Journal of Population Therapeutics & Clinical Pharmacology

IN-SILICO ANALYSIS AS TOOLS OF MOLECULAR SCREENING OF A MUTATION CAUSING CATARACT.

Najeeb Ullah[1,2], Munir Ahmad Bhinder[1], Bisma Rauff [3], Leena Moaaz Hashim Mohamedain[4,5], Hafsa Sunniya[6], Saba Adam[2], Faiza Munir[1,5], Jeand Baloch[2]

[1]Department of Human Genetics and Molecular Biology, University of Health Sciences, Lahore, Pakistan.

[2]Department of Natural & Basic Sciences, University of Turbat, Turbat (Kech), Pakistan.

[3] Department of Biomedical Engineering, University of Engineering and Technology Lahore, Narowal Campus, Pakistan.

[4]Beijing Key Laboratory for Separation and Analysis in Biomedicine and Pharmaceuticals, School of Medical Technology, Beijing Institute of Technology, China.

[5]Medical Research Centre, Hamad Medical Corporation (HMC), Doha, Qatar.

[6]Xi'an Jioatong University, Xi'an, Shaanxi, China.

Correspondence: Munir Ahmad Bhinder, Email: munirbhinder@yahoo.com

Abstract

Cataract is a disorder of eye lens where an opaque lens causes the blurred vision. Occurrence of cataract is common around the world and its causes range from environmental reasons to genetic mutations. Genetic cataract is heterogeneous in nature and several hundred genes are discovered to this day, reported to cause cataract. This study involves in-silico analysis of an already known cataract causing mutation (c.233G>A) in order to ascertain the impact of this mutation in mutated protein. Bioinformatics tools were employed to calculate and compare physical parameters of mutated protein with that of its normal copy. Calculating physical parameters illustrated that mutated protein has higher polarity, lower hydrophobicity and enhanced refractivity. Conservation analysis identified that resulting mutated residue p.G78G is residing in a highly conserved region and this mutation has damaging impact. Three dimensional structures of both wild and mutant MP19 protein were predicted using I-Tasser tool and structures with highest quality score were selected. Predicted structure of wild MP19 protein was validated through Ramachandran plot and Z-score calculations using Procheck and PROSA web tools, respectively. The mutation (p.G78D) was characterised in wild MP19 protein model and was viewed through Chimera tool. In-silico analysis predicted that mutation p.G78G may bring about changes of clinical significance.

Key Words: Cataract, Conservation Analysis, Protein Structure Analysis, Bioinformatics

1. Introduction

Cataract is amongst the well-known disorders of eyes where affected person has either blurred vision or total loss of vision (Chen et al., 2017). Cataract is common among all population groups and there is no difference amongst different socio-economic classes (Khan et al., 2015; Patel et al., 2017). There is no alternate therapy available for cataract, except for surgical removal of lens. Hence, an intensive range of research has been conducted around the world to explore different therapeutic avenues for the purpose. This has resulted in identification of hundreds of genes and identification of subsequent mutations in these genes associated with genesis of cataract ("Home | Cat-Map | Washington University in St. Louis," n.d.). However, functional analysis of these identifies mutations are not catching up.

This study is a bioinformatics approach where in-silico analysis are carried out for a cataract causing mutation (c.233G>A) re-identified in our previous study (Ullah et al., 2024). That study was an exome sequencing approach, used to screen cataract affected families from Pakistan. The mutation was reported to cause cataract in a family with multiple affected members (Irum et al., 2016). It was a missense mutation in LIM2 gene which encodes a protein named MP19.

Analysis in this study included the conservative analysis of the genomic region where this mutation (p.G78G) has been reported. Furthermore, Physical parameters of protein, both in wild and mutated form, were calculated and compared to gauge the impact of change on this level. The 3D structure of protein is important to ascertain its characteristic properties and functions. However, there was no MP19 protein 3D structure available at PDB database (Berman et al., 2000). Hence, homology modelling technique was utilized to predict the 3D structure of both wild and mutated form of MP19 protein (Ashraf et al., 2018). Predicted 3D structures of protein were submitted to various validation and evaluation tools in order to assure the quality of structure, before using in functional analysis. Protein-protein interaction networking (PPI) is important for functioning of a protein in complex environment and MP19 protein was assessed for functioning in complex molecular environment.

