CHAPTER – II

REVIEW OF LITERATURE
REVIEW OF LITERATURE

2.1 Tuberculosis – the global scenario

Tuberculosis (TB) is a major medical issue in the world. It causes disease among a huge number of individuals every year and positions besides the human immunodeficiency virus (HIV) infection as a main source of death across the world. In 2014, there was an estimated 9.6 million fresh TB cases that prevailed among 5.4 million men, 3.2 million women and 1.0 million youngsters. There were additionally 1.5 million TB death cases (1.1 million among HIV-adverse individuals and 0.4 million among HIV-constructive individuals), of which roughly 890,000 were men, 480,000 were women and 140,000 were kids (WHO, 2008; 2013). In addition, around 500,000 new multi-drug resistant TB (MDR-TB) cases were evaluated to happen each year (Cole et al., 1998). The number of TB death cases is quite high, however, with appropriate analysis and right treatment, all individuals with TB can be cured (WHO, 2008)

2.2 The organism

The causal organism of TB, Mycobacterium tuberculosis was first extracted by a German doctor and researcher, Robert Koch in 1882 (Cole et al., 1998). This intracellular obligate aerobic actinomycete is around 2-6 μm long with a regeneration time of roughly 20-24 hours under ideal conditions (Rajagopalan et al., 1995; Attorri et al., 2000). It frames tight, rope-like totals in liquid culture (Fregnan and Smith, 1962) and rich sporadic raised states with consolidated fibers in the middle in solid media (Shinnick and Great, 1994) which can take up to 7 days or more to take shape (Schürch and Soolingen, 2012).

According to the Bergy’s Manual of Orderly Bacteriology (edition 2, volume 1), the systematic groping for Mycobacterium tuberculosis is presented below (Garrity, 2012).

\[
\begin{align*}
\text{Domain:} & \quad \text{Bacteria} \\
\text{Phylum:} & \quad \text{Actinobacteria} \\
\text{Class:} & \quad \text{Actinobacteria}
\end{align*}
\]
Mycobacterium has been sub-divided into three groups; the slow-dividing pathogenic that are the *M. tuberculosis* complex (MTBC); the slow-dividing non-tuberculous mycobacteria and the rapidly-dividing non-tuberculous mycobacteria (Gillespie, 2007). The *M. tuberculosis* complex consists of human pathogens like *M. tuberculosis*, *M. canetti*, *M. africanum*, a wide host range pathogen- *M. bovis* and a rodent pathogen – *M. microti* (Gutierrez et al., 2005). Not at all like most bacterial clonal aggregates which have no less than 1% synonymous single nucleotide polymorphisms (sSNPs), the MTBC show between 0.01% to 0.03% sSNPs (Parsons et al., 2002).

Although the individuals from the MTBC show such high genetic homogeneity, they exhibit altogether different epidemiology and range of host, therefore separation between the individual members is important for remedial treatment (Brennan, 2003). Another significant distinction that separates mycobacteria from other microorganisms is the possession of a very complex cell wall along with capsule, a core and an internal membrane layer (Fig. 2.1).

The capsule contains free lipids and mycolates, for example, the phosphatidylinositol mannosides (PIMs) and lipoarabinomannan (LAM). The center comprises of peptidoglycan (PG) and mycolic corrosive associated by an arabinogalactan (AG) polysaccharide, all in all called the mycolyl arabinogalactan– peptidoglycan (mAGP) complex. It is important to mention that the inner membrane is comprises of a lipid bilayer (Kocincová et al., 2001). Due to the presence of large quantity of lipid in the cell wall promotes it to acid fast category and it is apparent when the cells are stained with
Ziehl-Neelsen satain. The characteristic red or pink colour is due to retaining of the primary carbol fuchsin stain because they resist decolorization with acid alcohol.

Fig. 2.1 Schematic diagram showing the structure of mycobacterial cell wall (Jackson, 2014)

2.3 Pathogenesis of Mycobacterium tuberculosis

Pathogenesis and virulence factors of M. tuberculosis are understood very poorly, its intracellular location almost receives protection from the immune systems presenting varied interactions between the pathogen and the host immune response due to infection.

The most commonly occurring type of TB is pulmonary tuberculosis (Mehta et al., 2006) as the basic course of disease is by means of inward breath of airborne beads containing the tubercle bacilli from a contaminated individual (Clark-Curtiss and Haydel, 2003). After entering the bacillus moves to the alveoli where it gets phagocytosed by alveolar macrophages (Saunders and Cooper, 2000; Mehta et al., 1996). In this stage, there could be several different outcomes of the TB-infection (Fig. 2.2) depending on the quantity of
bacteria upon infection and the immunological status of the host (Kaufmann, 2001). The first effect that takes place is the activation of the alveolar macrophages by T cells which have recognized mycobacterial antigens (Mehta et al., 1996). This detection causes the phagosomes within the macrophages to fuse with lysosomes resulting in the removal of the bacteria (Saunders and Cooper, 2000). However, the rate of elimination process is not well understood and may be very minuscule as complete elimination of the bacilli is very difficult to get (Russell, 2007).

