International Journal of Computational Bioinformatics and In Silico Modeling
Vol. 4, No. 4 (2015): 683-690
Research Article
Open Access

ISSN: 2320-0634

Molecular modeling and dynamic simulations of zinc transporter (*tzn-1*) protein from *Neurospora crassa*

P.V Divya Rupa[1], Kunal Zaveri[1], Dr. P. Kiranmayi[2,*]

[1] *Department of Biochemistry and Bioinformatics, Institute of Science, GITAM University, Rushikonda, Visakhapatnam-530045, Andhra Pradesh, India.*
[2] *Department of Biotechnology, Institute of Science, GITAM University, Rushikonda, Visakhapatnam-530045, Andhra Pradesh, India.*

*Corresponding author: P. Kiranmayi; email: kiranmayi.patnala@gmail.com

Received: 04 June 2015 Accepted: 26 June 2015 Online: 06 July 2015

ABSTRACT

Zinc-regulated transporter 1 (*tzn-1*) is a membrane protein that is mainly involved in transport of zinc and iron across the membrane, thereby helping in maintenance of the homeostasis. The previous *in silico* studies suggests that deletion of this protein coding gene results in aconidiation. Even though *tzn-1* gene is not related to conidial formation pathway, still it resulted in aconidiation. To understand this mechanism and its interacting partners the tertiary structure of the protein is essential and hence this study was initiated. In the present study Fold recognition based model of *tzn-1* was built using Photosynthetic Reaction Center from *Rhodopseudomonas viridis* (1PRC_M) as a template. The model was refined by optimization and step wise energy minimization process. The refined model was placed into a POPC/TIP3P membrane system and was simulated for 20ns at 300K temperature with NPT ensemble. Amino acid residues Asp 183, His 185, His 189 and Glu 179, His 199, His 200 are the key active sites for binding of zinc ion. The Root Mean Square Fluctuation from dynamic simulations at the active sites is less and the overall structure quality is 81.2% with 96.2% amino acids in the favorable regions of Ramachandran plot. This study reports the simulated tertiary structure of *tzn-1* and dynamics study in the membrane and aqueous environment.

Keywords: Zinc transporter; *Neurospora crassa*; Fold recognition; Docking; Molecular Dynamics.

1. INTRODUCTION

Neurospora crassa a filamentous fungus undergoes both asexual and sexual reproduction. Asexual reproduction involves the formation of multinucleate macroconidia and uninucleate microconidia. Many genes have shown their significant role in this basic biological process of conidiation. Mutations in any of these genes result in aconidiation process. Various genes are expressed and repressed in the process of conidiation and any mutations in these genes affects the basic sporulation pathway. The genes that block conidial development are *fluffy*, *fluffyoid*, *acon-2*, *acon-3*, *csp-1* and *csp-2* [1,2]. Whereas the genes *con-6*, *con-10*, *con-11*, *con-8*, *con-13* show some crucial role during the early stages of conidial development [3]. Some other genes C2H2 conidiation transcription factor, non-repressor of *con-1*, non-repressor of *con-2*, microcycle blastoconidiation have been observed to play vital role during conidial development [4] a hydrophobin rodlet gene, *eas* is found to be expressed during early stages of development and regulated on exposure to light i.e., circadian rhythm [5].

The accessibility of the genomic sequence of *Neurospora crassa* helped in compiling the Metal Transportome-a complete suite of genes involved in metal transport and homeostasis [6]. The *in silico* analysis of metal transporters showed seven transporters playing vital role in zinc transport. Of the seven, a study was done by Kiranmayi *et al.,* to make a comparative functional evaluation of two hypothetical zinc transporters namely *tzn-1* (NCU07621.7) and *tzn-2* (NCU11414.7). Phenotypic analysis of *tzn-1* knockout mutant resulted in no conidiation in the organism

which is a biological process in all filamentous fungi that reproduce asexually from spores. Supplementation with zinc restores growth in the mutant, but not conidiation suggesting that *tzn-1* gene may also play a key role in this basic biological process [7].

