



Target Identification of *Mycobacterium tuberculosis* Phenotypic Hits Using a Concerted Chemogenomic, Biophysical, and Structural Approach

OPEN ACCESS

Edited by:

Tea Lanisnik Rizner,
University of Ljubljana, Slovenia

Reviewed by:

Alberto Cassetta,
Istituto di Cristallografia (CNR), Italy
Anderluh Marko,
University of Ljubljana, Slovenia

***Correspondence:**

Grace Mugumbate
grace@ebi.ac.uk
Vitor Mendes
vgm23@cam.ac.uk

† Present Address:

George Papadatos,
GlaxoSmithKline, Medicines Research
Centre, Stevenage, United Kingdom

‡ These authors have contributed
equally to this work and co-first
authors.

Specialty section:

This article was submitted to
Experimental Pharmacology and Drug
Discovery,
a section of the journal
Frontiers in Pharmacology

Received: 24 July 2017

Accepted: 12 September 2017

Published: 26 September 2017

Citation:

Mugumbate G, Mendes V,
Blaszczuk M, Sabbah M,
Papadatos G, Lelievre J, Ballell L,
Barros D, Abell C, Blundell TL and
Overington JP (2017) Target
Identification of *Mycobacterium
tuberculosis* Phenotypic Hits Using a
Concerted Chemogenomic,
Biophysical, and Structural Approach.
Front. Pharmacol. 8:681.
doi: 10.3389/fphar.2017.00681

Grace Mugumbate^{1,2*‡}, Vitor Mendes^{2*‡}, Michal Blaszczuk^{2‡}, Mohamad Sabbah³,
George Papadatos^{1†}, Joel Lelievre⁴, Lluís Ballell⁴, David Barros⁴, Chris Abell³,
Tom L. Blundell² and John P. Overington^{1,5}

¹ European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge, United Kingdom, ² Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom, ³ Department of Chemistry, University of Cambridge, Cambridge, United Kingdom, ⁴ Diseases of the Developing World, GlaxoSmithKline, Madrid, Spain, ⁵ Medicines Discovery Catapult, Alderley Edge, United Kingdom

Mycobacterium phenotypic hits are a good reservoir for new chemotypes for the treatment of tuberculosis. However, the absence of defined molecular targets and modes of action could lead to failure in drug development. Therefore, a combination of ligand-based and structure-based chemogenomic approaches followed by biophysical and biochemical validation have been used to identify targets for *Mycobacterium tuberculosis* phenotypic hits. Our approach identified EthR and InhA as targets for several hits, with some showing dual activity against these proteins. From the 35 predicted EthR inhibitors, eight exhibited an IC₅₀ below 50 μM against *M. tuberculosis* EthR and three were confirmed to be also simultaneously active against InhA. Further hit validation was performed using X-ray crystallography yielding eight new crystal structures of EthR inhibitors. Although the EthR inhibitors attain their activity against *M. tuberculosis* by hitting yet undefined targets, these results provide new lead compounds that could be further developed to be used to potentiate the effect of EthA activated pro-drugs, such as ethionamide, thus enhancing their bactericidal effect.

Keywords: *Mycobacterium tuberculosis*, phenotypic hits, target identification, drug resistance, EthR, InhA

INTRODUCTION

The pursuit of new and alternative drugs for the treatment of tuberculosis has led to phenotypic screens taking center stage in identifying novel and effective drug candidates (Lechartier et al., 2014). This screening effort has resulted in thousands of bioactive compounds released into the public databases, by researchers. The open access ChEMBL database (Gaulton et al., 2012, 2016) houses (as of February 2017) more than 16 datasets containing phenotypic hits for Neglected Tropical Diseases (<https://www.ebi.ac.uk/chemblntd>). These include the GSK TCAMS dataset containing about 776 phenotypic compounds that have shown activity against *Mycobacterium bovis* (IC₅₀ < 10 μM), and have exhibited low hepatotoxicity levels (Ballell et al., 2013) in whole-cell

screening assays. In particular, 177 of these compounds were found to be non-toxic for HepG2 cells [(HepG2 IC₅₀/MIC) > 50] and have demonstrated the ability to also highly inhibit the growth of the related *Mycobacterium tuberculosis* (MIC against H37Rv of < 10 μM), the main causative agent of tuberculosis, a disease affecting about 9 million people every year.

The success of phenotypic hits, in preclinical and clinical drug development depends to a greater extent, on the knowledge of their mechanism of action. However, the protein targets for most of the reported phenotypic hits are still to be identified. Experimentally identifying targets for all the phenotypic hits involves a broad range of approaches including genetic, proteomic, and transcriptional profiling, and more direct chemical-proteomic methods (Hart, 2005). It is therefore advantageous, after phenotypic screening to use computational methods that can predict likely protein targets of a given active small molecule, providing target-ligand pairs as starting points, thereby narrowing the search for the target proteins (Szardenings et al., 2004; Jenkins et al., 2006; Rebollo-Lopez et al., 2015).

In chemogenomics predictive approaches, targets are proposed by considering the chemical structural features of the active compounds (Jenkins et al., 2006) available in databases such as ChEMBL and PubChem BioAssay and comparing these to features of known ligands of a set of targets. Some of these approaches include ligand-based 2D chemical similarity assessment, cluster analysis and the use of 3D descriptors when the orphan compound has low similarity to all database molecules (Jenkins et al., 2006; Bender et al., 2007). In addition, machine-learning methods are also being used to extract targets and their associated ligands automatically from target-ligand information stored in multiple-target models. The multiple-category Laplacian-corrected Naïve Bayesian Classifiers (MCNBC) trained on extended-connectivity fingerprint of 964 targets classes in the WOMBAT database were originally described by Nidhi et al. (2006) and have been previously used (Ekins et al., 2013; Martínez-Jiménez et al., 2013) to identify targets for the TCAMS anti-TB phenotypic hits. However, experimental confirmation of the predicted targets for these compounds is limited and in most instances, still to be reported.

The enoyl-acyl carrier protein reductase (InhA) is a well-known anti-TB target involved in the biosynthesis of mycolic acids and is the target of first- and second-line TB drugs, isoniazid, and ethionamide respectively. The two drugs require metabolic activation inside *M. tuberculosis* cells by specific proteins; the heme enzyme catalase peroxidase (KatG) for isoniazid, and FAD-containing monooxygenase (EthA) for ethionamide; interestingly much of the clinically observed resistance is associated with the drug activation mechanisms (Takayama et al., 1972; Timmins and Deretic, 2006). Studies of the resistance mechanism of ethionamide revealed that an increase in the expression of HTH-Transcriptional regulator, EthR, reduces transcription and the level of active EthA protein, and results in resistance by *M. tuberculosis* (DeBarber et al., 2000; Dover et al., 2004). The discovery of this transcriptional regulator led to the notion of controlling EthA expression levels by targeting EthR and thus boosting the bactericidal effect of ethionamide (Frenois et al., 2004; Willand et al.,

2009). Therefore, identifying compounds that bind to EthR and prevent the repression of EthA would not only circumvent some of the resistance mechanisms but also reduce the required dosage and therefore the toxicity associated with this drug. In this work, we report for the first time, inhibitors of two *M. tuberculosis* targets, (EthR and InhA), identified using predictive *in silico* methods and definitively confirmed through biophysical methods, biochemical assays, and X-ray crystallography.

METHODS

Identification of Targets of Anti-TB Phenotypic Hits

In previous work, we reported a detailed description of the target prediction protocol (Mugumbate et al., 2015) comprising of two ligand-based and one structure-based approach to identify targets for the 776 TCAMS-TB compounds (Ballell et al., 2013) retrieved from the ChEMBL database (<https://www.ebi.ac.uk/chembl/>). The same protocol and models were used in this work to identify targets for further anti-TB phenotypic hit compounds; hence a brief description of the ligand-based method is presented here. The dataset of anti-TB compounds was prepared using a protocol in Pipeline Pilot version 8.5 (Warr, 2012), by generating 2D coordinates for each compound, and standardizing stereochemistry, and formal charges for common functional groups. In addition, salts and single atom were removed.

Ligand-Based Target Identification

A dataset containing 698,401 human and bacterial target-ligand pairs was extracted from ChEMBL version 17 (<http://www.ebi.ac.uk/ChEMBL>). The dataset contained targets with an annotated ChEMBL confidence score (CS) ranging from 7 to 9, indicating that the target—ligand pairs have unambiguously defined protein complexes or single proteins. Filtering the dataset to collect targets with 10 or more annotated ligands that exhibited IC₅₀, EC₅₀, or K_i values ≤ 10 μM or inhibition ≥ 50% against their respective targets resulted in a total of 2,257 unique human and bacterial proteins. Amongst these were 50 proteins belonging to the pathogenic bacterial species, *M. tuberculosis*. Human proteins were included to enable identification of potential host off-targets for the anti-tuberculosis compounds. MCNBCs were generated by training the protein identifiers on the structural features of their compounds, defined by their extended-connectivity fingerprint of diameter 6 (ECFP₆) (Rogers and Hahn, 2010) in Pipeline Pilot.

