Rationalization and prediction of drug resistant mutations in targets for clinical anti-tubercular drugs

Jyothi Padiadpu¹, Sumanta Mukherjee¹ and Nagasuma Chandra*

Department of Biochemistry, Indian Institute of Science, Bangalore 560012, India

Communicated by Vsevolod J. Makeev

(Received 31 May 2011; final version received 25 March 2012)

Resistance to therapy limits the effectiveness of drug treatment in many diseases. Drug resistance can be considered as a successful outcome of the bacterial struggle to survive in the hostile environment of a drug-exposed cell. An important mechanism by which bacteria acquire drug resistance is through mutations in the drug target. Drug resistant strains (multi-drug resistant and extensively drug resistant) of Mycobacterium tuberculosis are being identified at alarming rates, increasing the global burden of tuberculosis. An understanding of the nature of mutations in different drug targets and how they achieve resistance is therefore important. An objective of this study is to first decipher sequence as well as structural bases for the observed resistance in known drug resistant mutants and then to predict positions in each target that are more prone to acquiring drug resistant mutations. A curated database containing hundreds of mutations in the 38 drug targets of nine major clinical drugs, associated with resistance is studied here. Mutations have been classified into those that occur in the binding site itself, those that occur in residues interacting with the binding site and those that occur in outer zones. Structural models of the wild type and mutant forms of the target proteins have been analysed to seek explanations for reduction in drug binding. Stability analysis of an entire array of 19 mutations at each of the residues for each target has been computed using structural models. Conservation indices of individual residues, binding sites and whole proteins are computed based on sequence conservation analysis of the target proteins. The analyses lead to insights about which positions in the polypeptide chain have a higher propensity to acquire drug resistant mutations. Thus critical insights can be obtained about the effect of mutations on drug binding, in terms of which amino acid positions and therefore which interactions should not be heavily relied upon, which in turn can be translated into guidelines for modifying the existing drugs as well as for designing new drugs. The methodology can serve as a general framework to study drug resistant mutants in other micro-organisms as well.

Keywords: mutability index; residue conservation index; TB drug targets; mutant structures; decrease in drug binding

Background

Tuberculosis, the largest killer among infectious diseases, has been an enemy of mankind since antiquity. Despite the availability of several anti-tubercular drugs, a preventive vaccine (Bacille Calmette-Guérin, BCG) and systematic concerted efforts by the global medical community, tuberculosis remains an unsolved problem. The major cause for concern is the emergence of the multi-drug resistant (MDR-TB) and extensively drug resistant (XDR-TB) strains of the disease (WHO, 2010). The majority of the clinical cases of drug resistance are considered to be due to mutations in the target or prodrg activating proteins (Morar & Wright, 2010; Wright, 2007). A wide array of mutants of different target proteins have in fact been recorded in clinically isolated drug resistant strains of the bacterium (Sandgren et al., 2009).

In the clinic, the problem of drug resistance is typically addressed by rotation of antibiotic combinations and enhanced medical supervision to ensure patient compliance (Kohanski, Dwyer, & Collins, 2010). This is especially important for managing tuberculosis, since the treatment period typically spans 6 months to a year (WHO, 2010). In recent years, there are indeed studies towards countering drug resistance from a new drug discovery perspective, such as the use of virulence factors as drug targets, with a belief that such targets are less mutable (Tomioka, Tatano, Sano, & Shimizu, 2011; Zhang, 2005). Yet another approach used is to screen for chemical analogues that turn out to retain the desired pharmaco-
logical properties in wild type as well as mutant versions of the target protein (Kaneko, Cooper, & Mdluli, 2011; Payne, Gwynn, Holmes, & Pompilio, 2007). Despite these, the available statistics indicate that the disease is still on the rise, warranting more research to identify new ways of countering emergence of drug resistance. In order to make these and any long-term approaches useful, a thorough understanding of the underlying mechanisms of emergence of drug resistance is essential (Brown et al., 2010; WHO, 2010).

It seems logical to think that any mutation that is acquired in the target protein for the purpose of achieving drug resistance ought to be so specific, that drug binding is altered, yet stability and more importantly the function of the protein should be left intact, so that the pathogen continues to survive unhindered by the drug. This is because a mutation that either destabilizes the protein or disrupts its natural function through altering substrate binding would by itself be harmful to the organism and therefore may not survive any selection pressure (Toprak et al., 2011). However, these aspects in drug resistant mutants have not been adequately studied. A key question would be to understand which amino acid mutations at which positions do not compromise the function or stability of the protein. Here we seek to address this question through a comprehensive structural analysis of the target proteins, also including clues from the wealth of information present in the naturally occurring homologues in diverse species. Mutations acceptable to the protein in terms of the function and stability can be screened for any possible effect on binding to the given drug. The objectives of this study are (a) to decipher sequence as well as structural bases for the resistance from a large database of drug resistant mutants in different drug targets from Mycobacterium tuberculosis and (b) to utilize the generated knowledge to prioritize interactions based on the likelihood of the residues to acquire a mutation. Such knowledge can be ultimately useful for deriving design or modification principles for the given drug, by reducing reliance on interactions of the drug with highly mutable residues.

Methods

Data-set of drug resistant mutations of different target proteins

The data-set of point mutations to TB drug targets, curated by Sandgren and co-workers (2009), form the bulk of the data-set used in this study. To keep it up-to-date, a literature search was carried out to obtain data on drug resistant point mutations (Abbadi, Sameaa, Morlock, & Cooksey, 2009; Morlock, Metchock, Sikes, Crawford, & Cooksey, 2003; Ramaswamy et al., 2003; Srivastava et al., 2006). The UniProt database was also screened to identify information on such mutants (The UniProt Consortium, 2010). Additional data on mutants thus obtained were added onto the original data-set. Only the point mutations related to the proteins involved in the pharmacological action of the nine major drugs in the TB treatment regime are considered here in this study. Put together, the data-set contains about 670 different drug resistant mutants of 37 different drug targets for clinically used anti-tubercular drugs. Nine different drugs are used clinically for the treatment of tuberculosis, of which five (isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin) are front-line drugs and four (fluoroquinolones, ethionamide, para-amino salicylic acid and amikacin/kanamycin) are used as second-line drugs (Shenoi & Friedland, 2009).

