



OPEN

Unraveling the function and structure impact of deleterious missense SNPs in the human OX1R receptor by computational analysis

Mahvash Farajzadeh-Dehkordi^{1,2}, Ladan Mafakher³, Abbas Harifi⁴, Hashem Haghdoost-Yazdi⁵, Hossein Piri⁵ & Babak Rahmani^{1,2}✉

The orexin/hypocretin receptor type 1 (OX1R) plays a crucial role in regulating various physiological functions, especially feeding behavior, addiction, and reward. Genetic variations in the OX1R have been associated with several neurological disorders. In this study, we utilized a combination of sequence and structure-based computational tools to identify the most deleterious missense single nucleotide polymorphisms (SNPs) in the *OX1R* gene. Our findings revealed four highly conserved and structurally destabilizing missense SNPs, namely R144C, I148N, S172W, and A297D, located in the GTP-binding domain. Molecular dynamics simulations analysis demonstrated that all four most detrimental mutant proteins altered the overall structural flexibility and dynamics of OX1R protein, resulting in significant changes in the structural organization and motion of the protein. These findings provide valuable insights into the impact of missense SNPs on OX1R function loss and their potential contribution to the development of neurological disorders, thereby guiding future research in this field.

The G protein-coupled orexin/hypocretin receptor type 1 (OX1R/HCRTR1), an integral component of the orexinergic system, is widely acknowledged for its important role in promoting healthy life^{1,2}.

The human *OX1R* gene is located on chromosome 1p35.2 and encodes a 425 amino acid protein³. The protein comprises seven helical transmembrane segments (TM1–7) linked by three extracellular and intracellular loops (ECL1–3 and ICL1–3), an intracellular C-terminus, and an extracellular N-terminus⁴.

Numerous studies have demonstrated that OX1R exhibits a relatively selective binding affinity for the hypothalamic neuropeptide orexin-A and a lower affinity for the orexin-B neuropeptide³. This binding stimulates diverse downstream signaling pathways and contributes to multiple physiological processes, including feeding behavior, breathing, sleep–wake rhythm, drug addiction, reward-seeking, arousal and motivation, nociception, energy homeostasis, stress, and the fight/flight response^{3,5–14}. Furthermore, the orexin system has been suggested to possess innate proapoptotic activity, and systemic orexin receptors exhibit neuroprotective and anti-inflammatory effects^{15–17}. In addition, neuropeptides and orexin receptors are involved in coordinating sympathetic and cardiovascular activities^{18,19}. In contrast, dysregulated OX1R signaling influences pathological conditions such as anxiolytic behaviors, narcolepsy, cataplexy, ischemic stroke, depression, attention deficit and hyperactivity disorder, panic-related anxiety, Alzheimer's dementia, cancer, and Parkinson's disease^{20–30}. Moreover, genetic polymorphisms in the human *OX1R* gene are associated with panic disorder, polydipsia-hyponatremia, major depressive disorder (MDD), chronic migraine, aggressiveness, and sleep disorders^{31–35}.

Single nucleotide polymorphisms (SNPs) represent a prevalent and stable type of genetic variation within the human genome. They not only serve as valuable biological markers but can also be linked to the development of complex diseases, abnormalities, and variations in drug responses^{36–40}. Nonsynonymous SNPs (nsSNPs) in protein-coding regions, which result in alterations to amino acid sequences and the potential creation of mutated proteins with new structural and functional properties, have garnered significant interest. Deleterious nsSNPs at

¹Student Research Committee, Qazvin University of Medical Sciences, Qazvin, Iran. ²Department of Molecular Medicine, Qazvin University of Medical Sciences, Qazvin, Iran. ³Thalassemia & Hemoglobinopathy Research Center, Health Research Institute, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran. ⁴Department of Electrical and Computer Engineering, University of Hormozgan, Bandar Abbas, Hormozgan, Iran. ⁵Cellular and Molecular Research Center, Research Institute for Prevention of Non-Communicable Disease, Qazvin University of Medical Sciences, Qazvin, Iran. ✉email: b.rahmanigene@gmail.com

the genomic and/or proteome levels can induce detrimental functions by destabilizing protein tertiary structures, altering the physicochemical characteristics of proteins, and modifying protein–protein interactions, ultimately posing threats to cellular structural integrity^{41,42}.

Given the role of OX1R in neuropathological processes, it is crucial to analyze the implications of its nsSNPs. One significant aim in identifying deleterious nsSNPs is to conduct functional and structural evaluations. A comprehensive understanding of the conformational changes experienced by a protein can provide valuable insights into the underlying mechanisms of disease phenotypes and facilitate the identification of potential therapeutic agents capable of modulating protein function^{42,43}.

Experimental investigation of the impact of various nsSNPs is an expensive, time-consuming, and challenging process. However, employing *in silico* analysis through web-based computational methods is a feasible, effective, and affordable approach for investigating numerous SNPs in specific genes^{44,45}.

Therefore, the objective of this study is to retrieve OX1R nsSNPs and narrow them to ascertain the deleterious SNPs and assess their pathogenic effects on the protein using various biophysics-based computational algorithms, along with molecular simulation, to gain a better understanding of the mutation effects under physiological conditions.

Overall, the findings of this study will aid in determining the genotype–phenotype association of diseases, contribute to identifying the most significant SNPs in OX1R for population-based research, and be valuable in designing personalized medicine-based treatments for disorders triggered by these genomic SNPs.

Result

Distribution of OX1R SNPs

In the *OX1R* gene, 4295 SNPs were identified (accessed October 22, 2022). These included 2908 intron variants, 177 synonymous variants, 329 missense SNPs, 17 nonsense (stop-gained) SNPs, 219 variants were located in 3′-untranslated regions (UTR), and 237 variants were located in 5′-UTR, in addition to other upstream and downstream variants. Only missense SNPs were selected for analysis and screening. The methodological approach is shown in the schematic diagram (Fig. 1).

Determining 3D protein structure

As the complete structure of the human OX1R protein did not exist in the PDB bank (Fig. S1), homology modeling of the OX1R was performed by the Robetta webserver. Several approaches were used to evaluate the quality of the model (Table S1). Based on the Ramachandran plot analysis using the Procheck web server, more than 90% of the amino acid residues were located in the core and generally allowed regions, indicating that the protein structure is of high quality. The ERRAT value of more than 95% showed the high resolution of the protein model. The verify3D score of 42% demonstrated moderate structural quality for the protein. A Molprobability score of 100% confirms the high resolution of protein structure compared to proposed structural characterizations based on X-rays. The ProSA plot showed the model has a z-score value (−5.9) in the range of proteins with the same length whose structures were identified by X-ray and NMR structure (Fig. 2A). Comparing the quality of OX1R homology modeling with its crystal structure (PDB ID: 6 to 7) revealed that this model had approximately the same quality as its crystal structure. The tertiary structure of the human OX1R model is shown in Fig. 2B. The structure of OX1R homology modeling and 6 to 7 represented high compatibility (Fig. S2).

Determining the pathogenic SNPs

To estimate the deleterious SNPs of OX1R, various computational tools were employed. Initially, 329 missense SNPs were assessed using 10 bioinformatics tools: FATHMM-MLK, SIFT, Mutation Assessor, SNAP-2, PolyPhen-2, Panther-PSEP, PON-P2, CADD, Align-GVGD, and VEST-4, to investigate their functional impacts. Among them, 21 missense SNPs were consistently predicted to have functional effects by all the utilized methods (Table S2). Subsequently, the selected SNPs were passed through five different servers, namely Pmut, Suspect, PhD-SNP, SNPs&GO, and InMeRF to evaluate their potential association with pathogenicity. Among the 21 evaluated missense SNPs, 17 SNPs were identified as disease-associated or pathogenic by at least four of the five pathogenicity prediction tools (Table 1). These 17 missense SNPs were selected for further analysis and screening.

Determining the domain of OX1R protein

We used four tools, InterPro, PROSITE, Pfam, and CDD, to determine the conserved domains and locations of 17 pathogenic missense SNPs in the OX1R protein. All three databases illustrated the existence of an important domain: a seven-helix transmembrane G-protein-coupled receptor (GPCR) domain at positions 63–358, whereas CDD predicted the same domain at positions 47–369. Furthermore, the positions of the 17 pathogenic missense SNPs were analyzed, and all variants were found to be located in the GTP-binding protein domain.

Determining SNPs' impact on the OX1R protein stability

The free energy change ($\Delta\Delta G$) of the native and mutated forms of a protein is a vital index for protein stability. By evaluating the effect of mutations on the free energy, the impact of single-site mutations on protein stability can be precisely identified. A $\Delta\Delta G$ score < 0 indicated a decrease in protein stability, whereas a score > 0 indicated the opposite result.