2. Methods

The experimental work was carried out in the Departments of Human Genetics & Molecular Biology at University of Health Sciences, Lahore. The protein sequence of MP19 of human origin was obtained from Ensemble Genome Browser (Martin et al., 2023) (protein ID ENST00000596399.2). The retrieved sequence was used to develop 3D structure of respective protein based on homology modelling. It was ascertained that PDB (Protein Data Bank) has no MP19 protein crystal or NMR structures registered.

Primary physical, chemical and structural properties were determined to get the general picture of protein nature and its relation with other molecules in working environment. Protparam (Wilkins et al., 1999) online tool was used to compute different parameters of primary structure. These features included the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). Additionally, ProtScale (Wiederschain, 2006) was used to calculate and compare polarity, hydrophobicity, hydropathecity and refractivity of protein molecules. These physical parameters were first examined for wild type protein(s) and then these parameters were compared with those of

mutated protein. In order to ascertain the impact that mutation may inflict, the residue position in primary sequence becomes of paramount importance. The mutation becomes of molecular significance if it lies in highly conserved region of the protein. For conservation analysis, protein sequences of various species were obtained from NCBI protein database. These sequences were aligned using ClustalW online tool (Thompson et al., 1994)

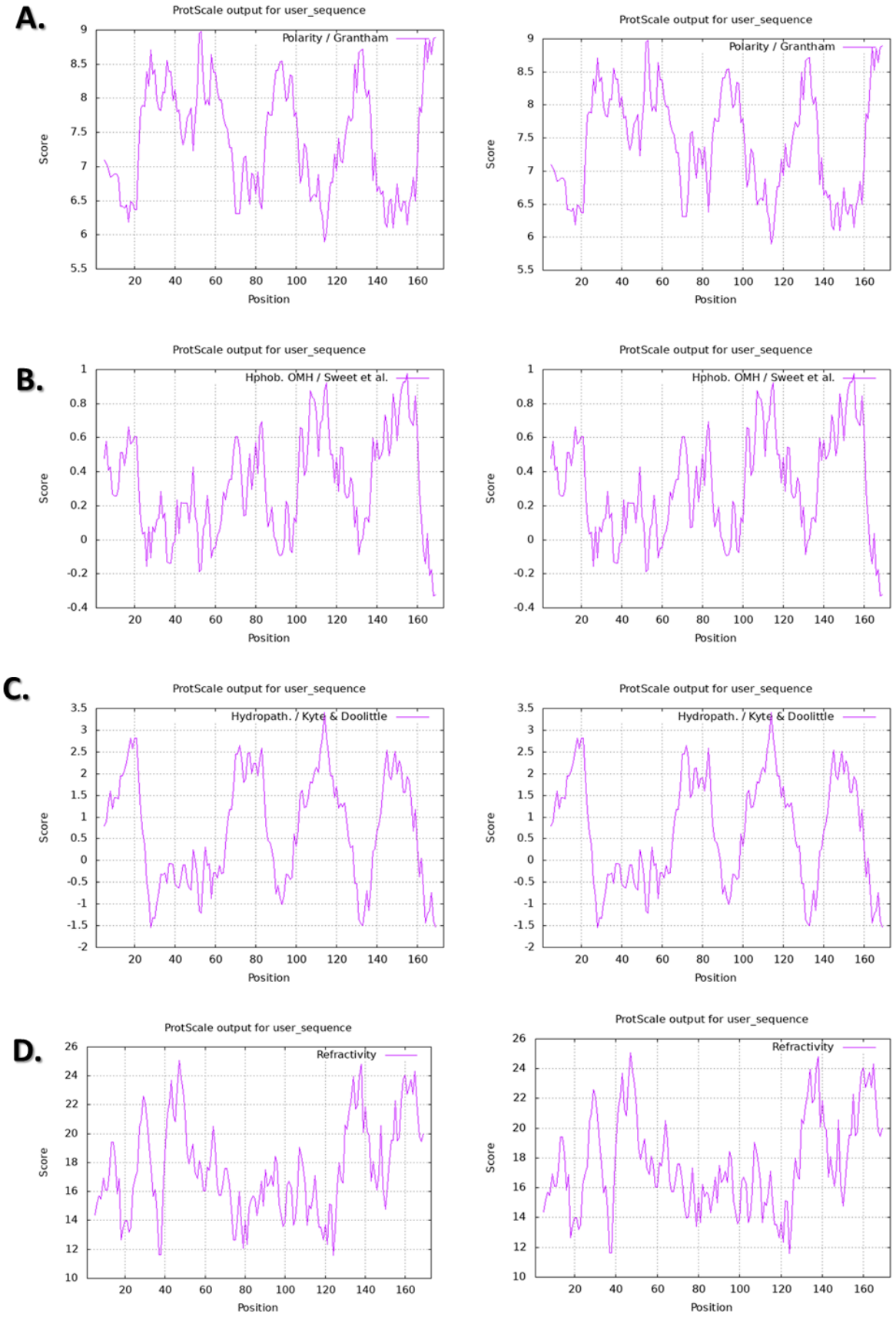
As mentioned before, to the best of our information there are no crystal or NMR structures of MP19 protein available on any of the protein structure databases including Protein Database (PDB). Keeping this information in view, 3D protein structures of both wild type and mutant copies were designed using I-Tasser online tool (Zhang, 2008). 3D models on I-Tasser online server are built based on multiple-threading alignments by LOMETS and iterative template fragment assembly simulations. To evaluate and validate the protein predicted structures, Psi/Phi Ramachandran plots for proteins were plotted via PROCHECK online server (Laskowski et al., 1996). The colouring/shading on the plot represents the different regions the darkest areas in (Figure. 4A & 4B) correspond to the "core" regions representing the most favourable combinations of phi-psi values.