Conservation index

The sequence information of the wild-type target proteins and their drug resistant mutants is collected from UniProt. The sequences of the homologous proteins for each target were identified by performing BLAST (BLASTP 2.2.25+) searches against (a) a database of mycobacterial sequences (MYC) and (b) against the entire non-redundant (NR) database (Altschul, Gish, Miller, Myers, & Lipman, 1990). Only those proteins in the database that had at least 70% similarity with the query target as obtained with the BLOSUM62 substitution matrix, covering at least 70% of the query length coverage and an e-value of less than .0001 were considered as homologues. Where the numbers of homologues were very high, only the top 100 hits were considered for this study, so as to maintain some uniformity in the number of homologues considered for different proteins and to consider only those variations that occur in the closest neighbourhood of the query in sequence space.

The conservation index (CI) is the measure of position specific relative information entropy, evaluated from the multiple sequence alignment of the query sequence with its sequence homologues (Ng & Henikoff, 2006). For the set of sequence homologues identified as described earlier, a multiple sequence alignment is performed using CLUSTALW (Thompson, Higgins, & Gibson, 1994). In order to have a relative understanding of the randomness, we represent a normalized score between 0–100, where 100 indicates that all the amino acids have been sampled equally at the given position. In other words, all 20 amino acids are found in that column in the multiple sequence alignment in some species or the other. CI thus captures functional importance of an amino acid at the given position (Strait & Dewey, 1996). Since CI is evaluated on the basis of natural sequence homologues, higher CI signifies low positional relevance.

The measure of information entropy is evaluated as

$$\eta(i) = - \sum_a p_i(a) \ln p_i(a)$$  

(1)
where $p_i(a) \rightarrow$ frequency of occurrence of amino acid $(a)$ at position $i$ in the alignment.

$$p_i(a) = \frac{n_i(a)}{\sum_a n_i(a)} \quad (2)$$

$n_i(a) \rightarrow$ number of occurrences of amino acid $(a)$ at position $i$ in the alignment.

The relative entropy or the CI is evaluated as

$$\mu(i) = \left( \frac{\eta(i)}{\eta_{\text{max}}} \right) \times 100 \quad (3)$$

where $\eta_{\text{max}}$ to maximum entropy value that can be achieved ($2.996 \approx 3.0$), i.e. when all amino acids are equally likely.

**Structural models of wild type and mutant target proteins**

Of the 38 targets, crystallographic structures were available for eight target proteins, from PDB (Berman et al., 2000). For an additional eight proteins, close structural templates were available in PDB, from which high confidence molecular models were obtained from the in-house Mtb structural proteome repository. Thus structural models for 16 proteins could be obtained. Models in the repository were derived using standard homology modelling protocols using ModPipe (Pieper et al., 2009) and chosen based on sequence identity and query length coverage, Discrete Optimized Protein Energy (DOPE) score and ModPipe Quality Score (MPQS) scores. The models have also been taken through stringent quality assessment measures as described earlier (Anand et al., 2011). Based on these wild-type protein structures, point mutant structures were modelled using SCRWL 4.0 (Krivov, Shapovalov, & Dunbrack, 2009) and energy minimized, leading to a total of 540 structural models. Energy minimization was carried out using GROMOS96 43A1 force field of GROMACS molecular dynamics simulation package which includes a conjugate gradient minimizer (Hess, Kutzner, van der Spoel, & Lindahl, 2008). Table S1 lists model building and quality check parameters for the models.

**Structural superposition**

In order to measure local structural changes due to a mutation, the minimized mutant structures were superposed onto their corresponding wild-type proteins. Superposition was achieved through a local implementation of the widely used Kabsch structural alignment algorithm (Kabsch, 1976), which has the following components: (a) translation of the set of coordinates, so that their centroids are aligned to origin, and (b) computation of a covariance matrix between translated coordinates, $\text{Cov} (X, Y) = X^T Y$, (c) singular value decomposition of the covariance matrix, $\text{SVD} (\text{Cov} (X, Y)) = USW^T$ and (d) computation of the rotation matrix $(WDU^T)$ where $D$ is a diagonal matrix, $D = [1, 1. \text{Sign}(\det(X^T Y))]$. Root mean square deviation (RMSD) of all atoms in the defined zones values are computed for each mutant-wild-type superposition.

**Binding site extraction**

Of the available target crystal structures, only InhA was available as a complex with the drug isoniazid and separately also with another drug ethionamide. A 4.5 Å zone around the drug atoms in either of the complexes was taken as the binding site in these structures. In other cases, however, the structure of the drug complex was not available. To delineate the boundaries of the binding site, we used PocketDepth, an algorithm developed and extensively validated in the laboratory (Kalidas & Chandra, 2008). In order to keep the definition criteria comparable in all structures, the pocket detection algorithm was run for the InhA structure also. This, in fact served as a positive control, since it identified all the residues surrounding the drug INH correctly.

**Docking drug molecules into target protein structures**

The three-dimensional structural models of the drugs were obtained from Pubchem (Li, Cheng, Wang, & Bryant, 2010), where available. Others were modelled using Virtual Computational Chemistry Laboratory implementation E-babel (Tetko et al., 2005). The drugs were docked onto the corresponding targets, using AutoDock Vina (Sousa, Fernandes, & Ramos, 2006; Trott & Olson, 2010). The receptor and the ligand files were taken in pdbqt format; and a search space of $60 \times 50 \times 50$ Å around the centre of the binding site was used. The interactions between the drug molecules and the corresponding targets were computed using MGL tools for each complex. The number of steps in a run is determined adaptively, according to the complexity of the problem, and several runs starting from random conformations are performed, after which the possible set of binding poses is clustered and ranked to output the largest clusters with highest binding potentials. The selection criteria of the docked solutions from AutoDock Vina are based on energies as well as the cluster sizes based on a RMSD measure of docked pose of the bound ligand. We have manually also checked and ascertained that known binding site residues are involved in binding in the final docked pose. The force field employed in the AutoDock is based on a comprehensive thermodynamic model for ligand binding free energies that also incorporates a charge-based method for evaluation of desolvation energies (Huey, Morris, Olson, & Goodsell, 2007). The scoring function combines certain advantages of knowledge-based potentials and empirical scoring functions. The binding energy scores obtained for a given drug...
with the mutant structures were then compared with the binding energy score of the corresponding wild-type protein to obtain a relative appreciation of the extent of binding (Huang, Shoichet, & Irwin, 2006; Hetenyi, Maran, Garcia-Sosa, & Karelson, 2007).