Of the 17 pathogenic missense SNPs, 12 and 15 SNPs were determined to decrease OX1R protein stability using I-Mutant 2.0, and INPS-MD tools, respectively. In contrast, the Mu-Pro server showed a decreasing effect on all the pathogenic SNPs (Table 2). The m-CSM, SDM, and DUET servers predicted 15, 13, and 14 SNPs, respectively, to be deleterious, as they caused a decrease in the stability of the OX1R protein (Table 2). According to the Dyna-Mut tool, seven SNPs of the 17 pathogenic SNPs had negative $\Delta\Delta G$ values, indicating that they

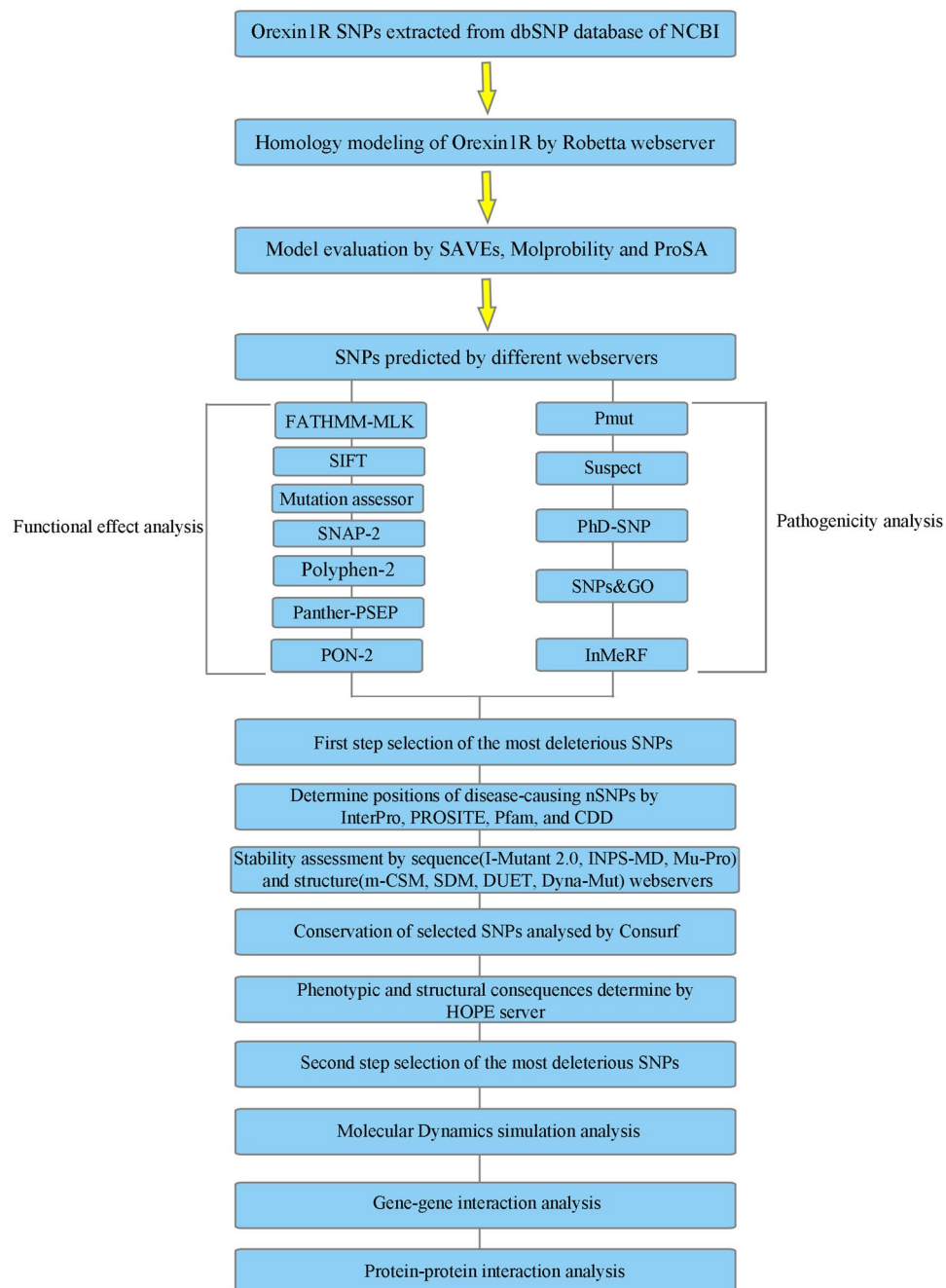


Figure 1. The methodological approach is shown in the schematic diagram.

reduced the stability of the OX1R protein (Table 2). Finally, five missense SNPs (R144C, I148N, S172W, A297D, and L312Q) among 17 pathogenic SNPs were identified as the most deleterious to OX1R protein stability ($\Delta\Delta G$ values less than zero) using all seven above-mentioned tools. We selected these five destabilizing SNPs for the phylogenetic conservation analysis.

Determining the phylogenetically conserved residues in the OX1R protein

The ConSurf tool was used to predict the level of evolutionary conservation at all the residue positions in OX1R. The OX1R protein structure was submitted as an input and the results are shown in Fig. S3. Based on the ConSurf outcomes for the OX1R protein, this protein is more evolutionarily conserved, with a greater number of conserved residues; 158 amino acid positions were estimated to have a conservation score between 7 and 9, and 93 of them had a conservation score of 9 (highly conserved residues). Furthermore, ConSurf analysis determined that out of the five destabilizing missense SNPs (R144C, I148N, S172W, A297D, and L312Q), three (R144C, I148N, and A297D) were highly conserved (score 9). Owing to their high conservation, residues I148 and A297

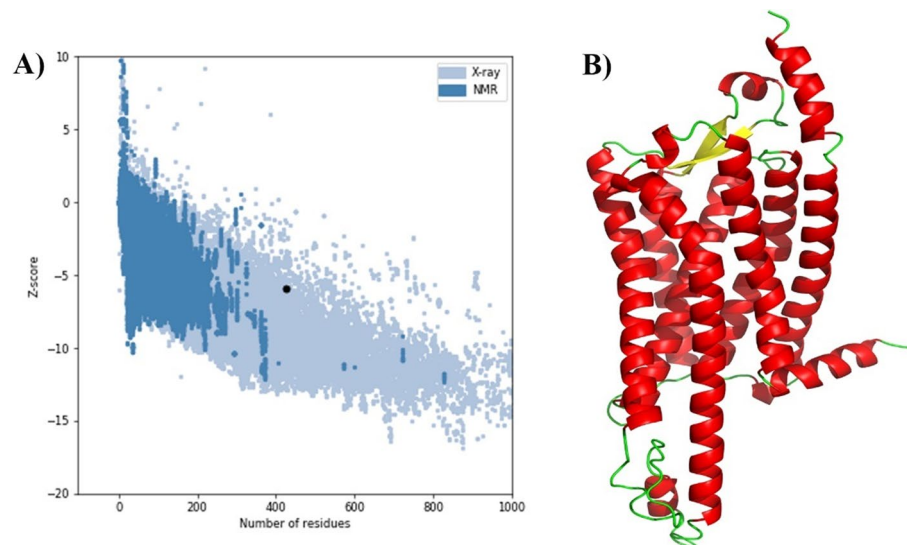


Figure 2. The ProSA plot and tertiary structure of human OX1R protein. (A) The ProSA plot of human OX1R protein (black dot) indicates comparable quality of X-ray and NMR. (B) The tertiary structure of human OX1R protein predicted by the PyMOL (Red: alpha helix; Yellow: beta-sheet; Green: random coils and other structures).

SNP ID	Protein change	Pmut	Suspect	PhD-SNP	SNPs&GO	InMeRF
rs768118872	I85T	Pathogenic	Pathogenic	Neutral	Disease	Normal
rs1368005577	N87Y	Pathogenic	Neutral	Disease	Disease	Pathogenic
rs1307483513	S89P	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
rs188180909	L90Q	Pathogenic	Pathogenic	Neutral	Disease	Normal
rs778134490	D92G	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
	D92V	Pathogenic	Neutral	Neutral	Disease	Pathogenic
rs1487077579	A141V	Pathogenic	Pathogenic	Neutral	Disease	Pathogenic
rs778794988	R144C	Pathogenic	Pathogenic	Neutral	Disease	Pathogenic
	R144L	Pathogenic	Neutral	Neutral	Disease	Pathogenic
rs201925923	A147D	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
rs1160159351	I148N	Pathogenic	Neutral	Disease	Disease	Pathogenic
rs200634924	P151L	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
rs1239300181	A161D	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
rs746566291	S172W	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
rs201360185	R293W	Pathogenic	Pathogenic	Disease	Disease	Pathogenic
rs767939002	A297D	Pathogenic	Pathogenic	Neutral	Disease	Pathogenic
rs1291095889	L312Q	Pathogenic	Pathogenic	Neutral	Disease	Pathogenic
rs746884559	P313H	Pathogenic	Neutral	Disease	Disease	Pathogenic
	P313L	Pathogenic	Neutral	Disease	Disease	Pathogenic
rs958763496	S351R	Pathogenic	Neutral	Disease	Disease	Pathogenic
rs1479886483	N354K	Pathogenic	Neutral	Disease	Disease	Pathogenic

Table 1. Screening of 21 deleterious missense SNPs associated with OX1R protein. Boldly highlighted mutations are found to be disease-associated or pathogenic by at least four out of five pathogenicity prediction tools.

were classified as structural and buried sites, while residue R144 was predicted to be a functional and exposed site of the OX1R protein. In addition, ConSurf estimated that S172W had a conservation score of 8 (conserved) and L312Q was found in an intermediately conserved site (score 6). Taken together, we selected four missense SNPs namely, R144C, I148N, S172W, and A297D that were located in conserved positions (score between 7 and 9), and the SNP in the L312 position was excluded from subsequent evaluations because it was revealed to have an intermediately conserved profile (score 6).

SNP ID	Protein change	I-Mutant 2.0	INPS-MD	Mu-Pro	m-CSM	SDM	DUET	Dyna-Mut
rs1368005577	N87Y	I	I	D	D	I	D	I
rs1307483513	S89P	D	D	D	I	D	D	I
rs778134490	D92G	D	D	D	D	D	I	I
rs1487077579	A141V	I	D	D	D	D	D	I
rs778794988	R144C	D	D	D	D	D	D	D
rs201925923	A147D	I	D	D	D	D	D	I
rs1160159351	I148N	D	D	D	D	D	D	D
rs200634924	P151L	D	D	D	D	I	D	I
rs1239300181	A161D	I	D	D	D	D	D	D
rs746566291	S172W	D	D	D	D	D	D	D
rs201360185	R293W	D	D	D	D	D	D	I
rs767939002	A297D	D	D	D	D	D	D	D
rs1291095889	L312Q	D	D	D	D	D	D	D
rs746884559	P313H	D	D	D	D	I	D	I
	P313L	D	D	D	D	I	I	I
rs958763496	S351R	I	D	D	D	D	D	I
rs1479886483	N354K	D	I	D	D	D	I	D

Table 2. Stability prediction analysis of 17 pathogenic missense SNPs of OX1R protein. Boldly highlighted mutations are found to be destabilizing from all seven tools. Abbreviations: I (Increase) and D (Decreased) of stability.

