Pakistan Academy of Sciences

Research Article

Bioinformatics Analysis of a 4bp Homozygous Deletion Mutation of EDAR Gene Identified as an Important Cause of Hypohidrotic Ectodermal Dysplasia in Pakistan

Abdul Hameed[1], Hafsah Muhammad[3*], Asif Mir[2], Muhammad Ajmal[1], and Nayyer Siddique[3]

[1]Institute of Biomedical and Genetic Engineering (IBGE), 24-Mauve Area,
G-9/1, Islamabad, Pakistan

[2]Department of Bioinformatics and Biotechnology, Faculty of Basic and Applied Sciences,
International Islamic University (IIU), H-10, Islamabad-44000, Pakistan

[3]Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan

Abstract: Hypohidrotic Ectodermal Dysplasia (HED) is a rare congenital disorder characterized by reduced hair, the absence of sweat glands, dental anomalies, and craniofacial malformations. This condition can be inherited through X-linked, autosomal recessive, or autosomal dominant modes of inheritance. Mutations in four specific genes (EDAR, EDA, WNT10A, and EDARADD) are known to cause HED. In this study for the identification of pathogenic mutations in a consanguineous Pakistani family with the autosomal recessive form of HED, microsatellite markers were utilized to genotype the known loci associated with the disease. The condition in this family was mapped to the EDAR gene locus located on chromosome 2q11-q13. Upon screening the EDAR gene, a novel homozygous 4bp deletion (718delAAGA) in exon 8 was identified, which segregated with the disease phenotype. This 4bp deletion in the EDAR gene results in a frameshift and early termination of translation, producing a truncated protein of 245 amino acids instead of the normal length of 539 amino acids. Various bioinformatics tools were employed to analyze the pathogenic mutation linked to a significant number of HED cases in Pakistan. I-TASSER was used to model the protein structure, CASTp facilitated the identification of various pockets, and STiTCH 3.1 determined EDA to be the ligand for EDAR. Docking analysis were conducted for both the mutant and wild-type EDAR proteins with EDA, revealing notable differences in the interaction sites between the docked complexes of the normal and mutant forms of EDAR with the ligand EDA. These analyses provided insights into the protein's structural features, active sites, interactions, and the overall impact of the mutation.

Keywords: HED, EDAR, Congenital Disorder, Ectodermal Dysplasia, Homozygous Deletion.

1. INTRODUCTION

The incidence of Ectodermal Dysplasia (ED), a rare congenital disorder affecting at least two or more ectodermal tissues (skin, teeth, hair, sweat glands, and nails), is approximately seven instances per 10,000 people [1]. The most prevalent kind of ED in humans, known as Hypohidrotic Ectodermal Dysplasia (HED), is characterized by faulty tooth, sweat gland, and hair development [2]. Most sufferers of HED have fewer or ineffective sweat glands, which reduces their capacity to sweat. The body regulates its temperature in part through sweat, which cools the body as it evaporates off the skin. In hot conditions, an inability to sweat can result in hyperthermia, which can have life-threatening health effects. HED is a Mendelian disorder and can be inherited as autosomal recessive, autosomal dominant and X-linked recessive trait. Mutations of four genes, namely Ectodysplasin A (EDA); Ectodysplasin A receptor (EDAR); ectodysplasin A receptor-associated death domain (EDARADD)

---

and WNT10A are responsible for causing HED. The four genes mentioned account for approximately 90% of the HED cases [3, 4]. A recurrent missense mutation (p.W434R) in exon 12 of EDAR gene has been identified in two consanguineous Kashmiri families [5]. The presence of a mutation at position 398 in the EDAR gene leads to a significant decrease in its binding capacity to EDARADD, resulting in the development of HED [6].

