

Cite this: *RSC Chem. Biol.*, 2021,  
2, 423

## The pursuit of mechanism of action: uncovering drug complexity in TB drug discovery

Tianao Yuan,<sup>†</sup> Joshua M. Werman<sup>†</sup> and Nicole S. Sampson<sup>\*</sup>

Whole cell-based phenotypic screens have become the primary mode of hit generation in tuberculosis (TB) drug discovery during the last two decades. Different drug screening models have been developed to mirror the complexity of TB disease in the laboratory. As these culture conditions are becoming more and more sophisticated, unraveling the drug target and the identification of the mechanism of action (MOA) of compounds of interest have additionally become more challenging. A good understanding of MOA is essential for the successful delivery of drug candidates for TB treatment due to the high level of complexity in the interactions between *Mycobacterium tuberculosis* (Mtb) and the TB drug used to treat the disease. There is no single “standard” protocol to follow and no single approach that is sufficient to fully investigate how a drug restrains Mtb. However, with the recent advancements in -omics technologies, there are multiple strategies that have been developed generally in the field of drug discovery that have been adapted to comprehensively characterize the MOAs of TB drugs in the laboratory. These approaches have led to the successful development of preclinical TB drug candidates, and to a better understanding of the pathogenesis of Mtb infection. In this review, we describe a plethora of efforts based upon genetic, metabolomic, biochemical, and computational approaches to investigate TB drug MOAs. We assess these different platforms for their strengths and limitations in TB drug MOA elucidation in the context of Mtb pathogenesis. With an emphasis on the essentiality of MOA identification, we outline the unmet needs in delivering TB drug candidates and provide direction for further TB drug discovery.

Received 5th December 2020,  
Accepted 23rd December 2020

DOI: 10.1039/d0cb00226g

rsc.li/rsc-chembio

### Introduction

#### The unmet need for TB drugs and current screening strategies for drug discovery

After almost 70 years since its discovery,<sup>1,2</sup> isoniazid (INH) has remained the most clinically successful drug for Tuberculosis (TB) treatment. In 1882 Robert Koch discovered the pathogen responsible for TB, yet in 2020, TB still represents the leading cause of death from a single infectious agent around the world.<sup>3</sup> There are an estimated 10 million people globally that developed active TB disease in 2018 and over half a million of those cases were diagnosed with drug resistant TB.<sup>3</sup> The increasing emergence of drug resistant *Mycobacterium tuberculosis* (Mtb) infection and the prolonged drug regimen required to treat drug resistant TB underscore the unmet need in TB drug discovery and development that currently exists.

Over the past 50 years, only two new drugs have been approved by the U.S. Food and Drug Administration (FDA) for

the treatment of TB. Bedaquiline (BDQ) was approved as a second-line drug for multidrug-resistant TB treatment in 2012<sup>4</sup> and pretomanid was approved in 2019 as part of a three-drug regimen for the treatment of highly drug resistant TB.<sup>5</sup>

At the beginning of the genomic era, the focus of antibacterial drug development shifted from traditional cell-based phenotypic screens to enzyme-based biochemical screens. Through the use of large compound libraries, compounds were screened in a high throughput manner for activity against the protein products of essential genes required for growth or survival of the bacterial organism. However, these screening campaigns have not proved very successful at delivering desirable drugs in areas of need.<sup>6,7</sup> Thereafter, the focus in the antibiotic drug discovery field has shifted back to whole cell-based phenotypic screens to identify bacterial growth inhibitors directly.<sup>8</sup> Notably, both newly approved TB drugs were discovered *via* whole cell-based phenotypic screens.<sup>9,10</sup>

The advantages that are provided by whole cell-based phenotypic screening include exclusion of compounds that are unable to penetrate through the hydrophobic mycobacterial cell wall, selection of compounds that inhibit vulnerable cellular targets, selecting hit compounds that target multiple cellular functions, and selecting prodrugs that require intracellular

Department of Chemistry, Stony Brook University, Stony Brook, NY 11794-3400, USA. E-mail: nicole.sampson@stonybrook.edu; Fax: +1-631-632-5738; Tel: +1-631-632-7952

<sup>†</sup> Contributed equally to this work.



bioactivation.<sup>8</sup> In fact, there have been analyses arguing that antibiotics that target a single protein are less successful than those directed against multiple molecular targets.<sup>11,12</sup>

However, whole cell-based phenotypic screening is considered to be largely empirical. While phenotypic screening is on the whole more efficient in translating compound activity into potential therapeutic impact, the mechanisms of action of these compounds are not immediately revealed. Compound libraries used in phenotypic screens have evolved to include compounds of greater structural diversity, for example, there has been renewed interest in natural products for their antibacterial properties. As new technologies have become available, screening strategies have become more comprehensive, leading to the identification of a greater number of hit compounds with novel structures and mechanisms. However, without consideration of an optimal mechanism of action (MOA), drug discovery programs generally carry high attrition rates downstream.<sup>13</sup> Attrition occurs because not all essential targets in laboratory-grown bacteria have equal potential to be exploited as antibacterial drug targets *in vivo*.

*In vitro* broth media conditions used in whole cell-based phenotypic screens greatly differ from the pulmonary micro-environment where *Mtb* resides in the host. This difference results in a high risk of selecting compounds that inhibit an essential target *in vitro* which is dispensable *in vivo*.<sup>14</sup> As a result, different culture models for whole cell-based phenotypic screens have been developed to more closely mimic the environment in which *Mtb* resides in the host, in the hope of improving drug discovery outcomes.<sup>6,8</sup> However, as these culture models increase in complexity, the identification and assignment of MOAs becomes more complex, and requires a longer time to unravel.<sup>15</sup>

### The essential role of MOA identification in successfully delivering TB drug candidates

In order to elucidate the MOA of an inhibitor, a vast amount of work is often required.<sup>16</sup> There is no generic method to elucidate

the MOA of all inhibitors and usually multiple approaches are required to uncover the mechanisms of a new compound's activity.<sup>17</sup>

Understanding MOA plays an essential role in antibiotic development. Although regulatory approval for antibiotics does not require a clearly defined MOA, there is strong motivation to understand the MOA or to identify the molecular target(s) of an antibiotic for advancing hit compounds and lead optimization.<sup>18</sup> Understanding enables exclusion of compounds that elicit bactericidal effects in a non-specific manner. Knowledge of the molecular target(s) also provides a basis for a comprehensive structure-activity relationship (SAR) study to improve activity and pharmacological properties, which can facilitate the rational design and development of next generation leads. This format of drug development simultaneously enables minimization of off-target effects in the host. Understanding MOA helps to lower the risk of compounds failing in the later stages of development due to host interference, general toxicity, and target drugability. As TB is generally treated with a combination of therapeutic agents, it is crucial to understand the molecular targets of each drug used so that an optimized combination may be designed to address the issues of bacterial persistence and drug resistance.<sup>19</sup> All of these factors highlight the need and importance of target identification and MOA assignment in antibiotic drug discovery (Fig. 1).

Prior to the genomic era, MOA studies for TB drugs were predominantly conducted through biochemical approaches. As detailed in several reviews outlining the history of TB drug development,<sup>20,21</sup> using INH as an example, the majority of MOA discovery focused on examining the intracellular levels of extractable cell components or the production of key metabolites, such as cell wall associated lipids and sugars.<sup>2</sup> Serendipity aside, the real breakthrough in TB drug MOA research occurred with the advent of the genomic era when DNA sequencing technology enabled entire genomes to be characterized and the introduction of foreign DNA through plasmid transformation became possible.<sup>22-24</sup> These approaches have provided researchers a

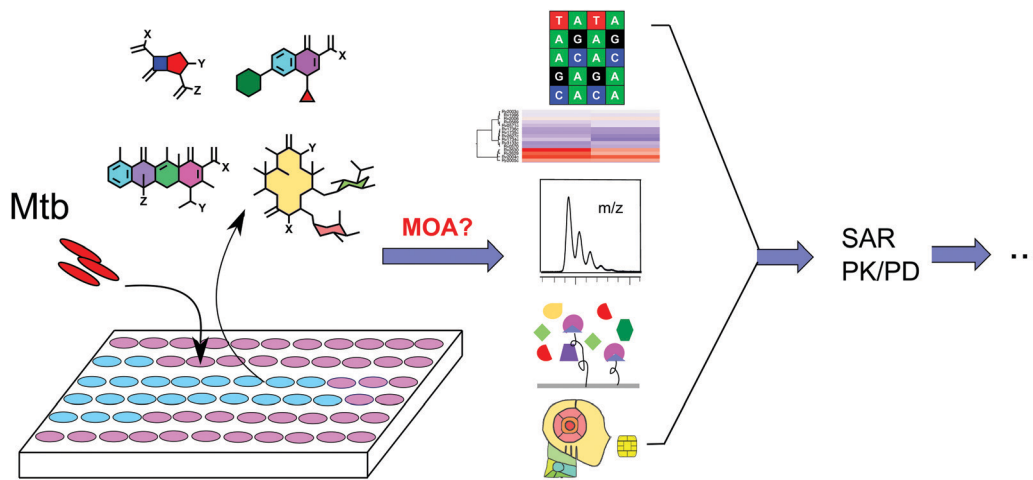


Fig. 1 Schematic representation of phenotypic screening-based TB drug discovery. Hit compounds are identified from whole cell-based phenotypic screens. Their MOAs can be assessed through genetic, metabolomic, biochemical, and computational approaches. A clearly defined MOA will lay a good foundation for the following SAR and pharmacokinetics (PK)/pharmacodynamics (PD) studies.



powerful tool in defining the primary molecular targets and the MOAs of TB drugs.

Benefiting from advancement in -omic technologies, the research strategies involved in the MOA identification of TB drugs have evolved from direct biochemical methods, to genetic interaction, chemical-genetics interaction profiling, metabolomic profiling, and computational machine learning. The integration of all these approaches from compound screening to MOA assessment has yielded over 10 new TB drug candidates that are currently undergoing clinical testing.<sup>25</sup>

While many such technologies are applied generally in drug MOA elucidation in the entire drug discovery field, in this review, we specifically summarize the methodologies applied in the assessment of MOA in TB drug discovery with an emphasis on Mtb physiology and Mtb infection pathogenicity. We highlight the latest technologies and advances for target identification in the field and discuss comparative advantages and disadvantages of each approach. In doing so, we hope to stimulate more innovative strategies to deliver drug candidates in TB drug discovery and expand our understanding of Mtb pathogenicity.

## Spontaneous resistant mutants and WGS

### Acquired drug resistance in Mtb

When streptomycin was introduced into the TB treatment regimen in the 1940s, the emergence of Mtb drug resistance in patients was quickly observed.<sup>26</sup> While intrinsic drug resistance plays a role in mediating antibiotic-induced stress in Mtb, acquired drug resistance accounts for the majority of the molecular mechanisms of drug resistance.<sup>27</sup> Acquired drug resistance in Mtb arises predominantly through spontaneous mutations in chromosomal genes, whereas mobile horizontal gene transfer (HGT), a common source of acquired drug resistance in many bacteria, is rare in Mtb.<sup>28</sup>

Spontaneous mutations arise in all prokaryotes at a constant rate of 0.0033 times per replication.<sup>29</sup> While most of the individual cells carrying mutations do not survive, the ones that persist provide advantages, which allow the organism to more effectively

adapt to the environment, and therefore accumulate in the population. In Mtb specifically, antibiotic resistance occurs at a rate of  $10^{-7}$ – $10^{-9}$  mutations per cell division.<sup>30</sup> Even though the causes and mechanisms for acquired drug resistance are different, isolating drug resistant colonies and following with genetic analysis provides a starting point for drug MOA hypotheses.

### Unraveling drug antibacterial mechanism through genome analysis

Mutations usually occur in the form of single nucleotide missense changes, premature termination, nonsense mutations, or frame-shift mutations which disrupt the bactericidal effect of antibiotics. The mechanisms for antibiotic resistance can be summarized into seven categories.<sup>31,32</sup> They are a gene mutation that results in:

- reduced binding of the drug but preserves function of the target
- increased expression of the target
- modified or increased catabolism of the drug
- decreased activation of the prodrug
- decreased drug uptake or increased drug export
- expression of a compensatory pathway
- loss of function in an enzyme that undoes the action targeted by the drug.