Till date many genes are reported in the conidiation process and their removal results in aconidiation process, but no metal ion transporter is shown to play a prominent role in aconidiation process and it is evident that by knocking a single gene resulted in aconidiating phenotype. To understand the process of aconidiation mediated by *tzn-1* gene at protein level the structure of protein is essential, as structure defines the function. Being a transmembrane protein, no structural data for *tzn-1* is available which lead to this study for prediction of protein structure by *in silico* approach. This predicted structural information will aid in understanding the mechanism of interaction of *tzn-1* with other conidiation genes which leads to aconidiation. The *in silico* approach for predicting the structure here used was Fold Recognition (Threading) approach instead of homology modeling, as no suitable homologous templates were identified for *tzn-1.* The increased availability of computational ability and increased potential functions have enabled the researchers to examine the spot at which the protein molecules generated for simulations are critical and these critical levels of examination determine the structure of a protein [8]. Here we have keyed out the prediction of *tzn-1* structure using threading approach and the predicted structure was set for membrane based molecular dynamics simulation.

2. MATERIALS AND METHODS

2.1 Sequence analysis

The sequence information of the protein *tzn-1* was retrieved from the Fungal Genetics Stock Center (FGSC) [9] in FASTA format. The most evolutionary conserved parts of proteins are known as domains or modules [10,11] thisdefines structure and function of a protein. Domain and family analysis of *tzn-1* domains were computed by ProDom [12], Pfam [13], PROSITE [14,15], InterPro [16]. Physicochemical properties of *tzn-1* were computed by ProtParam, Isotopident and Compute pI/Mw of ExPaSy (Expert Protein Analysis System) [17].

The spatial arrangement of amino acid residues that form intermittent assemblies of alpha helix and beta strands is known as secondary structure of protein. Predicting secondary structure is important as they play an important role in structure of protein and protein folding [18]. The secondary structure of *tzn-1* was computed using PSIPRED (Psi-blast based secondary structure prediction) server a Protein Sequence Analysis Workbench which constitutes different structure prediction methods from UCL (University College London) [19].

2.2 Protein Modeling and Validation

The structure of a protein plays a crucial role in determining the function of a protein, therefore understanding the structure of proteins is significant. Experimentally protein structures are predicted by X-ray crystallography and NMR (Nuclear magnetic resonance) studies, whereas *in silico* approaches include Homology modeling, Fold recognition and *Ab initio* modeling. Here the structure of *tzn-1* protein is predicted by *in silico* based approach. The *tzn-1* protein being a transmembrane protein no appropriate homologous template information is available hence Fold recognition approach is applied to predict the 3-dimensional structure of *tzn-1*.

Fold recognition also known as threading, and the basic concept behind this approach is it compares the target sequence with all the structures present in fold library. Then, based on the compatibility between the target and structures from fold library the best fit template is identified and the structure of the target is modeled. The 3D structure of *tzn-1* is modeled using Prime 3.0 (Schrodinger 2011) [20]. Prime first tries to identify homologous templates, when templates with low identity are retrieved the sequence is then subjected tofold recognition step, in which secondary structures are predicted for target sequence and, templates with similar folds are retrieved from fold library and the resultant folds are aligned, of them the best template is considered that with the highest score and then subjected to modeling.

After modeling the predicted structure was subjected to PROCHECK and ERRAT serversfor structure validation. PROCHECK server checks the stereo chemical quality of a protein structure mainly based on two criteria, hydrogen bonding energy and torsion angles (Phi, Psi and Chi) [21]. ERRAT validates the predicted protein structure by separating the correct and incorrect determined regions based on the interactions of atoms with respect to amino acid residues [22].

2.3 Prediction of metal binding sites

The modeled protein is a zinc ion transporter which binds zinc ion and transports the ions across the membrane. To identify the active pocket for binding the zinc ion, the modeled protein was subjected to CHED server [23] which identifies the binding sites of transition metal ions. CHED algorithm works on the principle of identification of a triad of amino acid residues of type Cysteine, Histidine, Glutamic acid and Aspartic acid. The stringent filtration was used in predicting the zinc ion metal binding site of the modeled *tzn-1* protein.

2.4 Molecular docking

To identify whether the zinc ion binds to the predicted binding sites, the modeled protein as a receptor and zinc ion as a ligand molecule was subjected to molecular docking. The docking was performed using a PATCHDOCK [24] with all the default parameters.

2.5 Molecular Dynamics Simulations

Desmond program [25] was used to perform all Molecular Dynamics (MD) calculations. This program is capable of exploiting high degree of computational parallelism as it uses a specific neutral territory method called Midpoint method [26]. To infer the interactions of amino acids in protein OPLS_2005 force-field [27] was used and *tzn-1* being a transmembrane protein initially membrane was set of computations. The orthorhombic system was selected with NPT Ensemble. All the default selections were selected to run the equilibration of the system which includes a series of process like restrained minimization and MD simulations which are designed to slowly relax the system without undergoing a change from the initial coordinates of the protein. The complete system was set to a temperature of 300K, 1 atmospheric pressure and 8664 water molecules were added to the system. The complete system with 45812 atoms was set for MD simulations for a time period of 20.002ns. Finally the RMSD (Root Mean Square Deviation) and energies were calculated to understand the structure and behavior of the protein molecule.