Here, in this study, we ascertained a consanguineous family demonstrating autosomal recessive inheritance of HED. By genetic linkage analysis with microsatellite markers for the known HED loci, it was mapped to EDAR locus on chromosome 2q11-q13. The NF-kB activating member of the tumor necrosis factor receptor family is encoded by the EDAR gene. The transmembrane receptor for the soluble ligand ectodysplasin A is the encoded protein. It is a component of a signaling pathway that is crucial for interactions between the ectoderm and mesoderm, two embryonic cell layers, and is necessary for skin formation before birth. The connections between the embryonic cell layers are crucial for the development of various ectoderm-derived features, including as the skin, hair, nails, teeth, and sweat glands [7-9]. The mutation screening of the EDAR gene revealed a homozygous 4bp deletion (718delAAGA) associated with the disease in our HED family. This deletion mutation induces a frameshift that leads to premature termination of the protein product. Consequently, the truncated protein is composed of 245 amino acids, compared to the normal 539 amino acids. This specific deletion has also been documented in other Pakistani families, resulting in a similar HED phenotype, and appears to be a significant contributor to HED in Pakistan [10]. Given its importance as one of the common causes of recessively inherited HED, this deletion mutation has been characterized using various bioinformatics tools to establish the relationship between protein structure and function, as well as to conduct active site analysis, interaction studies, and protein-protein docking analysis.

2. MATERIALS AND METHODS

2.1. Sample Collection

A consanguineous Pakistani family (Figure 1A) from the rural Sindh province (specifically Sukkur city) was selected for this study due to their diagnosis of autosomal recessive HED. This family had not previously been included in any studies. Ethical approval was obtained from the Institutional Review Board, and informed consent was secured from the patients, family members, and control individuals before sample collection. An experienced dermatologist conducted clinical examinations of the affected individuals at a local hospital. All affected subjects exhibited the hallmark clinical signs and symptoms of HED, including thin, dry, light-colored, and brittle skin,
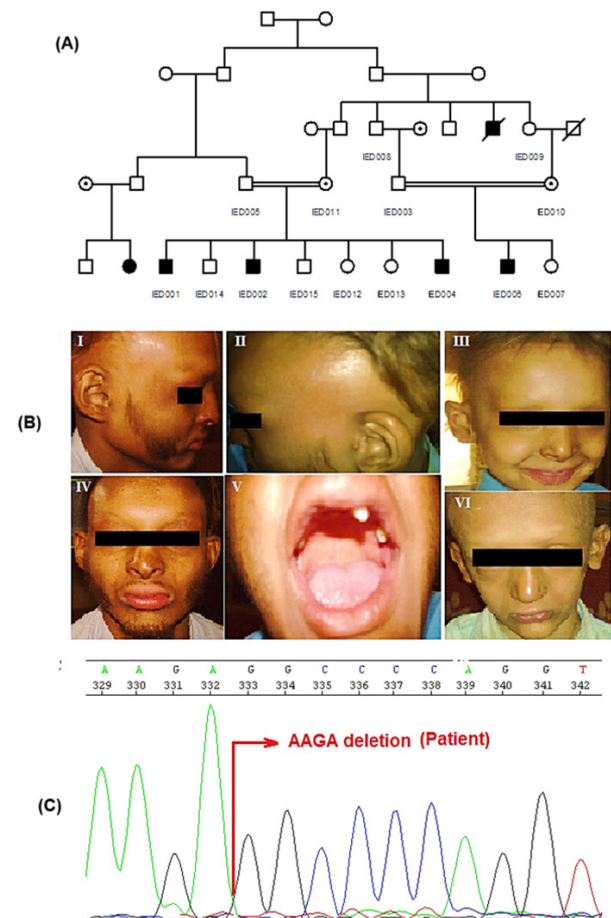

Fig. 1. (A) The pedigree of a Pakistani family affected by the autosomal recessive form of HED is depicted. Affected males and females are represented by filled squares and circles, respectively, while unaffected individuals are shown with open symbols. The study included individuals with Lab I.D. numbers, and double lines between symbols indicate consanguineous marriages. (B) Clinical observations of HED are displayed. Patients (I-IV & VI) exhibit characteristics such as fine scalp hair, absent eyebrows and eyelashes, a flattened nose, prominent lips, hyperpigmentation, and periorbital wrinkling. One of the affected family members (V) shows permanent conical teeth. (C) An electropherogram of a patient illustrates a homozygous 4bp deletion (del718AAGA) in exon 8 of EDAR gene.

a limited number of deformed teeth, decreased sweating capacity, periorbital wrinkles, and hyperpigmentation. Furthermore, the patients presented facial characteristics such as a prominent forehead, thick lips, and a flattened nasal bridge (Figure 1B). Family members seldom enter into consanguineous relationships, as they rarely marry outside the family.

