DRUG DISCOVERY AND RESISTANCE

A whole genome bioinformatic approach to determine potential latent phase specific targets in *Mycobacterium tuberculosis*

Lucas A. Defelipe [a], Dario Fernández Do Porto [b], Pablo Ivan Pereira Ramos [c, d], Marisa Fabiana Nicolás [d], Ezequiel Sosa [b], Leandro Radusky [a], Esteban Lanzarotti [a], Adrián G. Turjanski [a, **], Marcelo A. Marti [a, *]

[a] *Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina*
[b] *Plataforma de Bioinformática Argentina, Instituto de Cálculo, Pabellón 2, Ciudad Universitaria, Facultad de Ciencias Exactas y Naturales, UBA, Buenos Aires, Argentina*
[c] *Centro de Pesquisas Gonçalo Moniz, FIOCRUZ, Bahia, Brazil*
[d] *Laboratório Nacional de Computação Científica, Petrópolis, Rio de Janeiro, Brazil*

ARTICLE INFO

SUMMARY

Current Tuberculosis treatment is long and expensive, faces the increasing burden of MDR/XDR strains and lack of effective treatment against latent form, resulting in an urgent need of new anti-TB drugs. Key to TB biology is its capacity to fight the host's RNOS mediated attack. RNOS are known to display a concentration dependent mycobactericidal activity, which leads to the following hypothesis "if we know which proteins are targeted by RNOS and kill TB, we we might be able to inhibit them with drugs resulting in a synergistic bactericidal effect". Based on this idea, we performed an *Mtb* metabolic network whole proteome analysis of potential RNOS sensitive and relevant targets which includes target druggability and essentiality criteria. Our results, available at http://tuberq.proteinq.com.ar yield new potential TB targets, like I3PS, while also providing and updated view of previous proposals becoming an important tool for researchers looking for new ways of killing TB.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

It is estimated that one-third of the world population is infected with *Mycobacterium tuberculosis* (*Mtb*), resulting in about 2 million deaths worldwide each year [1]. The hosts immune response to tuberculosis (TB) infection relies in phagocytosis of the bacilli by the macrophages leading to granuloma formation that stops bacterial replication. Inside the macrophages bacteria face stressful conditions characterized by hypoxia, inducible Nitric Oxide (NO) synthase derived NO and nutrient deprivation. The bacilli in

response switches to a non replicative (latent) state, where it can remain hidden and alive for decades [2,3]. Reactivation of latent *Mtb* is a high risk factor for disease development, particularly in immunocompromised individuals [4]. Common treatment for tuberculosis involves a long regime with the front line drugs, isoniazid, rifampicin, pyrazinamide and ethambutol. However, the emergence of multi and extensively-drug-resistants (MDR and XDR) *Mtb* strains, and the negative drug–drug interactions with certain HIV (or other diseases) treatments, reveal the urgent need for new anti-TB drugs [5]. Moreover, it is important to highlight that the existing anti-TB drugs are ineffective against dormant *Mtb*, and that there is a lack of well-defined targets specific for this state [6,5]. The concentration dependent mycobactericidal activity of Reactive Nitrogen and Oxygen Species (RNOS) [7,pp215-235]; and the recent finding showing that Bicyclic nitroimidazol like compounds PA-824 and OPC-67683 (currently in clinical trials) kill non replicating *Mtb* by means of intracellular NO release after enzymatic reduction of the nitro group, underscores the relevance of RNOS for fighting TB infections [8–11]. Based on these observations an interesting hypothesis emerges: that identification of RNOS enzymatic targets would allow us to design inhibitors against these

---

* Corresponding author.
** Corresponding author.
E-mail addresses: adrian@qi.fcen.uba.ar (A.G. Turjanski), marcelo@qi.fcen.uba.ar (M.A. Marti).

enzymes to synergize with macrophages attack and kill TB in the latent phase [12]. With this in mind, and taking advantage of the recent developments in structural bioinformatic methods, we performed an *in-silico* detailed evaluation of *Mtb* druggable proteome and studied its relation with its sensitivity to RNOS stress conditions and its relevance in the general metabolism of the bacteria to find putative therapeutics targets of latent *Mtb*.

To determine which proteins are important in the dormant phase of *Mtb* is not only relevant in relation to the search for new targets, but also as an essential knowledge related to TB biology. Therefore, several works in the last decades have looked for stress related genes, using mainly genome wide DNA microarrays in a variety of conditions, that are supposed to mimic the conditions encountered by the bacilli inside the granuloma, like hypoxia, starvation, murine infection and exposure to a variety of RNOS, among other toxic chemicals [13–23]. These works allowed the identification of key regulons, whose expression is enhanced under nitrosative stress conditions, such as the dosRST regulon [24], as well as several RNOS detoxifying enzymes and cellular repair mechanisms [22]. The corresponding lists of genes have been thus used in several *in-silico* target identification pipelines [25–27], the last, and most comprehensive in relation to focus on nitrosative stress related genes, being that performed back in 2007 by Murphy et al. [6].

Little is known concerning the potential targets of RNOS in *Mtb*. In other words which enzymes would be totally or partially inhibited by these species [12]. Given the previously mentioned observations, predicting nitrosative stress sensitivity emerges as a promising strategy to fight TB, but requires an understanding of the chemical reactivity of RNOS that should be translated in an efficient *in-silico* search method, as briefly described below. The two main sources of macrophage derived RNOS are the NADPH oxidase complex, and the inducible NO synthase, which produce respectively superoxide anion radical (O_2^{*-}) [28] and NO [29]. Superoxide spontaneously dismutates into hydrogen peroxide, which yields the potent non selective oxidizing hydroxyl radical ($^{*}OH$). While the diffusion controlled reaction between NO and O_2^{*-} yields the nitrating and oxidizing peroxinitrite anion ($OONO-$) [30,31], which can further react with carbon dioxide to yield carbonate and NO_2^{*} radicals [32]. All these species have been shown to be cytotoxic against bacteria, due to reversible or irreversible reaction with key biomolecules [33]. NO main biological targets are metal centers, such as the heme and**/Fe** FeS clusters of various enzymes involved in aerobic respiration [34,35], a reactivity which could partially explain the effect of the nitroimidazol pro-drugs [8–11] Interestingly, under aerobic conditions these compounds have also shown to disrupt mycolic acid synthesis in *Mtb*, suggesting that other mechanisms must be operative. NO also reacts with thiols, leading to S-nitrosilation of key cysteines that may impart enzyme function [12,36], and other metals center such as Zn, Cu or Mn leading to inactivation of the corresponding metalloproteins. Recently, for example, NO has been shown to inhibit several of *Mtb* p450 enzymes at physiological concentrations, through the formation of highly stable heme nitrosil adducts [37]. Concerning the other RNOS molecular targets, much less is known. Peroxynitrite, and its derivatives, however, have been shown to be potent nitrating agents, in particular targeting Tyrosine (Tyr) residues [38,39]. Nitro-tyrosine formation and its inhibitory effects are well documented for several proteins, the paradigmatic cases being the Mn Superoxide Dismutase [40,41] and the SERCA-2a [42,41,43]. Interestingly, Tyr nitration is enhanced in the presence of metal cofactors, such as heme or Mn [44]. In summary, RNOS protein reactivity can be nailed down to the presence of reactive Cys, Tyr or transition metal centers. Although determining the functional outcome of the corresponding reaction is not straight forward; in

the present work, and based on our previous experience of RNOS chemical reactivity, with its expression pattern, under different infection mimicking conditions, its role in bacterial survival, its relevance in the metabolic network and protein structural druggability as determined by our recently developed genome-wide method, for the whole *Mtb* proteome. Our results show that this combined analysis results in a powerful methodology that is able to highlight and contextualize several new potential *Mtb* drug targets, expected to be critically sensitive under nitrosative stress conditions as those encountered *in-vivo*, while integrating and expanding knowledge for previous proposed targets. Our works thus, not only represents an update of the previously proposed and known *Mtb* targets, but also adds a new perspective for *Mtb* target selection.

2. Computational methods

All the analysis performed for the present work are based either on the *in-silico* calculation of selected properties for each protein or, on the integration and meta-analysis of publicly available data. Below, we present a reference to the data sources together with a short description of the applied protocols to compute the desired properties. Further details on the data sources and methods can be found in the corresponding references. The TuberQ pipeline is available to be applied to other genomes under request to the authors.

2.1. Download of Mt protein sequences and preliminary analysis

All Open Reading Frames (ORFs), or possible proteins, from *Mtb* H37Rv derived from the complete genome sequencing [45] were downloaded from the UniProt database (www.uniprot.org, organism code 3A1773) [46]. This results in a total of 4131 ORFs. All ORFs were analyzed with the HMMER program [47] and assigned to PFAM families or domains. For all cases where an ORF matches more than one domain in PFAM, structural properties were analyzed separately as well as together.

2.2. Origin and meta-analysis of microarray expression data under different stress and infection mimicking conditions

To determine which target proteins are relevant under stress conditions we performed a combined analysis of multiple published gene expression datasets derived from microarray experiments performed under a variety of conditions that model different stress conditions such as: hypoxia, starvation, macrophage culture among others [22,13–17,19–21,48,49,18]. The complete list of conditions and the corresponding references analyzed in the present work are listed in Table S1. The set is to our knowledge the most updated and complete set studied so far, and represents an update of the analysis performed by Murphy in 2007 [6].

