown with 100 µg/ml ampicillin and 50 µg/ml kanamycin in liquid media or on agar plates.
9. If compounds were dissolved in DMSO, controls were supplemented with equal amounts of DMSO, here 1% [6].
10. We recommend taking a 200 µl aliquot of the cell culture and transferring it into a clear, flat bottom 96-well plate to measure OD₆₀₀.
11. To save time, use this incubation step to start labeling 1.5 ml tubes with the sample IDs for **step 9** and aliquot 75 µl of 2× assay buffer into them.
12. When handling 2× assay buffer always work in a fume hood because it contains toxic β-mercaptoethanol.
13. The β-galactosidase activity is measured using the colorimetric change induced by the enzymatic cleavage of *o*-nitrophenyl-β-galactopyranoside (ONPG) [12]. ONPG is cleaved by β-galactosidase to yield galactose and *o*-nitrophenol which has a yellow color. When ONPG is in excess over the β-galactosidase enzyme in a reaction, the production of *o*-nitrophenol per unit time is proportional to the concentration of β-galactosidase enzyme. Therefore, the production of yellow color can be used to determine enzyme concentration and hence β-galactosidase expression. For this reporter example, when no compound is added, the *thiM* riboswitch is “on” and gene expression occurs leading to yellow color and high Miller Unit count in the assay (Fig. 1e, NC). When thiamine is added to the culture medium, the cells take it up and convert it to TPP. TPP binds to the *thiM* riboswitch and leads to a change in the RNA secondary structure that turns off gene expression by sequestering the ribosome binding site. Therefore, less β-galactosidase is produced and Miller Units decrease (Fig. 1e, T). Compounds that act in a similar way to thiamine on the *thiM* riboswitch lead to a decrease in gene expression (Fig. 1e, PT or TT).

Acknowledgments

G.M. was supported by a grant from the German Research Foundation (MA3442/1-2). C.E.L. was supported by the German Research Foundation grant LU1889-1 and NIH P01 grant GM022778 awarded to Professor Ronald R. Breaker, HHMI investigator, Yale University.

References

1. Breaker RR (2011) Prospects for riboswitch discovery and analysis. *Mol Cell* 43(6):867–879. doi:[10.1016/j.molcel.2011.08.024](https://doi.org/10.1016/j.molcel.2011.08.024)
2. Blount KF, Breaker RR (2006) Riboswitches as antibacterial drug targets. *Nat Biotechnol* 24(12):1558–1564. doi:[10.1038/nbt1268](https://doi.org/10.1038/nbt1268)
3. Lünse CE, Schüller A, Mayer G (2014) The promise of riboswitches as potential antibacterial drug targets. *Int J Med Microbiol* 304(1):79–92. doi:[10.1016/j.ijmm.2013.09.002](https://doi.org/10.1016/j.ijmm.2013.09.002)
4. Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR (2007) Antibacterial lysine analogs that target lysine riboswitches. *Nat Chem Biol* 3(1):44–49. doi:[10.1038/nchembio842](https://doi.org/10.1038/nchembio842)
5. Lünse CE, Schmidt MS, Wittmann V, Mayer G (2011) Carba-sugars activate the glmS-riboswitch of *Staphylococcus aureus*. *ACS Chem Biol* 6(7):675–678. doi:[10.1021/cb200016d](https://doi.org/10.1021/cb200016d)
6. Lünse CE, Scott FJ, Suckling CJ, Mayer G (2014) Novel TPP-riboswitch activators bypass metabolic enzyme dependency. *Front Chem* 2:53. doi:[10.3389/fchem.2014.00053](https://doi.org/10.3389/fchem.2014.00053)
7. Blount KF, Meggyola C, Plummer M, Osterman D, O'Connell T, Aristoff P, Quinn C, Chruscil RA, Poel TJ, Schostarez HJ, Stewart CA, Walker DP, Wuts PG, Breaker RR (2015) Novel riboswitch-binding flavin analog that protects mice against *Clostridium difficile* infection without inhibiting cecal flora. *Antimicrob Agents Chemother* 59(9):5736–5746. doi:[10.1128/AAC.01282-15](https://doi.org/10.1128/AAC.01282-15)
8. Kim JN, Blount KF, Puskarz I, Lim J, Link KH, Breaker RR (2009) Design and antimicrobial action of purine analogues that bind Guanine riboswitches. *ACS Chem Biol* 4(11):915–927. doi:[10.1021/cb900146k](https://doi.org/10.1021/cb900146k)
9. Mulhbacher J, Brouillet E, Allard M, Fortier LC, Malouin F, Lafontaine DA (2010) Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog* 6(4), e1000865. doi:[10.1371/journal.ppat.1000865](https://doi.org/10.1371/journal.ppat.1000865)
10. Howe JA, Wang H, Fischmann TO, Balibar CJ, Xiao L, Galgoci AM, Malinvern JC, Mayhood T, Villafania A, Nahvi A, Murgolo N, Barbieri CM, Mann PA, Carr D, Xia E, Zuck P, Riley D, Painter RE, Walker SS, Sherborne B, de Jesus R, Pan W, Plotkin MA, Wu J, Rindgen D, Cummings J, Garlisi CG, Zhang R, Sheth PR, Gill CJ, Tang H, Roemer T (2015) Selective small-molecule inhibition of an RNA structural element. *Nature* 526(7575):672–677. doi:[10.1038/nature15542](https://doi.org/10.1038/nature15542)
11. Winkler W, Nahvi A, Breaker RR (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419(6910):952–956. doi:[10.1038/nature01145](https://doi.org/10.1038/nature01145)
12. Miller JH (1992) A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Plainview, NY
13. Nelson JW, Plummer MS, Blount KF, Ames TD, Breaker RR (2015) Small molecule fluoride toxicity agonists. *Chem Biol* 22(4):527–534. doi:[10.1016/j.chembiol.2015.03.016](https://doi.org/10.1016/j.chembiol.2015.03.016)
14. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2(2006):0008. doi:[10.1038/msb4100050](https://doi.org/10.1038/msb4100050)
15. Simons RW, Houman F, Kleckner N (1987) Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene* 53(1):85–96
16. Mayer G, Raddatz MS, Grunwald JD, Famulok M (2007) RNA ligands that distinguish metabolite-induced conformations in the TPP riboswitch. *Angew Chem Int Ed Engl* 46(4):557–560. doi:[10.1002/anie.200603166](https://doi.org/10.1002/anie.200603166)
17. Tang X, Nakata Y, Li HO, Zhang M, Gao H, Fujita A, Sakatsume O, Ohta T, Yokoyama K (1994) The optimization of preparations of competent cells for transformation of *E. coli*. *Nucleic Acids Res* 22(14):2857–2858
18. Rentmeister A, Mayer G, Kuhn N, Famulok M (2007) Conformational changes in the expression domain of the *Escherichia coli* thiM riboswitch. *Nucleic Acids Res* 35(11):3713–3722. doi:[10.1093/nar/gkm300](https://doi.org/10.1093/nar/gkm300)

Chapter 14

Cell-Based Fluorescent Screen to Identify Inhibitors of Bacterial Translation Initiation

Federica Briani

Abstract

A strategy that can be applied to the research of new molecules with antibacterial activity is to look for inhibitors of essential bacterial processes within large collections of chemically heterogeneous compounds. The implementation of this approach requires the development of proper assays aimed at the identification of molecules interfering with specific cell pathways and potentially applicable to the high throughput analysis of large chemical library. Here, I describe a fluorescence-based whole-cell assay in *Escherichia coli* devised to find inhibitors of the translation initiation pathway. Translation is a complex and essential mechanism. It involves numerous sub-steps performed by factors that are in many cases sufficiently dissimilar in bacterial and eukaryotic cells to be targetable with domain-specific drugs. As a matter of fact, translation has been proven as one of the few bacterial mechanisms pharmacologically tractable with specific antibiotics. The assay described in this chapter is tailored to the identification of molecules affecting the first stage of translation initiation, which is the most dissimilar step in bacteria *vs.* mammals. The effect of the compounds under analysis is assayed in living cells, thus allowing evaluating their *in vivo* performance as inhibitors of translation initiation. Compared with other assays for antibacterials, the major advantages of this screen are its simplicity and high mechanism specificity.

Key words Translation initiation, Whole-cell assay, Gram-negative bacteria, Leaderless mRNA, S1 ribosomal protein, Ribosome, Antibacterial compounds

1 Introduction

Protein synthesis is initiated by the assembly on the mRNA of a translation-competent multi-subunit complex constituted by the ribosome, the translation initiation factors, and the initiator tRNA. This process occurs through very different multistep pathways in Bacteria *vs.* Eukarya [1, 2]. In bacteria the first step is the binding of the 30S small ribosomal subunit to the mRNA translation initiation region (TIR), a site within the 5'-untranslated region (5'-UTR) including the Shine-Dalgarno (SD) sequence and the AUG start codon. 30S binding is mediated by the interaction between the mRNA TIR and two main ribosomal elements: the anti-SD, i.e., a sequence at the 16S rRNA 3'-end complementary to the SD,

and the ribosomal protein S1. S1 has a prominent role in translation in gram-negative and high GC-content gram-positive bacteria [3–5]; it has been shown that in *Escherichia coli*, S1 is essential for the translation of most mRNAs [6], with the exclusion of a peculiar class of transcripts that are devoid of the 5'-UTR (i.e., leaderless mRNAs) [7–9]. Leaderless mRNAs are translated through a non-canonical, S1-independent pathway that occurs through the loading of a preassembled 70S ribosome on the mRNA 5'-end AUG [8, 10, 11]. Leaderless transcripts are very rare in *E. coli*. One of the few examples is represented by bacteriophage λ *cI* mRNA, which lacks a 5'-UTR when transcribed from the P_{RM} promoter.

S1 and the S1-dependent translation step appear as very good targets for new antibiotics. In fact, S1 is widely conserved among bacteria [4] and is encoded by essential genes in evolutionarily distant bacteria as *E. coli* and *Mycobacterium tuberculosis* ([12, 13]; TubercuList database). On the other hand, in mammals an S1 orthologue is absent from both the 40S ribosomal subunit and the mitochondrial ribosomes. No inhibitors of the essential S1 activity in translation have been identified so far.

I devised a simple whole-cell fluorescent assay aimed at identifying molecules interfering with the S1-dependent pathway of translation initiation [14]. The assay exploits the differences between leadered and leaderless translation to specifically target the S1 mechanism. In the first step, compounds of interest are screened for their effect on the expression of a reporter gene with a 5'-UTR (leadered reporter) cloned in *E. coli*. In the second step, inhibitors of the leadered reporter are analyzed for their effect on the expression of a leaderless variant of the same reporter. Molecules inhibiting the S1-dependent mechanism should not affect the leaderless reporter expression, whose translation does not require S1. This prediction was verified by exploiting the aminoglycoside antibiotic kasugamycin, which specifically inhibits leadered mRNA translation through a complex mechanism that involves the generation of minimal ribosomes lacking several ribosomal proteins, among which S1 [14–16]. The results of the analysis with kasugamycin demonstrated that in principle, compounds that selectively target the canonical pathway of bacterial translation initiation can be identified through the assay [14]. The application of the assay to the screening of the Prestwick library, a small collection of 1120 chemically heterogeneous compounds FDA-approved for clinical use, failed to provide any hits, a result that was expected given the small size of the library and the predicted high specificity of the assay [14]. Such specificity would be a relevant asset in the application of the assay to high-throughput screenings of large collections of compounds as it should allow to (1) identify and discard molecules generically toxic to the cells or acting through mechanisms different from translation initiation inhibition; (2) expedite the

identification of the molecular target(s) of the hits by restricting their research among cellular factors specifically involved in translation initiation.

2 Materials

2.1 Bacterial Strains

The AS19/pGM991, AS19/pGM999, and DH10B/pGM991 *E. coli* strains are exploited in the screening steps [14, 17, 18]. AS19 is an *E. coli* B derivative with increased permeability due to an unmapped mutation altering the outer membrane [17]. This defect facilitates the entry of compounds into the cells. DH10B is an *E. coli* K-12 laboratory strain with no reported permeability defects [18]. The strains carry either pGM991 or pGM999 plasmid, which are both derivatives of pGM963 [19], a multicopy shuttle vector conferring ampicillin resistance. The vector carries the *araC* gene and *araBp* promoter of the *E. coli* arabinose operon [20] inserted upstream of the eGFP open reading frame (ORF) devoid of the Shine-Dalgarno sequence and ATG start codon. In pGM991 and pGM999, DNA fragments encompassing either the 5'-UTR and first 9 codons of *E. coli recA* or the first 189 bp of phage λ *cI* ORF, respectively, are inserted in frame with the eGFP ORF. Transcription of pGM991 and pGM999 from *araBp* gives either a leadered *recA*-eGFP (pGM991) or a leaderless *cI*-eGFP mRNA (pGM999).

2.2 Culture Media and Growth Conditions

Prepare all media and solutions with analytical-grade water (hereafter indicated as aH₂O) and reagents. Prepare and store all solutions and media at room temperature, if not differently stated.

1. LD broth: LD is a variation of Lysogeny Broth (LB) [21] prepared by dissolving in 1 l of aH₂O 10 g of Bactotryptone, 5 g of Yeast extract, and 5 g of NaCl. Sterilize by autoclaving.
2. LD agar: dissolve 10 g of Agar in 700 ml of LD. Sterilize by autoclaving.
3. Trace elements mix: dissolve in 1 l of aH₂O 3.2 g FeCl₃·6H₂O, 0.06 g CaCl₂·6H₂O, 0.17 g Zn(Acetate)₂·2H₂O, 0.14 g MnSO₄·H₂O, and 0.05 g CuSO₄·5H₂O. Sterilize by filtration through 0.22 µm cellulose acetate membrane. Store at 4 °C.
4. M9 minimal medium: in 1 l of aH₂O dissolve 2 g NH₄Cl, 32 g Na₂HPO₄·12H₂O, 6 g KH₂PO₄, and 10 g NaCl. Sterilize by autoclaving. To 250 ml of the above sterile solution, add in the stated order the following sterile components: 0.09 ml 0.5 M CaCl₂, 0.54 ml 1 M MgSO₄, and 0.05 ml trace elements mix.
5. To provide proper aeration to bacterial cultures in glass tubes, the tubes were placed in a rotatory device. We used a New Brunswick Roller Drum TC-7 unit.

2.3 Chemicals

1. DMSO, molecular biology grade.
2. D(+)-Glucose. Dissolve 40 g in 100 ml of aH₂O. Sterilize by autoclaving.
3. Glycerol, molecular biology grade. Aliquot in glass bottles and sterilize by autoclaving.
4. L(+)-Arabinose, ≥99 % quality grade. Dissolve 10 g in 50 ml of aH₂O and sterilize by filtration through 0.22 µm cellulose acetate membrane. Store at 4 °C.
5. Dissolve ampicillin (Amp) in aH₂O to 50 g/l final concentration and sterilize by filtration through a 0.45 µm cellulose acetate membrane. Store at -20 °C.

2.4 Specific Equipment and Plasticware

1. 50 ml centrifuge tubes.
2. Black polystyrene 96-well microplates.
3. To provide proper aeration to bacterial cultures, a roller drum unit for glass tubes and a rotary flask shaker are needed.
4. Bench centrifuge accommodating 50 ml centrifuge tubes.
5. Fluorescence microplate reader and/or fluorescence imaging system.

3 Methods

3.1 Outline of the Assay Cascade

The screening cascade is outlined in Fig. 1. In **step 1**, the compounds of interest are assayed for their effect on the fluorescence of the permeable strain AS19/pGM991, which produces a leadered eGFP mRNA (Fig. 2), and fluorescence inhibitors are selected. All compounds that interfere with fluorescence expression, irrespective of their molecular target(s), and generically toxic molecules will be positively selected at this stage. However, since molecules that affect transcription and translation seem to be preferentially picked by the assay [14], they may be enriched among compounds passing **step 1**. The panel of inhibitors identified in **Step 1** is then assayed with the permeable strain AS19/pGM999, which expresses the leaderless eGFP mRNA variant (Fig. 2). It should be noted that the difference between the bacterial strains used in these stages of the screen cascade is limited to the translation initiation region of the reporter eGFP mRNAs. Thus, the vast majority of the molecules selected in **step 1** will inhibit fluorescence also in **step 2**. Conversely, inhibitors of the S1-dependent step of translation initiation (Canonical initiation inhibitors, CIIs) should not affect the synthesis of the CI-eGFP reporter protein and, in turn, strain fluorescence (Table 1). Thus, **step 2** confers high mechanism specificity on the screen. Lastly, in **step 3** the compounds of interest are assayed for their effect on the fluorescence of strain DH10B/pGM991 to assess their penetration in *E. coli* cells endowed with normal permeability.

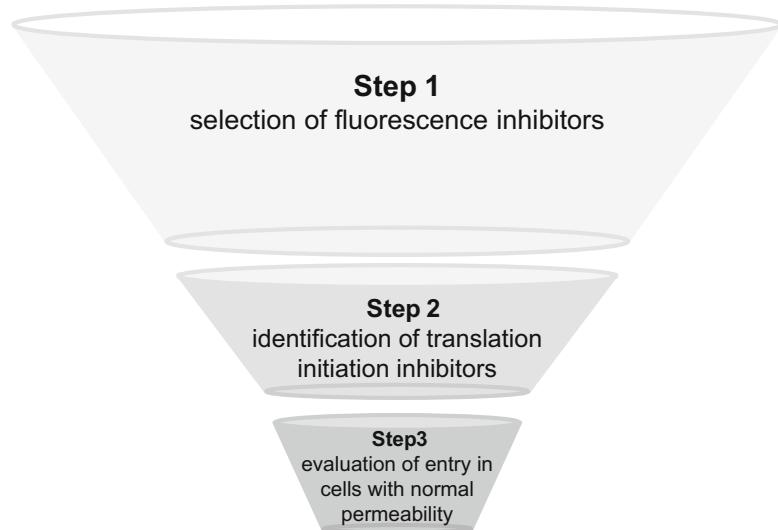


Fig. 1 Outline of the multi-step assay. In *Step 1*, the compounds of interest are screened with the permeable strain AS19/pGM991. Compounds inhibiting the fluorescence of this strain are then screened with strain AS19/pGM999 (*Step 2*). Molecules specifically targeting the canonical pathway of translation initiation should not affect AS19/pGM999 fluorescence. In *Step 3*, the entry of candidate compounds into cells with normal permeability is evaluated by assaying inhibition of DH10B/pGM991 fluorescence

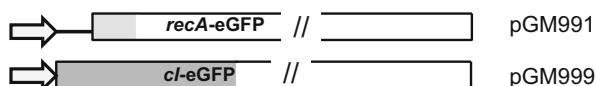


Fig. 2 Structure of reporter constructs. The structure of the cassettes expressing either the leadered (pGM991) or the leaderless (pGM999) mRNA is shown. Arrows, *araBp* promoter; line, *recA* 5'-UTR; boxes, chimeric *recA-eGFP* and *cl-eGFP* ORFs. The shadowed parts represent the portions of the two constructs encompassing *recA* and *cl* open reading frames

Table 1
Expected performance of translation initiation inhibitors in the assay steps

Step	Strain	eGFP mRNA	CII effect on fluorescence ^a
1	AS19/ pGM991	Leadered	<
2	AS19/pGM999	Leaderless	None
3	DH10B/pGM991	Leadered	<

^aCII canonical initiation inhibitors, < fluorescence decrease

3.2 Fluorescent Assay Protocol

1. Streak AS19/pGM991 from the glycerol stock stored at -80 °C on LD-Agar plates with 100 mg/l Amp. Incubate 16–20 h at 37 °C.
2. Transfer bacteria from the streak plate to a glass tube containing 5 ml of LD supplemented with 100 mg/l Amp by touching 3–4 colonies with a sterile loop and inoculating them into the broth. Incubate at 37 °C with aeration in a rotatory device for 16–20 h (*see Note 1*).
3. Dilute the stationary cultures to $OD_{600} = 0.1\text{--}0.12$ in 50 ml of M9 supplemented with 100 mg/l Amp and 0.4% glucose in a 250 ml flask. Incubate the flask at 37 °C with agitation at 120 rpm for about 3 h (*see Note 2*).
4. Pour the culture in a 50 ml centrifuge tube and harvest cells at $1091 \times g$ for 15 min at room temperature in a bench centrifuge. After carefully discarding the supernatant, resuspend the bacterial pellet in M9 to final OD_{600} between 0.8 and 1.
5. Dispense 110 µl aliquots of the cell resuspension into the wells of a black polystyrene 96-well microplate. Add 3 µl of DMSO in the first column wells (control samples) and the compounds to be tested diluted in DMSO in the other wells of the screening microplate. Incubate the microplate at 37 °C 15 min.
6. Add 10 µl of M9 supplemented with either 2.4% glycerol in four control wells (not induced samples) or 2.4% arabinose, which induces transcription of the reporter gene, in all the others. Incubate the plate 3 h at 37 °C in the dark with slow agitation (*see Note 3*).
7. After incubation, measure the fluorescence by means of a fluorescence microplate reader (*see Note 4*) and/or acquire the microplate image with a fluorescence imager (*see Note 5*). An example of a screening microplate image is shown in Fig. 3.
8. Compounds that affect the sample fluorescence (Table 1; *see Note 6*) enter **step 2**.
9. Screen the compounds selected in **step 1** with strain AS19/pGM999 according to the experimental procedure described above (from **points 1–7** of Subheading 3.2; *see Note 7*).
10. Identify compounds that do not affect sample fluorescence in **point 9** of Subheading 3.2 (Table 1). Such compounds, which should target translation initiation, can be then tested for their effect on the fluorescence of strain DH10B/pGM991 (once again, according to the experimental procedure described above from **points 1–7** of Subheading 3.2) to assess whether they can enter a strain with a normal permeability (Step 3 in Fig. 1; Table 1).

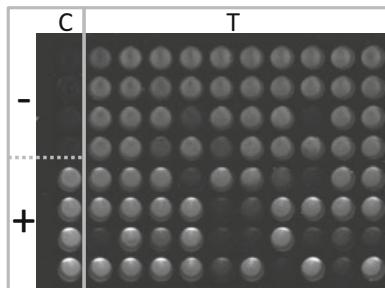


Fig. 3 Fluorescent whole cell assay: an example of a screening microplate. Cultures of AS19/pGM991 were grown and aliquoted in a 96-well microplate as described in the Subheading 3. 3 μ l of 80 heterogeneous compounds in DMSO solution were transferred into the test wells of the screening plate (T) and mixed with the cell resuspension. In the control column (C), 3 μ l of DMSO were added to the cell samples. After 3 h at 37 °C, the plate image was acquired with a VersaDoc Imaging system 4000MP. C – wells cells resuspended in M9 + 0.2 % glycerol, transcription of the *recA-eGFP* reporter off, C+ and T wells cells in M9 + 0.2 % arabinose, transcription of the reporter induced

4 Notes

1. In these conditions, cultures of AS19 and DH10B strains carrying either pGM991 or pGM999 usually reach an OD_{600} between 2.5 and 3.5 after the overnight incubation.
2. The AS19 derivatives exhibit a long lag phase and their OD_{600} less than double in this time span. Cultures exhibiting optical density decrease should be discarded as this could be due to cell lysis, an event that I have sporadically observed upon dilution in M9 of these strains.
3. I usually place the microplate on an orbital shaker set at 50 rpm in the dark. No significant growth of the cultures occurs in these conditions. This may explain why the assay seems to be more sensitive to inhibitors of gene expression than to compounds that affect the growth by interfering with other processes [14].
4. I have mainly used the Packard FluoroCount, but other microplate readers for fluorescence detection can be used. If you use a Packard FluoroCount, turn on the instrument and the Halogen Light Source (with Intensity set at 9–10) units at least 15 min before reading fluorescence. In the Control menu, click Read Settings to set the parameters for reading fluorescence. In particular, select the 485/530 nm excitation/emission wavelengths in the Filter pairs window. In the Sensitivity window, manually enter the position of one of the wells loaded with positive controls (i.e., induced samples, no

added compound) in the “Max Rfu well” empty space. I usually do not change other default settings of the instrument (Gain 1, PMT 1100, and Read Length 0.5).

5. When using a VersaDoc imaging system, open Quantity One and select VersaDoc in the Select Scanner window. Under Select Application, click on Fluorophore and then select Alexa 488 (530BP Blue Led). Put on the lower tray of the imager an empty microplate and click on Focus. The image of the microplate should appear on screen. Adjust focus by rotating the focus ring on the objective. Click Stop once focus has been adjusted. Remove the empty microplate and replace it with the screening microplate and close the imager door. Enter Exposure time (usually 500–600 s) and click Acquire.
6. I consider as inhibitors compounds that reduce fluorescence to the 30% or less of that of the positive controls (induced samples without any added compound).
7. Induced control cultures of AS19/pGM999 show a three- to fourfold reduction in the fluorescence with respect to AS19/pGM991 ones [14].

Acknowledgments

I thank lab members who were involved in the assay setup and in the screening of the Prestwick library, and in particular Matteo Raneri. The development of the assay was supported by the Italian Cystic Fibrosis Research Foundation (grant FFC#8/2013 sponsored by the FFC Delegation of Montebelluna “La bottega delle donne”).

References

1. Myasnikov AG, Simonetti A, Marzi S, Klaholz BP (2009) Structure-function insights into prokaryotic and eukaryotic translation initiation. *Curr Opin Struc Biol* 19(3):300–309. doi:[10.1016/j.sbi.2009.04.010](https://doi.org/10.1016/j.sbi.2009.04.010)
2. Benelli D, Londei P (2009) Begin at the beginning: evolution of translational initiation. *Res Microbiol* 160(7):493–501. doi:[10.1016/j.resmic.2009.06.003](https://doi.org/10.1016/j.resmic.2009.06.003)
3. Subramanian AR (1983) Structure and functions of ribosomal protein S1. *Prog Nucleic Acid Res Mol Biol* 28:101–142
4. Salah P, Bisaglia M, Aliprandi P, Uzan M, Sizun C, Bontems F (2009) Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res* 37(16):5578–5588. doi:[10.1093/nar/gkp547](https://doi.org/10.1093/nar/gkp547)
5. Roberts MW, Rabinowitz JC (1989) The effect of *Escherichia coli* ribosomal protein S1 on the translational specificity of bacterial ribosomes. *J Biol Chem* 264(4):2228–2235
6. Sørensen MA, Fricke J, Pedersen S (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli* *in vivo*. *J Mol Biol* 280(4):561–569. doi:[10.1006/jmbi.1998.1909](https://doi.org/10.1006/jmbi.1998.1909)
7. Tedin K, Resch A, Bläsi U (1997) Requirements for ribosomal protein S1 for translation initiation of mRNAs with and without a 5' leader sequence. *Mol Microbiol* 25(1):189–199. doi:[10.1046/j.1365-2958.1997.4421810.x](https://doi.org/10.1046/j.1365-2958.1997.4421810.x)

8. Moll I, Grill S, Gualerzi CO, Bläsi U (2002) Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol Microbiol* 43(1):239–246. doi:[10.1046/j.1365-2958.2002.02739.x](https://doi.org/10.1046/j.1365-2958.2002.02739.x)
9. Delvillani F, Papiani G, Dehò G, Briani F (2011) S1 ribosomal protein and the interplay between translation and mRNA decay. *Nucl Acids Res* 39(17):7702–7715. doi:[10.1093/nar/gkr417](https://doi.org/10.1093/nar/gkr417)
10. Moll I, Hirokawa G, Kiel MC, Kaji A, Bläsi U (2004) Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. *Nucleic Acids Res* 32(11):3354–3363. doi:[10.1093/nar/gkh663](https://doi.org/10.1093/nar/gkh663)
11. Udagawa T, Shimizu Y, Ueda T (2004) Evidence for the translation initiation of leaderless mRNAs by the intact 70 S ribosome without its dissociation into subunits in eubacteria. *J Biol Chem* 279(10):8539–8546. doi:[10.1074/jbc.M308784200](https://doi.org/10.1074/jbc.M308784200)
12. Briani F, Curti S, Rossi F, Carzaniga T, Mauri P, Dehò G (2008) Polynucleotide phosphorylase hinders mRNA degradation upon ribosomal protein S1 overexpression in *Escherichia coli*. *RNA* 14(11):2417–2429. doi:[10.1261/rna.1123908](https://doi.org/10.1261/rna.1123908)
13. Kitakawa M, Isono K (1982) An amber mutation in the gene *rpsA* for ribosomal protein S1 in *Escherichia coli*. *Mol Gen Genet* 185(3):445–447
14. Raneri M, Sciandrone B, Briani F (2015) A whole-cell assay for specific inhibitors of translation initiation in bacteria. *J Biomol Screen* 20(5):627–633. doi:[10.1177/1087057114566376](https://doi.org/10.1177/1087057114566376)
15. Schluelzen F, Takemoto C, Wilson DN, Kaminiishi T, Harms JM, Hanawa-Suetsugu K, Szaflarski W, Kawazoe M, Shirouzu M, Nierhaus KH, Yokoyama S, Fucini P (2006) The antibiotic kasugamycin mimics mRNA nucleotides to destabilize tRNA binding and inhibit canonical translation initiation. *Nat Struct Mol Biol* 13(10):871–878. doi:[10.1038/nsmb1145](https://doi.org/10.1038/nsmb1145)
16. Kaberdina AC, Szaflarski W, Nierhaus KH, Moll I (2009) An unexpected type of ribosomes induced by kasugamycin: a look into ancestral times of protein synthesis? *Mol Cell* 33(2):227–236. doi:[10.1016/j.molcel.2008.12.014](https://doi.org/10.1016/j.molcel.2008.12.014)
17. Sekiguchi M, Iida S (1967) Mutants of *Escherichia coli* permeable to actinomycin. *Proc Natl Acad Sci U S A* 58(6):2315–2320
18. Grant SG, Jesse J, Bloom FR, Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A* 87(12):4645–4649
19. Delvillani F, Sciandrone B, Peano C, Petiti L, Berens C, Georgi C, Ferrara S, Bertoni G, Pasini ME, Dehò G, Briani F (2014) Tet-Trap, a genetic approach to the identification of bacterial RNA thermometers: application to *Pseudomonas aeruginosa*. *RNA* 20(12):1963–1976. doi:[10.1261/rna.044354.114](https://doi.org/10.1261/rna.044354.114)
20. Schleif R (2000) Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet* 16(12):559–565. doi:[10.1016/S0168-9525\(00\)02153-3](https://doi.org/10.1016/S0168-9525(00)02153-3)
21. Bertani G (1951) Studies on lysogenesis. I The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* 62:293–300

Chapter 15

Bacterial Histidine Kinases: Overexpression, Purification, and Inhibitor Screen

Mike Gajdiss, Michael Türck, and Gabriele Bierbaum

Abstract

Bacterial histidine kinases are promising targets for new antimicrobial agents. In antibacterial therapy such agents could inhibit bacterial growth by targeting essential two-component regulatory systems or resensitize bacteria to known antibiotics by blocking stress responses like the cell wall stress response. However, (1) activity assays using the truncated phosphorylation domains have been shown to produce artifacts and (2) the purification of the full-length histidine kinases is complicated. Here, we describe a standard protocol for the recombinant expression and purification of functional full-length histidine kinases and other membrane proteins from gram-positive bacteria that do not harbor more than two trans-membrane domains using an *Escherichia coli* host. This guide also presents in vitro phosphorylation assays to screen for new antimicrobial compounds that target bacterial histidine kinases using radioactively labeled ATP and, as a novel approach, Phos-tag acrylamide gel electrophoresis to detect phosphorylated proteins by mobility shift in the polyacrylamide gel.

Key words Two-component regulatory system, Histidine kinase, Purification, Full-length, Kinase inhibitor, Antimicrobial compound, Phosphorylation, Phos-tag

1 Introduction

Excellent basic research during the last years led to a wider understanding of the complex regulation mechanisms of bacterial lifecycles that require two-component regulatory systems [1] and several studies proposed that essential two-component systems might be novel targets for antimicrobial compounds [2–4]. Furthermore, even the inhibition of a nonessential TCS can reduce the resistance of a pathogen by impairing its ability to react to antimicrobial stress, thereby enhancing the impact of established antibiotics that show insufficient efficacy under normal circumstances. For example, methicillin resistance in *Staphylococcus aureus*, which is still one of the most frequent nosocomial agents, is abolished after inactivation of the two-component system VraSR [5]. In conclusion, if all or several histidine kinases in a bacterium can be inhibited, this will offer new treatment options in combination with existing antibiotics.

Many histidine kinase inhibitors have been described to inhibit autophosphorylation after testing of the isolated cytoplasmic domains. However, false-positive test results have been obtained for several compounds [6, 7]. These assays had been performed using the cytoplasmic domains of the histidine kinases because it is easier to purify the truncated kinases and to perform phosphorylation assays without the need for detergents for stabilization of the membrane domains. However, because the decreasing intensity of the phosphorylated bands had been evaluated as signal for kinase inhibition, the fact that protein aggregation by the test compounds led to the loss of signal in the gel remained unrevealed. Therefore, detergent has always to be added if a true kinase inhibitor is to be discovered [6, 7] and use of the full-length enzyme might even be a better approach. The full-length histidine kinase is packed in detergent micelles and is able to yield a specific inhibition signal. The detergent molecules only arrange themselves in the vicinity of the hydrophobic parts of the protein, especially at the trans-membrane domains, and neither affect the ATP-binding region nor interfere with the binding of a specific compound. In addition, any other domains of the protein that could be putative target sites for potential inhibitors are present in the full length protein.

The following guide describes a standard protocol for the recombinant expression and purification of functional full-length histidine kinases and other membrane proteins from gram-positive bacteria that do not harbor more than two trans-membrane domains, using a detergent and ultracentrifugation to separate membrane-bound proteins from cytosolic proteins. The kinase can be used in phosphorylation assays to perform signaling studies or to test compounds that might be able to inhibit its function.

Phosphorylation assays can be performed using labeled ATP molecules that contain the isotope ^{32}P or ^{33}P at the γ -position. In a functional histidine kinase dimer, ATP binds to a specific part of the HATPase_c domain and a histidine residue in the HisKA domain is phosphorylated usually *in trans* with the radioactively labeled phosphate group. The phosphorylation activity is detected by running the sample in a sodium dodecyl sulfate polyacrylamide gelelectrophoresis (SDS-PAGE) and exposing the gel to autoradiography with a conventional X-ray film or a more sensitive phosphor imaging plate. Phosphotransfer to a response regulator can also be observed after adding the matching response regulator to the same reaction. This way, compounds that inhibit the phosphotransfer reaction can be detected.

