JACOBS PUBLISHERS

OPEN ACCESS

Jacobs Journal of Bioinformatics and Proteomics

Research article

In Silico Analysis and 3d Structure Prediction of mitochondrial RHO GTPase 2 Protein of *Danio rerio* (zebra fish) by Homology Modeling

Rumpi Ghosh[1], A.D. Upadhayay[1], A.K. Roy[1*] and Mamta Singh[1]

[1]*Bioinformatics Centre, College of Fisheries, CAU (I), Lembucherra, Tripura*

**Corresponding author: Dr. A.K. Roy, Bioinformatics Centre, College of Fisheries, CAU (I), Lembucherra, Tripura,*

Email: akroy1946@yahoo.co.in

Received: 09-03-2016

Accepted: 11-07-2016

Published: 11-21-2016

Abstract

The present study deals with Insilco characterization of homology-dependent modeling and 69739.7 kDa Mitochondrial RHO GTPase proteins of *Danio;rerio (Brachydanio rerio) not attempted earlier for this species*. The secondary structure prediction reveals the extended 43 coil Regions, 24 helix region, and 13 strands. The 3D structure of this protein was generated using Deep view/Swiss PDB viewer 3.7 by homology modeling. The predicted 3D model showed that 88.1% residues fall in the most favored region, 11.3% in additional allowed region, 0.6% generously allowed region and 0% fall in the disallowed region. Findings like isoelectric point (pI) 5.34, instability index 45.76 and aliphatic index 92.11.3D structure information of Rho GTPase protein will help us to know the role of protein in zebra fish and interaction of their ligands. The predicted structure can also be used for molecular docking studies for more insight into the structure. The rho GTPase protein analyzed in the study showed that this is an unstable protein. The resultant structure was evaluated by PROCHECK. Finally, the 3D structure was predicted by Swiss model server.

Keywords: Zebra fish *Danio rerio* (Brachydanio); Homology modeling; Mitochondrial Rho GTPase 2; Ramachandran plot; Secondary structure; 3D structure

Introduction

The zebra fish (*Danio rerio*) is a tropical freshwater fish belonging to the minnow family(Cyprinidae) of the order Cypriniformes. Native to the Himalayan region, it is a popular aquarium fish, which is frequently sold under the trade name zebra *Danio*. The species arose in the Ganges region in eastern India and commonly inhabits in streams, canals, ditches, ponds, and slow-moving or stagnant water bodies. Zebra fish is omnivorous, primarily eating zooplankton, phytoplankton, insects and insect larvae, although they can eat a variety of other foods. The zebra fish is also an important and widely used vertebrate model organism used in scientific research,(notably, developmental biology, toxicology, reproductive studies, teratology, genetics, neurobiology,environmental sciences, stem cell research, regenerative medicine, and evolutionary theory) and is the first vertebrate to be cloned. It is particularly notable for its regenerative abilities and has been modified by researchers to produce several transgenic strains. The strength of this fish includes its high degree of genetic conservation with humans and its simple, inexpensive maintenance. Additionally, gene expression can be easily manipulated in zebra fish embryos, and their transparency allows for observation of developmental processes.

Mitochondrial Rho GTPase 2 is an enzyme encoded by the RHOT 2 genes. RHOT2 is a member of the Rho GTPase family and one of two isoforms of the protein Miro: RHOT1 (Miro1) and RHOT2 (Miro2). Probably it is involved in the control of anterograde transport of mitochondria and their sub cellular distribution. Rho GTPase is known to play key roles in the modulation of a wide range of cellular processes like prolifera-

tion, apoptosis, cell migration, membrane trafficking, cytoskeleton rearrangements and transcriptional regulation. Through its key role in mitochondrial transport, RHOT2 is involved in mitochondrial homeostasis and apoptosis. On the basis of various structural and physicochemical parameters assessment, 3D structure information of Rho GTPase protein helped to understand the role of protein in zebra fish and interaction of their ligands. 3D structure prediction of protein requires X-ray crystallography and NMR spectroscopy which is very time-consuming and tedious. In silico method for predicting 3D structure reduces this gap. The predicted structure can also be used for molecular docking studies for more insight into the structure.