2.2. Extraction of Genomic DNA

Approximately 5cc venous blood samples were obtained from the affected and normal family members. Genomic DNA was isolated from peripheral blood following a standard organic method of DNA extraction from whole blood [11]. DNA was quantified by spectrophotometry, by measuring absorbance of optical density at 260 nm and dilution of 40ng/μL for each sample was prepared for PCR amplification.

2.3. Genotyping

The HED disease phenotype in family was tested for linkage by using microsatellite markers tightly linked the known loci (Table 1). The genomic DNA from each available family member was amplified through PCR in a 25 μL reaction volume. The PCR reaction mixture consisted of 40 ng genomic DNA, 200 mM of each dNTP, 20 pmol of each primer, 1X PCR buffer and 1 U of Taq DNA polymerase. The PCR amplification was conducted for 35 cycles using the following conditions: initial denaturation at 95°C for 1 minute, annealing at 55°C for 1 minute, and extension at 72°C for 1 minute. A final extension step at 72°C for 7 minutes was performed using a thermal cycler 9600. The amplified PCR

products were separated and visualized on an 8% non-denaturing polyacrylamide gel under UV-illumination, and the genotypes were assigned and recorded.

2.4. Mutation Screening of EDAR Gene

Primers were designed from intronic sequences of the EDAR gene to facilitate PCR amplification of all protein-coding exons (2–12) and exon/intron splice junctions. Genomic DNA was used as the template in a 30 litres reaction volume to screen for mutations. The ACCuPrep® PCR purification kit (Bioneer corporation, Seoul, Korea) was used to purify the PCR products, and the BigDye ver 3.1 Cycle Sequencing Kit was used to sequence them in the Applied Biosystems 3130 Genetic Analyzer (Life technologies, USA).

2.5. Bioinformatics Analysis

2.5.1. Sequence retrieval

The EDAR gene consists of 9 exons, and it encodes a protein with 539 amino acids. The gene's sequence was retrieved from the OMIM database using the MIM number 604095.

2.5.2. Structure modeling

An online technique, PSIPRED, was used to estimate the secondary structure of the EDAR protein McGuffin *et al*. [12]. The threading technique through I-Tasser [13, 14] was used to construct three-dimensional (3D) structures. Using RAMPAGE server [15], protein 3D models (wild and mutant) were assessed.

Table 1. Known Loci for ectodermal dysplasia.

Locus I.D. No.	Gene	Chromosomal location	Reported markers in the region	Markers available and analyzed
ED1	EDA	Xq12-q13	DXS7159	DXS7132, DXS6800
ED2	EDAR	2q11-q13	D2S1890 D2S2954 D2S1888	D2S436, D2S410
ED3	EDARDD	1q42.2-q43	D1S2680 D1S2850 D1S2678	D1S235, D1S547, D1S1609
ED4	WNT10A	18q22.1-18q22.3	D18S857-D18S815	D18S858 ATA7D07, ATA82B02

2.5.3. Pocket identification

Protein pockets were detected using Computed Atlas of Surface Topography of proteins (CASTp) [16].

2.5.4. Protein-protein docking

Using the STITCH3.1 database [17], the framework of protein interactions was examined. Using the PatchDock server [18, 19], protein-protein docking analysis was performed. After being collected, the first 10 docked complexes were sent to the FireDock server for improvement [20, 21]. Version 5.0 of the ViewerLite programme was used to display docked protein and ligand complexes. After the docking process, the interactions and 2D representations of protein-ligand complexes were analyzed using LIGPLOT [22].