2.3. Essentiality criteria

For comparative purposes and the sake of completeness we included four available whole TB genome essentiality criteria. The first three genome wide essentiality studies were done by Rubin's group using a genetic technique known as TraSH (Transposon Site Hybridization) where a random insertion of this mobile genetic element is produced in order to knockout a gene. Using this technique three different studies were done: I) The first one is an *in vitro* culture study done by Sassetti and coworkers [50]. II) The library from the former work was used by Sassetti et al., 2003b [49] and defied against an C57BL/6J mouse model and a subsequent culture assay was performed to determine the relative abundance of *Mtb*

different genetic lines. From this work only 192 genes predicted (p-value < 0.05) to be essential *in vivo* were added because of its relevance. The last study performed by Rengarajan et al., [18] is a macrophage survival study using the aforementioned TraSH library developed by Rubin's group. The last study is from Sassetti's group done using a himar1 based transposon mutation system and deep sequencing to determine the frequency of insertions, this study represents an update from the studies performed in 2003 by Rubin's group [51]. Based on these studies we assigned a given gene as essential if it was reported as such in any of these studies. Our essentiality criteria only covers *in vitro* growth.

2.4. Generation of structural homology based models

As of June 2015, 467 unique crystal structures are available for *Mtb* proteins. For all remaining ORFs, we tried to build homology based models using our previously developed structural genomic pipeline [27]. Presently we were able to build a total of 2061 models, which are freely available to download at the TuberQ website (http://tuberq.proteinq.com.ar/).

2.5. Structural assessment of druggability

Druggability of each potential target was assessed by determining (and characterizing) each target putative pocket ability to bind a drug like molecule, as determined by the fpocket program [52] and our recently developed criteria [27]. Briefly, the method is based on a Voronoi tessellation algorithm that performs pocket searches and computes several pocket physico-chemical descriptors (polar and apolar surface area, hydrophobic density, hydrophobic and polarity score) that are combined into a single number called the druggability score (DS), that ranges between 0 (non-druggable) to 1 (highly druggable) [52,27]. Based on previous analysis of druggability score distribution for all pockets that are found to host a drug like compound in the PDB, in relation to other (most likely non or less druggable pockets) (See http://tuberq.proteinq.com.ar/bioflux/information/Druggabilityanalysis.html for further details), we classified pockets in four categories, non druggable (ND), with DS between 0 and 0.2, possibly druggable (PD), with DS between 0.2 and 0.5, Druggable (D) with DS between 0.5 and 0.7, and highly druggable (HD) with DS between 0.7 and 1.

Active site identification. In order to identify the active site pocket and/or determine the relevance of a given pocket to the protein function, we used two different and complementary analysis based on the data from the Catalytic Site Atlas (CSA) and a PFAM position site importance criteria, both of which have been shown to increase a pocket likelyhood of being druggable in a broad sense [27].

2.6. Structure based assessment of nitro/oxidative stress sensitivity

The RNOS sensitivity criteria, is based on a combination of structural data and chemical reactivity knowledge. As mentioned in the introduction RNOS main targets are mainly metal centers of proteins, like the Heme group, Cysteine thiols, and also Tyrosine residues that can be nitrated. In the case of metalloproteins modification of the metal center oxidation/coordination state, results in partial or total inhibition of the activity, thus a protein was predicted to be sensitive to Nitro/oxidative stress if a Fe, Zn or Cu ion was present and reported to be essential for activity. Predicting the functional outcome of Tyr/Cys modification is not straightforward but a reasonable assumption is that if the Tyr or Cys residue is part of the active site residues (or active site pocket), their chemical modification may lead to decreased activity. Thus proteins were also marked as sensitive whenever a Tyr or Cys residue was present in the predicted druggable pocket. After this initial automatic analysis, each interesting case was analyzed in detail manually looking at the available literature and protein annotation data.

2.7. Construction of the whole genome metabolic network of M. tuberculosis H37Rv

To build the *Mtb* metabolic network (Mtb-MN), we used the PathoLogic algorithm within Pathway Tools v. 18.0 [53]. PathoLogic creates a Pathway/Genome Database (PGDB) containing the predicted metabolic pathways of a given organism using as input a Genbank file with the corresponding product annotations and gene coordinates along the genome. The Genbank entry AL123456.3 downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) was used as initial input for *Mtb*-MN reconstruction. The steps involved in the reconstruction include determining gene-protein-reaction associations, which are based in either the availability of the corresponding enzyme commission (EC) number or alternatively in the gene product annotation, using a custom dictionary within Pathway Tools which links products to reactions. We enriched the annotation file using EC number data from TubercuList v. 2.6 (http://tuberculist.epfl.ch) [54]. We also searched for sequence similarity with other bacteria in order to assign missing EC numbers. After automatic reconstruction a detailed manual curation was performed, which included the deletion of false-positive predicted pathways, the inclusion of missing pathways with experimental evidence and the filling of enzymatic "holes" in existing pathways using the pathway hole filler module. The reconstructed metabolic network was exported in systems biology markup language (SBML) format for downstream analyses. Reactions involving macromolecules (such as DNA, RNA and proteins, as per the BioCyc ontology) were filtered, and only the small-molecule complement of the *Mtb*-MN was considered. The metabolic Network information was now added to the previous TuberQ web site and freely available there [27].

2.8. Metabolic network analysis

After *Mtb*-MN reconstruction, we generated a list of all compounds present in the network, and we collected their frequency as reaction participants using a Python script. Those who most frequently appeared as reaction participants are considered currency compounds (such as ATP, cofactors, water) and were disregarded from the network since they may create artificial links on the graph-based representation of the network as they are involved in many reactions which are not necessarily related. A total of 51 compounds were filtered before the transformation of the metabolic network into a reaction graph, where nodes represent reactions (i.e usually enzymes) and there is an edge between two nodes if the product of one reaction is used as substrate on the reaction that follows. Cytoscape v. 2.8.3 was used for data visualization and further *Mtb*-MN analyses [55]. Choke-point analysis was conducted in order to identify potential drug targets from the metabolic perspective. Choke-point reactions are those that either uniquely produces or consumes a given product or substrate, respectively [56]. Therefore, it is assumed that the blocking of such reactions may either lead to the accumulation of a potentially toxic metabolite in the cell or the lack of production of an essential compound; thus, choke-point reactions may function as important metabolic nodes in drug targeting. We also calculated the betweenness centrality of every node in *Mtb*-MN. The betweenness centrality of a given node v, $C_B(v)$, in the graph $G = (V, E)$, where V is a set of vertices or nodes and E a set of edges is given by: $C_B(v) = \sum_{s \neq t \neq v \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$ [57] where σ_{st} is the number of shortest paths from $s \in V$ to $t \in V$ and $\sigma_{st}(v)$ represents the number of shortest paths from s to t that some node $v \in V$ lies in.

Thus, high values of node betweenness centrality from the metabolic perspective reflects the participation of a reaction as intermediary in many other transformations, and its blockage would generate disequilibrium in many different pathways.

Scoring function. As a last step of our analysis we classified the different metabolic pathways of the *Mtb*-MN with a scoring function to determine which pathways are relevant in the latent phase of *Mtb*. Firstly, we defined a pathway as druggable if at least one of the proteins involved was druggable. After ruling out nondruggable pathways, the scoring function is defined in equation (1) in order to assign a score to each pathway:

$$SF = \frac{1}{2} * \left(\frac{NR + Cy + Chk + Es + C}{5} + \frac{S + H + St + I}{4} \right)(2) \qquad (1)$$

Where NR is the ratio between the total number of reactions present in the pathway associated with a gene and the total number of reactions present in the pathway (spontaneous and enzymatic). Cy is the ratio between the node centrality and the node with the biggest centrality. Chk is the ratio between the number of chokepoints in the pathway and the total number of reactions present in the pathway. E is the ratio between the number of essentials genes present in the pathways and the total number of genes associated to the pathways. C = "completeness" is the ratio between the total number of reactions of the pathway associated with a gene and the total number of enzymatic reactions present in the pathway. S is the ratio between the number of genes associated to stress conditions [22,21,58] present in the pathway and the total number of genes associated to the pathway. H is the ratio between the number of genes associated to hypoxia [15,48,24]and the total number of genes associated to the pathway. St is the ratio between the number of genes associated to starvation [13,14] and the total number of genes associated to the pathway. Finally, I is the ratio between the number of genes associated to infection [16,17,19,20] and the total number of genes associated to the pathway. In other words the score sums up data on the relevance of individual genes to characterize the whole path. A value close to 1 would mean that 'all genes'in the pathway are essential, relevant for stress and so on.