The latent TB infection can be considered as the second scenario. The bacteria alter the intracellular environment of the phagosome by protecting its acidification, thus keeping aside normal maturation and phagocytosis (Saunders and Cooper, 2000; Riska and Carleton, 2002). The microbe continues to divide in the macrophage till the macrophage lyases. This causes the macrophages in the surrounding tissue to phagocytose the bacilli and this continues as a different cycles. It also causes the human immune system to react and try to phase out the bacteria as the lysis of the macrophage releases mycobacterial antigen. The fight between the immune system trying to eliminate the bacteria and the bacteria endeavoring to survive creates a caseating tubercle (Saunders and Cooper, 2000). This ultimately leads to the formation of granuloma and after formation of the granuloma, in a phase wise manner the bacteria is walled off from the effect of the immune system and it is termed as latent TB infection where there is no display of any clinically valid symptoms (Riska and Carleton, 2002; Flynn and Chan, 2001). The host system can contain the latent TB infection for the entire life and in case of any weakness of the immune system can lead to reactivation of tuberculosis. Chance of clinical manifestations is thus greatly increased in case of HIV infected patients (Fig. 2.2).

The third situation that can arise is infection followed by active disorder. The bacterium divides in the lungs, propagates to the lymph nodes and can get mixed into the bloodstream to infect the rest of the body (Kaufmann, 2001). Spreading of the bacilli in the blood can cause infection spreading to more organs like the bones, kidneys, brain etc.
The pathogen, \textit{Mtb} enters into the host body via the inhalation of respiratory droplet contaminated with the bacterium. After entering into the lungs, the dynamic interaction in between the host and the microbe can have any of the following outcomes.

1) The first host response becomes completely effective and kills the bacilli;
2) The \textit{Mtb} can grow and divide immediately after infection, causing primary TB;
3) The pathogen may remain dormant and never cause disease; and
4) The latent bacilli can reactivate themselves eventually and progress to disease condition (Bishai, 2000).

Probable outcomes of the disease, as discussed (Cole \textit{et al.}, 1998) are shown in Fig. 2.4.

\textbf{Figure 2.2.} Schematic depiction of TB in human. This tuberculosis disease features numerous stages, having two distinct forms of progression to active disease - primary and post-primary. The lungs are infected by \textit{Mtb} which begins to grow inside the pulmonary macrophages (immune cells responsible for ingesting foreign materials) during acute infection (Bishai, 2000).
2.4 Genome of *Mycobacterium tuberculosis*

The genome of *Mtb*, one of the first complete genomes to be sequenced, was deciphered in 1998 by Cole and co-workers (Camus *et al.*, 2002) which was later re-annotated in 2002 by Camus and coworkers (Smith, 2003). The *Mtb* genome is composed of 4,411,529 base pairs, contains around 4,000 genes, and has a very good guanine+ cytosine content that is reflected in the biased amino-acid composition of the proteins (Fig. 1.5). There is a huge difference in the coding potential of *Mtb* and the rest of the bacteria as a high percentage of coding process is devoted to the synthesis of enzymes responsible for causing lipolysis and lipogenesis, and also to the production of two new glycine-rich proteins, bearing a repetitive structure, suggestive of a reason for antigenic variation. Unique characteristic features of the tubercle bacillus include its slow and gradual growth, dormancy, complicated cell envelope, intracellular pathogenesis and genetic homogeneity. The H37Rv strain of *Mtb* has been very popular worldwide in the field of biomedical research since its first isolation in 1905, owing to the possession of complete virulence in the animal models of TB. It is seen that in contrast to few clinical isolates, the strain is also susceptible to drugs and responds well to genetic engineering (Camus *et al.*, 2002).

*Mtb* is a Gram-positive bacterium with GC rich genome and its cell envelope contains an additional layer after the peptidoglycan that is very rich in unusual lipids, glycolipids and polysaccharides. Complex and noval biosynthetic pathways generate cell-wall components such as mycolic acids, mycocerosic acid, phenolthiocerol, lipoarabinomannan and arabinogalactan, and several of these may contribute to mycobacterial longevity, trigger inflammatory host reactions and act in pathogenesis (Camus *et al.*, 2002). The mycobacterial cell wall is well-defined and is linked to the pathogenicity of *Mtb* (Barry *et al.*, 1998; Dubnau et al., 2000; Glickman *et al.*, 2000; Crick, *et al.*, 2001). The three polymers present in the cell wall, arabinogalactan-mycolate (Takayama *et al.*, 2005), covalently linked with peptidoglycan and trehalose dimycolate, provide a thick layer that provides immunity to the tubercle bacillus from general antibiotics and the host’s immune system (Draper and Daffé, 2005). The synthesis of
mycolic acids, that are long chain $\alpha$-alkyl-$\beta$-hydroxy fatty acids, the main constituents of this protective layer, has been regarded to be critical for the survival of $Mtb$ (Pasqualoto et al., 2004). InhA, an enoyl-ACP reductase, responsible for mycolic acid synthesis, is an established target for front-line anti-tubercular drugs (Lei et al., 2000) such as isoniazid (Banerjee et al., 1994) and ethionamide (Pasqualoto and Ferreira, 2001). The mycolic acid pathway was therefore chosen as the one of the target for structure-based drug discovery study here in this exercise.

