Science Research Reporter, 5(1):01-08, (April - 2015)

Online, Open Access Available at http://jsrr.net
ISSN: 2249-2321 (Print); ISSN: 2249-7846 (Online)
Research Article

In Silico Analysis of HMG CO-A Reductase of *Candida albicans* SC5314

Santosh Kodgire, Vikram Pawar, Nilesh Wagh, Laxmikant Kamble and Gajanan Zore*

School of Life Sciences, SRTM University, Nanded-431606
gbzore@rediffmail.com

Article Info

Received: 21-02-2015,

Revised: 19-03-2015,

Accepted: 27-03-2015

Keywords:

Candida albicans,
HMGCR,
Mevalonate pathway, Sterol biosynthesis,
drug target

Abstract

In the present study *In silico* comparative analysis of HMG CoA reductases (HMGCR) from *Candida albicans* and *Homo sapiens* were done. True orthologs of HMGCR from *C. albicans*, *H. sapiens* with percent protein similarity of 52 % and 34 % respectively to that of functionally characterized HMGCR from *Schizaosaccharomyces pombe* were found out. Conserved domains and motifs on these proteins was identified namely HPIH, MMPL superfamily and HMGCR superfamily. Out of these domains, the domain HPIH (4 to 154 amino acid) found in *C. albicans* was absent in *H. sapiens.* 3D protein analysis showed considerable differences in the structures of *C. albicans* and human HMG CoA reductases although amino acids TYR, GLU and ASN at catalytic domains are conserved. The position of these amino acids in ligand binding site of HMGCR of *C. albicans* is different than that found in human HMGCR. Phylogenetic analysis with respect to HMGCR reveals that *C. albicans* lies apart from *H. sapiens.* Thus this was an attempt to identify and understand structure, putative function and evolution of true ortholog of HMGCR from *C. albicans,* a pet target for cholesterol lowering drugs in humans by a detailed bioinformatics analysis.

INTRODUCTION

Candida albicans is one of the most frequent opportunistic fungal pathogen especially of immunocompromised individuals (Cutler *et al.*, 1991, Ruhnke *et al.*, 2006). Although it is a commensal member can colonize nearly every human tissue and organ, causing life threatening invasive infections upon compromised immune status and imbalance in body micro flora etc (Cutler *et al.*, 1991; Berman *et al.*, 2006; Calderone and Fonzi, 2001).

HMG-CoA reductase (HMGCR) is an integral ER enzyme that catalyzes the rate-limiting step in the sterol biosynthetic pathway (Roitelman *et al.*, 1992). In addition to sterols, this pathway also provides the cell with nonsterol metabolites, including isopentyl-adenine, dolichols, ubiquinone and prenyl groups for use in protein translation, glycosylation, electron transport and protein modification. Isoprenoids (or terpenoids) is the largest single family among natural products produced mostly in the plant kingdom. Membrane sterols, fragrance molecules, growth regulators (gibberellins and abscisic acids) and photosynthetic pigments (chlorophyll moiety and carotenoids) are playing important roles in cell growth, development and reproduction in eukaryotes (Kevin Davies 2004).

In mammalian cells, the amount of HMGCR is tightly regulated at multiple levels, including transcription, translation, and post-translation. For example, the degradation of HMGCR is modulated in response to the availability of both sterols and nonsterol metabolites (Lindgren *et al.*, 1985). All animals that have been examined contain a single HMGCR gene while *S. cerevisiae* contains two genes, HMG1 and HMG2, both encode functional HMGCR isozymes (Basson *et al.*, 2009). These genes are estimated to have arisen from a duplication event that occurred approximately 56 million years ago. Interestingly, although the catalytic function of the two isozymes is identical, expression of the two yeast isozymes is regulated differently (Koning *et al.*, 1996).