Phosphorylation of histidine kinases can also be detected using Phos-tag acrylamide as an additive to SDS-PAGE gels [8]. This compound complexes phosphorylated proteins using manganese ions, leading to a mobility shift of phosphorylated proteins. The bands are detected with conventional staining methods and the phosphorylated amount of the protein can be directly compared to the nonphosphorylated amount of protein (Fig. 1).

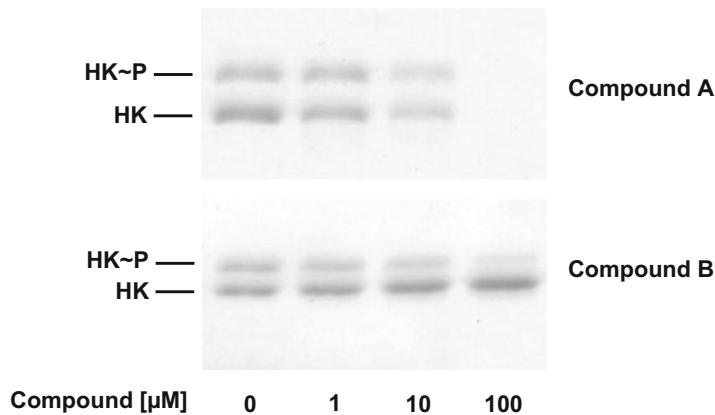


Fig. 1 Phos-tag PAGE of the truncated cytoplasmic variant of a histidine kinase with increasing concentrations of a putative inhibitory compound. The assay was performed as described in Subheading 3.3 and the gel was stained with Coomassie Brilliant Blue R250. The decreasing band intensity of both the phosphorylated (HK~P) and the nonphosphorylated (HK) histidine kinase results from an aggregation of the proteins caused by compound A, preventing migration in the acrylamide gel. The ratio of phosphorylated to nonphosphorylated protein does not change significantly, which indicates that the tested compound does not cause specific inhibition of phosphorylation. To demonstrate aggregation, a truncated cytoplasmic variant of the kinase in combination with Phos-tag gels should be used, since the addition of Triton X-100 to the full-length kinase would prevent aggregation. Compound B decreases the ratio of phosphorylated to nonphosphorylated protein and allows the kinase to migrate through the gel even at high concentrations indicating an inhibition of autophosphorylation without aggregation

2 Materials

2.1 Expression and Purification of a Recombinant Full-Length Histidine Kinase in *Escherichia coli*

Prepare all solutions in deionized water unless indicated otherwise.

1. Lysogeny broth (LB): 10 g/l tryptone, 5 g/l yeast, 10 g/l NaCl, pH 7.5. Autoclave (see Note 1).
2. 1 M Isopropyl- β -D-thiogalactoside (IPTG): Sterilize by filtration before adding to the culture.
3. 2 M Imidazole: Store at 4 °C.
4. 250 mM DDM (n-Dodecyl β -D-maltoside): Store at 4 °C (see Note 2).
5. Lysis buffer 1: 50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole, 2 mM β -mercaptoethanol, 30% (v/v) glycerol, pH 8. Dissolve NaH₂PO₄ and NaCl in 30% glycerol and add the appropriate amount of imidazole solution. Cool down to 4 °C, add β -mercaptoethanol prior to use and adjust to pH 8. Sterilize by filtration (see Note 3).

6. Lysis buffer 2: 50 mM NaH₂PO₄, 300 mM NaCl, 20 mM imidazole, 40 mM DDM, 2 mM β-mercaptoethanol, 30% (v/v) glycerol, pH 8. Dissolve NaH₂PO₄ and NaCl in 30% glycerol and add the appropriate amount of the imidazole and DDM solutions. Cool down to 4 °C, add β-mercaptoethanol prior to use, and adjust to pH 8. Sterilize by filtration.
7. Wash buffer 1: 50 mM NaH₂PO₄, 300 mM NaCl, 20 mM imidazole, 4 mM DDM, 2 mM β-mercaptoethanol, 30% (v/v) glycerol, pH 8. Prepare like lysis buffer 2.
8. Wash buffer 2: 50 mM NaH₂PO₄, 300 mM NaCl, 40 mM imidazole, 4 mM DDM, 2 mM β-mercaptoethanol, 30% (v/v) glycerol, pH 8. Prepare like lysis buffer 2.
9. Elution buffer: 50 mM NaH₂PO₄, 300 mM NaCl, 300 mM imidazole, 4 mM DDM, 2 mM β-mercaptoethanol, 30% (v/v) glycerol, pH 8. Prepare like lysis buffer 2.
10. Lysozyme: 100 mg/ml. Store at -20 °C.
11. Benzonase (*see Note 4*).
12. Ni-NTA affinity resin.
13. Polypropylene column (1 ml).
14. Dialysis buffer: 50 mM HEPES (N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid), 200 mM KCl, 50% (v/v) glycerol, pH 8. Autoclave.
15. Reagent using Bradford's method to determine the protein concentration.

2.2 Phosphorylation

Assay Using Radioactively Labeled ATP

1. 85 mM Triton X-100.
2. 1 M Dithiothreitol (DTT): Create small aliquots and store at -20 °C.
3. 5× Phosphorylation buffer: 250 mM HEPES, 2.5 M KCl, 25 mM MgCl₂, 2.5 mM DTT, 17.5% (v/v) glycerol, pH 8. Add DTT solution to an aliquot of buffer prior to use and discard the buffer afterward. Sterilize by filtration (*see Note 5*).
4. [γ -³³P]Adenosine 5'-triphosphate (ATP), Specific activity: >111TBq (3000 Ci)/mmol.
5. 2× Laemmli sample buffer: 65.8 mM Tris-HCl (pH 6.8), 26.3% (w/v) glycerol, 2.1% SDS, 0.01% bromphenol blue, 5% (v/v) β-mercaptoethanol. Add β-mercaptoethanol prior to use.
6. Quality precast 10% Bis-Tris gels, 1.0 mm, 10–15 well (*see Note 6*).
7. 20× MOPS SDS running buffer: 1 M MOPS, 1 M Tris, 2% SDS, 20 mM EDTA.
8. Prestained protein ladder.

9. Storage Phosphor Screen and exposure cassette or conventional X-ray films.
10. Phosphor imager.
11. Coomassie staining solution: 45% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R250. Mix vigorously after adding Coomassie Brilliant Blue and before every staining procedure.
12. Destaining Solution: 45% (v/v) methanol, 10% (v/v) acetic acid.

2.3 Phosphorylation Assay Using Phos-tag Acrylamide

1. 40% Acrylamide/bisacrylamide (37.5:1).
2. Resolving gel solution: 3 M Tris-HCl. Dissolve Tris in ultrapure water and adjust pH with 37% HCl to 8.5. Store at 4 °C.
3. 10 mM MnCl₂: Dissolve in ultrapure water. Store at 4 °C.
4. 5 mM Phos-tag acrylamide (Wako Pure Chemical Industries) in 3% methanol: Dissolve in methanol first by vortexing vigorously. Add ultrapure water, vortex and store at 4 °C (*see Note 7*).
5. 21 mg/ml Ammonium persulfate: Create solution in ultrapure water prior to use.
6. N,N,N',N'-Tetramethylethylenediamin (TEMED).
7. 2-Propanol.
8. Stacking gel solution: 1 M TRIS, 0.8% SDS, pH 6.8. Store at 4 °C.
9. 10 mM Adenosine 5'-triphosphate (ATP): Create a solution with ultrapure water.
10. 20% SDS. Store at room temperature (*see Note 8*).
11. SDS-PAGE running buffer: 25 mM TRIS, 192 mM glycine, 0.1% (w/v) SDS. Create a 10x stock solution without SDS and store at 4 °C. Add 0.1% SDS to the diluted working solution.
12. Coomassie staining solution: 45% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R250. Mix vigorously after adding Coomassie (*see Note 9*).
13. Destaining solution: 45% (v/v) methanol, 10% (v/v) acetic acid.

3 Methods

3.1 Recombinant Expression and Purification of a Full-Length Histidine Kinase in *Escherichia coli* (See Note 10)

1. In a 5 l Erlenmeyer flask, inoculate 1 l of LB containing the appropriate antibiotic(s) with 10 ml of an overnight culture of the expression strain and incubate in a water bath at 37 °C and constant shaking at 100 rpm. Observe the growth by measuring the absorbance at 600 nm (OD₆₀₀) with a spectrophotometer (*see Note 11*).

2. When the culture reaches an OD₆₀₀ of 0.5, induce the expression of the protein by adding 1 ml of a 1 M IPTG solution (final concentration: 1 mM).
3. Adjust the temperature in your water bath to 30 °C and incubate the culture for another 16–20 h.
4. Cool down the centrifuge to 4 °C.
5. Harvest the cells by centrifuging the culture at 7000×*g* for 10 min. Decant the supernatant.

Perform all the following steps with constant cooling on ice or in the cold room.
6. Resuspend the pelleted cells in 20 ml lysis buffer 1 and transfer the solution to a sterile 50 ml conical centrifuge tube. The volume should now be around 25 ml. At this point, the lysate can be stored at -20 °C if desired (*see Note 12*).
7. Add 250 µl of a 100 mg/ml lysozyme solution (final concentration: 200 µg/ml) and vortex.
8. Add Benzonase to a final concentration of 25 U/ml and vortex.
9. Incubate the lysate for 30 min on ice.
10. Lyse the cells by ultrasonication avoiding heating of the lysate. Keep the lysate on ice or turn on the cooling module if existent. 10–15 pulses of 10 s with 30 s of cooling breaks between the pulses are a basic protocol that needs to be adjusted to the sonicator available. Stop the sonication immediately when the lysate starts to foam.
11. You can add more Benzonase at this point, if the lysate is too viscous. Then an additional incubation for 30 min is necessary.
12. Centrifuge the lysate at 15,000×*g* for 10 min at 4 °C. Set the brake of the centrifuge to a low value to avoid detaching of the pellet from the wall of the tube.
13. Transfer the supernatant to a fresh 50 ml conical centrifuge tube and repeat the centrifugation step two times to get rid of as much of the cell debris as possible (*see Note 13*).
14. Transfer the supernatant to ultracentrifugation tubes (*see Note 14*).
15. Ultracentrifuge at 218,000×*g* and 4 °C for 60 min and decant the supernatant.
16. Pipette 4 ml lysis buffer 2 to the pellet and carefully release the pellet from the centrifuge tube wall using a clean glass rod or spatula. Transfer the released pellet by decanting it with the buffer into a fresh 50 ml conical centrifuge tube.
17. Resuspend the pellets with a small stir bar for 30–60 min (*see Note 15*).

18. Transfer the lysate to a fresh ultracentrifugation tube and centrifuge for 30 min at $218,000 \times g$. Transfer the supernatant to a fresh 50 ml conical centrifuge tube. At this point the lysate can be stored at -20°C if desired.
19. Add 1 ml Ni-NTA affinity resin to the lysate and let it stir gently for 2 h.
20. Equilibrate a polypropylene column with 1–2 ml lysis buffer 2, avoiding formation of air bubbles (*see Note 16*).
21. Load the lysate/resin mixture onto the polypropylene column containing lysis buffer 2. Open the lower cap of the column and let the buffer run out until the formation of a resin column is visible. Add the rest of the mixture and collect the flow-through (*see Note 17*).
22. Let the entire lysate run through but do not let the column run dry.
23. Wash the column with 5 ml of wash buffer 1 and then with 5 ml of wash buffer 2. Collect both fractions.
24. Elute the protein with 8 fractions of 200 μl elution buffer and collect them in small reaction tubes.
25. Immediately add 80 μl of glycerol to every elution fraction (final percentage: 50%) and freeze the samples at -20°C .
26. Perform a SDS-PAGE applying 10 μl of every fraction. For a normal SDS-PAGE you can use the protocol for casting the Phos-tag acrylamide gel described in Subheading 3.3, but without adding Phos-tag acrylamide and MnCl_2 .
27. When the dye front reaches the end of the gel, stop the electrophoresis and take the gel out of the cassette.
28. Wash the gel in deionized water for 5 min and incubate it for 30 min in the staining solution with gentle agitation.
29. Wash the gel in deionized water for 5 min and destain it for 2 h using the destaining solution. Leaving the gel overnight in deionized water with gentle agitation helps to reduce the background.
30. Choose the fractions with the highest yield of your protein for dialysis (*see Note 18*).
31. Pipette one or two eluted fractions into a dialysis cassette (Slide-A-LyzerTM Dialysis Cassettes, 10 K MWCO, 0.5 ml) using a syringe, attach the cassette to a float buoy, and incubate the cassette in 400 ml of dialysis buffer for 2 h at 4°C with constant gentle stirring.
32. Discard the dialysis buffer and repeat the dialysis step with new buffer for another 2 h.
33. Change the dialysis buffer again and incubate for 16 h.

34. Use a syringe to recover the protein from the dialysis cassette and store the dialyzed protein at -20 °C.
35. Determine the concentration of the dialyzed protein using the Bradford assay.

3.2 Phosphorylation Assay Using Radioactively Labeled ATP

The following protocol is an example for a radioactive phosphorylation assay with a sample volume of 10 µl to test an inhibiting compound at different concentrations (*see Note 19*).

1. Pipette ultrapure water, 1 µg of the purified kinase, 1 µl 85 mM Triton X-100 (final concentration: 8.5 mM), and 2 µl 5× phosphorylation buffer to each reaction tube (*see Note 20*).
2. Add 1 µl of different compound dilutions to each reaction and incubate for 10 min at room temperature (*see Note 21*).
3. Add the desired volume of ^{33}P -ATP to each reaction and incubate for 30 min at room temperature (*see Note 22*).
4. Stop the reaction with 10 µl 2× Laemmli SDS sample buffer (*see Note 23*).
5. Load the samples to a precast gel and load 2 µl of prestained protein ladder. For Bis-Tris precast gels use MOPS SDS running buffer. Let it run at a constant voltage of 180 V for approximately 1 h until the dye front reached the end of the gel.
6. Stop the electrophoresis and take the gel out of the cassette. Cut approximately 1 cm of the upper and the lower end. Use the prestained bands of the protein marker to ensure not to cut the region where you expect the protein bands (*see Note 24*).
7. Put the gel into a translucent small autoclave bag and seal it using a heat sealer.
8. Expose the gel to a storage phosphor screen or X-ray film. Phosphor imaging plates are much more sensitive, so an exposure time of 30–90 min should be sufficient. An overnight exposure is recommended for X-ray films (*see Note 25*).
9. Scan the phosphor imaging plate using a phosphor imager or develop the X-ray film.
10. After autoradiography, wash the gel in deionized water for 5 min and incubate it for 30 min in the staining solution with gentle agitation (*see Note 26*).
11. Wash the gel in deionized water for 5 min and destain it for 2 h using the destaining solution. Leaving the gel overnight in deionized water with gentle agitation helps to reduce the background.
12. Seal the gel in a translucent plastic bag and scan it to obtain a high quality digital image.
13. For a more precise analysis, use the software provided with the storage phosphor screen or any other quantification software

to identify the intensity of the bands. The phosphorylation activity can be defined as the intensity of the radioactive band per protein band. Use the control sample without the inhibitor as a reference to analyze any reduction of autophosphorylation activity (*see Note 27*).

3.3 Phosphorylation Assay Using Phos-tag Acrylamide

Casting of a resolving gel (7%) containing 50 μM Phos-tag acrylamide and 100 μM MnCl₂ with a total gel volume of 8 ml, and a stacking gel (4%) with a volume of 4 ml (*see Note 28*).

1. Mix 5.11 ml ultrapure water, 1.4 ml of 40% (w/v) acrylamide/bisacrylamide (37.5:1), 1.33 ml of the resolving gel solution, 80 μl of 10 mM MnCl₂, and 80 μl of 5 mM Phos-tag acrylamide in a 50 ml conical centrifuge tube and vortex. Adjust the volumes when you have a different resolving gel volume (*see Note 29*).
2. Ensure you have the gel casting apparatus prepared under a fume hood before continuing with the next steps (*see Note 30*).
3. Add 104 μl of 21 mg/ml APS and 5.36 μl of TEMED and vortex (*see Note 31*).
4. Pipette the mixture quickly between the glass plates, avoiding formation of bubbles.
5. Overlay the gel with 1 ml 2-propanol and let it polymerize for 30 min.
6. Decant the 2-propanol from the casting apparatus and let it evaporate for a few minutes.
7. Mix 3.12 ml ultrapure water, 0.4 ml of 40% (w/v) acrylamide/bisacrylamide (37.5:1), and 0.48 ml of the stacking gel solution in a 50 ml conical centrifuge tube and vortex.
8. Add 128 μl of 21 mg/ml APS and 3.2 μl of TEMED and vortex.
9. Pipette the mixture onto the resolving gel and fill up the gel cassette. Insert the well comb immediately and let the stacking gel polymerize for 30 min.
10. Transfer the Phos-tag gel from the casting apparatus into the electrophoresis cell. Fill the inner and the outer tank with SDS-PAGE running buffer and remove the well comb carefully from the gel.
11. Perform the phosphorylation of the kinase as described in Subheading 3.2 for the radioactive assay but use nonradioactive ATP.
12. Use 1 μl of the 10 mM ATP solution to start the phosphorylation reaction (*see Note 32*).
13. Load the samples into the wells and run the gel at a constant current of 30 mA. A Phos-tag gel runs more slowly than a normal acrylamide gel and takes about 2 h for a complete run. The voltage will increase continuously (*see Note 33*).

14. When the dye front reaches the end of the gel, stop the electrophoresis and take the gel out of the cassette.
15. Wash the gel in deionized water for 5 min and incubate it for 30 min in the staining solution with gentle agitation.
16. Wash the gel in deionized water for 5 min and destain it for 2 h using the destaining solution. Leaving the gel overnight in deionized water with gentle agitation helps to reduce the background.
17. Seal the gel in a translucent plastic bag and scan it to obtain a high quality digital image (*see Note 34*) (Fig. 1).

4 Notes

1. LB is a standard medium and works fine for most applications. However, when facing good purity but low yield of the purified protein, different media can be tested to increase the final cell weight. Try “terrific broth”: Dissolve 12 g tryptone, 24 g yeast extract, and 5 g glycerol in 900 ml deionized water. In another flask, dissolve 2.31 g KH_2PO_4 (170 mM) and 12.54 g K_2HPO_4 (720 mM) in 100 ml deionized water. Autoclave separately and after cooling mix the components prior to use. Note that a high cell density does not necessarily result in a higher protein purity.
2. Other detergents are also suitable and other groups use Triton X-100. If you are not sure, perform a small scale screening to check which detergent works best for your protein.
3. β -Mercaptoethanol decomposes within hours and has an effect on the pH, as well as temperature. Let the buffer cool down, add β -mercaptopethanol and adjust pH prior to use. There are several guides like this one that describe a purification method. You need to know the function of every component in your purification buffer and adjust it to your personal purification strategy. The method described here however worked well for all full-length histidine kinases from gram-positive bacteria we tested so far.
4. You can also use DNaseI and RNaseA (15 $\mu\text{g}/\text{ml}$ each) instead of Benzonase.
5. Similar to β -mercaptopethanol, DTT decomposes within hours. Prepare the 5 \times phosphorylation buffer without DTT and add DTT to a 1 ml aliquot prior to use. You can also create a phosphorylation buffer without KCl and add different concentrations of KCl to the reactions to check which concentration works best.
6. Precast gels are easier to handle while using radioactive material.

7. Add methanol directly to the tube containing Phos-tag acrylamide. It is sticky and there is no possibility of transferring it before resolving.
8. Without heating it will not dissolve. Use an incubator.
9. This staining method worked best for the gels we created so far. There are also other and more sensitive staining methods like silver staining or Colloidal Blue Staining that uses the more sensitive Coomassie G250 instead of R250.
10. The protein should be encoded on a vector with an inducible promoter (e.g., with IPTG). pET expression systems always worked well for our requirements and offer a broad range of different affinity tags and purification strategies [9]. A C-terminal 6xHis-tag without a leader peptide worked well for the histidine kinases we have tested so far. We also use a plasmid-encoded chaperone in some of our expression strains to prevent the formation of inclusion bodies. The so called “Walker strains” *E. coli* C41(DE3) and C43(DE3) are recommended as expression hosts [10].
11. There is no ideal protocol for expression and purification and it has to be optimized to different proteins. The expression can alter with incubation duration and temperature. Even other growth conditions like shaking frequency, flask shape, or the apparatus used for incubation can change the expression. It is recommended to search for the best expression conditions by performing a small scale screening. In our case, a big Erlenmeyer flask with a proportionate low culture volume incubated in a traditional water bath and a moderate shaking frequency gave the best results so far. Observe the growth of your culture after induction of the expression. When the culture stops growing after a short time, your protein seems to inhibit growth of the expression strain. Nevertheless, the expression can still work but formation of inclusion bodies can occur.
12. A good trick to overcome the problematic resuspension of big and sticky cell pellets is to shake two sterile glass marbles carefully in the centrifuge bottle.
13. Getting rid of as much of the cell debris as possible is crucial to increase the flow rate of the gravity-flow columns. Three centrifugation steps help in reducing the contamination with cell debris.
14. Split the lysate in several small centrifugation tubes if necessary and combine them afterward.
15. Any remaining cell debris visible as dark residue that cannot be resuspended will be pelleted in the next ultracentrifugation step.
16. Any air inclusions around the membrane inside the polypropylene column slow down the gravity-flow column.

17. Gravity-flow columns can take some time to run through. Do not use pressure or the resin gets even more compressed, which additionally slows down the flow rate.
18. Check for co-purified contaminant proteins. Sometimes fractions that are more pure rather than having the most protein are the better option to choose.
19. The sample volume is limited to the well size of the SDS-PAGE gel used and can be increased if needed. A time series measurement of the autophosphorylation activity can be performed using a bigger volume, taking small samples at different time points. To test different conditions like temperature or concentration of inhibitory substances, small scale single reactions in small reaction tubes are recommended.
20. Decide how much of the histidine kinase you want to use. 1 µg is a good amount to start with, making it easily detectable in the Coomassie stain afterward.
21. Dilute the tested compound to obtain a wide range of concentrations in the assay starting with 10 µg/ml and going up to 10 mg/ml. Because many compounds need a special solvent, use the solvent for dilution and create a control sample only using the solvent. For many hydrophobic compounds DMSO is used, which does not interfere with this assay in our experience.
22. The signal strength can be adjusted by altering the assay conditions. You can use more radioactive material but note that this will also increase the background signal. It is recommended to first screen for a suitable signal strength using different temperatures and incubation times. We have encountered different properties of each histidine kinase and how they react to assay conditions and buffer components like alkali salts. Follow the manufacturer's instructions to calculate the volume of labeled ATP needed and create a master mix with ultrapure water to simplify the pipetting. Try 1 µCi of ^{33}P -ATP for 1 µg of the kinase and adjust the concentration when the signal appears too low or too high. For very low activity ^{32}P -ATP is an option.
23. We experienced that using 2× Laemmli SDS sample buffer with 1:20 β-mercaptoethanol works best to denature the histidine kinase to an extent that lets it run mostly as a monomer in the polyacrylamide gel. Often a weak signal of the kinase dimer remains visible. Do not boil the samples since this can lead to dephosphorylation of the kinase.
24. Cutting of the gel is recommended to reduce the background signal that is present in the wells and the lower end of the gel, where the unbound ATP remains.
25. Increase the exposure time using the same gel when the signal appears to be too weak but note that the background will also increase.

26. It is important to analyze the gel using a conventional acrylamide gel staining like Coomassie Brilliant Blue R250 to ensure that the amount of the kinase does not differ in any of the samples.
27. Here, ImageQuant TL software was used. GelAnalyzer 2010a is an alternative software free of charge.
28. Cast the gels prior to use. According to the manufacturer's instructions Phos-tag acrylamide decomposes within hours.
29. Assemble your gel casting apparatus and check for the exact volume of the resolving gel using water. This will save some of the expensive Phos-tag acrylamide.
30. Using adhesive tape at the lower end of the glass plates helps to get it sealed properly.
31. The concentration of APS can be increased when facing problems with polymerization. Use a fume hood when pipetting TEMED.
32. A screening with different ATP concentrations is recommended to optimize the assay for each histidine kinase. A high ATP concentration will let the kinase work at full capacity.
33. Use a power supply that is able to run a constant current independently of the voltage. Nevertheless, we always limit the voltage to a maximum of 160 V to avoid heating of the gel.
34. When the separation of the bands is not satisfying, increase the concentration of Phos-tag acrylamide or/and decrease the total acrylamide/bisacrylamide concentration in the resolving gel. For very low density gels the addition of agarose is required.

References

1. Fritz G, Mascher T (2014) A balancing act times two: sensing and regulating cell envelope homeostasis in *Bacillus subtilis*. Mol Micro 94:1201–1207
2. Matsushita M, Janda KD (2002) Histidine kinases as targets for new antimicrobial agents. Bioorg Med Chem 10(4):855–867
3. Schreiber M, Res I, Matter A (2009) Protein kinases as antibacterial targets. Curr Opin Cell Biol 21(2):325–330
4. Gotoh Y, Eguchi Y, Watanabe T, Okamoto S, Doi A, Utsumi R (2010) Two-component signal transduction as potential drug targets in pathogenic bacteria. Curr Opin Microbiol 13(2):232–239
5. Boyle-Vavra S, Yin S, Daum RS (2006) The VraS/VraR two-component regulatory system required for oxacillin resistance in community-acquired methicillin-resistant *Staphylococcus aureus*. FEMS Microbiol Lett 262(2):163–171
6. Francis S, Wilke KE, Brown DE, Carlson EE (2013) Mechanistic insight into inhibition of two-component system signaling. Med Chem Commun 4(1):269–277
7. Stephenson K, Yamaguchi Y, Hoch JA (2000) The mechanism of action of inhibitors of bacterial two-component signal transduction systems. J Biol Chem 275(49):38900–38904
8. Kinoshita E, Kinoshita-Kikuta E, Takiyama K, Koike T (2006) Phosphate-binding tag, a new tool to visualize phosphorylated proteins. Mol Cell Proteomics 5(4):749–757
9. Türck M, Bierbaum G (2012) Purification and activity testing of the full-length YycFGHI proteins of *Staphylococcus aureus*. PLoS One 7(1), e30403
10. Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Bio 260(3):289–298

Part III

Response and Susceptibility

Chapter 16

Expression Profiling of Antibiotic-Resistant Bacteria Obtained by Laboratory Evolution

Shingo Suzuki, Takaaki Horinouchi, and Chikara Furusawa

Abstract

To elucidate the mechanisms of antibiotic resistance, integrating phenotypic and genotypic features in resistant strains is important. Here, we describe the expression profiling of antibiotic-resistant *Escherichia coli* strains obtained by laboratory evolution, and a method for extracting a small number of genes whose expression changes can contribute to the acquisition of resistance.

Key words Antibiotic resistance, Laboratory evolution, Transcriptome analysis, *Escherichia coli*

1 Introduction

Laboratory evolution of bacteria is a powerful tool for analyzing the response to antibiotic drug treatment and the acquisition dynamics of resistance [1–5]. In this approach, bacterial cells are exposed to antibiotic drug concentrations around which the cell growth is partially or completely inhibited such that a selective advantage for resistant strains is maintained. Although some essential factors in the evolution of drug resistance, such as horizontal gene transfer (HGT) [6] and interspecies communication [7], are difficult to investigate by laboratory evolution, this experimental system has several advantages in comparison with in vivo experiments for studying *de novo* acquisition of drug resistance, including a well-characterized ancestor strain, a defined environment, and parallel evolution experiments that discriminate necessary and unnecessary phenotypic or genetic changes.

Genome-wide phenotype-genotype analysis of resistant strains emerging through laboratory evolution provides clarity to the relationship between phenotypic and genotypic changes, and drug resistance. Whole-genome resequencing analysis of resistant strains obtained through selection with a single antibiotic drug can identify various mutations, including both mutations specific to

resistance to a particular antibiotic and mutations shared in resistance to more than one drug [3, 4]. Laboratory evolution under antibiotic selection pressure enables systematic screenings for cross-resistance and collateral sensitivity among antibiotics, which identifies physiological mechanisms and contributing mutations responsible for drug-drug interactions [5, 8, 9]. Also, integration of transcriptome and whole-genome resequencing analysis reveals that phenotype-genotype mapping is complex and includes various mutations that cause similar phenotypic changes [5].

In this chapter, we describe detailed methods for phenotypic analyses of antibiotic-resistant *Escherichia coli* strains for clarification of resistance mechanisms [5]. First, resistant strains are obtained by laboratory evolution under antibiotic selection pressure. For each strain resistant to a single antibiotic, resistance for other antibiotics is measured, to explore how the resistance acquisition to one drug changes the resistance and susceptibility to other drugs. Furthermore, changes in gene expression profile are quantified by using microarray analysis (Agilent microarray platform). By integrating the phenotype data of resistant strains, it can be demonstrated that resistance can be quantitatively predicted by the expression changes of a small number of genes. These analyses enable clarification of phenotypic changes contributing resistance acquisition and provide clues as to how genotypic changes cause antibiotic resistance.

2 Materials

Prepare all of the solutions for incubation steps using ultrapure water (prepared by purifying deionized water to attain a sensitivity of $18\text{ M}\Omega\text{ cm}$ at 25°C) and analytical grade reagents. Prepare all solutions for the molecular work using DNase/RNase-free water.

2.1 Experimental Evolution of Antibiotic Resistance and Quantification of Cross-Resistance and Collateral Sensitivity

1. Modified M9 medium: 47.7 mM Na_2HPO_4 , 22.0 mM KH_2PO_4 , 8.56 mM NaCl, 37.4 mM NH_4Cl , 0.5 mM MgSO_4 , 0.01 mM FeSO_4 , 0.1 mM CaCl_2 , 0.03 mM thiamine-hydrochloride, and 55.5 mM glucose, pH 7.0. Add 800 ml water to a beaker. Weigh 17.1 g $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 3 g KH_2PO_4 , 0.5 g NaCl, 2 g NH_4Cl , 123 mg $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2.78 mg FeSO_4 , 14.7 mg $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 10 mg thiamine-hydrochloride, and 5 g glucose and transfer to the beaker. Mix and adjust pH with H_3PO_4 . Add water to up to 1000 ml. Filter-sterilize with a bottle top filter with pore size 0.2 μm . Cover the bottle with aluminum foil for shading and store at 4°C .
2. Antibiotic stock solutions: Analytical grade antibiotics are used. Antibiotic stock solutions are made by dissolving powder stocks in specified solvents (Table 1). All antibiotic stocks

Table 1**List of antibiotics used for experimental evolution of antibiotic resistance**

Antibiotics name	Abbreviation	Class	Cellular target	Solvent	Concentration of stock (mg/ml)
Cefotaxime	CTX	Cephalosporin	Cell wall	Water	100
Cefoperazone	CPZ	Cephalosporin	Cell wall	Water	50
Ceftazidime	CAZ	Cephalosporin	Cell wall	Water	50
Cephalexin	CEX	Cephalosporin	Cell wall	Water	10
Cefixime	CFIX	Cephalosporin	Cell wall	EtOH	10
Streptomycin	SM	Aminoglycoside	Protein synthesis, 30S	Water	10
Kanamycin	KM	Aminoglycoside	Protein synthesis, 30S	Water	100
Amikacin	AMK	Aminoglycoside	Protein synthesis, 30S	Water	100
Gentamicin	GM	Aminoglycoside	Protein synthesis, 30S	Water	50
Neomycin	NM	Aminoglycoside	Protein synthesis, 30S	Water	100
Tetracycline	TC	Tetracycline	Protein synthesis, 30S	Water	20
Doxycycline	DOXY	Tetracycline	Protein synthesis, 30S	Water	100
Minocycline	MINO	Tetracycline	Protein synthesis, 30S	Water	10
Chloramphenicol	CP		Protein synthesis, 50S	EtOH	30
Azithromycin	AZM	Macrolide	Protein synthesis, 50S	EtOH	100
Trimethoprim	TP		Folic acid synthesis	DMSO	50
Rifampicin	RFP		RNA polymerase	DMSO	100
Nalidixic acid	NA	Quinolone	DNA gyrase	DMF	10
Norfloxacin	NFLX	Quinolone	DNA gyrase	Acetic acid	100
Oflloxacin	OFLX	Quinolone	DNA gyrase	0.5 M NaOH	100
Levofloxacin	LVFX	Quinolone	DNA gyrase	Acetic acid	2

(continued)

Table 1
(continued)

Antibiotics name	Abbreviation	Class	Cellular target	Solvent	Concentration of stock (mg/ml)
Enoxacin	ENX	Quinolone	DNA gyrase	0.5 M NaOH	100
Ciprofloxacin	CPFX	Quinolone	DNA gyrase	Water	20
Lomefloxacin	LFLX	Quinolone	DNA gyrase	0.5 M NaOH	50
Gatifloxacin	GFLX	Quinolone	DNA gyrase	0.5 M NaOH	50

dissolved in water are 0.2 µm filter-sterilized. The antibiotic stock solutions are stored at -80 °C prior to use.

3. Culture plates: 96-well microplates.
4. Incubator.
5. Microplate shaker.
6. Microplate reader.

2.2 Total RNA Preparation

1. Ice-cold ethanol containing 10% (w/v) phenol: Dissolve crystalline phenol in 99.5% ethanol. Keep on ice before use.
2. RNeasy Micro Kit (Qiagen).
3. RNase-Free DNase Set (Qiagen).
4. TE buffer: 10 mM Tris-HCl, 1 mM EDTA, pH 8.0.
5. Lysozyme in TE buffer: Dissolve lysozyme at 1 mg/ml in TE buffer.
6. Buffer RLT (RNeasy Micro Kit) containing 1% (v/v) 2-mercaptoethanol: Add 2-mercaptoethanol to Buffer RLT at 1% (v/v) just before use.
7. Spectrophotometer with microarray analysis mode (e.g., NanoDrop ND-1000).
8. (Optional) Agilent 2100 Bioanalyzer (Agilent).
9. (Optional) Agilent RNA 6000 Nano Kit (Agilent).

2.3 DNA Microarray and Image Analysis

1. Agilent Low Input Quick Amp WT Labeling Kit, one-color (Agilent, P/N 5190-2943).
2. Agilent One-Color RNA Spike-In Kit (Agilent, P/N 5190-2943).
3. RNeasy mini Kit (Qiagen).

4. NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific).
5. Agilent Gene Expression Hybridization Kit (Agilent, P/N 5188-5242).
6. Agilent SurePrint G3 Custom Microarray (Platform GPL18948), 8×60 K (Agilent, P/N G4863A) (*see Note 1*).
7. Agilent Hybridization gasket slide, 8 microarrays/slide (Agilent, P/NG2534-60014).
8. Agilent Hybridization Chamber, stainless (Agilent, P/N G2534A).
9. Agilent Hybridization oven (Agilent, P/N G2545A).
10. Agilent Gene Expression Wash Buffer 1 (Agilent, P/N 5188-5325).
11. Agilent Gene Expression Wash Buffer 2 (Agilent, P/N 5188-5326).
12. Agilent Microarray Scanner (Agilent, P/N G4900DA).
13. Agilent Feature Extraction software 9.5.3.1 or later.

3 Methods

Carry out all incubation steps in sterile conditions, and all molecular work in DNase/RNase-free conditions. In Subheading 3.1, the methods for laboratory evolution of *E. coli* under antibiotics are presented. In Subheading 3.2, we explain the protocol of quantification of cross-resistance and collateral sensitivity of the resistant strains. Subheadings 3.3–3.5 show the protocols of transcriptome analysis of the resistant strains by using microarray analysis. Subheading 3.6 presents an example of the integration of transcriptome data and the quantification of cross-resistance and collateral sensitivity in Subheading 3.2, by which the change of resistance to various antibiotics can be quantitatively predicted by expression levels of a small number of genes.

3.1 Laboratory Evolution of Antibiotic Resistance

1. Prepare 100 µl modified M9 medium with eight different concentrations of antibiotics in a 96-well microplate (*see Note 2*).
2. *E. coli* MDS42 [10] strain (*see Note 3*) were pre-cultured overnight in 200 µl modified M9 medium without drugs. The serial transfer culture was started by diluting the pre-culture up to an OD_{600 nm} of 3×10^{-5} .
3. Monitor growth of the cells by measuring the OD_{600 nm} of each well of the culture plate incubated for 23 h using the microplate reader.

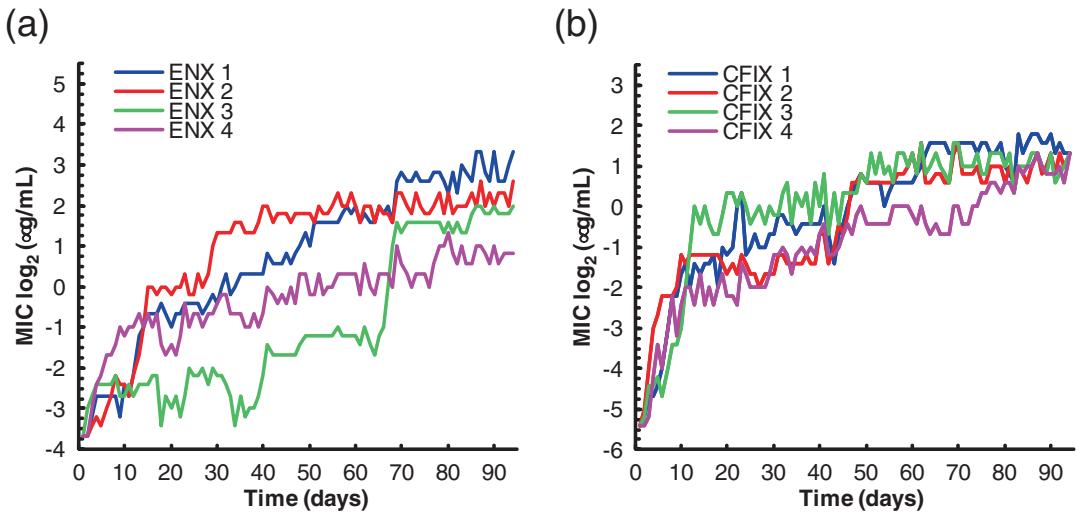


Fig. 1 Examples of laboratory evolution of antibiotic resistance. The time courses of the increase in MIC for Enoxacin (ENX) and Cefixime (CFIX) over 90 days of experimental evolution, (a, b) respectively. Day 0 corresponds to the parent strain before evolution. Four parallel series of experiments were performed. The figure is reproduced from [5] with permission

4. Dilute the cells in the well with highest antibiotic concentrations in which cells can grow with modified M9 medium to an $OD_{600\text{ nm}}$ of 3×10^{-5} (see Note 4).
5. Prepare a second 96-well microplate using 100 μl modified M9 medium per well with eight different concentrations of antibiotics (see Subheading 2.1, item 1) and inoculate 100 μl of the diluted cells to wells with the corresponding antibiotic.
6. Incubate the culture plate with shaking at 900 rpm on a microplate shaker at 34 °C for 23 h.
7. By repeating the daily propagation, a significant increase of minimum inhibitory concentration (MIC) can be observed (Fig. 1 for examples).
8. At appropriate time intervals (e.g., every 3 days), store the cells after the evolution experiments as glycerol stocks at -80 °C for further analysis.

3.2 Quantification of Cross-Resistance and Collateral Sensitivity

1. Thaw glycerol stocks of evolved and parent strains and add modified M9 medium to a volume of 210 μl .
2. After mixing the tubes, transfer 200 μl of the cells to a 96-well microplate.
3. Measure the $OD_{600\text{ nm}}$ of each strain using the microplate reader.
4. Dilute the cells with modified M9 medium to an $OD_{600\text{ nm}}$ of 1.5×10^{-5} .

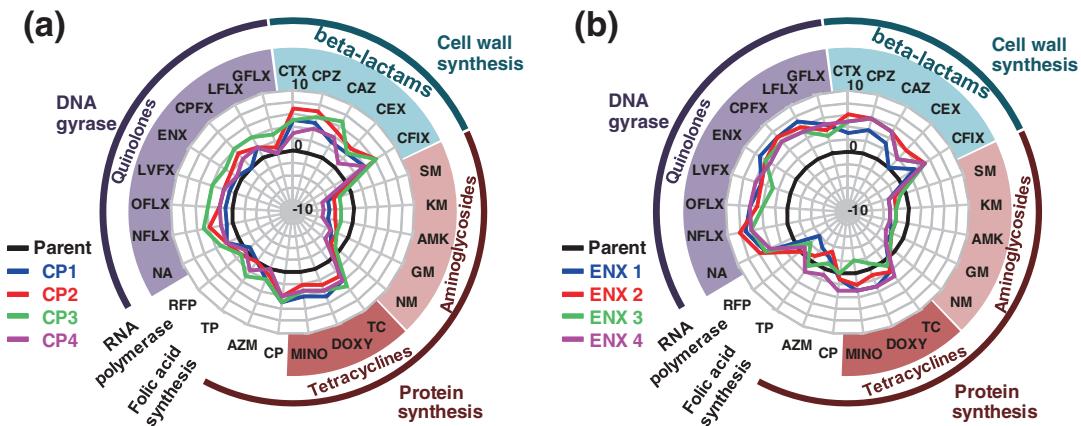


Fig. 2 Quantification of cross-resistance and collateral sensitivity. Changes in MICs for 25 antibiotics in Chloramphenicol (CP) and ENX resistant strains, (a, b) respectively. The radial axis depicts the log₂-transformed MIC relative to the parent strain. The black thick line indicates MICs of the parent strain, and the colored thick lines indicate relative MICs of four parallel-evolved resistant strains. The figure is reproduced from [5] with permission

5. Inoculate 200 μ l of the diluted cells to a 96-well microplate.
6. Incubate the culture plate with shaking at 900 rpm on a microplate shaker at 34 °C for 23 h.
7. Prepare 100 μ l modified M9 medium with 15 different concentrations of 25 antibiotics (Table 1) and without antibiotics in 96-well microplates (*see Note 5*). One microplate set for each strain to be tested.
8. Measure the OD_{600 nm} of each strain incubated for 23 h using the microplate reader.
9. Dilute the cells with modified M9 medium to an OD_{600 nm} of 3×10^{-5} .
10. Inoculate 100 μ l of the diluted cells to the prepared 96-well microplates (*see Subheading 3.2, step 7*). One microplate set for each strain to be tested.
11. Incubate the culture plates with shaking at 900 rpm on a plate shaker at 34 °C for 23 h.
12. Monitor growth of the cells by measuring the OD_{600 nm} of each well of the culture plate using the microplate reader (*see Note 6*).
13. Calculate and plot the MIC of the resistant strain relative to the parent strain (Fig. 2 for example).

3.3 Total RNA Preparation for Transcriptome Analysis

1. Thaw glycerol stocks of evolved and parent strains and add modified M9 medium to a volume of 210 μ l.
2. After mixing the tubes, transfer 200 μ l of the cells to a 96-well microplate.

3. Measure the OD_{600 nm} of each strain using the microplate reader.
4. Dilute the cells with modified M9 medium to an OD_{600 nm} of 1.5×10^{-5} .
5. Inoculate 200 μl of the diluted cells to a new 96-well microplate.
6. Incubate the culture plate with shaking at 900 rpm on a plate shaker at 34 °C for 23 h.
7. Measure the OD_{600 nm} of each strain incubated for 23 h using the microplate reader.
8. Dilute the cells with modified M9 medium to an OD_{600 nm} of 1×10^{-4} .
9. Inoculate 200 μl of the diluted cells to a 96-well microplate.
10. Incubate the culture plate with shaking at 900 rpm on a plate shaker at 34 °C to an OD_{600 nm} in the 0.072–0.135 range (equivalent of 10 generations) (*see Note 7*).
11. Harvest 180 μl of the cell solution from the well and place in an ultracentrifuge tube.
12. Immediately add an equal volume of ice-cold ethanol containing 10% (w/v) phenol to each cell solution, and mix gently but thoroughly.
13. Centrifuge it at 20,000 $\times g$ at 4 °C for 5 min.
14. Carefully remove and discard supernatant. Store pelleted cells at -80 °C until use.
15. Add 100 μl of 1 mg/ml lysozyme in TE buffer to the cells and mix by vortex for 10 s.
16. Incubate the lysozyme reactions at room temperature for 10 min. During incubation, vortex for 10 s every 2 min.
17. Add 350 μl of Buffer RLT containing 1% (v/v) 2-mercaptoethanol, vortex and pulse-spin tubes to collect samples.
18. Add 250 μl of 99.5% ethanol and pipet up and down several times. Do not pulse-spin tubes.
19. Load onto RNeasy Micro columns and purify total RNA using an RNeasy Micro Kit with on-column DNA digestion in accordance with the manufacturer's instructions.
20. Measure the total RNA concentration and quality using a spectrophotometer and electrophoresis (*see Note 8*).
21. Store the purified RNA at -80 °C until transcriptome analysis.

3.4 DNA Microarray Analysis

The total RNA preparation can be applied to other microarray platforms. Here, we give a concise account of our application to Agilent SurePrint G3 Custom Microarray, 8 × 60 K.

3.4.1 Preparation of Spike Mix

1. Create an appropriate spike mix dilution using Agilent One-Color RNA Spike-In Kit in accordance with the user manual, “Agilent One-Color Microarray-Based Exon Analysis - Low Input Quick Amp WT Labeling ver1.0” (*see Note 9*).
2. Thaw Spike Mix and Dilution Buffer provided in the Spike-In Kit at 37 °C for 5 min and mix thoroughly on a vortex mixer.
3. Put 2 µl of Spike Mix into a new tube and add 38 µl of Dilution Buffer provided in the Spike-In kit to make the First Dilution. Then, mix thoroughly on a vortex mixer.
4. Put 2 µl of the First Dilution into a fresh tube and add 48 µl of Dilution Buffer to make the Second Dilution. Then, mix thoroughly on a vortex mixer.
5. Put 2 µl of the Second Dilution into a fresh tube and add 38 µl of Dilution Buffer to make the Third Dilution (10,000-fold final dilution). Then, mix thoroughly on a vortex mixer.

3.4.2 cDNA Synthesis

1. Add 100 ng of total RNA to a 1.5 ml tube in a final volume of 2.3 µl (*see Note 10*).
2. Add 2 µl of the Third Dilution of Spike Mix and 1 µl of WT Primer provided in the Low Input Quick Amp WT Labeling kit.
3. Denature the primer and the template by incubating the reactions at 65 °C for 10 min.
4. Put the reactions on ice for 5 min.
5. Pre-warm the 5× First Strand Buffer provided in the Labeling kit at 80 °C for 4 min and thaw completely.
6. Prepare cDNA Master Mix containing 2 µl of 5× First Strand Buffer, 1 µl of 0.1 M DDT, 0.5 µl of dNTP mix, 1.2 µl of Affinity Script RNase Block Mix provided in the Labeling kit (*see Note 11*).
7. Add 4.7 µl of cDNA Master Mix to the sample tube and mix by pipetting.
8. Incubate the sample at 40 °C in a water bath for 2 h.
9. Move the sample to a 70 °C water bath for 15 min to inactivate reverse transcriptase.
10. Move the sample on ice and incubate for 5 min.

3.4.3 In Vitro Transcription

1. Prepare Transcription Master Mix containing 0.75 µl of Nuclease-free water, 3.2 µl 5× Transcription Buffer, 0.6 µl of 0.1 M DTT, 1 µl of NTP Mix, 0.21 µl of T7 RNA Polymerase Blend, and 0.24 µl of Cy3-CTP provided in the Labeling kit (*see Note 11*).
2. Add 6 µl of Transcription Master Mix to the sample tube.
3. Incubate the sample in a water bath at 40 °C for 2 h.

3.4.4 Purification

of Cy3-Labeled cRNA

Sample Using the RNeasy
Mini Kit

1. Add 84 μ l of Nuclease-free water to the sample to bring the total volume up to 100 μ l.
2. Add 350 μ l of Buffer RLT to the sample tube and mix by pipetting.
3. Add 250 μ l of 100% ethanol and mix thoroughly by pipetting.
4. Transfer the 700 μ l of the sample to an RNeasy spin column and centrifuge for 30 s at approx. $16,000 \times g$ in a microcentrifuge at 4 °C. Discard the flow-through.
5. Transfer the RNeasy spin column to a fresh tube and add 500 μ l of Buffer RPE containing ethanol. Centrifuge for 30 s at $16,000 \times g$ at 4 °C and discard the flow-through.
6. Add 500 μ l of Buffer RPE containing ethanol. Centrifuge for 60 s at $16,000 \times g$ at 4 °C and discard the flow-through.
7. Transfer the RNeasy spin column to a fresh tube and elute the cRNA sample. Add 30 μ l RNase-Free water onto the filter membrane in the column. Wait 60 s and then centrifuge for 30 s at $16,000 \times g$ at 4 °C.
8. Keep the cRNA sample on ice.

3.4.5 Assessment

of Cy3-Labeled cRNA

by Spectrophotometry

1. Start the NanoDrop software and set to Microarray Measurement mode.
2. Select RNA-40 as the Sample type.
3. Load 1.0–2.0 μ l of nuclease-free water and click Blank button.
4. Load 1.0–2.0 μ l of the sample and click Measure button to quantify (1) Cy3 dye concentration ($\text{pmol}/\mu\text{l}$), (2) RNA absorbance ratio (260 nm/280 nm), and (3) cRNA concentration ($\text{ng}/\mu\text{l}$).
5. Calculate specific activity as (concentration of Cy3)/(concentration of cRNA) $\times 1000 = \text{pmol Cy3 per } \mu\text{g cRNA}$.
6. Examine the yield and specific activity results. The recommended cRNA yield and specific activity for hybridization are 6 μg and 15 pmol Cy3 per μg cRNA, respectively.

3.4.6 Hybridization

1. Add 600 ng Cy3-labeled cRNA, 5 μ l of 10 \times Gene Expression Blocking Agent provided in the Gene Expression Hybridization Kit, nuclease-free water up to 24 μ l into a 1.5 ml fresh tube.
2. Add 1 μ l of 25 \times Fragmentation Buffer provided in the Gene Expression Hybridization Kit and incubate at 60 °C for exactly 30 min to fragment RNA. Immediately, cool on ice for one min.
3. Add 25 μ l of 2 \times Hi-RPM Hybridization Buffer provided in the Gene Expression Hybridization Kit to stop the fragmentation reaction.

4. Mix the sample well by careful pipetting. Do not introduce bubbles by the mixing.
5. Centrifuge for 1 min at room temperature at $16,000 \times g$.
6. Put sample on ice and apply it onto the array as soon as possible.
7. Load a clean gasket slide into the Agilent SureHyb chamber base.
8. Slowly apply 40 μl of hybridization sample onto the one of 8 gasket wells. Repeat this sample application eight times for the 8×60 K slide.
9. Grip the slide and slowly put it down on the SureHyb gasket slide. Make sure that the sandwich-pair is properly aligned.
10. Put the SureHyb chamber cover onto the sandwiched slides and firmly hand-tighten the clamp onto the chamber.
11. Load the chamber into Agilent Hybridization oven. Then hybridize at 65 °C for 17 h.

3.4.7 Washing Microarray Washing Slide

1. Fill slide-staining dish #1 with Gene Expression Wash Buffer 1 at room temperature.
2. Put a slide rack into slide-staining dish #2 containing a magnetic stir bar. Fill the dish with Gene Expression Wash Buffer 1 to cover the slide rack at room temperature. Put the dish on a magnetic stir plate.
3. Put the empty dish #3 containing a magnetic stir bar on the stir plate.
4. Remove one hybridization chamber from incubator. Then, slide off the clamp assembly and remove the chamber cover.
5. Remove the array-gasket sandwich from the chamber and submerge it into slide-staining dish #1 containing Gene Expression Wash Buffer 1.
6. Open the array-gasket sandwich in slide-staining dish #1 and remove the slide. And then put it into the slide rack in the slide-staining dish #2 containing Gene Expression Wash Buffer 1 at room temperature.
7. When all slides are put into the slide rack, stir the slide-staining dish #2 for 1 min.
8. Pour pre-warmed Gene Expression Wash Buffer 2 at 37 °C into the slide-staining dish #3.
9. Transfer the slide rack to slide-staining dish #3 containing Gene Expression Wash Buffer 2. Stir it using a moderate speed setting for 1 min.
10. Remove the slide rack from the slide-staining dish #3 slowly and put the slides into a slide holder.

3.4.8 Microarray Scanning

1. Carefully put the microarray slide into the slide holder. Then, put assembled slide holders into the scanner cassette.
2. Select the scanner protocol “AgilentG3_GX_1color.”
3. Click Start Scan. After scanning, a .tif image file is created.

3.5 Acquisition of Expression Levels and Normalization

1. To analyze the data, signal intensities of each probe are quantified from the .tif image file using Agilent Feature Extraction software 9.5.3.1 or later.
2. Acquisition of expression levels: In the microarray we designed for *E. coli* transcriptome analysis, 12 different probes are prepared for each coding region (see Note 1). The expression levels of each coding region are calculated as the median of signal intensities of these 12 probes (see Note 12). To compare gene expression profiles among multiple samples, the expression levels should be normalized in an appropriate way (see Note 13) as such as quantile normalization [11]. The goal of the quantile normalization method is to make the distribution of expression for each sample the same.
3. Given n expression profiles with m genes, form X of dimension $m \times n$ where each sample is a column.
4. Sort each column of X to give X_{sort} .
5. Take the averages across rows of X_{sort} and assign this average to each element in the row to get X'_{sort} .
6. Obtain $X_{\text{normalized}}$ by rearranging each column of X'_{sort} to realize the same ordering as original X .
7. After this procedure, each sample has the same distribution of gene expression.

3.6 Example of Data Analysis: Prediction of Antibiotic Resistance from Transcriptome Data

The purpose of transcriptome analysis is to extract genes whose expression changes contribute to the change of antibiotic resistance. Here, we show one example of such analyses, in which resistance levels (MIC) to various antibiotics are quantitatively predicted based on a small number of gene expression changes. For the details of the results of this analysis, see [5].

1. Prepare strains resistant to a single antibiotic by laboratory evolution (Subheading 3.1).
2. For each resistant strain, analyze cross-resistance and collateral sensitivity, i.e., quantify MICs to various antibiotics (Subheading 3.2).
3. For each resistant strain, quantify the expression profile by microarray analysis. To standardize the culture condition among the resistant strains, all transcriptome data should be obtained in a synthetic medium without addition of antibiotics.

4. In this analysis, we assumed that the drug resistance quantified by the MICs is determined as a function of gene expression levels. Also, for simplification, we neglected nonlinear effects and cross terms of the changes in gene expression. Thus, we assumed the following simple linear model to predict the MICs by the expression levels of N genes:

$$\text{MIC}_j^k = \sum_{i=1}^N \alpha_i^k X_{ij} + \beta^k$$

MIC_{jk} is the \log_2 -transformed relative MIC of the j th strain for the k th antibiotic, X_{ij} is the \log_{10} -transformed expression level of the i th gene in the j th resistant strain after standardization to zero mean and unit variance, and α_{ik} and β_k are fitting parameters.

5. Use the cross-validation method to avoid overfitting and to seek the number of genes with the highest predictive accuracy. For example, the resistant strains were randomly partitioned into four equally sized subgroups; one subgroup was used as the test data set for validation and the remaining three subgroups were used for the fitting of α_{ik} and β_k .
6. Sets of N genes that have high prediction accuracy for test data are obtained by a genetic algorithm (GA). First, prepare 1000 sets of randomly chosen N genes as an initial population.
7. Using each gene set, obtain α_{ik} and β_k by multiple regression method (*see Note 14*). The correlation coefficient between the predicted and observed MICs of the training data sets is used as the fitness of each gene set.
8. Among the 1000 gene sets, select those with fitness in the top 5% as the parent sets of the next generation. Then, generate mutant sets by randomly replacing a single gene without changing N .
9. Repeat more than 300 cycles of the generation and selection of mutant sets to obtain sets of N genes whose expression levels could represent the MICs for training data sets.
10. Repeat the selection of gene sets using 10,000 different training data sets prepared randomly by partitioning the total data set to obtain the frequency of genes selected after the GA. The expression levels of the frequently selected genes provide the most relevant information for predicting MIC changes.
11. To obtain the number of genes N which has highest prediction accuracy of test data set (*see Note 15*), perform the GA screening (more than 300 cycles for each) by changing N , for example, from 2 to 18.
12. After obtaining the optimal number of N , choose the appropriate gene set to represent the prediction accuracy of a test data set (Fig. 3 for example).

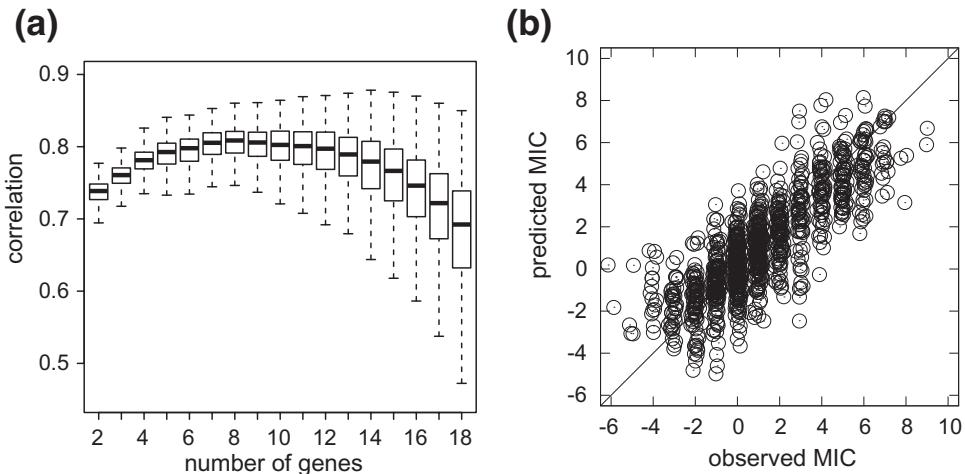


Fig. 3 Example of prediction of antibiotic resistance using transcriptomic data. **(a)** Box plot of prediction accuracy as a function of the number of genes, N , used for fitting. The prediction accuracy was quantified by the correlation coefficients between predicted MICs and observed MICs of the test data. **(b)** Comparisons between observed and predicted MICs for 25 antibiotics were calculated by fitting using the following eight genes: *acrB*, *ompF*, *cyoC*, *pps*, *tsx*, *oppA*, *folA*, and *pntB*. Only test data not used for the fitting are plotted. The error bars in the y-axis represent the standard deviation of predicted MICs calculated from 10,000 different sets of test data and training data. The figure is reproduced from [5] with permission

4 Notes

1. For transcriptome analyses of antibiotic-resistant strains in our previous study [5], we used an Agilent custom microarray. Here, we give a concise account of the design principles of the custom microarray. A fine-tuned expression 8×60 K custom microarray covering whole genes of *E. coli* K12 strain W3110 was newly developed according to the Agilent's expression array design guidance (<https://earray.chem.agilent.com/earray/>). First, we prepared all assumable 60-mer sequences corresponding to the entire 4499 coding regions as candidates for probes. It has been shown that microarray probes with stable secondary structures tend to show lower intensity [12] and quantification of this effect is difficult [13]. Therefore, we evaluated the stability of the secondary structures of all 60-mer candidates for probes using the UNAFold model [14] and then discarded candidates showing a lower free energy than a given threshold. At the same time, it should be appreciated that probes with high similarity to the complement sequences of other targets cause cross-hybridization and disturb the accurate estimation of expression level. Accordingly, we calculated the number of base pairs a candidate for a probe shared with other candidates for probes and then discarded candidates with 38 or more base pairs in common. We then selected 12 probes

from the candidates which conformed to the two conditions above for each coding region in such a way that GC contents of selected probes had widespread distribution. Although we could eventually set 12 probes for 4195 coding regions, we could not set 12 probes for the remaining 297 coding regions because they had high similarity to the sequence of another coding region or were very short. For the 291 coding regions that had high similarity to others, we selected 12 probes by easing the above conditions for probe selection and set as many probes as possible for six very short coding regions. As a result, we designed 53877 specific probes for analysis of expression levels of the 4499 coding regions and added 6070 non-control probes that were used for background correction and 1319 Agilent controls. The information of probe sequence is available in NCBI database (Platform ID: GPL18948).

2. The ranges of antibiotic concentrations used for the evolution experiments are in doubling dilution steps up and down from 1 µg/ml with three quartile concentrations according to minimum inhibitory concentrations (MICs) of each evolving culture line. The antibiotic serial diluted plates were stored at -80 °C prior to use. At a daily transfer, we thawed antibiotic serial diluted plates at 34 °C for 1 h and diluted the antibiotics with modified M9 medium according to MICs of each evolving culture line.
3. *E. coli* MDS42 is a reduced-genome strain that is free from transposon sequences (IS elements) [10]. The reason we used *E. coli* MDS42 strain here is to simplify the analysis of the genome sequences and the evaluation of identified mutations to fitness. Although identification of the transposon insertion site is possible by using high-throughput sequencers such as Illumina Hiseq, SOLiD, and Roche FLX+, to determine the precise sequence around the insertion often requires additional analysis as sequencing by the Sanger method. Thus, we adopted *E. coli* MDS42 to make the resequencing analysis easy and reliable.
4. We defined a well whose OD_{600 nm} was greater than 0.03 as the well in which cells could grow in our previous study [5].
5. Serial dilutions of each antibiotic are made in 96-well microplates using the modified M9 medium and stored at -80 °C prior to use. The ranges of antibiotic concentrations used for determining MICs are in doubling dilutions steps up and down from 1 µg/ml as required depending on each antibiotic.
6. The MICs were defined as the lowest concentration of antibiotic that reduced the growth to an OD_{600 nm} of less than 0.03 in our previous study [5].

7. We recommend identifying the best estimate time points, when growth of resistant and parent strains reach appropriate ranges for harvest, by analyses of growth curves in advance. Because most of the antibiotic-resistant strains sustain a reduction of fitness in the absence of antibiotics, known as the fitness cost, their growth rates vary significantly [15].
8. We recommend evaluating the quality of the purified total RNA using Agilent 2100 bioanalyzer and RNA 6000 Nano Kit, because the quality of the purified RNA is important for accurate estimation of the gene expression level. The RIN (RNA integrity number) should be approximately 9.0.
9. Agilent One-Color Microarray-Based Exon Analysis - Low Input Quick Amp WT Labeling was updated to version 2.0 in August 2015.
10. Select the amount of RNA input, 25, 50, or 100 ng.
11. Prepared cDNA Master Mix and Transcription Master Mix are set at room temperature until use to avoid formation of a precipitation. Use each Master Mix as soon as possible after adding enzymes.
12. To obtain gene expression levels from multiple probes designed for a coding region, we used median instead of mean signal intensity to avoid significant effects of outliers on expression levels. We confirmed that the quantification based on the median of 12 different probes provides a reliable quantification, for example, more than 99% of expression levels of replicates are within a 1.5-fold range.
13. The obtained expression levels are often biased for various reasons, such as differences in RNA quality, labeling efficiency, and so on. The expression bias can cause misinterpretations of expression changes; thus, to compare expression profiles, data should be normalized in an appropriate way. We recommend quantile normalization [11], which makes the distribution of expression levels for each data set the same.
14. Multiple regression method is available in various computer languages and libraries, for example the functions lm() in R and regress() in Matlab.
15. When the number of genes N is large, the prediction accuracy for the test data set became small due to overfitting. In contrast, when N was small, the linear combination of genes was insufficient to represent changes in the MICs and the prediction accuracy became small. The optimal number of N with highest prediction accuracy lies between these two cases.

Acknowledgment

This work was supported in part by a Grant-in-Aid for Scientific Research (S) [15H05746], a Grant-in-Aid for Scientific Research (B) [26290071,15H04733], and a Grant-in-Aid for Challenging Exploratory Research [26650138] from JSPS.

References

1. Lenski RE (1998) Bacterial evolution and the cost of antibiotic resistance. *Int Microbiol* 1(4):265–270
2. Conrad TM, Lewis NE, Palsson BO (2011) Microbial laboratory evolution in the era of genome-scale science. *Mol Syst Biol* 7:509. doi:[10.1038/msb.2011.42](https://doi.org/10.1038/msb.2011.42)
3. Lazar V, Nagy I, Spohn R, Csorgo B, Györkei A, Nyerges A, Horvath B, Voros A, Busa-Fekete R, Hrtyan M, Bogos B, Mehi O, Fekete G, Szappanos B, Kegl B, Papp B, Pal C (2014) Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nat Commun* 5:4352. doi:[10.1038/ncomms5352](https://doi.org/10.1038/ncomms5352)
4. Toprak E, Veres A, Michel JB, Chait R, Hartl DL, Kishony R (2012) Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat Genet* 44(1):101–105. doi:[10.1038/ng.1034](https://doi.org/10.1038/ng.1034)
5. Suzuki S, Horinouchi T, Furusawa C (2014) Prediction of antibiotic resistance by gene expression profiles. *Nat Commun* 5:5792. doi:[10.1038/ncomms6792](https://doi.org/10.1038/ncomms6792)
6. Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* 405(6784):299–304. doi:[10.1038/35012500](https://doi.org/10.1038/35012500)
7. Lee HH, Molla MN, Cantor CR, Collins JJ (2010) Bacterial charity work leads to population-wide resistance. *Nature* 467(7311):82–85. doi:[10.1038/nature09354](https://doi.org/10.1038/nature09354)
8. Lazar V, Pal Singh G, Spohn R, Nagy I, Horvath B, Hrtyan M, Busa-Fekete R, Bogos B, Mehi O, Csorgo B, Posfai G, Fekete G, Szappanos B, Kegl B, Papp B, Pal C (2013) Bacterial evolution of antibiotic hypersensitivity. *Mol Syst Biol* 9:700. doi:[10.1038/msb.2013.57](https://doi.org/10.1038/msb.2013.57)
9. Imamovic L, Sommer MO (2013) Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. *Sci Transl Med* 5(204):204ra132. doi:[10.1126/scitranslmed.3006609](https://doi.org/10.1126/scitranslmed.3006609)
10. Posfai G, Plunkett G 3rd, Feher T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M, Burland V, Harcum SW, Blattner FR (2006) Emergent properties of reduced-genome Escherichia coli. *Science* 312(5776):1044–1046. doi:[10.1126/science.1126439](https://doi.org/10.1126/science.1126439)
11. Bolstad BM, Irizarry RA, Åstrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19(2):185–193
12. Matveeva OV, Shabalina SA, Nemtsov VA, Tsodikov AD, Gesteland RF, Atkins JF (2003) Thermodynamic calculations and statistical correlations for oligo-probes design. *Nucleic Acids Res* 31(14):4211–4217
13. Ono N, Suzuki S, Furusawa C, Agata T, Kashiwagi A, Shimizu H, Yomo T (2008) An improved physico-chemical model of hybridization on high-density oligonucleotide microarrays. *Bioinformatics* 24(10):1278–1285. doi:[10.1093/bioinformatics/btn109](https://doi.org/10.1093/bioinformatics/btn109)
14. Markham NR, Zuker M (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* 33(Web Server Issue):W577–W581. doi:[10.1093/nar/gki591](https://doi.org/10.1093/nar/gki591)
15. Suzuki S, Horinouchi T, Furusawa C (2015) Phenotypic changes associated with the fitness cost in antibiotic resistant *Escherichia coli* strains. *Mol Biosyst.* doi:[10.1039/c5mb00590f](https://doi.org/10.1039/c5mb00590f)

Chapter 17

Sample Preparation for Mass-Spectrometry Based Absolute Protein Quantification in Antibiotic Stress Research

Florian Bonn, Sandra Maass, and Dörte Becher

Abstract

Absolute protein quantification is an essential tool for system biology approaches and elucidation of stoichiometry of multi-protein complexes. In this chapter, a universal protocol for gel free absolute protein quantification in bacterial systems is described, which can be used for sample preparation prior to miscellaneous mass-spectrometry-based quantification workflows like AQUA, Hi3, and emPAI. In addition, a focus has been set to the specific challenges in antibiotic stress research.

Key words Gel free proteomics, Sample preparation, Absolute protein quantification, In-solution digestion, AQUA, QconCAT, Hi3, emPAI, Antibiotic stress response

1 Introduction

Antibiotics trigger stress responses in bacteria when the cells try to adapt their metabolism to the treatment. Metabolism is driven by enzymes and therefore adaptations in the metabolism due to antibiotic stress or resistance are connected to changes in the cellular proteome. Proteomics is a valuable tool to understand antibiotic action and bacterial resistance mechanisms in multiple ways. For example, 2D gel-based signature libraries can help to elucidate the mode of action of new, antibacterial compounds. As the focus of this chapter is gel-free proteomics, an excellent review from the group of Julia Bandow is recommended for readers who are interested in 2D gel-based proteomics in antibiotic research [1].

Besides the development on new drugs, the elucidation of bacterial resistance mechanism is a main issue in antibiotic research. Proteomic studies can help to understand the physiological adaptations in resistant strains [2]. Another option to study antibiotic action and bacterial adaptation are antibiotic stress experiments, the treatment of nonresistant bacteria with antibiotics. For elucidation

of changes in the proteome of resistant strains compared to sensitive strains or the analysis of the antibiotic stress response, all kinds of quantitative proteomic techniques can be applied. An overview about the technical possibilities in quantitative microbial proteomics can be found in a review by Otto et al. [3].

