

# Predicting the *in Vivo* Mechanism of Action for Drug Leads Using NMR Metabolomics

Steven Halouska,<sup>†</sup> Robert J. Fenton,<sup>‡</sup> Raúl G. Barletta,<sup>‡</sup> and Robert Powers<sup>†,\*</sup>

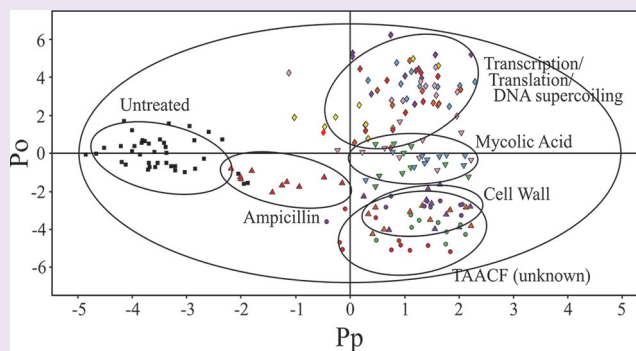
<sup>†</sup>Department of Chemistry, University of Nebraska-Lincoln, Lincoln, Nebraska 68588-0304, United States

<sup>‡</sup>School of Veterinary Medicine and Biomedical Sciences, University of Nebraska-Lincoln, Lincoln, Nebraska 68583-0905, United States

## Supporting Information

**ABSTRACT:** New strategies are needed to circumvent increasing outbreaks of resistant strains of pathogens and to expand the dwindling supply of effective antimicrobials. A common impediment to drug development is the lack of an easy approach to determine the *in vivo* mechanism of action and efficacy of novel drug leads. Toward this end, we describe an unbiased approach to predict *in vivo* mechanisms of action from NMR metabolomics data. *Mycobacterium smegmatis*, a non-pathogenic model organism for *Mycobacterium tuberculosis*, was treated with 12 known drugs and 3 chemical leads identified from a cell-based assay. NMR analysis of drug-induced changes to the *M. smegmatis* metabolome resulted in distinct clustering patterns correlating with *in vivo* drug activity.

The clustering of novel chemical leads relative to known drugs provides a mean to identify a protein target or predict *in vivo* activity.



Emerging and reemerging infectious disease outbreaks from numerous Gram-negative and Gram-positive pathogens have increased dramatically over the past decade.<sup>1</sup> Further, we are facing the serious likelihood that these pathogens will soon become resistant to all known antibacterial treatments, which may lead to worldwide pandemics.<sup>2</sup> Unfortunately, the development and approval of antibiotics have not kept pace with the growing emergence of resistant pathogens.<sup>3</sup> Instead, there has been a decline in the approval of new antibiotics.<sup>4</sup> Twenty novel classes of marketable antibiotics were produced between 1930 and 1962.<sup>5</sup> These classes of antibiotics inhibit a short list of cellular processes: cell wall biosynthesis, DNA supercoiling, transcription, translation, and folate biosynthesis. Since 1962, only two new antibiotic classes have received FDA approval: oxazolidinones, which inhibits protein synthesis, and cyclic lipopeptides, which destroys membrane potential. Both compounds are used in the treatment of Gram-positive bacteria, such as methicillin-resistant *Staphylococcus aureus* (MRSA).<sup>5</sup> However, additional antibiotics are needed to combat the prevalence of other multidrug-resistant pathogens, such as *Enterococcus faecium*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species that are infecting the majority of U.S. hospitals.<sup>6</sup> Also extreme drug-resistant strains of *Mycobacterium tuberculosis* are a rising threat in the world.

The Infectious Diseases Society (IDSA) has proposed an initiative to develop and approve 10 novel antibiotics by the year 2020.<sup>7</sup> However, existing drug discovery strategies may not

be able to meet these challenges. Drug discovery programs rely heavily on target-based high-throughput screening (HTS) of large chemical libraries followed by lead optimization.<sup>8,9</sup> Unfortunately, this approach has demonstrated an extremely high rate of failure and erroneous leads. Even when a valid HTS hit is found, it is uncertain if this chemical lead can penetrate into the bacterial cell and demonstrate *in vivo* activity.

NMR Metabolomics is evolving as a significant component of the drug discovery process and offers an inexpensive route to help overcome the multiple challenges faced by researchers.<sup>10</sup> Metabolomics is a relatively new field and is based on the identification and quantification of small molecules found in living cells or biofluids.<sup>11</sup> Since small molecules are downstream products of biomolecular processes, the identity and concentration of metabolites provide biochemical signatures for tracking the physiological effects of antibiotic efficacy, selectivity, and toxicity. Characterizing these biochemical signatures relies upon the global determination of numerous endogenous small molecules followed by pattern recognition using multivariate analysis.<sup>12</sup> Such comprehensive biochemical information can be readily obtained using <sup>1</sup>H NMR spectroscopy with minimal sample handling while providing highly reproducible data in an automated fashion.<sup>10</sup> Multivariate statistical analysis, such as orthogonal partial least-squares