With advancements in whole genome sequencing (WGS) technology, raising spontaneous resistant mutants *in vitro* on solid media has become the most common and straightforward approach to study the molecular mechanisms of TB drugs (Fig. 2). This approach, which can be widely applied, provides a high throughput platform to study the MOAs of the hits that arise from such screens.<sup>33</sup> Mutations that confer resistance then need to be introduced into the mycobacteria through a variety of genetic approaches<sup>34</sup> and the drug resistance of the engineered strain must be confirmed to have a similar drug phenotype to the spontaneously resistant strain. Other genetic techniques, including conditional expression of a target gene can be applied to further validate the interaction between the drug and its putative target.<sup>35</sup>

INH was identified in the 1950s.<sup>36</sup> However, even with many meticulously planned biochemical experiments examining the

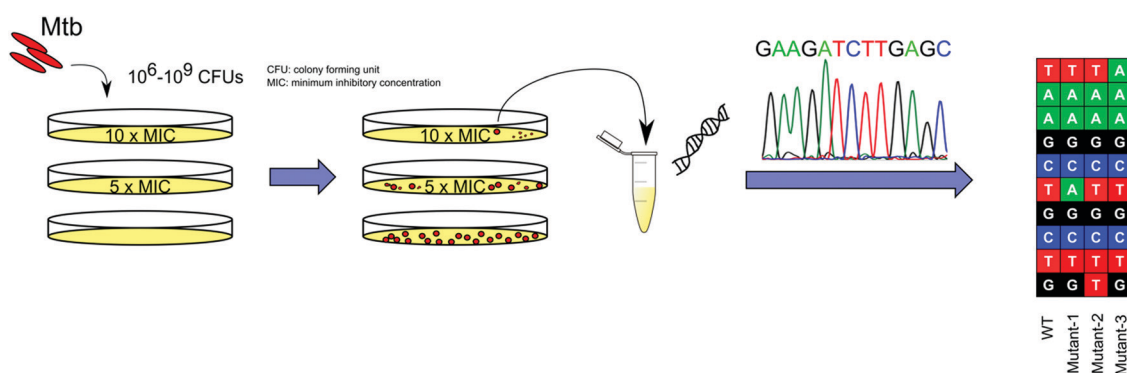


Fig. 2 Schematic representation of WGS of spontaneous drug resistant mutants for TB drug MOA identification.  $10^6$ – $10^9$  CFUs of Mtb are plated onto agar plates containing a range of concentrations of the drug, from  $2\times$ – $10\times$  MIC, for 8–10 weeks at 37 °C. Drug resistant colonies are picked and cultured in broth media, followed by the extraction of genomic DNA for WGS. CFU: colony forming unit. MIC: minimum inhibitory concentration.



mode-of-action of INH, it would not be until the late 1990s to early 2000, through the use of INH resistant Mtb strains, that researchers were able to reveal the MOA of INH. The Mtb catalase peroxidase KatG activates INH to an isonicotinoyl radical which forms an adduct with NAD, and the INH–NAD adduct inhibits InhA, the enoyl-ACP reductase of the fatty acid synthase type II system.<sup>2</sup> BDQ was the first TB drug approved by the FDA in nearly 50 years. The primary target of BDQ, the ATP synthase encoded by *atpE*, was identified from WGS of *in vitro* generated resistant strains.<sup>9</sup>

### Challenges in TB drug MOA assignment through resistant mutant generation

Due to the different mechanisms of drugs, and the essentiality and vulnerability of molecular targets or cellular machinery, spontaneous mutants are often difficult to generate. Mutants also may be only partially resistant to the drug thus the mutations may not indicate the direct molecular targets of the drugs. Moreover, non-specific mutations in the genome are commonly observed, further complicating identification of the molecular target. In many scenarios, a TB drug can work through a multifaceted mechanism, disrupting multiple biological pathways. As a result, the mutations may not point to a single target. Resistance mutations of prodrugs often only occur in the gene encoding the bioactivating enzyme, leaving the real target unidentified, such as the case for the newly approved TB drug pretomanid.<sup>37</sup>

**Polypharmacology and gene essentiality.** Pyrazinamide (PZA) was introduced into TB treatment in the 1970s, however its MOA has yet to be fully defined owing to the polypharmacologic nature of PZA. It binds to multiple cellular targets – contributing to the difficulty of defining its MOA. The challenging process of selecting resistant polymorphisms for PZA has also contributed to the difficulty in elucidating its MOA. The majority of Mtb PZA resistant strains, whether clinical or laboratory isolates, have mutations in the *pnca* gene, which encodes pyrazinamidase. Pyrazinamidase converts PZA to its bio-active form pyrazinoic acid.<sup>38</sup> It has long been believed that the activity of PZA requires low pH in the host, but raising PZA or pyrazinoic acid resistant mutants on acidic solid media turned out to be both unsuccessful and technically challenging.<sup>39,40</sup> All these factors have contributed to the challenge of unraveling the MOA of PZA. More recently there have been advancements in this effort. The identification of mutations in *panD* that contribute to pyrazinoic acid resistance offer more clues, but the precise MOA of PZA still remains undefined. A recent review by Gopal and colleagues summarized a list of mutations from 9 unique genes that have been reported to contribute to PZA resistance.<sup>39</sup>

For compounds with MOAs targeting multiple proteins, depending on the resistance rate of each target, detecting spontaneously resistant strains can be more difficult. This difficulty is most likely due to the low probability of multiple protective mutations occurring in a single cell within a manageable population size. First described in the 1950's, a method to induce mutations chemically by addition of a mutagen, such

as 5-bromouracil, enables elevation of the mutation rate within the population, thereby increasing the likelihood of isolating drug resistant strains.<sup>41</sup> This approach was applied in a recent study to unravel the molecular target of a class of quinoline compounds in Mtb, where with elevated mutation rates, single nucleotide polymorphisms (SNPs) were identified in *Rv2439c* which encodes a glutamate kinase in proline biosynthesis.<sup>42</sup> The quinoline compounds function by activating the glutamate kinase, which leads to an overproduction of proline, and the associated redox imbalance causes bacterial death.

It is worth noting that not all mutations that rescue mycobacteria from antibiotic treatment occur in an essential gene. A recent phenotypic screen identified a class of 4-hydroxyquinolines for their bactericidal activities against both replicating and non-replicating Mtb. Mechanistic investigations under replicating conditions revealed that SNPs in the non-essential gene *Rv1239c* encoding a putative Mg<sup>2+</sup>/Co<sup>2+</sup> transporter CorA in Mtb could rescue the bacteria from compound treatment in replicating conditions.<sup>43</sup> Transposon disruption of *Rv1239c* or direct chemical inhibition of CorA did not affect Mtb viability indicating CorA was not the molecular target of the 4-hydroxyquinolines. Instead, subsequent experiments showed that the compounds functioned by depleting the intracellular Mg<sup>2+</sup> and the altered function of CorA might favor the Mg<sup>2+</sup> influx into the bacteria to mitigate the bactericidal activity. However, SNPs in *Rv1239* were not able to protect Mtb from 4-hydroxyquinoline treatment under non-replicating conditions suggesting a different MOA. WGS of drug resistant strains raised under non-replicating conditions did not yield informative results reflecting the limitation of using resistance mutants to rationalize bacterial chemical phenotypes.

**Unknowns and further confirmation.** It may seem paradoxical that a high drug resistance rate can help to unravel the drug MOA, but from a drug development point of view, being able to select a resistant mutant easily in an essential gene is not necessarily a good indication for an ideal drug target candidate. A good drug candidate and an ideal drug target both require a low drug resistance rate, which in turn could result in unsuccessful formations of resistant strains. Consequently, other approaches for drug MOA elucidation are required.

Moreover, our understanding of Mtb's genome is still limited; of 4000 genes in the Mtb genome, 25% encode hypothetical proteins with unknown biochemical functions.<sup>44</sup> Thus, it is challenging to pinpoint the MOA of an individual compound solely relying on resistant mutant sequencing. As the puzzles of many drug mechanisms remain unsolved, researchers need to employ more comprehensive and multifaceted approaches to answer the complex questions posed in/by target validation.

## Transcriptional profiling

### Bacterial transcriptional adaptation to antibiotic perturbations

When generating resistant mutants is not possible or useful, what other genetic approaches can be applied to uncover mechanistic information? Bacteria often respond to chemical or physical perturbations by altering the level or activity of





transcription factors and initiate rewiring of cellular regulatory networks.<sup>45,46</sup> The alteration of the transcriptome often relies on DNA binding proteins, activators or repressors, which interact with a specific DNA sequence in promoter regions. Genome wide transcriptional responses of Mtb to an antibiotic can be used as a supplementary tool to understand the expression level of key genes or transcription of critical metabolic pathways, which are critical for adaption to the disruption caused by antibiotic treatment. This information helps to reveal the primary target and the events downstream of primary target inhibition that actively contribute to drug action. As such, transcriptional profiling has been widely utilized to predict and validate the MOA of TB drugs during lead optimization<sup>47,48</sup> and has been utilized to understand and predict drug synergy in order to yield better therapeutic outcomes.<sup>49</sup>

### Predicting drug MOAs by transcriptional profiling in Mtb

Transcriptional profiling has become a mainstay in TB drug discovery since its introduction to Mtb research in early 2000. Through microarray-based genome wide transcriptional profiling, Barry and colleagues provided an encyclopedic summary of Mtb's transcriptional response to 430 antibiotics and to inhibitory growth conditions.<sup>47</sup> This work revealed 150 gene clusters in Mtb's transcriptome that are jointly regulated to adapt to metabolic perturbations. These gene networks serve as signatures that are specific to certain types of metabolic inhibition, such as cell wall synthesis, cell respiration, DNA transcription, and translation. Drawing on the work of Barry *et al.*, with the broad application of RNA-seq technology, transcriptional profiling has been applied to study Mtb's pathogenesis in terms of how Mtb responds and adapts to different growth and stress conditions,<sup>50–52</sup> and more importantly, has laid an essential foundation for the use of transcriptional profiling to understand and predict the MOAs of TB drugs (Fig. 3).

### Advantages of transcriptional profiling in TB drug MOA prediction

When compared to WGS of resistant mutants, transcriptional profiling can reveal additional information on the MOA of an antibiotic and its impact on the global mycobacterial transcriptome. Pretomanid works against both replicating and non-replicating Mtb and its activity requires bio-reductive activation of the aromatic nitro group and concomitant nitric oxide release.<sup>10,37,53</sup> The bactericidal activity under aerobic conditions is associated with pretomanid's inhibition of mycolic acid biosynthesis.<sup>10</sup> To elucidate its killing effect under hypoxia, transcriptional profiles of Mtb under pretomanid treatment were examined.<sup>54</sup> Through this approach, cell wall synthesis inhibition was validated and a new aspect of the MOA: inhibition of respiration causing an energy production constraint, was discovered. Upon drug treatment, genes involved in respiration, such as the *cyd* operon encoding the cytochrome *bd* complex, were upregulated, and consequently the cellular redox and energy homeostasis were disrupted. These observations led to a proposed mechanism of pretomanid under hypoxia that the nitric oxide released from the nitro bio-reduction of pretomanid poisoned the respiratory electron transport chain and disrupted energy homeostasis.

Transcriptional profiling can help decipher how bacteria respond to mediate the bactericidal or bacteriostatic effects caused by antibiotic treatment. As demonstrated in a transcriptional network analysis study of BDQ on Mtb, after 96 h of BDQ treatment, in addition to genes in the ATP synthase operon which is the primary target of BDQ, upregulation of two gene regulators Rv0324 and Rv0880 aid in tolerance of Mtb to BDQ.<sup>55</sup> Upon their upregulation, Mtb enters into a BDQ-tolerant state, which manifests as delayed killing during the early stage of drug treatment.<sup>56</sup> Genetic disruption of these two regulators accelerated the killing effect of BDQ. Therefore, transcriptional profile analysis revealed two critical genes as nontraditional drug targets that have the potential to reduce the treatment

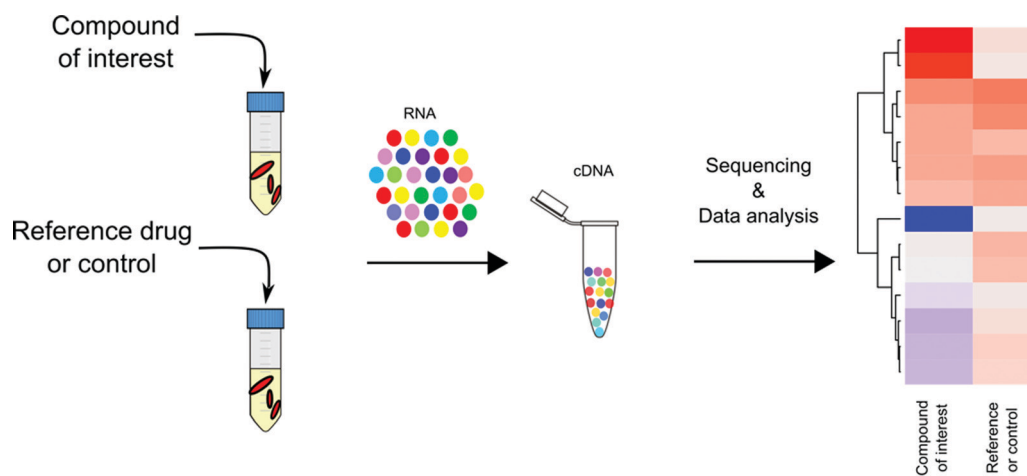


Fig. 3 General workflow of transcriptional profiling for TB drug MOA assessment. Mycobacterial total RNA is isolated from Mtb cultures with or without drug treatment and purified RNA is converted to cDNA. The resulting cDNA libraries are subjected to quantitative sequencing analysis for transcriptional profiling.



duration of BDQ TB therapy. These two genes would likely be dismissed as drug targets due to their non-essentiality to Mtb growth, yet here their importance for BDQ MOA suggests a potential therapeutic strategy to sensitize Mtb to BDQ.