We determined the robustness of our models by a guaranteed random selection of compounds for both the training and test sets and minimized bias by presenting the model with a test set of previously unseen compounds using *two* methods (Supporting Information). In both instances the models displayed 75–90% accuracy, quantified by the percentage of compounds with correctly assigned targets reported in ranked positions 1–5. During the production stage, a model MCNBC trained on 695,902 target-ligand pairs was used to identify targets for all 776 phenotypic hits. We retained pairs of targets and compounds with positive Bayesian (NB)-scores and Z-scores ≥ 1.5. The

Bayesian Score is defined as the sum (P_{total}) of the logarithm of Laplacian-corrected Bayes rule of conditional probability [$P(A|Fi)$] for each fingerprint feature of a compound (12, 21). The Z-score distinguishes the compound scores for a particular target from the influence of the background noise. The score was obtained by predicting targets for > 1,200,000 compounds in ChEMBL version 17 database and generated the background information by calculate mean NB scores (μ) and standard deviation (σ) for each target. For each predicted target, standard scores (Z-score) were calculated from a statistical analysis of the NB scores for each compound: $Z\text{-score} = (X - \mu)/\sigma$ where, X is the NB score of a target for a compound. We have similarly generated Bayesian models using ChEMBL versions 20 and 21 data, however here we report the results obtained using ChEMBL 17 based models used at the time of the study.

These predictions were complemented by utilizing a second and independent ligand-based approach, The Similarity Ensemble Approach (SEA), which exploits chemical similarity between two sets of ligands, was used to study the pharmacological relationships between drugs (Keiser et al., 2009; Lounkine et al., 2012) and to predict targets for all 776 GSK phenotypic hits. At the time of this study the latest version of SEA utilized ChEMBL version 16 data and extended-connectivity fingerprint of diameter 4 (ECFP_4) fingerprints. To validate the method, the tool was used to predict targets for TB drugs of known modes of action where the bioactivity data are not found in ChEMBL version 16. We selected predicted *M. tuberculosis* protein targets from target-ligand pairs displaying expectation (E) values < 10^{-1} . In general, the E-value describes the significance of similarity between an orphan compound and a set of ligands and is a good indicator for the most likely targets (Keiser et al., 2007). A cut-off of 10^{-1} guaranteed selection of all possible target-ligand pairs and was suitable for our purposes.

Structure-Based Target Identification

The MCNBC and SEA methods provided a set of potential predicted targets, from which targets to use in structure-based target identification were selected. The Internal Coordinate Mechanism (ICM) method developed by Molsoft L.L.C (Totrov and Abagyan, 1997) was used to estimate the binding affinities of the anti-TB phenotypic hits to two predicted *M. tuberculosis* targets, EthR and InhA, based on the ICM scoring function, and available structures in wwPDB. The scoring function takes into consideration the energy changes due to protein-ligand binding and is defined as:

$$\Delta G = \Delta E_{IntFF} + T\Delta S_{Tor} + \alpha_1 \Delta E_{HBond} + \alpha_2 \Delta E_{HBDsol} + \alpha_3 \Delta E_{SolEl} + \alpha_4 \Delta E_{HPhob} + \alpha_5 Q_{Size}$$

where: ΔE_{IntFF} is change in van der Waals interactions of ligand and receptor and the internal force-field energy of the ligand, $T\Delta S_{Tor}$ is the change in free energy due to conformational entropy and weighted ($\alpha_1 - \alpha_5$), ΔE_{HBond} is the hydrogen bond term, ΔE_{HBDsol} accounts for the disruption of hydrogen bonds with solvent, ΔE_{SolEl} is the solvation electrostatic energy change upon binding, ΔE_{HPhob} is the hydrophobic free energy gain and Q_{Size} is the ligand size correction term.

Preparation of the EthR and InhA Proteins for Docking

We extracted the crystal structure of *M. tuberculosis* EthR bound to the inhibitor BDM31369 (PDB code: 3Q0V, resolution 1.95 Å) from the Protein DataBank in Europe (PDBe). The structure is one of the EthR structures containing synthetic inhibitors that occupy the allosteric pocket. The structure is dimeric and consists of chain A with 188 ordered residues and chain B containing 194 ordered residues, bound to the inhibitor BDM31369 (PDBE code; LL4). Some of the key residues interacting with the ligand include L87, I107, F110, M142, Y148, T149, V152, N176, and W207. Based on the structure validation from PDBe, chain A with higher quality backbone and side chain geometry percentiles, was used in our docking calculations. Using standard ICM-docking (Neves et al., 2012) receptor preparation tools the EthR structure was prepared by deleting all water molecules, adding and optimizing hydrogen atoms, adding missing heavy atoms, and adjusting amide groups. The “setup receptor” tool was used to generate receptor maps that cover the entire binding pocket of EthR using the crystallized ligand to determine the grid center, a grid size of 0.5 Å, and grid dimensions (x, y, z) of 51, 43, and 40 Å. The performance of the docking procedure was evaluated by redocking the co-crystallized ligand at thoroughness value of 1, which regenerated the crystal conformation at root mean square deviation (RMSD) of 1.81 Å between the C α atoms of the docked and the co-crystallized ligands. This proved that the software regenerated the crystal structure conformation and accurately docked the ligand.

Similarly, we extracted the 3D coordinates of an InhA structure in complex with the inhibitor PT92 (PDBe code: 4OHU, resolution of 1.6 Å) with the best structure-quality score from the PDBe. In the presence of the co-factor, NAD, the inhibitor occupies the substrate-binding pocket and primarily interact with residues L218, A157, M199, M161, Y158, and the nicotinamide moiety of NAD. The structure is tetrameric each chain with 289 residues and chains A, B, and C had percentiles of 100%. Chain A was used in the docking calculations. ICM molecules were generated using the ICM-docking receptor tools to generate grid maps of dimensions (x, y, z) 49, 46, and 44 Å, using grid step of 0.50 Å. After preparation of the crystal ligand, it was re-docked into the binding site and its conformation was regenerated at root mean square deviation of 0.86 Å between the C α atoms of the docked and co-crystal ligands.

Preparation of Small Molecules and Docking

Three-dimensional coordinates of 776 anti-TB compounds were generated using a Pipeline Pilot protocol. The molecules were imported into ICM (Totrov and Abagyan, 1997), amide bonds were fixed, hydrogen atoms were built and the structures were converted into ICM molecules. The compounds were docked into EthR and InhA using thoroughness/effort value of 2 and other default ICM parameters. The ICM scores were standardized by calculating the ligand efficiency indices (LEI), defined as the ratio between the ICM score and the number of heavy atoms for each docked molecule (Abad-Zapatero et al.,

2010). Compounds with $LEI \geq 1.0$ were retained for further analysis for each target protein.

Compounds

All anti-tuberculosis phenotypic compounds used in this work were provided by GSK TCAMS project, Spain, as 10 mM DMSO solutions.

Bacterial Strains and Cloning

inhA and *ethR* genes were amplified from chromosomal DNA of *M. tuberculosis* H37Rv DNA acquired from ATCC (ATCC25618D-2) and cloned in a pET28a vector (Novagen), modified with an N-terminal 6xHis-SUMO tag between BamHI/HindIII sites and in pHAT5 vector (Peranen et al., 1996) between BamHI/EcoRI sites, respectively. The resulting constructs were confirmed by DNA sequencing and transformed into *E. coli* BL21(DE3) strain (Invitrogen).

Recombinant Expression and Protein Purification

E. coli BL21(DE3) transformed with *inhA* and *ethR* containing plasmids cells were grown to mid-exponential growth phase ($OD_{610} = 0.8$) in LB media (Invitrogen) containing 30 mg.l^{-1} kanamycin (*inhA*) or 100 mg.l^{-1} Ampicilin (*ethR*) at 37°C . Isopropyl β -D-1-thiogalactopyranoside (IPTG) was then added at a final concentration of 0.5 mM to induce gene expression and the temperature was lowered to 18°C . EthR purification followed the protocol previously described by us (Surade et al., 2014). For *InhA* cells were harvested 18–20 h later by centrifugation and re-suspended in 50 mM HEPES pH 7.5, 0.5 M NaCl, 10% glycerol (w/v) and 20 mM imidazole with protease inhibitor tablets (Roche), DNaseI and 5 mM MgCl_2 . Cells were lysed by sonication and the cell lysate was centrifuged at 27,000 g for 30 min to remove cell debris. Recombinant *InhA* was purified with a HiTrap IMAC Sepharose FF column (GE-Healthcare), equilibrated with 50 mM HEPES pH 7.5, 0.5 M NaCl, 10% glycerol (w/v), and 20 mM imidazole. Elution was performed in the same buffer with 500 mM imidazole. Imidazole was removed by dialysis into 50 mM HEPES pH 7.5, 0.5 M NaCl and 10% glycerol (w/v) and SUMO tag was cleaved overnight at 4°C by adding Ulp1 Protease at 1:100 ratio. SUMO tag, Ulp1 protease and uncleaved SUMO-*InhA* were removed with a HiTrap IMAC Sepharose FF column (GE-Healthcare), equilibrated with 50 mM HEPES pH 7.5, 0.5 M NaCl, 10% glycerol (w/v), and 20 mM imidazole. Flow through containing *InhA* was collected, concentrated and loaded in a Superdex 200 column equilibrated with 50 mM HEPES pH 7.5, 150 mM NaCl, and 10% glycerol (w/v). Fraction purity was determined by SDS-PAGE. The purest fractions were pooled, concentrated to $\sim 12 \text{ mg.ml}^{-1}$, flash frozen in liquid nitrogen and stored at -80°C . Protein concentration was determined using the Bradford method.