3. Results

Results are organized as follows, we will begin analyzing the structural druggability of TB proteome performing a broad classification of the proteins. Secondly we combined this data with a meta analysis of the expression profiles under different infection-mimicking conditions, essentiality, and RNOS sensitivity. Finally we integrated this data with a metabolic network analysis of *Mtb*. See Figure 1
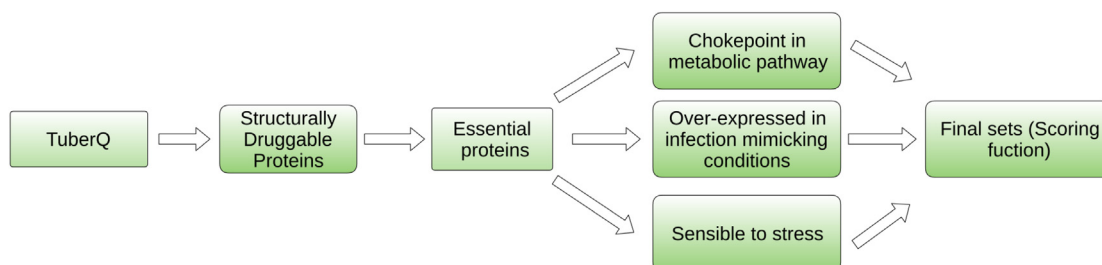
3.1. Classification of all Mtb domains structures according to their druggability and essentiality criteria

We begin our analysis by classifying all available domain structures (including those derived from X-ray and Homology Models) according to their structural druggability. For this sake we first divided the domains in four groups. The first group corresponds to the positive control group, i.e those proteins from *Mtb*, which have been crystallized in the presence of a drug-like compound or an inhibitor (we will call this the Crystallized With Drug or CWD group). The second groups is the group we call the Drug Binding Domains Group (DBD). Proteins are included in the DBD if there is at least one crystal structure, of another protein from the same PFAM domain, that was crystallized in the presence of an inhibitor or drug-like compound. Thus the DBD group includes all proteins that could be druggable. The remaining proteins correspond to all *Mtb* structures that bear no relation to any structure hosting a drug-like compound. This group is divided in X-ray structures from *Mtb* proteins and models. For all structures, we computed all the possible pockets and their corresponding Druggability Score(DS) using fpocket. Special attention was paid to the pocket hosting the drug-like compound when available. As expected this is usually the highest scoring pocket in the given protein. For the DBD group we selected the pocket that matches the drug-hosting pocket in the protein-drug complex structure of the same domain. Finally, for the remaining proteins we further analyzed the pocket that matches the active site prediction using CSA, PFAM or both. Having selected the relevant pockets, we classified each ORF in the four categories defined in methods according to their DS. The results are shown in Table 1 below.

The results show as expected that most of the*Mtb* available structures crystallized in the presence of a drug have high DS Scores. Moreover, among these proteins appear enoyl-ACP reductase InhA (P9WGR1), which is the primordial target of the first line drug for TB treatment isoniazid, as well as Hydroxymycolate synthase mmaA4 (Q79FX8) and Serine/threonine-protein kinase Pkn B (P9WI81) both proteins for which there are known inhibitors (S-adenosyl-N-decyl and Ser/Thr-mitoxantrone) that have been shown to stop mycobacterial growth [59]. Thus and in agreement with previous works [52,27] our method is able to predict to a high degree the likelihood of a protein to host a drug-like compound. From a general point of view, an interesting result is that over half of the analyzed structures (both crystals and models) are likely to bind a drug-like compound. This value is higher than the one computed using only a sequence based analysis [60] of about 21%, and possibly reflects a bias towards the determination of structures that are already known as drug targets.

The first group of interest, where novel targets can be found, concerns the DBD-HD group, the fact that both an association

Figure 1. Prioritization pipeline used in this work.

Table 1
***Mtb* proteins classified according to their druggability score**. Numbers in parentheses indicate number of essential proteins as defined in methods. CWD: Crystallized with drug, DDE: Druggable Domain Extended, RWC: Rest With Crystal RWM: Rest With Model.

Score/Group	CWD	DBD	RPWC	RPWM	Total
ND	2 (1)	8 (1)	5 (1)	11 (0)	26 (3)
PD	7 (5)	68 (32)	9 (5)	28 (15)	112 (57)
D	26 (19)	290 (120)	28 (12)	108 (40)	452 (191)
HD	50 (25)	1187 (495)	97 (37)	544 (211)	1878 (768)
Total	85 (50)	1553 (648)	139 (55)	691 (266)	2468 (1019)

criteria (assignation to DBD) and a structural criteria (the DS) match for several cases, is a strong argument for the selection of these 1187 proteins, of which 495 are reported to be essential for *in vitro* growth and thus interesting for further analysis. Also, there are about 641 proteins between crystals and models, of which 248 also are essential, which we predict are highly druggable solely from a structural point of view. Complete list of these HD proteins, are presented in SI (Tables S2 and S3). Moreover, their structures and pocket features are freely available online at TuberQ.

3.2. Prioritization of Mtb proteins according to their expression in infection-mimicking conditions

To further rank the over 600 (highly) druggable and essential proteins identified above as best possible targets, we performed an analysis of the available data concerning their expression level under infection mimicking conditions. The selected conditions, which group different reports, comprise Hypoxia, Starvation, RNOS stress and infection in mice. We first ranked all proteins according to the number of conditions where the protein was found to be over expressed, thus proteins with an expression score (ES) of 0 are not found to be over-expressed in any case and are called Non over-expressed group (NO), while proteins with ES of 4 or 3 are found over-expressed in three or four different conditions are called OMC (Overexpressed in most conditions) (This data is presented for each protein in Tables S2 and S3 in SI). Table 2 shows a global picture of the over expressed proteins.

In the DBD group data shows that there are 48 druggable proteins, of which 20 are essential, that belong to the OMC group. For example, there are proteins like the well known Redox sensor histidine kinase response regulator DevS (P9WGK3), known to be involved in RNOS sensing and signal transduction harboring a druggable kinase ATP binding domain. A more interesting case is presented by 3-methyl-2-oxobutanoate hydroxymethyltransferase (or Ketopantoate hydroxymethyltransferase KPHMT, UniProtID P9WIL7), a protein which has been shown to be relevant in both Hypoxia and Infection. Among the remaining crystallized proteins

Table 2
Number of proteins reported as over-expressed in (NO to OMC) infection-mimicking conditions. a) Numbers in parenthesis corresponds only to Druggable and Essential proteins. b) Expression Score (ES) describes the number of conditions where the protein was found to be over-expressed, ranging from NO (Non overexpressed) to OMC (Overexpressed in most conditions, protein is expressed in 4 or 3 conditions, namely Hypoxia, Starvation, RNOS stress, Infection in mice). OSC (Overexpressed in some conditions) are proteins in which reports indicate to be overexpressed in 1 or 2 different conditions. Numbers in parenthesis corresponds only to Druggable and Essential proteins. RPWC: Remaining proteins with crystal and RPWM: Proteins of unknown druggability with model.

ESb	DBD	RPWC	RPWM	Total
OMC	48 (20)	8 (4)	27 (9)	83 (33)
OSC	591 (192)	64 (20)	253 (82)	908 (294)
NO	914 (283)	67 (13)	410 (119)	1391 (415)

we find, for example, proteins like L,D-transpeptidase 2 (UniProtID I6Y9J2), the Alpha-beta hydrolase (UniProtID I6XU97) and DNApol III delta subunit (UniProtID O06363). Interestingly, almost half of the analyzed "essential" proteins are over-expressed in some conditions, an observation that possibly reflects the fact that *Mtb* protein expression is highly regulated and adaptable to variety of subtle different conditions or external stimuli, and that the up-regulated proteins perform key functions.

3.3. Incorporation of a stress-sensitivity prediction

As mentioned in the introduction, one key hypothesis to fight TB is to identify which proteins are already targeted by the RNOS attack of the immune response and try to inhibit them also pharmacologically. Therefore, and beside the expression analysis, we used sequence-structure information combined with chemical reactivity knowledge to predict possible sensitivity towards RNOS. As described above, these species main targets are metal centers of proteins, like the Heme group, Cysteine, and also Tyrosine residues that can be nitrated/oxidized. The sensitivity condition for previously identified druggable and essential proteins is also presented in Tables S2 and S3, while a global analysis is presented in Table 3 below.

The data from Table 3 show that there are over 800 proteins which are potentially sensitive to RNOS. Tyrosine presence is twice as common as Cysteine or metal ion, but it is should be noted that its predictive power as RNOS sensitive is moderate in the best case. Combination of RNOS sensitivity with previous criteria shows that there about 200 proteins that fulfill all of them and are thus ranked on top. Beyond the global analysis it is interesting to see what new proteins fill all criteria (Druggable, essential, Over-expressed and sensitive to RNOS) and thus emerge as possible targets from this analysis. One case, that caught our attention are Inositol-3-phosphate synthase (I3PS, ino1) (See Figure 2 to view the detected cavity (UniProtID P71703)). The druggable pocket harbors two oxidable tyrosines and one oxidable cysteine and matches the NAD binding site and has a considerable size.

3.4. Construction and incorporation of Mtb metabolic network analysis to prioritize protein targets

As a last step in our prioritization procedure, we used Pathway tools and manual revision to build a comprehensive *Mtb*-MN. One of our goals is to identify choke points, those reactions that either have a unique substrate or product, because genes associated to these reactions are supposed to relevant for *Mtb*. Another relevant characteristic is the centrality of the reactions, meaning the many pathways converge or emerge from them. Global characteristics of the network are presented in Table 4 and Figure 3. Overall we could assign 985 unique genes/proteins to 1369 enzymatic reactions that can be grouped in 257 pathways. Of a total of 1708 enzymatic reactions, 1305 were associated to previously reported essential genes. We failed to unequivocally assign additional 339 reactions that form part of the MN, possibly due to a lack of gene/protein

Table 3
***Mtb* proteins predicted to be sensitive to RNOS**. In a) Highly Druggable proteins b) Proteins are considered over-expressed if ES is greater than 3. Numbers in parenthesis corresponds only to druggable and essential proteins.