3. RESULTS AND DISCUSSION

3.1 Sequence Analysis

The translated protein sequence of *tzn-1* protein was retrieved from the FGSC database with the FGSC Locus number NCU07621.7 having the sequence length of 418 amino acids.Protein domains define the structure and function of a protein. Domains for *tzn-1* were analyzed by different web servers like ProDom, Pfam, InterPro and PROSITE. Analysis from these servers revealed that *tzn-1* constitute zinc transporter domain, belonging to the ZIP family. Functionally ZIP family domains transports the metal ions like cadmium, zinc, iron and manganese across the membrane and was first identified in plants [28]. The identified zinc transporter was found to have maximum similarity with the zinc transporter domain of *Lachancea thermotolerans* with domain ID C5DDE5_LACTC [29]. As predicted here within the stretch of 3-316 amino acid residues zinc transporter domain is present in *tzn1*.

The physicochemical properties of *tzn1* were computed using Isotopident, Compute pI/Mw and ProtParam from ExPaSy server (Table 1). The molecular weight for *tzn-1* was observed as 45385.7 and isoelectric point was calculated to be 5.08, theseproperties can be further helpful while isolating the proteins using electrophoresis. The Extinction coefficient was observed as 52870 and stability index to be 36.44 which indicates the protein is stable. This protein constitutes higher concentrations of valine, isoleucine, alanine and leucine as the aliphatic index is found to be very high. The GRAVY index hasfound to be 0.233, a positive value which indicates that the protein is hydrophobic nature.

Table 1. Predicted physicochemical properties of *tzn-1*

S.No	Physicochemical Properties	Predicted Property
1.	Number of amino acids	418
2.	Molecular weight	45385.7
3.	Theoretical pI	5.08
4.	Total number of negatively charged residues (Asp + Glu)	44
5.	Total number of positively charged residues (Arg + Lys)	23
6.	Formula	$C_{2059}H_{3150}N_{526}O_{600}S_{16}$
7.	Total number of atoms	6351
8.	Extinction coefficients, assuming all pairs of Cys residues form cystines	52870
9.	Extinction coefficients, Assuming all pairs of Cys residues are reduced	52370
10.	Estimated half-life	20-30 hrs
11.	Instability index	36.44
12.	Aliphatic index	96.41
13.	Grand average of hydropathicity (GRAVY)	0.233
14.	Monoisotopic mass	45356.768

Legend: *Tzn-1 protein constitutes of 418 amino acids with the molecular weight of about 45385.7 Daltons, half-life of the protein after synthesis is for about 20-30hrs, relative volume occupied by aliphatic side chains is given by aliphatic index and is about 96.41 and the hydropathy is 0.233.*

Secondary structures represent the number of helix, sheets, coils which defines the tertiary structure of a protein and hence the secondary structures are computed by PSIPRED server. The predictions revealed that 12 alpha helixes are present, of which eight helices are highly confident helix predicted for *tzn-1* and another four are short stretches of helices with low confidence value and the beta sheets that are predicted show very less confidence (Figure 1).