Enzymatic Assays (InhA)

InhA activity was assessed by a spectrophotometric assay that followed the oxidation of NADH at $\lambda = 340 \text{ nm}$ in the presence of 2-trans-octanoyl-CoA. Compounds were dissolved in 100% DMSO. The enzyme reaction contains 30 mM PIPES pH 7.5, 50

mM NaCl, 1% (v/v) DMSO and 0.1 mM EDTA. *InhA* (100 nM) was pre-incubated for 10 min at room temperature with 0.25 mM NADH and compounds at a final concentration of 100 μM in 150 μl reaction volume. The reaction is started by the addition of 2-trans-octanoyl-CoA at a final concentration of 1.5 mM, prepared as described previously (He et al., 2006). The reactions were followed for 20 min using a spectrophotometer plate reader (CLARIOstar—BMG LABTECH). Inhibition percentage was calculated based on the initial rates of reaction. All experiments were performed in triplicate.

Surface Plasmon Resonance Assay (EthR)

The surface plasmon resonance (SPR) assay was performed on a BIACore T100 with a protocol similar to what was previously described (Surade et al., 2014). Briefly, *M. tuberculosis ethA* promoter DNA (106 bp) was amplified with 5'-CGGTCATGGATCCACGCTATCAAC-3' and 5'-biotin-CTGACTGGCCGCGGAGGTGGT-3' and immobilized in a SA chip (GE Healthcare). A control DNA fragment from pUC19 (113 bp) was amplified in a similar fashion. Biotinylated DNA (promoter and control) was then flowed over different channels and similar levels of resonance units (RUs) were achieved. For the initial screening a single run was performed for each compound. A 2 μM EthR solution was then incubated individually with each of the compounds at 100 μM , in 2 mM MgCl_2 , 10 mM TRIS/HCL pH 7.5, 200 mM NaCl, 0.1 mM EDTA, and 2% DMSO and flowed over the chip at 20 $\mu\text{l}/\text{min}$ for 120 s. For dissociation 150 s were given using the same flow rate. Binding was determined by subtracting the response of the control channel from the experimental channel at steady-state. Regeneration of the chip was performed between samples by flowing 0.03% SDS in running buffer for 60 s at 20 $\mu\text{l}/\text{min}$. To calculate IC_{50} , EthR-DNA-binding response, in the presence of compounds, was measured at several compound concentrations between (0.78 and 100 μM). Data were analyzed in Prism 5 (Graphpad software) and IC_{50} were calculated using a non-linear regression model. Dose response experiments were performed in triplicate.

Crystallization

EthR crystallization and compound soaking of crystals was carried out as previously described (Surade et al., 2014). An exception to this was compound 3 where co-crystallization was required. For this purpose EthR at 20 mg.ml^{-1} was incubated for 30 min on ice with 1 mM compound and 10% DMSO. Co-crystallization was performed at 18°C , using the sitting drop method and several commercial crystallization screens were tested (PEG I, (Qiagen); JCSG-plus, PACT, wizard classic I&II and III&IV (Molecular Dimensions). The highest diffracting condition for compound 3 was found in PACT crystallization screen solution E12 (0.2 M sodium malonate dibasic and 20% PEG 3350) using 0.3 μl EthR-compound 3 solution and 0.6 μl reservoir solution. All the other compounds that were co-crystallized with EthR did not yield crystals or the crystals diffracted only at low resolution. A cryogenic solution was prepared by adding ethylene glycol up to 25% v/v to mother

liquor. Crystals were briefly transferred to this solution, flash frozen in liquid nitrogen and stored for data collection.

Data Collection, Structure Solution, and Refinement

All data sets were collected at stations I02, and I04, at Diamond Light Source (Oxford, UK). Data collection and refinement statistics are summarized in Table S1.

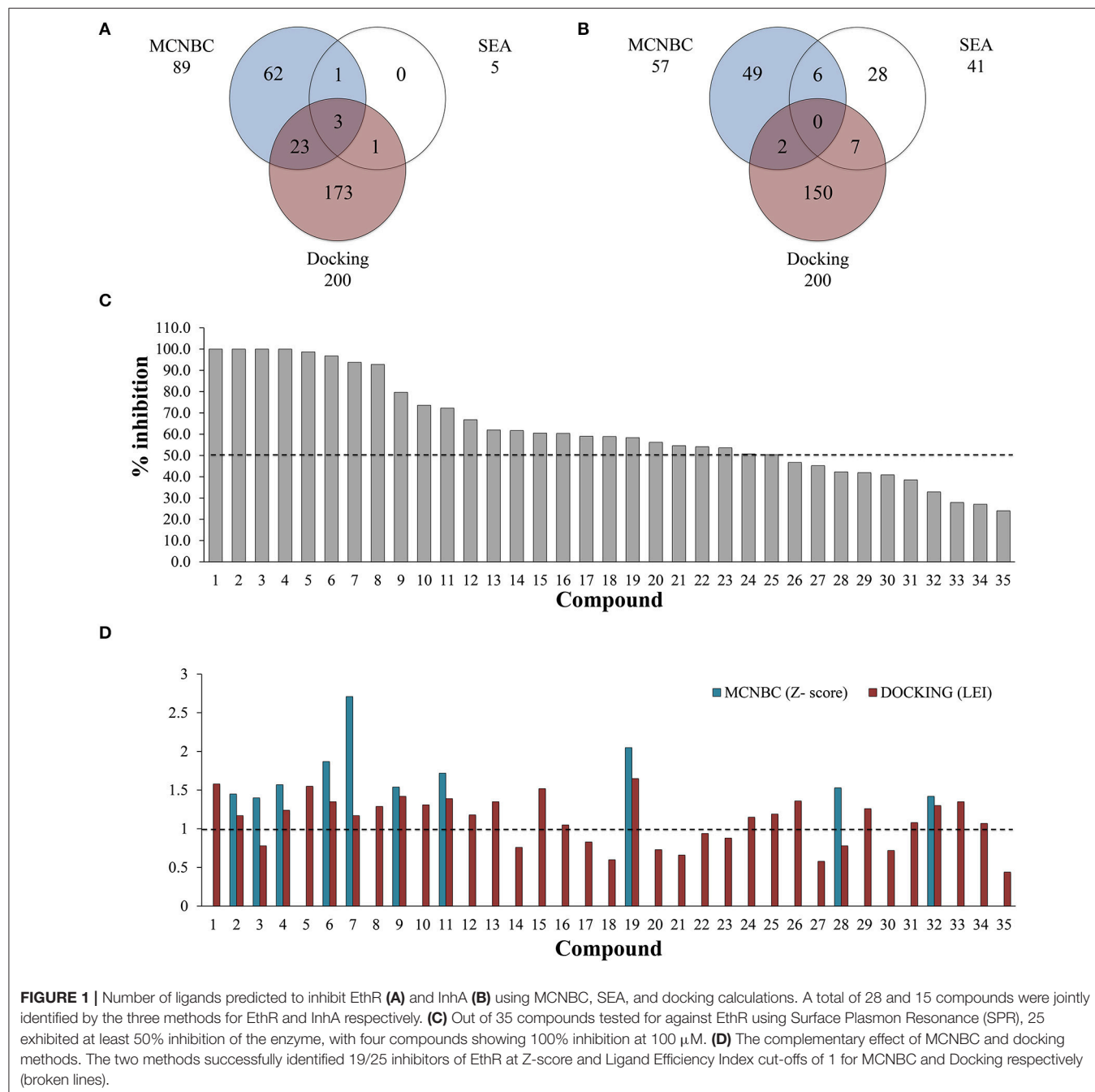
Diffraction data were indexed, integrated, and reduced using autoPROC from Global Phasing Limited (Vonrhein et al., 2011).

Molecular replacement was performed with Phaser (McCoy et al., 2007) using PDB structure 1T56 as a search model. Refinement was carried out iteratively with PHENIX (Adams et al., 2010) and Coot (Emsley et al., 2010), and ligand and water fitting was performed with Coot (Emsley et al., 2010).