**Fig. 2.3** Genome of *Mycobacterium tuberculosis* (Cole and Barrell, 1998)

### 2.5 TB and AIDS

HIV pandemic is one of the key factors behind the resurgence of TB, apart from the emergence of drug-resistant strains. In patients infected with HIV, probability of developing TB increases up to 30 times as the immune system is compromised due to HIV infection (Ducati et al., 2006). Latent TB infections could be progressed to active disease stage in case of HIV infection. Moreover, by production of stimulatory cytokines TB induces the AIDS in HIV-positive individuals.
Therefore, the HIV-\textit{Mtb} co-infection is a serious concern both to infected patients and the global human population (Brennan, 1997). It is noted that several problems such as resistance and HIV co-infection are highly associated with TB. For example, due to the high abandon rate of treatment by co-infected patients, there is a higher emergence of drug-resistant TB strains (WHO, 2002). Such opportunistic infections have a direct impact on the mortality rate in infected patients.

\section*{2.6 Current status of anti-tubercular drugs and MDR and XDR form of TB}

Although several drugs and the Bacillus Calmette-Guérin (BCG) vaccine is available, TB still exists a serious health matter throughout the world, and it justifies the requirement of discovery of new drug targets for the design of more effective drugs. TB is declared as a global health emergency in 1993 by WHO, because of its prevalence in different parts of the globe (Zhang, 2005).

Drugs that are currently used against Tuberculosis are either derived from antibiotics or are chemical derivatives. Streptomycin is regarded as the first effective TB drug that was isolated from \textit{Streptomyces griseus} by Albert Schatz and Selman Waksman in 1944, which led to the initiation of modern TB chemotherapy (Janin, 2007). The drugs currently in use also draws its origin to the sulpha drugs, developed to treat infections caused by Gram-positive bacteria. In 1938, sulphanilamide was reported to inhibit TB growth in pigs, leading to efforts to refine the sulpha drugs for TB treatment, and the synthesis of thiosemicarbazones, which were more effective than sulphanilamide, but were less effective, compared to streptomycin (Janin, 2007). Rooting from an observation that salicylate and benzoate stimulated oxygen consumption of \textit{Mtb}, in 1945, a new anti-TB drug, \textit{p}-aminosalicylic acid was discovered.

In 1952, isoniazid was discovered and it was regarded as major breakthrough in the treatment of tuberculosis. The discovery of isoniazid was on the basis of the nicotinamide activity against \textit{M. tuberculosis} in an animal model, and the reshuffling of chemical groups in thiosemicarbazone. In 1952 nicotinamide lead also led to the
development and discovery of the another new entity called pyrazinamide. The identification of the activity of polyamine and diamine against \textit{Mtb} and the synthesis of diamine analogs in 1961 showed the path to the discovery of ethambutol. Further analysis and screening for antibiotics from soil microbes led the way to discovery of many other anti-tuberculosis drugs, such as kanamycin and its derivative amikacin, cycloserine, capreomycin, viomycin, and rifamycins and its derivative rifampicin which is another effective and popular drug for TB (Janin, 2007). It can be mentioned that most of the TB drugs currently in use were developed and discovered during the 1950s to 1960s. Furthermore, broad-spectrum quinolones were developed in 1980s on the basis of the anti-bacterial activity of nalidixic acid which was discovered in the 1960s. The quinolones, were subsequently shown to have high activity against \textit{Mtb} and have been second-line drugs for the treatment of drug-resistant TB since the late 1980s (Janin, 2007). Thus, most of the TB drugs currently being used have been discovered by a combination of serendipity and novel chemical modifications of an available lead compound. Given that most of these discoveries were made years ago, there is a definite need for applying newer strategies and technologies for discovery to address the serious threat generated by TB and the increase of drug resistant strains.
Fig. 2.4 Multi-drug resistant tuberculosis (MDR-TB) is a category of tuberculosis resistant to at least two principal first line drugs rifampicin and isoniazid. (Source: National Institute of Allergy and Infectious Diseases)

Currently, about twenty drugs are available for treating TB, out of which four drugs, *viz.* rifampicin, isoniazid, pyrazinamide and ethambutol, are being used as first-line drugs. Injectable forms of drugs such as amikacin, kanamycin, viomycin and capreomycin are preferred next for treatment. For the treatment of MDR-TB, Fluoroquinolones such as ciprofloxacin and ofloxacin have been found to be very useful. Second-line drugs such as *p*-aminosalicylic acid, ethionamide and cycloserine, have established better clinical efficacy but also reported to have prominent side effects (Feng and Barletta, 2003). Ethionamide and isoniazid are found to be inhibitors of mycolic acid biosynthesis (Banerjee *et al.*, 1994; Pasqualoto and Ferreira, 2001), where as cycloserine and ethambutol blocks synthesis of peptidoglycan (Deng *et al.*, 1995) and cell wall arabinogalactan (Belanger *et al.*, 1996; Telenti *et al.*, 1993) respectively, weakening the cell wall of the bacterium (Table 2.1). Rifampicin and amikacin exert and their
pharmacological action kill *Mtb* by inhibiting bacterial RNA or the process of protein translation (Busscher *et al*., 2005; Maus *et al*., 2005; Zhang and Amzel, 2002).