Determining the molecular and phenotypic effects of the nsSNPs

We used the HOPE server to calculate the physicochemical properties of the four most detrimental missense SNPs (R144C, I148N, S172W, and A297D) in OX1R.

In the case of the R144C mutation, the substitution of a large positively charged residue, arginine, with a smaller neutrally charged residue, cysteine, at position 144 occurs. The wild-type residue is found in a domain that is essential for protein activity, but the mutation presents an amino acid with different characteristics, which can interrupt this domain and restrict its function. Furthermore, the analysis reveals that the wild-type residue, arginine, engages in hydrogen bonding and salt bridge interactions with aspartic acid at position 143. Conversely, the substituted residue, cysteine, is unable to establish the same interactions as arginine, resulting in a perturbation of the local stability of the protein.

The I148N, S172W, and A297D substitutions presented a larger residue than the wild-type moiety, which likely did not fit into the protein core. If the replaced amino acid does not fit the protein, it causes structural changes that are sometimes hazardous. Particularly, with S172W, it was predicted that the wild-type residue is found in a region annotated in the UniProt database as a transmembrane domain and forms hydrogen bonds with serine at position 131 and isoleucine at position 168, which would affect contact with the lipid membrane and cause loss of the same hydrogen bonds. Moreover, the wild-type residues I148N, S172W, and A297D were more hydrophobic than the mutant residues, which could have caused repulsive hydrophobic interactions in the core of the protein. The A297D substitution introduces a negatively charged residue (aspartic acid) in a buried moiety (alanine), which can lead to protein-folding problems. The A297D in the 3D structure was located in an α -helix. The difference in properties between aspartic acid and alanine can easily cause alanine to not prefer α -helices as secondary structures. In addition, conservation analysis showed that mutations of 100% conserved residues at positions 144 and 148 were probably harmful to the protein. The structural visualization of the four most detrimental missense SNPs was conducted using the HOPE server, as illustrated in Fig. 3.

The secondary structure and percentages of the alpha helices, coils, and extended strands, in the OX1R native structure, were evaluated using the SOPMA tool as mentioned in Table S3. SOPMA revealed that 199 residues were associated with alpha helix (46.82%), 53 with extended strand (12.47%), 14 with beta-turn (3.29%), and 159 with the random coil (37.41%). Moreover, an analysis of the most detrimental SNPs' effects on the alterations in the secondary structure of the OX1R protein was conducted as mentioned in Table S3. It was observed that in all mutant proteins, a slight reduction in the number of residues participating in the alpha helix conformation was accompanied by a slight increase in the extended strand conformation compared to the native structure.

Molecular dynamics simulations

MDs measurements for the four most detrimental missense SNPs (R144C, I148N, S172W, and A297D) were conducted for 200 ns simulations to characterize the structural perturbation of OX1R protein caused by these missense SNPs in normal physiological conditions.

Stability analysis

The RMSD value was estimated to comprehend the overall stability of protein during the simulation⁴⁶. RMSD of the backbone's atoms of the native structure and the mutant proteins were plotted against the time to evaluate

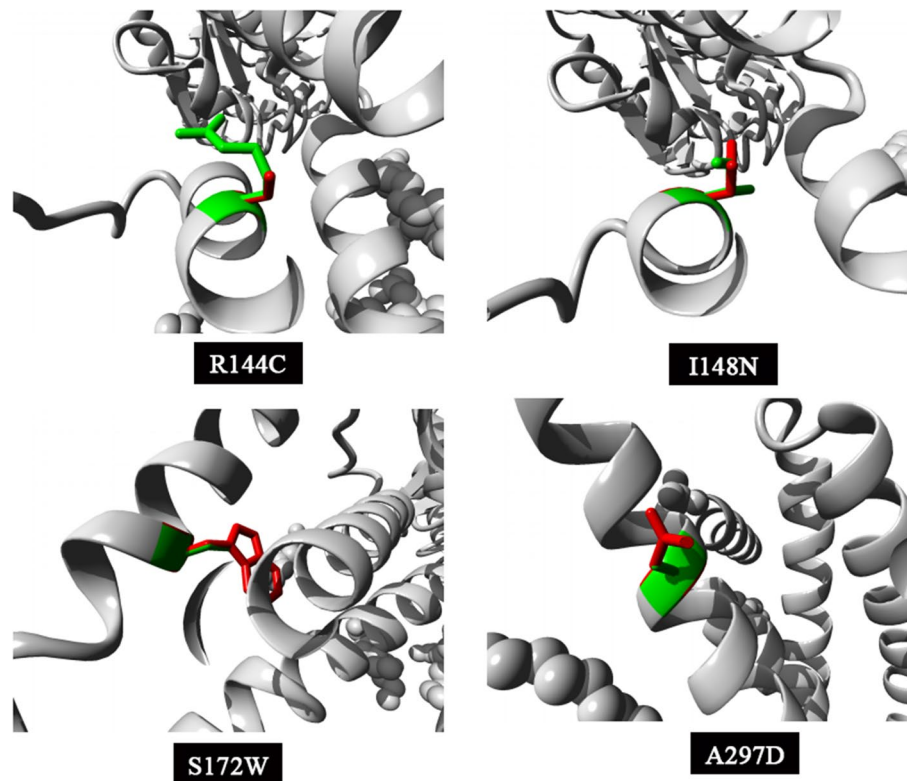


Figure 3. Visualization of wild-type (green) and mutated (red) amino acid residues for the four most detrimental missense SNPs (R144C, I148N, S172W, and A297D) using HOPE Server.

the alteration effects (Fig. 4A). Compared to the native structure, a significant increase in the average RMSD values for the R144C, I148N, S172W, and A297D mutant proteins was observed (Fig. 4A). The average RMSD for the native protein was found to be 0.542 nm, while for mutants R144C, I148N, S172W, and A297D, they were found to be 1.211 nm, 1.251 nm, 1.061 nm, and 1.473 nm, respectively. A higher RMSD value indicates a reduction in protein stability. Since stability is an important aspect of protein activity and function, decreased stability may have a damaging impact on the overall function of the protein⁴⁶. Data showed that all mutants showed higher RMSD value and fluctuation during molecular dynamics simulation which means the unstable structure of mutants compared to the native protein.

Compactness analysis

The Rg value indicates the overall dimension of the protein, thus acting as a crucial marker for assessing protein structure compactness⁴⁷. To analyze the impact of the four most detrimental missense SNPs on the structural integrity of OX1R, the Rg value was determined for native and mutant proteins.

As shown in Fig. 4B the mutants R144C, I148N, S172W, and A297D showed significant divergence in the structural compactness and rigidity of the native protein. The native protein had an average Rg value of 2.085 nm, whereas the mutants R144C, I148N, S172W, and A297D had average Rg values of 2.530 nm, 2.518 nm, 2.763 nm, and 2.857 nm, respectively. The results indicated that all mutants had fewer compact and more flexible structures, as their Rg values were higher compared to the native structure.

Solvent accessible surface area analysis

To determine the effect of the four most detrimental missense SNPs on the compactness of the hydrophobic core of the OX1R protein, SASA analysis was conducted. As displayed in Fig. 4C an alteration in SASA between the native and mutant proteins was observed. The average SASA values for the mutants R144C, I148N, S172W, and A297D were 247.796 nm², 209.098, 243.33 nm², and 202.499 nm², respectively, whereas, for the native structure, it was 192.189 nm². A greater SASA value signifies an expansion of a protein, suggesting that all mutant proteins may achieve more expanded structures compared to the native protein.

Flexibility analysis

To predict the impact of the four most detrimental missense SNPs on the dynamic behavior of residues within the OX1R protein structure, RMSF analysis was performed. The RMSF value is essential for understanding the overall flexibility of proteins during simulation. The backbone RMSF for each residue number was estimated for the native and mutant proteins (Fig. 5). The native protein had an average RMSF value of 0.117 nm. For mutants R144C, I148N, S172W, and A297D, the average RMSF values were 0.154, 0.139, 0.134, and 0.148 nm, respectively.