Increasing incidences of *Candida albicans* infections with increasing immune compromised individuals and emergence of drug resistance among the clinical isolates has dramatically enhanced mortality associated with candidiasis (Singh *et al*., 2013). Scientific community trying to improve the pipeline of novel and potent antifungal agents could not succeed because of availability of few fungal specific drug targets due to eukaryotic nature of *C. albicans*. HMG CoA reductase catalyzes a rate limiting step in a highly conserved pathway among eukaryotes i.e. mevalonate or sterol biosynthetic pathway (Lindgren *et al*., 1985). It is a pet target for cholesterol lowering drugs in humans.

Objective of the present study was to identify and understand structure, putative function and evolution of true ortholog of HMG-CoA from *Candida albicans* by a detailed bioinformatics analysis.

MATERIALS AND METHODS

Identification of true orthologs

HMGCR was functionally characterized in *Schizosaccharomyces pombe* (*S. pombe*) and the predicted protein sequence was used as reference to find "true ortholog" in *Candida albicans*. The methodology by Dhaliwal *et al.*, 2014 was used as the baseline in our entire study. In order to find the orthologs, criterion like highest level of sequence identity and query coverage in the protein sequence, presence of all domains and motifs (similar in size and distance) were considered (Dhaliwal *et al.*, 2014). Thus based on, cDNA and protein sequences of HMGCR from *S. pombe* (Fission yeast, Gene id: L76979, protein id: AAB39277) and its ortholog in *C. albicans* (*Candida*, Gene id: XM_708588, protein id: XP_713681), *Homo sapiens* (Human, Gene id: NM_000859, protein id: NP_000850) and two isoforms from *Saccharomyces cerevisiae* HMG1 (Yeast, Gene id: NM_001182434, protein id: NP_013636), HMG2 (Yeast, Gene id: NM_001182338, protein id: NP_013555) were retrieved from NCBI website. Orthologs were retrieved using query sequence as cDNA sequence of HMGCR from *S. pombe* and tBLASTx tool from NCBI.

Conserved Domain and Motifs Search

To identify the conserved domains present in the protein sequences, CD-Search using RPS-BLAST (Reverse Position-Specific BLAST) tool from NCBI compared a query protein sequence against position-specific score matrices that have been prepared from conserved domain alignments present in the Conserved Domain Database (CDD). Analysis of Conserved Domain and Motif in the retrived orthologs of *Candida albicans* and *Homo sapiens* for HMGCR protein was performed using NCBI CDD tool.

3-dimensional protein analysis of HMGCR from *C. albicans* and *Homo sapiens*

Predicted HMGCR protein sequence from *C. albicans* was used as the query sequence to generate the 3-D structure model using the Protein Homology Recognition Engine (PHYRE2) with default settings (Kelley *et al.*, 2009). The protein model generated using the PHYRE2 was further used to predict the ligand-binding site required for docking studies using the 3DLigandSite tool (Kelley *et al.*, 2009).

Multiple Sequence Alignment and analysis of predicted amino acid sequences of HMGCR

To analyze the differences and similarities in all the selected organisms, amino acid sequences of the predicted HMGCR proteins were aligned using the Clustal Omega multiple sequence alignment tool. Results obtained were viewed in the Jalview tool for multiple alignments. Prediction of the conservation pattern and scores of each amino acids, query in the form of output clustal analysis file generated from multiple alignments and .pdb structure of HMGCR (*C. albicans*) was submitted to the ConSurf server.

Phylogenetic analysis

Clustal Omega tool was used to determine the phylogenetic relationship among the queried sequences.