3. RESULTS AND DISCUSSION

3.1. Genotyping and Mutation Screening

Genotyping using microsatellite markers for the known loci showed linkage of the family to EDAR gene locus on chromosome 2q11-q13 (Figure 1A). On subsequent sequencing and mutation screening in 12 exons and splice site junctions of EDAR gene revealed a homozygous 4bp deletion (718delAAGA) mutation in exon 8 associated with the disease in our family (Figure 1B). The EDAR mutation, specifically 718delAAGA, was found to be inherited recessively in all affected family members. However, in the obligate carriers within the family, the mutation was present in a heterozygous state. The disease-association of the mutation was further confirmed by analyzing the ethnically matched 100 control samples for 4bp deletion of EDAR. None of the control sample had this change.

In individuals with HED, the EDAR gene has been found to include over 20 different mutations. While deletion mutations of the EDAR gene have been described, the majority of these variants impact only a single amino acid residue within the receptor protein. A small number of EDAR mutations lead to the creation of an atypical ectodysplasin A receptor. These genetic alterations can disrupt the signaling pathways essential for the development of ectodermal structures such

as hair follicles and sweat glands. This class of mutation results in the autosomal dominant form of HED, where each cell possesses one copy of the mutated gene. Conversely, additional mutations in the EDAR gene can completely inhibit the production of ectodysplasin A receptor protein. This inactivity of the receptor prevents the initiation of vital chemical signals necessary for ectoderm-mesoderm interactions and the proper development of ectodermal structures. When such mutations affect both copies of the EDAR gene in a cell, it leads to an autosomal recessive form of HED. A related study also identified the same mutation in the EDAR gene [23], but our research included an extensive analysis using various bioinformatics tools to characterize the differing outcomes of such mutations on the disease manifestation associated with a deletion mutation identified in Pakistani families.

3.2. Structure Prediction and Evaluation

Based on the PSIPRED results for protein secondary structure prediction (Figure 2), it is anticipated that the standard EDAR structure comprises 10 helices, 4 strands, and 14 coils. The number of helixes and coils in a mutant EDAR are altered and change to 2 and 7, respectively. There are still the same amount of strands. The size of the protein is impacted by frame shift mutation, which also affects the amount of structural characteristics.

I-Tasser was utilised to estimate the 3D structure of the normal and mutant EDAR using a threading technique. Helixes, beta sheets, and coils can be seen in the formations in Figure 3. Remains in favoured, allowed, and outlier regions have been identified by evaluation of these projected
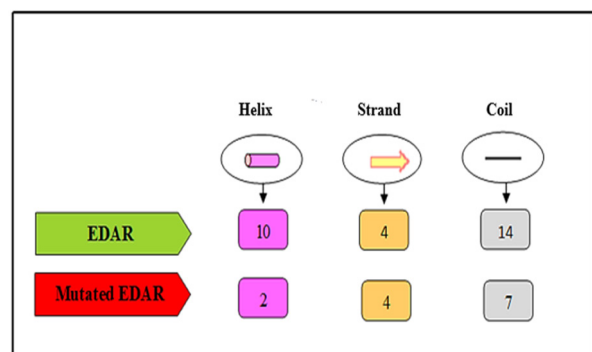

Fig. 2. A comparison of the PSIPRED results for the normal and mutated EDAR reveals the predicted number of structural features for both structures.

structures. For normal and mutant EDAR, the favoured, allowed, and outlier residue percentages are 334 (74.9%), 82 (18.4%), 30 (6.7%), and 161 (66.3%), 61 (25.1%), and 21 (8.6%), respectively. For both predicted structures, it is evident that a significant proportion of residues falls within the favored region, indicating the reliability of the predicted structures. Second, the difference in the amount of residues in favoured, allowed, and outlier regions demonstrates the difference between the two structures (wild and mutant). Protein function as well as size, structure, and conformation are all impacted by mutations, which results in a sick condition.