RESULTS AND DISCUSSION

Predicted *M. tuberculosis* Targets

We initially predicted 1,462 human and bacterial potential targets for the 776 GSK anti-TB compounds using the Multiple Category



Naive Bayesian Classifiers (MCNBC). Each potential target displayed positive Bayesian scores (NB) and standard deviation scores (Z-scores) ≥ 1.5 for respective ligands. In this predicted target set, 146 (10%) were bacterial proteins and 25 of these were *M. tuberculosis* proteins, which were assigned to 132 compounds. To obtain a wider coverage and possibly identify novel targets, we mapped functional data and chemical bioactivity data of all predicted human and bacterial targets across their *M. tuberculosis* orthologs as reported in the OrthoMCL database (Chen et al., 2006). This increased the number of predicted *M. tuberculosis* targets to 137 (Figure 1) for 517 compounds, and accounted

for 4,575 target-ligand pairs (Table S2A). Similarly, the second ligand-based method, SEA generated 36,607 target-ligand pairs for 1,346 proteins from all the organisms in ChEMBL version 16. After considering the *M. tuberculosis* orthologs in OrthoMCL, 110 targets were assigned to 428 compounds to give 1225 target-ligand pairs (Table S2B).

The predicted *M. tuberculosis* protein targets are involved in many known essential biological processes, with lipid metabolism having the highest number of predictions (Table 1). Although ~5% of the predicted protein targets for these compounds are established as essential for *M. tuberculosis* (Tables S2A,B), many others are not. Nevertheless, all compounds are predicted to hit multiple proteins with at least one essential protein per compound that could account for the observed compound activity in the phenotypic screen. Top predicted essential protein targets include many enzymes involved in lipid biosynthesis such as Pks13 (*Rv3800c*), a polyketide synthase involved in mycolic acid biosynthesis (Gavalda et al., 2009) and InhA (*Rv1484*) an anti-TB target of the approved drugs isoniazid and ethionamide, both part of the FASII system. Amongst the several non-essential proteins, is EthR (*Rv3855*) that is involved in ethionamide resistance (Tables S2A,B). For these reasons, we selected InhA and EthR to validate further the computational predictions. As aforementioned, targeting EthR presents a potential way to “boost” the anti-TB activity of ethionamide (Willand et al., 2009).

EthR and InhA Inhibitors Chemogenomic Predictions

The mycobacterial HTH-transcriptional regulator (EthR) has been reported to be a facilitator for the resistance of the

TABLE 1 | Target prediction according to functional category.

Functional category	Number of proteins
DNA replication, recombination, and repair	10
Transcription	2
Ribosome and translation	7
Post-translational modification, protein degradation, and chaperones	10
Signal transduction	2
Detoxification and defense	5
Energy production	11
Cell envelope biogenesis	8
Nucleotide metabolism	13
Amino acid metabolism	10
Co-factor metabolism	7
Lipid Metabolism	33
Inorganic ion metabolism	3
Secondary metabolites biosynthesis and catabolism	15
General function prediction	18
Unknown function	7

TABLE 2 | Inhibition, IC₅₀ values, MIC, ligand efficiency (LE), and PDB codes for inhibitors 1–7, 10, 12, and 19.

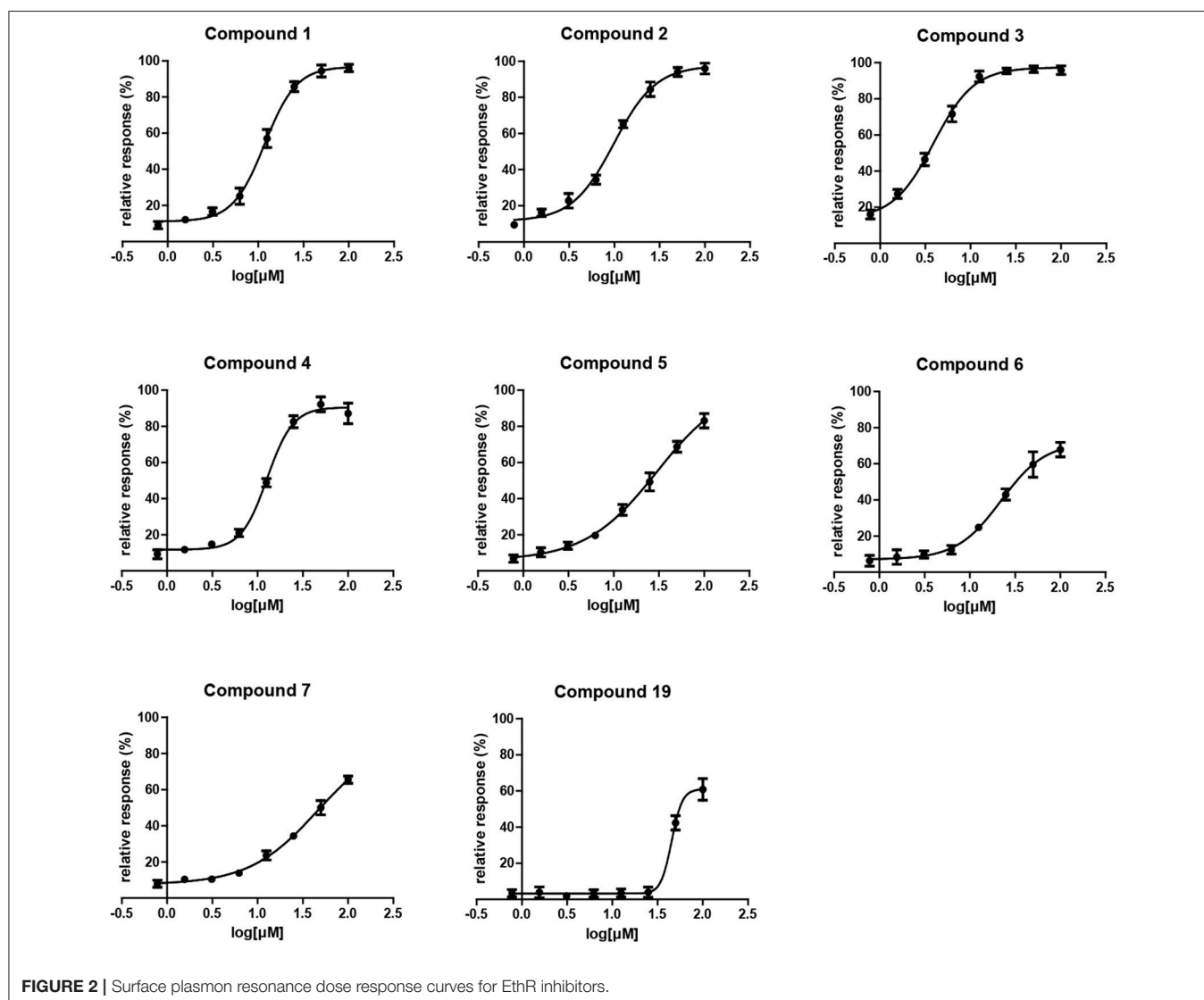
Compound	Inhibition (100 μ M) (%)	IC ₅₀	IC ₅₀ 95% confidence interval	MIC [#] (μ M)	LE	Structure
1 GSK1107112a	100	12	11–13	8.6	0.35	PDB code 5MXV
2 GSK1570606a	100	9.9	8.2–12	9.3	0.31	PDB code 5MYL
3 GSK2032710a	100	3.9	3.3–4.6	4.2	0.24	PDB code 5MYM
4 GSK735826a	100	13	11–15	2.7	0.28	Not obtained
5 GSK445886a	99	30	22–41	5.9	0.34	PDB code 5MYN
6 GSK735816a	97	22	20–25	2.3	0.30	PDB code 5MYR
7 GSK920684a	93	50	25–101	3.5	0.26	PDB code 5MYS
10 GSK921295a	74	*	*	8.7		PDB code 5MYT
12 SB- 435634	67	*	*	1.8		PDB code 5MYW
19 GSK2157753a	58	45	42–48	8.3	0.28	Not obtained

*IC₅₀ for these compounds was not possible to obtain in the tested conditions due to solubility issues. #MIC data was published before (Ballell et al., 2013).

second-line tuberculosis drug, ethionamide. The protein was amongst the *M. tuberculosis* proteins with the most compound predictions having 90 assigned ligands using both MCNBC and SEA (**Figure 1A**). EthR binding tunnel is a hydrophobic cavity with multiple aromatic residues decorating its surface and a few charged residues making it a promiscuous binding site known to bind molecules with different properties (Willand et al., 2009, 2010; Flipo et al., 2012; Surade et al., 2014; Villemagne et al., 2014; Nikiforov et al., 2016, 2017). It also possesses multiple hotspots regions where fragments bind (Surade et al., 2014; Villemagne et al., 2014; Nikiforov et al., 2016, 2017). It is therefore not surprising that this protein had a high number of predicted inhibitors as these characteristics make this site promiscuous. Extensive studies have already been performed on this protein. In excess of 20 crystal structures of EthR in complex with allosteric inhibitors have been reported. However, none of these have reached Phase I clinical trials, and therefore the hunt for inhibitors is still ongoing.

To complement the ligand-based predicted ligands for EthR, we used ICM dock (Neves et al., 2012) to dock all the 776 TCAMS compounds to a previously prepared crystal structure of *M. tuberculosis* EthR bound to an inhibitor BDM31369 (PDB code: 3Q0V). Most compounds docked into the long hydrophobic pocket of EthR and formed interactions with the side chains of W103, W207, and F110 residues and hydrogen bonds with N176 and N179. After ranking the compounds based on their calculated ligand efficiency index (LEI), 200 potential inhibitors displaying LEI greater than or equal to 1.0 were selected for further analysis. Three potential inhibitors of EthR were jointly predicted by the three complementary chemogenomic approaches, 23 compounds were identified by MCNBC and docking calculations whilst MCNBC/SEA and SEA/DOCKING each had one common compound (**Figure 1A**).