Table 1: Properties of HMGCR protein of *C. albicans* predicted using ProtParam tool

Sr. no	Property of protein	Description or Value
1.	Number of amino acids	1073
2.	Molecular weight	116417.6
3.	Theoretical pI	6.94
4.	Amino acid composition Ala (A) 101 Arg (R) 37 Asn (N) 49 Asp (D) 49 Cys (C) 15 Gln (Q) 33 Glu (E) 51 Gly (G) 70 His (H) 25 Ile (I) 83 Leu (L) 101 Lys (K) 61 Met (M) 22 Phe (F) 44 Pro (P) 41 Ser (S) 103 Thr (T) 68 Trp (W) 9 Tyr (Y) 30 Val (V) 81 Pyl (O) 0 Sec (U) 0	 9.4% 3.4% 4.6% 4.6% 1.4% 3.1% 4.8% 6.5% 2.3% 7.7% 9.4% 5.7% 2.1% 4.1% 3.8% 9.6% 6.3% 0.8% 2.8% 7.5% 0.0% 0.0%
5.	Total number of negatively charged residues (Asp + Glu)	100
6.	Total number of positively charged residues (Arg + Lys)	98
7.	Atomic composition Carbon C Hydrogen H Nitrogen N Oxygen O Sulfur S	 5208 8288 1386 1557 37
8.	Formula:	$C_{5208}H_{8288}N_{1386}O_{1557}S_{37}$
9.	Total number of atoms:	16476
0.	Extinction coefficients: (Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water) Ext. Coefficient: Abs 0.1% (=1 g/l):	 95075 0.817 (assuming all pairs of Cys residues form cystines)
1.	Estimated half-life:	30 hours (mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo)
2.	The N-terminal of the sequence considered:	M (Met)
3.	Instability index (II):	40.15 (This classifies the protein as unstable)
4.	Aliphatic index:	98.18
5.	Grand average of hydropathicity (GRAVY):	0.084

Clustal Omega is the multiple alignments sequencing tool for phylogenetic alignment tree developed by European Bioinformatics Institute (EBI). It provides information about various evolutionary relations between the genus in their protein or DNA sequence changes in evolution with highly consensus sequences region and phylogram. Clustal Omega uses seeded guide trees and HMM profile-profile technique. To study the phylogenetic analysis of *S. pombe*, *C. albicans*, *H. sapiens* and *S. cerevisiae* query in the form of protein sequences were submitted in the FASTA format to obtain phylogram (Dhaliwal *et al.*, 2014).

Prediction of protein properties

The properties of the proteins such as amino acid composition, theoretical pI value, molecular weight, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) were predicted using the EXPASY server protein analysis tool ProtParam. ProtParam is a tool which allows the computation of various physical and chemical parameters for a given protein sequence.

RESULTS AND DISCUSSION

Retrival of true orthologs of HMGCR

Using the HMGCR gene sequence of *S. pombe*, the reference organism as the query sequence and the tBlastx tool from NCBI, orthologous sequences of HMGCR from *C. albicans* (52%), *H. sapiens* (34%), *S. cerevisiae* both isoform Hmg1 and Hmg2 (84%) retrieved from NCBI database, exhibited the wide range of coverage and sequence identity. The cDNA and the protein sequences (predicted) were saved in FASTA format and used for further analysis (Eiglmeier *et al*., 2005).

Conserved Domain and Motifs search

Using the NCBI CD-search tool and the predicted protein sequence from *C. albicans* and *H. sapiens* as the query, conserved domains and motifs were identified in proteins from both organisms respectively (Paul 2002). The analysis revealed that the predicted HMGCR protein in *C. albicans* consists of three major domains namely HPIH, MMPL superfamily and HMGCR superfamily out of these, the domain HMGCR superfamily consists of five motifs which spans the region from 619 to 1027 amino acids.