Mass spectrometry-based proteomics is in principle only suitable for relative quantification, but several strategies for absolute protein quantification have been established. With these approaches protein copy numbers per cell can be calculated and used for determination of stoichiometry within physiological or functional protein complexes or for kinetic calculations of biochemical reactions. Global absolute proteome studies are therefore extremely helpful for the examination of complex metabolic adaptations like in antibiotic resistant bacterial isolates or in antibiotic stress experiments.

Absolute protein quantification in prokaryotic systems has recently been reviewed extensively [4]; therefore, in this chapter, only a brief overview is provided. For absolute protein quantification by mass spectrometry the addition of reference peptides or proteins of known amount is crucial. One commonly used approach to achieve highest accuracy is based on reference peptides, which were derived from the target protein, labeled with stable isotopes (^2D , ^{13}C , ^{15}N) and spiked into the digested cell extract. By comparison with the spiked heavy isotope of known amount the natural homologues of the isotopically labeled peptides can be quantified [5, 6]. However, due to the requirement of isotopically labeled peptides of the target proteins these approaches are limited to a rather small number of proteins. Maass et al. brought absolute quantification to a nearly global scale by calibrating 2D gels with reference proteins [7]. As 2D gel-based analyses are time consuming and in general limited to soluble proteins, gel-free approaches are preferred by most research groups. Gel-free absolute quantification of a large number of proteins in one sample can be achieved via the combined (and eventually corrected) intensities or spectral-counts of all associated peptides [8–10] or only the intensities of the three most intense peptides of each protein [11], which are then compared to one or more unlabeled reference proteins of known amounts to determine absolute intensities. All approaches have advantages and drawbacks, but in most cases the choice will be made on the availability of the necessary instrumentation and the scale of the experiment to be performed. As the protocol for sample preparation prior to absolute quantification of microbial proteins is dependent on the analyzed organisms, subproteomic fraction, and selected quantification approach, a general suggestion of a suitable workflow is not possible. In contrast to that, sample preparation requirements are very similar for all gel-free workflows. Complete extraction and digestion of all proteins from the cell without introducing a bias toward a specific class of proteins is crucial for absolute quantification. To determine copy numbers

per cell and protein concentration in the cell exact determination of cell number and cellular volume in a given sample is essential.

In this chapter, a workflow for complete extraction, digestion, and desalting of soluble microbial protein is provided, which can be used for all gel-free proteomic analyses. Although the protocol has been initially developed for gram-positive bacteria, especially *Bacillus subtilis* and *Staphylococcus aureus*, the protocol can be applied to other microorganisms. However, in these cases especially the cell lysis efficiency should be monitored carefully.

Furthermore, a focus is set on the specific challenges in the elucidation of antibiotic stress responses, which take effect especially in the first steps of the workflow. Antibiotics can induce morphological changes, which influence the number of cells per ml and OD unit. In addition, the composition and thickness of the cell wall and membrane can be influenced by antibiotics, which could lead to a decrease in cell disruption efficiency. For a calculation of proteins per cell the number of harvested cells, as well as the disruption efficiency, are crucial information, as it is absolutely essential to know the number of cells that were disrupted in the experiment. In this protocol, mechanical cell disruption by glass beat beating is suggested, which is an excellent choice for many gram-positive and gram-negative species. In some cases, other cell-disruption methods may be favorable. It is recommended to test different cell-lysis strategies and choose the most effective by cell disruption efficiency and protein yield.

2 Materials

Prepare all solutions in MilliQ quality water and also use MilliQ quality water for all dilutions. Prepare and store all reagents under ambient conditions if not indicated elsewhere.

2.1 Cell Counting

1. A Neubauer cell counting chamber or an alternative device, which allows for effective cell counting with your organism of choice.
2. Physiological saline: Solve 900 mg NaCl in 100 ml water. Autoclave or sterile filter the solution.

2.2 Cell Harvest and Lysis

1. 50 mM TEAB buffer: Dilute 0.5 ml of a 1 M Triethyl ammonium bicarbonate (TEAB) stock solution with 9.5 ml of water. Store at 4 °C for a maximum of 2 weeks.
2. A ribolyser.
3. Glass beads with a diameter of approximately 0.1 mm.
4. Low binding reaction tubes.

2.3 Determination of Protein Concentration

1. 10 mg/ml BSA stock solution: Solve 10 mg BSA in 1 ml of water and store in a low binding reaction tube. Otherwise dilute higher concentrated BSA solutions (*see Note 1*).
2. Concentrated hydrochloric acid (37%).
3. Stannous chloride solution: Dissolve 1 g SnCl₂ in 10 ml ethylene glycol (*see Note 2*).
4. Sodium acetate buffer: Solve 54.4 g sodium acetate in 20 ml acetic acid and fill up to 95 ml with water. Adjust pH to 5.5 with NaOH solution and fill up to 100 ml with water.
5. Ninhhydrin solution: Solve 1 g ninhydrin in 37.5 ml ethylene glycol and 12.5 ml sodium acetate buffer. Mix carefully in the dark until all ninhydrin is dissolved (at least 1 h). This solution should be stored only for few hours. Directly before adding the ninhydrin solution to the sample add 1.25 ml stannous chloride solution to the ninhydrin reagent (*see Note 2*).
6. Lightproof reaction tubes (for example brown reaction tubes).

2.4 Tryptic Digest

1. 0.5% RapiGest: Solve RapiGest (Waters) in 50 mM TEAB (see above) to a final concentration of 0.5% (w/v).
2. 500 mM TCEP: dissolve 10 mg Tris(2-carboxyethyl) phosphine (TCEP) in 70 µl 50 mM TEAB, prepare this solution freshly.
3. 500 mM IAA: Dissolve 10 mg iodoacetamide in 108 µl 50 mM TEAB prepare this solution freshly and store in the dark.
4. Trypsin-solution: Solve 20 µg of sequencing grade modified trypsin in 100 µl trypsin resuspension buffer (e.g., Promega, delivered together with the trypsin). Prepare shortly before use or prepare aliquots in low-binding reaction tubes and store at -20 °C.
5. Activated trypsin: Incubate trypsin solution at 30–37 °C with vigorous shaking for 10–15 min. Prepare directly before use.
6. Trifluoroacetic acid.

2.5 Desalting

1. Methanol (LC-MS quality).
2. Bulk C18 chromatographic material with a particle size of 5 µm or less.
3. StageTips (Thermo) (*see Note 3*).
4. Gelloader pipette tips (*see Note 4*).
5. Acetic acid or formic acid (LC-MS quality)
6. Buffer A: Add 10 µl acetic acid to 9.99 ml of water.
7. Acetonitrile (LC-MS quality).
8. Buffer B: Add 10 µl acetic acid to 9.99 ml of acetonitrile.
9. Elution solution: Mix 700 µl buffer B with 300 µl buffer A (*see Note 5*).

3 Methods

Handle cells and protein extracts at 4 °C or on ice whenever it is possible.

3.1 Cell Counting

1. Choose an antibiotic concentration for the experiments. We recommend using a concentration, which will decrease the growth rate by approximately 50% in the exponential growth phase (*see Note 6*).
2. Determine the ratio between optical density and cell count. Dilute an aliquot of the culture with physiological saline until you can count individual cells in the Neubauer chamber. It is recommended to count cells at least in four fields with 20–100 cells each in two replicates for every sample (*see Note 7*).
3. Calculate the number of cells per ml and OD unit for all samples.

3.2 Cell Harvest and Lysis

1. Take an aliquot of the culture to determine the optical density at the time point of cell harvest.
2. Fill the culture in centrifuge tubes, determine the exact volume either with a measuring cylinder or preferentially by weighing the tubes before and after filling (*see Note 8*).
3. Harvest the cells by centrifugation at 4 °C and $8000 \times g$. Choose a centrifugation duration, which is sufficient to achieve a stable cell pellet. The duration depends on organism, optical density, and media, but is usually between 3 and 10 min.
4. Wash the cells two times with TEAB. For each washing step use approximately 0.1 ml TEAB per ml and OD unit of the original culture. Be careful not to lose any cells during washing.
5. Resuspend the cells in approximately 0.02 ml TEAB per ml and OD unit of the original culture. Determine the number of cells per ml.
6. Transfer the cell suspension to tubes, which can be used in your ribolyser and are pre-filled with glass beads in a ratio of 1:2. Determine the weight of all tubes without and with cell suspension to determine the exact volume of the suspension.
7. Lyse the cells by beat beating. Three 30 s cycles at $0.66 \times g$ are in most cases sufficient to lyse at least 95 % of all cells. Between the cycles cool the cells for 5 min on ice. Determine the number of intact cells per ml.
8. Calculate the cell lysis efficiency (*see Note 9*).
9. Pellet cell debris and glass pearls by centrifugation at 4 °C and $8000 \times g$ for 2 min. Transfer the supernatant to a low binding reaction tube.

10. Pellet the remaining cell debris by centrifugation at 4 °C and $15,000 \times g$ for 15 min. Transfer the supernatant to a new low binding reaction tube. Determine the weight of all tubes before and after adding the protein extract to determine the exact volume.
11. Calculate the loss during the cell lysis process.
12. Calculate the number of cells which proteins are in your final sample on the base of the number of harvested cells, the cell lysis efficiency, and the cells lost during the cell lysis process.

3.3 Determination of Protein Concentration

1. Prepare BSA solutions with the following concentrations in TEAB buffer: 0.25, 0.5, 1.0, 1.5, 2.0, 4.0, 6.0, 8.0, and 10.0 µg/µl.
 2. Prepare reaction tubes for all concentrations and add 20 µl BSA solution to each tube. Prepare an additional tube for the blank and add 20 µl TEAB buffer.
 3. Prepare for each sample three reaction tubes with 20 µl protein extract each.
 4. Add 20 µl concentrated HCl to all samples (including BSA solutions and blank) and seal the reaction tubes (*see Note 10*).
 5. Incubate all samples for 24 h at 100 °C to achieve quantitative hydrolysis of all proteins.
 6. Cool the samples to room temperature on ice. Be careful as the samples are pressurized and corrosive!
 7. Centrifuge all samples for 30 s at $5000 \times g$ to remove all liquids from the lid of the tubes.
 8. Dilute 10 µl of every hydrolyzed sample with 90 µl water. Further dilute 50 µl of these solutions with 450 µl water to create 1:100 dilutions of the samples.
 9. Transfer 200 µl of each 1:100 dilutions to a lightproof reaction tube.
 10. Mix thoroughly with 200 µl water and 600 µl fresh ninhydrin solution (*see Note 11*).
 11. Incubate for 10 min at 100 °C.
 12. Cool the samples shortly on ice. Be careful as the samples are pressurized.
 13. Transfer the samples to micro cuvettes and measure absorbance at 575 nm against the blank.
 14. Use the BSA samples to create a calibration curve. Calculate the protein concentration of the other samples on the base of this curve.
1. Transfer samples with a total protein amount of 100 µg per samples to low binding reaction tubes. Fill up to 77 µl with 50 mM TEAB (*see Note 12*).

3.4 Tryptic Digest

2. Add 20 µl 0.5% RapiGest.
3. If necessary, add undigested protein references.
4. To reduce disulfide bridges, add 1 µl 500 mM TCEP and incubate the samples for 30–45 min at 60 °C.
5. Cool the samples shortly on ice and spin down all samples.
6. Add 2 µl 500 mM IAA and incubate for 15 min in the dark at ambient temperature.
7. Add 2.5 µl activated trypsin and incubate for 5–6 h at 37 °C and shaking (900 rpm).
8. To stop the tryptic digest, add 0.5–1 µl trifluoroacetic acid. Check the pH, which needs to be below 3 (*see Note 13*).
9. Incubate for 30 min at 37 °C without shaking to precipitate the RapiGest. Mix the samples every 5–10 min by inverting the tubes.
10. Spin down the RapiGest by centrifugation at $20,000 \times g$ for 12 min. Transfer the supernatant to a new low binding tube and repeat this step twice.
11. Store the supernatant at –20 °C or proceed with the desalting.

3.5 Desalting

1. Create a mount for every sample by puncturing a hole of 1.5–2.5 mm diameter in the lid of a 2 ml reaction tube. Cut the StageTip approximately 2 mm below the original C18 and place it in the mount.
2. Resuspend 10 µl chromatographic material per sample in twice the amount of methanol.
3. Add 15 µl of the C18 suspension to every StageTip and sediment it by centrifugation (2000–5000 $\times g$, 10–20 s). Add more C18 material until the lower tampered part over the original chromatographic material is filled with C18 material (*see Note 14*).
4. Wash the C18 material twice with 100 µl buffer A. Fill the buffer without air bubbles into the tip and squeeze it through by centrifugal forces (8000 $\times g$, 1–2 min).
5. Wash the chromatographic material twice with buffer B (*see Note 15*).
6. Equilibrate the C18 material with 100 µl buffer A. Squeeze the buffer until only 1–2 mm of the tip above the custom C18 is filled with liquid. Do not let the C18 material run dry in the following steps.
7. Fill the samples into the tips, place it directly over the C18 without an air bubble in between.
8. Load the peptides onto the C18 material by squeezing the sample through by centrifugation (8000 $\times g$, 1–2 min).
9. Wash the peptide loaded C18 with 100 µl buffer A. Repeat this step once.

10. Squeeze the rest of the buffer through the C18 with a syringe (syringes are supplied with the StageTips) until the upper layer of the buffer is directly over the C18.
11. Prepare one LC-MS vial for every sample and determine the weight of each vial.
12. Elute the sample with 30 µl elution solution. Use a syringe to squeeze the eluate directly into the prepared vials (*see Note 16*).
13. Add 10 µl buffer A to every sample.
14. Reduce the sample volume to 3–10 µl in a vacuum centrifuge (*see Note 17*).
15. Add peptide references for absolute quantification if necessary.
16. Determine the exact sample volume by weight and fill up with buffer A to a final volume of 100 µl.
17. Prepare aliquots for LC-MS analysis and analyze directly or store at -70 °C.

4 Notes

1. BSA is not easy to resolve completely. It is recommended to use a commercially available BSA solution with known amount.
2. Ethylene glycol is very viscous, cut the top of the pipette tips before use and pipette very slowly and cautiously.
3. Commercially available C18-based solid phase extraction kits for peptide desalting with a binding capacity of at least 100 µg can be used alternatively according to the manufacturer's instructions.
4. The low cost products normally have a rather wide capillary and are therefore better suited for this protocol.
5. Because acetonitrile evaporates faster than water this solution should not be stored for more than a few hours and the tube should be kept close.
6. Do not limit yourself to subinhibitory concentrations. If antibiotic stress is induced in the exponential growth phase or later higher doses of antibiotics can be necessary to induce considerable effects on proteome level.
7. Antibiotics can change the cellular shape dramatically. This also changes the ratio between cell numbers per ml and optical density. The ratio between OD and cell number can be determined in preliminary tests.
8. With standard media a mass density of 1 g per ml can be assumed.
9. If the cell lysis efficiency is below 95 % you should optimize the cell lysis process and eventually test other protocols.

10. There are special clamps available for sealing reaction tubes (e.g., MCP LidLocks, Sorenson), as an alternative you can put a glass plate with some weights on the top of the tubes if you boil them in a heating block.
11. The ninhydrin solution is viscous and needs to be pipetted slowly and cautiously.
12. The protocol can be scaled down easily to an initial protein amount of 20 µg per sample by linearly adapting the amounts of all reagents. Also upscaling is possible, but the maximum load of each StageTip should not exceed 150 µg.
13. The pH should not be below 1.5, to avoid acidic hydrolysis. An optimum is a pH between 2 and 3.
14. If C18 material is sticking at the side of the StageTip you can use either methanol or acetonitrile to flush it.
15. Because of the lower back pressure of acetonitrile the centrifugation times for squeezing buffer B are shorter than for buffer A.
16. Especially if you have a high number of samples you can put some tissue on the top of the forcer of the syringe to protect your finger.
17. Quantitative solvation of peptides without detergents or chaotropes is complicated; therefore, the samples should not run dry!

References

1. Wenzel M, Bandow JE (2011) Proteomic signatures in antibiotic research. *Proteomics* 11(15):3256–3268
2. Vranakis I, Goniotakis I, Psaroulaki A et al (2014) Proteome studies of bacterial antibiotic resistance mechanisms. *J Proteomics* 97:88–99
3. Otto A, Becher D, Schmidt F (2014) Quantitative proteomics in the field of microbiology. *Proteomics* 14(4–5):547–565
4. Maass S, Becher D (2016) Absolute protein quantification in bacterial systems. *J Proteomics* 136:222–233
5. Gerber SA, Rush J, Stemman O et al (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci* 100(12):6940–6945
6. Pratt JM, Simpson DM, Doherty MK et al (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat Protoc* 1(2):1029–1043
7. Maass S, Sievers S, Zühlke D et al (2011) Efficient, global-scale quantification of absolute protein amounts by integration of targeted mass spectrometry and two-dimensional gel-based proteomics. *Anal Chem* 83(7):2677–2684
8. Ishihama Y, Odai Y, Tabata T et al (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 4(9):1265–1272
9. Lu P, Vogel C, Wang R et al (2006) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* 25(1):117–124
10. Schwanhäusser B, Busse D, Li N et al (2011) Global quantification of mammalian gene expression control. *Nature* 473(7347):337–342
11. Silva JC, Gorenstein Marc V, Guo-Zhong L et al (2006) Absolute quantification of proteins by LCMS^E. *Mol Cell Proteomics* 5(1):144–156

Chapter 18

Label-Free Quantitation of Ribosomal Proteins from *Bacillus subtilis* for Antibiotic Research

Sina Schäkermann, Pascal Prochnow, and Julia E. Bandow

Abstract

Current research is focusing on ribosome heterogeneity as a response to changing environmental conditions and stresses, such as antibiotic stress. Altered stoichiometry and composition of ribosomal proteins as well as association of additional protein factors are mechanisms for shaping the protein expression profile or hibernating ribosomes. Here, we present a method for the isolation of ribosomes to analyze antibiotic-induced changes in the composition of ribosomes in *Bacillus subtilis* or other bacteria. Ribosomes and associated proteins are isolated by ultracentrifugation and proteins are identified and quantified using label-free mass spectrometry.

Key words Mass spectrometry, Ribosome heterogeneity, Stress, Proteomics

1 Introduction

Ribosomes are remarkable ribonucleoprotein complexes and essential for translation in all living cells. Due to their key role in cellular physiology, ribosomes are targeted by many clinically used antibiotics [1, 2]. To allow for protein synthesis, more than 50 ribosomal proteins and three ribosomal RNAs assemble into the bacterial 70S ribosome. Yet, the ribosome is not a fixed entity but is adapted to changing environmental conditions and stresses, resulting in ribosome heterogeneity. As recently reviewed [3–5], ribosome heterogeneity is, e.g., achieved by altered stoichiometry and composition of ribosomal proteins, the modification status of ribosomal proteins or rRNA, and interaction with additional proteins. For instance, the stoichiometry of ribosomal proteins in *Escherichia coli* was found to change depending on the growth rate [6–8]. Similarly, upon zinc-limiting conditions in *Bacillus subtilis*, ribosomal proteins L31 and S14 containing zinc-binding motifs are replaced by non-zinc-binding homologues [9, 10]. During cold shock, bulk protein translation is repressed as *E. coli* ribosomes are inactivated by binding of protein Y (pY), which is rapidly

released when growth conditions improve [11]. Yet, to allow for selective translation of cold shock proteins while bulk protein production is repressed, increased levels of translation initiation factors IF1 and IF3 are necessary [12, 13]. Ribosome hibernation due to dimerization into 100S particles is mediated by associating proteins and is a common response to environmental stress conditions in bacteria [14–16]. Ribosome hibernation is also important for tolerance toward aminoglycosides in stationary phase [17]. In response to the antibiotic kasugamycin, *E. coli* ribosomes lacking some ribosomal proteins of the small subunit like S1 occur. Although these ribosomes are unable to translate canonical mRNAs, translation of leader-less mRNAs (lRNAs) was confirmed *in vivo* and *in vitro* [18]. Ribosomes lacking S1 are also present upon normal growth conditions and lRNA translation might be relevant under stress conditions [19, 20]. The role of ribosomes as regulatory elements shaping the protein expression profile by ribosomal heterogeneity is discussed as the ribosome filter hypothesis [21].

The method described in this chapter is suitable for label-free quantitation of ribosomal proteins from *B. subtilis* and other bacteria for studying ribosome heterogeneity upon antibiotic treatment or other stress conditions. Studying ribosome heterogeneity can be helpful for understanding changes in protein expression counteracting antibiotic-induced stress or to analyze adaptation strategies in response to ribosome-targeting antibiotics. Using this method, we quantified 380 proteins of untreated *B. subtilis* ribosome isolations. Among those, 27 of 35 and 18 of 22 ribosomal proteins were quantified, which according to ref. [22] can be part of the large and small ribosomal subunit, respectively (Table 1). In *B. subtilis*, 35 ribosomal proteins are thought to be essential, as gene disruption was not possible [22]. All but one of these essential proteins were identified using the described method (Table 1). Furthermore, proteins typically associated with the ribosome were found like, e.g., initiation factors InfB and InfC; elongation factors TufA, FusA, Tsf, and Efp; trigger factor Tig; ribosome recycling factor Frr; and the GTPases involved in ribosome assembly YscC and YqeH.

2 Materials

Unless noted otherwise, prepare stock solutions in *A. dest*. Use UPLC/MS grade solvents, formic acid (FA), and trifluoroacetic acid (TFA) for mass spectrometry. Make sure to follow regional waste disposal regulations.

Table 1
Quantification of ribosomal proteins after ribosome isolation from untreated *B. subtilis*

Ribosomal protein^a		fmol					
		Replicate			Mean	SD	% SD^b
		1	2	3			
RplA	L1	86.5	78.6	92.1	85.7	5.5	6.5
RplB	L2	124.1	145.6	181.3	150.3	23.6	15.7
RplC	L3	87.8	84.5	79.6	83.9	3.4	4.0
RplD	L4	117.4	106.5	118.8	114.2	5.5	4.8
RplE	L5	130.7	113.1	126.8	123.5	7.5	6.1
RplF	L6	104.1	100.5	106.2	103.6	2.4	2.3
RplI	L9	7.7	8.1	9.0	8.3	0.5	6.2
RplJ	L10	43.1	40.1	38.0	40.4	2.1	5.2
RplK	L11	81.6	70.9	69.2	73.9	5.5	7.5
RplL	L7/L12	239.3	214.8	210.3	221.5	12.8	5.8
RplM	L13	119.4	100.4	126.2	115.3	10.9	9.5
RplN	L14	53.8	46.9	48.5	49.7	3.0	6.0
RplO	L15	90.6	77.3	86.5	84.8	5.6	6.6
RplP	L16	64.9	63.1	65.6	64.5	1.0	1.6
RplQ	L17	98.7	91.2	91.7	93.9	3.4	3.7
RplR	L18	76.3	60.0	53.8	63.4	9.5	14.9
RplS	L19	118.8	105.4	100.9	108.3	7.6	7.0
RplT	L20	65.8	72.3	77.1	71.7	4.6	6.4
RplU	L21	77.7	73.9	92.6	81.4	8.1	9.9
RplV	L22	107.2	95.0	100.2	100.8	5.0	4.9
RplW	L23	44.5	39.7	44.8	43.0	2.3	5.5
RplX	L24	NF ^c	63.1	86.9	75.0	11.9	15.8
Ctc	L25 homologue	ND ^d					
RpmA	L27	61.8	59.6	40.6	54.0	9.5	17.6
RpmB	L28	ND					
RpmC	L29	94.3	78.6	85.9	86.2	6.4	7.4
RpmD	L30	49.6	45.1	54.8	49.8	3.9	7.9
RpmEA	L31A	6.8	7.6	15.5	10.0	3.9	39.2
RpmEB	L31B	ND					

(continued)

Table 1
(continued)

Ribosomal protein ^a		fmol			Mean	SD	% SD ^b
		1	2	3			
RpmF	L32	ND					
RpmGA	L33.1	ND					
RpmGB	L33.2	ND					
RpmH	L34	ND					
RpmI	L35	NF	9.0	31.1	20.0	11.1	55.3
RpmJ	L36	ND					
ypfD	S1	2.3	2.1	1.0	1.8	0.6	31.5
RpsB	S2	52.5	51.0	76.6	60.0	11.7	19.5
RpsC	S3	61.1	53.3	70.8	61.7	7.2	11.6
RpsD	S4	103.8	87.4	142.1	111.1	22.9	20.6
RpsE	S5	124.0	121.7	181.2	142.3	27.5	19.3
RpsF	S6	73.3	60.2	82.0	71.8	9.0	12.5
RpsG	S7	108.0	92.1	129.9	110.0	15.5	14.1
RpsH	S8	85.1	74.1	98.1	85.8	9.8	11.5
RpsI	S9	44.1	40.7	58.6	47.8	7.7	16.2
RpsJ	S10	61.0	56.4	92.9	70.1	16.2	23.2
RpsK	S11	59.9	57.7	79.9	65.8	10.0	15.1
RpsL	S12	29.0	34.4	66.9	43.5	16.7	38.4
RpsM	S13	62.4	53.4	77.6	64.5	10.0	15.4
RpsN	S14	ND					
YhzA	YhzA homologue	ND					
RpsO	S15	48.7	40.9	69.6	53.1	12.1	22.8
RpsP	S16	48.5	43.5	66.7	52.9	10.0	18.8
RpsQ	S17	36.6	35.8	43.6	38.7	3.5	9.0
RpsR	S18	21.7	17.9	25.4	21.6	3.0	14.1
RpsS	S19	51.1	46.9	80.8	59.6	15.1	25.3
RpsT	S20	ND					
RpsU	S21	ND					

For ribosomal proteins identified in at least two of three replicates the normalized amount is shown as fmol on column. Arithmetic means (mean) and standard deviations (SD) are given

^aProteins that according to ref. [22] are thought to be essential in *B. subtilis* are marked in bold

^bValues ≥39 % are marked in bold (see Note 37)

^cProtein not found in this replicate

^dProtein not detected in at least two of three replicates

**2.1 Cultivation
of *B. subtilis*
in Belitzky Minimal
Medium (See Note 1)**

1. Water bath shaker.
2. Sterile Erlenmeyer glass flasks (100 ml, 500 ml) with cotton plugs.
3. Photometer.
4. Falcon centrifuge.
5. Belitzky base medium, 15 mM $(\text{NH}_4)_2\text{SO}_4$, 8 mM MgSO_4 , 27 mM KCl, 7 mM $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$, 50 mM Tris-HCl, pH 7.5 (see Note 2).
6. 200 mM KH_2PO_4 , autoclaved, stored at RT.
7. 1 M CaCl_2 , autoclaved, stored at RT.
8. 1 mM FeSO_4 , filter sterilized, stored at RT.
9. 25 mM MnSO_4 , filter sterilized, stored at RT.
10. 20% (w/v) D-glucose, filter sterilized, stored at RT.
11. 500 mM L-glutamic acid, filter sterilized, stored at 4 °C.
12. 39 mM L-tryptophan, filter sterilized, stored at 4 °C.
13. Supplemented BMM: 600 µM KH_2PO_4 , 2 mM CaCl_2 , 2 µM FeSO_4 , 10 µM MnSO_4 , 0.2% glucose, 4.5 mM L-glutamic acid, 780 µM L-tryptophan in base medium.
14. Glycerol stocks of *B. subtilis* 168 for inoculation (see Note 3).

**2.2 Ribosome
Isolation by
Ultracentrifugation**

To protect samples from RNase activity, only use autoclaved water for buffer preparation. This method was modified from ref. [23, 24].

1. French Press cell disruption system (see Note 4).
2. Ultracentrifuge equipped with fixed angle and swing-out rotors.
3. Appropriate ultracentrifugation tubes.
4. Magnetic stirrer.
5. Magnetic stir bars (≤ 0.6 cm).
6. Glass beaker.
7. 10% H_2O_2 in a squirt bottle.
9. 1.1 M sucrose, filter sterilized, stored at 4 °C.
10. 1 M Tris-HCl, pH 7.5, autoclaved, stored at RT.
11. 2 M NH_4Cl , autoclaved, stored at RT.
12. 0.5 M EDTA (see Note 5), autoclaved, stored at RT.
13. 1 M magnesium acetate (MgAc), filter sterilized, stored at RT.
14. 0.1 M DTT (see Note 6).
15. 10 mg/ml DNase I, stored as 100 µl aliquots at -20 °C.
16. Buffer A (see Note 7), 10 mM Tris-HCl, pH 7.5, 10 mM MgAc, 60 mM NH_4Cl , 0.5 mM EDTA, 3 mM β -mercaptoethanol (see Note 8).

17. Buffer B (*see Note 7*), 20 mM Tris-HCl, pH 7.5, 10.5 mM MgAc, 0.5 mM NH₄Cl, 0.5 mM EDTA, 3 mM β-mercaptoethanol (*see Note 8*).
18. Protease inhibitor cocktail of choice.

2.3 Tryptic Digest of Isolated Ribosomes

Use MS grade water and solvents for preparation of stock solutions. Wear protective gloves and clean lab coats, and clean the working bench thoroughly to avoid contaminations with keratin (*see Note 9*).

1. Refrigerated table-top centrifuge.
2. Thermomixer adjustable to 60 and 37 °C (*see Note 10*).
3. 1% (w/v) stock solution of RapiGest (Waters, Milford, USA), 100 µl aliquots stored at -20 °C.
4. 250 mM Tris-(2-carboxyethyl)phosphine hydrochloride (TCEP), 25 µl aliquots stored at -20 °C.
5. 500 mM iodoacetamide (*see Note 11*).
6. 0.5 µg/µl trypsin (*see Note 12*).
7. TFA.

2.4 UPLC-MS^E Analysis

1. Synapt G2-S HDMS ToF mass spectrometer equipped with an ESI nanoLockSpray source coupled online to a nanoAcquity UPLC system and operated with MassLynx software (e.g., version V4.1 SCN932; Waters, *see Note 13*).
2. NanoACQUITY UPLC Symmetry C18 Trap Column (pore size, 100 Å; particle diameter, 5 µm; inner diameter, 180 µm; length, 20 mm; Waters).
3. NanoACQUITY UPLC peptide CSH C18 Column (pore size, 130 Å; particle diameter, 1.7 µm; inner diameter, 75 µm; length, 150 mm; Waters).
4. Ultrasonic bath.
5. Glass vials for mass spectrometry.
6. Hi3 quantitation standard (Hi3 PhosB standard, Waters; *see Note 14*), stored as 10 µl aliquots of 10 pmol/µl in 3% acetonitrile/0.5% TFA at -80 °C.
7. 0.1% FA in MS-grade *A. dest.* for dilution of tryptic digests.
8. Solvent A, 0.1% FA in MS-grade *A. dest.*, degassed in an ultrasonic bath for 10 min.
9. Solvent B, 0.1% FA in acetonitrile, degassed in an ultrasonic bath for 10 min.
10. Lock mass, 0.25 µg/ml leucine enkephalin (*see Note 15*) in 50% acetonitrile containing 0.1% FA, degassed in an ultrasonic bath for 10 min.

2.5 Label-Free Quantitation of MS^E Data

1. ProteinLynx Global Server (PLGS, version 2.5.2, Waters).
2. Microsoft Excel (e.g., version 2010) or comparable software.

3 Methods

3.1 Cultivation of *B. subtilis* for Ribosome Isolation (See Note 16)

Be sure to work steriley while preparing media and handling cultures. Grow all cultures of *B. subtilis* at 37 °C in a shaking water bath under steady agitation at 200 rpm in sterile Erlenmeyer flasks with cotton plugs. The ribosome isolation and label-free quantitation method of course is not restricted to *B. subtilis*, but this organism is chosen here as an example for the experimental procedure. The method described here can also be combined with in vivo crosslinking techniques to stabilize transient protein interactions if necessary.

1. Use 10 µl of a glycerol stock of *B. subtilis* 168 to inoculate 20 ml of BMM in a 100 ml Erlenmeyer flask (see Note 17). Prepare 3–10 serial 1:10 dilutions of the same volume in Erlenmeyer flasks and incubate overnight at 37 °C in a shaking water bath (see Note 18). Store the remaining supplemented BMM at 4 °C for short-term usage.
2. After overnight incubation, measure the OD₅₀₀ against BMM. Choose one of the cultures, which is still in exponential phase (see Note 19; OD₅₀₀ 0.25–0.8) to inoculate the main culture to an OD₅₀₀ of 0.05. Use 100 ml (see Note 20) of BMM pre-warmed to 37 °C in a 500 ml Erlenmeyer flask for the main culture.
3. Let the culture grow to OD₅₀₀ 0.15 (see Note 21), split the volume and transfer to new pre-warmed Erlenmeyer flasks (see Note 22). Leave one of the cultures untreated as control and add your antibiotic of choice in appropriate concentrations to the second flask (see Note 23). Incubate in the water bath for 1 h to allow adaptation of the proteome to the antibiotic stress.
4. Transfer each culture to 50 ml falcon tubes and harvest the cells by centrifugation for 10 min at 4 °C and 3000×*g*. From now on, keep the samples cool at all times. Discard the supernatant and wash the cells by suspending in 25 ml buffer A. Centrifuge again (10 min, 4 °C, 3000×*g*), discard the supernatant, and store at –80 °C until further use.

3.2 Cell Disruption and Ribosome Isolation (See Note 24)

To protect your samples from RNases, use 10% H₂O₂ to clean the working bench and ultracentrifugation tubes. Make sure to exactly balance all samples for ultracentrifugation using a scale (see Note 25).

1. Resuspend each pellet in 5 ml buffer B and protease inhibitor of choice, DNase (15 µl of stock), and DTT (25 µl of stock).

