International Journal of Pharmaceutical Chemistry
ISSN: 2249-734X (Online)
Journal DOI: 10.7439/ijpc
CODEN: IJPCH3 (American Chemical Society)
Research Article

Homology modeling of *Candida Albicans* lanosterol 14 α-demethylase and validation of the homology model

Rani S. Kankate and Vikas C. Nathe*

Department of Pharmaceutical Chemistry, MET's Institute of Pharmacy,Pune University,422003, Nashik, India

Abstract

The cytochrome P450 sterol 14α-demethylase enzyme (CYP51) is the target of azole antifungals. Azoles block ergosterol synthesis, and thereby inhibit fungal growth, by binding in the active-site cavity of the enzyme and ligating the iron atom of the heme cofactor through a nitrogen atom of the azole. In this work, homology models of the CYP51 enzymes from *Candida albicans* were constructed based on the X-ray crystal structure of CYP51 from *Saccharomyces cerevisiae*.
Keywords: Homology Modeling, Antifungal, *Candida albicans* and Triazole

1. Introduction

Opportunistic and invasive fungal infections have increased dramatically in recent years to become important causes of morbidity and mortality. *Candida* and *Aspergillus* sp. are the most infamous fungal pathogens that account for majority of invasive and opportunistic fungal infections, respectively, occurring worldwide [1]. The antifungal agents available to treat these infections can be divided into four groups according to their modes of action. They impair membrane integrity (polyenes), interact with microtubules (griseofulvin), or inhibit macromolecule synthesis (flucytosine) or ergosterol biosynthesis (azoles, morpholines, and allylamines). For treatment of systemic *Candida* infections only polyenes, flucytosine, and azoles are used [2]. 1,2,4-triazole system and its analogs have been investigated as therapeutically interesting drug candidates because of their varied properties, such as selective COX-2 inhibitors, anti-acetylcholinesterase, and antimicrobial agents[3].

Candida albicans: The major pathogen has been *C. albicans*, normally a commensal of the oral cavity and gastrointestinal tract of humans. Non-albicans *Candida spp*. (e.g., *C. glabrata, C. tropkalis, C. krusei*), however, are also found with increasing frequency. The other Candida spp. encountered in human infections is *Candida tropicalis*, *Candida glabrata*, *Candida parapsilosis*, *Candida stellatoidea*, *Candida krusei*, and *Candida kyfer*. *C. albicans* can cause chronic, superficial infections in otherwise healthy individuals; a prime example is vaginal candidiasis. Although, *C. albicans* is recognized as the most common pathogen among the Candida spp., in recent years other *Candida spp.*, have been isolated with increasing frequency from clinical disease. New triazoles, such as voriconazole and posaconazole, and the echinocandins are active against these two species, although crossresistance was noted within the azoles in some *C. glabrata* strains [4]. Azole antifungal agents inhibit the cytochrome P450 sterol 14 α-demethylase (14DM, CYP51) by a mechanism in which the heterocyclic nitrogen atom (N-3 of imidazole and N-4 of triazole) binds to the heme iron atom in the binding site of the enzyme. Lanosterol-14 α-demethylase (CYP51) is a key enzyme of sterol biosynthesis in fungi. The resulting ergosterol depletion and the accumulation of precursor 14 α -methylated sterols disrupt the structure of the plasma membrane, making it more vulnerable to further damage, and alter the activities of several membrane-bound enzymes [5]. The importance of homology modeling has been steadily increasing because of the large gap that exists between the overwhelming number of available protein sequences and experimentally solved protein structures, and also, more importantly, because of the increasing reliability and accuracy of the method. Given a protein sequence, homology modeling usually consists of the following four steps1) identify the homologue of known structure from the Protein Data Bank; 2) align the query sequence to the template structure; 3) build the model based on the alignment; 4) assess and

* Correspondence Info
Vikas C. Nathe
Department of Pharmaceutical Chemistry,
MET's Institute of Pharmacy, Pune University,422003, Nashik, India.
E mail: vikasnathe43@gmail.com

refine the model. When the sequence identity is above 40%, the alignment is straight forward, there are not many gaps, and 90% of main-chain atoms could be modeled with an RMSD (root-mean-square distance) error of about 1 Å [6].