**Received:** September 7, 2011

**Accepted:** October 18, 2011

**Published:** October 18, 2011

Table 1. Description of Antimicrobial Compounds and Dosages Used in This Study

compound	class	mechanism of action	MIC <sup>a</sup> ( $\mu\text{g/mL}$ )	dosage <sup>b</sup> ( $\mu\text{g/mL}$ )
ampicillin	penicillins	inhibits transpeptidation and prevents cell wall formation	16.0 <sup>c</sup>	96.0
chloramphenicol	amphetamines	inhibits protein synthesis by binding to the 50S ribosomal subunit	6.0	6.0
ciprofloxacin	fluoroquinolones	inhibits DNA gyrase and prevents DNA supercoiling	0.2	2.0
D-cycloserine	oxazolidinones	inhibits alanine racemase and alanine ligase and prevents cell wall formation (different from other oxazolidinones that inhibit protein synthesis)	750	750
ethambutol	amino alcohols	disrupts arabinogalactan formation by inhibiting arabinosyl transferase	10.0	100.0
ethionamide	pyridine derivatives	inhibits mycolic acid formation similar to isoniazid	20.0	160.0
gentamicin	aminoglycosides	inhibits protein synthesis by binding to the 30S ribosomal protein S12 and 16S rRNA	2.0	8.0
isoniazid	pyridine derivatives	a prodrug that inhibits InhA and prevent mycolic acid synthesis	2.0	48.0
kanamycin	aminoglycosides	inhibits protein synthesis by binding to the 30S ribosomal protein S12 and 16S rRNA	4.0	4.0
rifampicin	rifampicins	inhibits RNA polymerase and prevent RNA synthesis	30.0	60.0
streptomycin	aminoglycosides	inhibits protein synthesis by binding to the 30S ribosomal protein S12 and 16S rRNA	0.25	1.5
vancomycin	glycopeptides	binds to the D-alanyl-D-alanine dipeptide and prevents cell wall formation	50	450
amiodarone	benzofurans	unknown	26.6	212.8
clofazimine	anilines	unknown	0.32	3.84
chlorprothixene	thioxanthines	unknown	36.0	216.0

<sup>a</sup>Literature values of minimum inhibitory concentrations against *M. smegmatis* used as a starting point to determine an optimal dosage for the NMR metabolomics study. <sup>b</sup>Actual dosage used to treat *M. smegmatis* cells to inhibit growth by ~50% following drug treatment. <sup>c</sup>Reported for *M. smegmatis*  $\beta$ -lactamase and ribosomal protein S12 mutants.

discriminant analysis (OPLS-DA), is typically employed to extract information from the large and complex NMR data sets.<sup>13</sup> Simply, OPLS-DA is used to identify clustering patterns from the major variations between NMR spectra.<sup>10</sup>

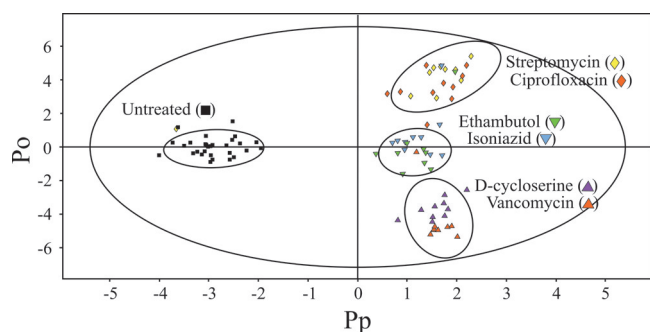
Herein, we describe a new method using <sup>1</sup>H NMR and OPLS-DA to profile the *in vivo* mechanism of action of known antibiotics used to treat *M. tuberculosis*. More importantly, we aim to use this information to classify compounds with unknown mechanisms of action, but demonstrated antitubercular activity. Our approach is predicated on the hypothesis that drugs with similar modes of activity or therapeutic targets will have a similar impact on the metabolome of *M. smegmatis* and will cluster together in an OPLS-DA scores plot. Thus, the mode of action of a novel chemical lead can be inferred from its clustering in an OPLS-DA scores plot relative to drugs with defined biological targets. Importantly, if the chemical lead is separated from known drugs in the OPLS-DA scores plot, then this result would infer a new mechanism of action and a potentially valuable, new antibiotic.

Our methodology was demonstrated using 12 antibiotics known to inhibit the growth of *M. tuberculosis* and *M. smegmatis* (Table 1). The mechanism of action for each antibiotic was identified from the Drug Bank Database,<sup>14</sup> and the minimum inhibitory concentrations (MICs) were obtained from the scientific literature.<sup>15–23</sup> In addition, three chemical leads were randomly selected from the Tuberculosis Antimicrobial Acquisition and Coordinating Facility (TAACF) library of compounds (<http://www.TAACF.org>). The compounds were screened against *M. tuberculosis* and have MICs comparable to those of known TB drugs, but the biological target or mechanism of action was not reported by TAACF. The non-pathogenic *M. smegmatis* was used as a model system for the NMR metabolomics study.