The clinical utility of the first-line anti-TB drug isoniazid (INH) against the enoyl acyl protein reductase InhA, is being challenged by the increasing drug resistance attributed to



mutations or deletions in the active site of katG gene, which encodes catalase-peroxidase KatG, an enzyme responsible for the metabolic activation of INH (Vilcheze and Jacobs, 2014). For this reason, it is imperative to identify new drugs that target InhA but circumvent the KatG activation pathway. We used the chemogenomic methods MCNBC and SEA to synergistically allocate more than 80 as potential inhibitors of InhA. From the docking calculations 159 compounds displayed $LEI \geq 1.0$ indicating favorable binding affinity to InhA (Figure 1B). Even though no compounds were jointly predicted by all three methods, about 15 compounds were commonly predicted by either docking and SEA, MCNBC, and docking or MCNBC and SEA approaches.

From the predicted EthR inhibitors, we selected a set of 26 compounds (1–16, 19, 24–26, 28–29, and 31–34), that were identified by one or more of the computational methods above, and for which physical samples were available for use in the *in vitro* assays. Approximately one third of these compounds displayed Z-scores ≥ 1.00 indicating similar structural features to those previously reported for EthR ligands in our training set. To investigate the ability of the chemogenomic methods to distinguish between true and false negatives our second set contained 9 compounds (17–18, 20–23, 27, 30, and 35) from the 776 GSK-TCAMS dataset, that fell outside the selection criteria and represent different molecular clusters. The target prediction scores for the 35 compounds are given in Table S3. About 51%

(17 compounds) of this set was predicted to inhibit both EthR and InhA using both ligand-based and structure-based methods. Hence, *in vitro* experiments were also performed to confirm the activities of the 35 compounds against InhA.

Validation of EthR Inhibitors

A single compound concentration run at 100 μM was performed using SPR to evaluate the inhibitory effect of individual compounds in EthR binding to DNA (Table S2A). From this run we could immediately observe that the 71% (25/35) of the compounds were active against EthR ($> 50\%$ inhibition) and that compounds that were outside the selection criteria all had low activities (inhibition $< 60\%$), (Figures 1C,D), confirming that these computational methods can distinguish true and false negatives. The top 20 compounds were then selected to determine IC_{50} values (Table 2 and Figure 2). Amongst these were compounds 1, 5, 8, 10, 12, and others identified solely from docking calculations, illustrating the complementarity of ligand-based and structure-based methods (Figures 1C,D). Eight of the tested compounds had an IC_{50} lower than 50 μM while for the others, accurate IC_{50} values were not possible to obtain due to solubility issues in the tested assay conditions. We then managed to obtain X-ray crystal structures for compounds 1, 2, 3, 5, 6, 7, 10, and 12 bound to EthR (Figures 3, 4). Omit maps for each of the ligands is available in Figure S1.

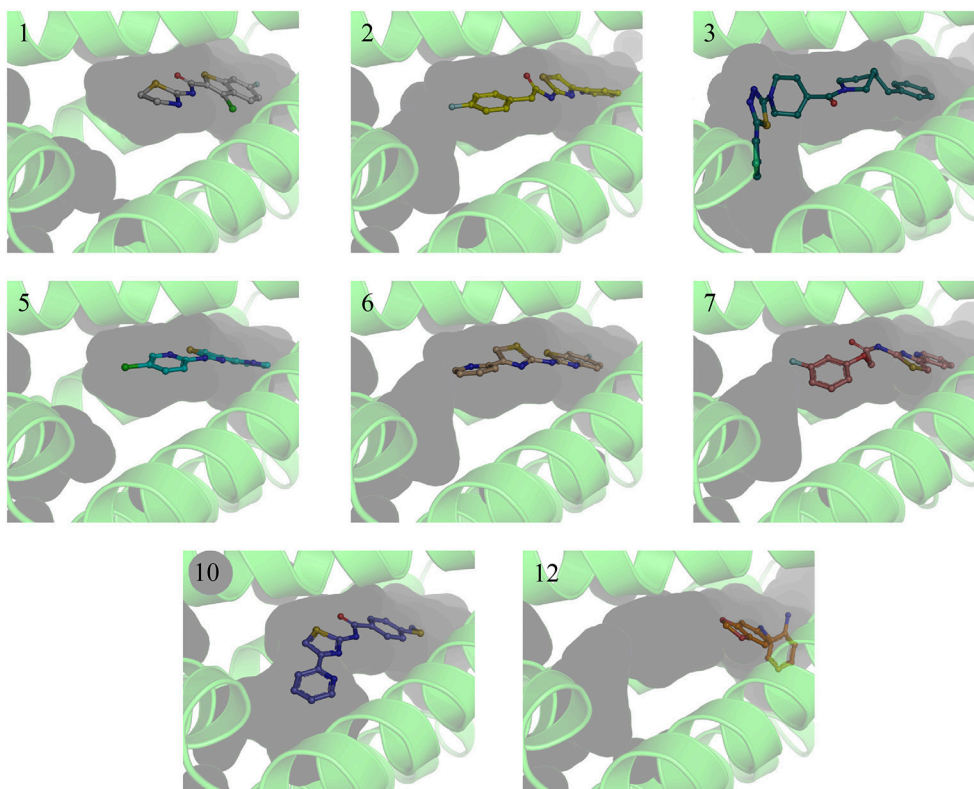


FIGURE 3 | X-ray crystal structures of EthR:inhibitor complexes showing binding cavity surface and binding modes of the ligands. Numbers reflect the respective compounds. Figures were made using Pymol.