Protein predicted models were assured for its quality by Z-score via ProSA web tool (Wiederstein and Sippl, 2007). It displays values in a plot (shown in Figure 4C) that contains the z-scores of all experimentally determined protein chains in current PDB. It can be used to check whether the z-score of the input structure is within the range of scores typically found for template proteins of similar size. After designing 3D structures, both wild type and mutant copies were superimposed to know the impact of the mutation on the structural changes and functional properties of protein (Pettersen et al., 2004).

STRING, an online database of known and predicted protein interactions (Szklarczyk et al., 2021), is used to calculate the physical contacts of the MP19 protein. The interactions include direct (physical) and indirect (functional) associations. We constructed an extended network based on a medium confidence score of 0.40, which implies that only interactions with more than low level of confidence were extracted from the database and considered as valid links for the PPI (protein-protein interaction) network.

3. Results

LIM2 gene is located on chromosome 19 at its q arm, and it translates into MP19 protein. It is a transmembrane protein holding great significance for the proper functioning of lens. Physical parameters of the protein were examined and it describes MP19 a protein size of 173



amino acids. It molecular weight is 19674.26 and its PI stands at 9.64. Negatively charged

Figure 1. ***Comparison of Physical Parameters between Normal MP19 (Right) and its p.G78D Mutant Copy (Left).*** On X-axis is the position of amino acid residues and on the Y-axis is the value of each amino acid. **A.** Comparison of polarity. **B.** Comparison of hydrophobicity. **C.** Comparison of hydropathicity. **D.** Comparison of Refractivity.

residues in protein sequence of MP19 are four whereas 13 amino acids are positively charged. Extinction coefficient of the protein is 51910 and its instability index is computed to be 37.34, suggesting a stable protein structure. This makes sense as MP19 protein activity in lens requires greater stability. Grand average of hydropathicity of MP19 protein is calculated to be 0.651.

To understand the physical changes that mutation p.G78D brings about in MP19 protein, mutant MP19 was compared with its wild type for four parameters including polarity, hydrophobicity, hydropathacity and refractivity (Gasteiger et al., 2003). The mutant MP19 (p.G78D) protein was found more polar, less hydrophobic and exhibiting more refractivity (Figure 1).

The mutation identified (p.G78D) is a missense mutation and it is of greater importance to know whether the mutation is located in the conserved regions or otherwise. For the purpose, the mutation was subjected to conservation analysis and it was found that p.78 is a highly conserved region and thus holds importance for the proper functioning of the MP19 protein. The conservative nature of the residue and that region of the protein was found un altered in many other organisms (Figure 2).

	p.74	p.75	p.76	p.77	p.78	p.79	p.80	p.81	p.82
HUMAN	G	G	A	W	G	A	Q	D	A
MOUSE	C	A	T	S	G	I	I	M	G
BOVINE	C	A	T	S	G	I	I	M	G
COW	C	A	T	S	G	I	I	M	G
RAT	C	A	T	S	G	I	I	M	G

Figure 2. ***Conservation Anlysis of Residue p.78 of MP19 Protein.*** p.78 is located in highly conserved region of MP19 protein. The residue is found conserved in across other species.

