Molecular modeling of 2-nitropropane dioxygenase domain of *Mycobacterium tuberculosis* H37Rv and docking of herbal ligands

K V Ramesh, B N Akhila and Sudha Deshmukh*

Department of Biotechnology, Center for Postgraduate Studies, Jain University, 18/3, 9th Main, Jayanagar III Block, Bangalore 560011, India

Received 28 September 2010; revised 10 May 2011

The 3D structure of enoyl reductase (ER) domain generated by the SWISS MODEL server contains the 2-nitropropane dioxygenase (2NPD) structure displaying the TIM barrel fold. Though TIM barrel fold is made up of both main and inserted domains, in our study, we could only predict the structure of the main domain, which had central barrel of eight β-strands surrounded by eight α-helices. Superimposition of the 2NPD region of ER domain of *Mycobacterium tuberculosis* H37Rv on to the corresponding region of 2UVA_G revealed a good structural alignment between the two, suggesting this template to be a good structural homologue. Among various herbal ligands that were screened as inhibitors, daucosterol was found to bind in closest proximity to the flavin mono nucleotide (FMN) binding site with the lowest docking energy.

**Keywords:** DOCK, Enoyl reductase domain, Fatty acid synthase, 2-Nitropropane dioxygenase, SWISS MODEL, *Mycobacterium tuberculosis* H37Rv

*Author for correspondence*

E-mail: sudhadeshmukh@yahoo.com

Tel: 080-43226508
Fax: 080-43226507

**Abbreviations:** ER, enoyl reductase; FAS, fatty acid synthase; FMN, flavin mono nucleotide; MDR, multidrug-resistant; 2NP, 2-nitropropane; 2NPD, 2-nitropropane dioxygenase.

*Mycobacterium tuberculosis* remains one of mankind’s deadliest pathogens responsible for approximately 2 million deaths worldwide every year and estimated to infect one-third of world’s population. The major factors responsible for the explosion in numbers of infections with *M. tuberculosis* are the deadly synergy with HIV, followed by development and spread of multidrug-resistant (MDR) strains of *M. tuberculosis*. The problem of MDR is compounded by inadequate drug supplies and poor TB control programmes, particularly in the third world.

Fatty acid biosynthesis in *Mycobacterium* is mediated by fatty acid synthase (FAS-I and -II). While FAS-II is a collection of individual enzymes, FAS-I is a polypeptide with multiple active sites that perform all the catalytic reactions in the pathway. As FAS-II are absent in humans, they are considered as valuable targets for drug development. The crystal structure of FAS-II protein of *Thermomyces lanuginosus* (PDB: 2UVA) reveals it a heterodacamic protein (α6β6) with enoyl reductase (ER) as one of the domains of its β polypeptide chain.

The increasing emergence of drug resistant TB and HIV infection, which compromises host defense and allows latent infection to reactivate or render individuals more susceptible to TB, pose further challenges for effective control of the disease. Although TB can be cured with chemotherapy, the treatment is exceedingly lengthy. Apart from significant toxicity, the lengthy therapy also creates poor patient compliance, which is a frequent cause for emergence of drug-resistant and often deadly multi-drug resistant TB (MDR-TB) bacteria. Currently, TB chemotherapy involves first-line drugs: isoniazid, rifampin, pyrazinamide and ethambutol. If the treatment fails as a result of bacterial drug resistance or intolerance to one or more drugs, second-line drugs are used, such as para-aminosalicylate, kanamycin, fluoroquinolones, capreomycin, ethionamide and cycloserine. Treatment is made quite difficult by the presence of metabolically silent, persistent or dormant bacteria within host lesions, which are not responsive to the anti-mycobacterial drugs that usually kill growing bacteria but not persistent bacteria.

The increasing problem of MDR-TB has focused attention on developing new drugs that are not only active against drug resistant TB, but also shorten the therapy. Although there is an urgent need and
significant interest in developing new TB drugs\textsuperscript{16-18}, no new class of TB drugs has been introduced in the market in the past 40 years. Identification of appropriate targets in the \textit{Tubercle bacillus} is a key step in the process of drug development and several studies are directed on identifying these targets\textsuperscript{18-22}.