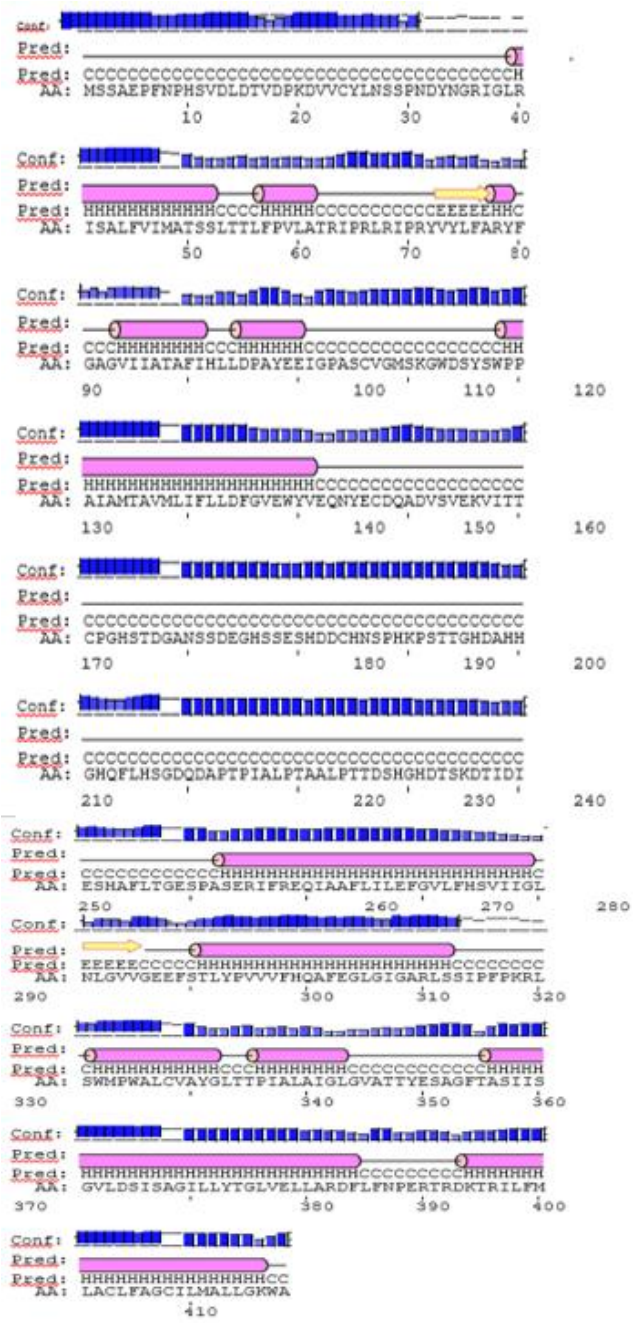

Figure 1. Secondary structure predicted for *tzn-1* using PSIPRED

Legend: In the image pink color represents alpha helices, yellow color arrows represents beta sheets, black lines represent coils and the blue bars represents the confidence levels of predicted secondary structure.

3.2 Protein Modeling and Validation

The protein structure for *tzn1* was modeled using the Prime module ofSchrodinger suite. The target protein *tzn-1* was subjected for the identification of the homologous template by BLAST search which could retrieve the homologous structures with very low identity, so instead of homology modeling the threading approach was considered to model the protein. In threading first the secondary structure of *tzn-1* is predicted by using run SSP option, the results retrieved here were similar to that of predictions from PSIPRED predictions. Then, based upon these secondary structure templates were identified from the fold library and the best template identified was the Photosynthetic Reaction Center from *Rhodopseudomonas viridis* (1PRC_M) [30] with the score of 325.

This identified 1PRC_M template was further used for modeling *tzn-1* protein by following homology modeling approach. Now again same first step is repeated but instead of finding the homologs, the template structure predicted by threading is used and then sequence of the template and targets were aligned and finally by using model build option the structure of *tzn-1* is predicted and the structure represents 12 helices (Figure 2) which are similar to that of secondary structure predicted by PSIPRED server. The protein molecule was modeled only for domain region (51-355) because of lack of homologous structures and folds similar to that of target protein.

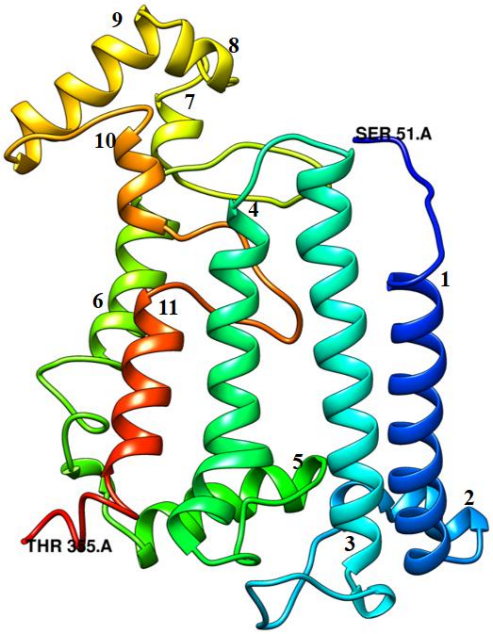

Figure 2. Modeled *tzn-1* protein structure

Legend: Modeled structure of *tzn-1* representing 12 helices as predicted by PSIPRED server.