3.3. Analysis of Active Sites

The active site in proteins is typically a hydrophobic region with side chain atoms. For the purpose of creating drugs and predicting 3D structures, it is useful to identify compounds that bind to target proteins. Additionally, since these pockets contain particular amino acids, their prediction is helpful for mutational research. Using CASTp, a total of 59 pockets for normal EDAR and 20 for mutant EDAR have been located. Mutation effected structure of
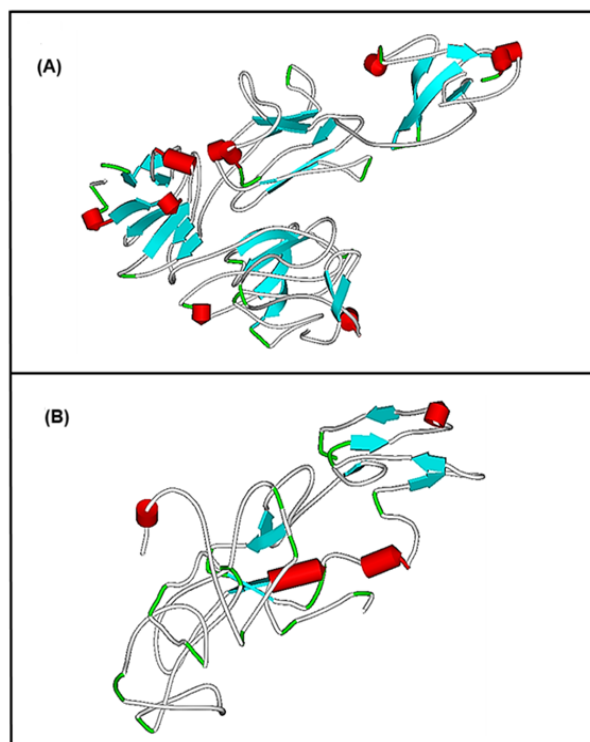

Fig. 3. The 3D structure prediction of both (A) Human EDAR and the (B) Mutated EDAR was performed using I-Tasser. The structures were displayed in a Schematic display style through ViewerLite v5.0.

protein and its active sites. The mutation causes the changes in the size, structure, conformation, active sites of the protein. All these changes ultimately affect the normal gene function.

3.4. Protein-Protein Docking

The EDAR functional partners were predicted using the Stitch3.1 server. Figure 4 displays predicted functional partners for EDAR gene. The protein that exhibited the highest interaction score (0.998) with EDAR was selected for further analysis and considered as the protein ligand, identified as EDA.EDA plays a role in epithelial-mesenchymal signalling during the development of ectodermal organs. To explore the docking procedure as well as the impact of mutation, this protein ligand was docked with both the normal and mutant structures of EDAR. The docking complex of the EDAR receptor and EDA ligand is shown in Figure 5(A), while the docking complex of the mutant EDAR receptor and EDA ligand is shown in Figure 5(B).

The EDAR wild type structure is altered by mutation, becoming a mutant structure that is smaller than usual. Change in the structure also affects the conformation, which affects the protein's interaction sites (i.e. active sites). The receptor protein's active site is where ligands engage and bind. The ligand interaction also alters when a wild type structure is altered because the active site changes. Table 2 presents the docking results, highlighting the receptor/ligand residues engaged in the interaction of both normal and mutant forms of EDAR with the ligand EDA. The docking contact involves certain residues on both the ligand and the receptor.
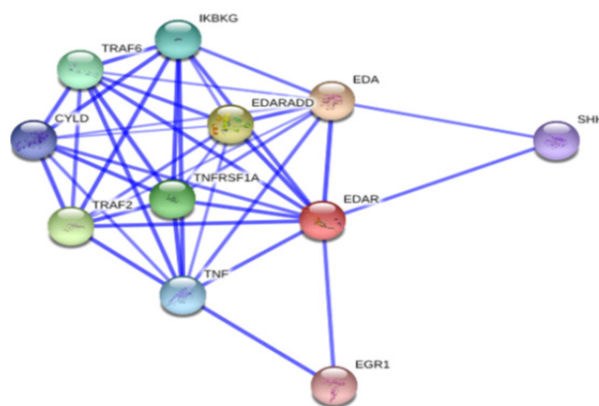