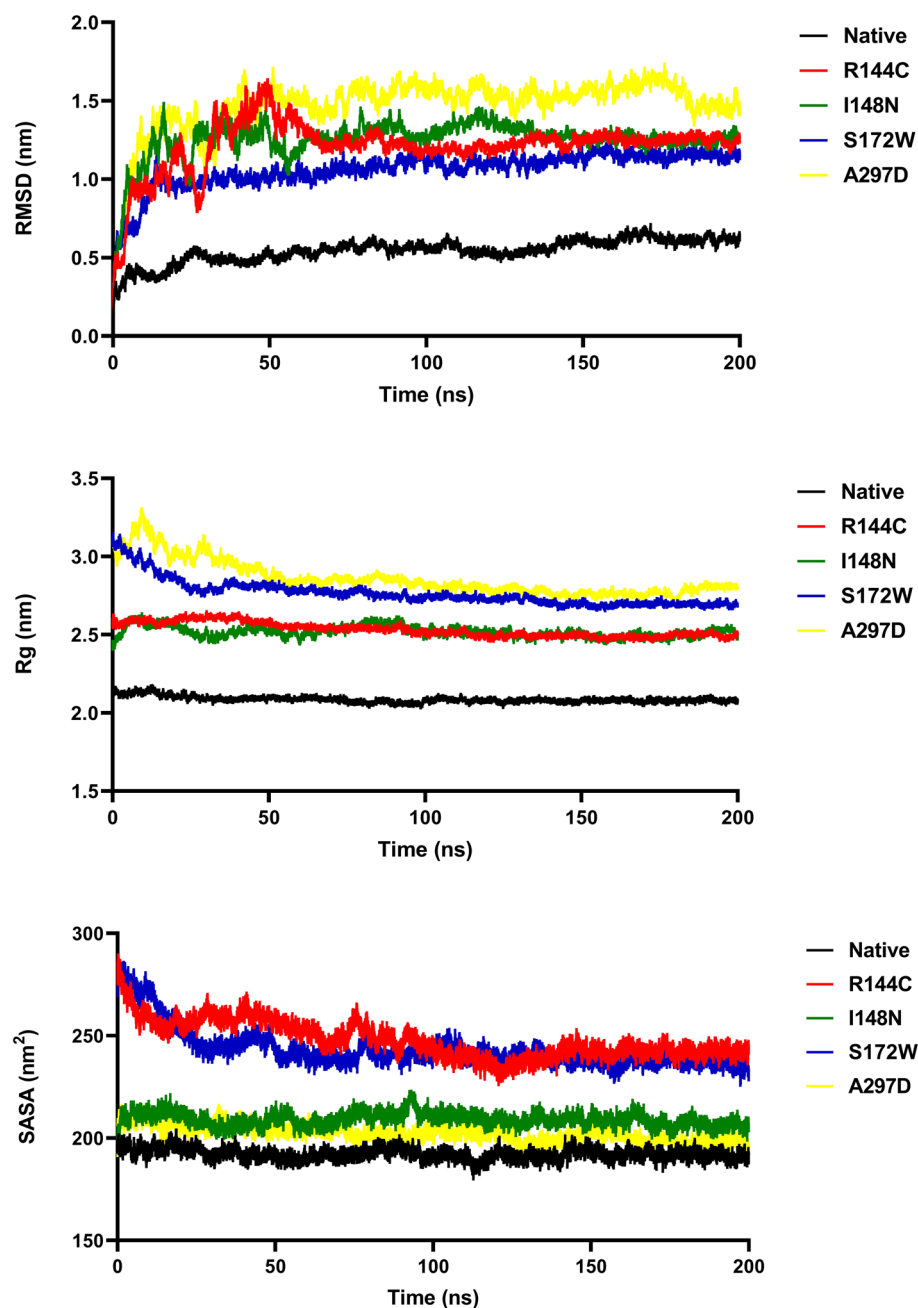


Figure 4. Analysis of RMSD, Rg, and SASA of native and four mutants of OX1R over 200 ns simulation. (A) RMSD values of Ca atoms of native and mutant structures. (B) Calculation of Rg, which is represented as a time-dependent change during the simulation. (C) Calculation of SASA represented as a time-depend. The color scheme is as follows: native (black color), R144C mutant (red color), I148N mutant (green color), S172W (blue color), and A297D mutant (yellow color).

Hence, slightly higher flexibility was observed for all mutant proteins compared to that of the native protein. Mutants R144C and I148N, compared to the native structure, showed higher fluctuations between residues 245 to 260 and 250 to 280, respectively (Fig. 5A). The S172W mutant showed a prominent peak of 0.58 nm at the very first residues (Fig. 5B). In mutant A297D, residues 250–280 and 400–425 were observed to have higher flexibility than the native structure (Fig. 5B). According to RMSF analysis, although there were no significant changes in the overall flexibility of all mutant proteins, residual changes were observed.

Secondary structure analysis

To identify the alteration in OX1R protein secondary structural arrangement caused by the four most detrimental missense SNPs, the secondary structure compositions of native and mutant proteins were analyzed, as shown in Fig. 6. The mutant R144C (Fig. 6B) showed a slight increase in the β -sheet content at residue 185 to 203, the other

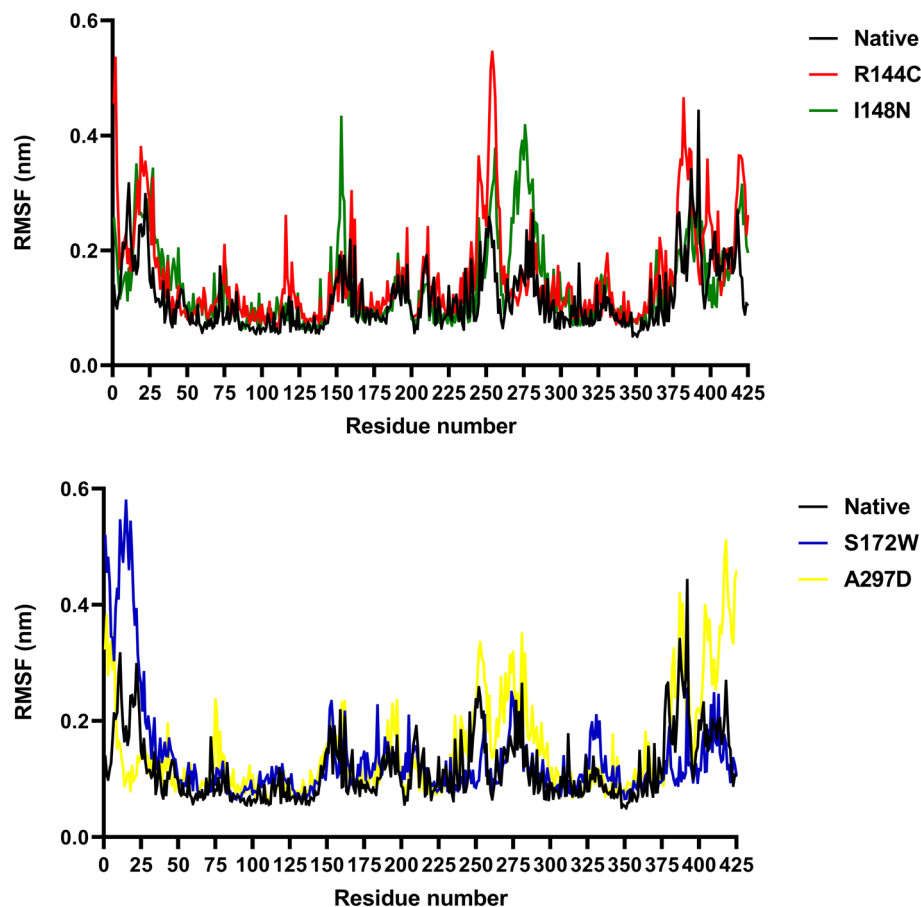


Figure 5. RMSF of Ca atoms of the native structure and mutant proteins over 200 ns simulation. (A) RMSF plot of native structure (black color) vs. R144C (red color) and I148N (green color) mutant proteins. (B) RMSF plot of native structure (black color) vs. S172W (blue color) and A297D (yellow color) mutant proteins.

elements were marginally similar to the native protein. The mutant I148N showed a tiny increase in the β -sheet trend at residues 185 to 203 and 185 to 261, while a transition from α -helix to bend and turn to residues 300 to 330 compared to the native protein (Fig. 6C). The S172W mutant showed α -helices turning into turns, bends, and 5-helices between residues 85–120 and an increase in the β -sheet trend at residues 185 to 203 compared to the native protein (Fig. 6D). In mutant A297D, many α -helices turned into bends, turns, 5-helices, and 5-helices between residues, 30–120 and 280–360. Furthermore, an increase in the β -sheet content at residue 185 to 203 compared to the native protein also was observed (Fig. 6E).

The evaluation of the percentage of each secondary structure in native and mutant proteins showed a moderate reduction in the structural profile of mutants R144C, I148N, S172W, and A297D compared to the native structure (Table S4). Significant alterations in the secondary structure composition of all mutants can be found as a result of inhibiting the α -helix conformation while increasing the β -sheet conformation compared to the native protein, which is in line with the SOPMA outcomes shown in Table S3, where the number of α -helices residues were altered in all mutants.

Principal component analysis

PCA was conducted to determine the effect of the four most detrimental missense SNPs on structural motion and locally restricted fluctuations of native and mutant proteins. Principal components (PCs) with larger eigenvalues are eigenvectors that play an important role in the overall concerted motion of the protein⁴⁸. The first two PCs (PC1 and PC2) were chosen to determine the collective motion in the phase space during the simulation. As depicted in Fig. 7 mutants R144C, I148N, S172W, and A297D occupied more space than the native protein. The PCA findings suggested that the concerted motions were perturbed after mutations, resulting in the loss of structural stability in mutant proteins, in agreement with RMSD, Rg, SASA, RMSF, and secondary structural analyses.