In order to analyze changes in the *M. smegmatis* metabolome, the drug dosage needs to be below lethal levels and affect only cell growth. Typically, a drug concentration that inhibits cell growth by approximately 50% of the growth rate of untreated

cells is desirable. While MIC values are available from the literature, these concentrations are based on standardized drug gradients, inoculum sizes, and readout end points. Additionally, the reported MICs were obtained with different bacterial strains, at different growth stages or cell densities, and under a variety of experimental conditions that includes either broth or agar methods. Further complicating the situation is the diversity of MICs values reported for a single drug. Thus, the literature MIC values listed in Table 1 were simply used as a starting point to determine an optimal dosage for the NMR metabolomics study under our experimental conditions. Each drug was titrated over a concentration range of 1 to 24 times the literature MIC values. The individual drug concentrations needed to achieve ~50% growth inhibition are reported in Table 1. An average growth inhibition of  $43.1 \pm 10.5\%$  was observed after the addition of each of the 15 drugs. Preparation of the *M. smegmatis* cell cultures for metabolomic analysis was then performed using the optimal dosage for each drug.

Due to the inherent variability of biological samples and to provide a robust statistical analysis, 10 cultures inoculated with each antibiotic and 40 cultures of untreated cells were prepared for the NMR metabolomics study. A 1D <sup>1</sup>H NMR spectrum was collected for each biological sample, which were normalized using center averaging and analyzed using OPLS-DA. A representative 2D OPLS-DA scores plot displaying a comparison between 6 antibiotics with known mechanisms of action is shown in Figure 1. The OPLS-DA model was cross-validated using a modified leave-one-out method. A quality assessment score ( $Q^2$ ) of 0.82 was obtained, which is an excellent result compared to an ideal score of 1. Thus, the cross-validation indicates a highly reliable model. Each point in the 2D OPLS-DA scores plot represents a single 1D <sup>1</sup>H NMR spectrum of a specific drug-treated or untreated cell culture. The 2D OPLS-DA scores plot consists of 4 separate clustering patterns, which demonstrates that each group has a considerably different impact on the metabolome of *M. smegmatis*. Importantly, all of the drug-treated *M. smegmatis*

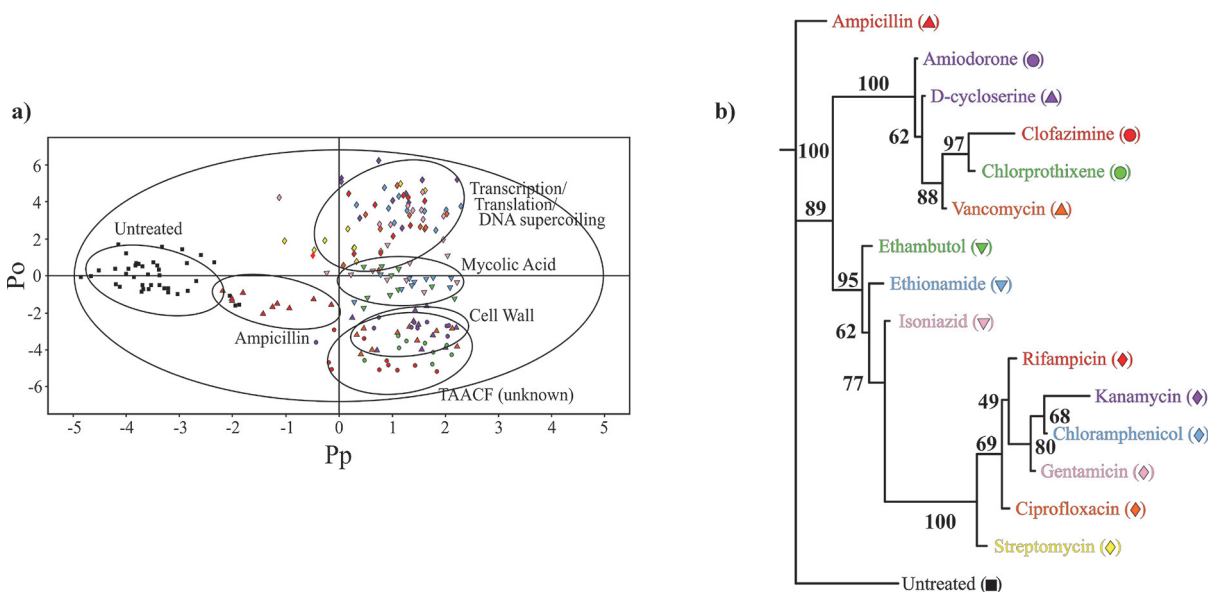


**Figure 1.** 2D OPLS-DA scores plot demonstrating the clustering pattern obtained for six different antibiotics with known and distinct biological targets: untreated *M. smegmatis* cells, ciprofloxacin, streptomycin, ethambutol, isoniazid, D-cycloserine, and vancomycin treated *M. smegmatis* cells. Symbols and labels are indicated on the plot. The ellipses correspond to the 95% confidence limits from a normal distribution for each cluster. The untreated *M. smegmatis* cells was designated the control class, and the remainder of the cells were designated as treated. The OPLS-DA used one predictive component and three orthogonal components to yield a  $R^2X$  of 0.610,  $R^2Y$  of 0.893, and  $Q^2$  of 0.82.