2. Disrupt cells by French Press using eight to ten passages (*see Note 26*).
3. Transfer the suspensions to appropriate ultracentrifugation tubes, and remove cell debris in three ultracentrifugation steps in a fixed-angle rotor at 4 °C (*see Note 27*). Centrifuge at $16,000 \times g$ for 8 min, discard the pellet, and subject the supernatant to a second centrifugation at $22,000 \times g$ for 8 min. Again, discard the pellet and centrifuge the supernatant at $43,000 \times g$ for 25 min.
4. Slowly load the supernatant onto 5 ml of sucrose in appropriate ultracentrifugation tubes and fill up completely with buffer B. Make sure that the weight of all tubes is exactly balanced before proceeding to ultracentrifugation in a swing-out rotor at 4 °C for 20 h at $80,000 \times g$.
5. Slowly decant the supernatant and rinse the ribosome pellet with buffer B (*see Note 28*). Add a magnetic stir bar (≤ 0.6 cm) and resolve the pellet on ice for ~2 h in 500 µl buffer B (*see Note 29*).
6. Determine the protein concentration, e.g., by using Bradford or Ninhydrin assays following manufacturer's instructions or standard lab protocols and store as 0.25 µg/µl aliquots with a volume of 50 µl in 1.5 ml tubes (*see Note 30*) at -80 °C until further use.
7. To confirm that ribosome isolation was successful, use 10 µg of the samples for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) and Coomassie staining using standard protocols. In addition, 16S and 23S rRNA can be extracted and analyzed by 3-(N-morpholino)propanesulfonic acid (MOPS) agarose gels using standard protocols. As a reference, Fig. 1 shows a Coomassie-stained SDS gel of a typical isolation of ribosomal proteins from *B. subtilis* (A) and a corresponding RNA gel for extracted 16S and 23S rRNA from the same sample (B).

3.3 Tryptic Digest (See Note 31)

1. Add 0.1% RapiGest (Waters) and 2.5 mM TCEP to each sample and incubate at 60 °C for 45 min.
2. Add 5 mM freshly prepared iodoacetamide and incubate at RT for 25 °C in the dark (*see Note 32*).
3. Add 0.25 µl trypsin to reach an enzyme to protein ratio of 1:100 and mildly shake samples incubating at 37 °C (*see Note 10*) for 5 h.
4. Add 2 µl TFA sample to precipitate RapiGest and to terminate the digestion. Remove RapiGest by centrifugation at $16,000 \times g$ for 10 min at 4 °C. Transfer the supernatant to a fresh tube and repeat until RapiGest is removed completely (*see Note 33*).
5. Store tryptic digest at -80 °C until further use.

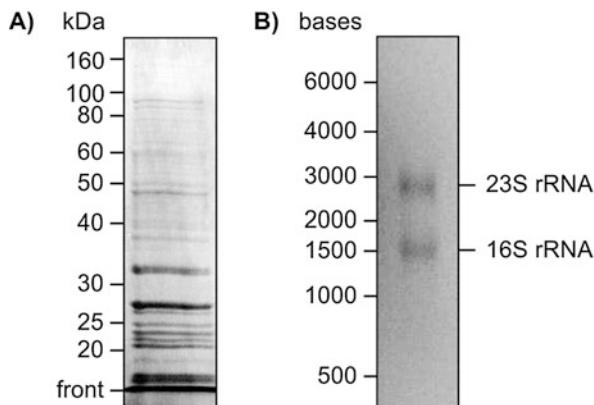


Fig. 1 Visualization of ribosomal proteins (a) and ribosomal RNA (b) after ribosome isolation from *B. subtilis*. Ribosomal proteins after SDS-PAGE and Coomassie staining are shown in (a). rRNA was extracted from isolated ribosomes, separated by MOPS agarose electrophoresis, and visualized by ethidium bromide (b)

3.4 Mass Spectrometry

1. Tryptic digests are diluted 1:10 in 0.1% FA containing 12.5 fmol/ μ l Hi3 quantitation standard and 10–20 μ l is transferred into glass vials. For analysis, 4 μ l is loaded onto the trap column (see Note 34). Samples are desalted with a flow of 10 μ l/min 0.5% solvent B for 3 min.
2. Peptides are eluted with a linear gradient from 0.5 to 60% B in 150 min. The column is subsequently washed and re-equilibrated with the following steps: linear gradient to 99.5% B in 5 min, hold at 99.5% B for 10 min, linear gradient to 0.5% B in 5 min, hold at 99.5% for 10 min.
3. In parallel, leucine-enkephalin is pumped with a flow rate of 0.5 μ l/min as lock mass.
4. Mass spectra in MS^E mode are recorded with MassLynx software using the following settings: start time, 0 min; end time, 180 min. Acquisition mode: polarity, positive; analyzer mode, resolution. MS mass range: low mass, 50 Da; high mass, 1200 Da; scan time, 1 s. Collision energy: in function 2 (high energy) ramp trap collision energy from 14 to 45 V. LockSpray properties: acquire LockSpray, do not apply correction; scan time, 1 s; interval, 60 s; scans to average, 3; mass window, \pm 0.3 Da.
5. Acquisition times: start time, 0 min; end time, 180 min.
6. Acquisition mode: polarity, positive; analyzer mode, resolution.
7. MS^E mass range: low mass, 50 Da; high mass, 1200 Da; scan time, 1 s.
8. Collision energy: in function 2 (high energy) ramp trap collision energy from 14 to 45 V.
9. LockSpray properties: acquire LockSpray, do not apply correction; scan time, 1 s; interval, 60 s; scans to average, 3; mass window, \pm 0.3 Da.

3.5 Label-Free Quantitation of Proteins

1. Load MS^E raw data into PLGS for processing.
2. Use the following Apex3D settings as processing parameters to generate mass spectra for database analysis: chromatographic peak width, automatic; MS ToF resolution, automatic; lock mass, 556.2771 Da/e; lock mass window, 0.25 Da; low energy threshold, 50 counts; elevated energy threshold, 15 counts, intensity threshold, 500 counts.
3. Chromatographic peak width, automatic.
4. MS ToF resolution, automatic.
5. Lock mass, 556.2771 Da/e; lock mass window, 0.25 Da.
6. Low energy threshold, 50 counts; elevated energy threshold, 15 counts, intensity threshold, 500 counts.
7. For database analysis, use a nonredundant database containing protein sequences of *B. subtilis* 168 (e.g., NCBI Reference Sequence: NC_000964.3), as well as the sequences for the Hi3 quantitation standard (refer to manufacturer's instructions), trypsin, and keratin. The following parameters are suitable for the database search: peptide and fragment tolerance, automatic; minimal fragment ion matches per peptide, 2; minimal fragment ion matches per protein, 6; minimal peptide matches per protein, 3; maximum protein mass, 250,000 Da; primary digest reagent, trypsin; secondary digest reagent, none; missed cleavages, 1; fixed modifications, carbamidomethyl C; variable modifications: deamidation N, deamidation Q, oxidation M; false positive rate, 4 %; calibration protein, PhosB; calibration protein concentration, 50 fmol.
8. Peptide and fragment tolerance, automatic.
9. Minimal fragment ion matches per peptide, 2; minimal fragment ion matches per protein, 6; minimal peptide matches per protein, 3.
10. Maximum protein mass, 250,000 Da.
11. Primary digest reagent, trypsin; secondary digest reagent, none; missed cleavages, 1.
12. Fixed modifications, carbamidomethyl C; variable modifications: deamidation N, deamidation Q, oxidation M.
13. False positive rate, 4 %.
14. Calibration protein, PhosB; calibration protein concentration, 50 fmol.
15. Activate the automation setup option for IdentityE (Ion accounting output) for automated export of peptide and protein lists for each sample to manually analyze the proteomics data as described below or use the automated Expression Analysis provided by PLGS following manufacturer's instructions.

3.6 Manual Analysis of Regulated Proteins (See Note 35)

1. Open all protein lists (typically three biologically independent per condition investigated) in Excel.
2. Using the filter function in Excel, reject proteins with <95 % probability (proteins with “green OK” symbol in PLGS or “Green” written in an excel sheet in column “protein. AutoCurate” are of at least 95 % probability), proteins found in the random reverse database (protein.dataBaseType random), and proteins that could not be quantified (no entry in protein. fmolOnColumn).
3. To compensate for possible variations in sample load, apply the following normalization. For each sample, calculate the sum for protein.fmolOnColumn (absolute protein amount) of all proteins, then normalize the value for each protein in the sample against this sum (see Note 36).
4. Combine the protein amount data of all samples in one spreadsheet. For further analysis, for each condition include only proteins for which data exists from at least two out of three biological replicates.
5. Calculate arithmetic means and standard deviations (SD) for each condition (see Note 37, Table 1). To select up- and down-regulated proteins, calculate the \log_2 ratio of the arithmetic means of control and antibiotic-treated conditions for each protein.
6. Select for up- or down-regulated proteins using a confidence interval of 95 % ($\text{mean}_{\log_2\text{ratio}} \pm 1.963 \times \text{SD}_{\log_2\text{ratio}}$), p -values below 0.05, and appropriate SD values (see Note 37). Proteins found in all three replicates representing one condition but in none of the replicates representing the other condition can be considered “unique” for this condition if the SD value is appropriate (see Note 37).

4 Notes

1. This protocol describes cell cultivation in chemically defined medium, which allows full control over medium composition. However, this is not critical for the successful application of the following steps of the protocol, and bacterial cultures can be grown in complex media.
2. Weigh in all components for 1 l of basal medium and dissolve in 800 ml *A. dest.* Adjust pH to 7.5 using HCl, fill up to 1000 ml with *A. dest.*, prepare 250 ml aliquots, autoclave, and store at room temperature (RT).
3. Let *B. subtilis* 168 grow to exponential phase in BMM at 37 °C and 200 rpm, dilute 1:1 with 100 % glycerol and store as 50 µl aliquots at -80 °C.

4. French Press is comparably gentle for cell disruption, yet other methods like bead beating or sonication might work as well for this protocol.
5. Adjust pH to 8 with NaOH; otherwise EDTA will not dissolve.
6. Store aliquots at -20 °C. DTT in solution is not suitable for long-term storage. Discard thawed aliquots and prepare fresh if aliquots lost the typical odor of DTT. Do not refreeze thawed aliquots.
7. Prepare fresh from stock solutions prior to use.
8. β-mercaptoethanol solutions of 99% purity have a concentration of 14.3 M.
9. Keratin contaminations will interfere with ionization of your samples during mass spectrometry. Use lint-free precision wipes for all cleaning purposes. If you refill pipette tips manually, wear clean gloves and wash out tip boxes with MS grade ethanol or methanol regularly. Do not autoclave tips for tryptic digest or mass spectrometry to minimize contaminations derived from the autoclaving process.
10. Use a thermomixer with ~100 rpm at 37 °C and without shaking at 60 °C, respectively. Alternatively, tape a tube rack onto an orbital shaker in a 37 °C climate chamber and use a standard heating block for 60 °C.
11. Prepare fresh, use on the same day only, and store at 4 °C in the dark until needed.
12. Use sequencing grade trypsin. Prepare stock solutions in the buffer recommended by the manufacturer and store as 10 µl aliquots at -80 °C. Thaw and refreeze for a maximum of three times.
13. In principle, ribosomes can of course be analyzed by alternative quantitation strategies and other types of instruments. As we use label-free quantitation based on the Hi3 technology of MS^E data with a Synapt setup, this protocol will only cover this approach.
14. This standard is suitable for bacterial proteomics, as these peptide sequences are not encoded in bacterial genomes.
15. Store 400 µg/ml stocks in 0.1% FA at -80 °C. Alternatively, use 500 fmol/µl [Glu1]-fibrinopeptide B (250 pmol/µl stocks in 0.1% FA) as lock mass, yet this compound is drastically less stable.
16. For quantitative analysis of antibiotic-induced changes of the ribosome composition, cultures are grown as biologically independent triplicates. All the following steps including cell disruption, ribosome isolation, and tryptic digest should be performed simultaneously and mass spectrometrical analysis should be performed *en bloc* to minimize technical variation.

17. Use a minimal ratio of medium to flask volume of 1:5 to allow aerobic growth of the cultures.
18. This is done to yield exponentially growing cultures for inoculation of the main culture. The number of dilutions needed may vary and depends on the used glycerol cultures and incubation times.
19. This is important to avoid a lag phase when starting the main culture, which will happen if stationary phase cells are used or if cells are cold-shocked by inoculating cold BMM.
20. The protocol works fine for a minimum volume of ~50 ml per sample for successful ribosome isolation from *B. subtilis*.
21. This cell density is suitable for monitoring proteome changes after 60 min of antibiotic stress with control cultures still in exponential phase at the time of harvest. This is critical, as the composition of the ribosome is altered upon entry into stationary phase.
22. Use pre-warmed and sterile falcons, pipettes, or measuring cylinders for this purpose.
23. The amount of antibiotic used should reduce growth rates to 50–70% but should not kill the cells. This makes sure that the cells can adapt their proteome to the stress caused by the antibiotic.
24. The method described here does not separate the 30S from the 50S subunits. To do so, a subsequent sucrose gradient centrifugation is needed.
25. Wipe the outside of the ultracentrifugation tubes to remove water stemming from keeping the samples on ice and place tubes on a scale in a glass beaker. To balance tubes against each other, remove the cap, place it on the scale next to the glass beaker, and use a pipette to add the appropriate amount of buffer B.
26. The number of passages needed for successful cell disruption has to be adjusted to the bacterial strain investigated and the volume of the pressure cell. Use as many passages as needed to reach a clear suspension. Make sure the pressure cell is cooled down prior to use (on ice or in a cooling chamber) to avoid heating of the sample. If necessary, interrupt the disruption process to cool the pressure cell on ice.
27. Please refer to the manual of your ultracentrifuge for detailed information on handling centrifugation tubes, as operating errors can be extremely dangerous and procedures may vary between centrifuges and rotors.
28. Gently rinse pellet with buffer B, but do not resuspend the pellet.

29. Place the ultracentrifugation tube in a glass beaker filled with ice and put it onto a magnetic stirrer in a cool room (4 °C).
30. We recommend using low-bind tubes for storage and handling of proteinaceous samples.
31. It is highly recommended to use pipettes suitable for accurately transferring very small volumes (0.1–2.5 µl-pipettes and 10 µl-pipettes, respectively).
32. Put the tubes into a lightproof box or wrap in an aluminum foil.
33. Make sure that no precipitated RapiGest remains in the solution—this may take additional one to two centrifugation steps. If RapiGest is not pelleted properly (solution is opaque or particles are visible in the solution), add few more µl of TFA and repeat the centrifugation steps.
34. The volume of sample loaded may depend on the sample loop. For this example, a 5 µl loop was used and an injection volume of 4 µl contains tryptic peptides of 100 ng isolated protein plus 50 fmol of Hi3 quantitation standard.
35. In Excel, conditional formatting and the “vlookup” command are useful for manual handling of proteomics data.
36. Standard techniques like Bradford or Ninhydrin assays for whole protein quantitation in the samples prior to tryptic digests are not very reliable when determining rather low protein concentrations. The total protein amount loaded onto the column as detected by MS is calculated as a quality check for all samples and serves as reference for the normalization. If the average deviation between samples is high (e.g., more than 30%), you should consider repeating the MS analysis or preparing additional replicates for quantitation.
37. For technical replicates, the standard deviation using the MS^E-based Hi3 method was shown to be less than 15% [25]. For quality control, calculate the mean standard deviation across the three biological samples representing each condition. E.g., we found that the mean standard deviation for protein amounts of ribosome isolations (representing the combined biological and technical variability) to be ~21 ± 18%. As an example, results for ribosomal proteins isolated from untreated *B. subtilis* are given in Table 1.

Acknowledgments

We thank Birgit Klinkert and Johanna Roßmanith for the practical introduction into ribosome isolation and for technical support. Jennifer Stepanek and Dominik Wüllner are acknowledged for critically reading the manuscript. Furthermore, we would like to

thank Dörte Becher and Knut Büttner for sharing mass spectrometry protocols. Funding from the German Federal State of North Rhine Westphalia (NRW) is acknowledged for the mass spectrometer (“Forschungsgroßgeräte der Länder”) used in this protocol. JEB acknowledges funding from NRW for the grant “Translation of innovative antibiotics from NRW.”

References

- McCoy LS, Xie Y, Tor Y (2011) Antibiotics that target protein synthesis. Wiley Interdiscip Rev RNA 2(2):209–232
- Wilson DN (2014) Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nat Rev Microbiol 12(1):35–48
- Byrgazov K, Vesper O, Moll I (2013) Ribosome heterogeneity: another level of complexity in bacterial translation regulation. Curr Opin Microbiol 16(2):133–139
- Starosta AL, Lassak J, Jung K, Wilson DN (2014) The bacterial translation stress response. FEMS Microbiol Rev 38(6):1172–1201
- Sauert M, Temmel H, Moll I (2015) Heterogeneity of the translational machinery: variations on a common theme. Biochimie 114:39–47
- Deusser E, Wittmann HG (1972) Ribosomal proteins: variation of the protein composition in *Escherichia coli* ribosomes as function of growth rate. Nature 238(5362):269–270
- Van Duin J, Kurland CG (1970) Functional heterogeneity of the 30S ribosomal subunit of *E. coli*. Mol Gen Genet 109(2):169–176
- Kurland CG, Voynow P, Hardy SJ, Randall L, Lutter L (1969) Physical and functional heterogeneity of *E. coli* ribosomes. Cold Spring Harb Symp Quant Biol 34:17–24
- Nanamiya H, Akanuma G, Natori Y, Murayama R, Kosono S, Kudo T, Kobayashi K, Ogasawara N, Park SM, Ochi K, Kawamura F (2004) Zinc is a key factor in controlling alternation of two types of L31 protein in the *Bacillus subtilis* ribosome. Mol Microbiol 52(1):273–283
- Natori Y, Nanamiya H, Akanuma G, Kosono S, Kudo T, Ochi K, Kawamura F (2007) A fail-safe system for the ribosome under zinc-limiting conditions in *Bacillus subtilis*. Mol Microbiol 63(1):294–307
- Agafonov DE, Kolb VA, Spirin AS (2001) Ribosome-associated protein that inhibits translation at the aminoacyl-tRNA binding stage. EMBO Rep 2(5):399–402
- Giangrossi M, Brandi A, Giuliodori AM, Gualerzi CO, Pon CL (2007) Cold-shock-induced *de novo* transcription and translation of *infA* and role of IF1 during cold adaptation. Mol Microbiol 64(3):807–821
- Giuliodori AM, Brandi A, Giangrossi M, Gualerzi CO, Pon CL (2007) Cold-stress-induced *de novo* expression of *infC* and role of IF3 in cold-shock translational bias. RNA 13(8):1355–1365
- Wada A, Yamazaki Y, Fujita N, Ishihama A (1990) Structure and probable genetic location of a “ribosome modulation factor” associated with 100S ribosomes in stationary-phase *Escherichia coli* cells. Proc Natl Acad Sci U S A 87(7):2657–2661
- Ueta M, Yoshida H, Wada C, Baba T, Mori H, Wada A (2005) Ribosome binding proteins YhbH and YfiA have opposite functions during 100S formation in the stationary phase of *Escherichia coli*. Genes Cells 10(12):1103–1112
- Tagami K, Nanamiya H, Kazo Y, Maehashi M, Suzuki S, Namba E, Hoshiya M, Hanai R, Tozawa Y, Morimoto T, Ogasawara N, Kageyama Y, Ara K, Ozaki K, Yoshida M, Kuroiwa H, Kuroiwa T, Ohashi Y, Kawamura F (2012) Expression of a small (p)ppGpp synthetase, YwaC, in the (p)ppGpp(0) mutant of *Bacillus subtilis* triggers YvyD-dependent dimerization of ribosome. Microbiologyopen 1(2):115–134
- McKay SL, Portnoy DA (2015) Ribosome hibernation facilitates tolerance of stationary-phase bacteria to aminoglycosides. Antimicrob Agents Chemother 59(11):6992–6999
- Kaberdina AC, Szaflarski W, Nierhaus KH, Moll I (2009) An unexpected type of ribosomes induced by kasugamycin: a look into ancestral times of protein synthesis? Mol Cell 33(2):227–236
- Delvillani F, Papiani G, Deho G, Briani F (2011) S1 ribosomal protein and the interplay between translation and mRNA decay. Nucleic Acids Res 39(17):7702–7715
- Vesper O, Amitai S, Belitsky M, Byrgazov K, Kaberdina AC, Engelberg-Kulka H, Moll I (2011) Selective translation of leaderless mRNAs by specialized ribosomes generated

- by MazF in *Escherichia coli*. *Cell* 147(1): 147–157
21. Mauro VP, Edelman GM (2002) The ribosome filter hypothesis. *Proc Natl Acad Sci U S A* 99(19):12031–12036
22. Akanuma G, Nanamiya H, Natori Y, Yano K, Suzuki S, Omata S, Ishizuka M, Sekine Y, Kawamura F (2012) Inactivation of ribosomal protein genes in *Bacillus subtilis* reveals importance of each ribosomal protein for cell proliferation and cell differentiation. *J Bacteriol* 194(22):6282–6291
23. Spedding G (1990) Isolation and analysis of ribosomes from prokaryotes, eukaryotes, and organelles. In: Spedding G (ed) Ribosomes and protein synthesis. Practical approach series. IRL Press, Oxford, UK, pp 1–29
24. Blaha G, Stelzl U, Spahn CM, Agrawal RK, Frank J, Nierhaus KH (2000) Preparation of functional ribosomal complexes and effect of buffer conditions on tRNA positions observed by cryoelectron microscopy. *Methods Enzymol* 317:292–309
25. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ (2006) Absolute quantification of proteins by LCMS^E: a virtue of parallel MS acquisition. *Mol Cell Proteomics* 5(1): 144–156

Chapter 19

Functional Metagenomics to Study Antibiotic Resistance

Manish Boolchandani, Sanket Patel, and Gautam Dantas

Abstract

The construction and screening of metagenomic expression libraries has great potential to identify novel genes and their functions. Here, we describe metagenomic library preparation from fecal DNA, screening of libraries for antibiotic resistance genes (ARGs), massively parallel DNA sequencing of the enriched DNA fragments, and a computational pipeline for high-throughput assembly and annotation of functionally selected DNA.

Key words Functional metagenomics, Antibiotic resistance genes, Resistome, Functional selections, Massively parallel DNA sequencing, High-throughput assembly, Profile HMM-based annotation, PARFuMS, Resfams

1 Introduction

The continued evolution and global spread of antibiotic resistance genes (ARGs) in pathogens has become a major clinical and public health problem [1]. The increase in number, diversity, and range of multidrug-resistant organisms limits therapeutic options to resolve infections. To effectively mitigate or counter the antibiotic resistance problem, identification and characterization of ARGs as well as their modes of transmission and mechanisms of action is crucial.

The extensive use of antibiotics has unarguably led to a widespread increase in diversity and spread of ARGs in environmental reservoirs and pathogenic bacteria [2]. However, antibiotic resistance is ancient, existing long prior to the first discovery of natural-product antibiotics by Fleming [3]. Bacteria from diverse habitats carry extensive reservoirs of ARGs, collectively termed the “resistome,” which have the potential for facile transmission to pathogens [2, 4–6].

Two conventional approaches, culture-based [7] and targeted PCR-based [8], have been frequently applied to study ARGs in complex microbial communities. While cultivation in the lab is the “gold standard” for identifying bacteria with antibiotic resistance,

a large proportion (70–80 %) of bacteria are difficult to culture in the laboratory [9]. This leads to a huge under-sampling of micro-organisms belonging to diverse habitats, and as such their ARGs remain unanalyzed [10]. Targeted PCR-based approaches are generally used to identify and quantify ARGs with known sequence, bypassing the need for culture. However, these methods are only able to detect previously described genes and often require cloning into expression vectors and subsequent experimentation to verify function. Furthermore, homology-based identification and characterization of ARGs in shotgun sequences of microbial communities is inherently limited to the low number of genes with high sequence similarity to previously identified genes. In addition, such *in silico* analyses are unable to confirm the function encoded in putative ARGs and therefore require additional experimentation.

A much more efficient and powerful technique for characterizing resistomes is functional metagenomics (Fig. 1a) [3, 11–13], wherein total community DNA is cloned into an expression vector and transformed into a susceptible (and easily cultured) indicator strain. The resulting transformant library is assayed for antibiotic resistance by plating on selective media, and surviving ARGs are sequenced and annotated. This allows analysis of 10^9 – 10^{10} bp of DNA in a single experiment while exploiting three key advantages over culture- or PCR-based studies [12]: (1) No need to culture recalcitrant microorganisms, (2) No prior knowledge required about ARG sequences, and (3) Resistance phenotypes are directly associated with cloned and sequenced ARGs. Recent developments in high-throughput functional metagenomics [12] allow researchers to multiplex up to 400 functional metagenomic selections on a single Illumina sequencing lane. With the custom-built tool PARFuMS (Parallel Annotation and Reassembly of Functional Metagenomic Selections, Fig. 1b) [2], researchers can now perform demultiplexing, quality-filtering, trimming, assembly of the reads into full-length metagenomic fragments, and annotation in a single automated step, substantially reducing experimental cost.

Functional metagenomics has proven to be the most efficient and powerful method for the study of antibiotic resistance mechanisms and their associated genes across a wide variety of habitats. Our lab actively uses the following protocol and pipeline to identify and characterize ARGs derived from samples collected from different environmental- and human-associated microbial communities.

2 Materials

Prepare all reagents and buffers in ultrapure water. Use nuclease-free water to set up all reactions that involve DNA. Prepare and store all reagents at room temperature unless indicated otherwise. Thoroughly follow all applicable waste disposal regulations when disposing of biological and chemical waste.

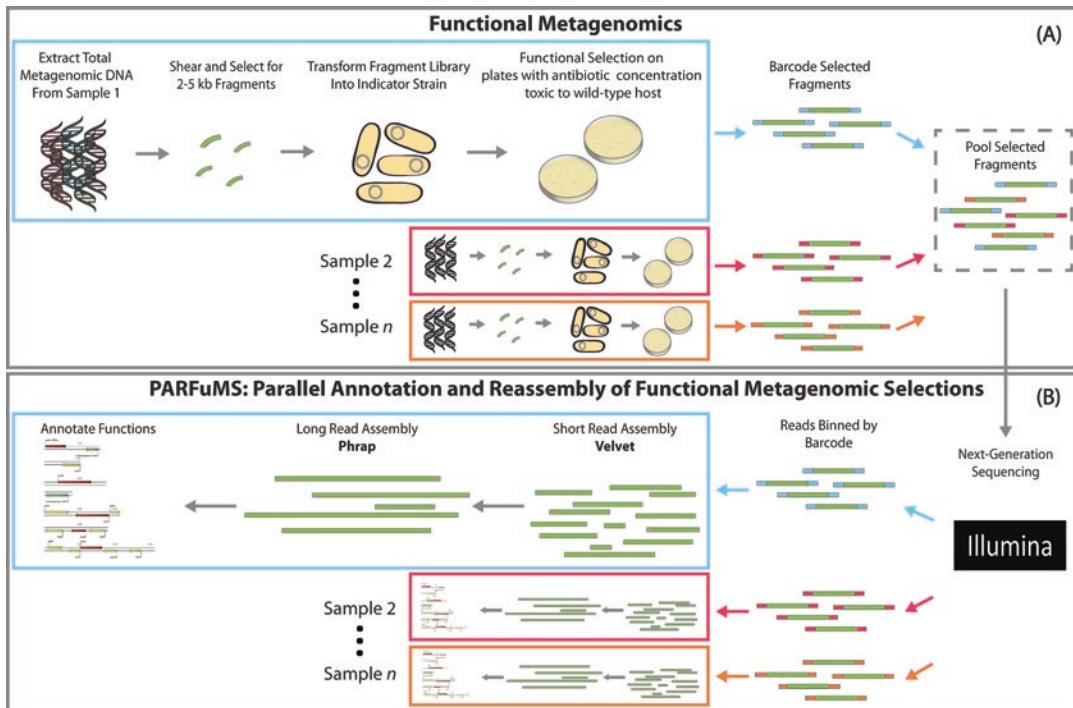


Fig. 1 Schematic representation of high-throughput functional metagenomic selection and resistome characterization. **(A)** DNA is extracted from the microbial community (e.g., feces), and 2–5 kb fragments are cloned into an expression system in an indicator strain (e.g., *Escherichia coli*). Transformants harboring resistance genes are selected using antibiotics at concentrations inhibitory to the wild-type indicator strain. Antibiotic resistance genes containing DNA fragments are PCR-amplified, sheared, bar-coded by end-ligation of oligonucleotide adapters with 5–8 bp unique sequences, and pooled in sets of up to 400 selections for next-generation sequencing. **(B)** Short (101 bp) sequencing reads from the Illumina platform are computationally binned by barcode, and assembled with computational pipeline PARFuMS, in which intermediate-length contigs from multiple rounds of assembly with the short-read assembler Velvet are assembled into full-length contigs using the long-read assembler Phrap. Contigs are annotated using custom profile Hidden Markov Models based Resfams database

2.1 DNA Extraction

1. DNA extraction buffer: 200 mM NaCl, 200 mM Tris, 20 mM EDTA in ultrapure water. Add 1.17 g of NaCl, 4 mL of 0.5 M EDTA (pH 8.0), and 20 mL of 1M Tris-HCl (pH 7.5) to ultrapure water. Make final volume to 100 mL.
2. 20% Sodium Dodecyl Sulfate (SDS).
3. 0.1 mm zirconium beads (BioSpec Products, Bartlesville, OK, USA).
4. 3 M sodium acetate, pH 5.5.
5. Phenol:Chloroform:Isoamyl alcohol (25:24:1), pH 8.0 (*see Note 1*).

6. Mini-Bead beater (BioSpec Products BioSpec Products, Bartlesville, OK, USA).
7. Phase-lock gel tube (PLG) (5Prime, Gaithersburg, MD, USA).
8. Molecular biology grade Isopropyl alcohol (Isopropanol).
9. Ethyl alcohol (Ethanol).
10. Tris-EDTA (TE) buffer, pH 8.0.
11. RNase A (100 mg/mL).
12. PCR Purification Kit.
13. Qubit® dsDNA BR Assay Kit (Invitrogen/ThermoFisher Scientific, Grand Island, NY, USA).
14. Qubit® dsDNA HS Assay Kit (Invitrogen/ThermoFisher Scientific, Grand Island, NY, USA).
15. Qubit® 2.0 Fluorometer (Invitrogen/ThermoFisher Scientific, Grand Island, NY, USA).

2.2 Construction of Metagenomic Library

1. 1% Low-melting point agarose gel in 0.5× Tris-Borate-EDTA (TBE) buffer with SYBR green I DNA binding dye.
2. 6× Gel loading dye.
3. pZE21 MCS-1 expression vector [14] (*see Note 2*).
4. High fidelity (HF) DNA Polymerase Kit (*see Note 3*).
5. Alkaline Phosphatase, Calf Intestinal (CIP), and corresponding 10× reaction buffer.
6. EB buffer (Qiagen, Hilden, Germany).
7. Gel Extraction Kit.
8. MinElute PCR Purification Kit (Qiagen).
9. MinElute Gel Extraction Kit (Qiagen).
10. End-It™ DNA End-Repair Kit (Epicentre, Madison, WI, USA).
11. Fast-Link™ DNA Ligation Kit (Epicentre).
12. 0.025 µm cellulose membrane (Millipore, Billerica, MA, USA).
13. 0.1 cm-gap Gene Pulser®/MicroPulser™ Electroporation Cuvettes (Bio-Rad, Hercules, CA, USA).
14. High efficiency ($\geq 4 \times 10^{10}$ cfu/µg) electrocompetent *E. coli* cells (*see Note 4*).
15. *Taq* DNA Polymerase with corresponding 10× reaction buffer.
16. 10 mM Deoxynucleotide (dNTP) Solution Mix.

2.3 Illumina Library Preparation and Sequencing

1. T4 DNA polymerase (New England Biolabs Inc, Ipswich, MA, USA).
2. T4 Polynucleotide Kinase (PNK) (New England Biolabs Inc, Ipswich, MA, USA).

Table 1
Commonly used antibiotics and their MICs are listed below

Antibiotics	Code	Selection concentration ($\mu\text{g/mL}$)	Antibiotic class	Antibiotic type
Aztreonam	AZ	8	β -Lactam	Synthetic
Chloramphenicol	CH	8	Amphenicol	Natural
Ciprofloxacin	CI	0.5	Fluoroquinolone	Synthetic
Colistin	CL	8	Polymyxin	Natural
Cefepime	CP	8	β -Lactam	Synthetic
Cefotaxime	CT	8	β -Lactam	Semisynthetic
Cefoxitin	CX	64	β -Lactam	Semisynthetic
d-Cycloserine	CY	32	Amino acid derivative	Natural
Ceftazidime	CZ	16	β -Lactam	Semisynthetic
Gentamicin	GE	16	Aminoglycoside	Natural
Meropenem	ME	16	β -Lactam	Semisynthetic
Penicillin	PE	128	β -Lactam	Natural
Piperacillin	PI	16	β -Lactam	Semisynthetic
Piperacillin-Tazobactam	PI-TZ	16-PI/4-TZ	β -Lactam	Semisynthetic
Tetracycline	TE	8	Tetracycline	Natural
Tigecycline	TG	2	Tetracycline	Semisynthetic
Trimethoprim	TR	8	Folate synthesis inhibitor	Synthetic
Trimethoprim-sulfamethoxazole	TR-SX	2-TR/38-SX	Folate synthesis inhibitor	Synthetic

3. T4 DNA ligase (New England Biolabs Inc, Ipswich, MA, USA).
4. 1 μM pre-annealed, Illumina barcoded sequencing adapters (see Note 5).
5. High-Fidelity PCR Master Mix.

2.4 Antibiotics and Media

1. *Antibiotics*: Purchase antibiotics in dry powdered form and prepare stock solutions as detailed in the manufacturer's MSDS. Detailed information on antibiotics is listed in Table 1 (see Note 6).
2. *Lysogeny Broth (LB) agar medium*: Add 5 g yeast extract, 10 g NaCl, 10 g tryptone, and 15 g agar in 900 mL of dH₂O. Mix well to dissolve and adjust pH to 7.5 using 1 N NaOH. Add dH₂O to 1 L. Autoclave on liquid cycle for 20 min or according to your autoclave's specification (see Notes 7 and 8).

3. *Lysogeny Broth (LB)*: Add 5 g yeast extract, 10 g NaCl, and 10 g tryptone in 900 mL of dH₂O. Mix well to dissolve and adjust pH to 7.5 using 1 N NaOH. Add dH₂O to 1 L. Autoclave on liquid cycle for 20 min or according to your autoclave's specifications (see Notes 7 and 8).
4. *Lysogeny Broth (LB) with 15% glycerol and 50 µg/mL kanamycin*: Mix 15 mL glycerol and 85 mL autoclaved LB broth in a clean sterile flask. Mix well and filter sterilize. Add kanamycin stock solution so that the final concentration of kanamycin is 50 µg/mL.
5. *Mueller–Hinton (MH) agar*: Add 2 g beef infusion solids, 1.5 g starch, 17 g agar, and 17.5 g casein hydrolysate in 900 mL of dH₂O. Mix well to dissolve and adjust pH to 7.4 using NaOH. Add dH₂O to 1 L. Autoclave on liquid cycle for 20 min or according to your autoclave's specifications (see Notes 7 and 8).