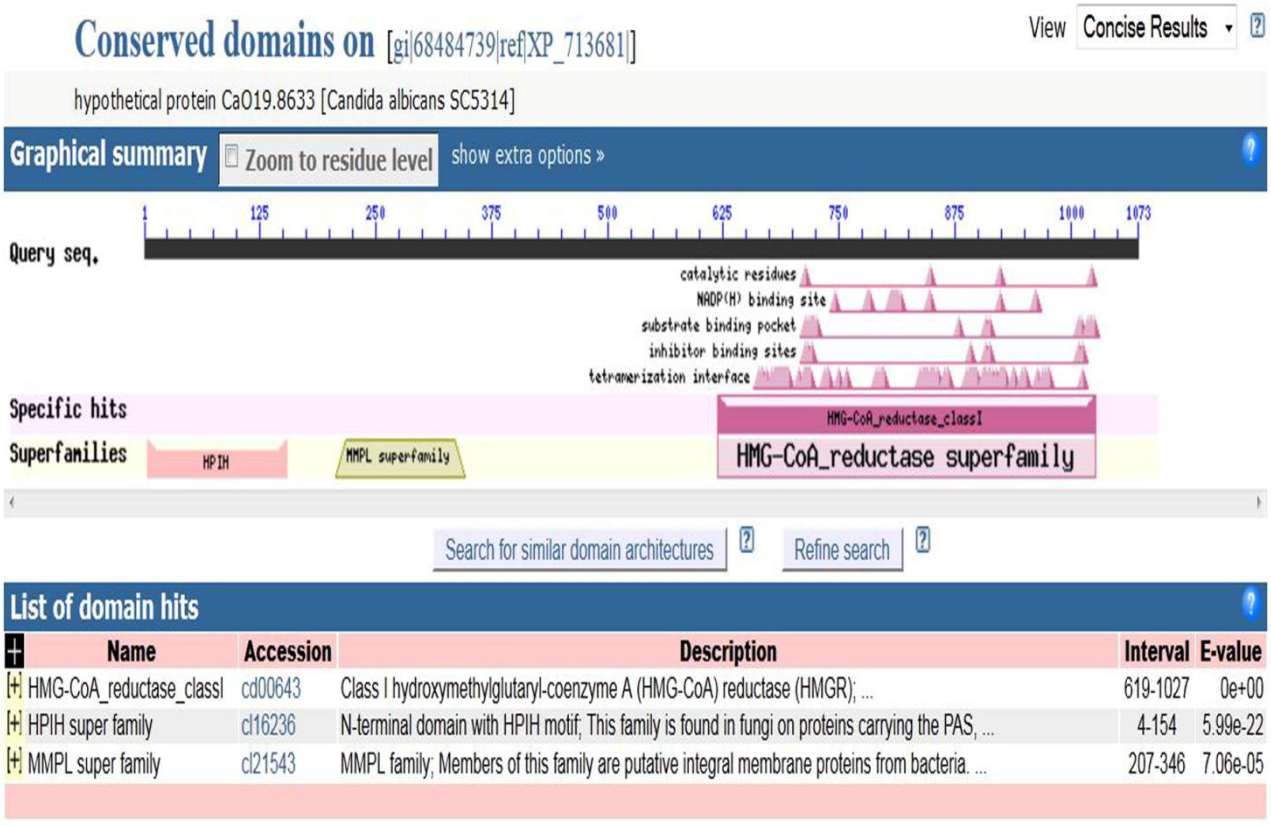

Figure 1: Conserved domains and motifs in *Candida albicans* HMGCR protein

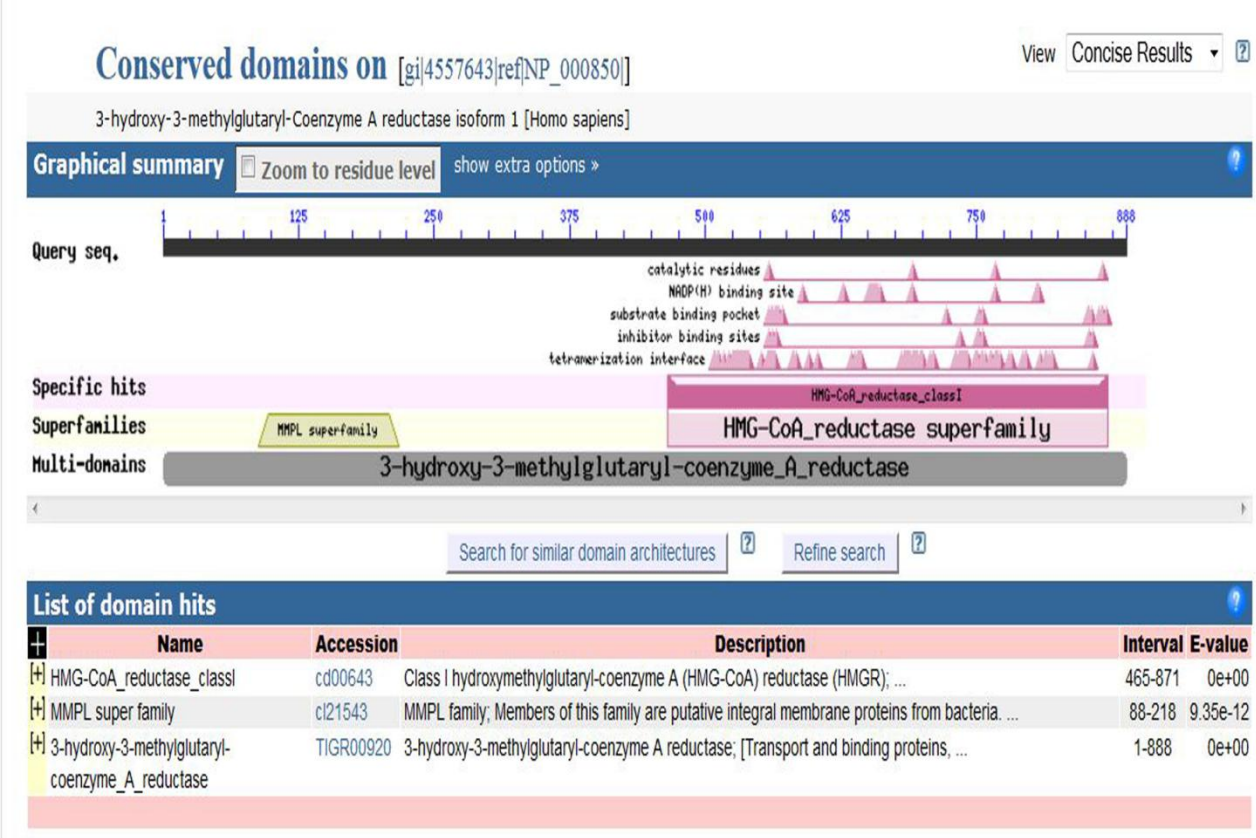