Small molecules like vitamin C and cysteine have been shown in recent studies to induce enhanced production of reactive oxygen species (ROS) in Mtb, such that vitamin C treatment can sterilize Mtb culture<sup>57</sup> and cysteine can sensitize Mtb to INH or rifampicin treatment.<sup>58</sup> As demonstrated by Vilcheze *et al.*, these small molecules may not have a primary target in Mtb, instead they rather act in a more general way by disrupting cellular machineries. Transcriptional analysis was used as a primary approach to investigate the mechanisms involved in the bacterial killing by these small molecules. Critical pathways and essential genes in the intermediary metabolism, respiration, and lipid biosynthesis that play significant roles in Mtb's response were identified. Subsequent biochemical experiments that were designed based on the information revealed by transcriptional profiling eventually led to the elucidation of the MOAs of these small molecules. Vitamin C causes bursts in oxidative stress levels in Mtb by generating Fe<sup>2+</sup> from Fe<sup>3+</sup> and subsequent production of ROS in the presence of oxygen *via* Fenton reactions, while cysteine enhances respiration and oxygen consumption in Mtb initiating elevated production of ROS. A similar mechanistic study was undertaken for a bicyclic 2-pyridone compound that was shown to reverse INH resistance in Mtb by disrupting respiration homeostasis.<sup>59</sup>

As bacterial strains can be engineered with controlled transcription levels of certain pathways, transcriptional profiling provides utility for target confirmation in drug mechanistic investigation. In a target-based whole cell screen, Zheng *et al.* selectively targeted the DosRST signaling pathway by using a DosRST-dependent fluorescent Mtb reporter strain.<sup>60</sup> After the hit compounds had been identified, transcriptional profiling studies were undertaken to investigate the inhibitory mechanism of the hit compounds by comparing the transcriptional response of wild-type Mtb upon compound treatment with the response of the  $\Delta dosR$  mutant. Among the 55 genes that were downregulated by hit compound HC102A, 48 genes were also down regulated by  $\Delta dosR$  mutation compared to wild-type Mtb. When the  $\Delta dosR$  mutant was treated with HC102A, no additional genes were repressed when compared to  $\Delta dosR$  mutant without compound treatment, validating that HC102A phenocopies  $\Delta dosR$  mutation and thus has a highly specific on target effect. Following the comprehensive transcriptional profiling studies, well designed biochemical based approaches were undertaken that elucidated the specific inhibitory mechanisms, which included inhibiting DosST sensor kinase heme and DosST autophosphorylation.

Transcriptional analysis and the related "real-time quantitative reverse transcription-PCR" (RT-qRT-PCR) method provide the versatility needed for drug MOA study in Mtb under a wide array of conditions. In addition to the heterogeneity of Mtb infection in the host,<sup>61</sup> the metabolic state and transcriptome of Mtb *in vivo* are different than in broth culture and dynamically change.<sup>51,62,63</sup> Thus, transcriptional and metabolic responses to TB drugs can be very condition dependent.<sup>64-66</sup> Therefore,

evaluations of drug MOA under *in vitro* growth conditions may not fully capture the entire spectrum of how a drug works or how Mtb responds to drug treatment during the *in vivo* infection. Transcriptional profiling of different *in vitro* models provides the possibility to evaluate Mtb's gene expression changes induced by antibiotics during different metabolic states, thus facilitating more comprehensive *in vivo* drug MOA studies.

### Limitations of transcriptional profiling in TB drug MOA prediction

Transcriptional profiling is a powerful tool in drug MOA evaluations to comprehend the effects of drug treatment on bacteria on a transcriptional level but is considered to be an indirect approach for drug MOA assessment. MOA predictions based on transcriptional profiling require further experimental validation or supporting evidence.

The transcriptional response of Mtb to antibiotics is dynamic during the course of drug treatment, and thus many factors affect the outcome. Different durations of drug treatment, different drug concentrations, and different growth conditions of Mtb may all yield different transcriptional responses.

Genome wide transcriptional profiling data analysis is based on the comparison of ~4000 measurements across the Mtb genome. Since it is a comparative analysis, the reference transcriptome or strain used in the experiment is essential for obtaining informative results. The reproducibility of transcriptome analyses pose a major challenge, as technical and procedural errors will dramatically affect experimental outcomes and complicate data analyses.<sup>67</sup> Furthermore, bacterial responses to antibiotic treatment are a combination of drug on-target inhibition and global stress responses; in order to understand the MOA of a drug from transcriptional profiling it is essential to distinguish the nonspecific stress responses from the real drug inhibitory response.

## Bacterial reporter strains

### Genetically engineered Mtb reporter strains for drug MOA prediction

Another aspect of transcriptional response-based drug MOA prediction and assessment is the utilization of genetically engineered bacterial reporter strains. As detailed in recent reviews,<sup>68,69</sup> specifically designed reporter strains of Mtb have been widely utilized in target-based whole cell screens for TB drug discovery and studies of pathogen–host interactions. A typical reporter strain contains a gene of interest or a promoter and a signal generating gene. The inhibition or induction of the gene of interest can be studied by monitoring the signal generated. The identification of inhibition-specific promoter genes in Mtb, such as *iniBAC* for cell wall synthesis inhibition<sup>70</sup> and *recA* for DNA damage,<sup>71</sup> have facilitated the exploitation of these promoters in constructing reporter strains. When a certain cellular metabolism is inhibited, the promoter is induced, the reporter is then expressed, and a readout signal such as luminescence or fluorescence is generated.



Three reporter strains generated by Naran *et al.* via the coupling of the bioluminescent *luxCDABE* operon to the Mtb *iniBAC*, *recA*, and *radA* promoters, which control expression of genes responsible for mycobacterial cell wall biosynthesis or DNA metabolism, provided a quantitative measurement in real time of the rate of induction of these promoters over a 10–12 day incubation period.<sup>72</sup> Based upon the time point at which luminescence occurs during the course of drug treatment, and the fold change of luminescence readout compared to control groups, this approach may indicate whether a compound is directly targeting a cellular process or if it is a secondary or downstream effect of the drug MOA. TB drugs directly targeting cell wall biosynthesis or DNA metabolism showed early induction of promoters during the first 4 days of treatment, whereas compounds that might have associated downstream effects on cell wall biosynthesis or DNA metabolism showed delayed responses. Moreover, the reporter strains based on *recA* and *radA* promoters distinguish between compounds that inhibit cell growth by inhibiting DNA metabolism from the general genotoxins, thereby eliminating compounds that possess cytotoxicity and act in a nonspecific manner.

### Strengths and limitations of Mtb reporter strains for drug MOA prediction

The readout of these inducible reporter strains is continuous and in real time, and unlike quantitative RT-PCR or transcriptome analysis, does not require cell disruption and can deliver results continuously over an extended duration. It is also possible to use reporter strains to study the dynamic changes of drug treatment during Mtb infection in a host system, by incorporating genes encoding fluorescent proteins and imaging tools, such that a better understanding of how an antibiotic acts against Mtb *in vivo* can be achieved.

Despite the advantages that these reporter strains provide in antibiotic MOA assessment such as rapid, quantitative and in real time detection, a positive signal from the reporter does not provide definitive proof of the MOA. Similar to transcriptional profiling, these reporter strains do not pinpoint the molecular target of the antibiotic and can only be used with a bias to examine suspected metabolic pathways that the compound is potentially affecting.

Moreover, the scope of the pathways that can be examined are limited and the construction of these reporter strains is highly dependent on the promoter genes chosen. As our understanding of Mtb's bacterial machinery and their regulatory mechanisms is still limited, available and suitable inhibition-specific inducible promoters that can be applied to this system are limited.<sup>73</sup>

## Chemical genetics interaction

### Reverse chemical genetics for antibiotic MOA assessment

Selecting spontaneous resistant mutants against compounds of interest and then linking the drug resistance phenotype to genotype is considered a forward chemical genetics approach.

Conversely, if a genome wide collection of mutants is available, the phenotype of drug susceptibility can be mapped to specific mutants by screening the compound of interest against the genome wide library of mutants. This strategy is considered a reverse chemical genetic approach<sup>74</sup> and is referred to as the approach of chemical genetics interaction in this article.

Mapping drug phenotype to gene deletion mutants was first established in yeast as a method to identify drug targets.<sup>75</sup> The same principles of chemical genetic interactions were soon adapted in the MOA assessment of antibiotics in bacterial organisms like *E. coli*<sup>76,77</sup> and *S. aureus*.<sup>74</sup> Genetic methods, such as antisense-based genetic overexpression or genetic knock-down were used to generate mutants with altered gene expression levels in order to test their fitness in the presence of compounds of interest at lethal concentrations. The central hypothesis of these approaches is that the chemical genetics interaction between a compound of interest and its molecular target will cause a change in drug susceptibility between wild-type and a genetic mutant related to the target. The hypo- or hypersensitivity to antibiotics will then lead to a fitness change in terms of bacterial growth and abundance, which can be quantified by sequencing of barcodes that are engineered with the genetic modification and unique to each specific mutant. Consequently, a chemical genetics interaction profile can be generated and the potential molecular target(s) of the compound of interest can be proposed.

### Mtb mutants with altered gene expression levels for biased drug MOA prediction

In biased drug screens targeting particular metabolic pathways, a small library of mutant strains from that targeting pathway can be used to assess with which potential proteins the hit compounds are interacting. For example, using a library of cholesterol analogs to target cholesterol metabolism in Mtb, Yang and coworkers discovered a number of azasteroids that could sensitize Mtb to existing TB drugs.<sup>78</sup> In order to elucidate the MOA of these compounds, a small library of knock out or transposon mutants of genes regulated by cholesterol were utilized to characterize bacterial fitness changes after drug treatment.<sup>78</sup> They discovered that azasteroid compounds lost their drug potentiation activities against *fadE18*, *melF*, and *melH* transposon mutants. These three genes are all transcriptionally repressed by the Mce3R regulator, indicating the involvement of Mce3R regulon in the MOA of azasteroid. From a growth inhibitor high-throughput screen in Mtb infected macrophages, VanderVen and coworkers identified a group of compounds that are conditionally active against Mtb in cholesterol-containing media indicating a cholesterol metabolism associated MOA. An Mtb  $\Delta icl1$  mutant was utilized to examine the compounds' ability to restore the growth of the  $\Delta icl1$  mutant in cholesterol-containing media. When grown on cholesterol, the  $\Delta icl1$  mutant is not able to detoxify propionyl-CoA accumulated from cholesterol catabolism. As a result, the  $\Delta icl1$  mutant is not able to grow in cholesterol-containing media by itself. Restored growth of the  $\Delta icl1$  mutant in cholesterol media thereby indicated that the MOA is associated with



inhibition of propionyl-CoA producing enzymes from either the cholesterol catabolism pathway or the downstream methylcitrate cycle.<sup>79</sup>

### Genome wide chemical genetics interaction profiles (CGIPs) in Mtb for drug MOA identification

The use of genetic methods to screen for inhibitors of essential genes in bacteria can be challenging to achieve. Prior to 2019, several groups undertook using saturated transposon mutant libraries for differential drug sensitivity screening to generate a CGIP in either *Mycobacterium smegmatis* (*M. smeg*)<sup>80,81</sup> or Mtb.<sup>82</sup> These CGIPs were used to identify the primary and secondary targets of TB drugs and to identify the genetic determinants of intrinsic antibiotic susceptibility. However, as transposon insertion leads to loss of function of gene products, some mutants carrying insertions in essential genes under certain culture conditions will not be represented in these libraries. As a result, the mutant pool size is limited and may not include essential targets that are most desirable.

An alternative approach is depletion of an essential gene rather than elimination of the target. Subsequent screening of compound libraries to find hits that target these hypersensitive strains provides a chemical-genetic interaction profile. In 2019, a library of 2014 hypomorphic strains representing 474 of approximately 625 essential genes in Mtb was constructed by Johnson and coworkers using conditional proteolysis or transcriptional control.<sup>83</sup> Using a pool of 152 of these hypomorphic alleles, they screened against a library of 50 000 bioactive small molecules and generated more than 7 million CGIPs in Mtb. This robust library generation strategy identified both antimycobacterial compounds and their candidate targets simultaneously (Fig. 4). The CGIPs generated from the screening uncovered compounds with novel targets, but also identified new chemical classes that act on the same targets as existing Mtb drugs.