The obtained 3D structure is then validated using PROCHECK and ERRAT. The validation tools represent that the protein structure that is modeled is satisfactory as the Ramachandran plot derived from PROCHECK analysis represents that about 96.2% of amino acid residues fall in favored region (Figure 3) and ERRAT validates the overall structural quality to be 81.22% (Figure 4).

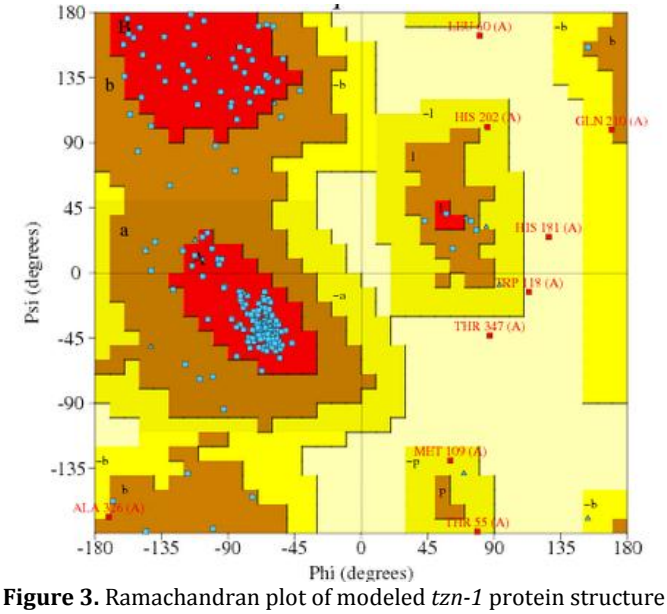

Figure 3. Ramachandran plot of modeled *tzn-1* protein structure

Legend: The image represents that about 83.8% of residues are in most favored regions and 12.4% residues are in additionally allowed region.

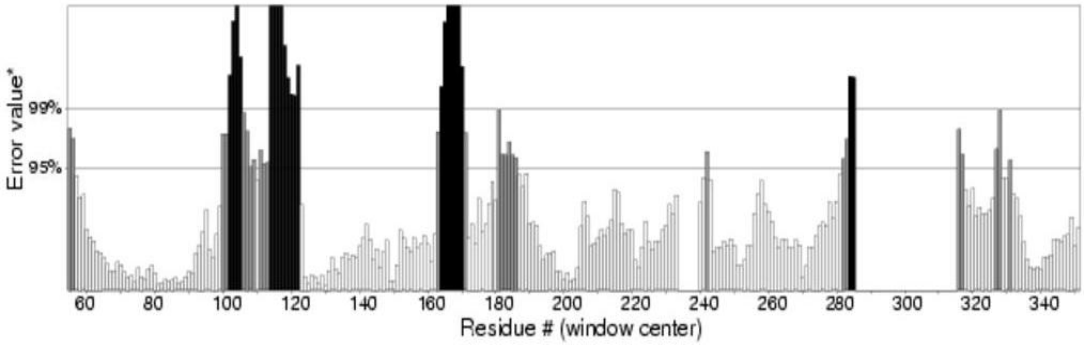

Figure 4. Overall structure quality of modeled *tzn-1* predicted by ERRAT

Legend: The image represents the overall structural quality of *tzn-1* is 81.57 %

3.3 Prediction of metal binding sites

As *tzn-1* is a zinc ion transporter protein, therefore to identify the residues that are active in binding zinc metal and act as active sites for the transport of zinc were predicted by CHED server. In CHED server the stringent filtration is used, and the results revealed that the residues that can bind zinc metal strongly are two triads they are **Asp 183, His 185, His 189** and **Glu 179, His 199, His 200** (Figure 5). According to Kiranmayi *et al.,* Zinc metal binds to Asp and His residues [7] and CHED server also predicted similar residues.

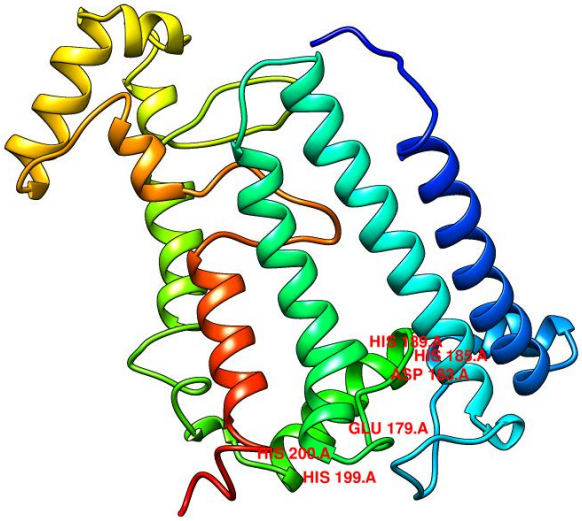

Figure 5. Zinc ion binding sites predicted by CHED server

Legend: The image represents two binding triads predicted by CHED – (**Asp 183, His 185, His 189)** and **(Glu 179, His 199, His 200)**