2.5 Instruments and Glassware

1. Thermocycler.
2. Centrifuges for 1.5 and 50 mL tubes.
3. Heat block.
4. Sonicator.
5. Electroporator.
6. Electrophoresis unit.
7. Gel trays and tank.
8. Gel imager and conversion screen.
9. Petri dish.
10. Cryo tubes.

2.6 Primers

1. pZE21 linearizing forward primer (5' GACGGTATCGATAA GCTTGAT 3').
2. pZE21 linearizing reverse primer (5' GACCTCGAGGGGG GG 3').
3. Colony PCR forward primer (5' GATACTGAGCACATC AGCAGGA 3').
4. Colony PCR reverse primer (5' CCTGATTCTGTGGATAA CCGTA 3').
5. Primer F1 (5' CCGAATTCAATTAAAGAGGGAGAAAG 3').
6. Primer F2 (5' CGAATT CATTAAAGAGGGAGAAAGG 3')
7. Primer F3 (5' GAATTCAATTAAAGAG GAGAAAGGTAC 3').
8. Primer R1 (5' GATATCAAGCTTATCGATACCGTC 3')
9. Primer R2 (5' CGATATCAAGCTTATCGATACCG 3').

10. Primer R3 (5' TCGATATCAAGCTTATCGATACC 3').
11. Illumina PCR Forward Primer (5' AATGATACGGCGA CCACCGAGATCTACACTCTTCCCTACACGACGCT CTTCCGATCT 3').
12. Illumina PCR Reverse Primer (5' CAAGCAGAACGG CATA CGA GAT CGGTCT CGGCATT CCTG CTG AACCGCTCTCCGATCT 3').

3 Methods

Carry out all procedures at room temperature unless otherwise specified. Procedures involving bacterial culture, media preparation, and transformation should be performed in a clean, sterile environment. Phenol:Chloroform:Isoamyl alcohol should be handled in the chemical hood. A PCR hood should be used to set up PCR reactions to avoid cross-contamination. Protocols may vary for different manufacturer reagents and kits. Follow manufacturer's instruction carefully and make necessary changes in protocol.

3.1 Metagenomic DNA Extraction

1. On dry ice, aliquot 50–100 mg of fecal material into a 2 mL sterile polypropylene tube (*see Note 9*).
2. Prepare samples for bead-beating by adding 250 μ L 0.1 mm zirconium beads, 210 μ L 20% SDS, 500 μ L DNA extraction buffer and 500 μ L phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) to the tube containing the fecal material. Keep samples on ice for about 5 min to cool down (*see Note 10*).
3. Lyse microbial cells by bead-beating using a Mini-Bead beater on “homogenize” setting for a total of 4 min (bead-beating for 2 min followed by cooling the samples on ice for 2 min, and bead-beating again for another 2 min) (*see Note 11*).
4. Centrifuge samples at $6081 \times g$ for 5 min at 4 °C.
5. Immediately prior to use, pellet Phase Lock Gel (PLG) tube at maximum speed (~ $16,058 \times g$) in a microcentrifuge for 20–30 s.
6. Taking care to avoid the pellet, transfer the top aqueous phase using a micropipette to a clean phase-lock gel tube and add 600 μ L of phenol:chloroform:isoamyl alcohol (25:24:1; pH 8.0) to the tube. Gently mix by inversion at least ten times. Do not vortex the tubes (*see Note 12*).
7. Centrifuge samples at ~ $16,058 \times g$ at room temperature for 5 min.
8. Transfer top aqueous phase from the phase-lock gel tube (~600 μ L) into a clean 1.5 mL reaction tube.

9. Add 1/10 volume of 3 M sodium acetate (pH 5.5) (~60 μ L) and 1 volume (~600 μ L) of -20 °C isopropanol to the tube containing the aqueous phase from the previous step. Mix thoroughly by inversion.
10. Store at -20 °C for at least 2 h or overnight. DNA precipitate should be visible immediately or after incubation. (*PAUSE POINT: The samples can be stored overnight at -20 °C*).
11. After incubation, centrifuge the tube containing DNA at ~16,058 $\times g$ at 4 °C for 20 min. Discard supernatant carefully without disturbing the DNA pellet.
12. Wash pellet with 1 mL of 100% ethanol. Vortex to loosen the DNA pellet from the tube. Centrifuge for 5 min at ~16,058 $\times g$ at 4 °C. Remove ethanol without disturbing the pellet.
13. Evaporate residual ethanol by placing the sample tube on a 55 °C heat block (*see Note 13*).
14. Add 150 μ L of TE (pH 8.0) buffer and incubate at 55 °C until DNA is completely dissolved. Gently vortex sample if needed.
15. Add 10 μ L of RNase A (100 mg/mL) to the DNA sample and incubate it for 5 min at room temperature.
16. Purify DNA using the PCR Purification kit as per manufacturer's protocol.
17. Quantify purified DNA using the Qubit® dsDNA BR or HS Assay Kit and the Qubit® 2.0 Fluorometer using the manufacturer's protocols (*see Note 14*).

3.2 Metagenomic Library Preparation

3.2.1 Plasmid Preparation

1. Prepare the pZE21 MCS-1 expression vector for ligation by linearizing at the *HincII* site using inverse PCR using the blunt-end HF DNA polymerase with the following reaction conditions (*see Notes 2 and 3*).
Mix the following components in a 50 μ L reaction volume.
 - 10.0 μ L 10x Polymerase reaction buffer,
 - 1.5 μ L 10 mM dNTP mix,
 - 1.0 μ L 50 mM MgSO₄,
 - 5.0 μ L Polymerase enhancer solution,
 - 1.0 μ L 100 pg/ μ L circular pZE21 DNA,
 - 0.75 μ L 10 μ M pZE21 linearizing forward primer,
 - 0.75 μ L 10 μ M pZE21 linearizing reverse primer,
 - 0.4 μ L HF DNA Polymerase, and
 - 29.6 μ L Nuclease-free H₂O to a final volume of 50 μ L.
2. Transfer 50 μ L of the above master mix to each PCR well. Run PCR on the thermocycler as follows: 95 °C for 5 min, then 35

cycles of [95 °C for 45 s, 55 °C for 45 s, 72 °C for 2.5 min], then 72 °C for 5 min (*see Note 15*).

3. Prepare 1% agarose gel by adding 1 g of agarose to 100 mL of 0.5× Tris-Borate-EDTA (TBE) buffer (10× stock concentration). Heat solution until agarose is completely dissolved. Let it cool to about 65 °C by placing at room temperature, and occasionally swirl the flask to let it cool evenly. Add SYBR green I DNA binding dye (10,000× in water) to a final concentration of 1× (10 µL). Prepare gel-casting tray with comb for sample loading and pour agarose solution into casting tray while it is still liquid. Allow to solidify. Place the gel in an electrophoresis chamber, add enough 0.5× TBE buffer to cover the surface of the gel, and remove the comb.
4. Add 6× gel loading dye to the PCR-amplified DNA. Run on a 1% low-melting-point agarose gel with SYBR green I DNA binding dye at 70 V for 120 min.
5. Excise gel slice corresponding to a ~2200 bp fragment and transfer to a clean tube. Purify DNA using the gel extraction kit. Elute DNA in 50 µL of nuclease-free molecular grade water (*see Notes 16 and 17*).
6. Dephosphorylate purified plasmid using calf intestinal phosphatase (CIP): For 50 µL reaction, add 40 µL of gel-purified DNA, 5 µL of CIP (10 U/µL), and 5 µL of the corresponding 10× reaction buffer. Incubate reaction mixture at 37 °C overnight and heat-inactivate reaction by incubating for 15 min at 70 °C (*see Note 18*).
7. Purify plasmid using the PCR purification kit as per manufacturer's protocol.
8. Quantify purified plasmid using Qubit® dsDNA HS Assay Kit and the Qubit® 2.0 Fluorometer. Store plasmid at -20 °C. Avoid multiple freeze-thaw cycles (*see Notes 14 and 19*).

3.2.2 Insert Preparation

1. Dilute up to 20 µg metagenomic DNA in EB buffer to a final volume of 200 µL. Shear the DNA to a size range of approximately 3 kb using a sonicator with manufacturer's recommended settings (*see Note 20*).
2. Add 6× gel loading dye to sheared DNA to a final concentration of 1×. Run sample through 1% low-melting point agarose gel with SYBR green I DNA binding dye as described above (70 V for 120 min). Excise a gel slice corresponding to 2–5 kb fragment size using a clean disposable knife (*see Notes 17, 21 and 22*).
3. Extract metagenomic DNA from the excised gel slice using the gel extraction kit and elute in 34 µL nuclease-free water.

4. Following purification, use END-It™ DNA End Repair kit to end repair the DNA: For each 34 µL volume of size-selected metagenomic DNA, add the following:

5 µL dNTP mix (2.5 mM),
5 µL ATP (10 mM),
5 µL 10× End-Repair Buffer, and
1 µL End-Repair Enzyme Mix to a final volume of 50 µL.

Mix gently and incubate at room temperature for 45 min. Heat-inactivate the reaction at 70 °C for 15 min.

5. Purify DNA using PCR purification kit. Elute the DNA with 30 µL of nuclease-free water.
6. Quantify purified DNA using the Qubit dsDNA HS Assay Kit and the Qubit 2.0 Fluorometer. Concentrate DNA using a vacuum concentrator or heat block at 55 °C to a final volume of 8–10 µL (*see Notes 14 and 23*).

3.2.3 Ligation and Dialysis

1. Perform ligation reaction using end-repaired metagenomic DNA and linearized vector using the Fast-Link™ DNA Ligation Kit using the following protocol. Maintain a 5:1 molar ratio of insert:vector for ligation. Prepare a 15 µL reaction tube with the following reagents:

1.5 µL 10× Fast-Link buffer,
0.75 µL 10 mM ATP,
1 µL Fast-Link DNA ligase (2 U/ µL),
Metagenomic DNA,
Linearized vector, and
Nuclease-free H₂O to a final reaction volume of 15 µL.

Along with each set of ligations, prepare a negative control ligation reaction without any insert (i.e., metagenomic DNA) (*see Notes 23–25*).

2. Incubate reaction at room temperature overnight.
3. Heat-inactivate reaction by incubating for 15 min at 70 °C.
4. After heat inactivation, dialyze ligation reactions as follows:
5. Fill clean petri dish with 20 mL of nuclease-free water. Place a 0.025 µm cellulose membrane on top of the water so that it floats. Carefully transfer entire volume of ligated product to the membrane and close lid. Incubate for 45–60 min and carefully collect sample in clean 1.5 mL tube. Use the full reaction volume for transformation (*see Note 26*).

*3.2.4 Electroporation,
Metagenomic Library
Amplification,
and Quantification*

1. Place a 0.1 cm-gap sterile electroporation cuvette, microcentrifuge tube, and ligated DNA on ice.
2. Thaw electrocompetent cells on ice. Mix by tapping gently (*see Note 4*).
3. Aliquot 25 μ L of electrocompetent cells to prechilled microcentrifuge tube on ice.
4. Add entire ligation reaction volume (~15 μ L) of sample to the aliquoted electrocompetent cells and stir briefly with pipet tip (*see Note 27*).
5. Perform electroporation using 0.1 cm cuvette with the following settings on electroporator: 10 μ F, 600 Ω , and 1800 V (*see Note 4*).
6. Within 10 s of pulse, add 975 μ L of recovery medium to cuvette and gently pipet up and down to resuspend cells. Transfer the cells and recovery medium to clean tubes.
7. Place the tube in a shaking incubator at 250 rpm for 1 h at 37 °C.
8. Repeat steps 4–7 for negative control ligation reaction.
9. After 1 h incubation, prepare 10⁻², 10⁻⁴, and 10⁻⁶ dilutions in LB-kanamycin broth of metagenomic sample libraries and negative control (no insert ligation). Also include a 10⁻¹ dilution for the negative control.
10. Plate 100 μ L of each dilution onto separate LB agar plates containing 50 μ g/mL kanamycin. Incubate plates overnight at 37 °C. The following day, count and record the number of colonies for each plate (*see Note 28*).
11. Inoculate the rest of the recovered cells into 50 mL of LB broth containing 50 μ g/mL kanamycin and grow overnight, shaking at 26 °C in an Erlenmeyer flask. Harvest the cells after optical density at 600 nm (OD₆₀₀) of the culture reaches 0.6–1.0.
12. The following day, centrifuge 50 mL overnight culture at 855 $\times g$ for 8 min to recover pellet. Discard liquid supernatant and resuspend pellet in 15 mL of LB broth containing 15% glycerol and 50 μ g/mL kanamycin. Aliquot metagenomic library into 2 mL Cryo-tubes and store at –80 °C for subsequent screening.
13. Pick 36 random colonies from the titer plates and resuspend in 50 μ L of nuclease-free water. Use this as a template for PCR reactions to estimate average insert size. Set up PCR reaction as follows: Prepare 25 μ L PCR reaction for each sample with the following reagents:
 - 2.5 μ L template DNA,
 - 2.5 μ L 10× reaction buffer,

1.0 μL 10 mM dNTP mix,
 1.0 μL 10 μM Colony PCR forward primer,
 1.0 μL 10 μM Colony PCR reverse primer,
 0.5 μL *Taq* DNA polymerase (5U/ μL), and
 16.5 μL nuclease-free water to bring the final reaction volume to 25 μL .

Use the following thermo-cycler settings to amplify DNA:
 94 °C for 10 min, 25 cycles of [94 °C for 45 s, 55 °C for 45 s, 72 °C for 5 min], and 72 °C for 10 min.

- Run PCR amplified fragments on 1% low melting point agarose as described above. Visualize DNA fragments using a gel imager and record the size of each fragment. Calculate the average insert size and estimate total library size using the following equation (*see Note 29*).

$$\text{Library Size (GB)} = \frac{\left\{ \text{TC} \times \left(\frac{\text{TR} - (\text{FR} + \text{NI} + \text{LS})}{\text{TR} - \text{FR}} \right) \times \text{AI} \right\}}{10^9}$$

where TC = total clones (cfu/mL of library determined from the titer plates), TR = total number of PCR reactions, FR = number of failed reactions, NI = number of colonies with no insert, determined by PCR reaction that produce ~300 bp band, LS = number of reactions that yield inserts less than 500 bp (estimated by subtracting 300 bp from band size), AI = average insert size after subtracting 300 bp from band size.

3.3 Screening for Antibiotic Resistance and Amplification of Antibiotic Resistance-Confering DNA Fragments

- Determine the minimum inhibitory concentration (MIC) of each antibiotic by plating negative control of the electrocompetent cells transformed with unmodified pZE21 on MH agar with 50 $\mu\text{g}/\text{mL}$ of kanamycin (MH-Kan) and additional antibiotic. Test each batch of antibiotic using negative control before screening library. Commonly used antibiotics and related MIC are listed in Table 1.
- Calculate the amount of library stock needed for screening. Adjust the concentration of frozen metagenomic library with LB-Kan broth such that each 100 μL aliquot of plating solution contains at least 10 times the total unique clones estimated in the library.
- To calculate the titer of the frozen metagenomic library stock, thaw one frozen aliquot on ice, prepare 10⁻², 10⁻⁴, and 10⁻⁵ dilutions, and plate on LB agar plates containing 50 $\mu\text{g}/\text{mL}$ kanamycin. Incubate plates overnight at 37 °C. Count and record the number of colonies on each plate.

4. Calculate amount of library stock needed for screening using the following formula:

$$\text{Amount of library stock required} = \left[\frac{\text{cfu / mL of library determined from the titer plates}}{\text{post electroporation} \times 10,000} \right] / \left[\frac{\text{titer of library stock determined in previous step (cfu / mL)}}{} \right]$$

(See Note 30).

5. Make enough diluted library stock to screen all antibiotic plates for each library. Prepare at least 100 μL extra to set up titer plates and compensate pipetting error.
6. Plate 100 μL of diluted library on MH agar with 50 $\mu\text{g}/\text{mL}$ of kanamycin (MH-Kan) and one additional antibiotic at the MIC for the negative control. Additionally, plate negative control of the electrocompetent cells transformed with unmodified pZE21 to ensure that the concentration of antibiotic used entirely inhibits the growth of clones with only pZE21 (without metagenomic insert). While setting up screening experiment, plate titers of your diluted library stock on LB-Kan plates to ensure that you plated stock with a titer in the expected range. 10^{-4} and 10^{-5} dilutions are appropriate.
7. Incubate plates for 24 h at 37 °C.
8. After incubation, inspect plates for any resistant colonies and record results.
9. Collect all resistant colonies by adding 750 μL of LB-Kan broth with 15% glycerol to the plates and gently scraping colonies with a sterile L-shaped cell-spreader. If required, repeat this step to collect any leftover colonies.
10. Collect the slurries of functionally selected clones in 2 mL Cryo-tubes by pipette aspiration and store at -80 °C.
11. To isolate the antibiotic-resistant metagenomic inserts, thaw the stock of antibiotic-resistant slurries from the above step on ice and aliquot 300 μL of cells into a new, clean 1.5 mL reaction tube. Pellet cells by centrifuging at $16,058 \times g$ for 5 min.
12. Discard supernatant. Gently wash pellet with 1 mL of nuclease-free H₂O and centrifuge at $16,058 \times g$ for 5 min.
13. Discard supernatant and resuspend cells in 30 μL nuclease-free H₂O. Freeze at -80 °C for 1 h and thaw to promote cell lysis.

14. Centrifuge lysed cells at $16,058 \times g$ for 2 min. Collect supernatant and use it as a template for amplification of resistance-conferring DNA fragments. For 25 μL PCR reactions, the following components are mixed together:

2.5 μL template DNA,
2.5 μL 10× polymerase reaction buffer,
0.5 μL 10 mM dNTP mix,
0.5 μL Taq DNA polymerase (5 U/ μL),
3.0 μL custom primer mix, and
16.0 μL nuclease-free H_2O to bring the final reaction volume to 25 μL .

The custom primer mix consists of three forward and three reverse primers. The following is the volume added from each primer stock solution (10 mM) to prepare the mix (*see Note 31*).

0.5 μL of primer F1,
0.5 μL of primer F2,
0.5 μL of primer F3,
0.21 μL of primer R1,
0.43 μL of primer R2, and
0.86 μL of primer R3

15. Perform PCR reaction using the following cycling conditions: 94 °C for 10 min, 25 cycles of [94 °C for 45 s, 55 °C for 45 s, 72 °C for 5.5 min], and 72 °C for 10 min.
16. Run PCR products on 1% agarose gel to confirm amplification.
17. Perform purification and quantification of amplified metagenomics inserts using a PCR Purification kit and Qubit dsDNA HS assay kit, respectively. Follow manufacturer's protocol for detailed instructions.

3.4 Illumina Library Preparation and Sequencing

1. Dilute 500–2000 ng of PCR-amplified metagenomics inserts from each selection to a total volume of 130 μL in EB buffer in a 96-well PCR plate and shear to 150–200 bp using a sonicator (*see Note 32*).
2. Purify and concentrate sheared DNA using the MinElute PCR Purification Kit and elute the DNA in 20 μL 55 °C nuclease-free H_2O . Use eluted DNA as input for Illumina library preparation.
3. End-repair sheared DNA by mixing 20 μL of eluted DNA with 2.5 μL T4 DNA ligase buffer, 1 μL 1 mM dNTPs, 0.5 μL T4 DNA polymerase, 0.5 μL T4 Polynucleotide Kinase, and 0.5 μL Taq DNA polymerase for a total reaction volume of 25 μL .

4. Incubate reaction mixture in a thermocycler at 25 °C for 30 min followed by 20 min at 75 °C.
5. To each end-repaired sample, add 5 µL of 1 µM pre-annealed barcoded sequencing adapters and 0.8 µL of T4 DNA ligase (*see Note 5*).
6. Incubate the reaction mixture on a thermocycler at 16 °C for 40 min followed by 10 min at 65 °C.
7. Run each sample on 2 % agarose gel, stained with SYBR green I DNA binding dye as previously described, in 0.5× TBE buffer at 120 V for 2 h. Add 6× loading dye to DNA before loading on the gel.
8. Remove gel slice corresponding to 300–400 bp using a clean disposable knife and purify DNA using the MinElute Gel Extraction kit. Elute into 12 µL Buffer EB.
9. Set up PCR reaction to enrich purified DNA using 12.5 µL 2× Phusion® High Fidelity Master Mix and 0.5 µL of 10 µM Illumina PCR forward Primer and 0.5 µL of 10 µM Illumina PCR reverse Primer in a 25 µL reaction using 2 µL of purified DNA as a template. Amplify DNA at 98 °C for 30 s followed by 18 cycles of [98 °C for 10 s, 65 °C for 30 s, and 72 °C for 30 s] with a final extension of 5 min at 72 °C.
10. Run each sample on 2 % agarose gel in 0.5× TBE, stained with SYBR green I DNA binding dye as previously described.
11. Remove gel slice in 300–400 bp size range using clean disposable knife and purify DNA using the MinElute Gel Extraction kit.
12. Perform DNA quantification using Qubit dsDNA HS assay kit.
13. Dilute each sample to 10 nM, and combine an equal volume of each 10 nM sample in one tube for Illumina HiSeq/MiSeq (*see Note 33*).

3.5 Computational Analysis of Sequencing Reads

We have developed a computational pipeline, PARFuMS as previously described in [2], for high-throughput assembly and annotation of resistance-conferring DNA fragments obtained from many independent functional selections. Below we have described a general workflow of the pipeline. It is important to note that the steps mentioned below include modifications that were made after the publication.

3.5.1 Demultiplexing and Preprocessing of Sequencing Reads

After sequencing, Illumina paired-end sequencing reads are demultiplexed based on barcode sequences. Each read is assigned to a sample-specific file by exact barcode matching via a mapping file. This step generates several smaller sequencing files, such that assembly and annotation for each sample can be performed in parallel.

Subsequently, reads corresponding to either expression vector or Illumina adapter sequences are removed or trimmed before proceeding to assembly (*see Note 34*).

3.5.2 De Novo Assembly of Short Read Sequences

In order to carry out functional characterization of the metagenomic insert library, quality-filtered short read sequences need to be assembled into longer contiguous sequences, commonly called contigs. Several assembly programs have been specifically developed to carry out this task, such as Meta-IDBA [15], Meta-Velvet [16], and InteMap [17]. We have implemented an iterative assembly approach as also described in [2]. In this method, short reads are first assembled into intermediate length contigs using three iterations of the short-read assembler Velvet [18]. Following each round of assembly, redundant contigs are collapsed to one sequence using CD-HIT [19] and chimeric sequences are removed by mapping raw reads against assembled contigs using FR-HIT [20]. The first iteration of Velvet takes all reads as an input, while in the second and third rounds, the reads not present in previously assembled contigs are utilized. The velvet-assembled, nonredundant contig set is then passed to the long-read assembler Phrap for two iterations. The first iteration assembles the Velvet output into more complete contigs that are subsequently linked together if two contigs are bridged by a sufficient number of raw paired-end reads. The final iteration of Phrap uses these linked contigs as input and provides a more complete assembly of linked contigs that are subsequently annotated using Resfams [21].

3.5.3 Annotation of Assembled Contigs with Antibiotic Resistance Functions

1. Identification of open reading frames in assembled contigs can be achieved by using gene prediction tools such as MGC [22], Metagenemark [23], MetaGenAnnotator [24], and GLIMMER-MG [25] (*see Note 35*).
2. Download the latest version of the Resfams database and supporting datafiles from <http://dantaslab.wustl.edu/resfams> (*see Note 36*).
3. Download the latest version of HMMER3 [26] from <http://hmmer.org/>; and follow instructions provided in the documentation to install HMMER3.
4. Run the hmmscan function of HMMER3 to annotate protein sequences using the Resfams database (Resfam.hmm) with the following parameters: -cut_ga, -tblout. Below is the basic usage to search against the Resfams database:

```
hmmscan --cut_ga --tblout -o <output_file>
Resfam.hmm <protein_seqs.fasta>
```
5. The output will list translated ORFs (antibiotic resistance proteins) present in functional selections (*see Note 37*).

4 Notes

1. Check pH of the packaged product before use. Some products are packaged at pH 6.7 and come with separate alkaline buffers. Add sufficient volume of alkaline buffer to Phenol: Chloroform:Isoamyl alcohol (25:24:1) (pH 6.7) to achieve pH 8.0. Store at 4 °C.
2. Other expression plasmids may be used to construct the library. While selecting another plasmid consider compatibility, copy number, and selective marker. Selecting different plasmid will also require redesigning of primers (for plasmid linearization, colony PCR, sequencing library preparation, etc.), and reestablishment of minimum inhibitory concentration for each antibiotic for *E. coli* with the new empty plasmid.
3. It is critical to use a high fidelity polymerase for efficient amplification of DNA fragments with a low error rate. The given reaction volumes and conditions for PCR in this protocol are optimized for Platinum® *Pfx* DNA Polymerase and may differ for other high fidelity polymerases. For a different polymerase, adjust PCR reaction conditions as per the manufacturer's recommendation.
4. Electroporation settings and minimum inhibitory concentration (Table 1) are optimized for *E. coli* E. cloni 10G SUPREME cells ($\geq 4 \times 10^{10}$ cfu/ μg transformation efficiency) (Lucigen Corporation, Middleton, WI, USA). If using different electrocompetent cells, follow manufacturer's instructions for optimal electroporation condition. Reestablish MIC against each antibiotic with electrocompetent cells transformed with empty plasmid.
5. The adapters consist of unique 7 bp oligonucleotide sequences (barcodes) specific to each sequencing sample (e.g., one barcode used per antibiotic selection plate), facilitating the demultiplexing of barcoded reads on a sequencing run to enable separate assembly of contigs corresponding to each sample. Forward and reverse sequencing adapters are annealed by heating 1 μM mixture to 95 °C followed by slow cooling (0.1 °C per second) to a final holding temperature of 4 °C. Store pre-annealed barcoded adapters at -20 °C.
6. Carefully read the MSDS for each antibiotic and store antibiotics accordingly. Some antibiotics need to be stored in a refrigerator or a desiccator (or both). If antibiotics are light-sensitive, store them in a dark place. Expiration dates for solutions are earlier than for the stock powder. Antibiotic stock solutions prepared in ultrapure water need to be filter-sterilized. Antibiotic solutions prepared in organic solvents do not need filter sterilization. It is advisable to prepare stock solutions just before use and, if necessary, store at -20 °C.

7. Pre-warm the water bath to 55 °C. After removing the media from the autoclave, allow it to cool to 55 °C by placing in a pre-warmed water bath. Add the appropriate amount of the desired antibiotic to this media and pour about 20 mL of media per 10 cm sterile petri dish. When pouring plates, keep your bench area sterile and clean.
8. If your lab has premixed powder, use the amount suggested by the manufacturer.
9. Although 10–20 µg of starting DNA is required for one metagenomic library preparation, it is recommended to extract DNA sufficient for at least three metagenomic library preparations (50–60 µg). Insert yield may vary from sample to sample, and lower yield may hinder the process of library preparation. Thus, it is advisable to begin the library preparation with 10–20 µg of DNA and if insert yield is insufficient, the insert preparation step can be repeated to achieve the requisite amount. It is not recommended to perform the complete DNA extraction process again for a subset of samples as that may introduce extraction-specific biases into the study. Also, DNA preparations from too much fecal sample may have inhibitors that will interfere with downstream steps.
10. Close cap tightly to avoid any leakage during bead beating.
11. Continuous bead beating for 4 min may overheat samples. Cooling samples before and during bead beating helps prevent heat-induced DNA damage.
12. Use of phase lock gel tubes helps eliminate interphase protein contamination during phenol extraction. The organic phase and the interphase materials are effectively trapped in or below the gel, thus allowing easy removal of the top aqueous phase containing DNA by pipetting or decanting.
13. Do not over dry DNA pellet; dried DNA is sometimes difficult to dissolve.
14. DNA quantification using spectrometric methods may overestimate DNA concentration due to the presence of other compounds in the solution. Measurement of DNA concentration using a fluorimeter thus may give a more accurate concentration.
15. Do not use a high amount of circular plasmid as template. Carryover circular plasmid (template) may give a high background of self-ligated plasmid (plasmid without insert) during metagenomic library preparation.
16. Gel purification and size selection is preferred over a simple PCR cleanup to remove any primer dimers and other forms of plasmid. During gel extraction, try to avoid DNA differing from 2200 bp in size. In the final step of the gel extraction protocol, elute DNA with nuclease-free water instead of

elution buffer. Elution buffer contains salts that may interfere with the following steps of plasmid preparation. Optional: Further, purify plasmid DNA using a PCR purification kit to concentrate the DNA sample. This step may reduce the amount of CIP needed for dephosphorylation step.

17. Use conversion screen and filter while acquiring images of DNA electrophoresis samples to avoid exposing samples or users to harmful UV radiation. Use face and eye protection while working under UV lights.
18. Dephosphorylation removes the terminal 5'-phosphate group from DNA and thus suppresses self-ligation and circularization of linearized plasmid DNA. This is a very important step for high-efficiency plasmid preparation. Maintain a proper buffer ratio for optimal results.
19. Prepare enough plasmid for one project. Different batches of pZE21 prepared in this way have different ligation efficiencies, even when the same protocol has been followed for all. Quality control of each batch is required before using any new batch of plasmid. The easiest way to do this is by preparing an insert library from salmon sperm DNA and ligating it into your new vector. Once you confirm the ligation efficiency (ideally >70% of transformants should contain insert), you may start with the real ligation.
20. If using Covaris E210, we recommend the following settings: Duty cycle: 20%, Intensity: 0.1, Cycles per burst: 1,000, Treatment time: 600 s.
21. Use a separate gel box for each sample and clean the gel box before use with 10% bleach for 10 min. Wash gel boxes at least three times with dH₂O between each run to avoid cross-contamination.
22. Some commercially available molecular DNA markers (ladders) consist of various DNA fragments and may contain ARGs. Make sure that the molecular marker used during gel purification does not contain any ARGs. You can also prepare a custom molecular DNA marker by amplifying specific sizes of nonbacterial DNA.
23. 200 ng or more insert is required for ligation. If a sufficient amount of insert is not obtained, repeat insert preparation step (as mentioned in Note 5). It is also advisable to set up two separate ligation reactions if the total amount of insert is more than 1 µg.
24. When setting up multiple ligation reactions, it is preferable to make a master mix of ligation reagents to avoid pipetting error.
25. In this case, average insert size and plasmid size are similar (~2200 bp); thus, the mass ratio approximates the molar ratio.

26. Dialysis of samples removes salt traces. Higher salt concentrations in the DNA sample may cause arcing during electroporation. Make sure that no part of the cellulose membrane sinks into the water. If multiple samples are applied on the same membrane, keep at least 2 cm space between each sample to avoid contamination.
27. Do not vortex or pipet up and down to mix the sample. This can introduce air bubbles and warm the cells. Air bubbles in sample may cause arcing during electroporation.
28. The pZE21 expression vector has the kanamycin resistance cassette.
29. Colony PCR primers are designed such that they will amplify insert at the cloning site and approximately 150 bp of plasmid on each side. Thus, when no insert is present, these primers will yield a band of 300 bp in size. This serves as a colony PCR control. Do not count PCR reactions in any calculation that failed to amplify the 300 bp fragment. While calculating fragment size, deduct 300 bp from the size on the gel. For example, if a 1500 bp band is observed on the gel, the actual fragment size is 1200 bp. Repeat colony PCR if more than 20% of reactions failed to amplify.
30. If the concentration of the frozen library stock is higher than desired, dilute library with LB broth with 50 µg/mL kanamycin; if it is lower than the desired concentration, pellet cells by gentle centrifugation at 855 g and reconstitute in an appropriate volume of LB-Kan broth. Dilute the libraries in LB broth with 50 µg/mL kanamycin so that 100 µL plating solution contains about 10 copies of each clone. For example, if colony count following electroporation indicates that the metagenomic library contains 500,000 clones, prepare 100 µL of plating solution that contains (10*500,000) clones. To avoid underestimation of amount of library stock you need to plate, use the highest estimated colony count following electroporation and the lowest estimated titer of your library stock in the formula.
31. The staggered primer mix ensures diverse nucleotide composition during early Illumina sequencing cycles.
32. If using Covaris E210, we recommend the following settings: Duty cycle: 10%, Intensity: 5, Cycles per burst: 200, and Treatment time: 180 s.
33. Contact your local sequencing center for specific DNA amount required for sequencing run.
34. We use cross_match from the Phrap package with the following options to remove vector and adapter sequences:
`-gap1_only -minmatch 6 -minscore 10 -gap_init -3`

35. We use the standalone version of Metagenemark with default parameters to identify open reading frames.
36. Resfams is a curated database of protein families and their associated profile HMMs that are confirmed for antibiotic resistance function using experimental methods and are organized in ontology. It is extensively used for high-throughput annotation of sequence-novel ARGs. There are two variants of the Resfams database: Resfam-core.hmm and Resfam-full.hmm. The core database of Resfams profile HMMs was trained using unique antibiotic resistance protein sequences from the Comprehensive Antibiotic Resistance Database (CARD) database, the Lactamase Engineering Database (LacED), and Jacoby and Bush's collection of curated beta-lactamase proteins. The core database of Resfams profile HMMs is supplemented with additional profile HMMs from the Pfam and TIGRFam databases to generate the full Resfams profile HMM database. The full version of the Resfams database should only be used when there is previous functional evidence of antibiotic resistance activity, such as in functional metagenomic selections.
37. The option “`--cut_ga`” requires that genes meet profile-specific gathering thresholds (rather than a global, more permissive, default log odds cutoff) before receiving annotation. Check HMMER3 documentation for detailed instructions on hmmscan.

Acknowledgments

The original implementation of functional metagenomics selections for interrogating resistance genes was described by J. Handelsman and colleagues in 1998. We thank A. Moore and B. Wang for protocol optimization of high-throughput versions of this method presented in this manuscript, A. Reyes and K. Forsberg for computational implementation of PARFuMS, M. Gibson for development of the Resfams database, and members of the Dantas lab for helpful general discussions and for comments on the manuscript.