**Estimation of structural stability**

Stability analysis was carried out by using PoPMuSiC (Dehouck et al., 2009), a novel tool that provides an estimate of change in structural stability caused by a single point mutation in the protein. It estimates the stability changes with linear combination of database derived potentials, whose coefficients depend on the solvent accessibility of the mutated residue. A systematic analysis of all possible mutations of the given protein structure at all residue positions is carried out. The resulting PoPMuSiC scores are then analysed to identify mutations that have the potential to render the protein unstable.

**Identification of residues critical for drug binding**

In order to systematically scan the binding site for possible mutations that may render the protein resistant to a given drug, we have computationally modelled mutations using a protocol that mimics ‘scanning alanine mutagenesis’ (Cunningham & Wells, 1989). Each residue of binding site (zone 1) is mutated to an alanine and its affinity to the drug computed by a docking exercise similar to that described above. This is carried out for all target proteins. Where possible, the procedure is repeated with its natural ligand for each protein. Individual residue contributions are computed from the set of mutants by calculating the difference in binding energies with the corresponding wild-type protein in each case.

**Computing ligand access channels**

In order to identify those residues that may be involved in providing an access channel to the ligand or the drug to the binding site in the protein structure, Caver (Cunningham & Wells, 1989) has been used. Caver is a tool, which analyses the void volume in the protein structures using reciprocal distance convex hull approximation to set the protein boundary and using Dijkstra algorithm to find shortest paths to access the protein interior through void volumes.

**Scoring and ranking strategy**

We selected five different parameters: (a) the CI, (b) solvent accessibility, (c) change in drug binding affinity, (d) stability and (e) binding site structural deformation measured in terms of RMSD. Seventy-five per cent cut-off criteria that explain most of the mutational evidence was selected as a decision boundary. In each case depending on the cut-off values, we have introduced a fuzzy measure to quantify the decision bias, where zero signifies no confidence and one signifies maximum confidence to correlate a particular observation with a resistant mutation. The correlation between different classifiers is used to devise a ranking scheme, incorporating the confidence measure to rank the observed mutations.

**Results and discussion**

**Overview**

A broad outline of the study is illustrated in the flow-chart in Figure 1. In a nutshell, a data-set of 540 drug resistant point mutants of 16 drug targets of clinically used anti-tubercular drugs of seven types has been studied to understand the structural basis for drug resistance. Following this, sequence conservation analyses against two different data-sets have been performed, from which, residue-wise, site-wise and whole protein conservation indices have been computed. Data obtained from these computations have then been used to analyse the nature of proteins in which observed mutations occur.

Contributions to drug binding by individual residues at the binding site in each protein have been computed using computational alanine scanning mutagenesis. Similarly, residues important for the stability of the structure are identified through a computation of a comprehensive cassette of all possible mutants at each position. We have also attempted to rationalize the loss of drug binding ability in the mutants through structural analysis of direct and indirect interactions of the residue with the drug. Finally, combined criteria that the observed mutations fall into have been analysed, which forms a framework for prediction of amino acid positions that have high propensity to acquire resistant mutations.

**Structural analysis of drug resistant mutants**

Table S2 indicates the number of drugs, the corresponding target proteins considered in the study. Numbers of mutations available for different targets are also listed. A total of 540 structural models were considered for further analysis. Binding sites are defined for each of the target proteins as described in the methods section. Mutations are then classified into those that occur in the binding site and those occur elsewhere. Of the mutations, 70 were found to be in the binding site itself in nine different proteins, while the remaining 470 mutations in 14 proteins were outside the binding site. Those in the binding site itself is referred to as zone 1 mutations hereafter, and can be considered as having a possible direct effect whereas those outside the site can only have an indirect effect on drug binding.

Those residues that directly interact with residues in zone 1 in the wild-type proteins are called zone 2 residues and mutations in them termed as zone 2 mutations,
so as to capture effects of a mutation in this region on the conformation of the residues in zone 1. The rest are referred to as zone 3+ mutants. The distribution of mutations into the three zones is illustrated in Figure 2.

The structural models were analysed in terms of the effect a mutation has, on binding the given drug. Binding energies were computed for the wild type as well as the mutant complexes and differences between them noted. An obvious analysis is to check whether the differences in the individual drug–protein interactions between the wild type and the mutant versions, can explain loss in drug binding in the latter. Of the 540 mutations, 70 are in zone 1 and readily explain loss of drug binding. Structural positions of a few selected mutations are illustrated in Figure S1.

**InhA**: 10 mutations are observed in InhA (enoyl-ACP reductase), (Rv1484) the primary target for the activated form of isoniazid (Rawat, Whitty, & Tonge, 2004).
2003), of which eight are in zone 1. The available crystal structure of the mutation S94A indicates that a hydrogen bond between Ser 94 and a water molecule hydrogen bonded to the O9 oxygen atom of the drug isoniazid is indeed disrupted by the loss of serine hydroxyl moiety in the vicinity, leading to significant weakening of drug binding (Vilcheze et al., 2006). The other binding site mutations such as I21V, I21T, I95P, I95T and I194T also show differences in the interactions with INH as compared to wild type.

KatG: Isoniazid is a prodrug, which gets converted to an active form (an adduct) by the enzyme KatG, catalase/peroxidase (Rv1908c) (Cade, Dlouhy, Medzhradszky, Salas-Castillo, & Ghiladi, 2010). As KatG is a crucial enzyme for the mechanism of action of isoniazid, it has been well studied (Ghiladi, Knudsen, Medzhradszky, & Ortiz de Montellano, 2005). As many as 191 point mutations are observed in clinically isolated drug resistant varieties. However, only seven of them are featured in zone 1 and around 20 of them in zone 2 and the rest are in zone 3. W107R, Y229F, mutations are in the active site of the protein, and result in steric hindrance for the drug to bind, compared to that in the wild type. S315T is considered to be at least the cause for 40% of INH resistance (Pym, Saint-Joanis, & Cole, 2002). Our results indicate that, in terms of binding affinity, this mutation does not show much variation with respect to the wild type. On the other hand, we observe a constriction in the channel size, reducing access to the drug, possibly the cause for resistance. Reduction of access to the drug has in fact been reported earlier in the literature (Cade et al., 2010), lending support to our observation.