3.4 Molecular Docking

Docking of zinc to predicted structure was carried out by an online server PATCHDOCK to check whether the zinc ion binds into the predicted binding sites. The zinc ion has interacted with His 199 (Figure 6) which indicates that zinc ion haseffectively bound with the second predicted catalytic triad (Glu 179, His 199, His 200). The PATCHDOCK score was observed to be 2056 with an area of 221.6 and Atomic Contact Energy (ACE) to be -38.9.

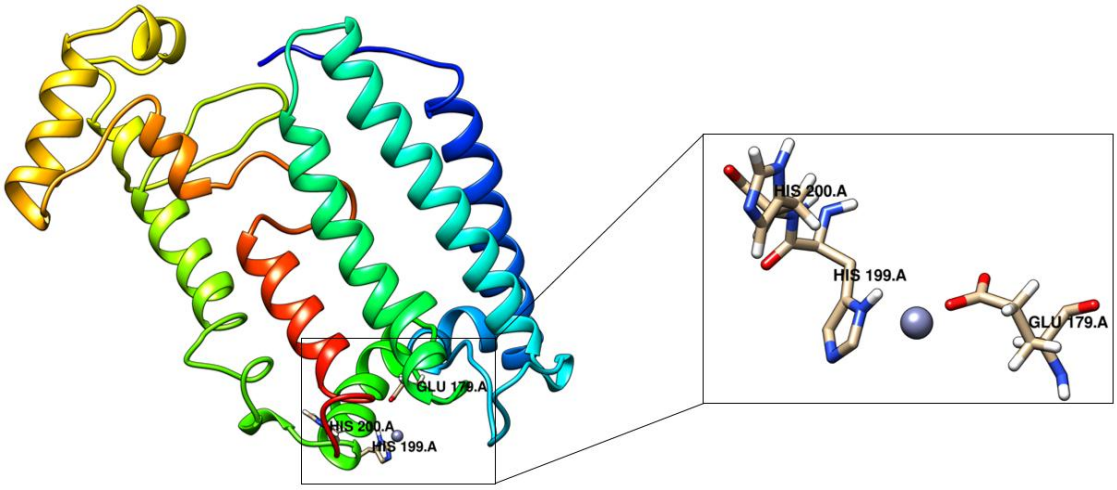

Figure 6. Molecular docking of modeled *tzn-1* with zinc ion

Legend: The image represents that zinc ion binds to the catalytic triad Glu 179, His 199 and His 200.

3.5 Molecular Dynamics Simulations

The modeled protein was simulated in the pre-equilibrated POPC/TIP3P membrane system with orthorhombic boundary conditions for a time period of 20.002ns by maintaining 300K temperature and 1 atmospheric pressure. The simulation system consisted of about 45812 atoms. To understand the quality, dynamic behavior and structure of the protein RMSD, RMSF and Energy are calculated.

Root mean Square Deviation (RMSD) was calculated to measure the global backbone deviation (Figure 7), which calculates the c-α coordinates of the optimal rigid body and for better understanding the local changes in the protein Root Mean Square Fluctuation (RMSF) was calculated (Figure 8). RMSF calculates specific positions of each and every atom in the macromolecule. RMSD of modeled *tzn-1* in simulation

studies with respect to time have shown to be in accepting atomic deviation of about 1-5 angstroms and RMSF fluctuations indicates that the ends of the protein N- and C- terminals with loop regions show higher fluctuations than the protein with secondary structures as they are rigid. This stands for the active site of the protein has a stable conformation with not much of fluctuations.