Figure 2: Conserved domains and motifs in *Homo sapiens* HMGCR protein

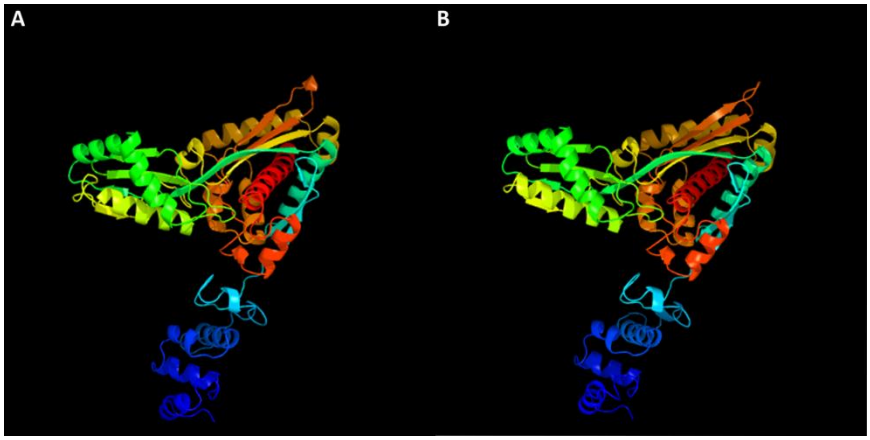

Figure 3: 3-dimensional structure of HMGCR protein of A. *Candida albicans* B. *Homo sapiens*