Feature	Crystal (models)	Essential (E)	HD[a] and E	HD,E and Over[b]
AS Metal	149	86	57	41
AS Cys	130 (164)	64 (49)	37 (37)	30 (28)
AS Tyr	269 (274)	135 (84)	82 (69)	58 (42)
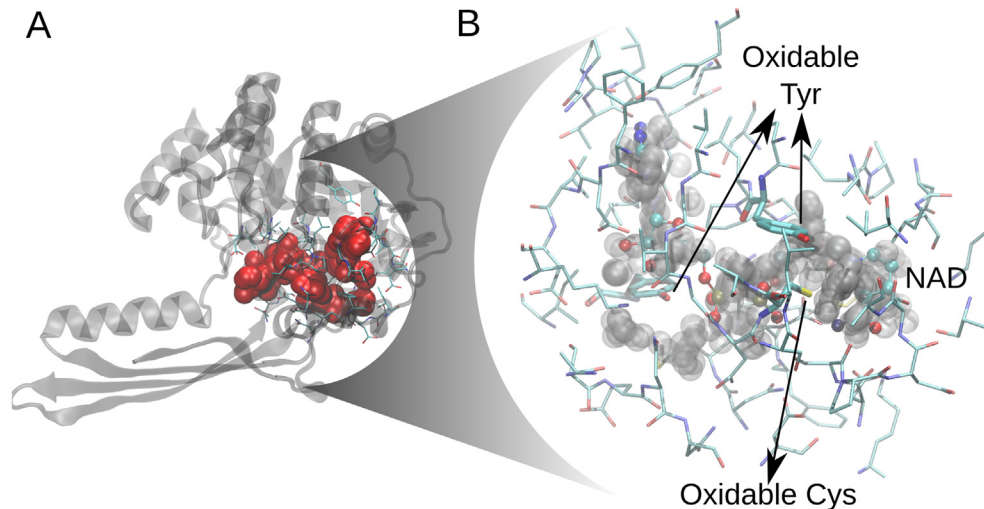
Figure 2. Inositol-3-Phosphate Synthase Structural properties A) View of the I3PS fold with the druggable pocket depicted in red spheres (PDBID 1GR0). B) Closeup view of the druggable pocket superimposed with NAD binding site. Oxidable Cysteine and Tyrosine residues of the site are drawn ticker. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

detailed characterization, the presence of spontaneous reactions and limitations of the algorithm used to reconstruct the network. We determined that 509 reactions correspond to choke points. From these, 77% are associated with essential genes, compared to 55% if all reactions are considered, thus as expected there is significant coincidence between essential genes and choke points. To prioritize genes/proteins we first decided to score each pathway according to their metabolic relevance, as determined by their number of choke points and their centrality, as well as their relevance to host infection mimicking conditions, as determined by the number of ES or over-expressed genes (see methods for details). These results are presented in SI Table S4.

MN analysis revealed several high scoring pathways, for example the responsible for mycothiol biosynthesis depicted in Figure 4. Mycothiol is crucial for the intracellular redox status regulation, and plays a pivotal role for *Mtb* survival in macrophages [61]. All the involved enzymes (Rv0046c, Rv00486, Rv1170, Rv2130c, Rv0819) are essential, and 2 of the 4 crystallized proteins (Ino1 -Rv0046c- and mshB -Rv1170-), have a DS greater than 0.7. Ino1 is also over-expressed in RNOS stress, hypoxia and starvation while mshB is over-expressed in RNOS stress conditions (Figure 4).

Another top ranking pathway is the one responsible for histidine biosynthesis, which has been suggested as harboring potential drug targets due to its absence in mammals [62]. The pathway is composed by eight essential proteins. Two of them have a structure/model and are druggable.

A further top-ranked path is the well characterized mycolate biosynthesis pathway. Mycolate is an integral cell wall component of *Mtb* that participates in survival ability of *Mtb* within infected hosts, virulence and evasion of the immune system. This pathway is targeted by first-line tuberculosis drugs such as isoniazid and ethambutol [63]. The pathway compromise 22 genes (Rv3804c Rv0470c Rv0242c Rv1483 Rv1483 Rv0242c Rv3720 Rv2524c Rv0974c Rv2247 Rv2502c Rv3280 Rv3799c Rv2524c Rv3720 Rv0636 Rv2245 Rv2246 Rv0644c Rv3372 Rv3801c Rv3800c). The importance of this pathway lies in the number of genes that are essential (83%) and druggable (60%). Also 90% of the reactions that participate in this pathway were found to be choke points. See Table 4 for further details.

MN analysis also reveals the relevance of sulfur metabolism which has been shown to be essential for the survival and virulence of many pathogens including *Mtb*. Moreover, most genes are absent

Table 4
Top pathways as classified by our scoring function. A is the number of genes involved in that particular pathway, B is the total number of reactions, C is the number of choke points and D is the centrality of the pathway.

Pathway	A	B	C	D	Best gene in pathway	Pathway importance
Alanine degradation IV	1	1	0	42,990	Rv2780	0.684747
[2Fe−2S] iron-sulfur cluster biosynthesis	3	10	9	25,887	Rv1465	0.643445
L-1-phosphatidyl-inositol biosynthesis (Mycobacteria)	2	3	3	4038	Rv0046c	0.606344
mycothiol biosynthesis	5	6	6	10,870	Rv0046c	0.537073
Alanine biosynthesis III	1	1	0	25887	Rv1464	0.530718
NAD biosynthesis I (from aspartate)	5	6	4	14,256	Rv1595	0.500894
Chorismate biosynthesis from 3-dehydroquinate	5	6	4	34,233	Rv3227	0.482852
Histidine biosynthesis	8	10	10	11,012	Rv2122c	0.479155
Homocysteine biosynthesis	2	2	1	5104	Rv3340	0.468182
Selenate reduction	2	5	5	6819	Rv1285	0.457486
Lipoate biosynthesis and incorporation I	2	2	2	17,349	Rv2217	0.452901
UDP-N-acetyl-D-glucosamine biosynthesis I	4	5	3	293,284	Rv1018c	0.452365
Mycolate biosynthesis	22	182	162	75,004	Rv2246	0.451791
Taxiphyllin bioactivation	1	1	1	0	Rv0186	0.450000
dTDP-L-rhamnose biosynthesis I	6	4	4	6857	Rv3464	0.445889
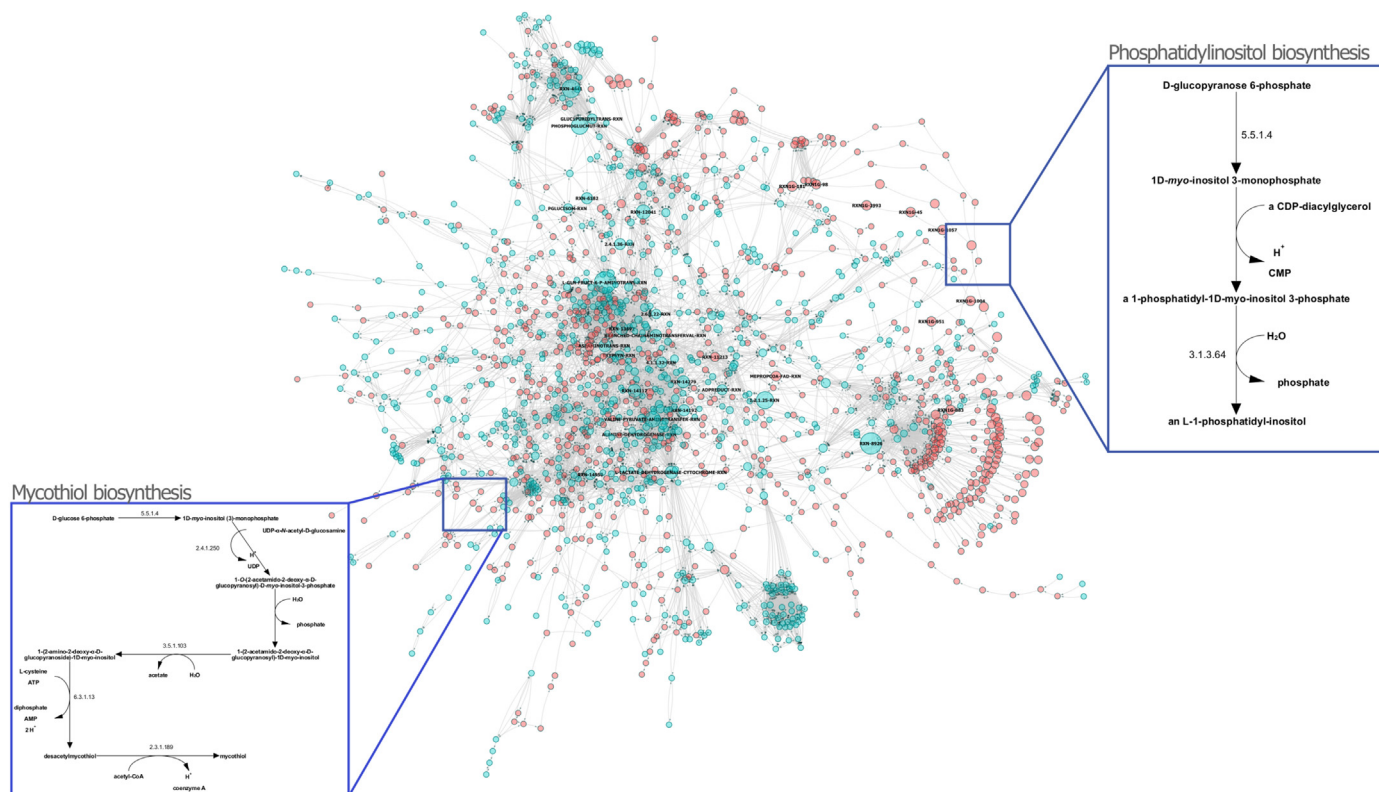
Figure 3. Reaction graph of the Mycobacterium tuberculosis metabolic network. Each node represents a predicted reaction in the *M. tuberculosis* metabolism, and there exists an edge between two nodes if the product of a given action is used as substrate for the following reaction. Node sizes were drawn proportional to their betweenness centrality on the reaction graph, and top-ranked reactions according to this metric are shown. Nodes in red represent choke points. Detailed processing of two pathways are highlighted, namely mycothiol biosynthesis and phosphatidyl-inositol biosynthesis, both having high scores according to our analyses. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in humans. Among these pathways, methionine metabolism, including degradation to homocysteine, performed by the drug-gable protein Rv3340 (overexpressed in most conditions) and Rv3341. Particularly interesting among the involved proteins is also sahH (Rv3248c), whose role in the regulation of ʟ-homocysteine has been recently elucidated [64]. Moreover, it has also been implicated as intermediate in *Mtb* resistance mechanisms, along with MetK (Rv1392) [65]. SahH catalyzes the hydrolysis reaction of SAH to homocysteine and adenosine with NAD + acting as a co-factor. The protein has a alpha/beta fold composed of a Ross-mann motif that binds NAD and SAH (as this fold is characteristic of nucleoside binding proteins) [66]. The druggable pocket is delimited mainly by polar and negatively charged residues and has a volume of 2284 Å^3, rather big as it has to accommodate both the ligand and the cofactor. It has 2 tyrosines (Tyr 493 and Tyr 495) in the binding site of NAD making it potentially sensible to RNOS. SahH has been determined to be essential in both high throughput knockout studies performed [50,51].