2. Materials and methods

2.1 Protein Preparation

The initial protein structure (4K0F) was in a PDB-format file and included a co-crystallized ligand (itraconazole) and did not include explicit hydrogen. The preparation of the protein was done in the following steps:

- The hydrogens were added.
- All the water molecules in the structure were deleted.
- The protein, metal ions, and cofactors were adjusted.
- The ligand bond orders and formal charges were adjusted.
- The bound itraconazole was deleted.
- Then PDB format of the file was converted to and saved as .Mol2 format.

3. Results and Discussion

3.1 Homology Modeling of *Candida Albicans SC5314*.

Homology modeling helps to build a three dimensional structure of a target based on a template that helps to identify the putative active sites and binding pockets, which further delves the probable ligand-protein interactions to understand the exact mechanism and function of the particular protein. Homology modeling involves taking a known sequence with an unknown structure (target) and mapping it against a known structure (template) of one or several similar (homologous) proteins. It would be expected that two proteins of similar origin and function would have reasonable structural similarity.

The primary sequence of target organism, *Candida albicans*, was obtained from the Universal Protein Resource (http://www.uniprot.org/uniprot/P10613) (Accession Code: P10613) (Fig. No. 1) and sequence homologous was obtained from Protein Data Bank (PDB) (http://www.rcsb.org/pdb/explore/explore.do?structureId=4k0F) using Blast search. Then, the chosen model was subjected to energy minimization and molecular dynamics simulations to obtain a stable and low energy conformation. The quality of the final refined model was assessed by a series of tests for its internal consistency and reliability. Finally, the best quality model of *C. albicans* was subjected to further calculations and molecular modeling studies. In literature the structure of cytochrome P450 lanosterol 14α-demethylase was developed homologically using crystal structure of lanosterol 14α-demethylase from *Saccharomyces cerevisiae* YJM 789 as template (531 amino acid residues). Based on the result of blast search, we used the crystal structure of *Saccharomyces cerevisiae* YJM 789 lanosterol 14α-demethylase (CYP51) with intact transmembrane domain bound to itraconazole as a template for homology modeling (PDB ID. 4K0F). The quality of homology models depend on the sequence similarity between template and target. If the sequence identity is greater than 30%, the two proteins probably have a common ancestor and are likely to share a common 3D structure. In general the process of homology modeling involves four sequential steps such as template selection, target-template alignment, model construction, and model assessment.

Fig.No.1: FASTA sequence for lanosterol 14α-demethylase of *Saccharomyces Cerevisiae.*

>gi|576865179|pdb|4K0F|A:12-524 Chain A, Crystal Structure of Lanosterol 14-alpha Demethylase With Intact Transmembrane Domain Bound To Itraconazole

LEYVNIGLSHFLALPLAQRISLIIIPFIYNIVWQLLYSLRKDRPPLVFYWIPWVVGSAVVYGMKPYEFFEECQKKYGDIFSFVL
LGRVMTVYLGPKGHEFVFNAKLADVSAEAAYAHLTTPVFGKGVIYDCPNSRLMEQKKFVKGALTKEAFKSYVPLIAEEVY
KYFRDSKNFRLNERTTGTIDVMVTQPEMTIFTASRSLLGKEMRAKLDTDFAYLYSDLDKGFTPINFVFPNLPLEHYRKRD
HAQKAISGTYMSLIKERRKNNDIQDRDLIDSLMKNSTYKDGVKMTDQEIANLLIGVLMGGQHTSAATSAWILLHLAERP
DVQQELYEEQMRVLDGGKKELTYDLLQEMPLLNQTIKETLRMHHPLHSLFRKVMKDMHVPNTSYVIPAGYHVLVSPG
YTHLRDEYFPNAHQFNIHRWNNDSASSYSVGEEVDYGFGAISKGVSSPYLPFGGGRHRCIGEHFAYCQLGVLMSIFIRTL
KWHYPEGKTVPPPDFTSMVTLPTGPAKIIWEKR

In the present study, we report the construction of the 3D structure of C. albicans by homology modeling using the sterol 14α-demethylase (CYP51) from *Mycobacterium tuberculosis* as a template. In order to clarify the binding mode of azole antifungal agents with modeled 14α -DM and provide straight forward information for further

rational drug design, molecular docking was used to dock azoles into the active site of enzyme. The sequence used is: gi/1169073/sp/P10613/CP51_CANA 528aa. (CA - CYP51).