Three dimensional structure of protein is important to know the impact of change and to the best of our information, there are no crystal or NMR structures of MP19 protein available on any of the protein structure databases including Protein Database (PDB). Keeping this information in view, 3D protein structures of both wild type and newly found mutant copies were designed using I-Tasser online tool (Zhang 2008). After designing and validating 3D structures, both wild type and mutant copies were superimposed to know the impact of the mutation on the structural changes and functional properties of protein (Figure 3).

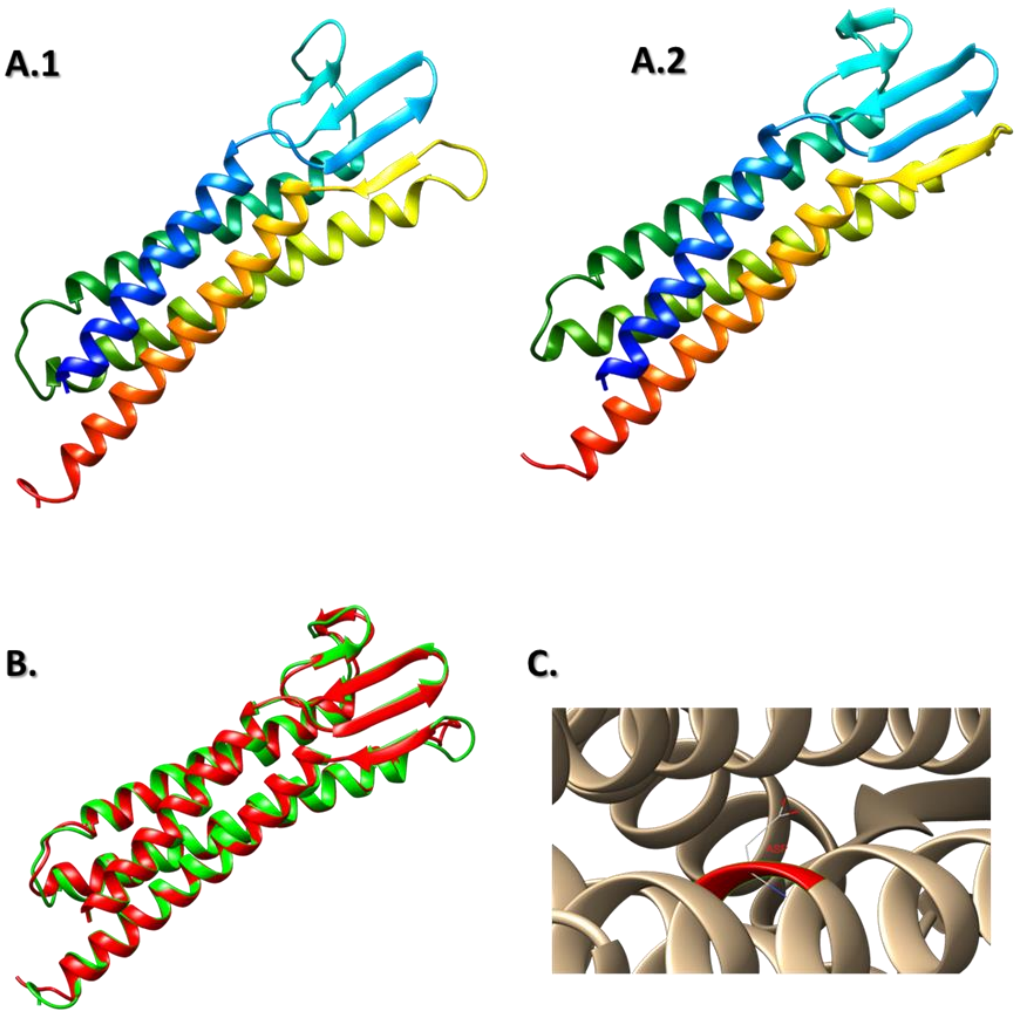
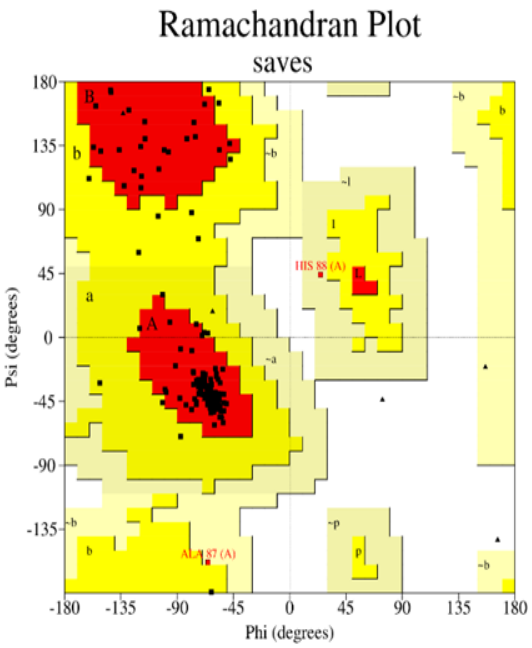