This work was supported in part by the NIH Director’s New Innovator Award (<http://commonfund.nih.gov/newinnovator/>), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK:<http://www.niddk.nih.gov/>), and the National Institute of General Medical Sciences (NIGMS: <http://www.nigms.nih.gov/>), of the National Institutes of Health (NIH) under award numbers DP2DK098089 and R01GM099538 to G.D. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

1. Levy SB, Marshall B (2004) Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* 10. doi: [10.1038/nm1145](https://doi.org/10.1038/nm1145)
2. Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MO, Dantas G (2012) The shared antibiotic resistome of soil bacteria and human pathogens. *Science* 337(6098):1107–1111. doi: [10.1126/science.1220761](https://doi.org/10.1126/science.1220761)
3. Perron GG, Whyte L, Turnbaugh PJ, Goordial J, Hanage WP, Dantas G, Desai MM (2015) Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PloS One* 10(3). doi: [10.1371/journal.pone.0069533](https://doi.org/10.1371/journal.pone.0069533)
4. Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J, Handelsman J (2010) Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 8(4):251–259. doi: [10.1038/nrmicro2312](https://doi.org/10.1038/nrmicro2312)
5. Sommer MOA, Church GM, Dantas G (2010) The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* 1(4):299–303. doi: [10.4161/viru.1.4.12010](https://doi.org/10.4161/viru.1.4.12010)
6. Monier J-M, Demanèche S, Delmont TO, Mathieu A, Vogel TM, Simonet P (2011) Metagenomic exploration of antibiotic resistance in soil. *Curr Opin Microbiol* 14(3):229–235. doi: [10.1016/j.mib.2011.04.010](https://doi.org/10.1016/j.mib.2011.04.010)
7. D'Costa VM, McGrann KM, Hughes DW, Wright GD (2006) Sampling the antibiotic resistome. *Science* 311(5759):374–377. doi: [10.1126/science.1120800](https://doi.org/10.1126/science.1120800)
8. Pérez-Pérez FJ, Hanson ND (2002) Detection of plasmid-mediated AmpC β -lactamase genes in clinical isolates by using multiplex PCR. *J Clin Microbiol* 40(6):2153–2162. doi: [10.1128/jcm.40.6.2153-2162.2002](https://doi.org/10.1128/jcm.40.6.2153-2162.2002)
9. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA (2005) Diversity of the human intestinal microbial flora. *Science* 308(5728):1635–1638. doi: [10.1126/science.1110591](https://doi.org/10.1126/science.1110591)
10. Daniel R (2005) The metagenomics of soil. *Nat Rev Micro* 3(6):470–478. doi: [10.1038/nrmicro1160](https://doi.org/10.1038/nrmicro1160)
11. Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5(10):R245–R249
12. Pehrsson EC, Forsberg KJ, Gibson MK, Ahmadi S, Dantas G (2013) Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs. *Front Microbiol* 4:145. doi: [10.3389/fmicb.2013.00145](https://doi.org/10.3389/fmicb.2013.00145)
13. Riesenfeld CS, Department of Plant Pathology, Microbiology Doctoral Training Program a, Goodman RM, Department of Plant Pathology, Microbiology Doctoral Training Program a, Gaylord Nelson Institute for Environmental Studies UoWM, Madison, WI 53706, USA, Handelsman J, Department of Plant Pathology, Microbiology Doctoral Training Program a (2015) Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol* 16(9):981–989. doi: [10.1111/j.1462-2920.2004.00664.x](https://doi.org/10.1111/j.1462-2920.2004.00664.x)
14. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 25(6):1203–1210
15. Peng Y, Leung HC, Yiu SM, Chin FY (2011) Meta-IDBA: a de novo assembler for metagenomic data. *Bioinformatics* 27(13):i94–i101. doi: [10.1093/bioinformatics/btr216](https://doi.org/10.1093/bioinformatics/btr216)
16. Namiki T, Hachiya T, Tanaka H, Sakakibara Y (2012) MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res* 40(20):e155. doi: [10.1093/nar/gks678](https://doi.org/10.1093/nar/gks678)
17. Lai B, Wang F, Wang X, Duan L, Zhu H (2015) InteMAP: integrated metagenomic assembly pipeline for NGS short reads. *BMC Bioinformatics* 16:244. doi: [10.1186/s12859-015-0686-x](https://doi.org/10.1186/s12859-015-0686-x)
18. Zerbino DR (2010) Using the Velvet de novo assembler for short-read sequencing technologies. Current protocols in bioinformatics/editorial board, Andreas D Baxevanis [et al] CHAPTER:Unit-11.15. doi: [10.1002/0471250953.bi1105s31](https://doi.org/10.1002/0471250953.bi1105s31)
19. Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28(23):3150–3152. doi: [10.1093/bioinformatics/bts565](https://doi.org/10.1093/bioinformatics/bts565)
20. Niu B, Zhu Z, Fu L, Wu S, Li W (2011) FR-HIT, a very fast program to recruit metagenomic reads to homologous reference genomes. *Bioinformatics* 27(12):1704–1705. doi: [10.1093/bioinformatics/btr252](https://doi.org/10.1093/bioinformatics/btr252)
21. Gibson MK, Forsberg KJ, Dantas G (2015) Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* 9(1):207–216. doi: [10.1038/ismej.2014.106](https://doi.org/10.1038/ismej.2014.106)

22. El Allali A, Rose JR (2013) MGC: a metagenomic gene caller. *BMC Bioinformatics* 14(Suppl 9):S6. doi:[10.1186/1471-2105-14-S9-S6](https://doi.org/10.1186/1471-2105-14-S9-S6)
23. Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res* 38(12):e132. doi:[10.1093/nar/gkq275](https://doi.org/10.1093/nar/gkq275)
24. Noguchi H, Taniguchi T, Itoh T (2008) MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res* 15(6):387–396. doi:[10.1093/dnares/dsn027](https://doi.org/10.1093/dnares/dsn027)
25. Kelley DR, Liu B, Delcher AL, Pop M, Salzberg SL (2012) Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic Acids Res* 40(1):e9. doi:[10.1093/nar/gkr1067](https://doi.org/10.1093/nar/gkr1067)
26. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 41(12):e121. doi:[10.1093/nar/gkt263](https://doi.org/10.1093/nar/gkt263)

Chapter 20

Epidemiological Surveillance and Typing Methods to Track Antibiotic Resistant Strains Using High Throughput Sequencing

Miguel Paulo Machado, Bruno Ribeiro-Gonçalves, Mickael Silva,
Mário Ramirez, and João André Carriço

Abstract

High-Throughput Sequencing (HTS) technologies transformed the microbial typing and molecular epidemiology field by providing the cost-effective ability for researchers to probe draft genomes, not only for epidemiological markers but also for antibiotic resistance and virulence determinants. In this chapter, we provide protocols for the analysis of HTS data for the determination of multilocus sequence typing (MLST) information and for determining presence or absence of antibiotic resistance genes.

Key words High-throughput sequencing, Microbial typing, Gene finding, Online databases, Antibiotic resistance

1 Introduction

The apparent inexorable rise in antibiotic resistance registered in multiple bacterial species has been a cause of increasing concern [1]. While resistance can be disseminated horizontally by mobile genetic elements, it is also well known that for a number of bacterial species, the rise in antibiotic resistance is due to the rapid spread and success of a few lineages that are frequently multidrug resistant. The ability to track lineages or clones provides not only important information for clinical microbiologists and epidemiologists to target interventions, but also for researchers determining the causes of increasing resistance or studying the dissemination of resistance mechanisms.

In recent years, the development of High-Throughput Sequencing (HTS) methodologies, leapfrogged the capacity of researchers to accurately and unequivocally identify bacterial clones since in a couple of days one is able to determine the sequence of 99 % of the genome of a bacterial strain, given one has access to an

axenic culture. This technology also allowed for an increase in discriminatory power in the identification of bacterial strains, a field known as Microbial Typing.

1.1 Microbial Typing Methods

Over the last 30 years several methods have been developed with the goal of discriminating bacterial strains. These have been used in outbreak investigation, for surveillance of antibiotic resistant strains, in the study of bacterial population dynamics and their response to human interventions and for characterizing the natural history of persistent infections [2]. The first methods were purely phenotypic, such as the assessment of antibiotic resistance (antibio-typing) or susceptibility to lysis by panels of different phages (phage typing) [3]. With the introduction of gel-based molecular biology techniques, several typing methods were developed based on band patterns after digestion with infrequent cutting endonucleases of an amplified target DNA (Restriction Fragment Length Polymorphism analysis –RFLP) or whole cell DNA (Pulsed-Field Gel Electrophoresis- PFGE). For over a decade, PFGE was the gold standard method for the identification of bacterial clones since, although time-consuming and labor-intensive, it was the most discriminatory and reliable technique. However, in 1998 Maiden and colleagues proposed a novel method based on the sequence of seven housekeeping genes to study the population structure of *Neisseria meningitidis* [4]. This method, which they designated Multilocus Sequence Typing (MLST), provided the first truly portable approach to microbial typing due to its use of sequence data. Contrary to the gel-based band patterns, sequence data could be unambiguously compared between laboratories without resorting to expensive and frequently proprietary specialized image analysis software. In addition, MLST can be readily applied to several bacterial and eukaryotic species using specific targets for each species. Attesting to the success of MLST, the largest online repository of MLST information hosted by University of Oxford, PubMLST has schemas and databases for 108 bacterial species and 9 eukaryotic species, (accessed in 26/01/2016), many of which store information on tens of thousands of strains [5].

Traditionally, MLST is based on the specific amplification by PCR of 400–700 bp internal fragments of seven housekeeping genes followed by Sanger sequencing. Since MLST analyzes only a fragment of the gene and not the complete gene, each sequence is usually referred to as locus. The sequence of each locus is compared to an online database and an allele number is attributed to each unique sequence. If a novel sequence is found, the curator of the database validates it as a novel allele based on the submitted trace files and attributes it a new allele identifier. A Sequence Type (ST) identifies a unique combination of alleles at each of the seven loci. Although one could use standard phylogenetic analysis methods based directly on the sequence of different loci, for microbial typing purposes the comparison between STs is fre-

quently done based on the number of different alleles between allelic profiles, independently of the actual number of differences found at sequences level. In this type of analysis and relative to a given ST, a Single Locus Variant (SLV) is any ST that only has one different allele in any of the loci and a Double Locus Variant (DLV) is any ST that has two different alleles in any of the loci. Although these are the two most frequently used terms, the same naming logic can be expanded to any number of differences, these being generically referred to as Locus Variants (LV).

In order to infer patterns of evolutionary descent among related STs, Feil et al. [6] proposed the eBURST algorithm for the analysis of MLST data, which allows for the creation of unrooted trees using the allelic profile data and to define Clonal Complexes (CCs) of closely related strains. More recently, Francisco et al. suggested improvements to eBURST by proposing the global optimal eBURST (goeBURST) algorithm [7], guaranteeing a global optimal and unique solution for the resulting trees. This work also established the close relationship between the application of eBURST and the analysis using Minimum Spanning Trees. The PHYLOViZ software [8], which allows for the overlay of epidemiological data on the resulting trees, allows users to perform a goeBURST analysis while also extending the eBURST algorithm to allow the analysis of any number of loci. Bayesian clustering algorithms are also increasingly used and BAPS is a popular software implementing such an approach [9]. Detailed protocols on how to perform MLST and analyze the data using BAPS are also available [10].

Several other sequence-based typing methods were proposed, always enabling the comparison of the DNA sequence of a local strain with online databases to attribute a universal classification and allow comparison between different studies. Among these, two single locus methods remain widely used: *spa* typing for *Staphylococcus aureus* [11], based on the sequence of repeats on the *spa* gene, and *emm* typing for *Streptococcus pyogenes*, based on the variation of a region of the *emm* gene encoding the serospecificity of the M protein [12].

Multilocus Variable Number of Tandem Repeat Analysis (MLVA) [13] is another multilocus method that found application in various bacterial species. MLVA is based on simultaneously evaluating the number of repeats found in a set of different loci. Although it is in essence a sequence-based method, the sequencing of the loci is not an absolute requirement since the number of repeats can be inferred by the length of the amplified product determined by capillary electrophoresis. In Table 1 are shown a list of available websites that host schemas and analysis options for the most relevant sequence-based typing methods. More comprehensive reviews on microbial typing methods and data analysis methodologies can be found in the recent literature [14, 15].

Table 1
List of the most used sequence-based typing methods online databases

Method	Database	URL
MLST/cgMLST/wgMLST	PubMLST Institut Pasteur MLST Enterobase	http://www.pubmlst.org http://bigsdb.web.pasteur.fr/ https://enterobase.warwick.ac.uk/
MLVA	MLVAbank MLVA Bacterial Genotyping MLVA.net	http://mlva.u-psud.fr/mlvav4/genotyping/ http://www.mlva.eu http://www.mlva.net
Spa typing	Ridom Spa-Server	http://spaserver.ridom.de/

1.2 High-Throughput Sequencing Methodologies in Microbial Typing

HTS technologies, and in particular benchtop sequencers, are transforming microbial pathogen research. This technology allows for a paradigm shift in the field of Microbial Typing and Molecular Epidemiology, since now researchers can readily assess variation in hundreds or thousands of targets in the genome simultaneously, instead of focusing in a single or only a few targets, as was the case with MLST or MLVA.

HTS technologies generate large sets of reads from DNA or RNA of a given sample. A read can be defined as a short DNA sequence, which can vary in length between 70 base pairs (bp) and 500 bp per run for short-read technologies (454, Illumina and Ion Torrent), but that can be as high as 200 kb for long-read technologies (PacBio and Oxford Nanopore).

Several competing technologies are available on the market, with different throughput capability (number of reads per run per hour), read length, and operation cost. An extensive description of the different technologies available is beyond the scope of this chapter and interested readers are referred to a recent review on the subject, focusing on bacterial genome sequencing [16]. More technical reviews on the comparison of the performance of benchtop sequencers for short-read technologies are also available [17, 18].

The assembly of the reads of a bacterial chromosome from short-read technologies usually results in several nonoverlapping contiguous fragments of sequence (contigs). These vary in size and number according to the technology used, the target genome coverage, the quality of the reads, and the assembly software used. In contrast, long-read technologies have the potential to provide the complete sequence of a bacterial chromosome as well as of any plasmids present. This is due to the fact that the length of a single large-read is able to bridge areas that short-read technologies struggle with, such as repeat regions larger than the read length. However, long-read technologies currently have a higher base

calling error rate and their cost is still higher per strain to obtain information about 99% of the genome than short-read technologies.

Currently, the challenge has shifted from the production of good quality sequencing data to the bioinformatics analysis of the generated data. In order to extract relevant genomic information from the reads, two major strategies can be employed: de novo assembly or reference mapping.

De novo assembly algorithms use methodologies to assemble the reads into longer contigs, without resorting to a reference genome. They use several different strategies to determine how the reads overlap to reconstruct a longer contiguous sequence with a confidence level determined by the read depth (coverage), which also conditions the error rates of the reconstructed sequence. A comprehensive benchmark of different assemblers on the context of the analysis of bacterial HTS data was recently presented [19]. De novo assembly strategies offer several advantages, such as: (1) providing a set of contigs when a finished genome is not available for reference mapping, (2) recovering genomic regions not present in the chosen reference (large insertions), and (3) identifying putative chromosomal or gene rearrangements. However, de novo assemblies require high sequencing coverage to produce the best contigs, they are prone to inaccuracies due to gene duplications, repetitive sequences or multiple copies of the same gene or mobile genetic element, and, most importantly, require high computing power and significant time to produce results [20].

Since de novo assemblies provide a draft genome that can be annotated using available software such as Prokka [21], these can be used to search for specific genes or reveal the presence of novel genes. Traditional phylogenetic approaches can also be used to analyze these draft genomes. In addition, in microbial typing draft genomes can be used in gene-by-gene approaches [22], which revisit the MLST concepts but extend them to multiple loci across the whole genome (wgMLST) or using only loci present in a core genome shared by most strains of a given species (cgMLST). A simple wgMLST approach, instead of using only a limited number of housekeeping genes, considers the entire catalog of genes recovered by annotating the contigs in a given draft genome to probe for gene presence and absence. In contrast, in cgMLST the allelic variation of the loci in the core genome is analyzed using the same approach used in traditional MLST. In the future, hybrid methodologies combining both approaches will certainly maximize the information we can derive from draft genomes. The freely available BIGsDB software [5] pioneered this approach and comprehensive wgMLST schemas for *N. meningitidis* and *Campylobacter* sp. are already hosted at the PubMLST database. This approach has also been adopted by commercial software such as Ridom Seqsphere+ from Ridom GmbH (<http://www.ridom.de/>) and Bionumerics

from Applied Maths (<http://www.applied-maths.com/>). Both these software offer a full suite of tools for using gene-by-gene approaches on HTS data.

Reference mapping algorithms rely on a different approach by using diverse strategies to align each read to a reference genome (or contigs) and determine any variation between the resulting alignment of the available reads and the reference sequence. These approaches usually report Single Nucleotide Polymorphisms (SNPs) and small insertions and deletions (indels). The confidence level of the call of variations is heavily dependent on the coverage at the variable region. Given the algorithmic complexity and the need for process optimization in terms of running time, several dozens of algorithms for reference mapping are available, each employing a different strategy to improve the compromise between performance and accuracy [23]. For bacterial strain discrimination and identification, this approach, being faster than de novo assembly, has been extensively used in several studies [24–26] as a way to determine SNPs and indels for subsequent phylogenetic analysis using traditional approaches. This can be particularly useful in outbreak analysis studies due to the low expected variation between outbreak related isolates.

However, for longitudinal population-based studies, this approach is best applied to monomorphic, infrequently recombining species with existing complete reference genomes, since these typically present a stable core genome that can be readily used as reference and that guarantees a good coverage of the genome of each strain [27]. More panmictic and highly recombinogenic species have a smaller core genome, and the choice of an adequate reference genome to use for mapping to maximize the indexing of variation within a bacterial species can prove challenging. The way this problem has been solved is by choosing multiple references but this also creates difficulties for the downstream analyses.

1.3 Antibiotic Resistance Determinants Databases

The intensive use of antibiotics in different areas of human activity has led to a strong selective pressure over the microorganisms, leading to an increase in antibiotic-resistant bacteria. Many different mechanisms of antibiotic resistance are known, ranging from single nucleotide mutations in particular genes, extensive remodeling of existing genes or the acquisition of novel genes. The latter can be easily spread through horizontal gene transfer since these can be found in plasmids or other mobile genetic elements, while recombination with exogenous DNA occurs in most species but is particularly important in naturally competent bacteria. On the other hand, expansion of resistant clones can also play a key role in increasing resistance. The construction of comprehensive databases that would collect data on the genetic mechanisms underlying antibiotic resistance, allowing researchers to easily

access and share this information, would be essential resources for resistome prediction from HTS data.

Presently, several publicly available antibiotic-resistant genes repositories are available, covering different antibiotic classes and with different tools to access the data. The Comprehensive Antibiotic Gene Database (CARD) (<http://arpcard.mcmaster.ca/>) [28] is an exhaustive knowledge resource regarding antibiotic resistance genes from a variety of organisms, genomes, and plasmids. The Antibiotic Resistance Genes Database (ARDB) (<http://ardb.cbcn.umd.edu/index.html>) [29] is a manually curated repository covering resistance genes for all antibiotic classes. The ResFinder repository from the Center for Genomic Epidemiology (<https://cge.cbs.dtu.dk//services/ResFinder/>) [30] provides an exhaustive information resource on antibiotic resistance genes from sequenced bacterial genomes. Compared to other databases, the ResFinder webtool accepts both HTS sequence reads and de novo assembled sequences. However, it only offers information about antibiotic resistance that results from the acquisition of specific resistance genes. The Lactamase Engineering Database (LacED) (<http://www.laced.uni-stuttgart.de/>) [31] specializes in β -lactamase resistance genes and provides curated sequence annotations and a BLAST service that easily allows the comparison of new data.

In this chapter, we present two distinct approaches to handling HTS data in the context of microbial typing and genomic epidemiology. In the first part, we will illustrate how the MLST information can be extracted from HTS read sets and how the analysis of the profiles can be performed using the goeBURST algorithm. In the second part, we will illustrate how available public datasets can be mined to detect the presence of antibiotic resistance genes, using reference mapping approaches and public antibiotic resistance determinants databases.

2 Materials

2.1 Computer Resources and Software Installation

The analysis that will be illustrated in this chapter will require the use of an appropriate computer setup in terms of hardware and software.

Currently, most medium-high end laptops with 8GB RAM and quad core CPUs satisfy the hardware requirements for the used software, since bacterial genome sizes typically range from one to eight megabases (Mb). The examples discussed here require the use of the Linux operating system or other UNIX-based systems such as Mac OS X, since current state-of-the-art software packages for HTS analysis are developed to run in this environment. Basic familiarity with UNIX commands and file manipula-

tion is also a requirement for the successful interpretation of the presented protocols.

The software packages required for the described analyses are provided in Table 2. Although we provide the installation procedure for the program versions available at the time of this writing, it is still advisable to visit each program web site and follow the guidelines found there.

The commands necessary to install the required software are presented in Table 3. The user must have sudo (superuser) privileges in the system. We assume that the programs will be installed in a directory named `~/methods_protocols/programs/` located in the home directory of the user. In the command line, the following commands should be executed to create the needed directory structure:

```
$ cd ~
$ mkdir methods_protocols
$ cd methods_protocols
$ mkdir programs
$ export PATH=~/methods_protocols/programs/
bin:$PATH
```

It is also required to create folders to store the input files and analysis results from the different examples:

```
$ cd ~/methods_protocols/
$ mkdir de_novo_assembly
$ mkdir reference_mapping
$ mkdir antibiotic_resistance
```

2.2 High-Throughput Sequencing Data

The aim of this chapter is not to provide a protocol for sample preparation and HTS sequencing using a particular technology. However, there are a few general requirements for the successful use of HTS data in characterizing bacterial genomes:

1. A coverage of 60 \times to 100 \times , i.e., in average each position of the genome should be represented by a depth of 60–100 reads. This is usually evaluated by mapping the reads onto the genome of a reference strain.
2. The read coverage should be uniformly distributed across the genome.
3. If possible, choose data from a paired-end (PE) sequencing protocol. This library preparation approach aims to sequence both ends of each DNA fragment (insert). The extra information provided by PEs can increase the reliability of an assembly or mapping analysis.
4. Confirm the resulting reads quality. The files containing the read information are commonly provided in FASTQ format. This file format contains the sequence itself and sequence quality information. We recommend the use of the FastQC software to assess the quality of the resulting reads. FastQC can be found at

Table 2
Description and websites of the tools used

Name	Description	Link
SPAdes [34]	Genome Assembler with multi k-mers and read correction	http://bioinf.spbau.ru/spades
BLAST+ [36]	Tollkit to find regions of local similarity between sequences (required for MLST 2)	http://www.ncbi.nlm.nih.gov/books/NBK279690/
MLST2	Type attribution to sequence files by scanning against PubMLST schemes	https://github.com/tseemann/mlst
Samtools/Bcftools [38]	A suite of programs for interacting with HTS data (required for ReMatCh)	http://www.htslib.org/
Bowtie2 [39]	An ultrafast and memory-efficient tool for aligning sequencing reads to reference sequences (required for ReMatCh)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Bedtools [40]	A swiss-army knife of tools for a wide-range of genomics analysis tasks (required for ReMatCh)	http://bedtools.readthedocs.org/en/latest/
GATK [41]	A software package for analysis of high-throughput sequencing data (required for ReMatCh)	https://software.broadinstitute.org/gatk/
Picard	A set of tools for working with next generation sequencing data in various formats (required for ReMatCh)	http://broadinstitute.github.io/picard/
Aspera connect	An install-on-demand Web browser plug-in that facilitates high-speed uploads and downloads with an Aspera transfer server. (required for ReMatCh)	http://downloads.asperasoft.com/connect2/
ReMatCh	A tool for read mapping, coverage control and consensus sequence production	https://github.com/miguelpmachado/ReMatCh/tree/antibiotics_book_chapter
PHYLOVIZ [8]	Software for the analysis of allelic profiles/SNP data using the goeBURST algorithm for the construction of trees	http://www.phylowiz.net
R	A software environment for statistical computing and graphics	https://www.r-project.org/

Table 3
Software installation procedures (see Note 1)

Name	Commands required for installation
SPAdes	<pre>cd ~/methods_protocols/programs wget http://spades.bioinf.spbau.ru/release3.9.0/SPAdes-3.9.0-Linux.tar.gz tar xf SPAdes-3.9.0-Linux.tar.gz cp -r ./SPAdes-3.9.0-Linux/bin/ ./ cp -r ./SPAdes-3.9.0-Linux/share/ ./ rm -rIf SPAdes-3.9.0-Linux.tar.gz SPAdes-3.9.0-Linux/</pre>
BLAST+	<pre>cd ~/methods_protocols/programs wget ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/ ncbi-blast-2.4.0+-x64-linux.tar.gz tar xf ncbi-blast-2.4.0+-x64-linux.tar.gz cp ./ncbi-blast-2.4.0+/bin/* ./bin/ rm -rIf ncbi-blast-2.4.0+-x64-linux.tar.gz ncbi-blast-2.4.0+/-</pre>
MLST 2	<pre>cd ~/methods_protocols/programs git clone https://github.com/tseemann/mlst.git cp ./mlst/bin/* ./bin/ cp -r ./mlst/db/./ cp -r ./mlst/perl5/ ./ rm -rIf mlst sudo cpan Moo List::MoreUtils</pre>
Samtools/Bcftools	<pre>sudo apt-get install zlib1g-dev sudo apt-get install libncurses5-dev cd ~/methods_protocols/programs wget https://github.com/samtools/samtools/releases/download/1.2/samtools-1.2.tar.bz2 tar xf samtools-1.2.tar.bz2 cd samtools-1.2 make make prefix=~/methods_protocols/programs install cd .. rm -rIf samtools-1.2.tar.bz2 samtools-1.2/ wget https://github.com/samtools/bcftools/releases/download/1.2/bcftools-1.2.tar.bz2 tar xf bcftools-1.2.tar.bz2 cd bcftools-1.2 make make prefix=~/methods_protocols/programs install cd .. rm -rIf bcftools-1.2.tar.bz2 bcftools-1.2/</pre>
Bowtie2	<pre>cd ~/methods_protocols/programs wget http://downloads.sourceforge.net/project/bowtie-bio/bowtie2/2.2.9/bowtie2-2.2.9-linux-x86_64.zip unzip bowtie2-2.2.9-linux-x86_64.zip cp ./bowtie2-2.2.9/bowtie* ./bin/ rm -rIf bowtie2-2.2.9-linux-x86_64.zip bowtie2-2.2.9/</pre>

Table 3 (Cont'd)
Software installation procedures (see Note 1)

Bedtools	<pre>cd ~/methods_protocols/programs wget https://github.com/arq5x/bedtools2/releases/download/v2.25.0/ bedtools-2.25.0.tar.gz tar xf bedtools-2.25.0.tar.gz cd bedtools2/ make cd .. cp ./bedtools2/bin/* ./bin/ rm -rIf bedtools-2.25.0.tar.gz bedtools2/</pre>
GATK	<pre>sudo apt-get install openjdk-8-jdk cd ~/methods_protocols/programs # Download GATK software to this folder tar xf GenomeAnalysisTK-3.5.tar.bz2 chmod +x GenomeAnalysisTK.jar mv GenomeAnalysisTK.jar ./bin/ rm GenomeAnalysisTK-3.5.tar.bz2</pre>
Picard	<pre>cd ~/methods_protocols/programs wget https://github.com/broadinstitute/picard/releases/download/2.1.0/ picard-tools-2.1.0.zip unzip picard-tools-2.1.0.zip chmod +x ./picard-tools-2.1.0/*jar mv ./picard-tools-2.1.0/* ./bin/ rm -r picard-tools-2.1.0.zip picard-tools-2.1.0/</pre>
Aspera connect	<pre>cd ~/methods_protocols/programs wget http://download.asperasoft.com/download/sw/connect/3.6.2/ aspera-connect-3.6.2.117442-linux-64.tar.gz tar xf aspera-connect-3.6.2.117442-linux-64.tar.gz sh aspera-connect-3.6.2.117442-linux-64.sh rm aspera-connect-3.6.2.117442-linux-64.sh aspera-connect-3.6.2.117442- linux-64.tar.gz cd ~/ cp -Rn .aspera/connect/* ~/methods_protocols/programs/ rm -rIf .aspera/</pre>
ReMatCh	<pre>sudo apt-get install python-dev sudo apt-get install python-pip pip install numpy cd ~/methods_protocols/programs git clone https://github.com/bfrgoncalves/ReMatCh.git chmod +x ./ReMatCh/rematch.py cd bin ln -s ..//ReMatCh/rematch.py ./</pre>
PHYLOViZ	<pre>cd ~/methods_protocols/programs wget https://bitbucket.org/phyloviz/phyloviz-main/downloads/phyloviz-1.1a.zip unzip phyloviz-1.1a.zip rm phyloviz-1.1a.zip</pre>
R	<pre>sudo apt-get install r-base</pre>

<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

This software provides a user-friendly interface to evaluate several of the parameters that evaluate the quality of the reads obtained, such as average base calling quality vs. read length, read length distribution, and G+C content distribution.

2.3 Databases and Data sets Used in This Chapter

2.3.1 European Nucleotide Archive (ENA)/Sequence Read Archive (SRA)

The European Nucleotide Archive (ENA) is supported by the European Molecular Biology Laboratory (EMBL) and offers a comprehensive record of the world's sequencing data. The National Center for Biotechnology Information (NCBI), which is supported by the U.S. National Library of Medicine, also provides HTS data storage through the Sequence Read Archive (SRA). Both ENA and SRA databases are synchronized regularly, which ensure that the data deposited into either of the databases will be available through both services.

We will exemplify how to mine data available in the ENA database by retrieving MLST information for a subset of samples studied in a previous publication [32] (namely *Streptococcus pneumoniae* serotype 3 isolates), as well as briefly explain how to do the same thing for the entire deposited data of a given species (we will use *S. pneumoniae* dataset as an example).

2.3.2 Comprehensive Antibiotic Resistance Database (CARD)

Using the Methicillin-resistant *S. aureus* (MRSA) isolates studied previously [33] and deposited in ENA database under study accession number PRJEB4980, we will exemplify how antibiotic resistance genes can be found using directly sequence reads using a reference mapping approach. To that goal, we will take advantage of the Comprehensive Antibiotic Resistance Database (CARD), which is a pilot project of the Antibiotic Resistance Pipeline (a Canada - UK Joint Health Research Program on Antibiotic Resistance) and includes an extensive set of reference genes involved in antibiotic resistance from a variety of organisms, genomes, and plasmids. For simplicity, a subset of the CARD database that excludes genes that confer resistance via specific mutations will be used.

3 Methods

3.1 Analysis of Multilocus Sequence Typing Data from HTS Data

3.1.1 De Novo Assembly

Recent developments in de novo assembly algorithms have increased assembly accuracy and speed. Therefore, we will demonstrate how to de novo assemble a genome of *S. pneumoniae* (ENA run accession number ERR026213) using SPAdes Genome Assembler [34], which can then be used to identify the MLST sequence type (ST). SPAdes is a true multi k-mer assembler (combines information from intermediate assemblies produced with k-mers of different length in

the final assembly) that can work with sequence data from different sources (single-cell or standard multicell) produced by different technologies (Illumina and IonTorrent, or hybrid assemblies combining PacBio, Oxford Nanopore or Sanger reads).

1. For the selected *S. pneumoniae* Illumina PE sequence data, it can be downloaded as follows:

```
$ cd ~/methods_protocols/de_novo_assembly/
$ wget ftp://ftp.sra.ebi.ac.uk/vol1/fastq/
ERR026/ERR026213/ERR026213_*
```

2. And the assembly can be performed by typing the following commands (see Note 2):

```
$ spades.py --careful --pe1-1 ERR026213_1.
fastq.gz --pe1-2 ERR026213_2.fastq.gz
--threads 1 --memory 4 -o ./spades
```

The de novo assembled contigs and scaffolds can be found inside the recently created spades/ folder as contigs.fasta and scaffolds.fasta, respectively.

3.1.2 Reference Mapping

Mapping reads to small (a few kb size) reference loci needs reduced computing power and time when compared with de novo assembly. Therefore, this approach allows the faster analysis of large amounts of data. However, going from reads to sequences requires the use of different bioinformatics tools: one to index the reference sequence; a mapper to map reads against the reference; a tool to convert the output into binary format, sort and index the output to speed up analysis; one for calling the allelic variants; and finally another one to produce the consensus sequence (the sequence incorporating the allelic variants). Users might also want to control for sequence coverage to ensure robustness in allele call and may want to download large amounts of publicly available sequence data from databases. In order to simplify and automate all these steps, we created a tool that combines all these steps: ReMatCh.

ReMatCh has different operating modes, ranging from harvesting the sequence of the loci of interest for all data available at ENA repository from a single species, to performing read mapping to loci of interest of a certain subset of ENA sequence data or to analyzing user supplied sequence data. In this section, we will mainly focus on how to determine the ST of a given ENA run accession list, and briefly describe how to determine the STs of all sequences available from a given species, using both ReMatCh and MLST 2 software.

The first step is to prepare the reference sequences. We will explain how to use the PubMLST database to retrieve MLST reference sequences for *S. pneumoniae*:

1. Access the *S. pneumoniae* PubMLST database (<http://pubmlst.org/spneumoniae/>).
2. Navigate to “Sequence/profile definitions database.”

3. Click on export “Sequences.”
4. In “Select STs” box type any ST number (we used ST 1); in “Select loci” box click on “All” button to automatically select all scheme loci; finally “Submit” your request.
5. After the job is processed, download the “XMFA output file (not aligned).”
6. The downloaded PubMLST XMFA file needs to be converted into a fasta file. Using text editor software, remove the equal sign line at the end of each sequence.
7. For convenience, change the sequences headers.
8. Save the modified file as `mlst.fasta` in the `~/methods_protocols/reference_mapping/` directory.

A previous publication [32] reports the sequence of 616 isolates of *S. pneumoniae* and provides complete metadata information that can be compared to the STs found using this approach. For the purpose of this exercise we will use a subsample of the strains by focusing on serotype 3 isolates. To obtain the metadata proceed as follows:

1. Access Nature Genetics article webpage via <http://dx.doi.org/10.1038/ng.2625> (or <http://www.nature.com/ng/journal/v45/n6/full/ng.2625.html>).
2. Navigate to the “Supplementary information” section.
3. Download the “Supplementary Table 1” (a Microsoft Excel file) (*see Note 3*).
4. Open the download file in a spreadsheet software (*see Note 4*).
5. Apply a filter to the data by choosing “Data” menu, then “Filter” submenu, and finally clicking on “Add Auto Filter.”
6. In “Serotype” column drop-down box choose “3.”
7. Select and copy the displayed ENA run accession numbers (that belong to serotype 3 isolates and should represent 11 isolates).
8. Paste the run accession list into a text file and save it as `Spneumoniae_serotype3.txt` in the `~/methods_protocols/reference_mapping/` folder.