cell cultures form distinct and separate clusters from the untreated cell cultures. This is consistent with all of the drugs being biologically active and inhibiting *M. smegmatis* cell growth. Antibiotics that share a similar or identical biological target were observed to cluster together in the OPLS-DA scores plot. For example, ethambutol and isoniazid inhibit mycolic acid biosynthesis that prevents the formation of the arabinogalactan-mycolic acid matrix. Streptomycin and ciprofloxacin form the second cluster. Streptomycin prevents protein

synthesis, and ciprofloxacin inhibits DNA supercoiling that affects replication, transcription, and repair, leading to a similar disruption in protein synthesis. Since these two antibiotics cluster together, it implies that the inhibition of transcription or translation results in a similar impact on the metabolome. Vancomycin and D-cycloserine both affect cell wall formation and form the third cluster. In a principal component analysis (PCA) of the data (see Supplemental Figure 1S) there is a more pronounced separation between vancomycin and D-cycloserine along PC2. This reflects a fundamental difference between PCA and OPLS-DA, where PCA is limited to a linear model and does not readily differentiate between within-class and between-class variations.<sup>13,24</sup> Correspondingly, OPLS-DA is preferred as long as cross-validation verifies a reliable model.

The NMR metabolomics analysis was then expanded to include a total of 12 drugs with known biological targets and 3 compounds randomly chosen from the TAACF library. Amiodorone, clofazamine, and chlorprothixene are active against TB but have unknown mechanisms of action according to the TAACF database. Nevertheless, the three compounds are known drugs, where amiodorone is an antiarrhythmic agent that affects potassium efflux, chlorprothixene is an antipsychotic drug that inhibits dopamine receptors, and clofazamine is a 40-year-old leprosy treatment with an ill-defined biological activity. The 2D OPLS-DA scores plot (Figure 2A) identified 4 distinct clusters and yielded a highly reliable cross validation  $Q^2$  score of 0.671. As before, the different clusters are correlated with distinct modes of action: inhibition of cell wall formation, inhibition of mycolic acid biosynthesis, and inhibition of transcription, translation or the overall effects of DNA supercoiling.



**Figure 2.** (a) 2D OPLS-DA scores plot demonstrating the clustering pattern for 12 antibiotics with known biological targets and three compounds of unknown *in vivo* activity: untreated *M. smegmatis* cells, chloramphenicol, ciprofloxacin, gentamicin, kanamycin, rifampicin, streptomycin, ethambutol, ethionamide, isoniazid, ampicillin, D-cycloserine, vancomycin, amiodorone, chlorprothixene, and clofazimine treated *M. smegmatis* cells. The symbols correspond with the coloring scheme and labeled symbols indicated on the tree diagram in (b). The ellipses correspond to the 95% confidence limits from a normal distribution for each cluster. The untreated *M. smegmatis* cells (black square) was designated the control class, and the remainder of the cells were designated as treated. The OPLS-DA used one predictive component and six orthogonal components to yield a  $R^2X$  of 0.715,  $R^2Y$  of 0.803, and  $Q^2$  of 0.671. (b) Metabolomics tree diagram determined from the OPLS-DA scores plot. The coloring scheme and associated symbol for each compound in the tree diagram correlates with colored symbols in the OPLS-DA scores plot. The bootstrap numbers for each node are indicated on the tree diagram.

The accompanying metabolomics tree diagram<sup>25</sup> (Figure 2B) clearly visualizes the relative groupings of the three antibiotic classes. The bootstrap numbers of 89–100 indicate a statistically significant separation between the five clusters and the reliability of the general drug and TAACF classifications. The metabolic tree diagram also provides a finer separation between the drugs within each cluster. These separations may reflect actual differences in the specific drug targets. For example, D-cycloserine and vancomycin are on separate branches in the cell wall node potentially because D-cycloserine inhibits alanine racemase and alanine ligase compared to vancomycin binding the D-alanyl-D-alanine dipeptide. Alternatively, the separation may result from differences in the relative activity of the drug. Streptomycin forms a separate branch in the transcription, translation, or DNA supercoiling drug cluster despite having a similar target (binding to the 30S ribosomal protein S12 and 16S rRNA) relative to other members within the cluster. However, streptomycin is one of the most active compounds tested, requiring a dosage of only 1.5  $\mu\text{g}/\text{mL}$  to inhibit *M. smegmatis* growth by approximately 50%. Also, overinterpreting these subtle separations may be erroneous since the within-cluster differences may simply reflect experimental variability and may not be biologically relevant. For instance, an average growth inhibition of  $43.1 \pm 10.5\%$  was observed after the addition of each of the 15 drugs. This dosage variability may lead to unintended separations in the 2D OPLS-DA scores plot. Essentially, the reliability of these finer cluster differences is dependent on additional supportive biological data.