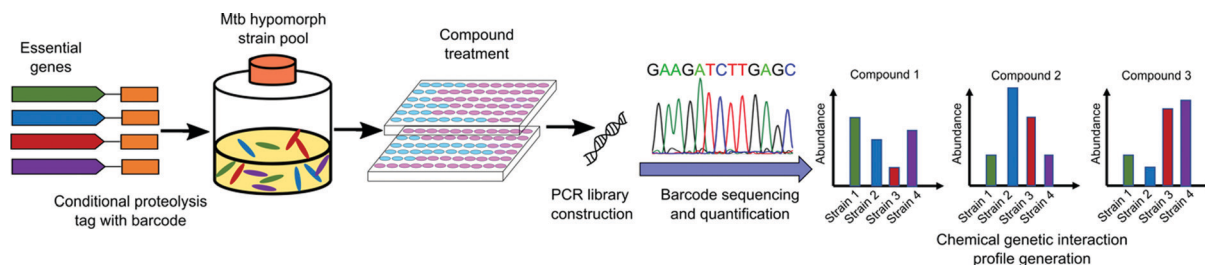
In this large-scale chemical genetics screening campaign, Johnson *et al.* prioritized compounds that did not strongly inhibit wild-type Mtb, but were strongly active against at least one hypomorph. Due to the static equilibrium conditions of MIC determination assays, solely MIC based compound prioritization strategy may bias toward selection of extracellular targets. Less stringent MIC cutoffs in initial whole-cell screens

may identify hits that can be developed into lead compounds with improved potency against vulnerable intracellular targets.<sup>8</sup> This approach led to the discovery of compound BRD-8000.3 which possesses a nanomolar MIC against wild-type Mtb by targeting an essential uncharacterized efflux pump in Mtb, EfpA.<sup>83</sup> Then, using the CGIP of BRD-8000.3 as a reference, Johnson and coworkers later reported a new compound, BRD-9327, which inhibited the same target in Mtb but displayed a different mode of action and resistance mechanism than compound BRD-8000.3. These two compounds are synergistic and in Mtb, resistance to one increases the sensitivity to the other, thereby presenting a potential drug combination to restrain acquired drug resistance in Mtb.<sup>84</sup>

More recently, CRISPR technology has been employed in chemical genetics to advance drug discovery and target validation. This technology can be utilized to knockout genes and to either inhibit or activate the expression of a gene without changing the genome. In drug-target validation, CRISPR interference editing provides an efficient and scalable platform for controllable gene expression silencing which is advantageous compared to other methods of transcriptional control due to lack of disruption of DNA sequence in the genome. It has been demonstrated in Mtb that CRISPR interference can achieve a 20 to 100-fold knockdown of target gene expression with minimum proteotoxicity and has been shown to be more effective than conditional proteolysis, thereby offering a great tool for generation of hypomorphic mutants of essential genes at genome scale.<sup>85</sup>

### Advantages and limitations for CGIP in identifying drug MOAs

Chemical genetics interaction profiling provides a systematic chemical biologic method to identify a molecular target from a genetic library. CGIP also affords insights into the secondary and tertiary molecular targets of a drug, which is integral to understanding a drug's MOA. As an approach to compound screening, CGIP facilitates rapid target identification and verification. CGIP allows hit compounds to be selected and prioritized on the basis of MOA, instead of potency alone. Additionally, CGIP enables a rapid expansion of the chemical entities inhibiting the same target, which may work through different modes of interaction, to tackle drug resistance.



**Fig. 4** Schematic representation of large-scale chemical-genetic interaction screening strategy for the discovery of new Mtb inhibitors. Mtb hypomorph strains are generated by integrating the conditional proteolysis tag, caseinolytic protease (Clp) DAS tag, with barcode at the C terminus of target genes of interest into the chromosome. Compounds are screened against pools of hypomorph strains. After compound treatment, chromosomal barcodes are quantified for changes in abundance relative to vehicle controls and sequenced by PCR amplification. The CGIP for each compound is represented as the change in strain abundance relative to control.<sup>84</sup>





However, some gene depletion strains can possess hypersensitivity to many bioactive agents, as manifested by the disruption of FecB, a putative iron di-citrate-binding protein, which confers resistance of Mtb to many antibiotic agents.<sup>82</sup> This may reflect the intrinsic synthetic lethality between the depletion of genes which interact with the drug target and the chemical inhibition of the target protein. Another limitation of this approach is that efflux pumps and detoxification mechanisms in the cells may prevent the primary molecular target from being exposed by reducing the effective cellular drug concentration.<sup>74</sup>

## Metabolomic profiling

### Metabolic adaptations to antibiotic perturbations can be studied

Metabolomics refers to the strategic combination of analytical methods in order to study the global change of small molecules in an organism as they are broken down or synthesized, otherwise known as metabolites.<sup>86</sup> The dynamic changes in these endogenous molecules are generally in response to standard biological perturbations, such as stress or nutrient shifts, but unique metabolic responses are observed after the introduction of antibiotics to the bacterial organism.<sup>87</sup> The total pool of small molecules in biological systems can be characterized and identified by analytical chemistry methodologies that generate signature liquid chromatography (LC) and mass spectrometry (MS) profiles.

Traditionally, metabolomic profiling has been applied in the biochemical characterization of novel metabolic enzymes and pathways. This approach has led to the discovery of the biochemical functions of many essential enzymes in Mtb, including enzymes in central carbon metabolism and lipid and amino acid biosynthesis.<sup>88</sup>

Conversely, the dynamic change and intracellular levels of small molecules can be mapped to a genotype as well. Understanding the changes in the levels of these biological molecules has been exploited as a useful tool in the identification of antibiotic function and more specifically in validating the function of an metabolic enzyme target.<sup>89</sup> In this vein, accumulation of substrates and decreases in metabolic products can generally signify that, either directly or indirectly, bacterial metabolic processes are being disrupted.

Metabolic enzyme inhibitors are optimal candidates for MS-based MOA studies due to the direct correlation between substrate or product concentration in the inhibited pathway and enzyme-target-identity. Popular drug targets in Mtb include folate and cell wall biosynthesis, transcription, translation, and DNA supercoiling.<sup>90</sup> These targets, among numerous other pathways, have been identified or verified by global metabolomic studies.

### Unraveling drug MOAs by metabolomic profiling in Mtb

Through early trials examining pretomanid efficacy in clinical isolates of Mtb drug resistant strains, it was apparent that pretomanid did not have the same resistance profile as other commonly used antibiotics, insinuating a novel mode of action.

Subsequent sequencing of pretomanid resistant mutants implicated F<sub>420</sub>-dependent glucose-6-phosphate dehydrogenase (G6PD) in susceptibility,<sup>37</sup> but the primary mode of action of pretomanid was not fully characterized. G6PD is not the direct target of pretomanid, rather G6PD is required to replenish the reduced F<sub>420</sub> cofactor required for deazaflavin dependent nitroreductase (Ddn) activation of pretomanid.<sup>91</sup>

More recently, the Mtb metabolome under treatment with pretomanid and nine other antibiotics was analyzed by gas chromatography (GC)-MS.<sup>92</sup> The analysis confirmed previous findings of a unique mode of action and provided further insights into the MOA of pretomanid. An entirely distinct set of responses was observed under pretomanid treatment in comparison to other classes of antibiotics, and the metabolomic data acquired illuminated the cause of the unique set of responses. The increased turnover of G6PD to recycle the F<sub>420</sub> cofactor leads to increased production of downstream metabolites of glucose-6-phosphate, phosphorylated aldehyde sugars, which through glycolytic mechanisms leads to the accumulation of methylglyoxal – a cytotoxic molecule detrimental to the overall fitness of the bacteria.<sup>92</sup> Whereas resistant mutant sequencing revealed the susceptible enzyme, G6PD, metabolomic profiling was required to identify the mode of bactericidal action.

Through WGS of resistant mutants, BDQ was thought to enact its antimycobacterial effects predominantly through the inhibition of mycobacterial ATP synthase. Drug-target inhibition is often oversimplified as the “sole basis of antibiotic induced cell death,” which is readily apparent in target-based drug design.<sup>93</sup> Wang *et al.* sought to elucidate secondary downstream effects of BDQ treatment *via* metabolomic analyses. The dose and time dependent profiles of 130 metabolites, most of which are involved in adenosine phosphate-dependent reactions, were examined in response to antibiotic treatment. A near quantitative correlation between ATP and glutamine levels revealed glutamine synthetase as a collateral target to ATP synthase inhibition, enhancing BDQ's anti-mycobacterial activity.<sup>94</sup> Whereas genomic based techniques were only able to identify mutations in the *atpE* gene, metabolite analysis exposed both ATP synthase inhibition as well glutamine synthesis vulnerability.

An amidino-urea scaffold was identified from a high-throughput screening for its antimycobacterial activity both *in vitro* and in TB animals by Ballinger and coworkers. To explore the MOA of compound 8918, its impact on the Mtb metabolome was examined in parallel with four TB drugs with known mechanisms.<sup>31</sup> Several molecules from coenzyme A (CoA) biosynthesis and CoA utilizing pathways were discovered to be affected specifically by the amidino-urea compound in a concentration dependent manner. Subsequently, WGS of spontaneous resistant mutants confirmed the target of 8918 to be the phosphopantetheinyl transferase (PptT) in lipid biosynthesis, that is responsible for pantothenate transfer from CoA to acyl carrier proteins (ACPs) and enables ACPs to synthesize structural lipids. Compound 8918 was the first reported bactericidal compound against Mtb which functioned by inhibiting the CoA metabolism pathway.



Some drugs in bacteria activate compensatory mechanisms to endure antibiotic stresses, this response type can be utilized to elucidate the genuine drug target and to understand drug resistance. In the case of fluoroquinolones (FQ) which inhibit bacterial DNA gyrase, metabolomic studies identified an increase in cellular dTDP-rhamnose levels as a survival mechanism to mediate resistance to FQs. Increased concentration of dTDP-rhamnose, an activated sugar metabolite used for modifying polysaccharide capsule, results in increased DNA gyrase A transcription, thereby increasing the amount of FQ target in the cell for sequestration of antibiotic. *gyrA* expression is induced by a respiratory metabolic transcription factor AroA. In turn, *gyrA* expression inhibits dTDP-rhamnose production.<sup>95</sup> This gene circuit is an example of the interplay between DNA replication and metabolism.

### Strengths and challenges of metabolomic profiling for unraveling TB drug MOAs

The overview of global biochemical events that is revealed through metabolomics enables researchers to obtain a clearer picture of the molecular events occurring during antibiotic treatment. Similar to transcriptome analysis, control parameters and experimental design are key to achieve comprehensive and informative global metabolomic studies due to the dynamic nature of metabolism during bacterial growth.<sup>96</sup> Coupled with extensive knowledge of metabolic pathways, “blockages” caused by exogenously inhibited gene function can be traced back to their source. It is through these efforts that MS metabolomics has been effectively utilized to uncover the mechanisms of action of current TB therapies.

Metabolomic analysis has the capability to be implemented in a high-throughput manner. Experiments can be automated on a  $\mu$ -volume scale in 96–384 well plates when appropriate methods are utilized, enabling drugs of interest to be examined in a robust manner.<sup>97</sup> The ability to customize metabolomics, specifically data set parameters, by screening for certain metabolites through the use of specialized mass spectrometry instruments adds to its utility.<sup>98</sup>

The ability to choose between targeted and untargeted studies facilitates the dual modality of metabolomics for use in both large- and small-sample size experiments. The global analyses of eight small molecule inhibitors with known MOAs by Vincent *et al.*<sup>96</sup> and the pseudo-targeted screening of nearly 1300 chemically diverse compounds by Campos *et al.*<sup>99</sup> are excellent examples of the targeted/untargeted dichotomy that is afforded by MS based metabolomics.

However, challenges still remain. In untargeted metabolomics, which is useful as a frontline strategy to uncover global information, the sheer size of data sets can be difficult to analyze. The use of pseudo- and targeted experiments, in which ions of particular molecules are selected for, assuages many of the complications associated with big data. Although metabolomic profiling is a useful tool in bridging the gap between intracellular biomolecular changes and drug–target inhibition, the accumulation of metabolic molecules is often a reflection of the disruption of a certain cellular machinery. Tracing these downstream effects

to their primary drug target is challenging and requires a combination of genetic and biochemical approaches.

## Biochemical affinity purification: target pull-down

### Direct capture of drug-interacting proteins through chemical strategies

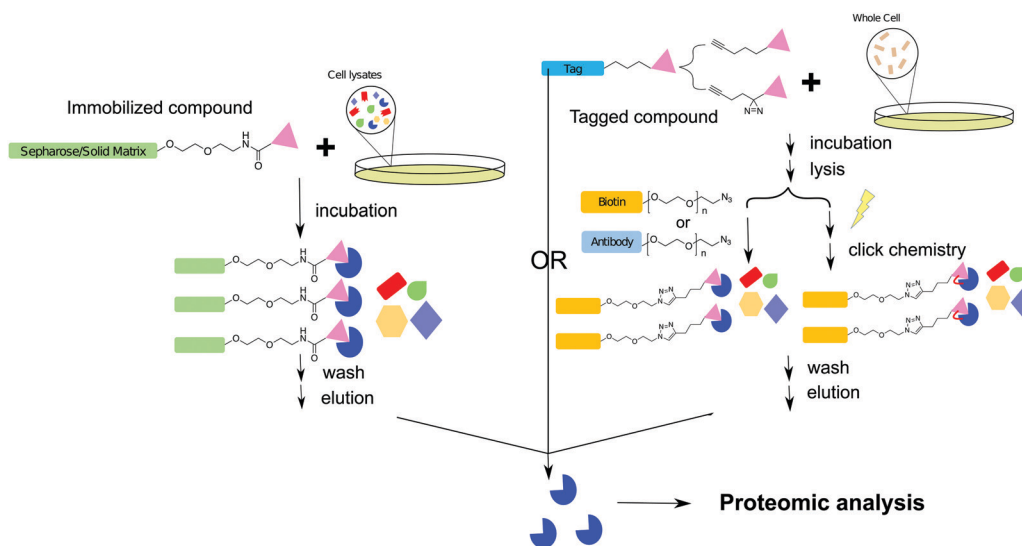
The biochemical affinity purification or target pull-down assay is a more tangible and direct method of drug–target identification in comparison to the other techniques described in this review. This technique has been applied in drug target identification and drug off-target evaluation since the 1950s, and this approach has been extensively reviewed in the field of chemical biology.<sup>100–102</sup> Like Paul Ehrlich’s magic bullet theory, the target pull-down approach is established on the concept that most drugs inhibit a target protein by direct binding in an exclusive and highly specific manner. The goal of a pull-down assay is to capture and characterize the protein(s) that an inhibitor strongly interacts with in a cellular context.

