Molecular Modeling of Discoidin Domain Receptor Tyrosine Kinase 2 (DDR2) and ATPase, Na+/K+ Transporting, Alpha 4 Polypeptide (ATP1A4):
Candidate Genes for Autosomal Recessive Cone-Rod Dystrophy (CORD8)

Naureen A K1, Naqsh-e-Zahra S2, Awan W A3

Abstract

Background & Objective: Cone Rod Dystrophy, (CORD8) is an inherited progressive disease that causes degeneration of cones and rod photoreceptor cells of retina, leading to blindness. CORD8 is characterized by loss of color vision, severe photophobia, and epiphora since childhood. The objective of this research study was to predict and evaluate the 3 dimensional structure of most likely candidate genes DDR2 and ATP1A4 for CORD8 disease.

Methods: MODELLER (9v10) was employed to construct three dimensional structures of DDR2 and ATP1A4 candidate gene. Structure evaluation was performed by PROCHECK and ERRAT tools.

Result: Suitable templates 4AG4 and 3B8E with homology of 54% and 79% and E-values of 7e-110 and 0.0 was selected from Protein databank (PDB) for candidate genes DDR2 and ATP1A4 respectively. PROCHECK evaluation tool showed 82.1 %, 76.1 % residues of DDR2 and ATP1A4 respectively in most favored region while ERRAT shows quality factor of 25.474 and 43.924.

Conclusion: The predicted structure might used to explore the structural insights of protein as well as it can lead to computer aided drugs designing. Furthermore, development of hybrid software based on specific data of inherited retinal disorder is major area of focus for identifying most suspected genes involved in CORD8.

Keywords: CORD8, Cone-rod dystrophies, Candidate gene analysis, Insilico analysis, Structure prediction, Bioinformatics.

Introduction

Cone rod dystrophies belong to pigmented retinopathies which are characterized by deposition of retinal pigments localized on macular region of the eye. These dystrophies are illustrated by cone and rod loss leading to color vision defects, night blindness and loss of visual acuity. CRDs are majorly non syndromic but may be part of some other syndromes like Bardet Biedl syndrome and Spinocerebellar Ataxia Type 7 (1). To date, seven loci have been identified for cone–rod dystrophies on chromosomes 17q (2), 19q (3), 18q (4), 17p13 (5,6), 6q (7), 1q12 (8) and 8p11 (9).

1. Department of Biochemistry, PMAS- Arid Agriculture University Rawalpindi, Pakistan
2. Department of Bioinformatics and Biotechnology, International Islamic University Islamabad, H-10 Sector Islamabad, Pakistan
3. IIRS, Isra University Islamabad Campus

Correspondence to:
Naureen Aslam Khattak
University Institute of Biochemistry and Biotechnology Department of Biochemistry, PMAS- Arid agriculture university, Rawalpindi

Mutations in the peripherin/RDS (10), CRX (11,12), RetGC-1 (13), ABCR (14), CNGA3 (15), RPGRIPI (16) and SEMA4A (17) genes have been shown to cause both autosomal dominant and recessive cone–rod dystrophies. DDR2 and ATP1A4 genes are common genes involved in an autosomal recessive CORD8. Discoidin domain receptors (DDRs) belong to the receptor tyrosine kinase family. DDR2 is primarily
found in cells of mesenchymal origin, such as fibroblasts and smooth muscle cells (18). Protein encoded by \textit{ATP1A4} gene belongs to the family of P-type cation transport ATPases. Na+/K+ -ATPases are responsible for establishment and maintenance of the electrochemical gradients of Na and K ions across the plasma membrane which are essential for osmoregulation, for sodium-coupled transport of a variety of organic and inorganic molecules, and for electrical excitability of nerve and muscle. Genome sequencing of various organisms have created a heap of raw data, management of which is time consuming. In order to manage bulk of data, field of bioinformatics arise in combination with different computational techniques. Bioinformatics aid in organization and data analysis of existing information and submission of new data (19,20). Structural data analysis include prediction of secondary and tertiary protein structures, tools for 3D structural alignment (21,22)which can assist in understanding of protein’s function and different protein fold (23,24). In this study, structural analysis of candidate genes \textit{DDR2} and \textit{ATP1A4} was performed by using comparative modeling.