Among other top-scoring pathways revealed by our MN analysis are those related to lipoate synthesis. The two key genes (Rv2218 [lipA] and Rv2217 [lipB]) are essential and lipB was also found to be druggable and expressed under starvation condition. Moreover, the druggable pocket of LipB has Cys 176, Tyr22 and Tyr 91 making the pocket potentially sensible to RNOS. Although not an ubiquitous process in bacteria, lipoate has been implicated with a role in microbial pathogenesis including immune response-induced oxidative and nitrosative stress in mycobacteria, and it has also been acknowledged that lipoylated proteins take part in important antioxidants processes [67–69], thus the attractiveness of this

high-scoring pathway from the target-finding aspect. Moreover, LipB has been structurally characterized and shown to have promising therapeutic properties [70].

Finally, another important pathway revealed by our MN analysis is "alanine degradation IV" performed by an unique gene (Rv2780). Although it was not reported as essential, this ʟ-alanine dehydro-genase is the first antigen reported to be present in the pathogenic *Mtb*, but not in the vaccine strain *Mycobacterium bovis* BCG [71]. Moreover, it was suggested that the lack of a functional alanine dehydrogenase impairs BCG replication in humans [72]. This pro-tein is druggable and appears to have a key function in nitrosative stress since it is over-expressed in all the studied conditions. In summary, our integrated analysis of *Mtb* metabolism along with expression, essentiality and druggability data has allowed the identification of important pathways harboring promising thera-peutic targets. All this information is freely and easy accessible at the TuberQ website, and presents a direct route for a fast evaluation of a given protein potentiality as drug target to fight TB.

4. Discussion

4.1. What's novel in TB targets?

Given the potential of whole genome target identification approaches, since the deciphering of the whole *Mtb* genome [45], several *in silico* based works have appeared on the subject [73–75,25,76]. They usually included some druggability prediction, a role in dormancy based on previous gene expression data, an essentiality and/or an off-target criteria to prioritize the protein

Mycothiol Biosynthesis
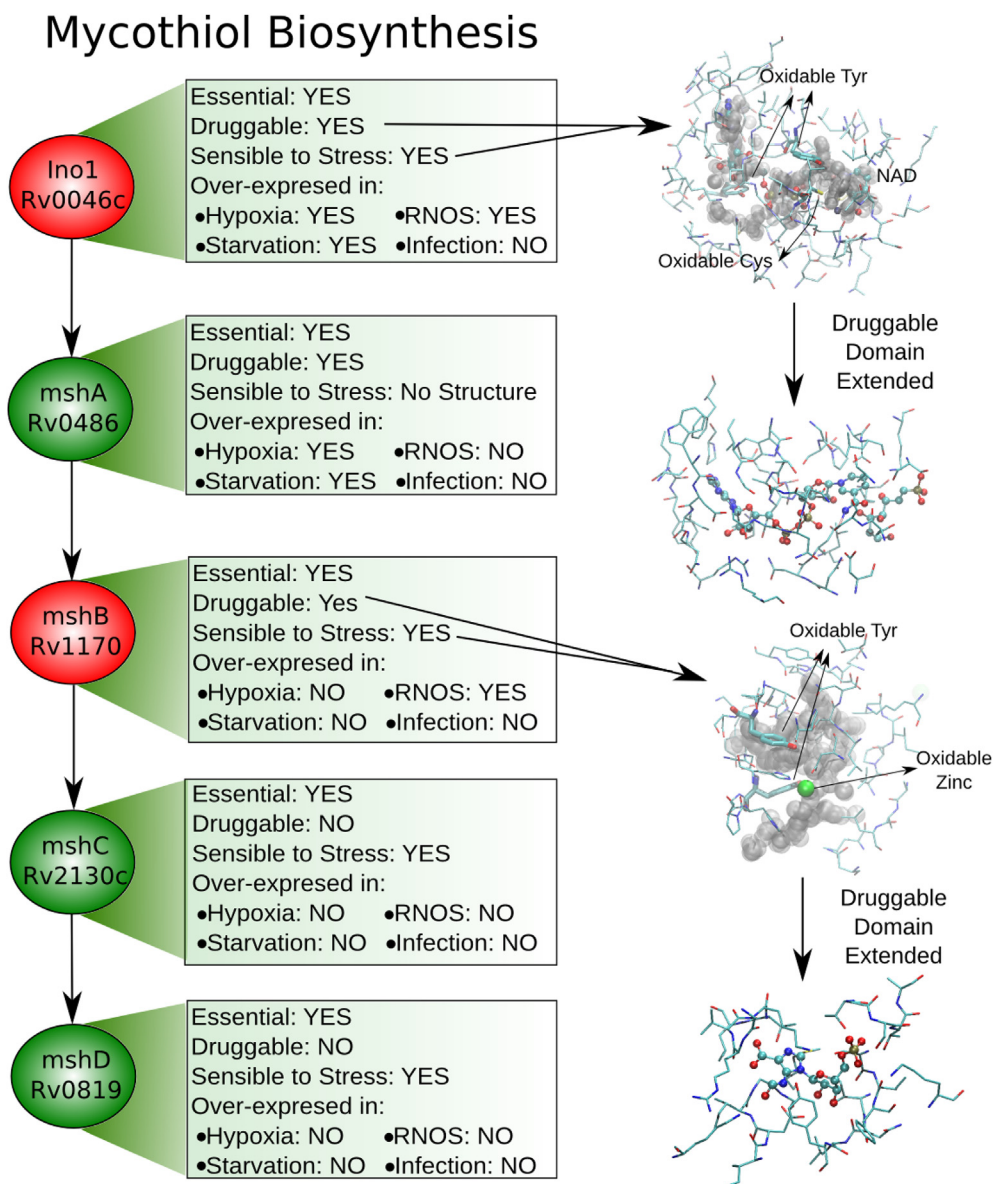
Figure 4. Representation of the Mycothiol biosynthesis pathway with two outstanding proteins pockets depicted with their corresponding DBD partner pockets shown. The most important proteins in this pathway Ino1 and mshB are shown in red. A representation of the pocket and of the known drugged pocket of the PFAM Family (DBD) are shown next to their respective protein. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

targets. Essentiality criteria [51,50] were based mainly on experimental mutagenesis studies which have the drawback that, due to redundant and compensatory functions, result in only 10—15% of genes being individually essential [77—80], but many more can be synthetically lethal when knocked down in combination [81]. In this context, we prioritized a target druggability and stress sensitivity over its essentiality, and look for its role in the context of *Mtb* metabolism, by highlighting whole pathways to be targeted instead of single proteins. Concerning the off-target criteria it is important to note that a desirable anti-TB drug should be specific for the bacteria and do not interfere with hosts proteins. This usually translates in either a sequence [73] or structural comparison [65] between potential targets and the host proteins, and those targets too similar to host proteins are discarded. The difficulty with this approach is that even a single aminoacid substitution may result in differential binding, as shown by the emergence of antibiotic resistance due to this type of mutations, and the selectivity of many

drugs [77,82]. A paradigmatic case are the diarylquinolines that target *Mtb* membrane bound subunit F0 of the ATP synthase (present in all organisms) that shows nonetheless a very narrow spectrum of activity, loosing potency even against other actinobacteria, being inactive against both gram positive and negative bacteria, and showing over 20 thousand fold selectivity versus mammals [83,84]. Therefore, we believe that off-target criteria should be very stringent and can be even considered in later stages of pharmacological development, usually in the context of structural ligand binding studies. In any case it is interesting to compare present results to previous target selection pipelines, to see what are common as well as new identified targets.

Among our top scoring targets, many proteins broadly mentioned in the literature as signaling proteins (pnkB, pnkG, DevS) or related to mycolic acid synthesis (IhnA, pcaA, pks13, fas, fad32D), panthenoate biosynthesis (panB) and Cytochromes (cyp 121 and cyp125) have already been highlighted by Chandra's Lab

Table 5
Newly and revalidated *Mtb* targets found using structural druggability, metabolic importance analysis and expression data in infection mimicking conditions. Revalidated targets are taken from [25].