EthA: active form of ETH, an InhA inhibitor is formed by a FAD-dependent monooxygenase, EthA (Rv3854c) (Morlock et al., 2003). Fourteen different point mutants of EthA are known in ETH resistant forms, of which three are in zone 1 (T184 K, T340 K and R461S) and five are in zone 2. T390A features specifically in zone 2. Binding appears to be unfavourable in the mutants, due to steric clashes.

RpoB: the target for rifampicin, RpoB features only one mutant in zone 1 while it has as many as 87 mutations in outer zones. The region between residues 507 and 533 (in zone 3) has been referred to as RRDR or ‘rifampicin resistance determining region’ in the literature (Lee, Lim, Tang, & Wong, 2005), but mutations in other regions are also linked with RIF resistance. The region has been reported to possibly alter the structure of the binding site for rifampicin and supposed to directly block the path of the elongating RNA when RpoB bound with rifampicin (Campbell et al., 2001; Jenkins, 2005).

PncA: while the precise target of PZA is not as yet known, PncA is known to be an activator of the prodrug PZA into a pharmacologically active species (Lemaître, Callebaut, Fenois, Jarlier, & Sougakoff, 2001). As many as 180 point mutations spanning almost all possible residues are observed in this protein (Scorpio et al., 1997; Sekiguchi et al., 2007), many of them retaining partial catalytic activity. Of these, 16 are in zone 1 and 88 are
in zone 2. The binding energy as calculated through a docking exercise by AutoDock was not favourable for any of these mutations. Experiments reported in the literature demonstrate that the mutant enzymes have reduced catalytic activities, thus validating predictions from docking studies (Lemaitre et al., 2001). Of the first group or zone 1, three mutants (K96T, D8G and S104R) show binding affinities drastically lower when compared to the wild-type PncA protein. C138Y shows high steric hindrance due to introduction of the aromatic group. C138S also shows a decrease in the binding affinity. The second group of mutants, F13S, Y103S, A146V, T61P and P69L, which are reported to show lower levels of activity, are in the neighbourhood of the binding site (and belong to zone 2), whereas T167I, T168N are located far away from the binding site.

GidB: a target of streptomycin (probable glucose-inhibited division protein B), Rv3919c, in which about 15 mutations are seen in resistant strains (Spies, da Silva, Ribeiro, Rossetti, & Zaha, 2008). The mutants are in zone 2 and in the vicinity of the binding site, possibly causing subtle structural changes that lead to suboptimal disposition of the binding site residues, reducing affinity to the drug. In ManB (Rv3264c), a target of ethambutol, resistant mutants such as D152N, show high steric hindrance due to consequential rearrangements at the site.

**Zone 2 and zone 3 + mutations**

Some of the zone 2 mutations could be rationalized by the fact that they lead to rearrangements of residues in zone 1 thereby affecting drug binding. To analyse this aspect systematically, we computed RMSD values at the sites after superposition of all Ca atoms in the energy minimized models and the wild-type structure. The change in binding affinities due to mutations has been analysed through rank correlation studies. Each observation is ranked on the basis of their change in drug binding affinity and plotted against computed binding site RMSDs. A correlation between reductions in affinity for drug binding with increase in RMSD values at the site atoms is clearly evident from Figure 3.

The effect of binding site residue rearrangement is evident for zone 1 mutations, while this is much more complex for zone 2 and zone 3 residues, since the induced rearrangement involves complex interactions among various residues. Hence we performed the analysis separately for these two scenarios (Figure 3(a) and (b) for direct and induced rearrangements, respectively).

Figure 3(a) illustrates the RMSD values at the site for zone 1 mutations, for each target. The results reveal that the binding site is being perturbed as the mutations occur in the binding site itself correlating clearly with a lesser binding affinity score for individual mutants. This is particularly evident in the figure for Inha, PncA, KatG.
and RpsL where a large number of mutants are known. Figure 3(b) is a similar plot for RMSD values of zone 1 residues due to zone 2 and zone 3 mutations. The correlation, although to a lesser extent than with zone 1 mutants, is apparent enough even in the outer zones, indicating that the binding site is perturbed, even due to outer zone mutations and in general correlates with reduction in drug binding affinity. In some of the cases of outer zone mutations, mechanistic explanations for rearrangement of zone 1 residues are clear. For example, in EthA the target for ETH, a mutation T390A in zone 2, leads to shifts in a spatially adjacent residue S380 and also in Q289 near zone 1, leading to loss of drug binding. In some cases, mutants appear to be located in domain interfaces and subunit interfaces, as has been noted earlier as well (Mathys et al., 2009), opening important questions as to how such mutations can lead to loss or reduction of drug binding. Although the effects of all zone 2 and zone 3 mutations are not fully comprehensible at this point in time, some structural explanations ought to be present. One possible aspect to consider would be to see if reduction in binding affinity correlates loss of flexibility in the protein molecule due to a zone 2 or zone 3 mutations. Perhaps, a dynamic view, such as through molecular dynamics simulations can throw light on the changes that occur in the binding site due to a mutation in a remote location.

Another important reason for resistance from outer zone mutations could be that they alter the access to the drug by modifying the access channel either by introducing steric hindrance or due to the loss of a transient interaction, perhaps required for positioning the drug initially before it reaches the binding site. Some of the zone 2 and zone 3 mutations indeed appear to occur in these regions, for example, in EthA, T390A and also in KatG S315T appear to be directly situated in the access channels (Figure S2). About 36 mutations from four proteins appear to occur in the access channels computed using Caver. In InhA, PncA, KatG and EthA, distinct access channels can be discerned from the structure, which houses some of the mutations outside zone 1. InhA mutation at zone 2 residue I47T shows that it has a direct impact on the binding site residue I15, which interacts with the ligand. Similarly in EthA, the residues such as T390 and R461 are important zone 2 residues mutations thus having effects on the binding site.