Free energy landscape analysis

Using the first two eigenvectors of the FELs, all possible conformation of atoms, present in native and mutant proteins, were explored. FEL color spectrum labels range from blue to red, where blue spots indicate folds and the highest stable state of the protein, while red spots represent unfolded states and a lower stable state. Figure 8 illustrates that all mutants R144C, I148N, S172W, and A297D exhibited fewer blue spots with structure expansion

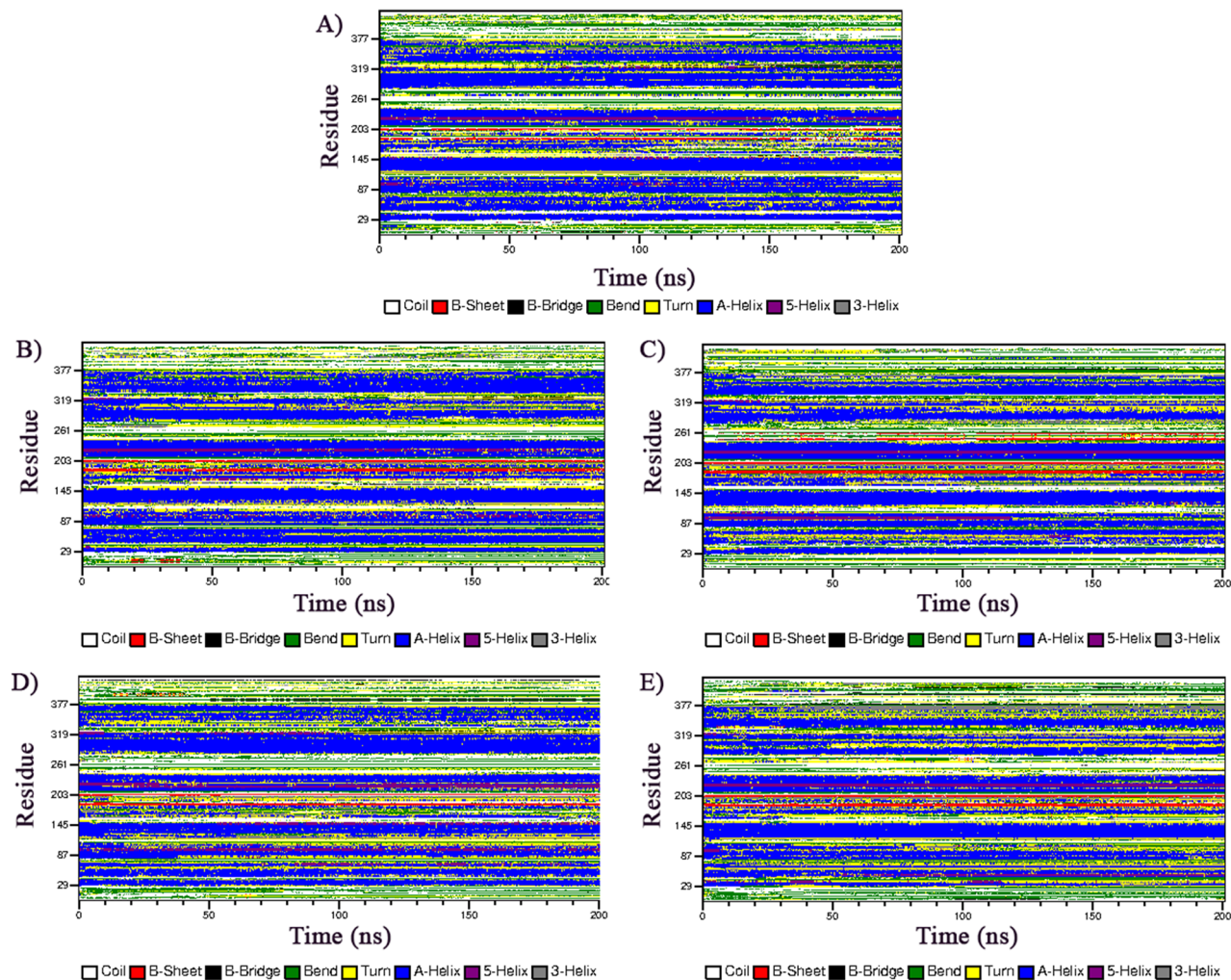


Figure 6. Effect of the four most missense SNPs in the stability of the secondary structure elements of the OX1R during the 200 ns simulation. (A) native structure, (B) mutant R144C, (C) mutant I148N, (D) mutant S172W, and (E) mutant A297D.

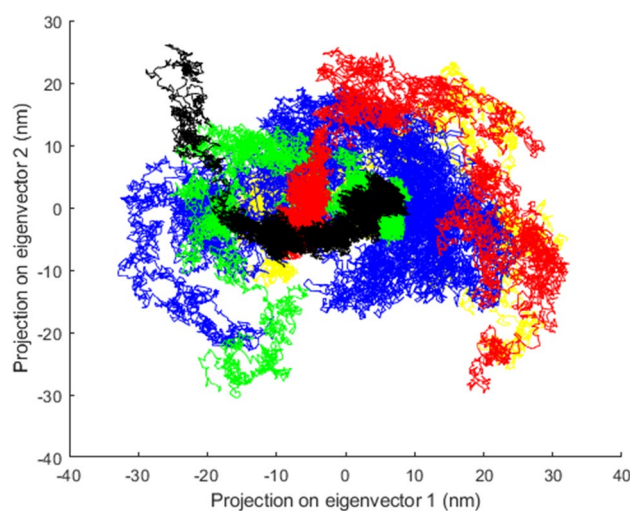


Figure 7. The principal component analysis (PCA) of the native structure and mutant proteins over 200 ns simulation. The color scheme is as follows: native (black color), R144C mutant (red color), I148N mutant (green color), S172W (blue color), and A297D mutant (yellow color).

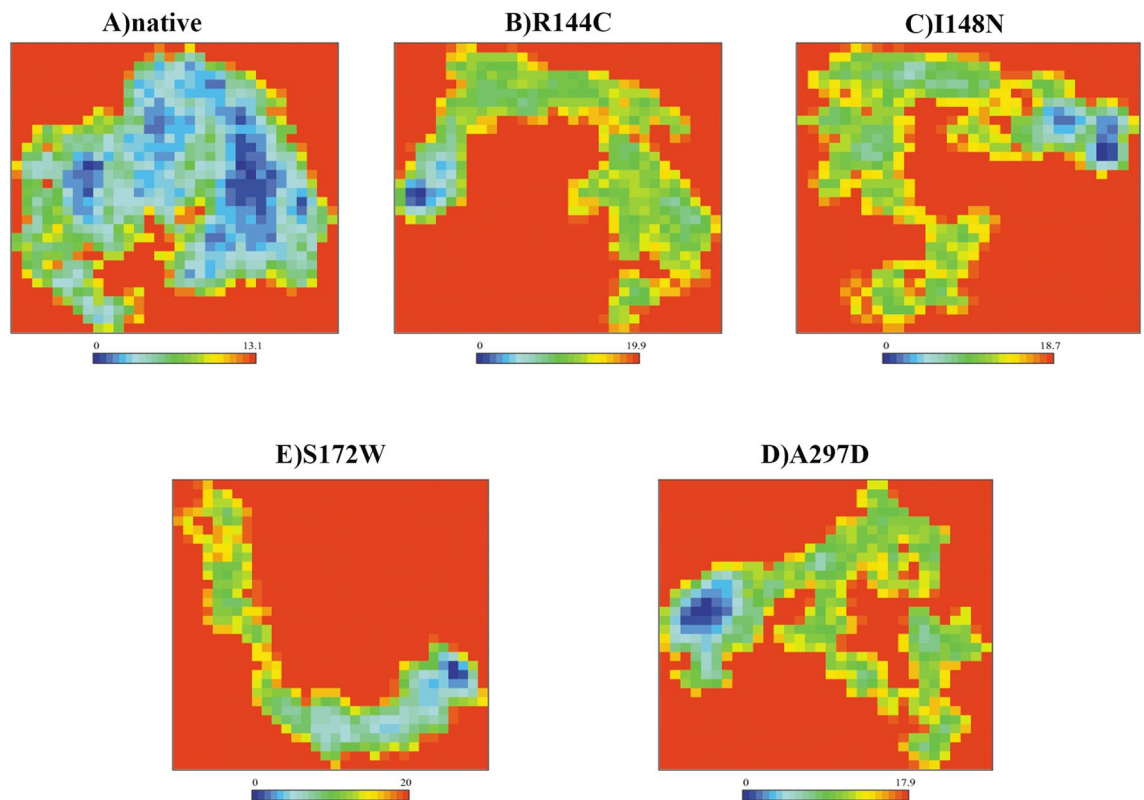


Figure 8. Free energy landscape (FEL) of the native structure and mutant proteins over 200 ns simulation. (A) native structure, (B) mutant R144C, (C) mutant I148N, (D) mutant S172W, and (E) mutant A297D.

and a greater range of Gibbs free energy than the native protein. As the folding pattern of a protein directly affects its stability, more unfolded states suggest a decrease in the stability of the protein structure resulting from the four most detrimental missense mutations. The results obtained by FEL analysis were consistent with those obtained by RMSD and Rg analysis.

Determining the effect of mutations on OX1R interaction

Previous studies have demonstrated that OX1R belongs to the GPCR protein family, which interacts with G protein complexes composed of Gi, Go, Gs, or G11 to transmit signals intracellularly⁴⁹. Therefore, any changes in amino acids in these regions may affect the interaction between OX1R and Gi protein. The structure analysis of OX1R mutants demonstrated that mutants R144C, I148N, and A297D except for S172W located in the region could affect OX1R signaling by changing the structure and the binding mode with Gi protein (Fig. 9A). In addition, the structure analysis of S172W after simulation demonstrated that this mutation altered the structure of the transmembrane alpha-helix and affected the position of OX1R in the membrane (Fig. 9B).

Protein–protein interaction analysis demonstrated that all four detrimental missense SNPs affected OX1R interaction with the Gi protein (Table S5). The data obtained from free binding energy calculations using the HAWDOCK web server, ΔG assessments with Prodigy, and the predictions of affinity changes ($\Delta\Delta G$) from the mCSM-PP2 tool collectively illustrate that mutants A297D and S172W had significantly lower free binding energy than the native, while mutants I148N and R144C had slightly higher free binding energy (Table 3). These results suggest that the complexes formed by mutants A297D and S172W exhibit greater stability with the Gi protein than those formed by I148N and R144C.