Protein name	Rv	Status	Druggability	Pathway (importance)	Profile expression
Inositol-3-phosphate synthase	Rv0046c	New target	0.719	Myo-inositol biosynthesis (0.3871) L-1-phosphatidyl-inositol biosynthesis (Mycobacteria) (0.6063). mycothiol biosynthesis (0.5370)	Str, Hyp, Sta
3-phosphoshikimate 1-carboxyvinyltransferase	Rv3227	New target	724	Chorismate biosynthesis from 3-dehydroquinate (0.4828)	Str
O-acetylhomoserine aminocarboxypropyltransferase	Rv3340	New target	635	Homocysteine biosynthesis (0.4681)	Hyp, Sta, Inf
3-oxoacyl-[acyl-carrier-protein] synthase 2	Rv2246	New target	709	Mycolate biosynthesis (0.4517) fatty acid biosynthesis initiation II (0.3883) 8-amino-7-oxononanoate biosynthesis I (0.3765)	Str, Hyp
Octanoyltransferase	Rv2217	New target	703	Lipoate biosynthesis and incorporation I (0.4529)	Sta
Bifunctional protein GlmU	Rv1018c	New target	911	UDP-N-acetyl-D-glucosamine biosynthesis I (0.4523)	Str, Hyp, Inf
Rv1465	Rv1465	New target	926	[2Fe–2S] iron-sulfur cluster biosynthesis	Str
1D-myo-inositol 2-acetamido-2-deoxy-alpha-D-glucopyranoside deacetylase	RV1170	Revalidated	781	Mycothiol biosynthesis (0.5370)	Str
Sulfate adenylyltransferase subunit 2	Rv1285	Revalidated	891	Selenate reduction (0.4579) sulfate activation for sulfonation (0.4326)	Str, Hyp, Sta
dTDP-glucose 4,6-dehydratase	Rv3464	Revalidated	676	dTDP-L-rhamnose biosynthesis I (0.4459)	Str
Enoyl-[acyl-carrier-protein] reductase [NADH]	Rv1484	Revalidated	919	8-amino-7-oxononanoate biosynthesis I (0.3765) stearate biosynthesis II (bacteria and plants) (0.3700)	
3-methyl-2-oxobutanoate hydroxymethyltransferase	Rv2225	Revalidated	937	Phosphopantothenate biosynthesis I (0.4351)	Str, Hyp, Inf
Mycocyclosin synthase	Rv2276	Revalidated	887	Mycocyclosin biosynthesis (0.4435)	Hyp

and TDR Targets [75,25,85] as potential drug targets (shown in Table 4). Thus, as expected, there are some common ussual findings. Interestingly, none of these reports point to our top scoring proteins in the mycothiol synthesis pathway (like mshB and ino1) relevant for the redox balance in mycobacteria and suggested as a relevent metabolic pathway to kill the pathogen (See Figure 4 Inositol-3-phosphate synthase (I3PS, ino1) (UniProtID P71703) is a member of this mycothiol pathway). I3PS is the protein that converts Glucose-6-P to 1D-myo-inositol 3-phosphate, it has been reported as essential in both high throughput [50,51] as well as single mutant studies [86]. It is part of the DosR regulon and is also highly over-expressed under starving conditions. As shown in Figure 2 I3PS structure (pdbid 1GR0) presents a druggable pocket (DS score of 0.719) which overlaps with the NAD binding site, a site also known to be able to host drug like compounds in other proteins like InhA. Interestingly it displays 2 RNOS sensitive residues Tyr145 and Cys26 as well as a "structural/catalytic Zn" whose role has not been completely elucidated. Clearly I3PS harbors all characteristics of an ideal target.

Another, not previously noted case is exemplified by L-D transpeptidase (LDTP1) which is involved in the cross-linking of peptidoglycan in the mycobacterial cell wall (and thus related to the Mycolic acid biosynthesis) fundamental for *in vivo* resistance. LDTP1 is an enzyme involved in formation of peptidoglican cross-linking bonds, and thus essential. It is over-expressed in all conditions, particularly in the presence RNOS. The druggable pocket (DS score of 0.701) which is also the active site holds the key Cys 226 which is the nucleophile in the enzymatic reaction, strongly arguing in favor of its potential inhibition by RNOS. Interestingly, it has also been suggested that LDTP1 catalytic activity could be inhibited by Beta lactamic compounds [87,88].

Nevertheless, there are targets predicted by TargetTB which are not predicted by our classification pipeline. In general, this is due to the fact that our pipeline prioritizes structural druggability and infection mimicking conditions in order to classify potential drug targets for the latent phase of Mtb infection. For example, both DdlA (Rv2981c) and EmbA(Rv3794), known targets of cycloserine [89] and ethambutol [90], respectively are druggable and essential but lack any upregulation during infection mimicking conditions making them less attractive as latent phase targets. On the other hand, two targets that belongs to the

mycolyl-arabinogalactan-peptidoglycan biosynthesis pathway, AftA (Rv3792) and AftB (Rv3805c), were not predicted for us owing to lack of structures or models available.

Finally, from a metabolic network point of view, again several known targets appear such as mycolate biosynthesis, important for *in vivo* immune system modulation. Moreover, pathways involved in some way to RNOS response are top ranked and pointed to be relevant like NAD biosynthesis, homocysteine biosynthesis or Iron–Sulfur cluster biosynthesis. Summarizing Table 5 presents a set of new 7 targets identified by our pipelines, as well as a reassessment of 6 previously identified targets, with their computed characteristics.

4.2. The potential of TuberQ as a target-finding and characterization tool

A final important contribution of the present work, is that all the data, including newly predicted RNOS sensitivity and Metabolic Network analysis is freely available on TuberQ web site (http://tuberq.proteinq.com.ar/). We expect that the TuberQ database becomes a useful resource for all researchers working in the field of drug development in TB. The main feature of TuberQ is to provide wide druggability analysis of *Mtb* proteins in a consistent and effective manner, contributing to a better selection of potential drug targets for screening campaigns and the analysis of targets for structure-based drug design projects. On the other hand, we hope TuberQ can be used as a repository where a researcher can easily access information for a specific protein of interest. For example, if the Rv3227 (3-phosphoshikimate 1-arboxyvinyltransferase) is used as a query, seven records are shown in the start page (see Figure 5 for further information), corresponding to different crystals. By clicking, in one of them, the corresponding record is expanded. On the left side of the screen, four main tabs can be displayed. The Initials-Tab, expands the identificatory data of the protein, such as the Uniprot ID, protein "common" name, PFAM domain and the PDB ID or homology based model ID. By analyzing the biological functions of the protein, we find it is essential and over-expressed under stress conditions. Also, in this page, the best BLAST hit against the human genome is shown. In our example, it corresponds to a protein with very low identity. Selecting the tab with the structure name will open the structure-related data, including

Search results

P9WPY5 - AROA_MYCTU 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE

Initials	SubSeq	Hmm Evalue	Domain	Blast Evalue	Structure	Druggability	Details
	[8-417]	1.6E-120	PF00275.15	0.0	2BJB	0.724	See Details
	[8-417]	1.6E-120	PF00275.15	0.0	2O0X	0.696	See Details
	[8-417]	1.6E-120	PF00275.15	0.0	2O15	0.526	See Details
	[8-417]	1.6E-120	PF00275.15	0.0	2O0B	0.494	See Details
	[8-417]	1.6E-120	PF00275.15	0.0	2O0E	0.432	See Details
	[8-417]	1.6E-120	PF00275.15	0.0	2O0Z	0.34	See Details
	[8-417]	1.6E-120	PF00275.15	0.0	2O0D	0.136	See Details

Structure tab

Initials tab

Expression and annotation data tab

Reaction pathways tab

Reactions grouped by pathway:

- chorismate biosynthesis from 3-dehydroquinate (PWY-6163)
  - 2_5_1_19-RXN
    - substrates
      - phospho-enol-pyruvate: Rv0211 Rv1617 Rv1023 Rv1127c
      - shikimate-5p: Rv2539c
    - products
      - 3-enolpyruvyl-shikimate-5p: Rv2540c

Figure 5. TuberQ as a target selection tool. The different information tabs that are available when searching for proteins in TuberQ, are shown. In the initials Tab, information about the assignment to the corresponding crystal structure or model and the PFAM family are depicted with the corresponding links to UniProt, PFAM and PDB (when available). The Structure tab allows the user to view and analyze the druggable pockets and any available drug or cofactor in the PDB or models. The metadata tab contains information about function annotation taken from UniProt and expression data in various infection mimicking conditions. Finally, the Reactions tab shows assignment(s) of the protein to various metabolic pathways.

the interactive pocket visualization module. This visualization allows to analyze the druggable pocket in the context of the protein like residues that delimit a given pocket, the pocket volume as defined by alpha spheres, the presence of drug-like compounds or cofactors (such as metals or NAD) which can also be co-visualized in conjunction with the detected druggable pockets. The third tab contains the metadata, which corresponds to information derived from other databases and bibliography. With this information, we may assume this protein plays a role in latency, since it is overexpressed in conditions of oxidative stress and is up-regulated in the dormant phase. The fourth and last tab (Reactions) allows to see the protein in its metabolism context. In this case we easily identify this protein as being a component of the chorismate metabolic pathway. Overall, TuberQ allows obtaining relevant information of a protein of interest in a user-friendly way. In the case of 3-phosphoshikimate 1-carboxyvinyltransferase we may select it to further study, since we found that is essential and druggable, it appears overexpressed under different conditions that mimic infections and does not show significant similarity with human sequences. A final remark is that TuberQ framework allows to upload new data into the resource so new analysis can be done,

once the information is obtained. Third party users are encouraged to provide feedback and request to the authors to upload new gene/protein properties that they consider useful for the resource.

4.3. Conclusions

Combining structural druggability, essentiality and nitro-oxidative stress sensitivity criteria, in the context of *Mtb* metabolic network and expression data, we built an *in-silico* target prioritization pipeline that allowed identification of new potential stress sensitive TB targets, shown in Table 5. The identified proteins in the phosphatidyl-inositol and mycothiol biosynthesis pathways, thus represent best potential candidates to display synergical mycobactericidal effect when pharmacologically inhibited in combination with nitro-oxidative stress generating compounds or as those encountered by the bacteria in the macrophage. We also made all performed analysis and gathered data freely available in the TuberQ web site to allow other researches to analyze any *Mtb* protein of their interest.