\textbf{Material and Methods}

\textit{Template Selection:} The amino acid sequence of \textit{DDR2} and \textit{ATP1A4} proteins were retrieved from NCBI in FASTA format with lengths of 855 and 1029 residues respectively. For template selection for target protein, BLAST was applied against Protein Databank (PDB). For \textit{DDR2} and \textit{ATP1A4}, template 4AG4 and 3B8E was selected for homology modeling. Alignment and 3D model was built by Modeller program (25). Detailed overview of template selection is listed in table 1. Table 2 shows the summary of tools used in current study.

\textbf{Comparative Modeling:} Three Dimensional structure of both proteins \textit{DDR2} and \textit{ATP1A4} are not resolved yet in Protein Databank (PDB). Comparative modeling techniques were employed to get 3D structure of candidate proteins.

\textbf{Evaluation of Predicted Structure:} Stereo chemical qualities of predicted structures of both proteins were evaluated using PROCHECK(26) and ERRAT (27).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Accession ID & Total Score & Query Coverage & E-value & Max-identity \\
\hline
4AG4 & 344 & 39\% & 7e-110 & 54\% \\
3B8E & 1658 & 96\% & 0.0 & 79\% \\
\hline
\end{tabular}
\caption{Table 1: Template selection for \textit{DDR2} and \textit{ATP1A4} candidate genes}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Sr.No. & Tools/Databases & Output/Function \\
\hline
1 & BLAST & Homology Modeling and Template Selection \\
2 & MODELLER & Structure prediction \\
3 & Chimera & Visualization, Superimposition, Interaction \\
4 & PROCHECK & Structure Evaluation \\
5 & ERRAT & Structure Quality prediction \\
\hline
\end{tabular}
\caption{Table 2: Summary of Tools used in current study}
\end{table}

\textbf{Results}

Role of \textit{DDR2} and \textit{ATP1A4} genes in causing autosomal recessive cone rod dystrophies has been widely researched and documented and is of considerable importance. Table 3 shows the details of templates selected for structure prediction of \textit{DDR2} and \textit{ATP1A4} proteins. Generated structure of \textit{DDR2}}
is displayed in figure 1 and figure 2 shows the structure of \textit{ATP1A4} protein.

![Figure I: Predicted structure of DDR2 protein using 4AG4 template. β sheets are displayed in blue color, α helix in red.](image)

![Figure II: Predicted structure of ATP1A4 using 3B8E template. β sheets are displayed in blue color, α helix in red.](image)

To evaluate the predicted structure, Ramachandran plot and overall quality factor were accessed through PROCHECK and ERRAT respectively (Table 3). Ramachandran plot showed the distribution of amino acids in favoured, allowed and outlier regions. As mentioned in Table 3, most of the residues lie in favoured regions so it gives indication that models are reliable.

![Table 3: Evaluation results for DDR2 and ATP1A4 through PROCHECK program and ERRAT](image)

**Discussion & Conclusion**

3-dimensional structure prediction of candidate genes DDR2 and ATP1A4 involved in CORD8 were performed to understand the mechanism of inherited retinal disorder. Mutational analysis might be performed on the basis of structure prediction to recognized the structure abnormalities raised in DDR2 and ATP1A4 which encodes functional proteins. Predicted 3 dimensional structures can aid in docking studies to proposed a suitable drugs for cure the respective disease.

Identification of causative gene is a necessary step towards the understanding the pathophysiology of CORD8. With the use of candidate gene identification software’s and genetic database it can be reasonably assumed that most genes responsible for autosomal recessive cone rod dystrophy will be known in next few year but this will not be a straightforward step in autosomal recessive case because all the reported genes in CORD8 have no similarity at functional level. Current study is an important step to understand the functional outcomes from structural insights and their relation to the phenotype of the disease with the help of software’s resulting in the development of scientific knowledge regarding Inherited retinal disorder.
References
14. Cremers FPM, van de Pol DJR, van Driel M, den Hollander AI, van Haren FJJ, Knoers NVAM,