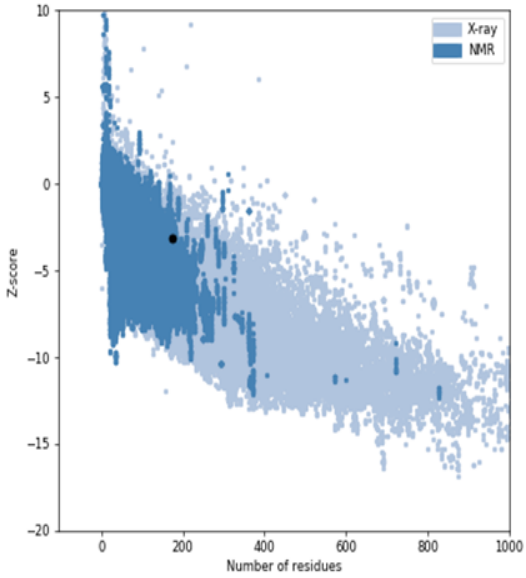
Figure 3. ***Protein Structure Models of MP19 Mutant Protein and its Normal Type*****. A.** Three-dimensional (3D) protein structure of normal MP19 protein(left) and its mutant (right) generated through homology modelling. **B.** Superimposition of MP19 protein(green) and its mutant (red). **C.** Depiction of mutation p.G78D in MP19 protein.

Quality and reliability of the structure was checked using Ramachandran plot and Z-score. The stereo-chemical quality of 3D structure was checked by Ramachandran plot (Figure 4) where analysing included residue by residue geometry and overall structure geometry. The red shaded areas correspond to the "core" regions representing the most favourable combinations of phi-psi values. The percentage of residues in the "core" regions is one of the better guides to stereo-chemical quality. The result of the Ramachandran plot (Figure 4A & 4C) showed 89.5% of residues in the favourable/core region whereas residues in disallowed regions were 9.2%. The stereo-chemical quality of the predicted models was found to be satisfactory. Global model quality was checked by ProsA tool. Z-score value measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations. The Z-score of the wild MP19 protein was -3.12 (Figure 4B).

A.



B.



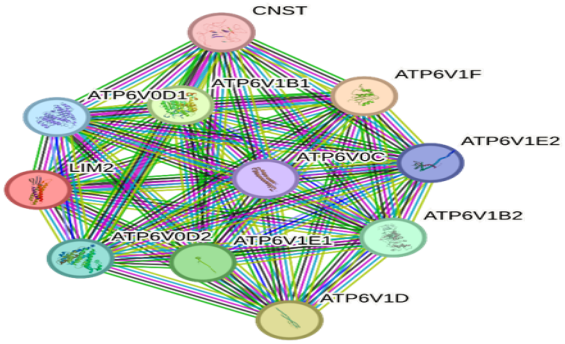
C.

Ramachandran plot statistics		
Residues in	**No. of residues**	**Percentage**
Most favored regions [A,B,L]	137	89.5%
Additional allowed regions [a,b,l,p]	14	9.2%
Generously allowed regions [~a,~b,~l,~p]	2	1.3%
Disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	**153**	**100.0%**
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	15	
Number of proline residues	3	
Total number of residues	173	

Figure 4. ***Model Evaluation of MP19 Protein*****. A.** Ramachandran plot showing the stereochemical spatial arrangement of amino acid residues. **B.** Z-score represents the overall quality of structure **C.** Statistical presentation of plot.