An extensive list of TB drugs could be found in the work of Fang and Barletta, (2003). Much detailed discussions of TB drugs and its targets have also been published and were studied during designing of the experiment (Janin, 2007; Zhang, 2007). There is a list of new chemical entities currently undergoing clinical trial and waiting for the final approval as new anti-TB drugs (Haile and Källenius, 2005). Vaccines for TB BCG is presently the only available vaccine against TB, and is widely administered within the WHO monitored programme for immunisation. The immune protective efficacy of BCG has been shown to be highly variable across different populations (Young and Dye, 2006). The BCG vaccine is widely used in many parts of the world to protect against severe childhood forms of TB but is comparatively unreliable against the highly infectious pulmonary form of TB characteristically found in adults (Gupta *et al*., 2007). Also, the duration of protection is found to be variable and it also does not appear to check the transmission of TB, which is considered to be another serious shortcoming (Gupta *et al*., 2007; Young and Dye, 2006). The current status of TB vaccines and the future prospects are reviewed in (Hoft, 2008; WHO, 2007). Modern therapy for TB depends on a combination of various potent anti-bacterials, such as rifampicin, isoniazid, and pyrazinamide, in a treatment of six months’ tenure. It may be necessary to add other first-line drugs such as ethambutol and streptomycin to the treatment, if initial resistance to isoniazid is observed. Whenever there is resistance to at least rifampicin and isoniazid, characterised as MDR-TB, it necessitates the extension of the treatment period and frequently, the use of second- and third-line drugs, despite their higher toxicity (Brennan, 1997).
<table>
<thead>
<tr>
<th>Drugs</th>
<th>Targets</th>
<th>Mechanism(s) of action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amikacin/Capreomycin</td>
<td>rrs (16S rRNA)</td>
<td>Inhibition of protein synthesis</td>
</tr>
<tr>
<td>Kanamycin/Viomycin</td>
<td>AlrA (rv3423c; D-alanine racemase), Ddl (Rv2981c; D-alanylalanine synthetase)</td>
<td>Inhibition of peptidoglycan synthesis</td>
</tr>
<tr>
<td>Cycloserine</td>
<td>cyclomycin A/B (Rv3793/Rv3794/Rv3795; Arabinosyl transferases)</td>
<td>Inhibition of cell wall arabinogalactan synthesis</td>
</tr>
<tr>
<td>Ethambutol</td>
<td>EmbC/A/B (Rv3793/Rv3794/Rv3795; Arabinosyl transferases)</td>
<td>Inhibition of cell wall arabinogalactan synthesis</td>
</tr>
<tr>
<td>Ethionamide, prothionamide and thiacetazone</td>
<td>InhA (Rv1484; Acyl carrier protein reductase)</td>
<td>Inhibition of mycolic acid synthesis</td>
</tr>
<tr>
<td>Fluoroquinolones</td>
<td>GyrA (Rv0006; DNA gyrase)</td>
<td>Inhibition of DNA gyrase</td>
</tr>
<tr>
<td>Isoniazid</td>
<td>InhA (Rv1484; acyl carrier protein reductase), KasA (Rv2245; β-ketoacyl synthase)</td>
<td>Inhibition of cell wall mycolic acid synthesis, and other potential multiple effects on DNA, lipids, carbohydrates, and NAD metabolism</td>
</tr>
<tr>
<td>p-aminosalicylic acid</td>
<td>unknown</td>
<td>Inhibition of folic acid synthesis and iron uptake</td>
</tr>
<tr>
<td>Pyrazinamide</td>
<td>Membrane function and energy metabolism, Fas (Rv2524c; Fatty Acid Synthase)</td>
<td>Disruption of membrane function and energy metabolism, Inhibition of fatty acid synthesis</td>
</tr>
<tr>
<td>Rifampicin and rifapentin</td>
<td>RpoB (Rv0667; RNA polymerase β-subunit)</td>
<td>Inhibition of RNA synthesis</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>RpsL (Rv0682; Ribosomal S12 protein), rrs (16S rRNA)</td>
<td>Inhibition of protein synthesis</td>
</tr>
<tr>
<td>Thiolactomycin</td>
<td>FabH (Rv0533c; β-ketoacyl synthase III), KasA (Rv2245; β-ketoacyl synthase), KasB (Rv2246; β-ketoacyl synthase)</td>
<td>Inhibition of fatty acid synthesis</td>
</tr>
</tbody>
</table>

Table 2.1 Current drugs for TB and their targets (Anishetty et al., 2005)