Discussion

Accumulation of data from numerous studies has demonstrated the contribution of various missense SNPs to the progression of diverse diseases^{36,37,40}. However, there is a paucity of knowledge regarding the alterations in OX1R structural and conformational dynamics caused by detrimental missense SNPs. In this study, we carried out a comprehensive in silico evaluation to forecast pathogenic SNPs and their potential impacts on the structure and function of OX1R protein by employing a combination of different computational approaches. To enhance prediction accuracy, we integrated tools from diverse categories, including homology-based, sequenced-based, consensus-based, and structure-based approaches. This strategy aims to increase confidence in the identification of potentially detrimental missense SNPs by mitigating biases in the results⁵⁰.

By filtering 4,295 SNPs located in the OX1R gene, a total of 329 missense SNPs were identified. Further evaluation in two steps, using 15 bioinformatics tools, revealed that 17 missense SNPs were pathogenic to OX1R. All

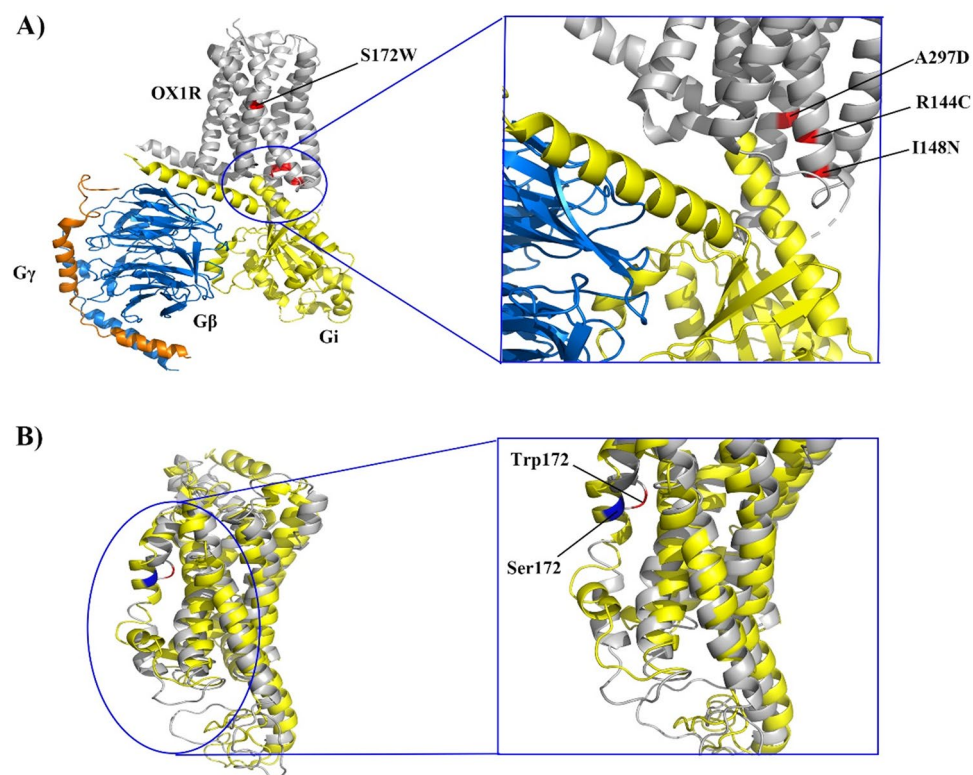


Figure 9. Structure analysis of OX1R mutant forms. (A) The mode of OX1R interaction with Gi protein. Data shows that mutant forms of I148N, A297D, and R144C except S172W are located in regions that directly interact with Gi protein. (B) The structural alignment of native (yellow) and mutant (gray) forms of OX1R after simulation. The result shows that changing Ser172 (red) to Trp caused changes in the alpha helix structure which is located transmembrane domain of the OX1R.

Model	HAWDOCK(MM/GBSA)					Prodigy		mCSM-PPI2	
	Free binding energy	VDW	ELE	GB	SA	ΔG (Kcal/mol)	Affinity	$\Delta\Delta G$ (kcal/mol)	Effect of affinity
Native	-53.73	-115.6	-829.9	907.5	-15.6	-10.7	1.3e-08	-	-
I148N	-50.43	-97.68	-967.24	1028.61	-14.12	-10.6	1.8e-08	-0.113	Decrease
A297D	-112.01	-223.4	-818.7	960.8	-30.5	-18.1	5.2e-14	0.318	Increasing
R144C	-47.4	-94.5	-833.6	893.5	-12.7	-9.6	8.7e-08	-0.462	Decrease
S172W	-88.19	-180.54	-1070.17	1187.11	-24.6	-16.8	4.8e-13	0.347	Increasing

Table 3. Protein–protein interaction analysis of OX1R with Gi protein. VDW Van der Waals potentials, ELE Electrostatic potentials, GB Polar solvation free energies, SA Nonpolar contribution to the solvation-free energy.

isolated 17 missense SNPs were found in the GTP-binding domain of the protein. This domain is responsible for catalytic activity. Therefore, mutations in this region are expected to affect protein activity and function^{4,31}.

In addition, stability analysis indicated that five SNPs (R144C, I148N, S172W, A297D, and L312Q) among the 17 pathogenic SNPs could lead to a decrease in OX1R stability, based on the results of all seven prediction tools. Protein stability is a critical characteristic that affects its structure, function, evolution, and biological activity, as well as its normal pathways⁵⁰. Any modification of protein stability may contribute to aberrant accumulation, misfolding, or protein degradation⁵¹.

Conservancy evaluation revealed that four of the five SNPs (R144C, I148N, S172W, and A297D) were found at crucial sites in the protein, as their conservation scores were high. The conservation level of a residue exhibits a strong correlation with its functional and structural significance. Generally, SNPs occurring in positions that are evolutionarily conserved are not well-tolerated, thereby resulting in the progression of diseases^{52,53}.

Hence, the four mutants, namely, R144C, I148N, S172W, and A297D, were identified as the most detrimental missense SNPs of the OX1R protein and selected for further computational analysis to determine their impact on the structure and function of the OX1R protein.

The structural impact of the four selected SNPs (R144C, I148N, S172W, and A297D) was assessed using the HOPE server. The R144C mutant leads to the substitution of arginine (positively charged residue) by cysteine (neutrally charged), where arginine is larger than cysteine. Charge and mass differences in proteins influence protein–protein spatiotemporal dynamics. The I148N, S172W, and A297D mutants had a larger residue than the wild-type moiety, which likely did not fit into the protein core. Consequently, these SNPs cause structural changes that are sometimes hazardous. Moreover, the wild-type residues of I148N, S172W, and A297D were more hydrophobic than the mutant residues, which could have caused repulsive hydrophobic interactions in the core of the protein, as predicted by the HOPE server. Generally, structural mutations cause changes in amino acid charge and size, hydrophobic propensity, salt bridges, and hydrogen bonds, which could cause loss of thermodynamic stability as well as aberrant folding and aggregation of proteins.

Secondary structure analysis conducted using the SOPMA server revealed that the most detrimental missense SNPs (R144C, I148N, S172W, and A297D) exhibited a modest decrease in the number of residues involved in the alpha-helix conformation, accompanied by a slightly higher proportion of extended strand conformation compared to the native structure. This observation suggests that these mutations may have a detrimental effect on OX1R protein stability.

Proteins possess a dynamic nature. We conducted 200 ns MDs to investigate the consequence of mutants R144C, I148N, S172W, and A297D on the structural dynamics and stability of the OX1R protein. The plateau of RMSD values revealed a significant decrease in protein stability for the mutants R144C, I148N, S172W, and A297D compared to the native structure throughout the MDs. Furthermore, a loss of stability and higher residue fluctuation was observed for all four mutant proteins compared to the native structure in the Rg and RMSF analysis. In support of these findings, SASA analysis showed that the high flexibility induced by all mutants may lead to expansion of the solvent-exposed area, leading to protein misfolding and being regarded as accountable for the loss of function. The secondary structure analysis also pointed out that the mutants R144C, I148N, S172W, and A297D compared to the native structure, introduced a slightly more β -strand conformation, whereas they disrupted the α -helix conformation, especially for the A297D mutant. These deviations might have an impact on protein folding thereby decreasing the stability of the protein. Results derived from the PCA confirmed that all four mutant proteins enhanced the overall flexibility and motion of the OX1R protein. Moreover, from FEL energy analysis, all mutants R144C, I148N, S172W, and A297D showed a significant difference in the stability of than native protein. A change in the protein stability could alter the protein function⁵⁴. Therefore, we indicate that the four most detrimental mutant proteins, R144C, I148N, S172W, and A297D might exhibit a significant effect on OX1R structure and function. This supposition is well-aligned with the outcomes derived from the investigations conducted by Saxena et al.⁵⁵, Agrahari et al.⁵⁶, and Shinwari et al.⁵⁷.

Signaling through OX1R is well established for its role in neuronal depolarization, which enhances excitability and firing rates through the activation of a non-selective cation current. Moreover, the orexin/OX1R system has been linked to inducing mitochondrial apoptosis, leading to a substantial reduction in cell growth across various cancer lines^{58,59}. An analysis of the structure positioning of the most detrimental mutant proteins on OX1R, in comparison with the μ -opioid protein complex as a GPCR receptor, revealed that the mutants R144C, I148N, and A297D are situated in regions directly interacting with the Gi protein⁶⁰. Protein–protein interaction studies further demonstrated that mutants A297D and S172W exhibited increased binding affinity with the Gi protein compared to the native form, while mutants I148N and R144C showed decreased binding affinity. The observed alterations in binding affinity with the Gi protein, as indicated by mutants R144C, I148N, S172W and A297D, suggest changes in Gi protein activation and subsequent downstream signaling pathways of OX1R. These structural changes have the potential to influence the regulation of homeostatic processes and antitumor activity associated with OX1R^{61,62}. Although S172W is not in close proximity to the Gi protein-binding site, secondary structure predictions suggest that it could induce conformational changes in OX1R protein, affecting its interaction with the Gi protein. This finding aligns with the work of Hong et al., indicating that interaction with extracellular region of OX2R can induce local and global conformational changes, impacting receptor activation⁴⁹.