Acknowledgments

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.tube.2015.11.009.

References

[1] WHO. Global tuberculosis report 2014. 2014. p. 171. http://apps.who.int/iris/handle/10665/91355.

[2] Wayne LG, Sohaskey CD. Nonreplicating persistence of mycobacterium tuberculosis. Annu Rev Microbiol 2001;55:139—63.

[3] Lillebaek T, Dirksen A, Vynnycky E, Baess I, Thomsen VO, Andersen AB. Stability of DNA patterns and evidence of Mycobacterium tuberculosis reactivation occurring decades after the initial infection. J Infect Dis 2003;188(7):1032—9.

[4] Barry CE, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, Schnappinger D, Wilkinson RJ, Young D. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. Nat Rev Microbiol 2009;7(12):845—55. http://dx.doi.org/10.1038/nrmicro2236. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4144869\&tool=pmcentrez\&rendertype=abstract.

[5] Koul A, Arnoult E, Lounis N, Guillemont J, Andries K. The challenge of new drug discovery for tuberculosis. Nature 2011;469(7331):483—90.

[6] Murphy DJ, Brown JR. Identification of gene targets against dormant phase Mycobacterium tuberculosis infections. BMC Infect Dis 2007;7:84. http://dx.doi.org/10.1186/1471-2334-7-84. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1950094&tool=pmcentrez/&rendertype=abstract.

[7] Nathan C, Ehrt S. Nitric oxide in tuberculosis. In: Rom W, Garay S, editors. Tuberculosis. 2nd ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2004.

[8] Matsumoto M, Hashizume H, Tomishige T, Kawasaki M, Tsubouchi H, Sasaki H, Shimokawa Y, Komatsu M. Opc-67683, a nitro-dihydro-imidazooxazole derivative with promising action against tuberculosis in vitro and in mice. PLoS Med 2006;3(11):e466.

[9] Singh R, Manjunatha U, Boshoff HI, Ha YH, Niyomrattanakit P, Ledwidge R, Dowd CS, Lee IY, Kim P, Zhang L, et al. Pa-824 kills nonreplicating mycobacterium tuberculosis by intracellular no release. Science 2008;322(5906):1392—5.

[10] Cellitti SE, Shaffer J, Jones DH, Mukherjee T, Gurumurthy M, Bursulaya B, Boshoff HI, Choi I, Nayyar A, Lee YS, et al. Structure of ddn, the deazaflavin-dependent nitroreductase from Mycobacterium tuberculosis involved in bio-reductive activation of pa-824. Structure 2012;20(1):101—12.

[11] Manjunatha U, Boshoff HI, Barry III CE. The mechanism of action of pa-824. Commun. Integr Biol 2009;2:215—4.

[12] Rhee KY, Erdjument-Bromage H, Tempst P, Nathan CF. S-nitroso proteome of mycobacterium tuberculosis: enzymes of intermediary metabolism and antioxidant defense. Proc Natl Acad Sci U. S. A 2005;102(2):467—72.

[13] Betts JC, Lukey PT, Robb LC, McAdam RA, Duncan K. Evaluation of a nutrient starvation model of mycobacterium tuberculosis persistence by gene and protein expression profiling. Mol Microbiol 2002;43(3):717—31.

[14] Hampshire T, Soneji S, Bacon J, James BW, Hinds J, Laing K, Stabler RA, Marsh PD, Butcher PD. Stationary phase gene expression of Mycobacterium tuberculosis following a progressive nutrient depletion: a model for persistent organisms? Tuberculosis 2004;84(3):228—38.

[15] Muttucumaru D, Roberts G, Hinds J, Stabler RA, Parish T. Gene expression profile of Mycobacterium tuberculosis in a non-replicating state. Tuberculosis 2004;84(3):239—46.

[16] Karakousis PC, Yoshimatsu T, Lamichhane G, Woolwine SC, Nuermberger EL, Grosset J, Bishai WR. Dormancy phenotype displayed by extracellular mycobacterium tuberculosis within artificial granulomas in mice. J Exp Med 2004;200(5):647—57.

[17] Ohno H, Zhu G, Mohan VP, Chu D, Kohno S, Jacobs WR, Chan J. The effects of reactive nitrogen intermediates on gene expression in mycobacterium tuberculosis. Cell Microbiol 2003;5(9):637—48.

[18] Rengarajan J, Bloom BR, Rubin EJ. Genome-wide requirements for mycobacterium tuberculosis adaptation and survival in macrophages. Proc Natl Acad Sci U. S. A 2005;102(23):8327—32.

[19] Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, et al. Transcriptional adaptation of mycobacterium tuberculosis within macrophages insights into the phagosomal environment. J Exp Med 2003;198(5):693—704.

[20] Talaat AM, Lyons R, Howard ST, Johnston SA. The temporal expression profile of mycobacterium tuberculosis infection in mice. Proc Natl Acad Sci U. S. A 2004;101(13):4602—7.

[21] Voskuil MI, Schnappinger D, Visconti KC, Harrell MI, Dolganov GM, Sherman DR, Schoolnik GK. Inhibition of respiration by nitric oxide induces a mycobacterium tuberculosis dormancy program. J Exp Med 2003;198(5):705—13.

[22] Voskuil MI, Bartek IL, Visconti K, Schoolnik GK. The response of mycobacterium tuberculosis to reactive oxygen and nitrogen species. Front Microbiol 2.

[23] Robinson JL, Adolfsen KJ, Brynildsen MP. Deciphering nitric oxide stress in bacteria with quantitative modeling. Curr Opin Microbiol 2014;19:16—24.

[24] Park H-D, Guinn KM, Harrell MI, Liao R, Voskuil MI, Tompa M, Schoolnik GK, Sherman DR. Rv3133c/dosr is a transcription factor that mediates the hypoxic response of mycobacterium tuberculosis. Mol Microbiol 2003;48(3):833—43.

[25] Raman K, Yeturu K, Chandra N. targettb: a target identification pipeline for mycobacterium tuberculosis through an interactome, reactome and genome-scale structural analysis. BMC Syst Biol 2008;2(1):109.

[26] Kinnings SL, Xie L, Fung KH, Jackson RM, Xie L, Bourne PE. The mycobacterium tuberculosis drugome and its polypharmacological implications. PLoS Comput Biol 2010;6(11):e1000976.

[27] Radusky L, Defelipe LA, Lanzarotti E, Luque J, Barril X, Marti MA, Turjanski AG. Tuberq: a mycobacterium tuberculosis protein druggability database. Database 2014;2014:bau035.

[28] Groemping Y, Rittinger K. Activation and assembly of the nadph oxidase: a structural perspective. Biochem J 2005;386:401—16.

[29] MacMicking J, Xie Q-w, Nathan C. Nitric oxide and macrophage function. Annu Rev Immunol 1997;15(1):323—50.

[30] Ferrer-Sueta G, Radi R. Chemical biology of peroxynitrite: kinetics, diffusion, and radicals. ACS Chem Biol 2009;4(3):161—77.

[31] Alvarez MN, Peluffo G, Piacenza L, Radi R. Intraphagosomal peroxynitrite as a macrophage-derived cytotoxin against internalized trypanosoma cruzi consequences for oxidative killing and role of microbial peroxiredoxins in infectivity. J Biol Chem 2011;286(8):6627—40.

[32] Denicola A, Freeman BA, Trujillo M, Radi R. Peroxynitrite reaction with carbon dioxide/bicarbonate: kinetics and influence on peroxynitrite-mediated oxidations. Arch Biochem Biophys 1996;333(1):49—58.

[33] Hugo M, Radi R, Trujillo M. In: cardona p-j, editor. Thiol dependent peroxidases in mycobacterium tuberculosis in understanding tuberculosis: deciphering the secret life of the bacilli; 2012.

[34] Nathan C, et al. Specificity of a third kind: reactive oxygen and nitrogen intermediates in cell signaling. J Clin Investig 2003;111(6):769—78.

[35] Ignarro LJ. Nitric oxide: a unique endogenous signaling molecule in vascular biology (nobel lecture). Angew Chem Int Ed 1999;38(13—14):1882—92.

[36] Hess DT, Matsumoto A, Kim S-O, Marshall HE, Stamler JS. Protein s-nitrosylation: purview and parameters. Nat Rev Mol Cell Biol 2005;6(2):150—66.

[37] Ouellet H, Lang J, Couture M, Ortiz de Montellano PR. Reaction of mycobacterium tuberculosis cytochrome p450 enzymes with nitric oxide†. Biochemistry 2009;48(5):863—72.

[38] Radi R. Nitric oxide, oxidants, and protein tyrosine nitration. Proc Natl Acad Sci 2004;101(12):4003—8.

[39] Souza JM, Peluffo G, Radi R. Protein tyrosine nitration—functional alteration or just a biomarker? Free Radic Biol Med 2008;45(4):357—66.

[40] Yamakura F, Taka H, Fujimura T, Murayama K. Inactivation of human manganese-superoxide dismutase by peroxynitrite is caused by exclusive nitration of tyrosine 34 to 3-nitrotyrosine. J Biol Chem 1998;273(23):14085—9.

[41] Xu S, Ying J, Jiang B, Guo W, Adachi T, Sharov V, Lazar H, Menzoian J, Knyushko TV, Bigelow D, et al. Detection of sequence-specific tyrosine nitration of manganese sod and serca in cardiovascular disease and aging. Am J Physiol-Heart Circ Physiol 2006;290(6):H2220—7.