The 3D structures of several \textit{M. tuberculosis} proteins and their role as potential drug targets have been exhaustively reviewed\textsuperscript{7}. There are several reports, which signify the importance of herbal ligands as prospective anti-microbial drugs, and few of them such as flavonoids\textsuperscript{23}, green tea\textsuperscript{24}, neem extracts\textsuperscript{25} are known to have inhibitory effects on different domains of FAS protein.

Structural information of an enzyme is very crucial in designing inhibitors. Therefore, in the absence of experimental data on the structure of FAS protein from \textit{M. tuberculosis} H37Rv, model building on the basis of known 3D structures of homologous protein is, at present, the most reliable method of obtaining structural information. Amidst various domains identified by the conserved domain database of NCBI, 3D structures of 5 different domains of FAS protein from \textit{M. tuberculosis} H37Rv have been predicted by us\textsuperscript{26}, which included 2NPD domain which is a part of ER domain. A further refinement of β-ketoacyl-ACP synthase domain using genetic algorithm operators and docking studies with a variety of herbal ligands has suggested eloemodin and nimbin as the best drug candidate against this target\textsuperscript{27}. The 2NPD domain, also a potential drug target\textsuperscript{13} has shown poor sequence identity with the best available template 2BMO_A in our earlier study\textsuperscript{26}.

In the present study, an attempt has been made to rebuild the 3D structure of the ER domain using a newly available template 2UV8 deposited in the PDB database, which gave a better sequence alignment. Putative drug binding sites have also been predicted using the DeepView package. A variety of herbal ligands deposited in Pubchem database were docked on to the homology modeled structure of 2NPD with a view to explore the best drug candidate.

\textbf{Materials and Methods}

\textit{3D Structure prediction}

Since 2-nitropropane dioxygenase (510-662) is found within the enoyl reductase (ER) domain (residue 354 to 1090) of FAS protein from \textit{M. tuberculosis} H37Rv, sequence boundary ranging from 141 to 1051 was used for model building in the present study. Additional residues (i.e., 141 to 353) were included anticipating active site residues to be present in these regions. The selection of appropriate PDB template for model building was done using the PSI-BLAST tool with the threshold E value set at 0.007 and setting remaining parameters to the default value\textsuperscript{28}. The sequences of few selected PDB templates producing significant alignment were extracted from the PDB databank. Multiple sequence alignment between 2NPD (with sequence boundary 510 to 662) and other selected PDB template sequences were generated using CLUSTAL W with all parameters set at default\textsuperscript{29}.

The ‘project mode’ of SWISS MODEL server\textsuperscript{30} was used for submitting the binary pdb file that was generated through DeepView package. This file consisted of primary sequence of ER threaded on to the PDB template structure 2UVA_G. Based on ERRAT\textsuperscript{31} analysis for the initial model, loop regions of the 3D models were refined using MODLOOP server\textsuperscript{32}. The model quality factor for the loop-refined structure was checked again using ERRAT server. Energy minimization was performed on the final loop refined model using DeepView package\textsuperscript{33}. The predicted ER model was evaluated using additional structure assessment tools like PROCHECK\textsuperscript{34}, which gives the G-value for the model based on Ramachandran plot analysis\textsuperscript{35}.

\textit{Docking studies}

On the basis of antibacterial/anti-tubercular activity, 57 herbal ligands and 15 ligands of neem origin were selected from the Pubchem database of NCBI. Further, drug similar ligands to the flavin mono-nucleotide (FMN) co-factor were selected randomly from the PDB databank and their structures were drawn using ISIS draw software\textsuperscript{36}. While the 3D structures of FMN and the substrate 2-nitropropane (2NP) were separated out from the PDB template 2UVA\textsuperscript{12}, chemical structures of these herbal ligands were drawn using ISIS Draw software package (http://www.mdli.com/) and converted into their 3D form using RASMOL\textsuperscript{37}. The DOCK program\textsuperscript{38} was used to dock the substrates as well as the herbal ligands on to the homology modeled 2NPD structure.