Fig. 4. Using the STITCH3.1 server, the predicted functional partners of EDAR were identified. Among these partners, EDA showed the highest interaction score with EDAR, which was 0.998.
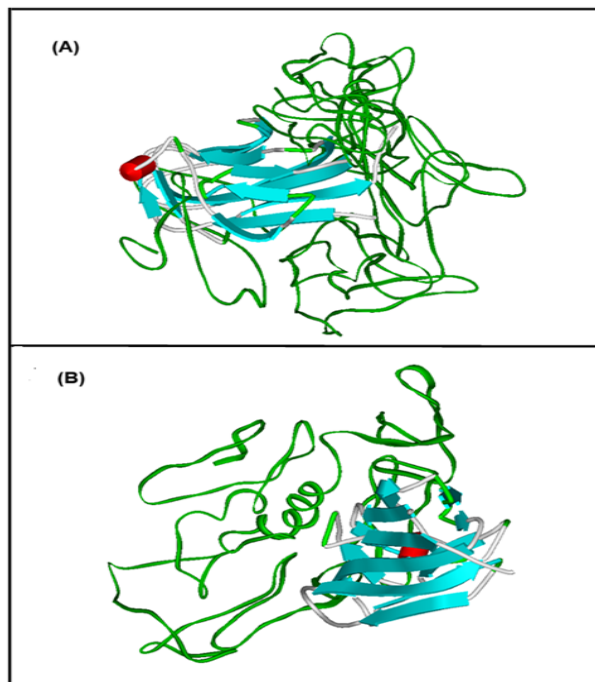
Fig. 5. Visualization of Docking Results using ViewerLite: (A) EDAR with EDR, (B) Mutated EDAR with EDR. Receptor: Green Color, Solid Ribbon display style, Ligand: Color by secondary type and Schematic display style.

Fig. 6. Dimplot Results for Docking Interactions; a) Normal EDAR with EDA, b) Mutated EDAR with EDA. Ligand Residues Involved in Hydrophobic Interactions are shown in blue color and represented by pink spoked arcs ( ). Receptor Residues Involved in Hydrophobic Interactions are shown in black color and represented by brick red spoked arcs ( ). Green dotted lines (..............) show Hydrogen Bonding. Receptor residues involved in H-Bonding are shown in Olive Green Color. Ligand residues involved in H-Bonding are shown in pink color.

Non-covalent interactions, such as hydrogen bonds and hydrophobic interactions, play a crucial role in facilitating the attachment of ligand/receptor residues. The docking analysis revealed the specific residues of both the receptor and the ligand that participated in hydrogen bonding and hydrophobic interactions, as listed in Table 2. The results indicate significant differences in the in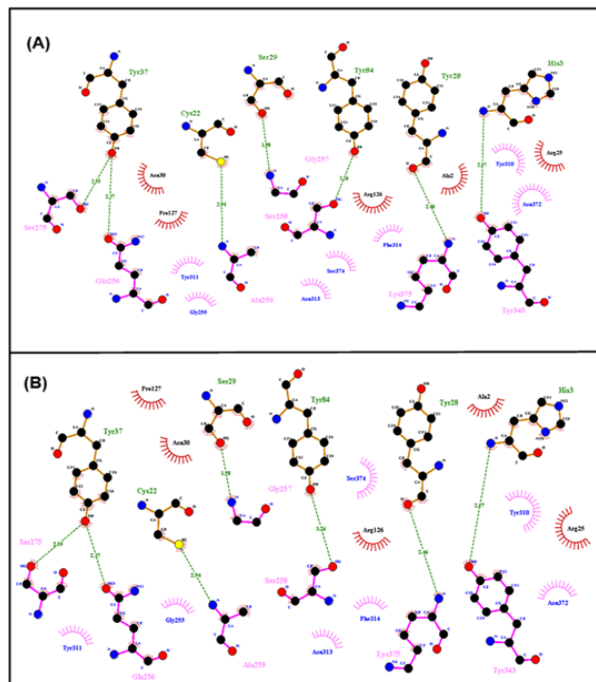teraction locations between the docked complexes of normal and mutant EDAR with the ligand EDA. The amino acids involved in these interactions at the binding site of the receptor protein provide valuable insights into the disparities in the interaction sites. Figure 6

Table 2: Receptor (EDAR) and ligand (EDA) residues involved in interactions.