Surprisingly, amiodorone, chlorprothixene, and clofazamine were found to cluster together in the 2D OPLS-DA scores plot and metabolic tree diagram. This was an unexpected result given that the three compounds were randomly selected from the large TAACF library and have diverse therapeutic usages. However, it also implies the three compounds share a similar mechanism of action in TB. Importantly, the three TAACF compounds also cluster with the antibiotics that disrupt cell wall formation, ampicillin, D-cycloserine, and vancomycin. This infers a similar mode of action between the three TAACF compounds and the antibiotics that are known to interfere with bacterial cell walls. A subsequent literature search indicated that the three TAACF compounds have been previously shown to disrupt bacterial membranes in organisms distinct from TB.<sup>26–30</sup> Thus, the literature results are consistent with our NMR metabolomics analysis, which support our general classification of the TAACF compounds as interfering with bacterial cell walls. It is important to note that while ampicillin is a member of this class of antibiotics, it is also skewed toward the untreated cells in the 2D OPLS-DA scores plot. Presumably, this is because of *M. smegmatis*  $\beta$ -lactamase activity that provides resistance to ampicillin.<sup>23,31</sup> The impact of ampicillin on the metabolome of *M. smegmatis* is significantly diminished such that ampicillin *M. smegmatis* is similar to untreated cells. As described previously, there are some differences between the OPLS-DA and PCA scores plot (see Supplemental Figure 2S). There is less discrimination between the untreated and drug-treated cells in the 2D PCA scores plot. This is not unexpected since PLS is preferred over PCA for discrimination between classes.<sup>24</sup> Also, there is a separation between the three TAACF compounds and the cell wall antibiotics in the PCA scores plot, but the TAACF compounds are still closer to the cell wall antibiotics in the associated metabolic tree diagram (see Supplemental Figure 2S). In fact,

the OPLS-DA and PCA metabolomic tree diagrams are quite similar despite these visible differences in the scores plots. Additionally, the quality of the OPLS-DA model is apparent from the fit to the data,  $R^2X > 0.610$  and  $R^2Y > 0.803$ , and the reliability of the model is apparent from the cross-validation  $Q^2$  scores  $> 0.617$ . Further validation of the OPLS-DA drug and TAACF classifications comes from the analysis of the metabolites identified as the major contributors to the OPLS-DA class separation (see Supplemental Figures 3S–6S).

The S-plots and loading plots determined from the OPLS-DA models identify the chemical shifts (and associated metabolites) that contribute to the observed separation between the untreated and treated cells in the 2D OPLS-DA scores plot. The metabolites and corresponding pathways predominately perturbed by the addition of each drug class are listed in Supplemental Tables 1S–3S. While there are some broad similarities in the metabolites and pathways affected by the drugs because the comparisons are all made relative to untreated cells, there are also some distinct differences. For example, proline, cytidine, uridine, and inosine (pyrimidine and purine pathways) are all uniquely decreased by drugs that affect transcription, translation, or DNA supercoiling. Obviously, nucleotides are essential metabolic precursors to DNA and RNA synthesis. Alternatively, choline phosphate, lysine, spermidine, citrulline, ascorbate, and dehydroascorbate (glycerophospholipid, lysine biosynthesis, arginine and proline metabolism, and ascorbate metabolism pathways) are decreased by drugs affecting the mycolic acid pathway. Ascorbate metabolism is directly linked to the mycolic acid pathway, where ascorbate leads to arabinose. Arabinose is the primary precursor for the arabinogalactan-mycolic acid complex. Also, the inhibition of spermidine synthesis has been previously observed for drugs targeting the mycolic acid pathway in mycobacteria.<sup>32</sup> Importantly, the set of metabolites affected by the TAACF compounds were identical to metabolites perturbed by D-cycloserine and vancomycin. Both show a decrease in oxaloacetate, glutamine, glutamate, methionine, and folate and an increase in isoleucine. Clearly, amino acids and their precursors are important components in peptidoglycan, cell wall, and cell membrane synthesis. There were some additional metabolites that are increased by the addition of D-cycloserine and vancomycin that were not observed with the TAACF compounds. These include other amino acids (alanine, lysine, serine, valine) and other precursors to peptidoglycan synthesis (*N*-acetyl-D-glucosamine, *N*-acetylneuraminic acid). Overall, the identity of the specific metabolites perturbed by each drug class is consistent with the 2D OPLS-DA scores plot clustering pattern and drug classifications.

In conclusion, we have demonstrated that different classes of antibiotics uniquely affect the metabolome of *M. smegmatis*. These metabolomic changes are directly correlated with broad mechanisms of action that are associated with each TB class of antibiotics, disruption of cell walls or membranes, inhibition of transcription, translation, or DNA supercoiling, or the inhibition of mycolic acid biosynthesis. Thus, NMR metabolomics provides an efficient, simple, and unbiased approach for providing rapid classification of promising drug leads that emerge from HTS. This is critical since HTS does not provide any information on mechanisms of action, only relative activity with a high-false positive rate. Instead, the *in vivo* biological activity of a novel lead can be inferred by its relative clustering to existing drug classes in an OPLS-DA scores plot derived from metabolomics data. Importantly, a chemical lead that

forms a distinct cluster from known drugs infers a potential new mechanism of action and a reason to prioritize the chemical lead for a detailed follow-up investigation. Of course, the induced metabolomic changes relative to untreated cells provide further confirmation of *in vivo* efficacy, which was implied from the HTS results. While the technique was demonstrated with *M. smegmatis*, it is generally applicable to bacterial pathogens and the effect of therapeutic agents on human cell lines in addition to the analysis of biofluids.