\textbf{Results}

\textit{3D Structure prediction}

Of all the templates retrieved by PSI-BLAST search, 2UVA_G and 2UV8_G showed the best sequence homology with the ER domain as indicated by the low E-values (Table 1), with 2UVA_G having
a maximum of 41% sequence identity and a gap of just 1%. The residues gly58,63, his64, asp69, leu75, gly94, lys105, gly114, asp121, gly126, glu134, asp147 and asp152 were found to be totally conserved in all the templates used. Further, gly94,126 and thr127 which constitute the phosphate-binding motif required for FMN-binding pocket were found to be totally conserved in all the templates used. The ER model generated using SWISS MODEL server with 2UVA_G as the template had 32 helices and 29 sheets. The initial model validated using ERRAT server showed a quality factor of 46.51; upon loop refinement, the quality factor increased to 68.93. Energy minimization of the final model after loop refinement resulted in a decrease of total energy from -24,336.18 to -27,383.73 kJ/mol. Validation of ER model by PROCHECK showed that 81.2% of the residues were in the core region of Ramachandran plot, followed by 17.4% in allowed region, 1.0% in generously allowed region and 0.4% in disallowed region. The overall G factor for the 2NPD model was -0.21 as against a value of -0.01 for 2UVA_G. Table 1—Summary of the best PDB templates generated at the end of the 20th iteration of PSI-BLAST analysis using 2 NPD domain of fatty acid synthase (NP_217040) from M. tuberculosis H37Rv (sequence boundary: 510-662)

<table>
<thead>
<tr>
<th>2-NPD region of FAS protein (M. tb H37Rv)</th>
<th>Template sequence ID</th>
<th>Region of template sequence aligned</th>
<th>Percent identity</th>
<th>Gap percent</th>
<th>E-Value</th>
<th>Annotation of the template</th>
</tr>
</thead>
<tbody>
<tr>
<td>510-662</td>
<td>*2UVA_G</td>
<td>672-821</td>
<td>41</td>
<td>1</td>
<td>5e-79</td>
<td>Crystal structure of fatty acid synthase from Thermomyces lanuginosus</td>
</tr>
<tr>
<td>510-662</td>
<td>*2UV8_G</td>
<td>675-823</td>
<td>38</td>
<td>1</td>
<td>2e-77</td>
<td>Crystal structure of yeast fatty acid synthase with stalled acyl carrier protein.</td>
</tr>
<tr>
<td>519-662</td>
<td>2GIN_A</td>
<td>93-229</td>
<td>19</td>
<td>12</td>
<td>0.44</td>
<td>Crystal structure of 2 NPD complexed with FMN and substrate.</td>
</tr>
</tbody>
</table>

* These two PDB templates (2UVA_G and 2UV8_G) had E-value better than the threshold of 0.07

Fig. 1—Multiple sequence alignment of 2NPD domain of M. tuberculosis H37Rv with selected PDB templates using CLUSTALW tool [(*) = completely conserved and identical residues, (,) and (.) = similar type of residues]

Superimposition of Cα backbone of ER and 2NPD models on to the corresponding region of 2UVA_G using DeepView package revealed that structural alignment was reasonably good. While the ER model showed root mean square deviation (RMSD) of 0.53 Å, 2NPD model had RMSD of 0.41 Å (Fig. 2a and b). FMN-binding pocket of 2NPD could be seen to be comprising of an (α/β)8 motif which is a folded barrel. The central barrel was made up of eight β strands (β1-8β) surrounded by eight α helices (α1-α8) (Fig. 3). The -NH2 terminal and -COO terminal residues for the FMN-binding pockets were met285 and asp522, respectively.