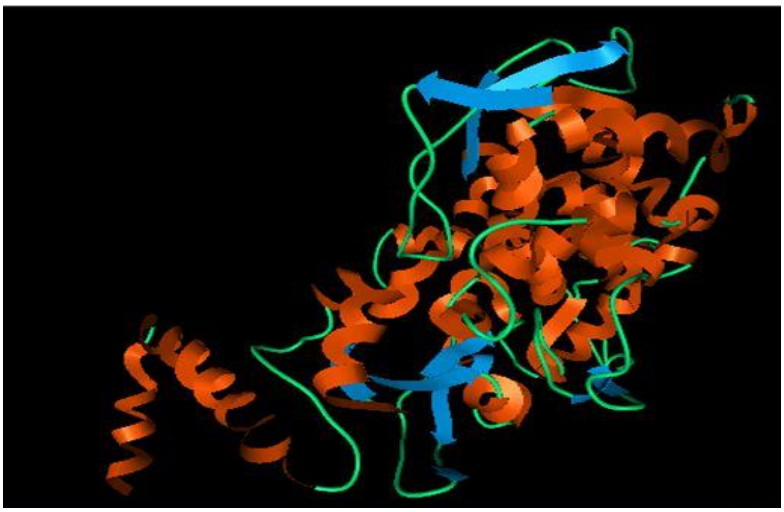
3.2 Using local BLAST to search for a template using target sequence

The sequence alignment of 4K0F was done using VLife MDS 4.3 showed identity 65%, similarity/ positives 78% and Gaps 1% with our target sequence (Fig. No. 2).

3.3 Model Optimisation

VLife MDS provides users with a utility to calculate energies and optimize geometries of molecules using different force field methods such as Dreiding, Universal Force Field (UFF) and Merck Molecular Force Field (MMFF). It also provides constrained optimization for small molecules and aggregated optimization for macromolecules like proteins and polypeptides, which energy minimizes the protein.The 3D model of developed homology model is shown in (Fig. No. 3).

Fig. No. 3: Homology modeled structure of lanosterol 14-alpha demethylase of *Candida Albicans.*



3.4 Model Validation

As at almost every step, choices had to be made there were a number of parameters which affected the final model. The model quality was assessed by checking their geometrical and physico-chemical parameters. VADAR, WHATIF (swift.cmbi.ru.nl/), PROCHECK, VERIFY 3D servers were used to assess the quality of bond lengths, dihedral values and angle distribution.

Ramachandran plot was in agreement with a good quality model which is carried out by RAMPAGE (Table No. 1) (Fig. No. 4. and 5) (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) [7] and Molprobity which are free online servers (http://kinemage.biochem.duke.edu) (Table No. 2) (Fig.No. 6) [8].

Table No. 1: Ramachandran plot values showing number of residue in the favored allowed, outlier region through Rampage evaluation server .

Structure	Number of residues in Favored region (%)	Number of residues in Allowed region (%)	Number of residues in Outlier region (%)
4K0F	98	2	0
MODEL BUILT	97.3	2.5	0.1

Table No. 2: Ramachandran plot values showing number of residue in the favored allowed, outlier region through Molprobity evaluation server.

Structure	Number of residues in Favored region (%)	Number of residues in Allowed region (%)	Number of residues in Outlier region
4K0F	98	99.8	0
MODEL BUILT	96.9	99.5	2

Fig. No. 4:Ramachandran plot of 4K0F.

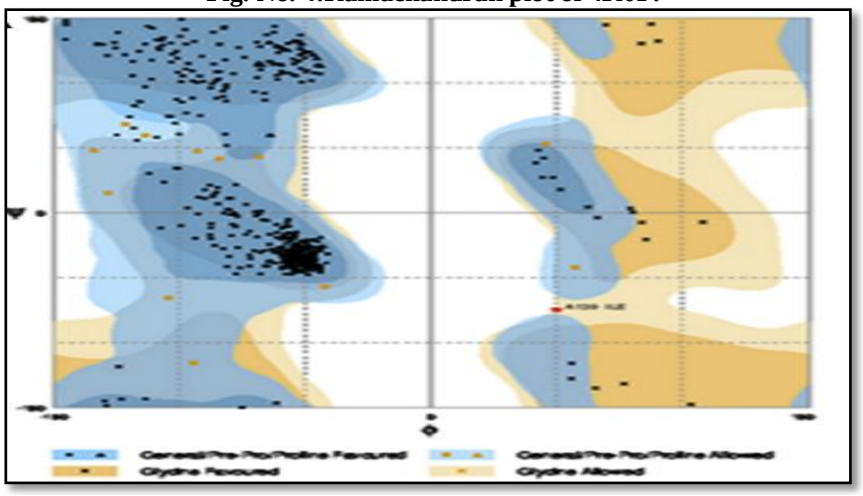


Fig. No. 5: Ramachandran plot of modeled structure.