Materials and Methods

For the prediction of 3D structure, the amino acid sequence of Mitochondrial Rho GTPase 2 is downloaded from the Uni prot KB (NP_001032768).The sequence length reported being 617 amino acids. A template selection search is performed using Geno 3D server, and Template pdb4c0jA was selected for homology modeling. The Template pdb4c0jA was verified by PDB sum server. The criteria used for template selection is maximum similarity (74%) with the query sequence. Therefore, template pdb4c0jA was selected because it showed highest similarity (74%) with the RHO GTPase protein. Geno 3D web server is an automated protein modeling web server for generating protein 3D model. Protein 3D modeling process in Geno 3D server is made using NPSA 3D sequence homology %. 3D run for similarity search with the PSI-BLAST program, generated up to 3 models. The sequence alignment was generated between Rho GTPase 2 and template protein using deep view/Swiss PDB viewer 3.7. Model quality was estimated by assessing the QMEAN score and Z –score. Model quality based on four main concepts objectives or goals, factors, criteria and evaluation process. QMEAN which stands for qualitative model energy analysis is composite scoring function describing the major geometrical aspects of protein structures. QMEAN was tested on several standard decoy sets including a molecular dynamics simulation decoy set as well as on a comprehensive data set .QMEAN shows a statistically significant improvement over nearly all quality measures describing the ability of the scoring function to identify the native structure and discriminate good from bad models. A Z-score is a measure of how many standard deviations below or above the population mean a row score. A Z-score is also known as a standard score and it can be placed on a normal distribution curve. $Z=x-\mu/\sigma$ where Z is the Z-score x is the value of the element, μ is the population mean, and σ is the standard deviation. The Template pdb4c0jA was selected by PDB sum server and it was verified based on the highest similarity (45%) with a query sequence and based on highest Z-score. Physicochemical properties like molecular weight, Theoretical pI, % Total number of negative positive residue, the composition of amino acids, extinction co- efficient, half-life, instability index, Grand average of hydropathicity (GRAVY) of a linear sequence of Rho GTPase 2 protein was calculated using Expasy's program server (http:/expasy.org/chi-bin/protparam). Finally, the structure was predicted by using Swiss model server.

Results and Discussion

Primary structure prediction

Primary structure of Rho GTPase 2 protein was predicted and its physicochemical properties were analyzed by Expasy's prot param server. The result showed that Rho GTPase protein has 617 amino acid residues and estimated molecular weight is 69739.7 kDa, theoretical isoelectric point (pI) 5.34 which shows that protein is acidic in nature. The isoelectric point is the pH at which the surface of the protein is covered with charge but the net charge of the protein is zero [4]. The maximum and a minimum number of amino acid present in the sequence are Leucine 11.5% and tryptophan 1.1%, the total number of positive and negatively charged residues are (Asp+-Glu)-81 and (Arg+Lys)-59 respectively, the estimated instability index (ll) of protein observed to be 45.76 which classifies the protein as unstable. High aliphatic index (92.11) of protein predicts its stability under a wide range of temperature. The negative value of the grand average of hydropathicity (GRAVY) (-0.181) indicates that protein is non-polar, hydrophilic protein is <-1 and better interaction of the protein with water.

Secondary structure prediction

The secondary structure of Rho GTPase was predicted by using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred) as presented in figure 1. The secondary structure of Rho GTPase 2 revealed the presence of 43 coil Regions and 13 beta-Strand, total 24 alpha helix region was highlighted in the predicted structure. The Rho GTPase 2 domain was divided into a coil (54.44%), helix (16.44%), and strand (29.12%).

Template identification

For homology modeling template was searched through Geno 3D server, and Template pdb4c0jA was selected for homology modeling. And it was verified by PDB sum server (http://www.ebi.ac.uk/pdbsum/). The experimental structures used for the construction of the model where Crystal structure of Drosophila Miro EF-hand and cGTPase domains in the Apo state (Apo-MiroS) (pdb4c0jA) which had 41% identity with target protein was used as a template for comparative homology modeling. Detailed output presented in figure 2.