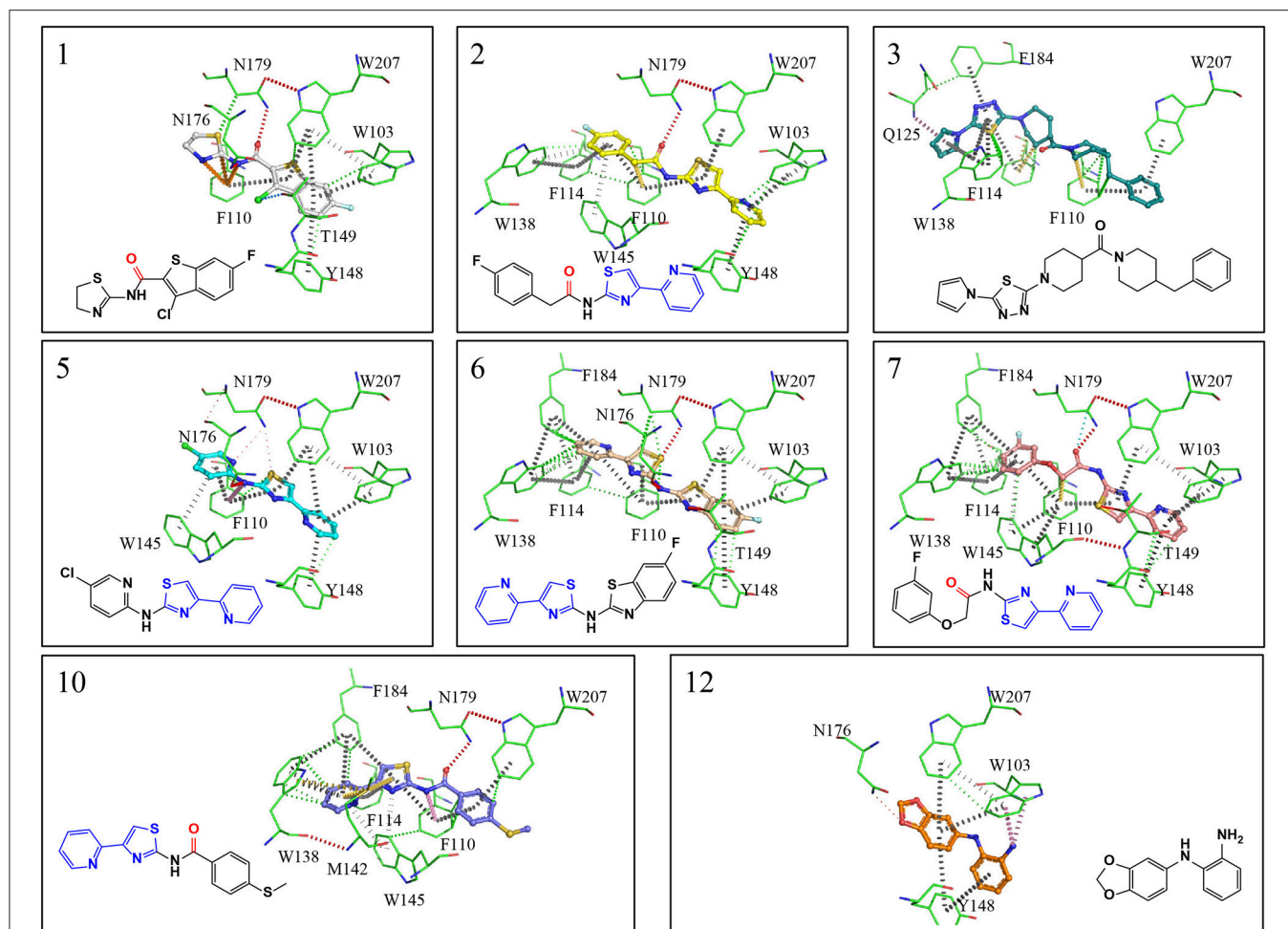
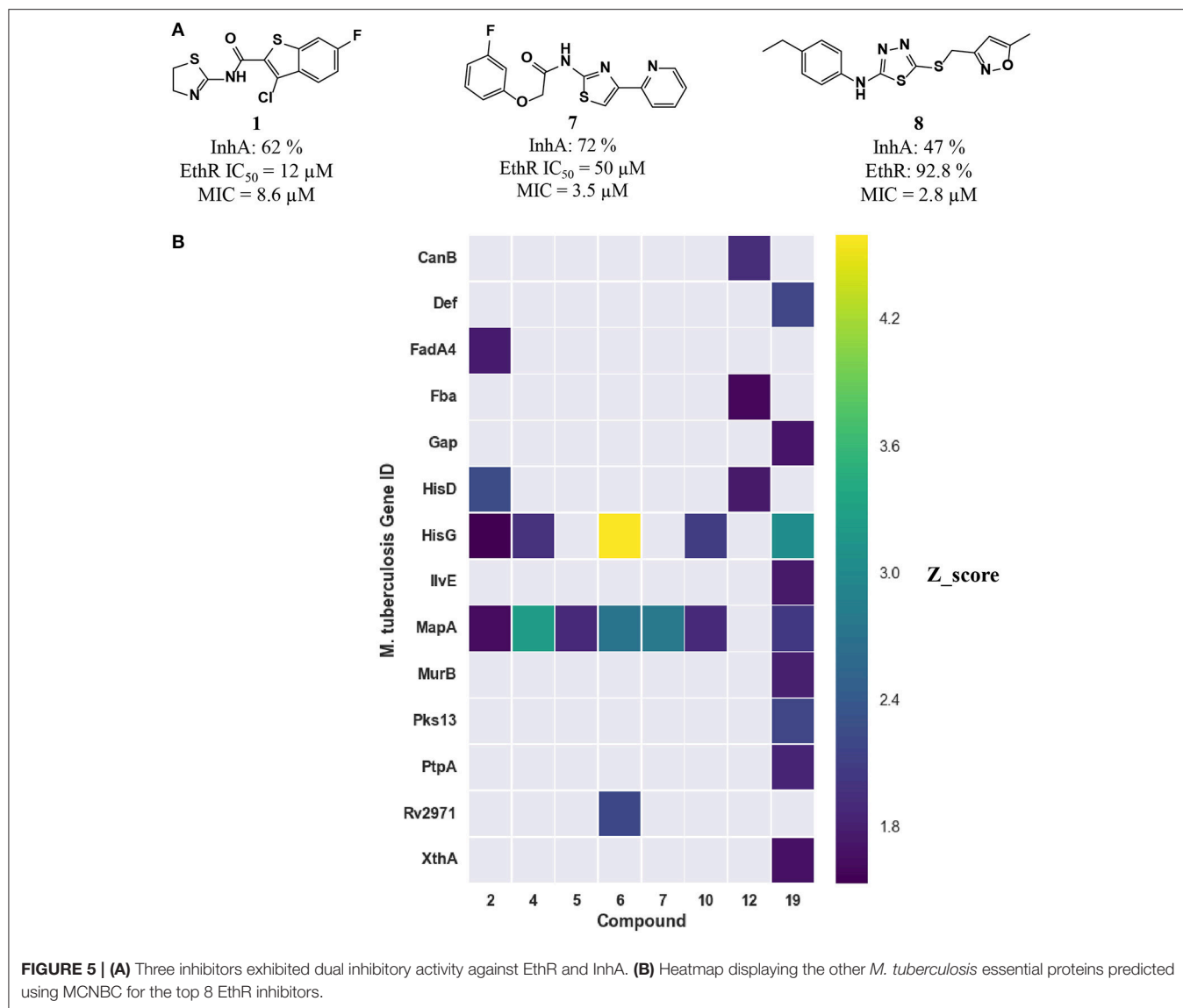


FIGURE 4 | X-ray crystal structures showing interactions maps between EthR and inhibitors. Red disks represent hydrogen bonds, blue dots halogen bonds, gray disks depict π - π interactions, green disks hydrophobic interactions, yellow disks sulphur- π and carbon- π interactions, pink disks donor- π interactions, orange disks cation- π interactions and cyan dots carbonyl interactions. Interactions were calculated using Intermezzo plugin for Pymol (Ochoa et al. unpublished). The carbonyl groups highlighted in red form hydrogen bonds with asn176 and are a common feature in many published EthR inhibitors. The common substructure (4-(pyridine-2-yl)thiazole) is highlighted in blue.

EthR binding site is characterized by a long hydrophobic tunnel surrounded mostly by hydrophobic residues, many of which are aromatic, plus four polar amino acids and a basic residue (Dover et al., 2004). The interactions of our compounds with EthR were dominated π - π interactions with (W103, F110, F114, W138, W145, Y148, F184, and W207), although hydrogen bonds were also present, with all but compound 3 forming at least one hydrogen bond with one of the polar residues (T149, N176, and N179) as seen in **Figure 4**. Other interactions such as hydrogen donor- π or cation- π (observed in compound 1 (F110), 3 (Q125), 5 (F110), 10 (F110), and 12 (W103), sulfur- π (observed in compound 10 with M142), carbon- π (observed in compound 2, 3 and 7), and halogen bond (observed in compound 1 with T149) were also present but less prevalent (**Figure 4**). Hydrophobic interactions, extending from W103 at the entrance of the binding site to W138 at the inner hydrophobic site are also observed for many of the compounds (**Figure 4**). This dense network of protein-ligand interactions appears to stabilize the

conformational changes in EthR and inhibition of DNA binding observed by SPR. Interestingly six of the best compounds (2, 4, 5, 6, 7, 10) share a common substructure composed by a 4-(pyridine-2-yl)thiazole (**Figure 4**), that for compounds 2, 5, and 7 occupies the same part of EthR binding site forming π -interactions with F110, W103, Y148, and W207, while for compounds 6 and 10 the same substructure forms π -interactions with F110, F114, W138, W145, and F184 (**Figure 4**). Another common feature shared between several of the compounds (1, 2, 3, 7, and 10) is the presence of a ketone that forms a hydrogen bond with N176 in all but compound 3. This ketone is a feature shared by many other EthR binding compounds reported in the literature (**Figure 4** and **Figure S2**). Although we did not obtain a structure for one of the most potent EthR inhibitors identified in this work (compound 4, $IC_{50} \approx 13 \mu M$) (**Table 2**), it is quite likely that this compound, like others reported here, interacts with EthR by forming similar interactions at the binding site.



Validation of InhA Inhibitors

To assess whether the predicted compounds are active against InhA, an enzymatic assay was performed with compounds tested at a single concentration of 100 μM. Only compounds **1**, **7**, and **8** showed inhibition greater than 40% (**Figure 5A**). IC₅₀s were not possible to calculate due to solubility issues in the required assay conditions. An attempt to address this issue by increasing DMSO concentration, to improve compound solubility, led to very low enzymatic activity that prevented obtaining reliable data.

Dual Activity against EthR and InhA

The computational target prediction approaches described here have highlighted the possibility of inhibiting two or more *M. tuberculosis* proteins by one of the phenotypic hits. Amongst the 35 compounds selected for *in vitro* assays against EthR, 17 compounds were also predicted to hit InhA. Three compounds displayed inhibitory activity against both EthR and InhA

(**Figure 5A**). Compound **1** is active against both EthR (IC₅₀ ≈ 12 μM) and InhA (62% inhibition at 100 μM) (**Table 2**), confirming the computational predictions. Interestingly, this compound was identified solely from structure-based predictions. Even though an attempt to obtain a co-crystal structure of the compound with InhA was unsuccessful, the docking LEI values for the compound to both enzymes (EthR: 1.58 and InhA: 1.11) indicate high binding affinity. The EthR-compound **1** co-crystal structure confirms the occupancy of the compound in the EthR binding site.

Although targeting EthR has been reported to boost the activity of the second-line drug ethionamide against InhA (Willand et al., 2009), EthR itself is not an essential protein and compounds that solely hit EthR will not show sterilizing activity without the concomitant use of ethionamide. Furthermore, the dual activity compounds reported here, are not sufficiently potent *in vitro* inhibitors of InhA to explain their observed MIC value. This indicates that although many of our predictions

were confirmed, the reasons behind whole cell activities of the compounds most likely lie elsewhere.