Receptor-ligand	Hydrogen bond interactions		Hydrophobic interactions		Figure
	Ligand residues	Receptor residues	Ligand residues	Receptor residues	
EDAR-EDA	Gln256, ly257, Ser258, la259, Ser275, yr343, Lys375	His3,Cys22, Tyr28, Ser29, Tyr37, Tyr84	Gly255, Tyr310, Tyr311, Asn313, Phe314, Asn372, Ser374	Ala2, Arg25, Arg126, Pro127, Asn30	5(A)
Mutated EDAR-EDA	Gln256, Gly257, Ser258, Ala259, Ser275, Tyr343, Lys375	His3, Cys22, Tyr28, Ser29, Tyr37, Tyr84,	Gly255, Tyr310, Tyr311, Asn313, Phe314, Asn372, Ser374,	Ala2, Arg25, Asn30, Arg126, Pro127,	5(B)

presents the dimplot data, illustrating the docking interactions between EDAR (normal and mutant) and the ligand EDA. It also assesses the differences between the docking interactions of the normal/ wild type structure of EDAR and its mutant form with EDA ligand in our docking results. Protein size, structure, conformation, and interaction sites have all undergone significant change as a result of mutation. The protein's function is significantly altered as a result of these changes in structure, confirmation, and interaction site, which results in the disease state.

4. CONCLUSIONS

Structure prediction, interaction, and docking analyses of wild-type and mutant EDAR revealed significant structural and functional alterations resulting from a 4-bp homozygous deletion. PSIPRED results indicated a decrease in the number of helices (from 10 to 2) and coils (from 14 to 7) in the mutant EDAR. I-Tasser and RAMPAGE analyses confirmed variations in the three-dimensional structures, highlighting changes in residues in favored, allowed, and outlier regions. CASTp analysis demonstrated a significant reduction in active pockets (from 59 in normal EDAR to 20 in mutant EDAR), suggesting disrupted active sites. Docking studies involving the ligand EDA revealed altered interaction sites alongside hydrogen bonding and hydrophobic interaction changes. These structural and functional modifications compromise protein activity, contributing to disease conditions associated with the mutation.

5. ACKNOWLEDGEMENTS

6. CONFLICT OF INTEREST

Authors declare no conflict of interest.

7. ETHICAL STATEMENT

The study was approved by the ethical committee of Institute of Biomedical and Genetic Engineering (IBGE), Islamabad, Pakistan, and Quaid-i-Azam University, Islamabad, Pakistan.

8. REFERENCES

1. G.G. Meshram, N. Kaur, and K.S. Hura. A case report of hypohidrotic ectodermal dysplasia: A mini-review with latest updates. *Journal of Family Medicine and Primary Care* 7(1): 264-266 (2018).

2. B. Zeng, X. Xiao, S. Li, H. Lu, J. Lu, L. Zhu, D. Yu, and W. Zhao. Eight Mutations of Three Genes (EDA, EDAR, and WNT10A) Identified in Seven Hypohidrotic Ectodermal Dysplasia Patients. *Genes* 7(9): 65 (2016).

3. M.L. Mikkola. Molecular aspects of hypohidrotic ectodermal dysplasia. *American Journal of Medical Genetics Part A* 149(9): 2031-2036 (2009).

4. C. Cluzeau, S. Hadj-Rabia, M. Jambou, S. Mansour, P. Guigue, S. Masmoudi, E. Bal, N. Chassaing, M.C. Vincent, and G. Viot. Only four genes (EDA1, EDAR, EDARADD, and WNT10A) account for 90% of hypohidrotic/anhidrotic ectodermal dysplasia cases. *Human Mutation* 32(1): 70-72 (2011).

5. J.N. Foo, C.C. Khor, M. Jelani, and G. Ali. A recurrent missense mutation in the EDAR gene causes severe autosomal recessive hypohidrotic ectodermal dysplasia in two consanguineous Kashmiri families. *The Journal of Gene Medicine* 21(9): e3113 (2019).