Docking studies

The substrate 2-nitropropane (2NP), the cofactor FMN, a few drug similar ligands and a variety of herbal ligands were successfully docked on to
the homology modeled 2NPD domain of FAS protein using DOCK program. While the residues involved in FMN binding pocket were gly^{432, 433}, his^{434}, asp^{439}, gly^{462, 464}, leu^{475}, asp^{491}, gly^{484, 496}, thr^{497}, gl^{504}, lys^{512} and asp^{517, 522}; ala^{374} to asp^{378}, lys^{400} to gly^{402}, glu^{427}, ala^{431} to ser^{436}, gly^{462}, leu^{494}, gly^{496} and thr^{497} were the residues found in the vicinity of the substrate, 2NP.

The docking energy of 2NP and FMN was -17.89 and -33.12 kcal/mol, respectively. Among the different neem based ligands that were docked, salannin^{39} (-38.58 kcal/mol) and gedunin and its derivatives^{40} had the lowest docking energy (-38.21 and -34.74 kcal/mol). Though all the herbal ligands as well as a number of neem based ligands got docked on to the 2NPD model, daucosterol from *Cichorium intybus*^{41} showed the highest affinity (-38.98 kcal/mol). Visualization of the docked structures using PyMOL package showed daucosterol to be in close proximity of FMN. Further, daucosterol was found complexed with FMN, suggesting it to be a potential inhibitor. The residues noticed in the vicinity of 2NP and daucosterol were pro^{401}, gly^{428, 433}, his^{434}, asp^{439, 491, 522}, arg^{430} and thr^{521} (Fig. 4).
Though TIM barrel fold is made up of both main and the ER domain of X-ray determined FAS II complex. domain for 2NPD model which had central barrel of conformity with the observations made by Jennie displaying the TIM barrel fold. This was in five residues required for FMN binding. asp leu of ER as well as 2NPD of 2UVA_G as the template for predicting the structure package. 2NPD and 2UVA_G structures are shown in red and blue colors, respectively [Amino acids - gly432, 462, 464, 496 and thr497 (black) are the conserved residues required for FMN binding. While O atom of amino acid pro401 of protein was involved in H-bonding with O5 atom of 2-nitropropane, NH1 atom of arg210 and O atom of thr521 had polar contact with O3P and O3' atoms of FMN respectively. The image was generated using PyMol package. 2NPD and 2UVA_G structures are shown in red and blue colors, respectively]

Discussion

In the present study, the main reason for choosing 2UVA_G as the template for predicting the structure of ER as well as 2NPD of M. tuberculosis H37Rv was that this protein is complexed with FMN. Fiser and Sali42 have suggested that the criteria for selecting a template should depend on the purpose of comparative model built on that. If a protein–ligand model is to be constructed, then choice of a template with a similar ligand will be an added advantage for predicting the ligand binding pockets.

Multiple sequence alignment of the PDB templates 2UVA and 2GJN with 2NPD sequence suggested that 14 residues were conserved (gly58,63, his64,65, asp69, gly94, leu105, gly114, asp121, gly126, thr127, glu134, lys142, asp147,152). Among these, gly94,126 and thr127 which were conserved, were the ones signifying the phosphate binding motif. This was in agreement with the studies made by Jun et al44, who observed similar amino acids being present at the phosphate-binding motif. Though H-bonding was noticed between arg430 and thr521 with FMN docked on to 2NPD, this was not in agreement with the studies carried out by Jun et al44. In their studies, polar contact was observed between FMN and gly22, gln24, thr75, lys124, asp143 and ser176 of the protein.

The anti-microbial herbal ligands that were used in the present study got docked on to the 2NPD model with docking energies ranging from -19.36 to -38.98 kcal/mol. Among these, daucosterol got docked with the lowest docking energy at a closer proximity of FMN binding site, suggesting this ligand to be a
potential inhibitor for 2NPD. Based on these studies, it is possible to speculate that daucosterol from *Cichorium intybus* can be considered as a prospective herbal drug for 2NPD of *M. tuberculosis*.

The present model of ER domain has been a static model with no consideration to the solvent environment. Refinement of the model using molecular dynamics approach could lead to a more realistic model for drug development.

References

1. World Health Organization, www.who.int/gtb