Having the reference sequences and the list with the ENA run accession numbers, ReMatCh can be run by typing the following commands (*see Notes 5 and 6*):

```
$ cd ~/methods_protocols/reference_mapping/
$ rematch.py ReMatCh -r mlst.fasta -d ../
rematch_run/ -cov 10 -qual 10 -mul 0.75 -l
Spneumoniae_serotype3.txt --threads 1 -rm-
Fastq -bowtieBuild -clean --asperaKey ~/.
methods_protocols/programs/etc/asperaweb_
id_dsa.openssh
```

Consensus sequences (as well as all the intermediate files produced) will be stored in specific folders for each run accession number within the rematch_run folder.

The same approach can be used to determine the STs of all publicly available genomes of a given species (see Note 7). To achieve this use the following commands (see Note 8):

```
$ cd ~/methods_protocols/reference_mapping/
$ rematch.py ReMatCh -r mlst.fasta -d ../
rematch_run_all_Spneumoniae/ -cov 10 -qual
10 -mul 0.75 -l rematch_run.list_runIDs.txt
--tax "Streptococcus pneumoniae" --threads
1 -rmFastq - bowtieBuild -clean --asperaKey
~/methods_protocols/programs/etc/asperaweb_
id_dsa.openssh
```

3.1.3 Defining MLST Sequence Types

For determining the STs of a small number of genomes, a practical online service is available from the Center for Genomic Epidemiology from the Danish Technical University [35]. In the website (<https://cge.cbs.dtu.dk//services/MLST/>) users can upload either reads from the majority of used HTS technologies (single-end or paired-end) or already assembled contigs and obtain the determined sequence type and allele information for all available MLST schemas. However, using this resource is impractical for more than a few strains.

The MLST 2 software provided by Torsten Seemann scans sequence files against PubMLST typing schemes using NCBI BLAST+ [36] blastn software. Users can run MLST 2 against a certain scheme using *--scheme* option, but it can be run in auto-detection mode that returns the scheme to which the query sequences are most likely to belong. Using the de novo assembled sequences we will exemplify how to run MLST 2 against a certain scheme (the auto-detection mode will be explained in the reference mapping example against the entire dataset of *S. pneumoniae* in the ENA database).

1. For the de novo assembled scaffolds of ENA run accession number ERR026213, MLST 2 can be run as follows (see Note 9):

```
$ cd ~/methods_protocols/de_novo_assembly/
$ mlst --scheme spneumoniae ./spades/scaffolds.fasta > ERR026213_mlst.txt
```

2. The STs of the serotype 3 isolates, whose sequences at the loci of interest were recovered using ReMatCh, can also be determined using MLST 2 as follows (see Note 10):

```
$ cd ~/methods_protocols/reference_mapping/
$ mlst --scheme spneumoniae ./rematch_run/* ./rematch_results/*_sequences.fasta >
Spneumoniae_serotype3.mlst.txt
```

3. In the case of all sequences of *S. pneumoniae* present in ENA, MLST 2 can be run with auto-detection mode to check for

possible miss-labeled sequence data, using the following commands:

```
$ cd ~/methods_protocols/reference_mapping/
$ mlst ./rematch_run_all_Spneumoniae/*
rematch_results/*_sequences.fasta >
Spneumoniae.mlst.txt
```

While the analysis in this section focused on determining STs of traditional MLST schemas, the same steps can be followed to use ReMatch for read mapping to any available or user-defined core genome MLST schema.

3.1.4 goeBURST Analysis of MLST Data

In this section, we provide an example of a goeBURST analysis using the PHYLOViZ software. We will use this software to explore the *S. pneumoniae* MLST database and how the metadata available on penicillin susceptibility of the isolates can be represented onto the population structure inferred from the MLST data. Similar analyses can be done with the profiles obtained in the previous sections and any available auxiliary data for the strains analyzed. Furthermore, SNP data or cgMLST data can also be analyzed following a similar process:

1. Access the *S. pneumoniae* PubMLST database (<http://pubmlst.org/spneumoniae/>).
2. Navigate to the “Isolates Database”.
3. Under Export select “Export Dataset”.
4. Click on the “None” button to clear all selections. Select only “penicillin” for this example by clicking on the appropriate checkbox.
5. Under Schemes select MLST by clicking on the appropriate checkbox.
6. Press the “Submit” button. Wait for the file to be ready to download. When ready, the Output area will display options to download the query results as text file, Excel file or a compressed tar file with both text and Excel files. Select the Excel file format and save it in your working directory.
7. The Excel file should contain 10 columns with the following headers: “penicillin”, the seven loci of MLST schema (“aroE”, “gdh”, “gki”, “recP”, “spi”, “xpt”, “ddl”) and the “ST(MLST)”. A final column with “clonal_complex(MLST)” as header is also present but should be deleted for this analysis.
8. Penicillin susceptibility minimal inhibitory concentrations (MIC) should then be manually curated and converted to the SIR convention following the Clinical and Laboratory Standards Institute (CLSI) recommended breakpoints before 2008 as epidemiological breakpoints [37]: Susceptible (S) for $\text{MIC} \leq 0.06 \text{ mg/L}$; Intermediate (I) for $0.12 \text{ mg/L} < \text{MIC} < 1 \text{ mg/L}$; Resistant (R) for $\text{MIC} \geq 2 \text{ mg/L}$.

9. Save the curated file as “PneumoPenSIR.txt” in `~/methods_protocols/`. A copy of the manually curated file obtained from the database in 26/02/2016 can be found in <https://figshare.com/s/a095d76c69e387b605dd>. This file will be used in the next steps.
10. To start PHYLOViZ type the following commands:

```
$ cd ~/methods_protocols/programs/phyloviz/
bin/
$ ./phyloviz &
```
11. On the File menu, select “Load dataset from MLST DBs” (*see Note 11*).
12. In the Dataset name type “Spneumoniae MLST DB”. Any name could be chosen since it works only as an identifier on the Dataset tab. On the “Public DB Name” dropdown menu scroll down until you find “pubmlst.org - Streptococcus pneumoniae” and click to select it. Press the “Next >” button.
13. On the “Typing Data” window press the “Start/Stop” button once to start the download of all MLST allelic profiles available in the database. You can monitor the progress through the progress bar. Wait a few seconds until “Done!” appears on the progress bar. Press “Next >”.
14. On the “Isolate Data” tab press the “Browse” button on “File:” and locate the PneumoPenSIR.txt that should be in `~/methods_protocols/`. Select the file. On the “Key:” dropdown menu, make sure that “ST” is selected. This links the allelic profiles to the auxiliary data we are providing. Press “Next >”.
15. On the “Sequence Data” tab, Press “Finish” (*see Note 12*).
16. On the “Datasets Tab”, the dataset should have appeared as “Spneumoniae MLST DB”. Click on the selector on the left of the item to show the Isolate Data and the MLST data. Double clicking on either will display the respective data.
17. To run the goeBURST algorithm on the entire dataset, using right mouse button click on “Multi-Locus Sequence Typing (MLST)” and select “Compute > goeBURST”.
18. In the “goeBURST Configuration” window, in the “Distance” tab, select “goeBURST distance” and press “Next >”.
19. In the “Level#” tab, make sure that “SLV” is selected in the slider option and press “Finish”.
20. A tab should appear with a log of the goeBURST algorithm operation. It should take a few seconds for the calculations to complete. At the end, it should display the date and time and “goeBURST algorithm: done”.
21. On the “Datasets Tab” under “Multi-Locus Sequence Typing (MLST)” a new item should appear named “goeBURST (Level 1; goeBURST distance)”. If not visible click on the

selector on the left of “Multi-Locus Sequence Typing (MLST)”. Double click on the “goeBURST (Level 1; goeBURST distance)” to open the tree window.

22. In the tree window, the optimization of the tree position will start. Zoom out with the mouse wheel until you see the whole tree. You can speed it up by moving the “Animation Speed” slider to its maximum (100). Due to the large size of the tree this may take some minutes, but you can accelerate the process by manually helping the untangling of tree branches. You can do this by clicking with the left mouse button on a node, and keeping it pressed, dragging it to the desired position. After the tree is correctly displayed reduce the animation speed (a speed of around 25 is recommended).
23. To optimize the display further, press the “Options” button and select “Control.” On the “Force Control” window, decrease “GravitationalConstant” to -4.6, and increase Default SpringLength to 100. Close the “Force Control” window.
24. To display the penicillin SIR classification onto the tree, double click on the “Isolate Data” item on the “Datasets tab”. The table containing the penicillin SIR and all the alleles should be displayed. A final column named “goeBURST[1]” represents the goeBURST group number attributed by the goeBURST algorithm.
25. While pressing the Control key, right click on the column Header “penicillin SIR”. Click the “Select” button, and then the “View” button. A colored pie chart with the various categories is shown. Customize the colors used by clicking on each color square on the legend. A pop-up appears with a color picker. Select Green for S, Orange for I, Red for R, and grey for NA.
26. On the “Datasets” tab double click again on “Multi-Locus Sequence Typing (MLST): “goeBURST (Level 1; goeBURST distance)” item to display the tree again. For increasing the display quality, click on “Options” and select the “High quality” checkbox.
27. The resulting image should be similar of what is represented in Fig. 1.
28. This tree shows the emergence of penicillin resistance more commonly on STs on the outer branches of the goeBURST tree. You can search for any ST in the tree by typing its number in the “search>>” box (lower right corner). The selected ST should be represented in the middle of the display area. When zooming, the selected ST will stay in the middle of the display facilitating a more direct visual exploration of that region of the tree.
29. A tutorial video for PHYLOViZ is available at <http://www.phyloviz.net/wiki/videos/>, demonstrating, on a smaller data-set, the features of the software.

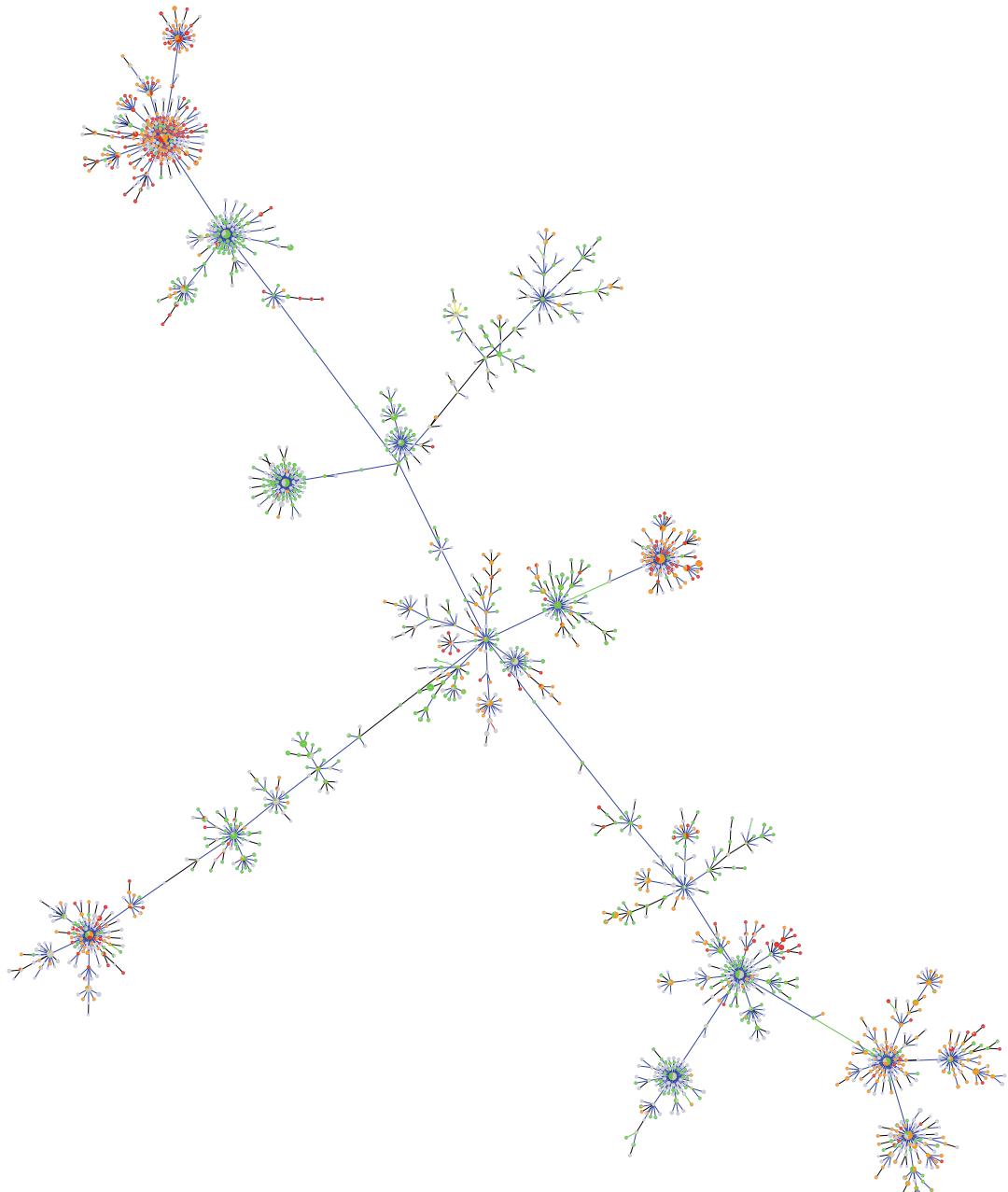


Fig. 1 Snapshot of the goeBURST largest *Streptococcus pneumoniae* Clonal complex created with MLST allelic data from PubMLST (<http://pubmlst.org/spneumoniae/>—accessed on 26/Feb/2016) colored with penicillin susceptibility data obtained from the isolates database. The *green* nodes represent penicillin susceptibility, *orange* nodes intermediate penicillin non-susceptibility and *red* nodes penicillin non-susceptibility, following Clinical and Laboratory Standards Institute (CLSI) recommended breakpoints prior to 2008. *Grey* nodes represent not available data (NA) on penicillin susceptibility

3.2 Finding Resistance Genes

3.2.1 Antibiotic Resistance Databases

3.2.2 Detecting Genes in HTS Datasets with Reference Mapping

In this example, we will query a set of MRSA genomes [33] for resistance genes present in the CARD database. The first step is to download the CARD database sequences file:

```
$ cd ~/methods_protocols/antibiotic_resistance/
$ wget http://arpcard.mcmaster.ca/blast/db/
nucleotide/ARmeta-genes.fa.gz
$ gunzip ARmeta-genes.fa.gz
```

The ENA run accession numbers of the MRSA [33] must be retrieved and provided, together with the CARD sequences file, to ReMatCh. Users can get this information using two different approaches: (a) via the ENA website (**steps 1–4**) or, (b) using UNIX commands (**step 5**) (*see Note 13*):

1. Access study PRJEB4980 on the site through the link <http://www.ebi.ac.uk/ena/data/view/PRJEB4980>.
2. Navigate to the “Read Files” tab and then click on “TEXT” to download the required information.
3. Open the downloaded file in spreadsheet software and copy the run accession numbers located in column “Run Accession.”
4. Paste those in a text file and save with name file mrsa_Leopold_2014_JCM.txt in `~/methods_protocols/antibiotic_resistance/` folder.
5. Using UNIX commands:

```
$ cd ~/methods_protocols/
antibiotic_resistance/
$ wget --output-document=ENA_study_
information_MRSA.txt "http://www.ebi.ac.uk/
ena/data/warehouse/filereport?accession=PRJE
B4980&result=read_run"
$ sed 1d ENA_study_information_MRSA.txt |
cut -f 6 > mrsa_Leopold_2014_JCM.txt
```
6. Now, with both files, ReMatCh can be run as follows:

```
$ cd ~/methods_protocols/antibiotic_resistance/
$ rematch.py ReMatCh -rnucleotide_fasta_pro-
tein_homolog_model.fasta -d ./rematch_run_AR/
-cov 10 -qual 10 -mul 0.75 -l mrsa_Leopold_2014_
JCM.txt --threads 1 -rmFastq -bowtieBuild
-clean --asperaKey ~/methods_protocols/pro-
grams/etc/asperaweb_id_dsa.openssh
```

The ReMatCh results can then be easily examined with the rematch.py mergeResults command. This command analyzes the coverage information for each locus and, based on the percentage of bases with the minimum coverage parsed by ReMatCh, it reports the genes present in a given dataset. Since the presence of several antibiotic resistance genes was previously reported [33] (Table 4), users can investigate the presence of the same genes using the approach described here. It can be run as follows:

Table 4
Antibiotic resistance genes previously reported

Antibiotic	Susceptibility	Genes
Clindamycin	Resistant	<i>ermA</i> , { <i>ermC</i> }
Erythromycin	Resistant	<i>msrA</i> ^b , <i>msrB</i> ^b
Gentamicin and tobramycin	Resistant	<i>aac6'-aph2"</i>
Linezolid	Susceptible	{ <i>cfr</i> }
Methicillin	Resistant	<i>mecA</i>
Mupirocin	Susceptible	{ <i>mupA</i> } ^b
Vancomycin	Susceptible	{ <i>vanA</i> } ^b

In the gene column, brackets highlight genes searched for and absent in the studied isolates according to the previous publication [33]

^aDenotes a gene not found using the approach described here

^bIndicates genes absent from the CARD database annotated as *Staphylococcus aureus* antibiotic resistance genes

1. As a first step, a list file containing the genes investigated previously [33] can be created as *mrsa_genes.txt* file and saved in *~/methods_protocols/antibiotic_resistance/* folder with the following content (see Note 14): *ermA ermC msrA msrB aacA-aphD aac cfr mecA mupA vanA*
2. Run the *rematch.py mergeResults* command as follows (see Note 15)

```
$ cd ~/methods_protocols/antibiotic_resistance/
$ rematch.py mergeResults --mrWorkdir ./rematch_run_AR/ --mrSequenceCoverage 0.85.
```
3. To retrieve information about *Staphylococcus aureus* annotated genes, and specifically those referred above, execute the following commands:

```
$ cd ~/methods_protocols/antibiotic_resistance/
$ grep "#" ./rematch_run_AR/merged_results/mergedResults.transposed.tab > mrsa_genes.mergedResults.tab
$ grep --ignore-case "Staphylococcus" ./rematch_run_AR/merged_results/mergedResults.transposed.tab | grep --ignore-case -- file=mrsa_genes.txt >> mrsa_genes.mergedResults.tab
```
4. For *Staphylococcus aureus* genes present in CARD database, the antibiotic resistance genes found in the isolates in this dataset, replicate the majority of results obtained by [33]. Inspection of the *mrsa_genes.mergedResults.tab* file reveals one difference from the previous reported results: the ERR375866 run acces-

sion number (corresponding to P13 isolate) lacks the ermA gene in our analysis. There may be multiple reasons for this discrepancy, such as the choice of resistance genes included in the CARD. Nevertheless, the mergedResults.tab file can be further explored in order to assess the presence of other antibiotic resistance genes.

4 Notes

1. MLST 2 requires the installation of some Perl modules that can be easily set up with *cpan* command. Samtools and Bcftools need to be installed independently, but both tools expect some libraries to be installed which can be achieved using the *apt-get* command. The newest version of GATK tool (which must be obtained through their website, <https://www.broadinstitute.org/gatk/>, and placed in *~/methods_protocols/programs* folder) requires Java Development Kit (JDK) version 8 to run, and it can also be installed through the *apt-get* command. The JDK will also be used by PHYLOViZ. Since ReMatCh requires the Python package Numpy to run, it can be set up using the *pip* command, which can be installed by its turn using the *apt-get* command. If users prefer to self contain Python dependencies, the installation of a Virtual Python Environment builder might be advisable (<https://pypi.python.org/pypi/virtualenv>). Also, ReMatCh uses Samtools/Bcftools v1.2 instead of the new version.
2. SPAdes uses algorithms for read error correction and for reducing the number of mismatches and short indels to obtain high-quality assemblies. Those can be activated using the *--careful* option. It is also recommended to set an adequate memory limit (in GB, for example *--memory 4*) to avoid exceeding the machine's capacity.
3. Users can simply try the following command to download Supplementary Table 1:
\$ wget <http://www.nature.com/ng/journal/v45/n6/extref/ng.2625-S2.xlsx>
4. For the example we used Linux Gnumeric software, but the same steps can be applied to other software like Microsoft Excel.
5. In ReMatCh *-cov* option sets the minimum position coverage required to call the allelic variants and to control mapping coverage; *-mul* sets the minimum coverage for the alternative allele, therefore allowing ReMatCh to control for multiple alleles that could result in gene duplication or contamination of the original sample with DNA from multiple isolates;

`-rmFastq` option tells ReMatCh whether to remove or not the fastq files after determining the allele variants (when analyzing large datasets it is recommended to set this option to conserve disk space).

6. For reference mapping, users can add extra sequences to both ends of the reference sequences. This “extra” sequence will be used by ReMatCh to provide reference support for read mapping avoiding coverage decrease at sequence end. Adding both upstream and downstream sequence to the region of interest is highly recommended since it allows proper read mapping to the entire region of interest and increased accuracy and robustness in allelic variant calling. ReMatCh can then be configured to ignore this extra sequence at both ends using the `--xtraSeq`.
7. When using ReMatCh for analysing all public available data from a certain species, users must be aware that in case of large datasets, such as that of *S. pneumoniae* which already has more than 30,000 sequenced libraries, the ReMatCh analysis will take longer than would be desirable. Furthermore, internet connection problems to ENA may occur during the process due to the high volume of data traffic.
8. In ReMatCh, `-l` option specifies the name of the file where ReMatCh will store the ENA run accession numbers of the species of interest when `--tax` option is set; `--tax` specifies the taxon to be downloaded from the ENA database. The user can choose a taxon definition at any level, for instance, instead of a single species (in this case *S. pneumoniae*), user may want to analyze all sequences of a given genus (in this case *Streptococcus*), a given family (in this case *Streptococcaceae*) or another higher level taxon. Beware that the volume of data increases quickly with higher taxa and this may pose problems in accessing ENA.
9. The MLST 2 tabular output file can then be parsed using a spreadsheet software. Users can check the available schemes using `$ mlst --longlist` command.
10. MLST 2 results from [32] serotype 3 isolates stored in the *Spneumoniae_serotype3.mlst.txt* tabular file can be opened in a spreadsheet software, by choosing open in the “File” menu or by importing the data to the spreadsheet software through the “Data” menu and “Import Data” submenu. Using the described approach it was possible to recover the sequence type of one isolate (ERR069750) that was listed as unavailable in the previous publication [32] metadata.
11. Users need to have Internet access to download the datasets. PHYLOViZ allows users to directly interact with publicly accessible MLST databases, and enable users to easily download the latest version of the database.

12. In PHYLOViZ, “Sequence Data” tab would allow downloading the fasta sequences for the loci of the MLST schema, or to provide these from local files, for subsequent analysis.
13. Although the publication [33] provided 18 ENA Secondary Sample Accession numbers relative to their *S. aureus* sequenced isolates, only 16 samples are in fact present in ENA database (ERS372434 and ERS372447 are missing). Four of the isolates seem to have been sequenced in duplicate (ERS372431, ERS372432, ERS372439, and ERS372445).
14. The *aacA-aphD* gene name is an alias to the *aac6'-aph2*" gene reported by [33]. For database search, it was also included aac to ensure that any gene name containing aac was selected by the search.
15. The rematch.py mergeResults command will create the mergeResults.tab file inside~/methods_protocols/antibiotic_resistance/rematch_run_AR/merged_results/ folder that will report which genes are present in the different sample. In the case of genes being present (genes with equal to or more than 85 % of nucleotides with 10 reads coverage minimum, ---mrSequenceCoverage option), the script will provide the mean sequence coverage, otherwise will report “Absent” for genes not present, or “Mul_Allele” for those genes that might have multiple alleles (depending on ReMatCh -mul option).

References

1. Courvalin P (2016) Why is antibiotic resistance a deadly emerging disease? *Clin Microbiol Infect*
2. Struelens MJ (1996) Consensus guidelines for appropriate use and evaluation of microbial epidemiologic typing systems. *Clin Microbiol Infect* 2:2–11
3. Baggesen DL, Sorensen G, Nielsen EM et al (2010) Phage typing of *Salmonella* Typhimurium—is it still a useful tool for surveillance and outbreak investigation. *Euro Surveill* 15(4):19471
4. Maiden MC, Bygraves JA, Feil EJ et al (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 95:3140–3145
5. Jolley KA, Maiden MCJ (2010) BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:595
6. Feil EJ, Li B, Aanensen D et al (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* 186:1518–1530
7. Francisco AP, Bugalho M, Ramirez M et al (2009) Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics* 10:152
8. Francisco AP, Vaz C, Monteiro PT et al (2012) PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics* 13:87
9. Corander J, Waldmann P, Marttinen P et al (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20:2363–2369
10. Thomas JC, Robinson DA (2014) Multilocus sequence typing of *Staphylococcus epidermidis*. *Methods Mol Biol* 1106:61–69
11. Koreen L, Ramaswamy SV, Graviss EA et al (2004) spa typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation. *J Clin Microbiol* 42:792–799
12. Beall B, Facklam R, Thompson T (1996) Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. *J Clin Microbiol* 34:953
13. Lindstedt B (2005) Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis* 26(13):2567–2582

14. Sabat AJ, Budimir A, Nashev D et al (2013) Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill* 18:17–30
15. Carrico JA, Sabat AJ, Friedrich AW et al (2013) Bioinformatics in bacterial molecular epidemiology and public health: databases, tools and the next-generation sequencing revolution. *Euro Surveill* 18:20382
16. Loman NJ, Constantinidou C, Chan JZM et al (2012) High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat Rev Microbiol* 10:599–606
17. Loman NJ, Misra RV, Dallman TJ et al (2012) Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 30:434–439
18. Jüemann S, Sedlazeck FJ, Prior K et al (2013) Updating benchtop sequencing performance comparison. *Nat Biotechnol* 31:294–296
19. Jüemann S, Prior K, Albersmeier A et al (2014) GABenchToB: a genome assembly benchmark tuned on bacteria and benchtop sequencers. *PLoS One* 9:e107014
20. Martin JA, Wang Z (2011) Next-generation transcriptome assembly. *Nat Rev Genet* 12:671–682
21. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30: 2068–2069
22. Maiden MCJ, van Rensburg MJJ, Bray JE et al (2013) MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol* 11:728–736
23. Hatem A, Bozdağ D, Toland AE et al (2013) Benchmarking short sequence mapping tools. *BMC Bioinformatics* 14:184
24. Chewapreecha C, Marttinen P, Croucher NJ et al (2014) Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. *PLoS Genet* 10:e1004547
25. Harris SR, Feil EJ, Holden MTG et al (2010) Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327:469–474
26. Gardy JL, Johnston JC, Sui SJH et al (2011) Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 364:730–739
27. Medini D, Donati C, Tettelin H et al (2005) The microbial pan-genome. *Curr Opin Genet Dev* 15:589–594
28. McArthur AG, Waglechner N, Nizam F et al (2013) The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* 57:3348–3357
29. Liu B, Pop M (2009) ARDB—antibiotic resistance genes database. *Nucleic Acids Res* 37: D443–D447
30. Zankari E, Hasman H, Cosentino S et al (2012) Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67: 2640–2644
31. Thai QK, Bös F, Pleiss J (2009) The Lactamase Engineering Database: a critical survey of TEM sequences in public databases. *BMC Genomics* 10:390
32. Croucher NJ, Finkelstein JA, Pelton SI et al (2013) Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat Genet* 45:656–663
33. Leopold SR, Goering RV, Witten A et al (2014) Bacterial whole genome sequencing revisited: portable, scalable and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J Clin Microbiol* 52(7):2365–2370
34. Bankevich A, Nurk S, Antipov D et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477
35. Larsen MV, Cosentino S, Rasmussen S et al (2012) Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* 50:1355–1361
36. Camacho C, Coulouris G, Avagyan V et al (2009) BLAST plus: architecture and applications. *BMC Bioinformatics* 10:421
37. CLSI. Performance standards for antimicrobial susceptibility testing; Eighteenth Informational Supplement. CLSI document M100-S18. Wayne, PA: Clinical and Laboratory Standards Institute; 2008
38. Li H, Handsaker B, Wysoker A et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079
39. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359
40. Quinlan AR (2014) BEDTools: the Swiss-army tool for genome feature analysis. Wiley, Hoboken, NJ, USA
41. Van der Auwera GA, Carneiro MO, Hartl C et al (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 11:11.10.1–11.10.33

INDEX

A

- Absolute protein quantification 281–289
Acydepsipeptide antibiotics 16
ADEPs 16–18
Antibacterial compounds 121–130, 281
Antibiotics 3–18, 23, 25, 32, 63–83, 175–199, 265, 281, 283, 288, 311–312
Antibiotic resistance 8, 86, 189, 264, 265, 267, 268, 274–275, 307–327, 331, 332, 336–337, 342, 350–352, 354
Antibiotic resistance genes (ARGs) 307, 309, 337, 342, 350–352
Antibiotic stress response 281–289
Antimicrobial agents 5, 7, 8, 11, 63, 107–117, 134
Antimicrobial compound 49–61, 109, 123, 159, 163–166, 172, 247
Antimicrobial peptides 7, 10, 129, 133–135, 141
AntiSMASH 31–34, 37–41, 43
AQUA 281
Avibactam 14

B

- Bio-assay 122
Biosensors 121–130, 145–156, 178–180
Biosynthesis 7–11, 15, 24, 25, 29, 32–34, 39, 50, 60, 86, 122, 123, 146, 179, 228–230

C

- Callyaerin A 64, 66–76, 83
Cell envelope stress 123, 129
Cell proliferation reagent WST-1 108, 110, 113–114
Cell wall 7, 9, 10, 14, 86, 121–130, 146, 183, 265, 283
Cluster boundaries 37, 39, 43, 44
Computer-aided drug design 85–101
Cytotoxicity assay 107–117, 178

D

- Depolarization 10, 134, 139
DNA recombination 180
DNA replication 12, 13, 178, 180, 183
Docking 10, 89, 92–96, 133
Drug 4, 5, 13, 16, 63, 85
Drug discovery 5, 25, 63, 87, 121

E

- Escherichia coli* (*E. coli*) 134, 160, 161, 164, 167, 171, 172, 178, 180, 186–188, 195, 204, 227–231, 233, 238–240, 249–254, 257, 264, 267, 274, 277, 291, 309, 310, 323

F

- Fatty acid disorder 160
Fluorescence anisotropy 202
Fluorescence polarization 177, 188, 190
Force fields 88, 90, 93
FtsZ 15–18
Full-length 26, 248, 249, 251–254, 308, 309
Functional metagenomics 307–327
Functional selections 321, 322

G

- Gel free proteomics 281, 283
Gene cluster families 29, 41–43
Gene finding 350–352
Genome mining 23–44
Gram-negative bacteria 7, 8, 14, 35, 140, 187, 238
Gram-positive bacteria 7, 14, 16, 35, 86, 87, 140, 171, 238, 248, 256, 283

H

- Hemolysis assay 111, 115
Hi3 296, 299, 300, 302, 304
High-throughput assay 175–177
High-throughput assembly 321
High-throughput screening (HTS) 175, 197–198, 203, 205, 213, 238
High throughput sequencing 331–354
Histidine kinase 247–259

I

- Identification of Natural compound Biosynthesis pathways by Exploiting Knowledge of Transcriptional regulation (INBEKT) 31, 34–36
Inhibitor 11–18, 204, 207
In-solution digestion 282

K

Kinase inhibitor 183, 248

L

Laboratory evolution 263–278
 Lactate dehydrogenase release 108, 110, 112, 116
 Laurdan 159–174
 Leaderless mRNA 238
 Lipid adaptation 163
 Lipid domains 159
 Lipid packing 160
 Lipid II cycle 87, 124, 129
 Luminescence 122, 123, 125–130, 179, 182

M

Massively parallel DNA sequencing 307
 Mechanism of action 13, 14, 17, 63
 Membrane fluidity 159–174
 Membrane permeabilization 133
 Membrane potential 133–141, 159
 Membrane targeting antimicrobials 159, 160
 Membrane viscosity 153, 156
 Methicillin-resistant *Staphylococcus aureus* (MRSA) 9, 10, 14–16, 86, 342, 350
 Microbial typing 332–337
 Model membranes 146–150, 153, 155, 156
 Mode of action (MoA) 10, 14, 16, 50, 146, 155, 188, 281
 Mode of inhibition 183, 217–218
 Molecular dynamics (MD) 86

N

Natural products 13, 16, 34, 42, 49, 53, 63, 64, 76, 77
 Neomycin 205, 210, 212, 213, 265
 New Drugs 4 Bad Bugs (ND4BB) 5
 Nuclear magnetic resonance (NMR) 63–83, 87, 88

O

Online databases 332–334
P
 PARFuMS 308, 309, 321, 327
 PC190723 15, 16
 Pharmacophore 86, 89, 92, 95, 96, 98, 101
 Phosphorylation 9, 229, 230, 248–251, 254–256
 Phos-tag 248, 249, 251, 253, 255–257, 259
 Potassium efflux 0, 135, 138, 141

Prioritization 29, 41–43
 Profile HMM-based annotation 327
 Purification 49, 191–194, 249–254, 272, 310, 314, 320

Q

QCONCAT 281

R

Reporter gene 122, 227–229, 231, 233, 238, 242
 Resazurin based cell viability assay 110
 Resfams 309, 322, 327
 Resistance 4, 5, 13, 14, 18, 23, 85, 86, 264–269, 274–276, 307–327, 336–337, 342, 350–352
 Resistome 307–309, 337
 Ribocil 14, 15
 Ribosome 11, 12, 292, 295–298
 RNase P 201, 202, 204, 208–211, 213–217, 219, 221, 222

S

S1 ribosomal protein 238
 Sample preparation 65, 156, 162–163, 168, 281–289, 338
 Secondary metabolite gene cluster 23–44
 Site identification by ligand competitive saturation (SILCS) 89, 91–92, 95, 96, 100
Staphylococcus aureus (*S. aureus*) 8, 9, 14–18, 86, 134, 139, 160, 187, 188, 247, 283, 333, 342, 354
 Stress response 10, 13, 281–283
 Structure-activity relationship (SAR) 85, 87, 97–99
 Structure elucidation 49, 63–83

T

Teixobactin 14
 Tetraphenylphosphonium bromide 133–138
 Transcriptome analysis 267, 269–270, 274, 276
 Translation initiation 11, 237–244
 tRNA 11, 12, 201, 202, 204–206, 215, 222, 237
 Two-component regulatory system 247

V

Virtual screening 89

W

Whole-cell assay 177, 178, 238, 243