The general workflow of a traditional target pull-down assay consists of immobilizing the small molecule of interest onto a solid matrix through a linker. After cell lysates have been passed through the solid matrix, and the matrix washed, the immobilized protein can be eluted, isolated, and characterized. With the development of analytical technologies and chemical synthetic approaches, several alternative methodologies have been developed, where in place of immobilization, the small molecule can be modified with an analytical handle or tag. In this scheme, the intact cell is treated with the small molecule which binds the target protein. The target is then isolated from cell lysates *via* the handle or tag (Fig. 5).<sup>103</sup> Once the inhibitor–enzyme complex is separated from cell lysates, the sequestered protein can be identified through a variety of methods. Briefly, these include examination of whole proteins *via* LC-MS, SDS-PAGE, or native-PAGE, as well as trypsin digestion followed by identification of peptide fragments. Peptide fragments may be analyzed *via* MS methods such as Matrix Assisted Laser Desorption Ionization (MALDI) and with the aid of software, protein identity can be ascertained.<sup>104</sup> This approach is also often referred as chemical proteomics.

### Direct target pull-down in the identification of TB drug targets

As demonstrated by Schmitt and coworkers, a target pull-down approach was undertaken to identify the protein target of the natural product cyclomarin in *Mtb*.<sup>105</sup> Initial generation of spontaneous resistant mutants was not successful, indicating a low mutation frequency, and thus, the need for an alternative approach to target identification. From cell lysates the target protein was captured *via* sepharose immobilized cyclomarin and identified to be ClpC1 (Rv3596) by proteomic analysis. ClpC1 is a part of the ATPase superfamily and functions as a regulatory subunit with chaperone activity that delivers proteins to the protease subunits of a protease complex. Interestingly, cyclomarin is also an effective antibiotic for treatment of





**Fig. 5** Workflow schemes of affinity purification for target protein isolation. The direct pull-down is achieved by immobilizing the compound of interest on a solid support. After incubating with cell lysates and washing out non-specifically bound entities, the target protein is isolated for proteomic analysis. In click chemistry or photo crosslinking assisted target pull-down, the functionalized compound of interest can be used in the context of whole cells. After incubation and cell lysis, the compound bound proteins are captured through covalent interactions with the capture ligand. After subsequent washing, capture and elution steps, the identities of isolated proteins are characterized by proteomic analysis.

*P. falciparum*, the causative parasite of malaria. The MOA was discovered through an analogous pull-down experiment [compared to the Mtb MOA assessment] and implicated a diadenosine triphosphate hydrolase, unrelated to the Clp protease complex previously confirmed as the target in Mtb.<sup>106</sup> Thus, the cyclomarin MOA is a curious twist on the concept of polypharmacology because it targets two unrelated proteins that are present in two taxonomically unrelated pathogens.

#### Click chemistry and photo crosslinking enabled target pull-down in identifying TB drug targets

With the discovery and development of click chemistry,<sup>107</sup> azido-alkyne functionalization often is applied to small molecules for affinity purification of their drug targets. The compound of interest is modified to contain either an azide or alkyne functional group with the aim of not reducing or changing the original bioactivity. The fundamentals of copper catalyzed azide-alkyne click chemistry have been extensively reviewed.<sup>108,109</sup> Copper free click cycloaddition is an alternative biorthogonal method often utilized under biological conditions which has also been adapted for target pull-down assays.<sup>110</sup> The complimentary functional group can be coupled to a secondary molecule such as an antibody or biotin, which is then, respectively, captured *via* a primary antibody or streptavidin bound immobilized surface (Fig. 4).<sup>111</sup>

Zhao *et al.* used a click chemistry based pull-down experiment to verify the target of a benzoxazinone inhibitor discovered *via* a target-based biochemical inhibitor screen against MarP, a serine protease that regulates intra-bacterial pH homeostasis in Mtb.<sup>112</sup> After treating *Mycobacterium bovis* BCG and performing the target pull-down protocol, the alkyne modified antibiotic bound to target was “clicked” to an affinity handle and the complex captured with streptavidin beads. MarP was

identified by immunoblotting with antibody raised against the purified protein. Analysis of the SDS-PAGE gel with fluorescently tagged streptavidin revealed several different bands. Peptide mass fingerprinting of these bands identified high temperature requirement A1 (HtrA1) as an additional protein that interacts with the benzoxazinone inhibitor. This alternative target, which was undetected in earlier genetic screens against benzoxazinone inhibitors, is an essential enzyme in Mtb and thus warrants further study as an antimycobacterial target.

Suicidal inhibitors modify the target protein through covalent interactions, thereby presenting a class of suitable candidates for affinity pull-down without cross-linker modifications. Suicidal inhibitors can be designed from a target-based approach. Thus, the pull-down may bias towards the designated protein target, but they can still be explored in the verification of MOA intracellularly. There are many examples of suicidal inhibitors against Mtb including benzothiazinones (BTZs) inhibiting DprE1<sup>113</sup> and oxathiazol-2-ones inhibiting Mtb proteasomes.<sup>114</sup> In contrast, noncovalent interactions between the small molecule and target protein in the cell are affected by many factors, such as drug residence time and thermodynamic binding affinity. A pitfall of noncovalent inhibitors is that interactions are not permanent and so compounds that do not covalently bind to their target may not function as optimal candidates for pull-down experiments as target dissociation will prevent the identification of the target. To circumvent this issue, photo crosslinking chemical moieties have also been included in the construction of small molecule probes for target pull-down assays. Photoreactive covalent linkers, such as benzophenone or diazirine moieties, which upon irradiation with ultraviolet light insert into C–H, N–H, or O–H bonds,<sup>115</sup> are highly practical reagents for resolving the matter of covalent attachment of inhibitors with some caveats.



Although not in Mtb, a combination of click chemistry and photo crosslinking based affinity pull-down was used to reveal the antibiotic mechanism for a class of kinase inhibitors against methicillin resistant *Staphylococcus aureus*.<sup>116</sup> To determine target identity and MOA, the lead compound PK150 was modified to contain a diazirine photo-linker and alkyne tag. After incubation with intact bacteria, photoreaction and conjugation to a rhodamine–biotin–azide tag through click chemistry facilitated protein capture. Enrichment *via* avidin beads and subsequent analysis by tandem LC-MS implicated SbsB, an essential membrane serine endopeptidase from the protein secretion pathway and MenG, which catalyzes the final step in menaquinone synthesis, a vitamin crucial for bacterial respiration and energy metabolism as the targets of PK150.<sup>116</sup> This dual functionalization by combination of diazirine photo-linker and alkyne tag ensured a successful capture of the potential target proteins and can be applied to pull-down experiments in Mtb for non-covalent inhibitors.

### Advantages of target pull-down in drug MOA unraveling and considerations in experimental design

While the pull-down assay is often considered the proverbial “fishing expedition” in the drug–target validation space, it has the ability to discern novel mechanistic information that may be overlooked by other target identification techniques. Tetrahydropyrazolo[1,5- $\alpha$ ]pyrimidine-3-carboxamide (THPP) was originally identified for its antimycobacterial activity through a large phenotypic screening campaign.<sup>117</sup> Original MOA assignment through WGS of drug resistant mutants revealed that the compound functioned *via* targeting MmpL3, an essential trehalose monomycolate lipid transporter in Mtb.<sup>118</sup> However, through a chemical proteomics strategy where a THPP analogue was covalently immobilized on sepharose beads, EchA6 was identified as the only protein that bound selectively to the compound compared to negative controls.<sup>119</sup> A series of biochemical evaluations validated the conclusion that EchA6 is the molecular target of THPP. In contrast, MmpL3 may act as a THPP transporter which accounts for the drug resistance obtained from spontaneous mutagenesis.

The design of target pull-down probes must be meticulous. Key to a successful target pull-down, the modification of the small molecule not only has to retain the antibacterial activity of the parental compound, but also must minimize nonspecific interactions with other proteins. For photo crosslinking type probes, the covalent linker must be small enough to not interfere with binding, and the irradiation resulting in photoreaction should be rapid and not interfere with biological functions. Benzophenones undergo photoreaction at long wavelengths, avoiding interfering absorbance by abundant nucleotides and their subsequent damage, but due to their bulkiness may interfere with inhibitor binding and require longer irradiation times. Aryl-azides are small but undergo photoreaction at short wavelengths that are harmful to biological molecules, and the reactive nitrene species formed by photoreaction often decreases photolabeling yields. Diazirines represent the optimal balance between wavelength and size of

the covalent linkers, their predominant drawback being the intricate synthetic route to make them.<sup>120</sup>

The tag used to capture the enzyme-inhibitor complex may also complicate the pull-down assay. When directly linked to the inhibitor, the tag can cause decreased target affinity due to its size. The concern of reduced affinity can be resolved by separating the tag from the inhibitor and instead, including a small reactive functionality to generate the tag after the enzyme-inhibitor complex has been formed. The azide–alkyne click reaction, as previously described, is an exemplary reactive pair for this type of tag attachment.

As nonspecific binding is difficult to avoid in affinity purification, control experiments and comparison groups are critical for identifying the real protein target in the cell. Enrichment and statistical analysis are often used to eliminate the interfering components. After the potential molecular target has been identified, the chemotype needs to be confirmed by the genotype, *i.e.* manipulation of the tentative target's gene expression level should yield a change in drug susceptibility. We refer to several reviews detailing the efforts that can be taken to ensure the success of the experiment.<sup>121,122</sup> Overall, the target pull-down coupled with MS proteomic analysis offers the most comprehensive and direct method of determining the interactions of an inhibitor and its corresponding targets and is a valuable tool in MOA discovery.

## Machine learning and computational inference

### Computational machine learning in predicting antibiotics' MOAs

Given recent advancements in artificial intelligence technologies, machine learning has become more and more popular in the field of antibiotic drug discovery,<sup>123</sup> to virtually screen antimicrobial agents and predict the MOAs and efficacies of hits. Machine learning (ML) involves the feeding of data sets, called training datasets, to programmer generated computer algorithms. These data sets are utilized to train the program in order to make conclusions based on experimentally obtained data and generally are derived from existing bioinformatic databases. Adaptive improvements are more effectively and successfully accomplished with large data sets of high-quality information. Deep learning (DL) is a subset of ML that has been gaining traction due to increased computing power and the increasing scale and complexity that is afforded by such advancements. The principles of ML algorithms are summarized in depth in reviews by Chen *et al.*<sup>124</sup> and Jung.<sup>125</sup>

This section will be centered on the utilization of ML techniques to determine or predict the target of antibacterial compounds from phenotypic screening assays, although not directly in Mtb. Prediction of targets can be made using a multitude of different types of data, from metabolomic, proteomic, and transcriptional, to other forms of genetic data.<sup>126</sup> A wide range of suitable data enables the application of ML/DL to a variety of biochemical experiments and makes this data





analysis technique versatile and in reach of many laboratories through collaborative efforts.

The general approach of ML is to monitor the global changes in biochemical state caused by the introduction of the antibiotic and to relate these results to the causative molecular target(s). These computational results can later be verified by experimentally by altering the gene expression level of the proposed target(s). Protein function annotation can guide the prediction of causal events. The genomic sequencing of the bacterial genome and the accurate annotation of the proteins encoded contributes greatly to the success of ML. Application of this combination has been successful in the mechanistic verification of antibiotics in *E. coli*. Yang *et al.* utilized machine learning with carefully curated biological network models in order to elucidate direct mechanistic insights into how antibiotics perturb the biological environment of an organism.<sup>127</sup> Their approach differs from traditional modern ML predictive techniques which generally lack a biological network model to enrich predictive models. The lethality of three antibiotics with distinct mechanisms of action were screened in *E. coli* using MS to analyze metabolomic perturbations. Benchmarking studies were used to compare lethality contributions with and without biological network modeling. Inclusion of networks allowed depiction of the contributions of specific metabolic pathways to the lethality of the antibiotics.

Gardner *et al.* utilized machine learning to identify drug MOA via multiple linear regression modeling of a nine-transcript subnetwork of the SOS pathway in *E. coli*. Quantitative RT-PCR was used to monitor transcript levels of the set of nine genes under a variety of perturbation combinations and to train the algorithm. The small molecule chemotherapeutic mitomycin C (MMC) was then used to treat *E. coli* culture, and the same nine gene transcripts were monitored in order to understand the transcriptional effects of MMC on the SOS pathway. All genes in the network that were examined underwent large transcriptional changes, but the *recA* gene was correctly identified as the MMC transcriptional target by the ML method.<sup>128</sup>

ML has yet to be widely applied in drug mechanistic investigations in Mtb. However, with more and more data available from the systematic studies of the Mtb metabolome, transcriptome, genome, and their corresponding phenotypes, ML will be on the forefront of TB drug discovery. By utilizing a “systems” approaches to reveal underlying drug mechanisms, ML, unlike other methods, can provide information on how events downstream of primary target inhibition actively participate in antibiotic lethality.<sup>129</sup> This tool is most powerful with comprehensive experimental data sets of systems under study, thus highlighting the potential of ML for deepening understanding of Mtb pathogenesis and drug MOA.

### Computational inference in predicting drug MOAs in Mtb

Computationally aided target inference explores chemical and structural databases, to predict drug MOA based on existing similar experimental results. At GlaxoSmithKline researchers

exploited an integrative three-dimensional computational analysis, based on chemical structure, protein ligand complex structure, and historical in-house registrations from biochemical assays, to predict drug targets from two large phenotypic screening campaigns in Mtb.<sup>130,131</sup> The algorithm compares the chemical structural similarity of input compounds to compounds with experimentally validated targets from the ChEMBL database. It searches the structural similarity of input compounds to ligand structures from protein complexes in the Protein Data Bank, and it also predicts target information based on existing in-house biochemical assays. This integrative approach allows the target prediction of hundreds of hit compounds from large screening campaigns.