The domain MMPL superfamily covers the region from 207 to 346 amino acids while, 4 to 154 amino acids in HPIH domain (Figure 1). However in case of *H. sapiens* only two domains were found. Domain HMGCR superfamily consists of five motifs, similar to *C. albicans* however ranging from amino acid 465 to 871 while second domain MMPL superfamily ranges from amino acid 88 to 218 (Figure 2). The third domain HPIH found in *C. albicans* is absent in *H. sapiens*.

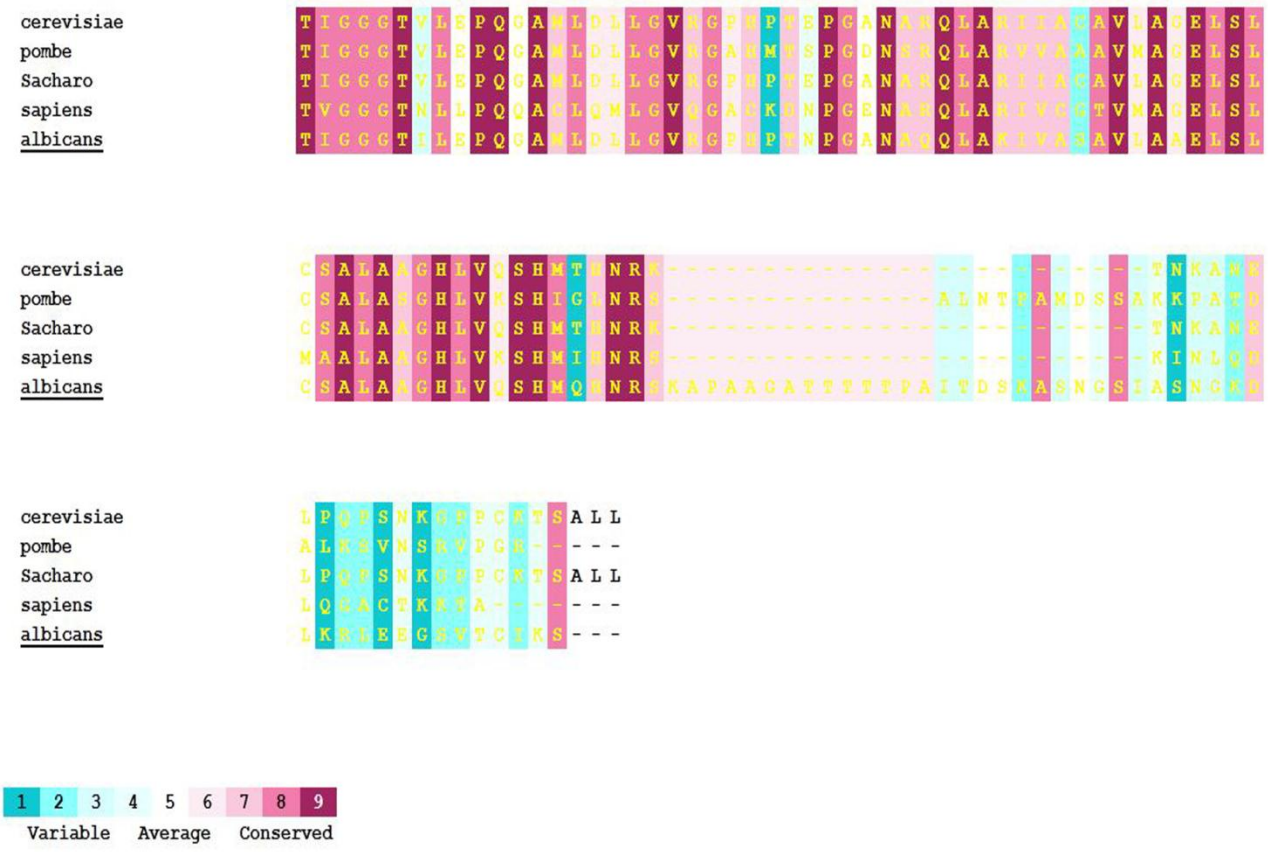

Figure 4: Representative figure of conservation pattern of amino acids of HMGCR protein in selected organisms using ConSurf and Clustal Omega tool.