A drug resistant mutation should not only retain function of the protein but should also reduce drug binding, which means affinity to the drug should be reduced, yet retaining wild-type affinity to the natural substrate. In addition, the stability of the protein should also be maintained (Figure S3). Therefore, not all positions are available for the emergence for such a mutation. Docking with natural substrates as well as the drugs to each of the mutants indicates that in general the mutants retain affinity to the natural substrates, indicating that this criterion is important for the emergence of a drug resistance mutation.

Binding energy scores of the mutants to the drugs are generally poorer as compared to their wild-type structures, indicating that drug binding is reduced in number of mutants (Figure 4(a)). Figure 4(b) clearly shows a right shift in the drug binding energy scores as compared to the substrate binding energy scores. Binding energy scores of a substrate is much higher than that of a drug for the targets studied here. The range for the natural substrate binding energy score is within −10.4 to −3.9 kcal/mol with an average of −6.43 kcal/mol and a standard deviation of 1.84, whereas drug molecule binding energy scores are within a range of −5.7 to 1.2 kcal/mol with an average of −3.31 kcal/mol and a standard deviation of 2.135, as obtained from the docking exercise using Autodock Vina. It must be noted that, expectedly, the differences in binding energies are subtle for some mutations, and prominent for some others. It is still clear that, the 13 percentile of mutations show better affinity to the natural substrate as compared to the drugs indicating that accepted point mutations for causing drug resistance affects the drug binding more than the natural substrate binding.

Conservation indices

To address the question of why mutations are observed in certain positions of a given protein, it is useful to approach the problem from the angle of which residues in a given target are the most poised to acquire mutations. Towards this, we carried out sequence conservation analyses and computed conservation indices of each residue (Larijani, Frieder, Basit, & Martin, 2005). Two different databases were used to identify homologues of the query target: (a) a database of 99 available mycobacterial genomes (MYC) and (b) the entire NR sequence database from NCBI containing 11,505,486 protein sequences (version of May 2011). Homologues identified through BLAST searches were then multiple-aligned using CLUSTALW for each target protein and the number of substitutions counted for each residue, from which the CI is computed as described in the methods section. Two different databases were used, so as to capture (a) those positions prone to substitutions even within the same genus and (b) those positions that show high variability when compared against the entire NR sequence database. For each of the proteins studied here, radial graphs are plotted as shown in Figure 5 for two examples (InhA and PncA). Individual residues in the order of the primary sequence, line the outer side of the circle, while the extent of conservation of the residue in the two databases (MYC and NR) are shown as radii of each residue in blue and red lines, respectively. The locations of observed mutations are also reflected in the figure as thin sharp radiating lines (no length significance), mutations in zone 1 are shown in
black and those of outer zones are in green. In the Figure 5 (b), we can see a large number of mutations throughout the sequence length of the protein PncA as indicated by the green lines. Average conservation indices of different proteins are observed to be clearly different, from this analysis. This indicates that different proteins have different proportion of mutable positions. Knowledge of this would be useful
in the rational choice of a drug target, by prioritizing possible targets based on their overall mutability values.

The trend in the conservation indices computed using homologues identified both in the set of 99 mycobacterial species as well as the large NR database are clearly similar. This suggests that for all the proteins studied here, the conservation patterns seen in the specific mycobacterial database hold good even for the larger NR database.

The residues are colour-coded based on their CI ranges as shown in Figure S4. As expected, the core of the protein, including some residues in the binding site are highly conserved and therefore have a low variability index, whereas residues at the surface have high variance among the set of homologues and therefore have high variability indices. Although this is common knowledge based on the current understanding of protein structures (Ng & Henikoff, 2006), it has rarely been used in relating to propensity of a given amino acid residue in acquiring a drug resistant mutation and the implications it will have for designing the drug.

In view of this, we checked if observed mutations occur at the most variable positions in the protein. Surprisingly, only a few of the zone 2 and zone 3 mutations were in the highly variable positions. A number of them are in fact in the high conservation positions. The trend for zone 1 mutations to occur in least variable positions seems to be even more pronounced (Figure 6(a)). This suggests that such mutations especially in zone 1 and also in outer zones ought to be a result of some type of guided evolution so as to attain the goal of acquiring resistance to drugs. It is also possible that zone 2 and zone 3 mutations on the other hand could be those that get selected for fitness in terms of drug resistance from a large pool of possible mutants that satisfy the criteria listed earlier. It would be useful to know which residues are more prone to mutations in each protein. Surprisingly, our analysis indicates a correlation of observed drug resistant mutations with position of high conservation indices in all the targets studied here. Although this might seem surprising at the outset, it appears logical when viewed from the perspective of evolution, since positions of low conservation generally have a lower influence on structure and function than high conservation positions, making the former less attractive for acquiring drug resistant mutations.

**Stability analysis**

In order to understand which positions when mutated alter stability of the protein substantially, systematic analyses of all possible mutations at all positions were carried out and the differences in internal energy computed using the PopMusIc tool. A small but clear drift in the distribution of pop stability score in the mutants can be seen from Figure 6(b). It was interesting to observe that all the observed mutations have only marginally lower stability in a number of cases. There are some positions in each target which when mutated, alter the stability drastically, which are not likely sites for drug resistant mutations. The analysis thus indicates which residues are likely and which are not likely to acquire a mutation.

We then checked if solvent accessibility played a role in the selection of positions for emergence of drug resistant mutations. Figure 6(c) indicates that majority of the observed mutations are in the lower accessibility positions. There was no clear pattern in the secondary structure distribution of the observed mutations. About 126 are seen in α-helices and 92 in β strands.
Integrated mutability index (IMI)

Of the six different parameters (SASA, Change in drug binding energy score, CI, Access Channel, Binding site RMSD and Protein structure stability) analysed, four of them in combination explain 473 of the mutations of 12 target proteins, which are conservation indices, change in drug binding affinity, stability and solvent accessibility. Explanation of the mutational incidence at a particular position on the basis of one parameter only will introduce a high rate of false positives. Combinations of three or more parameters together are observed to be useful in obtaining correlations in a large number of mutations, with higher confidence.

