Three dimensional structure modeling of lanosterol 14-α demethylase of Candida albicans and docking studies with new triazole derivatives

Abstract

Objective: In this report, three dimensional structure of lanosterol 14-α demethylase (CYP51) protein of Candida albicans was modeled and used for docking studies with eighteen synthetic triazole derivatives.

Methods: The model was generated by multiple threading alignment and iterative structure assembly simulation by I-TASSER and validated by Ramachandran plot using PROCHECK and MolProbity servers. The eighteen 1,2,3-triazole compounds were previously synthesized using two different synthetic routes by our research group in the search of new antifungal agents. The binding affinity and interaction of all triazole compounds with predicted model of CYP51 were determined by AutoDock Vina.

Results: I-TASSER generated five models in which Model 1 was selected as best predicted structure on the basis of C-score 1.30, TM-score 0.89 ± 0.07, and RMSD 4.7 ± 3.1 Å. It showed 89.2% residues in most favoured region and 7.1% residues in additional allowed regions. Docking results showed that all the derivatives have good binding energy in the range -9.8 to -7.4 kcal/mol. Moreover, compound NT-03 (-9.8 kcal/mol) was found as best inhibitor as it binds nicely into the active site of the enzyme.

Conclusion: The study suggests that I-TASSER is very easy and reliable tool for three dimensional structure predictions. The N-atom in quinoline ring of compound NT-03 involved in the additional interaction with predicted model of CYP51. Finally, I-TASSER might be used as an important tool in drug designing process in search of more effective therapies.

Keywords: Lanosterol 14-α demethylase; homology modelling; I-TASSER; Docking studies; Triazoles; Candida albicans

Introduction

The increased incidences of candidiasis and alarming rate of drug resistance has prompted researchers to develop novel and more effective antifungal agents. Opportunistic yeast Candida, mainly C. albicans, is responsible for this complicated fungal infection with 50-90% human candidiasis [1,2]. Commonly used antifungal drugs belong to various classes and characterised by different mechanisms of action and spectrum of activity. Among them, azoles are the first choice of drugs for the treatment of fungal infection [3]. They interfere with the synthesis of ergosterol by inhibiting CYP51 enzyme of its biosynthetic pathway in C. albicans [4]. CYP51, a member of Cytochrome P450 family, is a heme thiolate enzyme which converts lanosterol into 4,4′-dimethyl cholesta-8,14,24-triene-3-beta-ol [5]. The activity of azole drugs is attributed to the co-ordination binding of the heterocyclic nitrogen atom (N-3 of imidazole and N-4 of triazole) to the heme iron atom in the binding site of CYP51 enzyme. Due to inhibition of CYP51, depletion of ergosterol coupled with an accumulation of 14-methyl sterols which results in impaired fungal cell growth [6].

Mohammad Irfan and Mohammad Abid

Medicinal Chemistry Laboratory, Department of Biosciences, Jamia Millia Islamia (A Central University), New Delhi-110025, India

Corresponding author: Mohammad Abid

mabid@jmi.ac.in

Medicinal Chemistry Laboratory, Department of Biosciences, Jamia Millia Islamia (A Central University), New Delhi-110025, India.

Tel: +91-8750295095
Fax: +91-11-2698 0229

Citation: Irfan M, Abid M. Three dimensional structure modeling of lanosterol 14-α demethylase of Candida albicans and docking studies with new triazole derivatives. Chem Inform. 2015, 1:1.
The essential role of CYP51 enzyme in fungi and yeast metabolism make it an important target for antifungal drug design [7]. Although the advance in bioinformatics tools may help to develop new drug molecules, the absence of crystal structure of CYP51 enzyme of *C. albicans* is a major obstacle in the drug design efforts. Currently 38 X-ray coordinates of CYP51 available in PDB in which, one is for *Saccharomyces cerevisiae* (4LXI) [8] and others for non-membrane bound CYP51 like *Trypanosoma brucei* (3G1Q) [9], *Mycobacterium tuberculosis* (2B29) [10], *Homo sapiens* (3JUV) [11], *Leishmania infantum* (3L4D) [12] etc. In recent years, the advance research in three dimensional structure and function prediction have resolved this problem to an extent and helps to obtain valuable information of desired protein [13].

The three dimensional structure prediction by computational methods has divided into three categories: i) comparative modeling based on evolutionarily related homologous template [14], ii) threading methods which use template of proteins from different evolutionary origin with similar sequences [15], and iii) *ab initio* modeling in which no structural similarities for query protein available in PDB [16,17].