MP19 protein was queued at STRING online database to predict the interaction of this protein with other related proteins during the course of action. STRING database was selected for evidence option. Major proteins predicted in the interaction network were ATPases while MP19 was also found in interaction with CNST consortin, connexin sorting protein.

A.



B.

STRING Network Statistics	
Number of nodes	11
Number of edges	54
Average node degree	9.82
Average local clustering coefficient	0.982
Expected number of edges	11
PPI enrichment p-value	<1.0e-16

Figure 5. ***Protein-protein Interaction Network*****. A.** Proteins are represented by nodes and interaction is represented by edges. **B.** Statistical data of MP19 in interaction with other proteins.

4. Discussion

Lately, advances in Genetics and development of technology have yielded into a period of rapid gene identifications which are associated with genetic disorders (Bhinder et al., 2019). However, not all genetic disorders follow straight Mendelian pattern (Bhinder, 2017). In order to comprehend the causative pathway and finding suitable therapeutics options, there is further need to perform functional analysis of the genetic mutation.

There are score of reasons for the unmatched progress of functional analysis including the high cost and expertise required for functional analysis by means of characterizing the mutation in cell lines or animal models. The tools and techniques of bioinformatics have been utilized as alternate approach to fill the gap (Ashraf et al., 2018). Recently, advances in bioinformatics have greatly enhanced the significance of this subject and bioinformatics have become an adequate solution.

This study was aimed to perform functional analysis of mutation (c.233G>A) in LIM2 gene, using tools and techniques of Bioinformatics. This mutation was re-identified in our previous study using exome sequencing as an alternative method. LIM2 gene translates into a protein called MP19 which is a transmembrane protein with great functional importance for lens (Rl and Jh, 1993). The mutation causes MP19 protein to change its amino acid glycine into aspartic acid at position 78 (Irum et al., 2016). The wild type glycine residue is a known non polar amino acid whereas aspartic acid is a negatively charged residue. This may explain the change in the physical nature of mutant MP19 protein and thus the genesis of cataract.

The mutant MP19 (p.G78D) protein was found more polar than its normal counterpart (Figure 4A). Polarity is an essential feature of a protein molecule to not only determine its particular three dimensional shape but also plays role in proper functioning of protein (Aadland et al., 2019). A change in the polarity of a protein may change its structure and hinder its activity. Additionally, the change of amino acid glycine to aspartic acid has brought down the

hydrophobicity of the mutant protein (Figure 1B). In turn, this may disrupt the hydrophobic interactions of protein thus causing greater consequences for protein structure. Refractivity is the property of a protein molecule to bend light (Bassnett and Costello, 2017) and this property becomes very important when a protein is discussed in the backdrop of its functioning in the eye lens. The change of a hydrophobic glycine residue to a negatively charged aspartic acid at position 78 of MP19 protein greatly disturbs the refractivity. An increase of two folds is observed at position 78 of the mutant MP19 copy (Figure 1D).

Features of protein structure such as ratio of hydrophobic to hydrophilic amino acids, folds, motifs, ligand binding sites and overall three dimensional structure are of paramount importance for its normal working(Rollins et al., 2019). A slight change may disrupt the function of a protein, may end up disturbing a molecular cascade, and thus causing a disorder. Proteins are activated by ligands which get attach to receptor proteins at specific sites. The residues at ligand binding sites are crucial to the normal reception of ligands. To characterise the change in ligands specific to MP19 protein, there was need of both normal and mutant 3D structure of MP19 protein. However, there was no crystal or NMR MP19 protein structure available in PDB database. Therefore, both normal and mutant 3D structures of MP19 protein were designed using homology modelling (Figure 3A.1 & 3A.2). The mutation c.233G>A was characterised in predicted 3D model of mutated MP19 protein (Figure 3C). A change of hydrophobic amino acid into hydrophilic aspartic acid is self-explanatory about the change in phenotype.

MP19 protein is encoded by LIM2 gene and it is a transmembrane protein which functions in cell- cell connections (Rl and Jh, 1993). The protein interaction network of MP19 protein is populated by ATPases and CNST consortin proteins. The well characterised ATPases include the various motor proteins responsible for cargo transfers, cell motilities, and muscle contractions; the proteasome; the ATP synthase, F-ATPase; and the chaperone systems (Jessop et al., 2021). Other ATPases include DNA helicases and DNA replication complex and certain gene regulators (Duderstadt and Berger, 2008). Consortin (CNST) is an integral membrane protein that acts as a binding partner of connexins, the building blocks of gap junctions, and acts as a trans-Golgi network (TGN) receptor involved in connexin targeting to the plasma membrane and recycling from the cell surface (del Castillo et al., 2010).