## Concluding remarks and future perspectives

Over the past five decades, only two new small molecule therapeutics have been approved for the treatment of TB disease. With the hundreds to thousands of compounds in the development pipeline, it is difficult to perceive this reality. Why is it so difficult to deliver new TB drugs to market, especially considering the largely unmet needs that exist in TB treatment?

There are multiple contributing factors that complicate TB drug development. When we consider the history of the tuberculosis pathogen, which has historical mentions dating back nearly 3300 years to the first written records in India and China,<sup>132</sup> it is considered to be one of the most successful pathogens in human history. Mtb manages to escape eradication by human immune defenses to reside in a niche within the hosts lungs.<sup>133</sup> This niche, characterized as a granulomatous cellular environment, is difficult for antibiotics to penetrate.<sup>134</sup> Coupled with a slow growth rate, metabolic heterogeneity, and the formation of persistent mycobacteria, it is exceedingly challenging for antibiotics to sterilize infected tissue.<sup>135</sup> In addition, there is a high attrition rate of compounds in the pipeline due to unfavorable PK/PD and safety properties. Comprehension of the MOA of hit/lead compounds in the pipeline and how the MOA correlates to physiological metabolic state of Mtb is crucial for development of improved compounds.

Novel targets and thus MOAs are essential to develop improved compounds and to combat existing drug resistance, and target validation for clinical use is equally crucial. MOA assignments for compounds from phenotypic screens can unveil novel drug targets, such as MmpL3 and DprE1, that are vulnerable to small molecule inhibition and essential for bacterial growth.<sup>8</sup> These molecular targets can potentially facilitate a new series of target-based drug discovery programs to deliver new chemical entities with improved drug profiles.

Each strategy described in this review for MOA identification of TB drugs has associated strengths and limitations (Table 1). Some techniques may be more accessible, while others may require established platforms and prove to be technically challenging. A single appropriate approach most likely only serves to provide a starting point for hypothesis and validation.



Table 1 Summary of approaches for TB drug MOA investigation

| Approaches                                   | Advantage   | Limitations  |
|--|---|--|
| WGS of spontaneous resistant mutant          | <ul style="list-style-type: none"> <li>• Potential to reveal the direct drug target</li> <li>• Commonly used and relatively straightforward</li> <li>• With next generation sequencing, the approach has the potential for high-throughput screening</li> </ul> | <ul style="list-style-type: none"> <li>• Direct drug target mutations are not guaranteed</li> <li>• Non-specific mutations may arise</li> <li>• Mutations in drug modifying genes <i>i.e.</i> pro-drugs</li> <li>• Mutations in related but non-target genes <i>i.e.</i> transporters</li> <li>• Resistant mutants can be hard to raise</li> <li>• Further confirmation necessary</li> </ul> |
| Transcriptional profiling                    | <ul style="list-style-type: none"> <li>• Reveals global response to the drug</li> <li>• Reveals the downstream effect of primary drug target inhibition</li> </ul>  | <ul style="list-style-type: none"> <li>• May not reveal the direct drug target</li> <li>• Nonspecific drug-stress response interference</li> <li>• Experimental design can dramatically affect the outcome</li> </ul>  |
| Reporter strain                              | <ul style="list-style-type: none"> <li>• Fast and continuous readout</li> <li>• Can be used to dynamically study the drug MOA in real time</li> </ul>   | <ul style="list-style-type: none"> <li>• The inhibition-specific promoter genes in <i>Mtb</i> are limited</li> <li>• Does not pinpoint the molecular target of a drug</li> </ul>   |
| Chemical genetics interaction profiling      | <ul style="list-style-type: none"> <li>• High throughput</li> <li>• Can be used to access genome wide drug–target interactions</li> <li>• Reveals downstream molecular targets of a drug</li> </ul>   | <ul style="list-style-type: none"> <li>• Efflux pump and detoxification mechanism interference</li> <li>• Nonspecific hyper- or hyposensitivity</li> </ul>   |
| Metabolomic profiling                        | <ul style="list-style-type: none"> <li>• Insight into target inhibition at molecular level</li> <li>• High throughput</li> <li>• Global metabolome analysis reveals additional information secondary to primary target inhibition</li> </ul>                    | <ul style="list-style-type: none"> <li>• Tracing upstream metabolites to primary drug target can be challenging</li> <li>• Based on comparison and requires an annotated database that includes identified metabolites</li> <li>• Data analysis can be challenging</li> <li>• Requires extensive mass spec. setup for analysis conditions</li> </ul>   |
| Target pull-down                             | <ul style="list-style-type: none"> <li>• Direct and straightforward – capture and analysis of target</li> </ul>   | <ul style="list-style-type: none"> <li>• Chemical modification of the drug is needed</li> <li>• Non-specific binding interference</li> <li>• Requires secondary analysis <i>i.e.</i> proteomics</li> </ul>   |
| Machine learning and computational inference | <ul style="list-style-type: none"> <li>• Fast</li> <li>• Wet lab experiments not required</li> <li>• High throughput</li> </ul>   | <ul style="list-style-type: none"> <li>• Established on the basis of large data sets</li> <li>• Further confirmation is needed</li> <li>• Prior knowledge/expertise in computer languages required</li> </ul>  |

A comprehensive and indisputable assessment requires several lines of evidence including biochemical, genetic, metabolomic and structural-biological confirmations. The mechanisms with which TB drugs impart their activity are often complex. Even after years of clinical use, new aspects of how TB drugs function are still being uncovered. A single approach to uncover MOA, such as those described in this review, will not be sufficient, and a combined and comprehensive approach is required to effectively identify the mechanisms of a single agent.

Looking back, an antibiotics' PK/PD properties and safety profiles are directly linked to its MOA. The on target and off target effects of an antibiotic are reflective of the cellular machineries that are perturbed. Looking ahead, it is essential to continue current work to further understand the physiology and pathogenicity of *Mtb* to develop more comprehensive databases for systems-based assessment of drug MOA. Completing the cycle, drug MOA identification in *Mtb* will improve our understanding of the pathogenic bacterial machineries of *Mtb* and motivate more advanced screening strategies for new inhibitory chemical entities.<sup>136</sup> Novel approaches, such as exploiting the synthetic lethality gene network in bacteria,<sup>137</sup> utilizing machine-learning and artificial intelligence for revealing lethality mechanisms,<sup>127,138</sup> and computationally aided compound analysis<sup>139,140</sup> should all be used to elucidate MOA of *Mtb* drug candidates. Wider application of these approaches will advance translation of compounds into the

clinic for treatment of TB, a disease which remains persistent and deadly.

## Glossary of abbreviations

|                |                                       |
|----------------|---------------------------------------|
| ACPs           | Acyl carrier proteins                 |
| BDQ            | Bedaquiline                           |
| CFU            | Colony forming unit                   |
| CGIP           | Chemical genetics interaction profile |
| CoA            | Coenzyme A                            |
| DL             | Deep learning                         |
| FDA            | Food and Drug Administration          |
| FQ             | Fluroquinolone                        |
| G6PD           | Glucose-6-phosphate dehydrogenase     |
| GC             | Gas chromatography                    |
| INH            | Isoniazid                             |
| LC             | Liquid chromatography                 |
| <i>M. smeg</i> | <i>Mycobacterium smegmatis</i>        |
| MIC            | Minimum inhibitory concentration      |
| ML             | Machine learning                      |
| MMC            | Mitomycin C                           |
| MOA            | Mechanism of action                   |
| MS             | Mass spectrometry                     |
| <i>Mtb</i>     | <i>Mycobacterium tuberculosis</i>     |
| PD             | Pharmacodynamics                      |



|            |  |
|------------|--|
| PK         | Pharmacokinetics   |
| PZA        | Pyrazinamide   |
| ROS        | Reactive oxygen species  |
| RT-qRT-PCR | Real-time quantitative reverse transcription-polymerase chain reaction |
| SAR        | Structure–activity relationship  |
| SNPs       | Single nucleotide polymorphisms  |
| SOS        | Cellular global response to DNA damage                                 |
| TB         | Tuberculosis   |
| THPP       | Tetrahydropyrazolo[1,5- $\alpha$ ]pyrimidine-3-carboxamide             |
| WGS        | Whole genome sequencing  |

## Conflicts of interest

The authors have patent interest in tuberculosis therapeutic treatments.

## Acknowledgements

The author's research described in this review is funded by NIH grant RO1AI134054 to N. S. S.

## References

- H. H. Fox, *Science*, 1952, **116**, 129–134.
- C. Vilcheze and W. R. Jacobs Jr, *Annu. Rev. Microbiol.*, 2007, **61**, 35–50.
- WHO, Global Tuberculosis Report, [https://www.who.int/tb/publications/global\\_report/en/](https://www.who.int/tb/publications/global_report/en/), (accessed 30th December, 2019).
- FDA, SIRTURO (bedaquiline) label, [https://www.accessdata.fda.gov/drugsatfda\\_docs/label/2012/204384s000lbl.pdf](https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/204384s000lbl.pdf), (accessed 30th December, 2019).
- FDA, Drug Approval Package: Pretomanid, [https://www.accessdata.fda.gov/drugsatfda\\_docs/nda/2019/212862Orig1s000TOC.cfm](https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/212862Orig1s000TOC.cfm), (accessed 30th December, 2019).
- M. N. Gwynn, A. Portnoy, S. F. Rittenhouse and D. J. Payne, *Ann. N. Y. Acad. Sci.*, 2010, **1213**, 5–19.
- H. Brotz-Oesterhelt and P. Sass, *Future Microbiol.*, 2010, **5**, 1553–1579.
- T. Yuan and N. S. Sampson, *Chem. Rev.*, 2018, **118**, 1887–1916.
- K. Andries, P. Verhasselt, J. Guillemont, H. W. Gohlmann, J. M. Neefs, H. Winkler, J. Van Gestel, P. Timmerman, M. Zhu, E. Lee, P. Williams, D. de Chaffoy, E. Huitric, S. Hoffner, E. Cambau, C. Truffot-Pernot, N. Lounis and V. Jarlier, *Science*, 2005, **307**, 223–227.
- C. K. Stover, P. Warren, D. R. VanDevanter, D. R. Sherman, T. M. Arain, M. H. Langhorne, S. W. Anderson, J. A. Towell, Y. Yuan, D. N. McMurray, B. N. Kreiswirth, C. E. Barry and W. R. Baker, *Nature*, 2000, **405**, 962–966.
- L. L. Silver, *Nat. Rev. Drug Discovery*, 2007, **6**, 41–55.
- A. L. Hopkins, *Nat. Chem. Biol.*, 2008, **4**, 682–690.
- D. C. Swinney and J. Anthony, *Nat. Rev. Drug Discovery*, 2011, **10**, 507–519.
- K. Pethe, P. C. Sequeira, S. Agarwalla, K. Rhee, K. Kuhen, W. Y. Phong, V. Patel, D. Beer, J. R. Walker, J. Duraiswamy, J. Jiricek, T. H. Keller, A. Chatterjee, M. P. Tan, M. Ujjini, S. P. Rao, L. Camacho, P. Bifani, P. A. Mak, I. Ma, S. W. Barnes, Z. Chen, D. Plouffe, P. Thayalan, S. H. Ng, M. Au, B. H. Lee, B. H. Tan, S. Ravindran, M. Nanjundappa, X. Lin, A. Goh, S. B. Lakshminarayana, C. Shoen, M. Cynamon, B. Kreiswirth, V. Dartois, E. C. Peters, R. Glynne, S. Brenner and T. Dick, *Nat. Commun.*, 2010, **1**, 57.
- L. L. Silver, *Clin. Microbiol. Rev.*, 2011, **24**, 71–109.
- M. Schenone, V. Dancik, B. K. Wagner and P. A. Clemons, *Nat. Chem. Biol.*, 2013, **9**, 232–240.
- M. A. Farha and E. D. Brown, *Nat. Prod. Rep.*, 2016, **33**, 668–680.
- K. Katsuno, J. N. Burrows, K. Duncan, R. Hooft van Huijsduijnen, T. Kaneko, K. Kita, C. E. Mowbray, D. Schmatz, P. Warner and B. T. Slingsby, *Nat. Rev. Drug Discovery*, 2015, **14**, 751–758.
- M. Tyers and G. D. Wright, *Nat. Rev. Microbiol.*, 2019, **17**, 141–155.
- S. Chakraborty and K. Y. Rhee, *Cold Spring Harbor Perspect. Med.*, 2015, **5**, a021147.
- Global Alliance for TB Drug Development, *Tuberculosis*, 2008, **88**, 85–86.
- J. M. Heather and B. Chain, *Genomics*, 2016, **107**, 1–8.
- S. B. Snapper, L. Lugosi, A. Jekkel, R. E. Melton, T. Kieser, B. R. Bloom and W. R. Jacobs Jr, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 6987–6991.
- S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead and B. G. Barrell, *Nature*, 1998, **393**, 537–544.
- Clinical Pipeline, <https://www.newtbdrugs.org/pipeline/clinical>, (accessed 6Apr2020, 2020).
- J. Crofton and D. A. Mitchison, *Br. Med. J.*, 1948, **2**, 1009–1015.
- P. E. Almeida Da Silva and J. C. Palomino, *J. Antimicrob. Chemother.*, 2011, **66**, 1417–1430.
- S. H. Gillespie, *Antimicrob. Agents Chemother.*, 2002, **46**, 267–274.
- J. W. Drake, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 7160–7164.
- H. L. David, *Appl. Microbiol.*, 1970, **20**, 810–814.
- E. Ballinger, J. Mosior, T. Hartman, K. Burns-Huang, B. Gold, R. Morris, L. Goullieux, I. Blanc, J. Vaubourgeix, S. Lagrange, L. Fraisse, S. Sans, C. Couturier, E. Bacque, K. Rhee, S. M. Scarry, J. Aube, G. Yang, O. Ouerfelli, D. Schnappinger, T. R. Ioerger, C. A. Engelhart, J. A. McConnell, K. McAulay, A. Fay, C. Roubert, J. Sacchettini and C. Nathan, *Science*, 2019, **363**, eaau8959.