## METHODS

**Determining Optimal Drug Dosage for NMR Metabolomics Experiments.** *M. smegmatis* mc<sup>2</sup>155 cells were grown at 37 °C with shaking at 200 rpm in 50 mL of Middlebrook 7H9 media until an average optical density at 600 nm (OD<sub>600</sub>) of 0.6 was achieved. Each drug was titrated over a concentration range of 1 to 24 times the literature MIC values and the cells were grown for an additional 2 h. The optical density was recorded, and the growth rate inhibition was calculated by comparing the optical density of the treated cells to the untreated cells in the 2 h time period. The desired drug dosage was determined where a drug concentration inhibits cell growth by approximately 50% of the growth rate of untreated cells.

**Sample Preparation.** A total of 190 *M. smegmatis* mc<sup>2</sup>155 cultures were grown in 50 mL of Middlebrook 7H9 at 37 °C with shaking at 200 rpm until an OD<sub>600</sub> of 0.6 was achieved. A total of 40 untreated cultures were used as a control, and 10 cultures were inoculated with each antibiotic at the optimal dosage needed to inhibit cell growth by ~50% as described in Table 1. The cells were then grown for an additional 2 h. The used media was removed, and the cells were washed 3 times and resuspended with 1 mL of ice-cold double distilled water. The cells were lysed using a FastPrep-24 instrument for 60 s at 6 m/s, and the supernatant was extracted and frozen in a dry ice/ethanol bath. The samples were lyophilized and then resuspended with 700 μL of 99.8% D<sub>2</sub>O solution containing 50 mM phosphate buffer (pH 7.2, uncorrected) and 50 μM 3-(trimethylsilyl)propionic acid-2,2,3,3-*d*<sub>4</sub> (TMS-*d*<sub>4</sub>) as an internal standard for chemical shift referencing. The samples were then centrifuged for 5 min to remove any insoluble material, and 600 μL of the supernatant was transferred to an NMR tube.

**NMR Data Collection and Processing.** The NMR spectra were collected on a Bruker 500 MHz Avance spectrometer equipped with a triple resonance and *z* axis gradient cryoprobe. A BACS-120 sample changer was used for automated data collection. 1D <sup>1</sup>H NMR spectra were collected using excitation sculpting to remove the solvent signal and maintain a flat spectral baseline.<sup>33</sup> A total of 32 K data points with a spectral width of 5482.5 Hz, 16 dummy scans, and 128 scans were used to obtain each spectrum. The data was processed automatically using ACD/1D NMR Manager (Advanced Chemistry Development). Intelligent bucketing was used to integrate each spectral region with a bin size of 0.025 ppm. Each NMR spectrum was center averaged to minimize any experimental variations between bacterial cultures as follows:<sup>34</sup>

$$Z = \frac{X_i - \bar{X}}{\sigma} \quad (1)$$

where  $\bar{X}$  is the average signal intensity,  $\sigma$  is the standard deviation in the signal intensity, and  $X_i$  is the signal intensity within a bin. Noise regions of the spectra were omitted from the PCA analysis by setting the corresponding bins to zero.<sup>35</sup>

OPLS-DA and PCA was performed using Simca-11.5+ (Umetrics), where each <sup>1</sup>H NMR spectra was reduced to a single point in the 2D OPLS-DA and PCA scores plot. The OPLS-DA was calculated with two classes, untreated *versus* drug-treated cell cultures, for the Y matrix with the NMR data incorporated into the X matrix. The OPLS-DA model was cross validated using a modified version of the leave-one-out technique, where 1 out of every 7 samples (spectra) were left out to calculate a model and predict the left out data.<sup>36</sup> The procedure was sequentially repeated leaving out a different one-seventh of the data.

The predicted data was then compared to the original data, where the quality assessment ( $Q^2$ ) score provides a qualitative measure of the predictability of the model based on the consistency between the predicted and original data. An ideal value for  $Q^2$  is 1, where a typical value for a biological model is  $\geq 0.4$ .

Metabolomic tree diagrams with corresponding bootstrap values were created using our PCAtoTree program to interpret the OPLS-DA clustering pattern.<sup>25</sup> The metabolomics tree diagram is based on the Euclidean distances between the cluster centers from the 2D OPLS-DA scores plot. Standard bootstrapping techniques are used to generate a set of 100 distance matrices by randomly resampling the cluster centers and Euclidean distances. The set of 100 distance matrices are then used by PHYLIP (<http://www.phylip.com>),<sup>37</sup> phylogenetic software package, to generate 100 tree diagrams and a consensus tree diagram. The bootstrap numbers on the consensus tree diagram indicates the number of times each node was present in the set of 100 tree diagrams, where a bootstrap number below 50% indicates a generally insignificant node or insignificant separation between the clusters.