Figure 6. Pie charts showing the distribution of (a) conservation indices of all targets across NR and MYC data-sets, (b) ΔΔG values obtained from PopMuSiC indicating stability for all comprehensive cassettes of mutants and (c) ratios of residue-wise solvent surface accessible area to the total surface area. The left panel illustrates charts computed by taking the whole protein into consideration, for all the targets and the right panel illustrates charts computes by taking only observed mutation positions into account.
We have defined six different sets where in each set represents the mutation occurrence explained by particular selection criteria. Out of all possible set combinations, the selected set of four criteria explained maximum number of observed mutations. The scoring scheme is devised is based on the combination of the four parameters so as to get maximum intersection. The Venn diagram shown in Figure 7 illustrates that as many as 473 mutations are explained by one of the combinations. In order to derive an overall measure of the propensity to acquire a mutation, a metric termed as IMI has been computed using the following fuzzy rule

$$\text{IMI}(O_i) = (C_I(O_i) \cap \text{SASA}(O_i)) \cup (\Delta \Delta G(O_i)$$

$$\cap \text{SASA}(O_i) \cup (\Delta \Delta G(O_i) \cap C_I(O_i))$$

(4)

where, $O_i$ is the $i$th observation, IMI ($O_i$) is the overall predictability measure on the observation $O_i$, CI ($O_i$), SASA($O_i$), DBA($O_i$), $\Delta \Delta G(O_i)$ the confidence measure of observation $O_i$ classified properly using CI, solvent accessibility, change in drug binding affinity and change stability, respectively. $\cap$ and $\cup$ signifies logical AND and OR operation.

Through this each mutation is given a score, which indicates overall confidence, and also the number of the criteria that explains the classification. Mutations are ranked based on the integrated scores which explain how well the mutational event is explained. For example, KatG (L141F, N138F) satisfies all the criteria tested here and hence has the top most rank. Similar explanations can be obtained for about 294 drug resistant mutations in all, with confidence greater than 1. Table S2 shows the different parameters used here and a full list of best explained mutations.

For some mutations, such correlations were not observed, AhpC (alkylhydroperoxide reductase C, Rv2428) (L3S, K70L) is one such example. In summary, the parameters tested here are helpful in understanding which residues are highly mutable and therefore which proteins are more mutable. Information about this is likely to be extremely useful in target discovery, so that the best targets can be ultimately selected.

**Interaction indices of individual residues in zone 1**

Next, we carried out an analysis of the importance of the relative contribution of each interacting residue at the binding site towards drug binding (Borea, Varani, Gessi, Gilli, & Dalpiaz, 1998). This was achieved through mutating each residue individually to an alanine, mimicking that done experimentally for many proteins by the alanine scanning mutagenesis approach (Morrison & Weiss, 2001; Van Petegem, Duderstadt, Clark, Wang, & Minor, 2008). The difference in interaction energies between the wild type and the alanine mutant for a given residue is obtained as $\Delta \Delta G$ values and taken as a measure of the contribution of that residue towards drug binding in the wild-type protein. It is obviously important to know those residues which contribute significantly to drug binding. However, if such residues turn out to be highly mutable residues, then the chances of drug resistance would be higher and reliance on such residues should therefore be reduced. Examples of such residues in different targets are shown in Table 1. This type of analysis is conceptually new and will help in prioritizing interactions and hence residues in the binding site to rely on.

In summary, the analysis presented here reveals structural bases for several mutations, especially in zone 1. For others in zone 2 and zone 3, some are explained through rearrangements of zone 1 residues or through restricting access to the drugs as shown in Table 2.

Contributions to the binding energy by individual residues at the binding sites of different proteins have
been computed through modelling alanine mutations at each position. An understanding of interactions and hence residues based on these metrics, will help in deriving design principles to modify existing drugs and also to design new drugs.

There are perhaps multiple factors at play in the process of the emergence of a drug resistant mutation. Nevertheless, in the absence of any other determining information at this point in time, the analysis presented here is useful for understanding which residues have a higher likelihood of being substituted and yet retaining function. Of these, some will have an impact on drug binding and may get selected for drug resistance. While it is difficult to correlate different parameters with the position of the observed mutations in a linear fashion, some trends are clear. Where two or more of the defined criteria are satisfied, the likelihood of that position acquiring a drug resistant mutation is higher than those positions that do not satisfy the criteria. From the set of 540 observed mutations, we notice that (a) the stability of the protein due to a mutation is largely maintained but may be marginally reduced, (b) drug binding affinity is reduced in the mutants, (c) but no significant difference is observed for natural substrate binding affinity and (d) mutations occur at highly conserved positions in the protein that have low solvent accessibility. The analyses presented here is useful for rationalizing the loss of drug binding by the large set of mutations. While the subtle structural rearrangements that occur in zone 1 are easier to comprehend and predict, we also observed significant changes in the vicinity of the binding site in zone 2 and zone 3 mutations as well, suggesting that such mutations have a strong influence in the fine structure at the binding site. Further, based on the mutability propensities values, different residues in a target can in principle be ranked and prioritized to choose a drug such that it has interactions with least mutable residues at the site.

The methodology used is fairly generic and can be easily adapted to study other pathogenic microorganisms. The concepts about ranking targets and ranking target sites within a target are new. They can be easily implemented and have the potential to be incorporated into drug discovery pipelines. There is also scope for improvement by way of systematically exploring parameter space at several points. Those putative targets with high mutability propensities can be avoided, while those with low mutability propensities should get high priority.

| Target  | Number of mutations in zone 1 (a direct interactions with the drug) | Binding site rearrangement (RMSD) due to zone 2/3 mutations | Resistant mutations observed at conserved positions (CI) | Solvent accessible surface area | Access channel |
|---------|---------------------------------------------------------------|-----------------------------------------------------------|-----------------------------------------------------|-----------------------------|---------------|
| EthA    | 1                                                             | 9 (14)                                                    | 7 (15)                                             | 12 (15)                    | 1 (15)        |
| GidB    | 1                                                             | 7 (9)                                                     | 8 (10)                                             | 7 (10)                     | 0 (10)        |
| InhA    | 7                                                             | 1 (3)                                                     | 1 (10)                                             | 9 (10)                     | 1 (10)        |
| KatG    | 4                                                             | 51 (145)                                                  | 127 (149)                                          | 114 (149)                  | 31 (149)      |
| PncA    | 48                                                            | 59 (127)                                                  | 164 (175)                                          | 147 (175)                  | 3 (175)       |
| RpoB    | 1                                                             | 54 (77)                                                   | 78 (78)                                            | 64 (78)                    | 0 (78)        |
| RpsL    | 8                                                             | 8 (8)                                                     | 10 (16)                                            | 5 (16)                     | 0 (16)        |