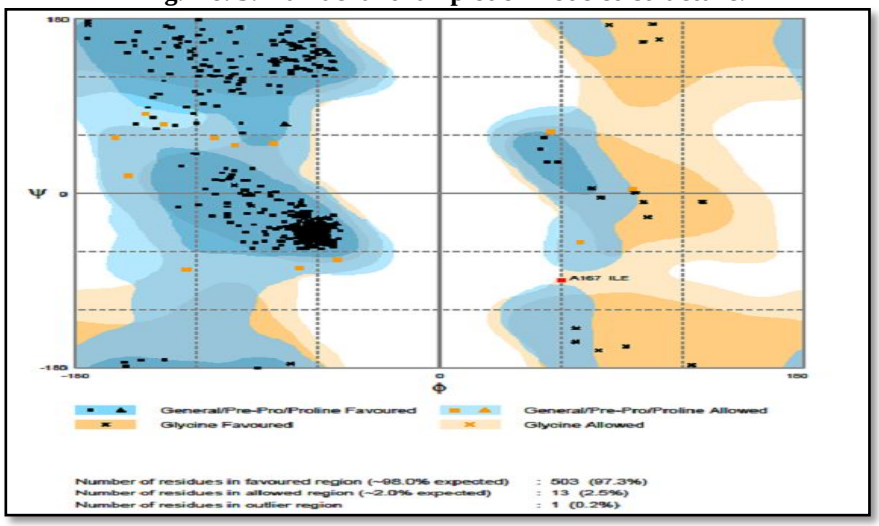
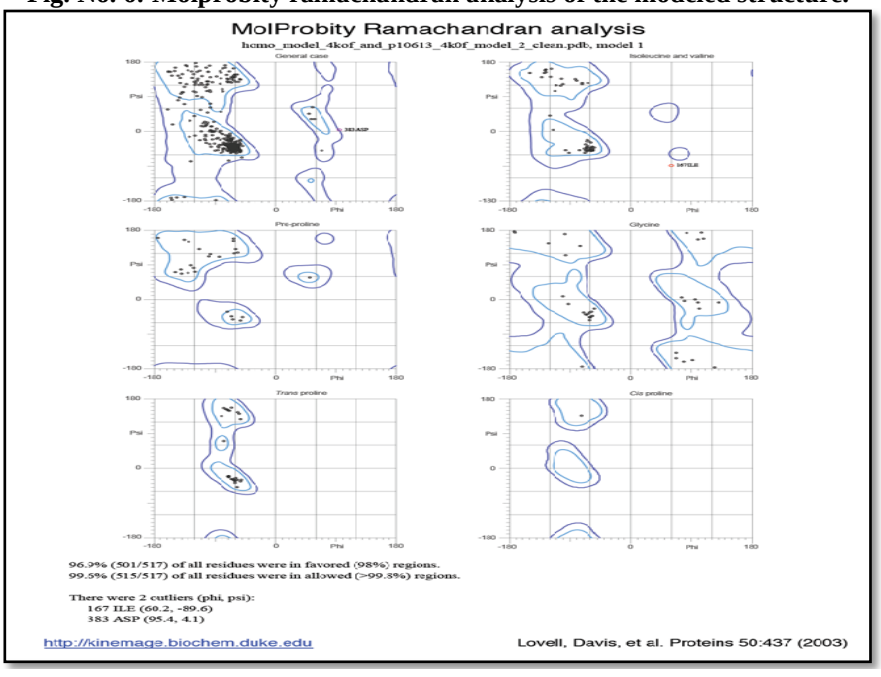


Fig. No. 6: Molprobity ramachandran analysis of the modeled structure.