5. Conclusion

Cataract is among the leading causes of blindness and molecular comprehension of genetics responsible for development of cataract has yet to be unfolded. Tools and techniques of bioinformatics may help understand these complex molecular mechanisms and provide affordability in the absence of wet lab approach.

References

Aadland, K., Pugh, C., Kolaczkowski, B., 2019. High-Throughput Reconstruction of Ancestral Protein Sequence, Structure, and Molecular Function. Methods Mol Biol 1851, 135–170. https://doi.org/10.1007/978-1-4939-8736-8_8

Ashraf, A., Ullah, N., Ahmad, G., 2018. Rat as an animal model for studying glutaredoxin-5 gene. Pak J Pharm Sci 31, 1301–1310.

Bassnett, S., Costello, M.J., 2017. The cause and consequence of fiber cell compaction in the vertebrate lens. Exp Eye Res 156, 50–57. https://doi.org/10.1016/j.exer.2016.03.009

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Research 28, 235–242. https://doi.org/10.1093/nar/28.1.235

Bhinder, M.A., 2017. Consanguineous marriage: A peril for coming generation. Journal of Genetic Disorders 1, 8.

Bhinder, M.A., Sadia, H., Mahmood, N., Qasim, M., Hussain, Z., Rashid, M.M., Zahoor, M.Y., Bhatti, R., Shehzad, W., Waryah, A.M., Jahan, S., 2019. Consanguinity: A blessing or menace at population level? Ann Hum Genet 83, 214–219. https://doi.org/10.1111/ahg.12308

Chen, J., Wang, Q., Cabrera, P.E., Zhong, Z., Sun, W., Jiao, X., Chen, Y., Govindarajan, G., Naeem, M.A., Khan, S.N., Ali, M.H., Assir, M.Z., Rahman, F.U., Qazi, Z.A., Riazuddin, S., Akram, J., Riazuddin, S.A., Hejtmancik, J.F., 2017. Molecular Genetic Analysis of Pakistani Families With Autosomal Recessive Congenital Cataracts by Homozygosity Screening. Invest. Ophthalmol. Vis. Sci. 58, 2207. https://doi.org/10.1167/iovs.17-21469

del Castillo, F.J., Cohen-Salmon, M., Charollais, A., Caille, D., Lampe, P.D., Chavrier, P., Meda, P., Petit, C., 2010. Consortin, a trans-Golgi network cargo receptor for the plasma membrane targeting and recycling of connexins. Hum Mol Genet 19, 262–275. https://doi.org/10.1093/hmg/ddp490

Duderstadt, K.E., Berger, J.M., 2008. AAA+ ATPases in the initiation of DNA replication. Crit Rev Biochem Mol Biol 43, 163–187. https://doi.org/10.1080/10409230802058296

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31, 3784–3788.

Home | Cat-Map | Washington University in St. Louis [WWW Document], n.d. URL https://cat-map.wustl.edu/ (accessed 8.5.23).

Irum, B., Khan, S.Y., Ali, M., Daud, M., Kabir, F., Rauf, B., Fatima, F., Iqbal, H., Khan, A.O., Al Obaisi, S., Naeem, M.A., Nasir, I.A., Khan, S.N., Husnain, T., Riazuddin, S., Akram, J., Eghrari, A.O., Riazuddin, S.A., 2016. Deletion at the GCNT2 Locus Causes

Autosomal Recessive Congenital Cataracts. PLoS ONE 11, e0167562. https://doi.org/10.1371/journal.pone.0167562

Jessop, M., Felix, J., Gutsche, I., 2021. AAA+ ATPases: structural insertions under the magnifying glass. Curr Opin Struct Biol 66, 119–128. https://doi.org/10.1016/j.sbi.2020.10.027