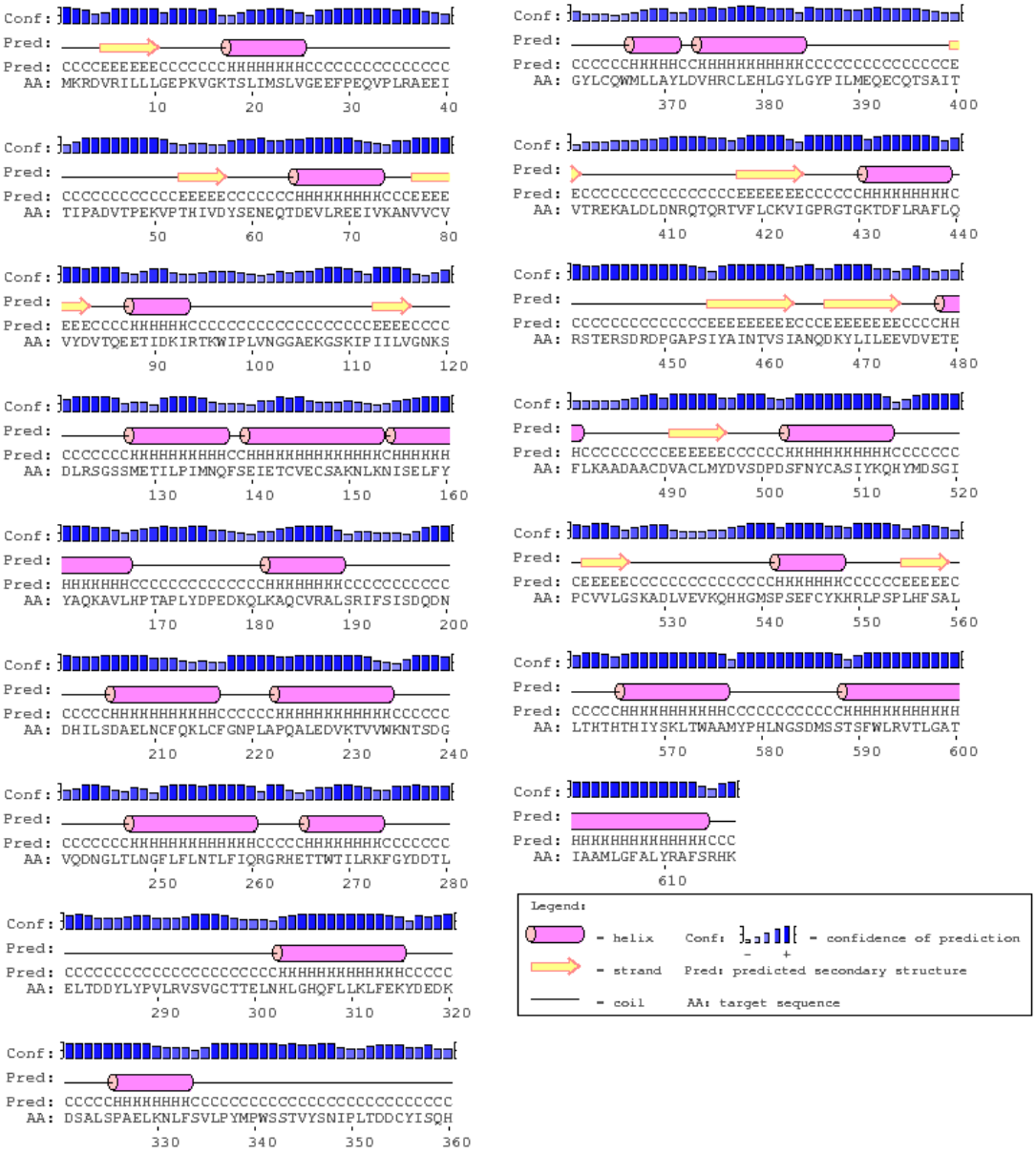

Figure 1. Representation of secondary structure of Rho GTPase protein by PSIPRED.