In this study, we selected I-TASSER (iterative threading assembly refinement) server for 3D structure prediction of CYP51 enzyme of *C. albicans*. I-TASSER has composite approaches of threading, structural refinement, and *ab initio* modeling to generate reliable model which is different from earlier reported methods of CYP51 structure prediction [18]. I-TASSER has been ranked as the best server for automated protein structure prediction in the last five consecutive CASP7 [19], CASP8 [20], CASP9 [21], CASP10 [22], and CASP11 experiments. The predicted model was used for docking studies with eighteen newly synthesized triazole derivatives from naturally bioactive scaffolds to explore their interactions and binding affinity in search of more potent drug like molecules.

Methods

Three dimensional structure prediction by I-TASSER

The sequence of 528 amino acids of *C. albicans* lanosterol 14-α demethylase protein (CYP51) was retrieved from Swiss Prot database (http://www.uniprot.org/uniprot/P10613) in FASTA format as follows:

MAIVETVIDGINYFLSLSVTQISILLSLVGLPGFVYNYLWQYLSSLRKRPLRFVFLFYWIPWFGS AASYGQQPVEFEFCRQKGYDVFSFMLLGKIMTVYGPKGFHVFBNAKLDSVADAYKHLTTPPFGKVIYDCPNISRLMEQKKFAKFAINTDSFMRVKYVIREILEINYVTFDSTSEKLKTEKTHGVANVMKTQPEITIFASRSLFGDEMRRIRFDSFAQLYSDDKGTFTPINPVPNLPLPHYWRDAAQKKSATYMKKEKSRERRGDIDPRNRLDSSLHSTYKDGVMKTQIEANLLILGMMGQHTSASTSASWFLHLGKEKPLQDVIYQEVVELLKEKGGLDNLTYEDLQKPLSVNNTIKETLMHPMLHSIFRKVTNPIREPNTVIVPKGYLYLVSPGAHTSERDYFNPDFTDPRTRWD TAAAANSVSVNNSDEVYDFGKSVSGYPLFGGGRHCIGEQFAYVQLGTLTTFTVYNLRWITDYGKVPPDPDSSMVVLPMLEPAEIWEKRET MGFRLTVTFVYNLRWITDYGKVPPDSSMVVLPMLEPAEIWEKRET MF

The three dimensional model was generated using I-TASSER server which generates 3D model of query sequence by multiple threading alignments and iterative structural assembly simulation [17]. The reasons behind the selection of this server were its availability, composite approach of modeling and performance in CASP competition. I-TASSER methodology includes general steps of threading, structural assembly, model selection, refinement, and structure based functional annotation [18]. Firstly, the evolutionary related protein against query sequence was checked by running PSI-BLAST which created sequence profile and used to predict secondary structure by PSIPRED [23]. Then, the query sequence was threaded through the representative PDB structure library using LOMETS [24]. The quality of template alignment was checked by Z-score and the best one used for the further consideration. In next step, the continuous fragments in threading alignments were excised to from an assembly structure model of aligned regions. The structural assembly of unaligned regions was built by *ab initio* modeling [17].

The modeling accuracy of unaligned regions is generally low while threading aligned regions have high accuracy so these template fragments keep rigid in simulation process to obtained high resolution structure. The replica exchange Monte Carlo simulation technique was used for fragment assembly [25]. The simulation include Ca/side chain correlation, H-bonds, hydrophobicity, spatial restraints from threading templates [24], and sequence based contact predictions from SVMSEQ [26]. The conformations generated during refinement simulation process were clustered by SPICKER [27] and the average of three dimensional coordinates of all the clustered structure was calculated to obtained cluster centroids.

In refinement process, the selected cluster centroids were again used to perform further fragment assembly simulation which helps to remove steric clashes and to refine the global topology of the cluster centroids. The PDB structures, structurally closed to the cluster centroids, were identified by TM-align [28]. The final structural models were generated by REMO [29] in which cluster centroids of second round simulation used as input. In last step, the function of three dimensional model of query protein was predicted by matching the proteins of known structure and function in PDB. The functional analogs were ranked on the basis of TM-score, RMSD, sequence identity, and coverage of the structure alignment. The quality of predicted model was determined by C-score (confidence score) which is ranged as –5 to 2. It depends on quality of threading alignment and the convergence of structural assembly refinement simulations.