Khan, A.O., Aldahmesh, M.A., Alkuraya, F.S., 2015. PHENOTYPES OF RECESSIVE PEDIATRIC CATARACT IN A COHORT OF CHILDREN WITH IDENTIFIED HOMOZYGOUS GENE MUTATIONS (AN AMERICAN OPHTHALMOLOGICAL SOCIETY THESIS). Trans Am Ophthalmol Soc.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., Thornton, J.M., 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8, 477–486. https://doi.org/10.1007/BF00228148

Martin, F.J., Amode, M.R., Aneja, A., Austine-Orimoloye, O., Azov, A.G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., Bhurji, S.K., Bignell, A., Boddu, S., Branco Lins, P.R., Brooks, L., Ramaraju, S.B., Charkhchi, M., Cockburn, A., Da Rin Fiorretto, L., Davidson, C., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi, T., Fatima, R., Giron, C.G., Genez, T., Ghattaoraya, G.S., Martinez, J.G., Guijarro, C., Hardy, M., Hollis, Z., Hourlier, T., Hunt, T., Kay, M., Kaykala, V., Le, T., Lemos, D., Marques-Coelho, D., Marugán, J.C., Merino, G.A., Mirabueno, L.P., Mushtaq, A., Hossain, S.N., Ogeh, D.N., Sakthivel, M.P., Parker, A., Perry, M., Piližota, I., Prosovetskaia, I., Pérez-Silva, J.G., Salam, A.I.A., Saraiva-Agostinho, N., Schuilenburg, H., Sheppard, D., Sinha, S., Sipos, B., Stark, W., Steed, E., Sukumaran, R., Sumathipala, D., Suner, M.-M., Surapaneni, L., Sutinen, K., Szpak, M., Tricomi, F.F., Urbina-Gómez, D., Veidenberg, A., Walsh, T.A., Walts, B., Wass, E., Willhoft, N., Allen, J., Alvarez-Jarreta, J., Chakiachvili, M., Flint, B., Giorgetti, S., Haggerty, L., Ilsley, G.R., Loveland, J.E., Moore, B., Mudge, J.M., Tate, J., Thybert, D., Trevanion, S.J., Winterbottom, A., Frankish, A., Hunt, S.E., Ruffier, M., Cunningham, F., Dyer, S., Finn, R.D., Howe, K.L., Harrison, P.W., Yates, A.D., Flicek, P., 2023. Ensembl 2023. Nucleic Acids Research 51, D933–D941. https://doi.org/10.1093/nar/gkac958

Patel, N., Anand, D., Monies, D., Maddirevula, S., Khan, A.O., Algoufi, T., Alowain, M., Faqeih, E., Alshammari, M., Qudair, A., Alsharif, H., Aljubran, F., Alsaif, H.S., Ibrahim, N., Abdulwahab, F.M., Hashem, M., Alsedairy, H., Aldahmesh, M.A., Lachke, S.A., Alkuraya, F.S., 2017. Novel phenotypes and loci identified through clinical genomics approaches to pediatric cataract. Hum Genet 136, 205–225. https://doi.org/10.1007/s00439-016-1747-6

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605–1612. https://doi.org/10.1002/jcc.20084

Rl, C., Jh, W., 1993. The human lens fiber-cell intrinsic membrane protein MP19 gene: isolation and sequence analysis. Current eye research 12. https://doi.org/10.3109/02713689309033503

Rollins, N.J., Brock, K.P., Poelwijk, F.J., Stiffler, M.A., Gauthier, N.P., Sander, C., Marks, D.S., 2019. Inferring protein 3D structure from deep mutation scans. Nat Genet 51, 1170–1176. https://doi.org/10.1038/s41588-019-0432-9

Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., Jensen, L.J., von Mering, C., 2021. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 49, D605–D612. https://doi.org/10.1093/nar/gkaa1074

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673–4680.

Ullah, N., Munir, F., Iqbal, S., Ilyas, M., Achakzai, J.K., Bari, A., Bhinder, M.A., 2024. Exome sequencing as a tool of molecular screening of cataract affected families. Journal of Population Therapeutics and Clinical Pharmacology 31, 849–859.

Wiederschain, G., 2006. The proteomics protocols handbook: Walker, J. (ed.) Humana Press, 2005, 988 p., $175. Biochemistry (Moscow) 71. https://doi.org/10.1134/S0006297906060150

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35, W407–W410. https://doi.org/10.1093/nar/gkm290

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 112, 531–552. https://doi.org/10.1385/1-59259-584-7:531

Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9, 40. https://doi.org/10.1186/1471-2105-9-40