PDBsum

Sequence search results

	PDB code	Model	Length	%-tage identity	a.a. overlap	z-score	Ligands	Protein name
1.	4c0k (A)	X-ray 2.80Å	404	45.0%	407	1496.7	HSE, UNX, SO4.	Crystal structure of drosophila miro ef hand and cgtpase domains bound to one calcium ion (ca-miros)
2.	4c0j (A)	X-ray 2.82Å	404	45.0%	407	1496.7	UNX, SO4, HSE.	Crystal structure of drosophila miro ef hand and cgtpase domains in the apo state (apo-miros)
3.	4c0l (A)	X-ray 3.00Å	404	45.0%	407	1496.7	UNX, SO4, GDP, HSE.	Crystal structure of drosophila miro ef hand and cgtpase dom bound to one magnesium ion and mg:gdp (mggdp-miros)
4.	4z8y (A)	X-ray 1.90Å	160	27.8%	162	250.2	GDP.	Crystal structure of rab gtpase sec4p mutant - s29v
5.	2e9s (A)	X-ray 1.78Å	164	29.7%	165	246.4	NO3, GDP.	Human neuronal rab6b in three intermediate forms
6.	2ffg(A)	X-ray 1.78Å	164	29.7%	165	246.4	GSP.	The crystal structure of human neuronal rab6b in its active gtpgs-bound form
7.	2fe4 (A)	X-ray 2.30Å	163	29.7%	165	246.4	NO3, GDP.	The crystal structure of human neuronal rab6b in its inactive gdp-bound form
8.	1q16 (C)	X-ray 1.80Å	167	28.7%	164	246.3	GDP.	Crystal structure of sec4-gdp
9.	1q17 (A)	X-ray 2.00Å	168	28.7%	164	246.2	GNP.	Crystal structure of sec4-guanosine-5'-(beta,gamma)- imidotriphosphate

Figure 2. Representation of PDB sum server sequence verified the result.

Homology modeling

The Swiss model is an online tool for 3D Structure prediction based on Homology modeling [2]. The structural alignment was generated using deep view/Swiss PDB viewer 3.7[3] and Swiss model server from sequence alignment between Rho GTPase 2 and template protein. To evaluate backbone conformation, Ramachandran plot (phi/psi) was obtained from PROCHECK analysis (Laskowski. et al., 1993) [6]. The stereo chemical parameters of the protein were calculated by using program PROCHECK [3]. Procheck results for the model using pdb4c0jA as a template are furnished in table-1.

Residues in most favored region		Residues in additional allowed regions		Residues in generously allowed regions		Residues in disallowed region	
%	number	%	number	%	number	%	number
88.1	319	11.3	41	0.6	2	0.0	0

Table 1. Procheck results for a model produced using pdb4c0jA as a template.

From table 1 it can be seen that out of 617 amino acids in the sequence, no one amino acid fell in the disallowed region, representing the accuracy in the predicted structure.88.1% (319). Residues fallen in the most favored region or allowed region is 11.3% whereas 0.6% fall in generously allowed regions and 0% fall in the disallowed region. This is indicative of the degree of accuracy in the predicted structure by check (figure 3). In the case of Soybean Trypsin inhibitor (SBTI) the procheck result shows that out of 168 amino acid in the sequence ,only 1 amino acid has fallen in the disallowed region , which representing the degree of accuracy in the predicted structure. 116 amino acid fell in most favored regions, 11 residues fallen in additional allowed regions, 5 fallen in generously allowed regions [1]. In the case of human gastrin protein residues in the most favored region is 86.2%, allowed region, and outlier region 6.9% which indicate a good quality of predicted model [5].

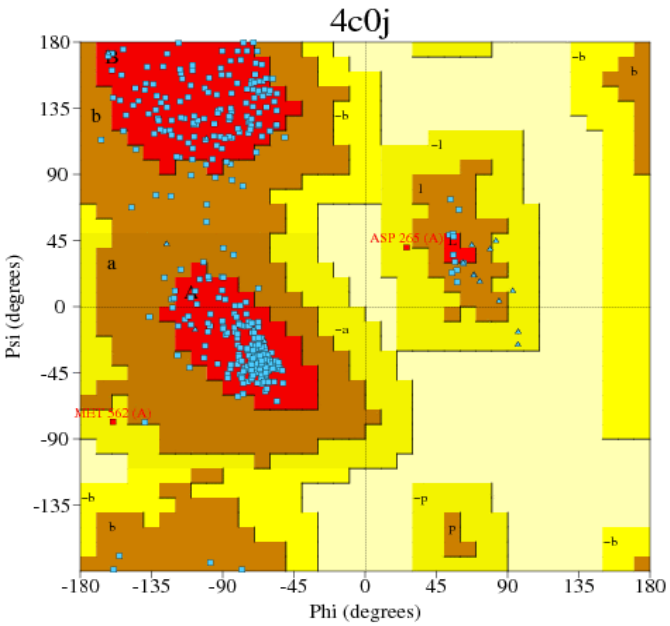

Figure 3. PROCHECK result for best model created using modeller for RHO GTPase.