Validation of predicted 3D model (model 1)

The conformation of best 3D model (model 1) predicted by I-TASSER was further validated by Ramachandran plot. The conformation of the predicted model was calculated by analyzing the phi (Φ) and psi (Ψ) torsion angles using PROCHECK online server. The Ramachandran plot obtained from PROCHECK describes a good quality model which has over 90% residues in most favoured region. The assessment of model was also confirmed by MolProbity online server [30].

Inhibitors for docking studies

*In silico* screening of selected triazole derivatives with the predicted model of *C. albicans*’s CYP51 enzyme was done in order to find their interactions and binding energy. Triazole derivatives
used in the study were recently synthesized by our group (Figures 1 and 2). 1,2,3-Triazole functionality showed various biological activities and present in different antifungal drugs like fluconazole, itraconazole, voriconazole etc [31]. Compounds CT-01 to CT-10 were synthesised by one pot three-component stepwise approach based on azide-chalcone oxidative cycloaddition and post-triazole arylation (Scheme 1) [32]. While compounds NT-01 to NT-08 was obtained by modification of some naturally occurring alcohols using [3+2] azide-alkyne cycloaddition strategy (Scheme 2) [33].

**Docking studies**

In this study, AutoDock Vina was used to determine the binding energy of designed triazole derivatives with predicted model obtained from I-TASSER and compared with the standard drug, fluconazole. AutoDock Vina is advanced version which is much faster and accurate in binding mode prediction if compared with its older version AutoDock 4.0 [34]. The affinity grid box was centred at the binding site and the grid spacing fixed to 0.375 Å which is highly recommended. The program automatically calculates the grid map and grid centre coordinates (X=-2.0, Y=0.98, Z=0.41) considering the predicted binding site of enzyme. After generating receptor and all ligands file in PyMOL AutoDock/Vina plugin, Vina was ran to obtained docking score (binding affinity in kcal/mol) and RMSD value for all compounds. PyMOL was used for visualization of the polar contacts in between ligands and binding site of protein [35]. The best docking result can be considered to be the conformation with the lowest docking score or binding energy and lowest RMSD.

**Results and Discussion**

**Three dimensional structure prediction**

The predicted model of *C. albicans* CYP51 protein and its three dimensional coordinate file in PDB format were successfully obtained from I-TASSER. The results obtained from server includes predicted secondary structure with confidence score (range 0 to 9), predicted solvent accessibility, five predicted structures with C-score, top ten template from PDB used in alignment, top ten PDB structural analogs, functional analogs protein, and binding site residues. Model 1 (Figure 3) was selected as best predicted model with C-score 1.30, TM-score 0.89 ± 0.07, and RMSD 4.7 ± 3.1 Å. C-score with higher value reflect a model of better quality [18].

Top ten threading templates for query protein sequence were identified by LOMETS meta-server (Table 1). The threading alignment is generally estimated by normalized Z-score. However, a normalized Z-score >1 value reflect a confident alignment but in case of small alignment of large query sequence, it does not give significant indication of modeling accuracy. Then percentage sequence identity in the threading aligned region (Iden1) and in the whole chain (Iden2) considered for the good homology (Table 1). The structural alignment program, TM-align, identified 4lxjA in PDB library as best structural analog of the top scoring model of I-TASSER with the TM-score of 0.975 (Table 2). The ligand binding site of Model 1 was predicted and aligned with the template (best hit 3juvA) binding site by TM-align. The binding score (BS) and confidence score of ligand binding (CscoreLB) for predicted binding site with template (3juvA) were measured as 1.38 and 0.73. The BS score is based on structure similarity between binding site of template and predicted model of I-TASSER. The BS score >1 and CscoreLB >0.5 indicates more reliable ligand binding site prediction [18].

**Assessment of predicted model**

The Ramachandran plots of best predicted model were obtained from PROCHECK and MolProbity servers which showed the reliability of model. The PROCHECK Ramachandran plot showed 89.2% residues in most favoured regions and 7.1% residues in additional allowed regions i.e., the total of 96.3% residues in allowed regions which indicates a good quality model. MolProbity Ramachandran plot also showed 96.6% residues in allowed regions which again confirmed the quality of predicted model (Figure 4).
The molecular docking of all the triazole derivatives with best predicted model of *C. albicans* CYP51 protein was successfully completed using AutoDock Vina. All the docked compounds showed good binding energy with a range of -9.8 to -7.4 kcal/mol (Table 3). The binding energy of standard drug fluconazole was found to be -8.6 kcal/mol. Among all the eighteen triazole derivatives, only five compounds showed equal or more binding energy than the standard drug fluconazole. Compound CT-2 of scheme 1 was the
Table 3 Binding affinity of docked compounds with predicted model of *Candida albicans* lanosterol 14-α demethylase protein. FLC=Fluconazole.