- 32 S. M. Gygli, S. Borrell, A. Trauner and S. Gagneux, *FEMS Microbiol. Rev.*, 2017, **41**, 354–373.
- 33 T. R. Ioerger, T. O'Malley, R. Liao, K. M. Guinn, M. J. Hickey, N. Mohaideen, K. C. Murphy, H. I. Boshoff, V. Mizrahi, E. J. Rubin, C. M. Sasseti, C. E. Barry 3rd, D. R. Sherman, T. Parish and J. C. Sacchettini, *PLoS One*, 2013, **8**, e75245.
- 34 W. R. Jacobs Jr, *Microbiol. Spectrum*, 2014, **2**, MGM2–0037–2013.
- 35 J. C. Evans and V. Mizrahi, *Front. Microbiol.*, 2015, **6**, 812.
- 36 J. Bernstein, W. A. Lott, B. A. Steinberg and H. L. Yale, *Am. Rev. Tuberc.*, 1952, **65**, 357–364.
- 37 U. H. Manjunatha, H. Boshoff, C. S. Dowd, L. Zhang, T. J. Albert, J. E. Norton, L. Daniels, T. Dick, S. S. Pang and C. E. Barry 3rd, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 431–436.
- 38 A. Scorpio and Y. Zhang, *Nat. Med.*, 1996, **2**, 662–667.
- 39 P. Gopal, G. Gruber, V. Dartois and T. Dick, *Trends Pharmacol. Sci.*, 2019, **40**, 930–940.
- 40 R. M. Anthony, A. L. den Hertog and D. van Soolingen, *J. Antimicrob. Chemother.*, 2018, **73**, 1750–1754.
- 41 S. Benzer and E. Freese, *Proc. Natl. Acad. Sci. U. S. A.*, 1958, **44**, 112–119.
- 42 G. G. Makafe, M. Hussain, G. Surineni, Y. Tan, N. K. Wong, M. Julius, L. Liu, C. Gift, H. Jiang, Y. Tang, J. Liu, S. Tan, Z. Yu, Z. Liu, Z. Lu, C. Fang, Y. Zhou, J. Zhang, Q. Zhu, J. Liu and T. Zhang, *Cell Chem. Biol.*, 2019, **26**, 1187–1194.e1185.
- 43 L. Lopez Quezada, S. Silve, M. Kelinske, A. Liba, C. Diaz Gonzalez, M. Kotev, L. Goullieux, S. Sans, C. Roubert, S. Lagrange, E. Bacque, C. Couturier, A. Pellet, I. Blanc, M. Ferron, F. Debu, K. Li, J. Aube, J. Roberts, D. Little, Y. Ling, J. Zhang, B. Gold and C. Nathan, *mBio*, 2019, **10**, e01405.
- 44 Z. Yang, X. Zeng and S. K. Tsui, *BMC Genomics*, 2019, **20**, 394.
- 45 A. N. Brooks, S. Turkarslan, K. D. Beer, F. Y. Lo and N. S. Baliga, *Wiley Interdiscip. Rev.: Syst. Biol. Med.*, 2011, **3**, 544–561.
- 46 J. C. Perez and E. A. Groisman, *Cell*, 2009, **138**, 233–244.
- 47 H. I. Boshoff, T. G. Myers, B. R. Copp, M. R. McNeil, M. A. Wilson and C. E. Barry 3rd, *J. Biol. Chem.*, 2004, **279**, 40174–40184.
- 48 P. Murima, P. F. de Sessions, V. Lim, A. N. Naim, P. Bifani, H. I. Boshoff, V. K. Sambandamurthy, T. Dick, M. L. Hibberd, M. Schreiber and S. P. Rao, *PLoS One*, 2013, **8**, e69191.
- 49 S. Ma, S. Jaipalli, J. Larkins-Ford, J. Lohmiller, B. B. Aldridge, D. R. Sherman and S. Chandrasekaran, *mBio*, 2019, **10**, e02627.
- 50 M. I. Voskuil, D. Schnappinger, K. C. Visconti, M. I. Harrell, G. M. Dolganov, D. R. Sherman and G. K. Schoolnik, *J. Exp. Med.*, 2003, **198**, 705–713.
- 51 I. Keren, S. Minami, E. Rubin and K. Lewis, *mBio*, 2011, **2**, e00100.
- 52 W. Lin, P. F. de Sessions, G. H. Teoh, A. N. Mohamed, Y. O. Zhu, V. H. Koh, M. L. Ang, P. C. Dedon, M. L. Hibberd and S. Alonso, *Infect. Immun.*, 2016, **84**, 2505–2523.
- 53 R. Singh, U. Manjunatha, H. I. Boshoff, Y. H. Ha, P. Niyomrattanakit, R. Ledwidge, C. S. Dowd, I. Y. Lee, P. Kim, L. Zhang, S. Kang, T. H. Keller, J. Jiricek and C. E. Barry 3rd, *Science*, 2008, **322**, 1392–1395.
- 54 U. Manjunatha, H. I. Boshoff and C. E. Barry, *Commun. Integr. Biol.*, 2009, **2**, 215–218.
- 55 E. J. R. Peterson, S. Ma, D. R. Sherman and N. S. Baliga, *Nat. Microbiol.*, 2016, **1**, 16078.
- 56 A. Koul, L. Vranckx, N. Dhar, H. W. Gohlmann, E. Ozdemir, J. M. Neefs, M. Schulz, P. Lu, E. Mortz, J. D. McKinney, K. Andries and D. Bald, *Nat. Commun.*, 2014, **5**, 3369.
- 57 C. Vilcheze, T. Hartman, B. Weinrick and W. R. Jacobs Jr, *Nat. Commun.*, 2013, **4**, 1881.
- 58 C. Vilcheze, T. Hartman, B. Weinrick, P. Jain, T. R. Weisbrod, L. W. Leung, J. S. Freundlich and W. R. Jacobs Jr, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 4495–4500.
- 59 K. Flentie, G. A. Harrison, H. Tukenmez, J. Livny, J. A. D. Good, S. Sarkar, D. X. Zhu, R. L. Kinsella, L. A. Weiss, S. D. Solomon, M. E. Schene, M. R. Hansen, A. G. Cairns, M. Kulen, T. Wixe, A. E. G. Lindgren, E. Chorell, C. Bengtsson, K. S. Krishnan, S. J. Hultgren, C. Larsson, F. Almqvist and C. L. Stallings, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 10510–10517.
- 60 H. Zheng, C. J. Colvin, B. K. Johnson, P. D. Kirchhoff, M. Wilson, K. Jorgensen-Muga, S. D. Larsen and R. B. Abramovitch, *Nat. Chem. Biol.*, 2017, **13**, 218–225.
- 61 A. M. Cadena, S. M. Fortune and J. L. Flynn, *Nat. Rev. Immunol.*, 2017, **17**, 691–702.
- 62 K. H. Rohde, D. F. Veiga, S. Caldwell, G. Balazsi and D. G. Russell, *PLoS Pathog.*, 2012, **8**, e1002769.
- 63 D. Schnappinger, S. Ehrh, M. I. Voskuil, Y. Liu, J. A. Mangan, I. M. Monahan, G. Dolganov, B. Efron, P. D. Butcher, C. Nathan and G. K. Schoolnik, *J. Exp. Med.*, 2003, **198**, 693–704.
- 64 K. N. Adams, K. Takaki, L. E. Connolly, H. Wiedenhoft, K. Winglee, O. Humbert, P. H. Edelstein, C. L. Cosma and L. Ramakrishnan, *Cell*, 2011, **145**, 39–53.
- 65 Y. Liu, S. Tan, L. Huang, R. B. Abramovitch, K. H. Rohde, M. D. Zimmerman, C. Chen, V. Dartois, B. C. VanderVen and D. G. Russell, *J. Exp. Med.*, 2016, **213**, 809–825.
- 66 J. P. Sarathy, L. E. Via, D. Weiner, L. Blanc, H. Boshoff, E. A. Eugenin, C. E. Barry 3rd and V. A. Dartois, *Antimicrob. Agents Chemother.*, 2018, **62**, e02266.
- 67 D. Schnappinger, *Cold Spring Harbor Perspect. Med.*, 2015, **5**, a021139.
- 68 R. B. Abramovitch, *IUBMB Life*, 2018, **70**, 818–825.
- 69 N. J. MacGilvary and S. Tan, *Pathog. Dis.*, 2018, **76**, fty017.
- 70 D. Alland, A. J. Steyn, T. Weisbrod, K. Aldrich and W. R. Jacobs Jr, *J. Bacteriol.*, 2000, **182**, 1802–1811.
- 71 K. L. Smollett, K. M. Smith, C. Kahramanoglou, K. B. Arnvig, R. S. Buxton and E. O. Davis, *J. Biol. Chem.*, 2012, **287**, 22004–22014.
- 72 K. Naran, A. Moosa, C. E. Barry 3rd, H. I. Boshoff, V. Mizrahi and D. F. Warner, *Antimicrob. Agents Chemother.*, 2016, **60**, 6748–6757.
- 73 H. P. Fischer, N. A. Brunner, B. Wieland, J. Paquette, L. Macko, K. Ziegelbauer and C. Freiberg, *Genome Res.*, 2004, **14**, 90–98.