Four additional OPLS-DA models were generated to identify specific metabolites associated with drug activity: (i) inhibition of translation, transcription, or DNA supercoiling drug-treated cells *versus* untreated cells, (ii) inhibition of mycolic acid synthesis drug-treated cells *versus* untreated cells, (iii) inhibition of cell wall synthesis drug-treated cells *versus* untreated cells, and (iv) the three TAAFC compounds *versus* untreated cells. S-plots and loading plots were generated from each OPLS-DA model. Bins (chemical shift values) demonstrating a covariance of greater than 0.10 or less than -0.10 were identified as major contributors to the class separation. Metabolites were identified from this list of chemical shifts using the Human Metabolome Database (HMDB, <http://www.hmdb.ca/>)<sup>38</sup> with a chemical shift tolerance of 0.02 ppm. Metabolic network maps were then generated using Cytoscape (<http://www.cytoscape.org/>)<sup>39</sup> with the MetScape<sup>40</sup> plugin for the top 100 metabolite predicted by HMDB. Metabolites were excluded if not part of a network or not present in *M. smegmatis*.

## ASSOCIATED CONTENT

### Supporting Information

Two figures of PCA scores plots and the associated metabolomic tree, four figures of OPLS-DA S-plots and loading plots, and four tables listing metabolites affected by each drug class. This material is available free of charge *via* the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [rpowers3@unl.edu](mailto:rpowers3@unl.edu).

## ACKNOWLEDGMENTS

We acknowledge O. Chacon for helpful and valuable review of this manuscript. S. Halouska was partially supported by O. Chacon's grant (R21 AI087561) to standardize NMR techniques included in this publication. This work was supported in part by funds from the America Heart Association (0860033Z) and by the National Institute of Allergy and Infectious Diseases (R21 AI087561) to R.P. The research was performed in facilities renovated with support from the National Institutes of Health (NIH, RR015468-01).

## REFERENCES

- (1) Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., and Daszak, P. (2008) Global trends in emerging infectious diseases. *Nature* 451, 990–993.
- (2) Meya, D. B., and McAdam, K. P. W. J. (2007) The TB pandemic: an old problem seeking new solutions. *J. Intern. Med.* 261, 309–329.