Our study identified pathogenic missense SNPs in the OX1R structure and explore their functional consequences, which can be further verified by experimental analysis to determine their role more precisely. However, insights into the structural positioning of detrimental mutants and their effects on Gi protein interactions offer opportunities for the development of targeted drugs, precision medicine strategies, and novel approaches to neurological disorders and cancer therapeutics.

Conclusion

In the current study, we employed a combination of in silico tools, MDs, and protein–protein interaction analyses to identify the most deleterious missense mutations affecting the structure and function of the OX1R protein. We detected 17 pathogenic mutations among 329 missense SNPs within the OX1R gene, primarily located in the GTP-binding domain. Four highly conserved missense SNPs, namely R144C, I148N, S172W, and A297D among the pathogenic SNPs were identified as significantly altering the dynamic stability and structure conformation of OX1R. Our computational pipeline offers a promising approach for prioritizing potential therapeutic targets, particularly for neurological disorders and cancer therapeutics associated with OX1R mutations.

Methods

Data retrieval

The list of all SNPs related to the OX1R gene was collected from the National Center for Biotechnology Information (NCBI) dbSNP database and mapped on genome assembly GRCh37.p13 (hg19) through the utilization of Variation Viewer⁶³. “OX1R” or “HCRT1R” was utilized as our search keyword and filtered for SNPs (<https://>

www.ncbi.nlm.nih.gov/variation/view/?q=HCRTR1). The human OX1R protein's FASTA sequence was obtained from the UniProt database (UniProt ID: O43613)⁶⁴.

Determining 3D protein structure and its validation

As the complete structure of the OX1R protein is not found in the PDB bank, homology modeling was performed by the Robetta web server (<https://rosetta.bakerlab.org/>)⁶⁵. To validate the tertiary structure of the OX1R model, the SAVES web server which compromised ERRAT, Verify3D, and Procheck tools was applied (<https://saves.mbi.ucla.edu/>)⁶⁶. The ERRAT web tool computes the quality of non-bonded interaction. The ERRAT value of around 91% to 95% means good quality protein structure⁶⁶. The Verify3D web server investigates the compatibility of the secondary structure of a protein based on one-dimension (1D) with the three-dimensional (3D) structure of the protein structure. The Verify3D value of more than 80% illustrates compatibility between the two structures⁶⁷. The Procheck web tool analyzes the stereochemical quality of each residue geometry in protein structure according to the Ramachandran plot⁶⁸. Based on this web server, the higher percent of residues located in the core and generally allowed region indicate a higher quality of protein structure. The Molprobity web server calculates the resolution of protein structure by combing and evaluating the clash score, rotamer, and Ramachandran into a single score and normalizing the scale as an X-ray resolution⁶⁹. The higher Molprobity score illustrates the better quality of protein structure. The ProSA web server evaluates the quality of protein structures by comparing their quality with protein structures whose structure was identified by X-ray and NMR⁷⁰. Structural representation of human OX1R protein was done by using PyMOL software.

Determining the pathogenic nsSNPs

We used a multistep strategy employing various computational algorithms to identify deleterious nsSNPs in the human *OX1R* gene. In the first step, we predicted the functional effects of the nsSNPs on the protein using 10 different in silico prediction tools. These tools included FATHMM-MLK⁷¹, SIFT⁷², Mutation assessor⁷³, SNAP-2⁷⁴, Polyphen-2⁷⁵, Panther-PSEP⁷⁶, PON-P2⁷⁷, CADD⁷⁸, Align-GVGD⁷⁹, and VEST-4⁸⁰. In the next step, we utilized five computational tools, namely Pmut⁸¹, Suspect⁸², PhD-SNP⁸³, SNPs&GO⁸⁴, and InMeRF⁸⁵, to assess the pathogenicity and disease association of the identified nsSNPs.

Determining the domain of OX1R protein

To ascertain the domains and positions of disease-causing nsSNPs in the OX1R protein, four databases, namely InterPro⁸⁶, PROSITE⁸⁷, Pfam⁸⁸, and CDD⁸⁹, were used. For all four tools, the FASTA sequence of the protein was provided as input.

Determining SNP's impact on the OX1R protein stability

To understand whether the OX1R protein will be in a stable or denatured form due to amino acid substitutions, we applied eight different in silico algorithms: three sequenced-based tools, I-Mutant 2.0⁹⁰, INPS-MD⁹¹, Mu-Pro⁹², and five structure-based servers, m-CSM⁹³, SDM⁹⁴, DUET⁹⁵, and Dyna-Mut⁹⁶.

Determining the phylogenetically conserved residues in the OX1R protein

The evolutionary pattern of conservation at each amino acid position in OX1R was determined using the ConSurf server. For each residue of the target protein, the level of conservation was calculated on a scale of 1 to 4 as a variable, 5–6 as an average, 7–8 as conserved, and 9 as extremely conserved^{97,98}. Functional and structural residues have also been identified. The 3D structure of the OX1R protein was used for ConSurf evaluation.

Determining the phenotypic and structural consequences of selected SNPs

To further investigate the structural and functional effects of amino acid modifications on the OX1R protein, we utilized the HOPE server⁹⁹.

Alterations in the secondary structure of the OX1R protein due to pathogenic SNPs, based on the amino acid sequences of the native and mutant proteins, were provided by the SOPMA server¹⁰⁰.

Molecular dynamics simulations

To analyze the effect of each mutation on protein structure changes, molecular dynamics simulations (MDs) were performed for native and mutant forms by Gromacs 2020.7¹⁰¹. The OPLSA force field was applied during the simulation. A cubic simulation box was filled with SPCE water type. The simulation system was neutralized with Na⁺ and Cl⁻. The system was energy minimized by the steep descent minimization algorithm for a maximum of 50,000 steps. For equilibration of the simulation system, a Berendsen temperature (tcouple) of 300 K and a Parrinello-Rahman pressure (pcouple) of 1 bar were selected. The electrostatic energy was computed by the Partial Mesh Ewald (PME) algorithm and the LINCS algorithm was applied to constrain all bonds. MDs were performed for 200 ns and the trajectory was given at intervals of 1 ps. The trajectory was used to analyze the structure changes by Root Mean Square Deviation (RMSD), Gyrate (Rg), Root Mean Square Fluctuation (RMSF), Hydrogen bond (H-bond), Solvent Access Surface Area (SASA), Principal component analysis (PCA), and Free energy landscapes (FELs) for both native and mutant forms. To compute the secondary structure changes during MDs, DSSP software was used.

Determining the effect of mutations on OX1R interaction

To analyze the effect of mutations on the activation of the OX1R pathway, protein–protein interactions were performed using the PPI3D web server¹⁰². To do this, the conformation mode of PDB ID:7u2k of C6-guano

bound Mu Opioid Receptor–Gi Protein Complex was applied. The results of the PPI3D tool were assessed using the Ligplot package to illustrate the amino acids that play a role in the interactions¹⁰³. Furthermore, the HAWDOCK tool was used to determine the types and total free-binding energies of the protein–protein complexes. The binding affinity of the protein–protein complex was calculated using the Prodigy server. The effect of each mutation on protein–protein was studied by the mCSM–PPI2 tool¹⁰⁴.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 24 August 2023; Accepted: 12 December 2023