3D structure prediction

Tertiary structure or 3D structure Prediction was done using Swiss model automated mode for homology modeling, Swiss model server searched for the solved template with similar sequences and best templates were aligned with target amino acid sequence. Templates with best E-value, Percentage similarities and a maximum number of query sequence covered were selected for homology modeling [2]. In the present case of this predicted 3D model, the modeled residue range from 180 -580, sequence identity is 44.47%, E-value is 000e-1 and QMEAN Z-Score is -2.32. The 3D structure of Rho GTPase protein of *Danio rerio* produced by Swiss model used as a template pdb4c0jA presented in Figure 4.

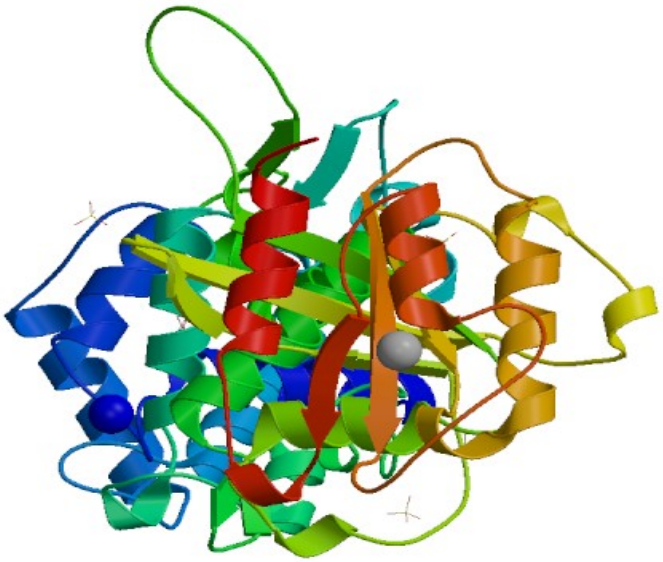

Figure 4. 3D structure representation of Rho GTPase protein of *Danio rerio* generated by Swiss model used as a template pdb4c0jA.

Conclusion

On the basis of various structural and physicochemical parameters assessment, 3D structure information of Rho GTPase protein developed. This will help us to know the role of protein in zebra fish and interaction of their ligands. 3D structure prediction of protein require X-ray crystallography and NMR spectroscopy which is very time consuming, tedious method and generate a large amount of data creating a gap between available sequences and solved structure. In silico method for predicting 3D structure reduces this gap. The new predicted structure can also be used for molecular docking studies for more insight into the structure.

Acknowledgement

Authors are thankful to the Dean, College of Fisheries, Central Agricultural University, Lembucherra, Agartala for encouragement and support. The Financial assistant by DBT, GOI, New Delhi, India for BIF Project under which this study has been carried out, is duly acknowledged.

References

1. Pohekar PG, Chikhale NJ. In silico sequence analysis and prediction of 3D structure of soybean Trypsin inhibitor (SBTI) by homology modeling. IJIRSET. 2013, 2(11): 6505-6508.

2. Bhalla AS, Lal R. In silico sequence analysis, Structure prediction and evaluation of fatty acid Syntheses protein (FAS) of Gallus Gallus (Chicken) by homology modeling. IJSR. 2014, 4(9): 843-847.

3. Singh M, Barman AS. 3d structure prediction of COI protein of air-breathing fish channa marulius. Bioinfolet 2014, 11(1B): 207-209.

4. Vidhya VG, Bhaskar A, Purushothaman P. In silico analysis and 3D modeling of SORD Protein in diabetic Retinopathy. J Comput Method Mol Design. 2011, 1(4): 22-27.

5. Bandari M, Swaroop RSK, Kodavaty J, Readdy CVK. In silico analysis and 3D structure prediction of human gastrin protein. 2013.

6. Laskowski RA, MacArthur MW, Moss DS, Thornton JMJ. App Cry's 26: 283, 1993.