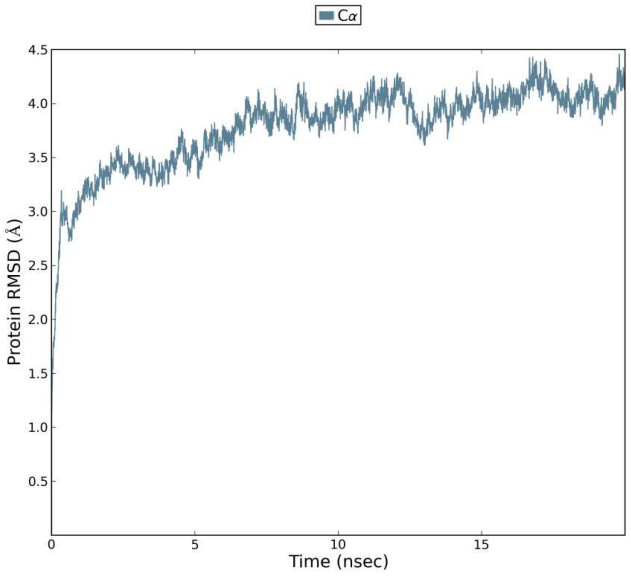

Figure 7. RMSD calculated for modeled *tzn-1* in simulation studies
Legend: RMSD of *tzn-1* is in between 1.5-4.5 angstroms

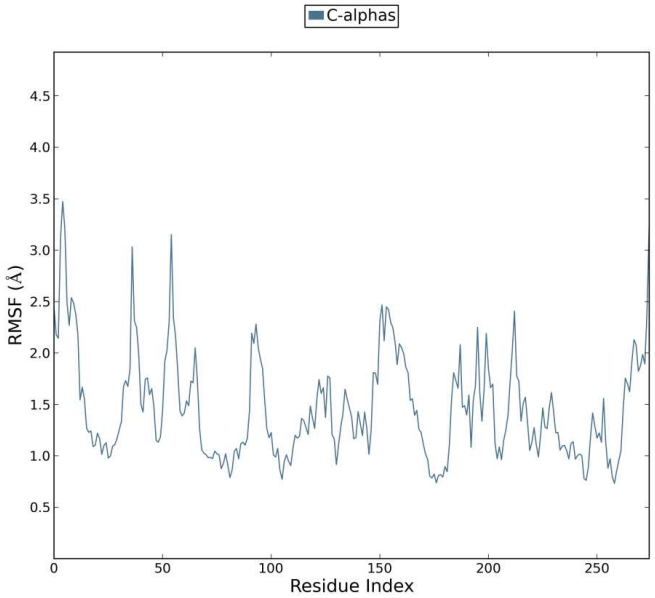

Figure 8. RMSF calculated for modeled *tzn-1* in simulation studies
Legend: RMSF of *tzn-1* ranging from 0.8-3.5 angstroms, the active region of the protein shows less fluctuations which indicates the stable region.

4. CONCLUSION

In this study we were able to model a reliable structure of *tzn-1* protein regardless of low sequence identity with template sequence. The knowledge of folds onproteins was used tomodel *tzn-1* protein and the modeled structure represented 12 helices predicted by PSIPRED server. As per Kiranmayi *et al.,* the zinc metal ions bind to the catalytic triad constituting Asp, His and Glu residues, the modeled structure also represented the same region as predicted by CHED server. The docking of modeled protein with zinc ion also showedthat zinc binds to His 199 which is a key residue of the catalytic triad comprisingGlu 179, His 199 and His 200. MD simulations showed a stable RMSD of 1-5 angstroms and the RMSF showedthe active regions having stable conformation with low fluctuations between 150-200 amino acid residues. This modeled protein would further help us to understand the aconidiation process by protein interaction studies with proteins involved in the conidiation pathway.

5. ACKNOWLEDGMENTS

This work was supported and funded by the Department of Science and Technology, Government of India with grant number SR/FT/LS-105/2012 to PK. The computational part of this work was executed using the Sun Workstation provided by the Bioinformatics facility from Department of Biochemistry and Bioinformatics, Institute of Science, GITAM University. PK acknowledges the support from a Project operated within the DST-Fast track for the Young Scientists Program. A commercial version of Prime 3.0 module from Schrodinger software which was procured under UGC-Project (F.No.42-669/2013) was used for modeling of protein. PK also acknowledges the constant support provided by Schrodinger team. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