6. T. Okita, N. Asano, S. Yasuno, and Y. Shimomura. Functional studies for a dominant mutation in the EDAR gene responsible for hypohidrotic ectodermal dysplasia. *The Journal of Dermatology* 46(8): 710-715 (2019).

7. J. Laurikkala, J. Pispa, H.-S. Jung, P. Nieminen, M. Mikkola, X. Wang, U. Saarialho-Kere, J. Galceran, R. Grosschedl, and I. Thesleff. Regulation of hair follicle development by the TNF signal ectodysplasin and its receptor Edar. *Development* 129(10): 2541-2553 (2002).

8. D.J. Headon, S.A. Emmal, B.M. Ferguson, A.S. Tucker, M.J. Justice, P.T. Sharpe, J. Zonana, and P.A. Overbeek. Gene defect in ectodermal dysplasia implicates a death domain adapter in development. *Nature* 414(6866): 913-916 (2001).

9. M. Moya-Quiles, M. Ballesta-Martínez, V. López-González, G. Glover, and E. Guillén-Navarro. A compound heterozygous mutation in the EDAR gene in a Spanish family with autosomal recessive hypohidrotic ectodermal dysplasia. *Archives of*

Dermatological Research 302: 307-310 (2010).

10. M. Naeem, D. Muhammad, and W. Ahmad. Novel mutations in the EDAR gene in two Pakistani consanguineous families with autosomal recessive hypohidrotic ectodermal dysplasia. *British Journal of Dermatology* 153(1): 46-50 (2005).

11. P. Guha, A. Das, S. Dutta, and T.K. Chaudhuri. A rapid and efficient DNA extraction protocol from fresh and frozen human blood samples. *Journal of Clinical Laboratory Analysis* 32(1): e22181 (2018).

12. L.J. McGuffin, K. Bryson, and D.T. Jones. The PSIPRED protein structure prediction server. *Bioinformatics* 16(4): 404-405 (2000).

13. A. Roy, A. Kucukural, and Y. Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* 5(4): 725-738 (2010).

14. Y. Zhang. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9(1): 40 (2008).

15. S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I. De Bakker, J.M. Word, M.G. Prisant, J.S. Richardson, and D.C. Richardson. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics* 50(3): 437-450 (2003).

16. T.A. Binkowski, S. Naghibzadeh, and J. Liang. CASTp: computed atlas of surface topography of proteins. *Nucleic Acids Research* 31(13): 3352-3355 (2003).

17. M. Kuhn, D. Szklarczyk, A. Franceschini, C. von Mering, L.J. Jensen, and P. Bork. STITCH 3: zooming in on protein–chemical interactions.

Nucleic Acids Research 40(D1): D876-D880 (2012).

18. D. Duhovny, R. Nussinov, and H.J. Wolfson. *E*fficient Unbound Docking of Rigid Molecules. In: Algorithms in Bioinformatics. R. Guigó and D. Gusfield (Eds.). *Lecture Notes in Computer Science, vol 2452. Springer, Berlin, Heidelberg* (2002).

19. D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, and H.J. Wolfson. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Research* 33(WS): W363-W367 (2005).

20. E. Mashiach, D. Schneidman-Duhovny, N. Andrusier, R. Nussinov, and H.J. Wolfson. FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Research* 36(WS): W229-W232 (2008).

21. N. Andrusier, R. Nussinov, and H.J. Wolfson. FireDock: Fast interaction refinement in molecular docking. *Proteins: Structure, Function, and Bioinformatics* 69(1): 139-159 (2007).

22. A.C. Wallace, R.A. Laskowski, and J.M. Thornton. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering, Design and Selection* 8(2): 127-134 (1995).

23. Z. Azeem, S.K.-U.-H. Naqvi, M. Ansar, A. Wali, A.K. Naveed, G. Ali, M.J. Hassan, M. Tariq, S. Basit, and W. Ahmad. Recurrent mutations in functionally-related EDA and EDAR genes underlie X-linked isolated hypodontia and autosomal recessive hypohidrotic ectodermal dysplasia. *Archives of Dermatological Research* 301(8): 625-629 (2009).