cerevisiae=Hmg1 isoform of HMGCR from *Saccharomyces cerevisiae*, pombe=HMGCR from *Schizosaccharomyces pombe*, Sacharo= Hmg2 isoform of HMGCR from *Saccharomyces cerevisiae*, sapiens=HMGCR from *Homo sapiens*, albicans= HMGCR from *Candida albicans*.

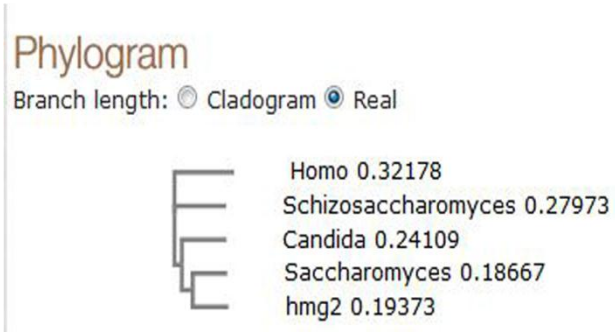

Figure 5: Phylogenetic analysis of selected organisms with respect to HMGCR.

Homo=HMGCR from *Homo sapiens*, Schizosaccharomyces=HMGCR from *Schizosaccharomyces pombe*, Candida=HMGCR from *Candida albicans*, Saccharomyces=Hmg1 isoform of HMGCR from *Saccharomyces cerevisiae*, Hmg2=Hmg2 isoform of HMGCR from *Saccharomyces cerevisiae*.

3 Dimensional protein analysis

3 Dimensional structures of the *C. albicans* and *H. sapiens* HMGCR proteins were generated from the Protein Homology Recognition Engine (PHYRE2). The 3-D structures were generated with 100% confidence using predicted protein sequence (Figure 3A and 3B). Ligand binding site analysis of the HMGCR protein for both the organisms was also performed using 3D ligandsite tool. Analysis revealed that HMGCR protein in *C. albicans* consists of TYR (633), GLU (683) and ASN (684) in the ligand binding, active site while HMGCR of *H. sapiens* although consists of same amino acids in ligand binding site, positions were changed i.e. 479, 528 and 529 respectively.

Multiple Sequence alignment and analysis of predicted amino acid sequences of HMGCR

Multiple sequence analysis for HMGCR protein from *C. albicans*, *H. sapiens*, *S. pombe* and two isoforms of *S. cerevisiae* Hmg1 and Hmg2 of amino acids 1073, 888, 1053, 1054 and 1045 respectively was performed using the Clustal Omega tool (Pais *et al*., 2014). It showed variable as well as conserved amino acid alignment in HMGCR protein, similarly the ConSurf server analysis gave the conservation pattern of amino acids (Armon *et al*., 2001). Representative image for the conservation pattern has been shown in Figure 4. The ConSurf server analysis also gave conservation score (1-9) of the amino acids at each position in the HMGCR proteins, where score 9 denotes highly conserved and 1 being the lowest. Amino acids in the HMGCR domain were highly conserved compared to remaining two domains of HMGCR protein (remaining one domain in case of *H. sapiens*).

Phylogenetic analysis

Phylogram for the predicted protein sequences of HMGCR was deduced using Clustal Omega tool in this study (Pais *et al*., 2014). Relatedness, between the selected organisms with their phylogenetic scores has been shown in the phylogram (Figure 5). Phylogenetic analysis reveals that *C. albicans* is closer to *S. cerevisiae* followed by *S. pombe*, while it lies apart from *H. sapiens*.