[42] Viner R, Ferrington D, Williams T, Bigelow D, Schoneich C. Protein modification during biological aging: selective tyrosine nitration of the serca2a isoform of the sarcoplasmic reticulum ca2+-atpase in skeletal muscle. Biochem J 1999;340:657—69.

[43] Knyushko TV, Sharov VS, Williams TD, Schöneich C, Bigelow DJ. 3-nitrotyrosine modification of serca2a in the aging heart: a distinct signature of the cellular redox environment. Biochemistry 2005;44(39):13071—81.

[44] Schopfer MP, Mondal B, Lee D-H, Sarjeant AA, Karlin KD. Heme/o2/• no nitric oxide dioxygenase (nod) reactivity: phenolic nitration via a putative heme-peroxynitrite intermediate. J Am Chem Soc 2009;131(32):11304—5.

[45] Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon S, Eiglmeier K, Gas S, Barry C r, et al. Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence. Nature 1998;393(6685):537—44.

[46] U. Consortium, et al. The universal protein resource (uniprot). Nucleic acids Res 2008;36(Suppl. 1):D190—5.

[47] Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, et al. The pfam protein families database. Nucleic Acids Res 2004;32(Suppl. 1):D138—41.

[48] Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Schoolnik GK. Regulation of the mycobacterium tuberculosis hypoxic response gene encoding α-crystallin. Proc Natl Acad Sci 2001;98(13):7534—9.

[49] Sassetti CM, Rubin EJ. Genetic requirements for mycobacterial survival during infection. Proc Natl Acad Sci 2003;100(22):12989—94.

[50] Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol 2003;48(1):77—84.

[51] Griffin JE, Gawronski JD, DeJesus MA, Ioerger TR, Akerley BJ, Sassetti CM. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. PLoS Pathog 2011;7(9):e1002251.

[52] Schmidtke P, Barril X. Understanding and predicting druggability. a high-throughput method for detection of drug binding sites. J Med Chem 2010;53(15):5858—67.

[53] Karp PD, Paley S, Romero P. The pathway tools software. Bioinformatics 2002;18(Suppl. 1):S225—32.

[54] Lew JM, Kapopoulou A, Jones LM, Cole ST. Tuberculist—10 years after. Tuberculosis 2011;91(1):1—7.

[55] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498—504.

[56] Yeh I, Hanekamp T, Tsoka S, Karp PD, Altman RB. Computational analysis of plasmodium falciparum metabolism: organizing genomic information to facilitate drug discovery. Genome Res 2004;14(5):917—24.

[57] Brandes U. A faster algorithm for betweenness centrality*. J Math Sociol 2001;25(2):163—77.

[58] Voskuil MI, Visconti K, Schoolnik G. *Mycobacterium tuberculosis* gene expression during adaptation to stationary phase and low-oxygen dormancy. Tuberculosis 2004;84(3):218—27.

[59] Wehenkel A, Fernandez P, Bellinzoni M, Catherinot V, Barilone N, Labesse G, Jackson M, Alzari PM. The structure of pknb in complex with mitoxantrone, an atp-competitive inhibitor, suggests a mode of protein kinase regulation in mycobacteria. FEBS Lett 2006;580(13):3018—22.

[60] Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov 2002;1(9):727—30.

[61] Newton GL, Fahey RC. Mycothiol biochemistry. Arch Microbiol 2002;178(6):388—94.

[62] Lunardi J, Nunes ES, Bizarro CV, Augusto Basso LA, Santiago Santos D, Machado P, et al. Targeting the histidine pathway in mycobacterium tuberculosis. Curr Top Med Chem 2013;13(22):2866—84.

[63] Barry CE, Crick DC, McNeil MR. Targeting the formation of the cell wall core of m. tuberculosis. Infect Disord-Drug Targets (Former Curr Drug Targets-Infect Disord 2007;7(2):182—202.

[64] Singhal A, Arora G, Sajid A, Maji A, Bhat A, Virmani R, Upadhyay S, Nandicoori VK, Sengupta S, Singh Y, Regulation of homocysteine metabolism by mycobacterium tuberculosis s-adenosylhomocysteine hydrolase, Sci Rep 3.

[65] Raman K, Chandra N. Mycobacterium tuberculosis interactome analysis unravels potential pathways to drug resistance. BMC Microbiol 2008;8(1):234.

[66] Rao ST, Rossmann MG. Comparison of super-secondary structures in proteins. J Mol Biol 1973;76(2):241—56.

[67] Allary M, Lu JZ, Zhu L, Prigge ST. Scavenging of the cofactor lipoate is essential for the survival of the malaria parasite plasmodium falciparum. Mol Microbiol 2007;63(5):1331—44.

[68] Bryk R, Lima C, Erdjument-Bromage H, Tempst P, Nathan C. Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein. Science 2002;295(5557):1073—7.

[69] Spalding MD, Prigge ST. Lipoic acid metabolism in microbial pathogens. Microbiol Mol Biol Rev 2010;74(2):200—28.

[70] Ma Q, Zhao X, Eddine AN, Geerlof A, Li X, Cronan JE, Kaufmann SH, Wilmanns M. The mycobacterium tuberculosis lipb enzyme functions as a cysteine/lysine dyad acyltransferase. Proc Natl Acad Sci 2006;103(23):8662—7.

[71] Chan K, Knaak T, Satkamp L, Humbert O, Falkow S, Ramakrishnan L. Complex pattern of mycobacterium marinum gene expression during long-term granulomatous infection. Proc Natl Acad Sci 2002;99(6):3920—5.

[72] Scandurra GM, Ryan AA, Pinto R, Britton WJ, Triccas JA. Contribution of l-alanine dehydrogenase to in vivo persistence and protective efficacy of the bcg vaccine. Microbiol Immunol 2006;50(10):805—10.

[73] Hasan S, Daugelat S, Rao PS, Schreiber M. Prioritizing genomic drug targets in pathogens: application to mycobacterium tuberculosis. PLoS Comput Biol 2006;2(6):e61.

[74] Raman K, Rajagopalan P, Chandra N. Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. PLoS Comput Biol 2005;1(5):e46.

[75] Agüero F, Al-Lazikani B, Aslett M, Berriman M, Buckner FS, Campbell RK, Carmona S, Carruthers IM, Chan AE, Chen F, et al. Genomic-scale prioritization of drug targets: the tdr targets database. Nat Rev Drug Discov 2008;7(11):900—7.

[76] Jamshidi N, Palsson BØ. Investigating the metabolic capabilities of mycobacterium tuberculosis h37rv using the in silico strain inj661 and proposing alternative drug targets. BMC Syst Biol 2007;1(1):26.

[77] Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol 2008;4(11):682—90.

[78] Zambrowicz BP, Sands AT. Modeling drug action in the mouse with knockouts and rna interference. Drug Discov Today TARGETS 2004;3(5):198—207.

[79] Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, et al. Functional characterization of the s. cerevisiae genome by gene deletion and parallel analysis. Science 1999;285(5429):901—6.

[80] Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B, et al. Functional profiling of the saccharomyces cerevisiae genome. Nature 2002;418(6896):387—91.

[81] Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, Proctor M, Onge RPS, Tyers M, Koller D, et al. The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science 2008;320(5874):362—5.

[82] Barry CE, Blanchard JS. The chemical biology of new drugs in the development for tuberculosis. Curr Opin Chem Biol 2010;14(4):456—66.

[83] Koul A, Dendouga N, Vergauwen K, Molenberghs B, Vranckx L, Willebrords R, Ristic Z, Lill H, Dorange I, Guillemont J, et al. Diarylquinolines target subunit c of mycobacterial atp synthase. Nat Chem Biol 2007;3(6):323—4.

[84] Haagsma AC, Abdillahi-Ibrahim R, Wagner MJ, Krab K, Vergauwen K, Guillemont J, Andries K, Lill H, Koul A, Bald D. Selectivity of tmc207 towards mycobacterial atp synthase compared with that towards the eukaryotic homologue. Antimicrob agents Chemother 2009;53(3):1290—2.

[85] Anand P, Chandra N, Characterizing the pocketome of mycobacterium tuberculosis and application in rationalizing polypharmacological target selection, Sci Rep 4.

[86] Movahedzadeh F, Smith DA, Norman RA, Dinadayala P, Murray-Rust J, Russell DG, Kendall SL, Rison SC, McAlister MS, Bancroft GJ, et al. The mycobacterium tuberculosis ino1 gene is essential for growth and virulence. Mol Microbiol 2004;51(4):1003—14.

[87] Dubée V, Triboulet S, Mainardi J-L, Ethève-Quelquejeu M, Gutmann L, Marie A, Dubost L, Hugonnet J-E, Arthur M. Inactivation of mycobacterium tuberculosis l, d-transpeptidase ldtmt1 by carbapenems and cephalosporins. Antimicrob agents Chemother 2012;56(8):4189—95.

[88] Cordillot M, Dubée V, Triboulet S, Dubost L, Marie A, Hugonnet J-E, Arthur M, Mainardi J-L. In vitro cross-linking of mycobacterium tuberculosis peptidoglycan by l, d-transpeptidases and inactivation of these enzymes by carbapenems. Antimicrob agents Chemother 2013;57(12):5940—5.

[89] Feng Z, Barletta RG. Roles of mycobacterium smegmatis d-alanine: D-alanine ligase and d-alanine racemase in the mechanisms of action of and resistance to the peptidoglycan inhibitor d-cycloserine. Antimicrob agents Chemother 2003;47(1):283—91.

[90] Belanger AE, Besra GS, Ford ME, Mikusová K, Belisle JT, Brennan PJ, Inamine JM. The embab genes of mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proc Natl Acad Sci 1996;93(21):11919—24.