6. REFERENCES

1. Lori Bailey-Shrode and Daniel J(2004). Ebbole1- The fluffy Gene of Neurospora crassa Is Necessary and Sufficient to Induce Conidiophore Development Genetics. 166: 1741–1749
2. Alejandro correa and Deborah bell-pedersen (2002). Distinct signalling pathways from the circadian clock participate in regulation of Rhythmic conidiophore development in n.crassa- Eukaryotic cell 1(2), 273-280
3. Anne Roberts and Charles yanofsky (1988). Genes expressed during conidiation in n.crassa: characterisation of con-8, nucleic acid research. 17(1):197-214
4. Gregory O. Kothe and Stephen J(1998). free the isolation and characterisation of nrc-1 and nrc-2, two genes encoding protein kinases that control growth and development in *N.crassa* genetics 149: 117-130
5. Frank-Roman Lauter, Vincenzo E.A. Russoand Charles yanofsky(1992). Developmental and light regulation of eas, the structural gene for the rodlet protein of N.crassa. Genes and development. 6:2373-2381
6. Kiranmayi P and Maruthi Mohan P (2006). Metal Transportome of Neurospora crassa. Insilico Biology. 6:169-180
7. Kiranmayi P, Anand T, SagarKP, et al.,(2009). Functional characterization of *tzn1* and *tzn2*-zinc transporter genes in *Neurospora crassa.* BioMetals. 22(3):411-420
8. Karplus M and Kuriyan J. Molecular dynamics and protein function (2005). PNAS. 102(19):6679–6685
9. "The *Neurospora crassa* e-Compendium" Created by Alan Radford, Tim Beck, John Davies and Nuria Navarro-Coy and Curated by Alan Radford, The University of Leeds
10. Doolittle RF and Bork P (1993). Evolutionarily mobile modules in proteins. Sci Am. 269(4):50–6
11. Kahn D, Rezvoy C and Vivien F (2008). Parallel large scale inference of protein domain families Proceedings of the 14th International Conference on Parallel and Distributed Systems. Melbourne, Australia, IEEE. 72-79
12. Catherine Bru, Emmanuel Courcelle, Se´bastienCarre`re, et.al.,(2005). The ProDom database of protein domain families: more emphasis on 3D. Nucl Acids Res.33:212–215
13. Punta M, Coggill PC, EberhardtRYet.al., (2012). The Pfam protein families database. Nucl Acids Res. 40(D1):290-301
14. Sigrist CJA, de Castro E, Cerutti L, et.al., (2012). New and continuing developments at PROSITE. Nucl Acids Res.1–4
15. de Castro E, Sigrist CJA, Gattiker A, et.al., (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucl Acids Res.34:362–365
16. McDowall J and Hunter S (2011). InterPro Protein Classification Bioinformatics for Comparative Proteomics Methods in Molecular Biology.694:37-47
17. Gasteiger E, Hoogland C, Gattiker A, et.al., (2005). Protein Identification and Analysis Tools on the ExPASy Server; (In) John M Walker (ed): The Proteomics Protocols Handbook, Humana Press.571-607
18. Yong-Yun Ji and You-Quan Li (2010). The role of secondary structure in protein structure selection. The European Physical Journal E. 32(1):103-107
19. Jones DT (1999). Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol.292: 195-202
20. Jacobson MP, Friesner RA, Xiang Z, et.al., (2002). On the Role of Crystal Packing Forces in Determining Protein Sidechain Conformations. J Mol Biol. 320:597-608
21. Morris AL, MacArthur MW, Hutchinson EG, et.al., (1992). Stereochemical quality of protein structure coordinates. Proteins.12(4):345-64
22. Colovos C and Yeates TO (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci.2(9):1511–1519
23. Babor M, Gerzon S, Rahev B, et.al.,(2008). Prediction of transition metal-binding sites from apo protein structures. Proteins. 70:208-217.
24. Schneidman-Duhovny D, Inbar Y, Nussinov R, et.al., (2005). PatchDock and SymmDock: servers for rigid and symmetric docking.Nucl. Acids. Res. 33:363-367.
25. Kevin JB, Edmond C, Huafeng X, et.al., (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. In: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. ACM,Tampa
26. Bowers KJ, Dror RO, Shaw DE (2006). The midpoint method for parallelization of particle simulations. J Chem Phys.124:184109–184111
27. Jorgensen WL, Maxwell DS, Tirado-Rives J(1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc.118:11225–11236
28. Guerinot ML (2000). The ZIP family of metal transporters. Biochimica et Biophysica Acta (BBA) – Biomembranes. 1465(1–2):190–198
29. Souciet JL, Dujon B, Gaillardin C, et.al., (2009). Comparative genomics of protoploid Saccharomycetaceae. Genome Res.;19:1696-1709
30. Deisenhofer J, Epp O, Sinning I, et.al., (1995). Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from Rhodopseudomonas viridis. J Mol Biol.246(3):429-57

*****