<table>
<thead>
<tr>
<th>Code</th>
<th>Binding energy (kcal/mol)</th>
<th>RMSD</th>
<th>Code</th>
<th>Binding energy (kcal/mol)</th>
<th>RMSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT-01</td>
<td>−8.4</td>
<td>0</td>
<td>CT-10</td>
<td>−8.4</td>
<td>0</td>
</tr>
<tr>
<td>CT-02</td>
<td>−9.1</td>
<td>0</td>
<td>NT-01</td>
<td>−9.1</td>
<td>0</td>
</tr>
<tr>
<td>CT-03</td>
<td>−8.7</td>
<td>0</td>
<td>NT-02</td>
<td>−9.7</td>
<td>0</td>
</tr>
<tr>
<td>CT-04</td>
<td>−8.4</td>
<td>0</td>
<td>NT-03</td>
<td>−9.8</td>
<td>0</td>
</tr>
<tr>
<td>CT-05</td>
<td>−8.7</td>
<td>0</td>
<td>NT-04</td>
<td>−8.7</td>
<td>0</td>
</tr>
<tr>
<td>CT-06</td>
<td>−7.4</td>
<td>0</td>
<td>NT-05</td>
<td>−9.4</td>
<td>0</td>
</tr>
<tr>
<td>CT-07</td>
<td>−8.6</td>
<td>0</td>
<td>NT-06</td>
<td>−8.9</td>
<td>0</td>
</tr>
<tr>
<td>CT-08</td>
<td>−8.3</td>
<td>0</td>
<td>NT-07</td>
<td>−9.1</td>
<td>0</td>
</tr>
<tr>
<td>CT-09</td>
<td>−8.9</td>
<td>0</td>
<td>NT-08</td>
<td>−9.3</td>
<td>0</td>
</tr>
<tr>
<td>FLC</td>
<td>−8.6</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The best compound in its series with lowest binding energy -9.1 kcal/mol. However all the compounds of scheme 2 showed less binding energy than the fluconazole and compound NT-03 with binding energy -9.8 kcal/mol was selected as best inhibitor. It also showed good interactions with the binding site residues of target protein. The N-atom of triazole ring was interacted with the TYR-132 residue of binding site. Compound NT-03 having quinoline ring also showed additional interactions with GLY-307 and HIS-310 residues in binding pocket due to N-atom of the quinoline ring (Figures 5 and 6).

Figure 2 Structure of triazole compounds synthesized by Scheme 2

Figure 3 a) Three dimensional structure of *C. albicans* CYP51 protein predicted by I-TASSER. b) Alignment of query protein (cyan) with structural analog (green) 4lxjA in PDB library.
Figure 4  Three dimensional structure prediction validation of top score model of I-TASSER by a) PROCHECK Ramachandran plot and b) MolProbity Ramachandran plot.

Figure 5  Interaction of best inhibitor (NT-03) with the binding site residues of predicted model.

Figure 6  Compound NT-03 inside the binding site of predicted model of lanosterol 14-α demethylase
Conclusion

In summary, three dimensional model of C. albicans CYP51 protein was built by multiple threading alignment and iterative structure assembly simulation by I-TASSER. The crystal coordinate 4lxjA in PDB library was found very close structural resemblance with predicted model. The reliability and accuracy of model were confirmed by Ramachandran plot using PROCHECK and MolProbity server. The docking studies with predicted model showed that compound NT-03 has good binding affinity (~9.8 kcal/mol). Moreover, it interacted with TYR-132 and GLY-307 and HIS-310 residues in binding site of predicted protein which is in accordance with the experimental bioactivity data. It can be concluded that natural compound based triazole derivatives can be developed as more effective antifungal agent.

Acknowledgement

Mohammad Abid gratefully acknowledges the funding support in the form of Major Research Project by University Grant Commission (UGC), Govt. of India (Grant No. 41-277/2012 (SR)). MI would like to acknowledge UGC, INDIA for fellowship.

Conflict of interest

Authors declare no conflict of interest.
References