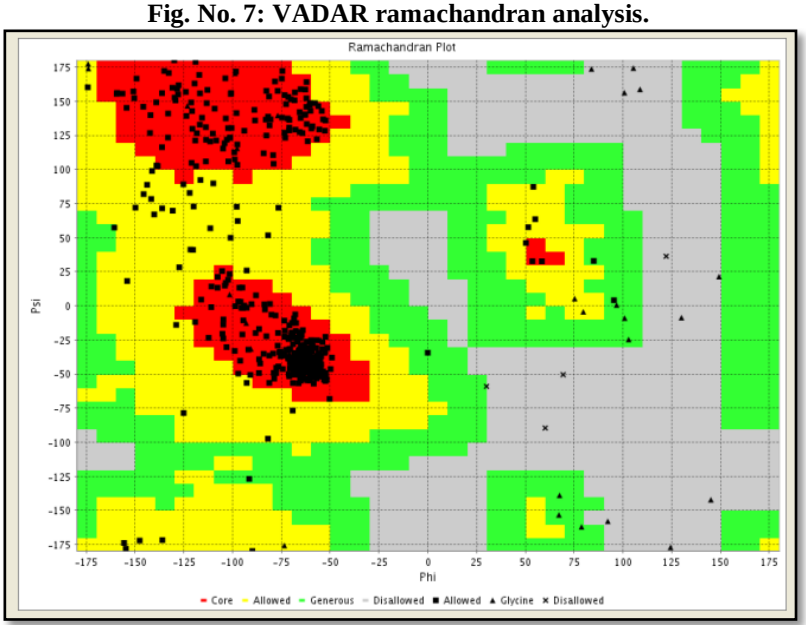
VADAR (Volume Area Dihedral Angle Reporter) is a comprehensive web server for quantitative protein structure evaluation (Table No. 3).

Table No. 3: Ramachandran plot values showing number of residue in the favored, allowed, generous and outside region through VADAR evaluation server.

Statistics	Observed	Expected
Res. in phipsi core	485 (91%)	479 (90%)
Res. in phipsi allowed	39 (7%)	37 (7%)
Res. in phipsi generous	4 (0%)	5 (1%)
Res. in phipsi outside	4 (0%)	0 (0%)
Res. in Omega Core	502(94%)	511(96%)
Res. in Omega allowed	19(3%)	16(3%)
Res. in Omega generous	4(0%)	0(0%)
Res. in Omega outside	7(1%)	5(1%)

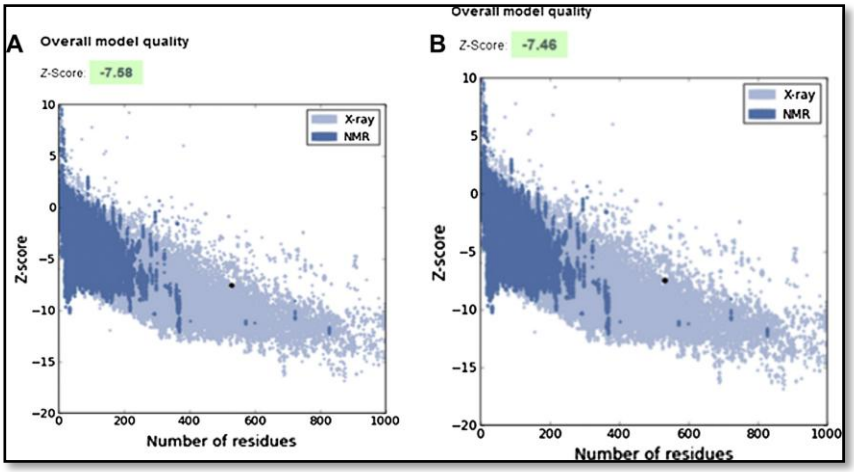
These include excluded volume, accessible surface area, backbone and side chain dihedral angles, secondary structure, hydrogen bonding partners, hydrogen bond energies, steric quality, solvation free energy as well as local and overall fold quality.The VADAR web server is freely accessible at (http://redpoll.pharmacy.ualberta.ca/vadar) (Fig. No.7) [9].

Fig. No. 7: VADAR ramachandran analysis.