2.7 The emergence of “extensively drug resistant” TB (XDR-TB)

XDR-TB defined as TB caused by strains resistant to isoniazid, rifampicin, and at least three main classes of second-line drugs, has been reported more recently (Fig. 2.5). XDR-TB is widespread, including occurrence in the developed countries such as the USA, where TB had been considered to be under control. The worldwide occurrence of XDR-TB raises the prospect of almost incurable TB (Brennan, 1997; Belanger et al., 1996).
Fig. 2.5 XDR-TB is defined as TB that is repellent to any fluoroquinolone, and at least one of three injectable second-line drugs (capreomycin, kanamycin, and amikacin), besides isoniazid and rifampicin (Source: National Institute of Allergy and Infectious Diseases; https://www.niaid.nih.gov/)

The frontline drugs show early bactericidal activity against actively metabolising bacilli, whereas the second-line bacteriostatics and restrained to increase the efficacy of the treatment in the presence of resistance. Among all the first-line anti-tuberculars, isoniazid shows a substantial bactericidal activity against the micro-organisms growing actively in cavities, followed by other drugs like rifampicin, streptomycin and quinolones. However, isoniazid is a drug with excessive toxicity and consumption of this drug results in frequent fever (Brennan, 1997).
2.8 Current TB therapy (DOTS)

The directly observed treatment, short course (DOTS) therapy is the first of its kind anti-tubercular therapy, approved by WHO for treating every TB patient. The initial phase includes/requires treatment with four drugs, viz. isoniazid, rifampicin, pyrazinamide and ethambutol, for two months daily, followed by a continued phase of treatment with isoniazid and rifampicin for another four months, three times a week. DOTS have a cure rate of up to 95%, although conditional on patient compliance (Zhang, 2007). This treatment focuses originally at a bacteriostatic action, preventing the cell wall synthesis, nucleic acids and mycobacterial proteins and thereby, leading to a fast removal of most infecting bacilli. The therapy also aims at a subsequent bactericidal action to strengthen the treatment through the elimination of all remaining bacilli. DOTS combine five foundational elements: political commitment, microscopic services, drug supplies, surveillance systems and direct treatment observation (Brennan, 1997). This strategy stops the development of newer infections and minimizes the emergence of MDR- and XDR-TB. In 2006, WHO launched the new Stop TB strategy (http://www.stoptb.org/). The uniqueness of this strategy is DOTS, the TB control approach launched by WHO in 1995. Since its launch, DOTS-based services have treated as many as 22 million patients under it. The new strategy lies on this victory, while identifying the significant hurdles of TB/HIV and MDR-TB. It also addresses access, equity and quality constraints, which follows evidence-based innovations to engage with private health-care care takers, empowering afflicted people and communities thereby helping in reinforcing health systems and promoting research (Johnson et al., 2006).

Drawbacks of the current therapy

Even though DOTS can cure TB, the prolonged six month therapy subsequently results in patient non-compatibility, a major source of drug-resistant strains. Although TB chemotherapy contributes in patients being non-infectious for a few weeks after starting of therapy, the therapy has to be continued for a considerable period to prevent reoccurrence. Several factors may be made responsible for this lengthy TB chemotherapy,
particularly the phenotypic resistance in non-replicating persisters (Janin, 2007). Persistence is one of the critical and complex aspects of TB, posing a serious challenge to the effectiveness of anti-tubercular drugs. Antibiotics are active only against growing bacteria, whereas they are ineffective against growth dormant bacteria. There are at least three types of non-growing bacteria that exhibit phenotypic resistance to antibiotics: (a) bacteria in the stationary phase, (b) bacteria in the dormant phase and (c) residual survivors or persisters not killed during antibiotic exposure when a growing culture is treated (Janin, 2007). The mechanism of phenotypic resistance in \textit{Mtb} is unknown. Another appalling fact is that most of the drugs used to treat TB have been discovered between 1950 and 1970, close to about forty years ago. The detrimental aspects of co-infection with HIV have also been extensively studied by the researchers. All the above factors, particularly the emergence of MDR-TB and XDR-TB, which are spreading to many countries and creating a major threat to TB eradication programmes (WHO, 2006; Pozzan, A. 2006), and creates urgent need for the exploration of newer and more rational strategies for the identification of newer drugs and drug targets for TB. The major need is to identify more rational targets, or to produce more effective drugs, which lead to a shortening of the therapy for treatment, as well as check the emergence of resistance, in tandem.