Values given in each field indicate the number of mutant forms that are explained by the indicated parameter. Numbers in parentheses indicate total number of mutants that could have shown that effect. For example for EthA, in the second field, 9 (14) refer to 9 mutant of a total of 14, show rearrangement in zone 1 due to mutations in zone 2 or zone 3.
Abbreviations

MDR-TB  Multi-drug resistant tuberculosis
XDR-TB  Extensively drug resistant tuberculosis
TB     Tuberculosis
INH    Isoniazid
EMB    Ethambutol
ETH    Ethionamide
PZA    Pyrazinamide
SM     Streptomycin
PAS    Para amino salicylic acid
CI     Conservation index
IMI    Integrated mutability index
RIF    Rifampicin

Acknowledgements

We thank the Department of Biotechnology (DBT) and Department of Science and Technology (DST), Mathematical Biology Initiative and the Government of India for financial support. The use of facilities at the Bioinformatics Centre, Indian Institute of Science is also gratefully acknowledged.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2012.691361.

References

Abbadia, S.H., Sameaa, G.A., Morlock, G., & Cooksey, R.C. (2009). Molecular identification of mutations associated with anti-tuberculosis drug resistance among strains of Mycobacterium tuberculosis. *International Journal of Infectious Diseases*, 13, 673–678.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410.

Anand, P., Sankaran, S., Mukherjee, S., Yeturu, K., Laskowski, R., Bhardwaj, A., Brahmachari, S.K., & Chandra, N. (2011). Structural annotation of Mycobacterium tuberculosis proteome. *PLoS ONE*, 6, e27044.

Berman, H.M., Westbrook, J., Feng, Z., Gililand, G., Bhat, T. N., Weissig, H., ... Bourne, P.E. (2000). The protein data bank. *Nucleic Acids Research*, 28, 235–242.

Borea, P.A., Varani, K., Gessi, S., Gilli, P., & Dalpiaz, A. (1998). Receptor binding thermodynamics as a tool for linking drug efficacy and affinity. *Farmaco*, 53, 249–254.

Brown, K.M., Costanzo, M.S., Xu, W., Roy, S., Lozovskiy, E. R., & Hartl, D.L. (2010). Compensatory mutations restore fitness during the evolution of dihydrofolate reductase. *Molecular Biology and Evolution*, 27, 2682–2690.

Cade, C.E., Dlouhy, A.C., Medzihradzsky, K.F., Salas-Castillo, S.P., & Ghiladi, R.A. (2010). Isoniazid-resistance conferring mutations in Mycobacterium tuberculosis KatG: Catalase, peroxidase, and INH-NADH adduct formation activities. *Protein Science*, 19, 458–474.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., & Darst, S.A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell*, 104, 901–912.

Cunningham, B.C., & Wells, J.A. (1989). High-resolution epitope mapping of bGH-receptor interactions by alanine-scanning mutagenesis. *Science*, 244, 1081–1085.

Dehouck, Y., Grosfils, A., Folch, B., Gilis, D., Bogaerts, P., & Rooman, M. (2009). Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0. *Bioinformatics*, 25, 2537–2543.

Ghiladi, R.A., Knudsen, G.M., Medzihradzsky, K.F., & Ortiz de Montellano, P.R. (2005). The Met-Tyr-Trp cross-link in *Mycobacterium tuberculosis* catalase-peroxidase (KatG): Autocatalytic formation and effect on enzyme catalysis and spectroscopic properties. *Journal of Biological Chemistry*, 280, 22651–22663.

Hess, B., Katzung, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4, 435–447.

Hetenyi, C., Maran, U., Garcia-Sosa, A.T., & Karelon, M. (2007). Structure-based calculation of drug efficiency indices. *Bioinformatics*, 23, 2678–2685.

Huang, N., Shoichet, B.K., & Irwin, J.J. (2006). Benchmarking sets for molecular docking. *Journal of Medicinal Chemistry*, 49, 6789–6801.

Huey, R., Morris, G.M., Olson, A.J., & Goodsell, D.S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*, 28, 1145–1152.

Jenkins, C. (2005). Rifampicin resistance in tuberculosis outbreak, London, England. *Emerging Infectious Diseases*, 11, 931–934.

Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica*, A32, 922–923.

Kalidas, Y., & Chandra, N. (2008). PocketDepth: A new depth based algorithm for identification of ligand binding sites in proteins. *Journal of Structural Biology*, 161, 31–42.

Kaneko, T., Cooper, C., & Mduli, K. (2011). Challenges and opportunities in developing novel drugs for TB. *Future Medicinal Chemistry*, 3, 1373–1400.

Kohanski, M.A., Dwyer, D.J., & Collins, J.J. (2010). How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology*, 8, 423–435.

Krivov, G.G., Shapovalov, M.V., & Dunbrack, R.L.Jr. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, 77, 778–795.

Larijani, M., Frieder, D., Basit, W., & Martin, A. (2005). The mutation spectrum of purified AID is similar to the mutability index in Ramos cells and in ung(−/−) msh2(−/−) mice. *Immunogenetics*, 56, 840–845.

Lee, A.S., Lim, I.H., Tang, L.L., & Wong, S.Y. (2005). High frequency of mutations in the rpoB gene in rifampin-resistant clinical isolates of *Mycobacterium tuberculosis* from Singapore. *Journal of Clinical Microbiology*, 43, 2026–2027.

Lemaître, N., Callebaut, I., Frenois, F., Jarlier, V., & Sougakoff, W. (2001). Study of the structure-activity relationships for the pyrazinamidase (PncA) from *Mycobacterium tuberculosis*. *Biochemical Journal*, 353, 453–458.

Li, Q., Cheng, T., Wang, Y., & Bryant, S.H. (2010). PUBChem as a public resource for drug discovery. *Drug Discovery Today*, 15, 1052–1057.