Prediction of protein properties

Properties of the HMGCR protein from *Candida* were predicted using the ProtParam tool from ExPasy server (Goli *et al*., 2013). Properties of the predicted protein sequence have been tabulated in table 1.

In silico analysis using the homology, phylogenetic analysis and 3-D structure prediction approach could help the computational biochemists to understand the fundamental information of HMGCR in *C. albicans* compared to other organisms.

Acknowledgement

Authors are thankful to Prof. Dr. Pandit Vidyasagar, Honorable Vice Chancellor, SRTM University, Nanded (MS) India for constant encouragement and support.

References

Armon A, 2001. ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J Mol Biol*. **307:** 447-463.

Basson ME, 2009. *Saccharomyces cerevisiae* contains two functional genes encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc. Natl. Acad. Sci. USA*. **83**: 5563-5567.

Berman J, 2006. Morphogenesis and cell cycle progression in *Candida albicans*. *Curr. Opin. in Microbiol*. **9** (6): 595-601.

Calderone RA and Fonzi WA, 2001. Virulence factors of *Candida albicans*. *Trends Microbiol*. **9**: 327-35.

Cutler JE, 1991. Putative virulence factors of *Candida albicans*. *Annu Rev Microbiol*. **45**: 187-218.

Dhaliwal AK, 2014. Comparative analysis of ABCB1 reveals novel structural and functional conservation between monocots and dicots. *Front Plant Sci.* **5**: 1-10.

Eiglmeier K, 2005. Comparative analysis of BAC and whole genome shotgun sequences from an *Anopheles gambiae* region related to *Plasmodium encapsulation*. *Insect Biochem Mol Biol.* **35** (8): 799-814.

Goli S, 2013. *In Silico* characterization and comparative analysis of PTEN in different species involved in nsclc. *Trends Life Sci*. **2** (1): 26-31.

Kelley LA, 2010. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Prot.* **4** (3): 363-371.

Kevin MD, 2004. Plant pigments and their manipulation. *Annu. Plant Rev*. **14:** 1-342.

Koning AJ, 1996. Different subcellular localization of *Saccharomyces cerevisiae* HMG-CoA reductase isozymes at elevated levels corresponds to distinct endoplasmic reticulum membrane proliferations. *Mol Biol Cell*. **7** (5):769-89

Lindgren V, 1985. Human genes involved in cholesterol metabolism: chromosomal mapping of the loci for the low density lipoprotein receptor and 3-hydroxy 3-methylglutarylcoenzyme A reductase with cDNA probes. *Proc. Natl. Acad. Sci. USA.* **82** (24): 8567–71.

Pais FS, 2014. Assessing the efficiency of multiple sequence alingment programs. *Algor. Mol. Biol.* **9** (4): 1-8.

Paul GY, 2002. Exploring Genomes finding conserved domains. *W. H. Freeman and Company.* 1-4.

Roitelman J, 1992. Immunological evidence for eight spans in the membrane domain of 3-hydroxy 3-methylglutaryl coenzyme A reductase: implications for enzyme degradation in the endoplasmic reticulum. *J. Cell Biol.* **117** (5): 959–73.

Ruhnke M, 2006. Epidemiology of *Candida albicans* infections and role of non-*Candida-albicans* yeasts. *Curr. Drug Targets*. **7**: 495-504.

Singh 2013. Antifungal screening of various spice extracts on azole resistant strains of *Candida*. *Curr. Discov*. **2** (1): 46-52.

How to Cite this Article:

Santosh Kodgire, Vikram Pawar, Nilesh Wagh, Laxmikant Kamble and Gajanan Zore, 2015. *In Silico* Analysis of HMG CO-A Reductase of *Candida albicans* SC5314. *Science Research Reporter*, **5**(1):01-08.