\textbf{2.9 Structure-based drug designing and its applications in tuberculosis drug discovery}

\textbf{2.9.1 Computer-Aided Drug Designing}

It is widely accepted drug discovery and development are time and resources intensive processes. There is an ever increasing effort to apply computational tool to the combined chemical and biological space in order to make drug discovery, design, development and its relevant optimization processes flexible. Use of computational techniques in drug discovery and development process is rapidly gaining popularity and universally adopted in implementation and appreciation. Different terms are being applied to this area, including computer-aided drug design (CADD), computational drug design, computer-
aided molecular design (CAMD), computer-aided molecular modeling (CAMM), rational drug design, *in silico* drug design, computer-aided rational drug design etc. The term Computer-Aided Drug Discovery and Development (CADDD) is more commonly used to give overview of the area and to cover the entire process. Both the computational and experimental techniques have significant roles in drug discovery and development and represent complementary approaches. CADDD entails the following

1. Use of power of computing to streamline drug discovery and development process
2. Leverage of biological and chemical information about ligands and targets to identify and optimize new drugs
3. Development of computational filters to eliminate compounds with undesirable physicochemical and pharmacokinetic properties and select the most promising candidates (Green, 2003).

Extensive progress in the area of CADDD has been made possible by advance level development in software and hardware computational tool and sophistication, identification of molecular targets, and an increasing database of publicly available target protein structures. CADDD is being applied to identify hits (active drug candidates), select lead molecules (most likely candidates for further evaluation), and optimize leads i.e. transform biologically active compounds into suitable drugs by improving their physicochemical, pharmaceutical, ADMET/PK (pharmacokinetic) properties. New drugs are discovered by virtual screening which classifies drug candidates from different chemical scaffolds by searching commercial, public, or private chemical structure databases. It is intended to minimize the size of chemical space and thereby allow focus on more promising candidates for lead discovery and optimization. The major goal is to enrich set of molecules with desirable properties (active, drug-like, lead-like) and to eliminate compounds with unfavorable properties. In another words, *in silico* modeling is used to significantly minimize time and resource requirements of chemical synthesis and biological testing. Rapid growth in the area of virtual screening is also felt as the number of citations in this area is very quickly increasing (Green, 2003). In a review article published in 2003, Green of GlaxoSmithKline mentioned that: “The future is bright. The
future is virtual”. Commonly used computational approaches include ligand-based drug design (pharmacophore, a 3-D spatial arrangement of chemical features essential for biological activity), structure-based drug design (drug-target docking), and quantitative structure-activity and quantitative structure-property relationships (Wang and Eisenberg, 2006). The pharmaceutical industry are aggressively involved in development of computational tools that will improve effectiveness and efficiency of drug discovery and development process, decrease use of animals, and increase predictability. In this study, the structure-based drug designing approach was used for screening of the compounds.

2.9.2 Docking

Molecular docking calculations predict the structure of the intermolecular complex formed between two or more constituent molecules. In the field of molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a energetically stable complex (Meng et al., 2004). Molecular docking screens large databases of small molecules by orienting and scoring them in the active site of the protein. Knowledge of the preferred orientation in turn may be used to predict the strength of association or binding affinity between two molecules using for example scoring functions.

2.9.3 Application of docking

Various applications of the docking can be mentioned as below.

a) Virtual screening (hit identification)

b) Drug Discovery (lead optimisation)

c) Binding-site identification (blind docking)

d) De-orphaning of a receptor

e) Protein – Protein (or Protein – Nucleic Acid) interactions

f) Structure-function studies

g) Enzymatic reactions mechanisms

h) Protein engineering
There are two approaches used for the molecular docking. One approach uses a matching technique that describes the protein and the ligand as complementary surfaces (Morris et al., 1998). The second approach simulates the actual docking process in which the ligand-protein pair wise interaction energies are calculated (Irwin, et al., 2002). Both approaches have significant advantages as well as some limitations.

There are many problems exist in the currently implemented molecular docking procedures. The conformational and the orientation sampling of putative drug in the protein site are crude. The model of the protein site itself even cruder, often, the receptor is kept completely rigid. The scoring function used to rank good ligands above poor compounds is problematic too, the net energy of binding is small difference of large values of large uncertainties and the calculation of desolating the ligand when it binds to the protein is complex to calculate. Despite these considerable problems, molecular docking has been used successfully in different times to identify promising lead compounds that may be used to begin a process of lead optimization towards a drug candidate (Feig et al., 2004). Two main problems of docking are as follows:

2.9.3.1 Posing

This is the process of determining whether a given conformation and orientation of a ligand fits the active site. This procedure usually responds to many alternative results.

2.9.3.2 Scoring

The pose score is a measure of the fit of a ligand into the active site. Scoring during the posing phase usually involves energy calculations (van der Waals, ligand strain, electrostatic etc.). Further re-scoring may attempt to estimate more accurately the free energy of binding (G, and therefore Kₐ) perhaps including properties such as entropy and solvation.

\[
[EI]_{aq} \rightleftharpoons [E]_{aq} + [I]_{aq}
\] ........................(1)
The free energy of binding ($G$) is related to binding affinity by the following equations:

$$
\Delta G = -RT \ln K_A
$$

$$
K_A = K^{-1}_A = \frac{[EI]}{[E][I]}
$$

Prediction of the correct structure (posing) of the $[E+I]$ complex does not require information about $K_A$. However, prediction of biological activity requires this information; scoring terms can therefore be divided in the following order. When considering the term $[EI]$, the following factors are important: steric, electrostatic, hydrogen bonding, inhibitor strain (if flexible) and enzyme strain. When considering the equilibrium shown in equation 1, the following factors are also important: desolvation, rotational entropy and translational entropy.
2.9.3.3 Scoring function in AutoDock