Interestingly some of the compounds reported here have chemical groups that are known to be modified by EthA. Thiophenes and thioethers are present in compounds **1**, **8**, and **10** and are known to be substrates for EthA (Mori et al., 2015; Grant et al., 2016). Furthermore, EthA was found to have a much wider substrate preference than initially thought and is known to modify also thiocarbamides and thioureas (Dover et al., 2007; Nishida and Ortiz de Montellano, 2011; Gopal and Dick, 2015; Grzegorzewicz et al., 2015). Although speculative it is possible that these compounds hit EthR but are also pro-drugs that are activated by EthA self-increasing their own activation by raising the levels of EthA available inside *M. tuberculosis* and subsequently hitting other targets.

A more conservative explanation is that the observed MICs are explained by a direct dual or polypharmacology by hitting other *M. tuberculosis* targets. To that effect, the heatmap in **Figure 4** displays the 16 essential *M. tuberculosis* proteins predicted for eight EthR inhibitors (compounds **2**, **4**, **5**, **6**, **7**, **10**, **12**, and **19**) using ligand-based approach. The other three inhibitors, **1**, **3**, and **8** were identified through a structure-based approach, hence are not shown in **Figure 4**. We observe that each compound has the potential to hit at least one essential protein, with MapA assigned for all but compound **12** (**Figure 5B**) and compound **19** potentially targeting 10 proteins. Furthermore, compounds **5** and **12** selected for both EthR and InhA from the structure-based predictions were assigned to further essential *M. tuberculosis* proteins by the ligand-based MCNBC method. A summary of the all the other predicted *M. tuberculosis* proteins for these compounds is given in Table S4. However, it is also possible that the observed MIC values are a result of a synergistic effect that also includes their activities against InhA and EthR and on other enzyme(s).

A designed multi-targeted approach in drug discovery has been attracting attention in recent years, with large scale functional genomics studies in several model organisms. Many single gene knockouts by themselves show no, or small, phenotypic effect with as little as ~19% of the genes being essential in a number of model organisms (Hopkins, 2008). Compounds that hit multiple targets have the capacity to explore synthetic lethal interaction that further expand the number of available targets in an organism (Hopkins, 2008). The fact that many of these compounds show strong EthR inhibition together with very good MIC values raises the prospect of further enhancing ethionamide killing effect with a combined therapy that uses an ethionamide boosting compound that also kills *M. tuberculosis* by hitting one or more enzymes.

Several other TB drugs thiacetazone, thiocarlide (isoxyl), and the recently developed perchlozone have been shown to be activated by EthA and hit the same target, HadABC, a group of enzymes involved in mycolic acid synthesis (Gopal

and Dick, 2015; Grzegorzewicz et al., 2015). Thiacetazone use has declined due to severe and sometimes fatal side effects (Miller et al., 1972), and isoxyl is no longer used because of failures in clinical outcomes due to its low plasma concentrations (Wang et al., 2012). Potentiating the activation of these drugs by increasing the amount of EthA available in the cells with EthR inhibitors might allow some of these drugs to be viable in modern TB chemotherapy by lowering the dosage required to kill *M. tuberculosis*. Perchlozone is a new TB drug approved in the Russian Federation to treat MDR-TB. Again, EthR inhibitors may potentiate the killing effects of this drug by increasing the amount of EthA available making it a much stronger anti-TB drug.

AUTHOR CONTRIBUTIONS

GM and GP generated the multi-category naive Bayesian models. GM designed the chemogenomic target prediction protocol, performed all ligand-based and structure-based target prediction experiments, analyzed the results, and selected target-compounds pairs for experimental validation. JL, LB, and DB prepared and provided all samples of the *M. tuberculosis* phenotypic hits. VM and MB cloned and purified the proteins, designed, and performed all biophysical and structural experiments. MS performed InhA enzymatic assays and synthesized InhA substrate. JPO, TLB, and CA supervised the work. GM and VM wrote the manuscript. All authors revised the manuscript.

FUNDING

GM was supported by a fellowship from the EMBL Interdisciplinary Postdoc (EIPOD) programme under Marie Skłodowska-Curie Actions COFUND (grant number 291772). VM and MB acknowledge Bill and Melinda Gates Foundation [subcontract by the Foundation for the National Institutes of Health (NIH)] (OPP1024021). VM further acknowledges Bill and Melinda Gates Foundation [subcontract by the Foundation for the National Institutes of Health (FNIH)] (OPP1158806). VM and MS acknowledge the European Community's Seventh Framework Programme (grant number 260872). GP would like to acknowledge the Wellcome Trust and the European Molecular Biology Laboratory for funding. JPO was funded by the member nation states of the European Molecular Biology Laboratory. TLB acknowledges The Wellcome Trust for funding and support grant number (200814/Z/16/Z).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fphar.2017.00681/full#supplementary-material>

REFERENCES

Abad-Zapatero, C., Perisic, O., Wass, J., Bento, A. P., Overington, J., Al-Lazikani, B., et al. (2010). Ligand efficiency indices for an effective mapping

of chemico-biological space: the concept of an atlas-like representation. *Drug Discov. Today* 15, 804–811. doi: 10.1016/j.drudis.2010.08.004
Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for

- macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 2), 213–221. doi: 10.1107/S0907444909052925
- Ballell, L., Bates, R. H., Young, R. J., Alvarez-Gomez, D., Alvarez-Ruiz, E., Barroso, V., et al. (2013). Fueling open-source drug discovery: 177 small-molecule leads against tuberculosis. *ChemMedChem* 8, 313–321. doi: 10.1002/cmdc.201200428
- Bender, A., Young, D., Jenkins, J., Serrano, M., Mikhailov, D., Clemons, P., et al. (2007). Chemogenomic data analysis: prediction of small-molecule targets and the advent of biological fingerprints. *Comb. Chem. High Throughput Screen.* 10, 719–731. doi: 10.2174/138620707782507313
- Chen, F., Mackey, A. J., Stoekert, C. J. Jr. and Roos, D. S. (2006). OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* 34, D363–D368. doi: 10.1093/nar/gkj123
- DeBarber, A. E., Mdllui, K., Bosman, M., Bekker, L. G., and Barry, C. E. III. (2000). Ethionamide activation and sensitivity in multidrug-resistant *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9677–9682. doi: 10.1073/pnas.97.17.9677
- Dover, L. G., Alahari, A., Gratraud, P., Gomes, J. M., Bhowruth, V., Reynolds, R. C., et al. (2007). EthA, a common activator of thiocarbamide-containing drugs acting on different mycobacterial targets. *Antimicrob. Agents Chemother.* 51, 1055–1063. doi: 10.1128/AAC.01063-06
- Dover, L. G., Corsino, P. E., Daniels, I. R., Cocklin, S. L., Tatituri, V., Besra, G. S., et al. (2004). Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR implicated in ethionamide resistance. *J. Mol. Biol.* 340, 1095–1105. doi: 10.1016/j.jmb.2004.06.003
- Ekins, S., Reynolds, R. C., Franzblau, S. G., Wan, B., Freundlich, J. S., and Bunin, B. A. (2013). Enhancing hit identification in *Mycobacterium tuberculosis* drug discovery using validated dual-event Bayesian models. *PLoS ONE*:8:e63240. doi: 10.1371/journal.pone.0063240
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 4), 486–501. doi: 10.1107/S0907444910007493
- Flipo, M., Willand, N., Lecat-Guillet, N., Hounsou, C., Desroses, M., Leroux, F., et al. (2012). Discovery of novel N-phenylphenoxyacetamide derivatives as EthR inhibitors and ethionamide boosters by combining high-throughput screening and synthesis. *J. Med. Chem.* 55, 6391–6402. doi: 10.1021/jm300377g
- Frenois, F., Engohang-Ndong, J., Loch, C., Baulard, A. R., and Villeret, V. (2004). Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis. *Mol. Cell* 16, 301–307. doi: 10.1016/j.molcel.2004.09.020
- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., et al. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 40, D1100–D1107. doi: 10.1093/nar/gkr777
- Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., et al. (2016). The ChEMBL database in 2017. *Nucleic Acids Res.* 45, D945–D954. doi: 10.1093/nar/gkw1074
- Gavalda, S., Leger, M. B., van der Rest, S., Tella, A., Bardou, F., Montrozier, H., et al. (2009). The Pks13/FadD32 crosstalk for the biosynthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Biol. Chem.* 284, 19255–19264. doi: 10.1074/jbc.M109.006940
- Gopal, P., and Dick, T. (2015). The new tuberculosis drug Perchlorone(R) shows cross-resistance with thiacetazone. *Int. J. Antimicrob. Agents* 45, 430–433. doi: 10.1016/j.ijantimicag.2014.12.026
- Grant, S. S., Wellington, S., Kawate, T., Desjardins, C. A., Silvis, M. R., Wivagg, C., et al. (2016). Baeyer-Villiger Monooxygenases EthA and MymA are required for activation of replicating and non-replicating *Mycobacterium tuberculosis* inhibitors. *Cell Chem. Biol.* 23, 666–677. doi: 10.1016/j.chembiol.2016.05.011
- Grzegorzewicz, A. E., Eynard, N., Quemard, A., North, E. J., Margolis, A., Lindenberg, J., et al. (2015). Covalent modification of the *Mycobacterium tuberculosis* FAS-II dehydratase by Isoxyl and Thiacetazone. *ACS Infect. Dis.* 1, 91–97. doi: 10.1021/id500032q
- Hart, C. (2005). Finding the target after screening the phenotype. *Drug Discov. Today* 10, 513–519. doi: 10.1016/S1359-6446(05)03415-X
- He, X., Alian, A., Stroud, R., and Ortiz de Montellano, P. R. (2006). Pyrrolidine carboxamides as a novel class of inhibitors of enoyl acyl carrier protein reductase from *Mycobacterium tuberculosis*. *J. Med. Chem.* 49, 6308–6323. doi: 10.1021/jm060715y
- Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* 4, 682–690. doi: 10.1038/nchembio.118
- Jenkins, J. L., Bender, A., and Davies, J. W. (2006). In silico target fishing: predicting biological targets from chemical structure. *Drug Discov. Today Technol.* 3, 413–421. doi: 10.1016/j.ddtec.2006.12.008
- Keiser, M. J., Roth, B. L., Armbruster, B. N., Ernsberger, P., Irwin, J. J., and Shoichet, B. K. (2007). Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* 25, 197–206. doi: 10.1038/nbt1284
- Keiser, M. J., Setola, V., Irwin, J. J., Laggner, C., Abbas, A. I., Hufeisen, S. J., et al. (2009). Predicting new molecular targets for known drugs. *Nature* 462, 175–181. doi: 10.1038/nature08506
- Lechartier, B., Rybniker, J., Zumla, A., and Cole, S. T. (2014). Tuberculosis drug discovery in the post-post-genomic era. *EMBO Mol. Med.* 6, 158–168. doi: 10.1002/emmm.201201772
- Lounkine, E., Keiser, M. J., Whitebread, S., Mikhailov, D., Hamon, J., Jenkins, J. L., et al. (2012). Large-scale prediction and testing of drug activity on side-effect targets. *Nature* 486, 361–367. doi: 10.1038/nature11159
- Martínez-Jiménez, F., Papadatos, G., Yang, L., Wallace, I. M., Kumar, V., Pieper, U., et al. (2013). Target prediction for an open access set of compounds active against *Mycobacterium tuberculosis*. *PLoS Comput. Biol.* 9:e1003253. doi: 10.1371/journal.pcbi.1003253
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40(Pt 4), 658–674. doi: 10.1107/S0021889807021206
- Miller, A. B., Nunn, A. J., Robinson, D. K., Fox, W., Somasundaram, P. R., and Tall, R. (1972). A second international cooperative investigation into thioacetazone side effects. *Bull. World Health Organ.* 47, 211–227.
- Mori, G., Chiarelli, L. R., Esposito, M., Makarov, V., Bellinzoni, M., Hartkoorn, R. C., et al. (2015). Thiophenecarboxamide derivatives activated by EthA Kill *Mycobacterium tuberculosis* by Inhibiting the CTP Synthetase PyrG. *Chem. Biol.* 22, 917–927. doi: 10.1016/j.chembiol.2015.05.016
- Mugumbate, G., Abrahams, K. A., Cox, J. A., Papadatos, G., van Westen, G., Lelievre, J., et al. (2015). Mycobacterial dihydrofolate reductase inhibitors identified using chemogenomic methods and *in vitro* validation. *PLoS ONE* 10:e0121492. doi: 10.1371/journal.pone.0121492
- Neves, M. A., Totrov, M., and Abagyan, R. (2012). Docking and scoring with ICM: the benchmarking results and strategies for improvement. *J. Comput. Aided Mol. Des.* 26, 675–686. doi: 10.1007/s10822-012-9547-0
- Nidhi, Glick, M., Davies, J. W., and Jenkins, J. L. (2006). Prediction of biological targets for compounds using multiple-category Bayesian models trained on chemogenomics databases. *J. Chem. Inf. Model.* 46, 1124–1133. doi: 10.1021/ci060003g
- Nikiforov, P. O., Blaszczyk, M., Surade, S., Boshoff, H. I., Sajid, A., Delorme, V., et al. (2017). Fragment-sized EthR inhibitors exhibit exceptionally strong Ethionamide boosting effect in whole-cell *Mycobacterium tuberculosis* assays. *ACS Chem. Biol.* 12, 1390–1396. doi: 10.1021/acscchembio.7b00091
- Nikiforov, P. O., Surade, S., Blaszczyk, M., Delorme, V., Brodin, P., Baulard, A. R., et al. (2016). A fragment merging approach towards the development of small molecule inhibitors of *Mycobacterium tuberculosis* EthR for use as ethionamide boosters. *Org. Biomol. Chem.* 14, 2318–2326. doi: 10.1039/C5OB02630J
- Nishida, C. R., and Ortiz de Montellano, P. R. (2011). Bioactivation of antituberculosis thioamide and thiourea prodrugs by bacterial and mammalian flavin monooxygenases. *Chem. Biol. Interact.* 192, 21–25. doi: 10.1016/j.cbi.2010.09.015
- Peranen, J., Rikonen, M., Hyvonen, M., and Kaariainen, L. (1996). T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. *Anal. Biochem.* 236, 371–373. doi: 10.1006/abio.1996.0187
- Rebollo-Lopez, M. J., Lelievre, J., Alvarez-Gomez, D., Castro-Pichel, J., Martínez-Jiménez, F., Papadatos, G., et al. (2015). Release of 50 new, drug-like compounds and their computational target predictions for open source anti-tubercular drug discovery. *PLoS ONE* 10:e0142293. doi: 10.1371/journal.pone.0142293
- Rogers, D., and Hahn, M. (2010). Extended-connectivity fingerprints. *J. Chem. Inf. Model.* 50, 742–754. doi: 10.1021/ci100050t
- Surade, S., Ty, N., Hengrung, N., Lechartier, B., Cole, S. T., Abell, C., et al. (2014). A structure-guided fragment-based approach for the discovery of allosteric inhibitors targeting the lipophilic binding site of transcription factor EthR. *Biochem. J.* 458, 387–394. doi: 10.1042/BJ20131127

- Szardenings, K., Li, B., Ma, L., and Wu, M. (2004). Fishing for targets: novel approaches using small molecule baits. *Drug Discov. Today Technol.* 1, 9–15. doi: 10.1016/j.ddtec.2004.08.009
- Takayama, K., Wang, L., and David, H. L. (1972). Effect of isoniazid on the *in vivo* Mycolic acid synthesis, cell growth, and viability of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* 2, 29–35. doi: 10.1128/AAC.2.1.29
- Timmins, G. S., and Deretic, V. (2006). Mechanisms of action of isoniazid. *Mol. Microbiol.* 62, 1220–1227. doi: 10.1111/j.1365-2958.2006.05467.x
- Totrov, M., and Abagyan, R. (1997). Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins* 29(Suppl. 1), 215–220.
- Vilcheze, C., and Jacobs, W. R. Jr. (2014). Resistance to isoniazid and ethionamide in *Mycobacterium tuberculosis*. Genes, mutations, and causalities. *Microbiol. Spectr.* 2:MGM2-0014-2013. doi: 10.1128/microbiolspec.MGM2-0014-2013
- Villemagne, B., Flipo, M., Blondiaux, N., Crauste, C., Malaquin, S., Leroux, F., et al. (2014). Ligand efficiency driven design of new inhibitors of *Mycobacterium tuberculosis* transcriptional repressor EthR using fragment growing, merging, and linking approaches. *J. Med. Chem.* 57, 4876–4888. doi: 10.1021/jm500422b
- Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., et al. (2011). Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* 67(Pt 4), 293–302. doi: 10.1107/S0907444911007773
- Wang, C., Garcia-Contreras, L., Muttill, P., and Hickey, A. J. (2012). Isoxyl assays in plasma. *J. Pharm. Biomed. Anal.* 60, 1–6. doi: 10.1016/j.jpba.2011.10.029
- Warr, W. A. (2012). Scientific workflow systems: pipeline pilot and KNIME. *J. Comput. Aided Mol. Des.* 26, 801–804. doi: 10.1007/s10822-012-9577-7
- Willand, N., Desroses, M., Toto, P., Dirie, B., Lens, Z., Villeret, V., et al. (2010). Exploring drug target flexibility using *in situ* click chemistry: application to a mycobacterial transcriptional regulator. *ACS Chem. Biol.* 5, 1007–1013. doi: 10.1021/cb100177g
- Willand, N., Dirie, B., Carette, X., Bifani, P., Singhal, A., Desroses, M., et al. (2009). Synthetic EthR inhibitors boost antituberculous activity of ethionamide. *Nat. Med.* 15, 537–544. doi: 10.1038/nm.1950

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AM and handling Editor declared their shared affiliation.

Copyright © 2017 Mugumbate, Mendes, Blaszczyk, Sabbah, Papadatos, Lelievre, Ballell, Barros, Abell, Blundell and Overington. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.