Mathys, V., Wintjens, R., Lefevre, P., Bertout, J., Singhal, A., Kiass, M., ... Bifani, P. (2009). Molecular genetics of para-aminosalicylic acid resistance in clinical isolates and...
spontaneous mutants of Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy*, 53, 2100–2109.

Morar, M., & Wright, G.D. (2010). The genomic enzymology of antibiotic resistance. *Annual Review of Genetics*, 44, 25–51.

Morlock, G.P., Metchock, B., Sikes, D., Crawford, J.T., & Cooksey, R.C. (2003). EthA, inhA, and katG loci of ethionamide-resistant clinical Mycobacterium tuberculosis isolates. *Antimicrobial Agents and Chemotherapy*, 47, 3799–3805.

Morrison, K.L., & Weiss, G.A. (2001). Combinatorial alanine-scanning. *Current Opinion in Chemical Biology*, 5, 302–307.

Ng, P.C., & Henikoff, S. (2006). Predicting the effects of amino acid substitutions on protein function. *Annual Review of Genomics and Human Genetics*, 7, 61–80.

Payne, D.J., Gwynn, M.N., Holmes, D.J., & Pompliano, D.L. (2007). Drugs for bad bugs: Confronting the challenges of antibacterial discovery. *Nature Reviews Drug Discovery*, 6, 29–40.

Perron, G.G., Lee, A.E., Wang, Y., Huang, W.E., & Barracough, T.G. (2012). Bacterial recombination promotes the evolution of multi-drug-resistance in functionally diverse populations. *Proceedings: Biological Sciences*, 279, 1477–1484.

Pieper, U., Eswar, N., Webb, B.M., Ermanian, D., Kelly, L., Barkan, D.T., … Sali, A. (2009). MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, 37, D347–354.

Pym, A.S., Saint-Joanis, B., & Cole, S.T. (2002). Effect of katG mutations on the virulence of *Mycobacterium tuberculosis* and the implication for transmission in humans. *Infection and Immunity*, 70, 4955–4960.

Raman, K., & Chandra, N. (2008). *Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance. *BMC Microbiology*, 8, 234.

Ramawat, S.V., Reich, R., Dou, S.J., Jaspers, L., Pan, X., Wanger, A., … Gravis, E.A. (2003). Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, 47, 1241–1250.

Rawat, R., Whitty, A., & Tonge, P.J. (2003). The isoniazid–NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase: Adduct affinity and drug resistance. *Proceedings of the National Academy of Sciences of the USA*, 100, 13881–13886.

Sandgren, A., Strong, M., Muthukrishnan, P., Weiner, B.K., Church, G.M., & Murray, M.B. (2009). Tuberculosis drug resistance mutation database. *PLoS Medicine*, 6, e2.

Scorpio, A., Lindholm-Levy, P., Heifets, L., Gilman, R., Siddiqi, S., Cynamon, M., & Zhang, Y. (1997). Characterization of pncA mutations in pyrazinamide-resistant *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*, 41, 540–543.

Sekiguchi, J., Miyoshi-Akiyama, T., Augustynowicz-Kopec, E., Złoszka, Z., Kirikae, F., Toyoda, E., … Kirikae, T. (2007). Detection of multidrug resistance in *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, 45, 179–192.

Shenoi, S., & Friedland, G. (2009). Extensively drug-resistant tuberculosis: A new face to an old pathogen. *Annual Review of Medicine*, 60, 307–320.

Sousa, S.F., Fernandes, P.A., & Ramos, M.J. (2006). Protein-ligand docking: Current status and future challenges. *Proteins*, 65, 15–26.

Spies, F.S., da Silva, P.E., Ribeiro, M.O., Rossetti, M.L., & Zaha, A. (2008). Identification of mutations related to streptomycin resistance in clinical isolates of *Mycobacterium tuberculosis* and possible involvement of efflux mechanism. *Antimicrobial Agents and Chemotherapy*, 52, 2947–2949.

Srivastava, S., Garg, A., Ayyagari, A., Nyati, K.K., Dhole, T. N., & Dwivedi, S.K. (2006). Nucleotide polymorphism associated with ethambutol resistance in clinical isolates of *Mycobacterium tuberculosis*. *Current Microbiology*, 53, 401–405.

Strait, B.J., & Dewey, T.G. (1996). The Shannon information entropy of protein sequences. *Biophysical Journal*, 71, 148–155.

Tetko, I.V., Gasteiger, J., Todeschini, R., Mauri, A., Livingstone, D., Ertl, P., … Prokopenko, V.V. (2005). Virtual computational chemistry laboratory—design and description. *Journal of Computer-Aided Molecular Design*, 19, 453–463.

The UniProt Consortium (2010). The universal protein resource (UniProt) in 2010. *Nucleic Acids Research*, 38, D142–D148.

Thompson, J.D., Higgins, D.G., & Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22, 4673–4680.

Tominka, H., Tatano, Y., Sano, C., & Shimizu, T. (2011). Development of new antituberculous drugs based on bacterial virulence factors interfering with host cytokine networks. *Journal of Infection and Chemotherapy*, 17, 302–317.

Toprak, E., Veres, A., Michel, J.B., Chait, R., Hartl, D.L., & Kishony, R. (2011). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nature Genetics*, 44, 101–105.

Trott, O., & Olson, A.J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31, 455–461.

Van Petegem, F., Duderstadt, K.E., Clark, K.A., Wang, M., & Minor, D.L. Jr. (2008). Alanine-scanning mutagenesis defines a conserved energetic hotspot in the CaValpha1AIT-CaValpha2beta interaction site that is critical for channel modulation. *Structure*, 16, 280–294.

Vilechez, C., Wang, F., Araiz, M., Hazbon, M.H., Colangeli, R., Kremer, L., … Jacobs, W.R.Jr. (2006). Transfer of a point mutation in *Mycobacterium tuberculosis* inhA resolves the target of isoniazid. *Nature Medicine*, 12, 1027–1029.

WHO (2010). Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 global report on surveillance and response. Wright, G.D. (2007). The antibiotic resistance: The nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, 5, 175–186.

Zhang, Y. (2005). The magic bullets and tuberculosis drug targets. *Annual Review of Pharmacology and Toxicology*, 45, 529–564.