Molecular mechanics terms of AutoDock are mentioned below.

\[ \Delta G_{vdW} = W_{vdW} \sum_{i,j} (A_{ij} / r_{ij}^{12} - B_{ij} / r_{ij}^{6}) \]

\[ \Delta G_{H-bond} = W_{H-bond} \sum_{i,j} E(t) \left( C_{ij} / r_{ij}^{12} - D_{ij} / r_{ij}^{10} + E_{hbond} \right) \]

\[ \Delta G_{elec} = W_{elec} \sum_{i,j} \frac{q_i \cdot q_j}{r_{ij}^{2}} \]

\[ \Delta G_{desolv} = W_{desolv} \sum_{i,j} \left( S_{ij} \cdot V_{ij} \cdot \exp\left(-\frac{r_{ij}^{2}}{2 \cdot \sigma^{2}}\right) \right) \]

Change in Torsional Free Energy when the Ligand goes from Unbound to Bound

Torsional \( \Delta G_{tor} = W_{tor} \cdot N_{tor} \)

2.9.3.4 Limitations of current docking methodologies

Flexible ligands \( \rightarrow \) Rotatable bonds \( \rightarrow \) Combinatorial explosion

Entropic effects \( \rightarrow \) Rotatable bonds

Solvation / desolvation \( \rightarrow \) Accurate computation is expensive

Water molecules (and ions)

Tautomers

Protein flexibility \( \rightarrow \) Induced fit

Specificity of binding \( \rightarrow \) Understanding important interactions

(Currently larger ligands are favoured by the scoring functions)

Pharmacokinetic effects, allosteric effects, biomolecule-biomolecule interactions (molecular context), etc. (Marrakchi et al., 2000).

2.9.3.5 Analysis of docking result and short-listing of the favorable compounds

Drug discovery projects are incomplete without analysis of the drug like properties. As a standard, the following properties are calculated in this study so that the compounds could be selected.
structural properties
- Hydrogen bonding
- Polar surface area
- Lipophilicity
- Shape
- Molecular weight
- Reactivity
- pKa

physicochemical properties
- Solubility
- Permeability
- Chemical stability

biochemical properties
- Metabolism (phases I and II)
- Protein and tissue binding
- Transport (uptake, efflux)

Pharmacokinetics (PK) and toxicity
- Clearance
- Half-life
- Bioavailability
- Drug–drug interaction
- LD50

2.10 Identification and selection of drug targets

Out of several established drug targets of *M. tuberculosis*, four protein targets are discussed below and also selected for the current study based on literatures available.

2.10.1 Pantothenate synthetase

Pantothenate biosynthesis is necessary for the virulence of *Mtb*, and the pathway presents potential drug targets against tuberculosis. The pantothenate biosynthetic pathway is best characterized in *E. coli*. It involves four steps catalyzed by enzymes encoded by panB, panC, panD, and panE genes (Merkel and Nichols, 1996). PanC encodes a pantothenate synthetase (PS) (EC 6.3.2.1), which catalyzes the last step of pantothenate biosynthesis, the ATP-dependent condensation of pantoate, and β-alanine to form pantothenate. PS catalyzes the ATP-dependent condensation of pantoate and β-alanine to form pantothenate. The crystal structure of PS from *M. tuberculosis* and its complexes with AMPCPP, pantoate, and pantoyl adenylate was studied thoroughly. It has been found that this enzyme complex with AMP and its last substrate, β-alanine, and show that the phosphate group of AMP serves as an anchor for the binding of β-alanine. This structure confirms that binding of β-alanine in the active site cavity can occur only
after formation of the pantoyl adenylate intermediate (Hasan et al., 2006). Pantothenate (Vitamin B5) is a key precursor for the biosynthesis of coenzyme A (CoA) and acyl carrier protein (ACP). Both of these are necessary cofactors for cell growth and are involved in essential biosynthetic pathways. The pantothenate pathway is not present in mammals and consequently represents an exciting target for the development of novel antibiotics and herbicides. The pathway is best understood in *E. coli*, where it comprises four enzymatic reactions (Fig. 2.6).

![Fig. 2.6](image)

**Fig. 2.6** The biosynthesis of pantothenate in bacteria, yeast and plants. Pantothenate synthetase is the last enzyme in the pathway, it catalyses the condensation of pantoate and β-alanine in the presence of ATP to give pantothenate. The overall reaction consists of two sequential steps, initial formation of a cofactor-substrate intermediate, followed by subsequent nucleophilic attack on the activated carbonyl by β-alanine (Lengauer and Rarey, 1996).
The whole metabolic pathway of the synthesis of pantothenate and the CoA is as follows (Fig. 2.7).