- (3) Gwynn, M. N., Portnoy, A., Rittenhouse, S. F., and Payne, D. J. (2010) Challenges of antibacterial discovery revisited. *Ann. N.Y. Acad. Sci.* 1213, 5–19.
- (4) Talbot, G. H., Bradley, J., Edwards, J. E. Jr., Gilbert, D., Scheld, M., and Bartlett, J. G. (2006) Bad bugs need drugs: an update on the development pipeline from the Antimicrobial Availability Task Force of the Infectious Diseases Society of America. *Clin. Infect. Dis.* 42, 657–668.
- (5) Coates, A. R., Halls, G., and Hu, Y. (2011) Novel classes of antibiotics or more of the same? *Br. J. Pharmacol.* 163, 184–194.
- (6) Fischbach, M. A., and Walsh, C. T. (2009) Antibiotics for emerging pathogens. *Science* 325, 1089–1093.
- (7) Infectious Diseases Society of America (2010) The 10 × '20 Initiative: pursuing a global commitment to develop 10 new antibacterial drugs by 2020. *Clin. Infect. Dis.* 50, 1081–1083.
- (8) Brown, D. (2007) Unfinished business: target-based drug discovery. *Drug Discovery Today* 12, 1007–1012.
- (9) Bunnage, M. E. (2011) Getting pharmaceutical R&D back on target. *Nat. Chem. Biol.* 7, 335–339.
- (10) Powers, R. (2009) NMR metabolomics and drug discovery. *Magn. Reson. Chem.* 47 (Suppl 1), S2–11.
- (11) Baker, M. (2011) Metabolomics: From small molecules to big ideas. *Nat. Methods* 8, 117–121.
- (12) Lindon, J. C., Holmes, E., and Nicholson, J. K. (2001) Pattern recognition methods and applications in biomedical magnetic resonance. *Prog. Nucl. Magn. Reson. Spectrosc.* 39, 1–40.
- (13) Bylesjoe, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., and Trygg, J. (2007) OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemom.* 20, 341–351.
- (14) Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., and Wishart, D. S. (2010) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res.* 39, D1035–1041.
- (15) Chung, G. A., Aktar, Z., Jackson, S., and Duncan, K. (1995) High-throughput screen for detecting antimycobacterial agents. *Antimicrob. Agents Chemother.* 39, 2235–2238.
- (16) Stephan, J., Mailaender, C., Etienne, G., Daffe, M., and Niederweis, M. (2004) Multidrug resistance of a porin deletion mutant of *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* 48, 4163–4170.
- (17) Malik, M., Lu, T., Zhao, X., Singh, A., Hattan, C. M., Domagala, J., Kerns, R., and Drlica, K. (2005) Lethality of quinolones against *Mycobacterium smegmatis* in the presence or absence of chloramphenicol. *Antimicrob. Agents Chemother.* 49, 2008–2014.
- (18) Lu, T., Zhao, X., Li, X., Hansen, G., Blondeau, J., and Drlica, K. (2003) Effect of chloramphenicol, erythromycin, moxifloxacin, penicillin and tetracycline concentration on the recovery of resistant mutants of *Mycobacterium smegmatis* and *Staphylococcus aureus*. *J. Antimicrob. Chemother.* 52, 61–64.
- (19) Halouska, S., Chacon, O., Fenton, R. J., Zinniel, D. K., Barletta, R. G., and Powers, R. (2007) Use of NMR metabolomics to analyze the targets of D-cycloserine in mycobacteria: role of D-alanine racemase. *J. Proteome Res.* 6, 4608–4614.
- (20) Burguiere, A., Hitchen, P. G., Dover, L. G., Dell, A., and Besra, G. S. (2005) Altered expression profile of mycobacterial surface glycopeptidolipids following treatment with the antifungal azole inhibitors econazole and clotrimazole. *Microbiology* 151, 2087–2095.
- (21) Mick, V., Rebollo, M. J., Lucia, A., Garcia, M. J., Martin, C., and Ainsa, J. A. (2008) Transcriptional analysis of and resistance level conferred by the aminoglycoside acetyltransferase gene *aac(2)-Id* from *Mycobacterium smegmatis*. *J. Antimicrob. Chemother.* 61, 39–45.
- (22) Chacon, O., Feng, Z., Harris, N. B., Caceres, N. E., Adams, L. G., and Barletta, R. G. (2002) *Mycobacterium smegmatis* D-alanine racemase mutants are not dependent on D-alanine for growth. *Antimicrob. Agents Chemother.* 46, 47–54.
- (23) Flores, A. R., Parsons, L. M., and Pavelka, M. S. Jr. (2005) Characterization of novel *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* mutants hypersusceptible to  $\beta$ -lactam antibiotics. *J. Bacteriol.* 187, 1892–1900.
- (24) Barker, M., and Rayens, W. (2003) Partial least squares for discrimination. *J. Chemom.* 17, 166–173.
- (25) Werth, M. T., Halouska, S., Shortridge, M. D., Zhang, B., and Powers, R. (2009) Analysis of metabolomic PCA data using tree diagrams. *Anal. Biochem.* 399, 58–63.
- (26) Rosa, S. M. L. J., Antunes-Madeira, M. C., Jurado, A. S., and Madeira, V. V. M. C. (2000) Amiodarone interactions with membrane lipids and with growth of *Bacillus stearothermophilus* used as a model. *Appl. Biochem. Biotechnol.* 87, 165–175.
- (27) Oliva, B., O'Neill, A. J., Miller, K., Stubbings, W., and Chopra, I. (2004) Anti-staphylococcal activity and mode of action of clofazimine. *J. Antimicrob. Chemother.* 53, 435–440.
- (28) Kristiansen, J. E., Thomsen, V. F., Martins, A., Viveiros, M., and Amaral, L. (2010) Non-antibiotics reverse resistance of bacteria to antibiotics. *In Vivo* 24, 751–754.
- (29) O'Neill, A. J., Miller, K., Oliva, B., and Chopra, I. (2004) Comparison of assays for detection of agents causing membrane damage in *Staphylococcus aureus*. *J. Antimicrob. Chemother.* 54, 1127–1129.
- (30) Hurdle, J. G., O'Neill, A. J., Chopra, I., and Lee, R. E. (2011) Targeting bacterial membrane function: An underexploited mechanism for treating persistent infections. *Nat. Rev. Microbiol.* 9, 62–75.
- (31) Kwon, H. H., Tomioka, H., and Saito, H. (1995) Distribution and characterization of  $\beta$ -lactamases of mycobacteria and related organisms. *Tuber. Lung Dis.* 76, 141–148.
- (32) Paulin, L. G., Brander, E. E., and Poso, H. J. (1985) Specific inhibition of spermidine synthesis in *Mycobacteria* spp. by the dextro isomer of ethambutol. *Antimicrob. Agents Chemother.* 28, 157–159.
- (33) Hwang, T.-L., and Shaka, A. J. (1995) Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients. *J. Magn. Reson., Ser. A* 112, 275–279.
- (34) Craig, A., Cloarec, O., Holmes, E., Nicholson, J. K., and Lindon, J. C. (2006) Scaling and normalization effects in NMR spectroscopic metabolomic data sets. *Anal. Chem.* 78, 2262–2267.
- (35) Halouska, S., and Powers, R. (2006) Negative impact of noise on the principal component analysis of NMR data. *J. Magn. Reson.* 178, 88–95.
- (36) Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J. J., van Duinhoven, J. P. M., and van Dorsten, F. A. (2008) Assessment of PLS-DA cross validation. *Metabolomics* 4, 81–89.
- (37) Retief, J. D. (2000) Phylogenetic analysis using PHYLIP. *Methods Mol. Biol.* 132, 243–258.
- (38) Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De, S. A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazzyrova, A., Shaykhtudinov, R., Li, L., Vogel, H. J., and Forsythe, I. (2009) HMDB: a knowledge-base for the human metabolome. *Nucleic Acids Res.* 37, D603–D610.
- (39) Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432.
- (40) Gao, J., Tarcea, V. G., Karnovsky, A., Mirel, B. R., Weymouth, T. E., Beecher, C. W., Cavalcoli, J. D., Athey, B. D., Omenn, G. S., Burant, C. F., and Jagadish, H. V. (2010) Metscape: a Cytoscape plugin for visualizing and interpreting metabolomic data in the context of human metabolic networks. *Bioinformatics* 26, 971–973.