- 74 T. Roemer, J. Davies, G. Giaever and C. Nislow, *Nat. Chem. Biol.*, 2011, **8**, 46–56.
- 75 G. Giaever, P. Flaherty, J. Kumm, M. Proctor, C. Nislow, D. F. Jaramillo, A. M. Chu, M. I. Jordan, A. P. Arkin and R. W. Davis, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 793–798.
- 76 R. Pathania, S. Zlitni, C. Barker, R. Das, D. A. Gerritsma, J. Lebert, E. Awuah, G. Melacini, F. A. Capretta and E. D. Brown, *Nat. Chem. Biol.*, 2009, **5**, 849–856.
- 77 M. A. Farha and E. D. Brown, *Chem. Biol.*, 2010, **17**, 852–862.
- 78 X. Yang, T. Yuan, R. Ma, K. I. Chacko, M. Smith, G. Deikus, R. Sebra, A. Kasarskis, H. van Bakel, S. G. Franzblau and N. S. Sampson, *ACS Infect. Dis.*, 2019, **5**, 1239–1251.
- 79 B. C. VanderVen, R. J. Fahey, W. Lee, Y. Liu, R. B. Abramovitch, C. Memmott, A. M. Crowe, L. D. Eltis, E. Perola, D. D. Deininger, T. Wang, C. P. Locher and D. G. Russell, *PLoS Pathog.*, 2015, **11**, e1004679.
- 80 D. C. Alexander, J. R. Jones and J. Liu, *Antimicrob. Agents Chemother.*, 2003, **47**, 3208–3213.
- 81 R. L. Campen, D. F. Ackerley, G. M. Cook and R. F. O’Toole, *Tuberculosis*, 2015, **95**, 432–439.
- 82 W. Xu, M. A. DeJesus, N. Rucker, C. A. Engelhart, M. G. Wright, C. Healy, K. Lin, R. Wang, S. W. Park, T. R. Ioerger, D. Schnappinger and S. Ehrhart, *Antimicrob. Agents Chemother.*, 2017, **61**, e01334.
- 83 E. O. Johnson, E. LaVerriere, E. Office, M. Stanley, E. Meyer, T. Kawate, J. E. Gomez, R. E. Audette, N. Bandyopadhyay, N. Betancourt, K. Delano, I. Da Silva, J. Davis, C. Gallo, M. Gardner, A. J. Golas, K. M. Guinn, S. Kennedy, R. Korn, J. A. McConnell, C. E. Moss, K. C. Murphy, R. M. Nietupski, K. G. Papavinasasundaram, J. T. Pinkham, P. A. Pino, M. K. Proulx, N. Ruecker, N. Song, M. Thompson, C. Trujillo, S. Wakabayashi, J. B. Wallach, C. Watson, T. R. Ioerger, E. S. Lander, B. K. Hubbard, M. H. Serrano-Wu, S. Ehrhart, M. Fitzgerald, E. J. Rubin, C. M. Sasseti, D. Schnappinger and D. T. Hung, *Nature*, 2019, **571**, 72–78.
- 84 E. O. Johnson, E. Office, T. Kawate, M. Orzechowski and D. T. Hung, *ACS Infect. Dis.*, 2020, **6**, 56–63.
- 85 J. M. Rock, F. F. Hopkins, A. Chavez, M. Diallo, M. R. Chase, E. R. Gerrick, J. R. Pritchard, G. M. Church, E. J. Rubin, C. M. Sasseti, D. Schnappinger and S. M. Fortune, *Nat. Microbiol.*, 2017, **2**, 16274.
- 86 I. du Preez, L. Luies and D. T. Loots, *Tuberculosis*, 2019, **115**, 126–139.
- 87 D. G. Robertson and U. Frevert, *Clin. Pharmacol. Ther.*, 2013, **94**, 559–561.
- 88 M. Drapal and P. D. Fraser, *Microorganisms*, 2019, **7**, 148.
- 89 R. S. Jansen and K. Y. Rhee, *Trends Pharmacol. Sci.*, 2017, **38**, 393–405.
- 90 Z. S. Bhat, M. A. Rather, M. Maqbool and Z. Ahmad, *Biomed. Pharmacother.*, 2018, **103**, 1733–1747.
- 91 A. E. Mohamed, F. H. Ahmed, S. Arulmozhiraja, C. Y. Lin, M. C. Taylor, E. R. Krausz, C. J. Jackson and M. L. Coote, *Mol. Biosyst.*, 2016, **12**, 1110–1113.
- 92 R. Baptista, D. M. Fazakerley, M. Beckmann, L. Baillie and L. A. J. Mur, *Sci. Rep.*, 2018, **8**, 5084.
- 93 P. Belenky, J. D. Ye, C. B. Porter, N. R. Cohen, M. A. Lobritz, T. Ferrante, S. Jain, B. J. Korry, E. G. Schwarz, G. C. Walker and J. J. Collins, *Cell Rep.*, 2015, **13**, 968–980.
- 94 Z. Wang, V. Soni, G. Marriner, T. Kaneko, H. I. M. Boshoff, C. E. Barry 3rd and K. Y. Rhee, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 19646–19651.
- 95 M. Zampieri, M. Zimmermann, M. Claassen and U. Sauer, *Cell Rep.*, 2017, **19**, 1214–1228.
- 96 I. M. Vincent, D. E. Ehmann, S. D. Mills, M. Perros and M. P. Barrett, *Antimicrob. Agents Chemother.*, 2016, **60**, 2281–2291.
- 97 M. Rohman and J. Wingfield, *Methods Mol. Biol.*, 2016, **1439**, 47–63.
- 98 K. Dettmer, P. A. Aronov and B. D. Hammock, *Mass Spectrom. Rev.*, 2007, **26**, 51–78.
- 99 A. I. Campos and M. Zampieri, *Mol. Cell*, 2019, **74**, 1291–1303.e1296.
- 100 B. Lomenick, R. W. Olsen and J. Huang, *ACS Chem. Biol.*, 2011, **6**, 34–46.
- 101 W. Zheng, G. Li and X. Li, *Arch. Pharmacol. Res.*, 2015, **38**, 1661–1685.
- 102 F. Cong, A. K. Cheung and S. M. Huang, *Annu. Rev. Pharmacol. Toxicol.*, 2012, **52**, 57–78.
- 103 A. McFedries, A. Schwaid and A. Saghatelian, *Chem. Biol.*, 2013, **20**, 667–673.
- 104 R. Aebersold and D. R. Goodlett, *Chem. Rev.*, 2001, **101**, 269–295.
- 105 E. K. Schmitt, M. Riwanto, V. Sambandamurthy, S. Roggo, C. Miault, C. Zwingelstein, P. Krastel, C. Noble, D. Beer, S. P. Rao, M. Au, P. Niyomrattanakit, V. Lim, J. Zheng, D. Jeffery, K. Pethe and L. R. Camacho, *Angew. Chem., Int. Ed.*, 2011, **50**, 5889–5891.
- 106 N. Burstner, S. Roggo, N. Ostermann, J. Blank, C. Delmas, F. Freuler, B. Gerhartz, A. Hinniger, D. Hoepfner, B. Liechty, M. Mihalic, J. Murphy, D. Pistorius, M. Rottmann, J. R. Thomas, M. Schirle and E. K. Schmitt, *ChemBioChem*, 2015, **16**, 2433–2436.
- 107 H. C. Kolb and K. B. Sharpless, *Drug Discovery Today*, 2003, **8**, 1128–1137.
- 108 L. Liang and D. Astruc, *Coord. Chem. Rev.*, 2011, **255**, 2933–2945.
- 109 S. I. Presolski, V. P. Hong and M. G. Finn, *Curr. Protoc. Chem. Biol.*, 2011, **3**, 153–162.
- 110 J. C. Jewett and C. R. Bertozzi, *Chem. Soc. Rev.*, 2010, **39**, 1272–1279.
- 111 L. I. Willems, W. A. van der Linden, N. Li, K. Y. Li, N. Liu, S. Hoogendoorn, G. A. van der Marel, B. I. Florea and H. S. Overkleeft, *Acc. Chem. Res.*, 2011, **44**, 718–729.
- 112 N. Zhao, C. M. Darby, J. Small, D. A. Bachovchin, X. Jiang, K. E. Burns-Huang, H. Botella, S. Ehrhart, D. L. Boger, E. D. Anderson, B. F. Cravatt, A. E. Speers, V. Fernandez-Vega, P. S. Hodder, C. Eberhart, H. Rosen, T. P. Spicer and C. F. Nathan, *ACS Chem. Biol.*, 2015, **10**, 364–371.
- 113 C. Trefzer, H. Skovierova, S. Buroni, A. Bobovska, S. Nenci, E. Molteni, F. Pojer, M. R. Pasca, V. Makarov, S. T. Cole, G. Riccardi, K. Mikusova and K. Johnsson, *J. Am. Chem. Soc.*, 2012, **134**, 912–915.



- 114 G. Lin, D. Li, L. P. de Carvalho, H. Deng, H. Tao, G. Vogt, K. Wu, J. Schneider, T. Chidawanyika, J. D. Warren, H. Li and C. Nathan, *Nature*, 2009, **461**, 621–626.
- 115 L. Dubinsky, B. P. Krom and M. M. Meijler, *Bioorg. Med. Chem.*, 2012, **20**, 554–570.
- 116 P. Le, E. Kunold, R. Macsics, K. Rox, M. C. Jennings, I. Ugur, M. Reinecke, D. Chaves-Moreno, M. W. Hackl, C. Fetzer, F. A. M. Mandl, J. Lehmann, V. S. Korotkov, S. M. Hacker, B. Kuster, I. Antes, D. H. Pieper, M. Rohde, W. M. Wuest, E. Medina and S. A. Sieber, *Nat. Chem.*, 2019, **12**, 145–158.
- 117 L. Ballell, R. H. Bates, R. J. Young, D. Alvarez-Gomez, E. Alvarez-Ruiz, V. Barroso, D. Blanco, B. Crespo, J. Escribano, R. Gonzalez, S. Lozano, S. Huss, A. Santos-Villarejo, J. J. Martin-Plaza, A. Mendoza, M. J. Rebollo-Lopez, M. Remuinan-Blanco, J. L. Lavandera, E. Perez-Herran, F. J. Gamo-Benito, J. F. Garcia-Bustos, D. Barros, J. P. Castro and N. Cammack, *ChemMedChem*, 2013, **8**, 313–321.
- 118 M. J. Remuinan, E. Perez-Herran, J. Rullas, C. Alemparte, M. Martinez-Hoyos, D. J. Dow, J. Afari, N. Mehta, J. Esquivias, E. Jimenez, F. Ortega-Muro, M. T. Fraile-Gabaldon, V. L. Spivey, N. J. Loman, M. J. Pallen, C. Constantinidou, D. J. Minick, M. Cacho, M. J. Rebollo-Lopez, C. Gonzalez, V. Sousa, I. Angulo-Barturen, A. Mendoza-Losana, D. Barros, G. S. Besra, L. Ballell and N. Cammack, *PLoS One*, 2013, **8**, e60933.
- 119 J. A. Cox, K. A. Abrahams, C. Alemparte, S. Ghidelli-Disse, J. Rullas, I. Angulo-Barturen, A. Singh, S. S. Gurcha, V. Nataraj, S. Bethell, M. J. Remuinan, L. Encinas, P. J. Jarvis, N. C. Cammack, A. Bhatt, U. Kruse, M. Bantscheff, K. Futterer, D. Barros, L. Ballell, G. Drewes and G. S. Besra, *Nat. Microbiol.*, 2016, **1**, 15006.
- 120 M. W. Halloran and J. P. Lumb, *Chemistry*, 2019, **25**, 4885–4898.
- 121 M. Fonovic and M. Bogyo, *Expert Rev. Proteomics*, 2008, **5**, 721–730.
- 122 M. H. Wright and S. A. Sieber, *Nat. Prod. Rep.*, 2016, **33**, 681–708.
- 123 H. Chen, O. Engkvist, Y. Wang, M. Olivecrona and T. Blaschke, *Drug Discovery Today*, 2018, **23**, 1241–1250.
- 124 R. Chen, X. Liu, S. Jin, J. Lin and J. Liu, *Molecules*, 2018, **23**, 2208.
- 125 A. Jung, *arXiv e-prints*, 2018, arXiv:1805.05052.
- 126 T. Ching, D. S. Himmelstein, B. K. Beaulieu-Jones, A. A. Kalinin, B. T. Do, G. P. Way, E. Ferrero, P. M. Agapow, M. Zietz, M. M. Hoffman, W. Xie, G. L. Rosen, B. J. Lengerich, J. Israeli, J. Lanchantin, S. Woloszynek, A. E. Carpenter, A. Shrikumar, J. Xu, E. M. Cofer, C. A. Lavender, S. C. Turaga, A. M. Alexandari, Z. Lu, D. J. Harris, D. DeCaprio, Y. Qi, A. Kundaje, Y. Peng, L. K. Wiley, M. H. S. Segler, S. M. Boca, S. J. Swamidass, A. Huang, A. Gitter and C. S. Greene, *J. R. Soc., Interface*, 2018, **15**, 20170387.
- 127 J. H. Yang, S. N. Wright, M. Hamblin, D. McCloskey, M. A. Alcantar, L. Schrubbers, A. J. Lopatkin, S. Satish, A. Nili, B. O. Palsson, G. C. Walker and J. J. Collins, *Cell*, 2019, **177**, 1649–1661.e1649.
- 128 T. S. Gardner, D. di Bernardo, D. Lorenz and J. J. Collins, *Science*, 2003, **301**, 102–105.
- 129 J. H. Yang, S. C. Bening and J. J. Collins, *Curr. Opin. Microbiol.*, 2017, **39**, 73–80.
- 130 F. Martinez-Jimenez, G. Papadatos, L. Yang, I. M. Wallace, V. Kumar, U. Pieper, A. Sali, J. R. Brown, J. P. Overington and M. A. Marti-Renom, *PLoS Comput. Biol.*, 2013, **9**, e1003253.
- 131 M. J. Rebollo-Lopez, J. Lelievre, D. Alvarez-Gomez, J. Castro-Pichel, F. Martinez-Jimenez, G. Papadatos, V. Kumar, G. Colmenarejo, G. Mugumbate, M. Hurle, V. Barroso, R. J. Young, M. Martinez-Hoyos, R. Gonzalez del Rio, R. H. Bates, E. M. Lopez-Roman, A. Mendoza-Losana, J. R. Brown, E. Alvarez-Ruiz, M. A. Marti-Renom, J. P. Overington, N. Cammack, L. Ballell and D. Barros-Aguire, *PLoS One*, 2015, **10**, e0142293.
- 132 I. Barberis, N. L. Bragazzi, L. Galluzzo and M. Martini, *J. Prev. Med. Hyg.*, 2017, **58**, E9–E12.
- 133 C. J. Martin, A. F. Carey and S. M. Fortune, *Semin. Immunopathol.*, 2016, **38**, 213–220.
- 134 V. Dartois, *Nat. Rev. Microbiol.*, 2014, **12**, 159–167.
- 135 B. Gold and C. Nathan, *Microbiol. Spectrum*, 2017, **5**, TBTB2–0031-2016.
- 136 Q. Wang, H. I. M. Boshoff, J. R. Harrison, P. C. Ray, S. R. Green, P. G. Wyatt and C. E. Barry 3rd, *Science*, 2020, **367**, 1147–1151.
- 137 L. Pasquina, J. P. Santa Maria Jr, B. McKay Wood, S. H. Moussa, L. M. Matano, M. Santiago, S. E. Martin, W. Lee, T. C. Meredith and S. Walker, *Nat. Chem. Biol.*, 2016, **12**, 40–45.
- 138 J. M. Stokes, K. Yang, K. Swanson, W. Jin, A. Cubillos-Ruiz, N. M. Donghia, C. R. MacNair, S. French, L. A. Carfrae, Z. Bloom-Ackerman, V. M. Tran, A. Chiappino-Pepe, A. H. Badran, I. W. Andrews, E. J. Chory, G. M. Church, E. D. Brown, T. S. Jaakkola, R. Barzilay and J. J. Collins, *Cell*, 2020, **180**, 688–702.e613.
- 139 Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda and M. Kanehisa, *Bioinformatics*, 2008, **24**, i232–240.
- 140 Z. He, J. Zhang, X. H. Shi, L. L. Hu, X. Kong, Y. D. Cai and K. C. Chou, *PLoS One*, 2010, **5**, e9603.