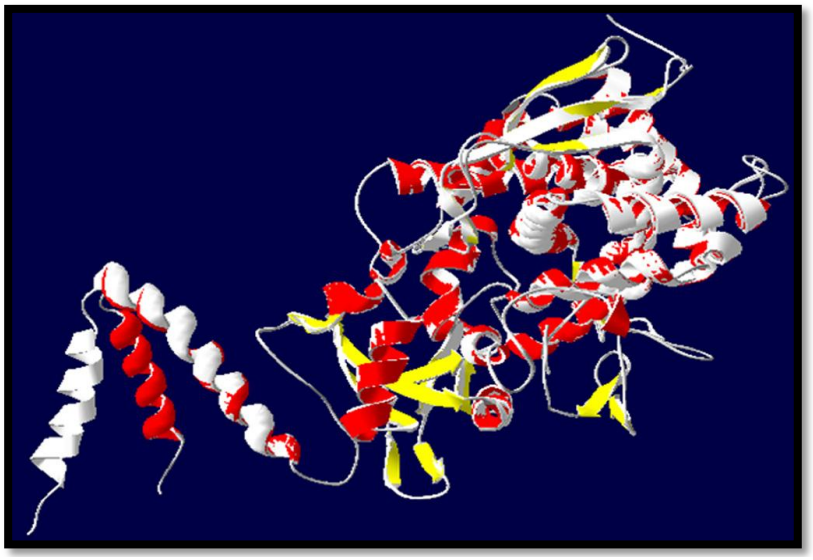
ProSA: Fold reliability analysis. Validation was carried out using ProSA to obtain the Z-score value for the comparison of compatibility (https://prosa.services.came.sbg.ac.at/prosa.php) (Fig. No. 8).

Fig. No. 8: (A) Plot of Z-Score of 4K0F and (B) Plot of Z-Score of modeled structure.



The Z-score plot showed spots of Z score values of proteins determined by NMR (represented in dark blue color) and by X ray (represented in light blue color)[10,11]. The two black dots represent Z-scores of our model (-7.46) and template (-7.58). These scores indicate the overall quality of the modeled 3D structure of lanosterol 14-alpha demethylase. RMSD (0.38 A˚) was calculated between the main chain atom of model and template (Fig. No. 9).

Fig. No. 9: RMSD between modeled structure and template of 0.38 A˚. red color indicates modeled structure of lanosterol 14 alpha demethylase, and white indicates template structure K0F.



It indicated close homology. This ensured the reliability of the model. Superimposition between target and template structure was done using SPDBV (http://spdbv.vital-it.ch/).

PROCHECK specializes in stereochemical quality evaluation with a particular focus on reporting torsion angle parameters [12]. The *G*-factor provides a measure of how "normal", or alternatively how "unusual", a given stereochemical property is. G-factors provide a measure of how unusual, a property values below -0.5 is unusual and values below -1.0 is highly unusual. Overall average of G-factors of modeled (u017) C. albicans SC5314 was -0.88 which indicates that it was highly unusual model (Fig. No. 10) (Table No. 4) [13].

Fig. No. 10: Ramachandran Analysis from Procheck statistics and G Factor Calculation for the build homology model.

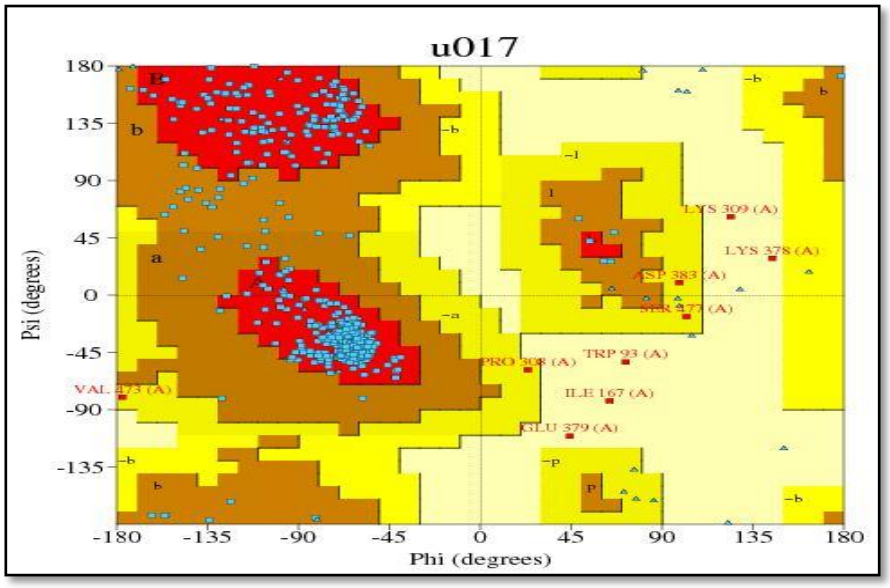


Table No. 4: G-Factor of homology model.