**Fig. 2.7** Metabolic pathway of pantothenate and coA synthesis of *Mycobacterium tuberculosis* ([http://www.genome.jp/kegg/pathway/map/map00770.html](http://www.genome.jp/kegg/pathway/map/map00770.html))

2.10.2 InhA

One such well known target is 2-*trans*-enoyl-acyl carrier protein reductase, called InhA which is acted upon by the first line anti-tuberculosis drug isoniazid. InhA displays a long-chain fatty acid elongation activity with the characteristic properties described for the FAS-II (fatty acid synthetase II) system. Inhibition of this activity by InhA inhibitors, namely isoniazid, octadecynoyl-CoA or hexadecynoyl-CoA, showed that InhA belongs to the FAS-II system. Moreover, the InhA inhibitors also blocked the biosynthesis of mycolic acids, which are major lipids of the mycobacterial envelope. The data suggest that isoniazid acts on the mycobacterial cell wall by preventing the FAS-II system from producing long-chain fatty acid precursors for mycolic acid biosynthesis.
2.10.3 Mycolic Acid Cyclopropane Synthase (Cma2)

Mycobacteria have a different type of cell wall in which mycolic acids play a critical role in its structure and function. Important characteristics conferred by this structure are resistance to chemical injury, low permeability to antibiotics, resistance to dehydration, and ability to thrive within the hostile environment of the macrophage phagolysosome (Banerjee et al., 1994). Mycolic acids are major components of the cell wall of *Mycobacterium tuberculosis*, and have long chain α-alkyl, β-hydroxy fatty acids containing 70–90 carbons in total (Feng and Barletta, 2003; Deng et al., 1995; WHO, 2015). The α-branch is a saturated alkyl chain, typically 24 carbons long. The meromycolate chain contains 40–60 carbons with two positions of modification, the distal position (close to the ω-end) and the proximal position (close to the β-hydroxy end). There are three classes of mycolic acids synthesized by *M. tuberculosis* called α-, keto-, and methoxymycolates, that are classified according to their modifications at these two positions. The α-mycolates have a cis cyclopropane ring at both positions, whereas keto and methoxymycolates have oxygenated groups at the distal position and a cyclopropane ring at the proximal position that can be in either a cis or trans conformation. There are at least three mycolic acid cyclopropane synthases (*PcaA, CmaA1*, and *CmaA2*) that are responsible for these site-specific modifications of mycolic acids. *CmaA2* have a seven-stranded α/β fold similar to other methyltransferases with the location and interactions with the cofactor S-adenosyl-L-methionine conserved. The structures of the ternary complexes demonstrate the position of the mycolic acid substrate binding site. Close examination of the active site reveals electron densities that are believed to represent a bicarbonate ion. The structures support the hypothesis that these enzymes catalyze methyl transfer via a carbocation mechanism in which the bicarbonate ion acts as a general base. In addition, comparison of the enzyme structures reveals a possible mechanism for substrate specificity (WHO, 2015). Therefore, the structure can provide a foundation for rational-drug design, which may lead to the development of new inhibitors effective against persistent bacteria.
2.10.4 Glutamine synthetase

Glutamine and glutamate are central molecules in nitrogen metabolism. Glutamine is used as the nitrogen donor for many nitrogen-containing molecules in the cell and is synthesized from L-glutamate, ammonia, and ATP by the enzyme glutamine synthetase (GS).

2.11 Strategic selection of the natural compounds for potential inhibitor of \textit{Mtb} drug targets

Proper selection of plant-based natural compounds against is an important parameter and based on information available in traditional medicine, following plants can be targeted to screen the potential natural inhibitor of \textit{Mtb} i.e., Neem (\textit{Azadirachta indica}), Tea (\textit{Camellia sinensis}) and \textit{Ginkgo biloba}. However, antibacterial compounds already existed in the DrugBank or similar databases (http://www.drugbank.ca/) can be used as reference molecules.

Neem is traditionally known to be effective as antibacterial agent during this multitude of uses in treatment of human health disorders. Due to the antibacterial properties of the compounds available in the leaf extracts of \textit{Azadirachta indica}, it was decided to use the compounds for searching potent anti-mycobacterial agents. Similarly Tea polyphenols especially Green Tea Polyphenols are also known to have antibacterial properties. Here in this study Green Tea Polyphenols which are commonly known for antibacterial activities were used to search potential lead molecule for anti-tuberculosis drug discovery. The \textit{Ginkgo biloba} has tremendous medicinal, spiritual, and horticultural importance in Chinese culture. The supplements are bestselling herbal medications with a long history of use in traditional medicine to treat blood disorders; these are known to improve memory and offer the best-known way to keep the mind sharp. Leaves and seeds of \textit{G. biloba} have been used in Chinese herbal medicine for thousands of years including treatment of tuberculosis (Isah, 2015).
DrugBank (www.drugbank.ca) is a richly annotated resource that combines detailed drug data with comprehensive drug target and drug action information. Since its first release in 2006, DrugBank has been widely used to facilitate *in silico* drug target discovery, drug design, drug docking or screening, drug metabolism prediction, drug interaction prediction and general pharmaceutical education (Wishart *et al.*, 2007).