Parameter	Score	Average score
Dihedral angle		
Phi–psi distribution	-0.05	-
Chi1-chi2 distribution	-0.100**	-
Chi1 only	-0.031	-
Chi3 & chi4	0.51	-
Omega	-0.74	-
		-0.42
Main-chain covalent forces		
Main-chain bond lengths	-1.77**	-
Main-chain bond angles	-1.42**	-
		-1.57
Overall average		-0.88*
Values below -0.5—unusual **Values below -1.0—highly unusual**		

4. Conclusion

Candida albicans has a worldwide distribution and is one of the most common causes of invasive fungal infections. In this study we had developed 3D model of *C. albicans*, which is important for the discovery of novel antifungal agents with broad spectrum activity. The *C. albicans* structure was modeled on the base of the most accurate crystal structure of 4K0F. The developed homology model in this study can be used for azole optimization, virtual screening or for de novo inhibitor design for the discovery of new antifungal agents.

Conflict of Interest

The author(s) declare(s) no conflict of interest.

Acknowledgement

The authors are thankful to Dr. Kundan Ingle (Application Scientist) of Vlife sciences for his help and guidance for homology modeling.

References

[1] Gollapudy R, Ajmani S, Kulkarni S. Modeling and interactions of Aspergillus fumigatus lanosterol 14-alpha demethylase "A" with azole antifungals. *Bioorg Med Chem.* 2004; 12:2937–50. http://dx.doi.org/10.1016/j.bmc.2004.03.034

[2] Marichal P, Bossche H Vanden, Odds FC, Nobels G, Warnock DW, Timmerman V, *et al*. Molecular Biological Characterization of an Azole-Resistant Candida glabrata Isolate. *Antimicrob Agents Chemother*.1997; 36:2229–37. http://dx.doi.org/10.1128/aac.36.12.2602 Actions

[3] Sangshetti JN, Deepak K, Sarkate AP, Shinde B. Synthesis, Antifungal Activity, and Docking Study of Some New 1, 2, 4-triazole Analogs. *Chem Biol Drug Des.* 2011; 78:800–809 http://dx.doi.org/10.1111/j.1747-0285.2011.01178.x

[4] Kathiravan MK, Salake AB, Chothe AS, Dudhe PB, Watode RP, Mukta MS, *et al*. The biology and chemistry of antifungal agents : A review. *Bioorg Med Chem.* 2012; 20:5678–98. http://dx.doi.org/10.1016/j.bmc.2012.04.045

[5] Iman M, Davood A. Homology modeling of lanosterol 14 a -demethylase of Candida albicans and insights into azole binding. *Med Chem Res*. 2014; 2890–99. http://dx.doi.org/10.1007/s00044-013-0769-z

[6] Zhexin Xiang. Advances in Homology Protein Structure Modeling. *Curr Protein Pept Sci*. 2006 June; 7: 217–227. http://dx.doi.org/10.2174/138920306777452312

[7] Lovell SC, Davis IW, Arendall III WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS & Richardson DC. Structure validation by Cα geometry: φ/ψ and Cβ deviation. *Proteins: Structure, Function & Genetics*. 2002; 50:437-450. http://dx.doi.org/10.1002/prot.10286

[8] Chen *et al*. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica* 2010; D66:12-21.

[9] Willard L, Ranjan A, Monzavi H *et al*. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* 2003; 31: 3316-3319. PMCID: PMC168972 http://dx.doi.org/10.1093/nar/gkg565

[10] Wiederstein & Sippl. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007; 35: W407-W410. http://dx.doi.org/10.1093/nar/gkm290

[11] Sippl MJ. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins.*1993; 17:355-362. http://dx.doi.org/10.1002/prot.340170404

[12] Laskowski R A, MacArthur M W, Moss D S and Thornton J M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, 1993; 26: 283-291.

[13] Engh RA and Huber R. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst.* 1991; A47:392-400. http://dx.doi.org/10.1107/S0108767391001071