Published online: 08 January 2024

References

- Villano, I. *et al.* Physiological role of orexinergic system for health. *Int. J. Environ. Res. Public Health* **19**, 8353 (2022).
- Katzman, M. A. & Katzman, M. P. Neurobiology of the orexin system and its potential role in the regulation of hedonic tone. *Brain Sci.* **12**, 150 (2022).
- Sakurai, T. *et al.* Orexins and orexin receptors: A family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573–585 (1998).
- Yin, J. *et al.* Structure and ligand-binding mechanism of the human OX1 and OX2 orexin receptors. *Nat. Struct. Mol. Biol.* **23**, 293–299 (2016).
- Young, J. K. *et al.* Orexin stimulates breathing via medullary and spinal pathways. *J. Appl. Physiol.* **98**, 1387–1395 (2005).
- Kuwaki, T. Orexin (hypocretin) participates in central autonomic regulation during fight-or-flight response. *Peptides* **139**, 170530 (2021).
- Mieda, M. The roles of orexins in sleep/wake regulation. *Neurosci. Res.* **118**, 56–65 (2017).
- Georgescu, D. *et al.* Involvement of the lateral hypothalamic peptide orexin in morphine dependence and withdrawal. *J. Neurosci.* **23**, 3106–3111 (2003).
- Harris, G. C., Wimmer, M. & Aston-Jones, G. A role for lateral hypothalamic orexin neurons in reward seeking. *Nature* **437**, 556–559 (2005).
- Sakurai, T. The role of orexin in motivated behaviours. *Nat. Rev. Neurosci.* **15**, 719–731 (2014).
- Nevárez, N. & de Lecea, L. Recent advances in understanding the roles of hypocretin/orexin in arousal, affect, and motivation. *F1000Research* **7**, (2018).
- Inutsuka, A. *et al.* The integrative role of orexin/hypocretin neurons in nociceptive perception and analgesic regulation. *Sci. Rep.* **6**, 29480 (2016).
- Milbank, E. & López, M. Orexins/hypocretins: Key regulators of energy homeostasis. *Front. Endocrinol.* **10**, 830 (2019).
- Johnson, P. L., Molosh, A., Fitz, S. D., Truitt, W. A. & Shekhar, A. Orexin, stress, and anxiety/panic states. *Prog. Brain Res.* **198**, 133–161 (2012).
- Liu, M.-F. *et al.* Orexin-A exerts neuroprotective effects via OX1R in Parkinson's disease. *Front. Neurosci.* **12**, 835 (2018).
- Xu, D. *et al.* Orexin-A alleviates astrocytic apoptosis and inflammation via inhibiting OX1R-mediated NF- κ B and MAPK signaling pathways in cerebral ischemia/reperfusion injury. *Biochim. Biophys. Acta (BBA) Mol. Basis Dis.* **1867**, 166230 (2021).
- Ammoun, S., Lindholm, D., Wootz, H., Åkerman, K. E. O. & Kukkonen, J. P. G-protein-coupled OX1 orexin/hcrtr-1 hypocretin receptors induce caspase-dependent and-independent cell death through p38 mitogen-/stress-activated protein kinase. *J. Biol. Chem.* **281**, 834–842 (2006).
- Murakami, M. *et al.* Involvement of the orexin system in sympathetic nerve regulation. *Biochem. Biophys. Res. Commun.* **460**, 1076–1081 (2015).
- Samson, W. K., Bagley, S. L., Ferguson, A. V. & White, M. M. Hypocretin/orexin type 1 receptor in brain: Role in cardiovascular control and the neuroendocrine response to immobilization stress. *Am. J. Physiol. Integr. Comp. Physiol.* **292**, R382–R387 (2007).
- Akça, Ö. F., Uzun, N. & Kılınc, İ. Orexin A in adolescents with anxiety disorders. *Int. J. Psychiatry Clin. Pract.* **24**, 127–134 (2020).
- Chemelli, R. M. *et al.* Narcolepsy in orexin knockout mice: Molecular genetics of sleep regulation. *Cell* **98**, 437–451 (1999).
- Feng, Y. *et al.* Neuroprotection by Orexin-A via HIF-1 α induction in a cellular model of Parkinson's disease. *Neurosci. Lett.* **579**, 35–40 (2014).
- Overeem, S. *et al.* The clinical features of cataplexy: a questionnaire study in narcolepsy patients with and without hypocretin-1 deficiency. *Sleep Med.* **12**, 12–18 (2011).
- Xiong, X. *et al.* Mitigation of murine focal cerebral ischemia by the hypocretin/orexin system is associated with reduced inflammation. *Stroke* **44**, 764–770 (2013).
- Ji, M.-J., Zhang, X.-Y., Chen, Z., Wang, J.-J. & Zhu, J.-N. Orexin prevents depressive-like behavior by promoting stress resilience. *Mol. Psychiatry* **24**, 282–293 (2019).
- Baykal, S. *et al.* Decreased serum orexin A levels in drug-naive children with attention deficit and hyperactivity disorder. *Neurol. Sci.* **40**, 593–602 (2019).
- Johnson, P. L. *et al.* A key role for orexin in panic anxiety. *Nat. Med.* **16**, 111–115 (2010).
- Davies, J. *et al.* Orexin receptors exert a neuroprotective effect in Alzheimer's disease (AD) via heterodimerization with GPR103. *Sci. Rep.* **5**, 12584 (2015).
- Hu, S. *et al.* Orexin A associates with inflammation by interacting with OX1R/OX2R receptor and activating prepro-Orexin in cancer tissues of gastric cancer patients. *Gastroenterol. Hepatol.* **43**, 240–247 (2020).
- Alexandre, D. *et al.* The orexin type 1 receptor is overexpressed in advanced prostate cancer with a neuroendocrine differentiation, and mediates apoptosis. *Eur. J. Cancer* **50**, 2126–2133 (2014).
- Annerbrink, K. *et al.* Panic disorder is associated with the Val308Iso polymorphism in the hypocretin receptor gene. *Psychiatr. Genet.* **21**, 85–89 (2011).
- Harro, J. *et al.* The orexin 1 receptor (HCRTR1) Gene as a susceptibility gene Contributing to polydipsia-hyponatremia in schizophrenia. *Neuropharmacology* **156**, 107527 (2019).
- Cengiz, M. *et al.* Orexin/hypocretin receptor, Orx1, gene variants are associated with major depressive disorder. *Int. J. Psychiatry Clin. Pract.* **23**, 114–121 (2019).
- Hamit, G., Ayca, O. A., Omer, B. I., Nevra, O. & Aynur, O. Association of circadian locomotor output cycles kaput rs1801260 and hypocretin receptor 1 rs2271933 polymorphisms in patients with chronic migraine and sleep disorder. *Neurol. Sci. Neurophysiol.* **39**, 98 (2022).
- Center, I. L. Association of single nucleotide polymorphisms of endothelin, orexin and vascular endothelial growth factor receptor genes with obstructive sleep apnea among Thai ethnic. *J Med Assoc Thai* **99**, S150–S157 (2016).

36. Karimi, S. *et al.* Impact of SNPs, off-targets, and passive permeability on efficacy of BCL6 degrading drugs assigned by virtual screening and 3D-QSAR approach. *Sci. Rep.* **12**, 21091 (2022).
37. Khan, N. *et al.* Investigating pathogenic SNP of PKC ϵ in HCV-induced hepatocellular carcinoma. *Sci. Rep.* **13**, 12504 (2023).
38. Farajzadeh-Dehkordi, M. *et al.* Evaluation of a warfarin dosing algorithm including CYP2C9, VKORC1, and CYP4F2 polymorphisms and non-genetic determinants for the Iranian population. *Pharmacol. Rep.* **1**, 1–10 (2023).
39. Alamri, A., Alkhalaiwi, F., Khan, N. U. & Tasleem, M. In silico screening and validation of *achyranthes aspera* as a potential inhibitor of BRAF and NRAS in controlling thyroid cancer. *Anticancer. Agents Med. Chem.* (2023).
40. Tasleem, M. *et al.* Computational analysis of PTP-1B site-directed mutations and their structural binding to potential inhibitors. *Cell. Mol. Biol.* **68**, 75–84 (2022).
41. Ahammad, I. *et al.* Impact of highly deleterious non-synonymous polymorphisms on GRIN2A protein's structure and function. *PLoS One* **18**, e0286917 (2023).
42. Saxena, S. *et al.* A bioinformatics approach to the identification of novel deleterious mutations of human TPMT through validated screening and molecular dynamics. *Sci. Rep.* **12**, 18872 (2022).
43. Jayaraj, J. M. & Muthusamy, K. Role of deleterious nsSNPs of klotho protein and their drug response: A computational mechanical insights. *J. Biomol. Struct. Dyn.* **1**, 1–11 (2023).
44. Ogun, O. J., Soremekun, O. S., Thaller, G. & Becker, D. An in silico functional analysis of non-synonymous single-nucleotide polymorphisms of bovine CMAH gene and potential implication in pathogenesis. *Pathogens* **12**, 591 (2023).
45. Hoda, A., Bixheku, X. & Lika, M. Computational analysis of non-synonymous single nucleotide polymorphism in the bovine PKLR gene: Computational analysis of bovine PKLR gene. *J. Biomol. Struct. Dyn.* **1**, 1–14 (2023).
46. Hashemi-Shahraki, F., Shareghi, B. & Farhadian, S. Investigation of the interaction behavior between quercetin and pepsin by spectroscopy and MD simulation methods. *Int. J. Biol. Macromol.* **227**, 1151–1161 (2023).
47. Hashemi-Shahraki, F., Shareghi, B., Farhadian, S. & Yadollahi, E. A comprehensive insight into the effects of caffeic acid (CA) on pepsin: Multi-spectroscopy and MD simulations methods. *Spectrochim. Acta A.* **289**, 122240 (2023).
48. Yang, L., Song, G., Carriquiry, A. & Jernigan, R. L. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. *Structure* **16**, 321–330 (2008).
49. Hong, C. *et al.* Structures of active-state orexin receptor 2 rationalize peptide and small-molecule agonist recognition and receptor activation. *Nat. Commun.* **12**, 815 (2021).
50. Khalid, Z. & Almaghrabi, O. Mutational analysis on predicting the impact of high-risk SNPs in human secretory phospholipase A2 receptor (PLA2R1). *Sci. Rep.* **10**, 1–11 (2020).
51. Clausen, L. *et al.* Protein stability and degradation in health and disease. *Adv. Protein Chem. Struct. Biol.* **114**, 61–83 (2019).
52. Greene, L. H. *et al.* Role of conserved residues in structure and stability: Tryptophans of human serum retinol-binding protein, a model for the lipocalin superfamily. *Protein Sci.* **10**, 2301–2316 (2001).
53. Williamson, K. *et al.* Catalytic and functional roles of conserved amino acids in the SET domain of the *S cerevisiae* lysine methyltransferase Set1. *PLoS ONE* **8**, e57974 (2013).
54. Bromberg, Y. & Rost, B. Correlating protein function and stability through the analysis of single amino acid substitutions. *BMC Bioinform.* **10**, 1–9 (2009).
55. Saxena, S., Murthy, T. P. K., Chandramohan, V., Yadav, A. K. & Singh, T. R. Structural and functional analysis of disease-associated mutations in GOT1 gene: An in silico study. *Comput. Biol. Med.* **136**, 104695 (2021).
56. Aghahari, A. K. *et al.* Understanding the structure-function relationship of HPRT1 missense mutations in association with Lesch-Nyhan disease and HPRT1-related gout by in silico mutational analysis. *Comput. Biol. Med.* **107**, 161–171 (2019).
57. Shinwari, K. *et al.* In-silico assessment of high-risk non-synonymous SNPs in ADAMTS3 gene associated with Hennekam syndrome and their impact on protein stability and function. *BMC Bioinform.* **24**, 251 (2023).
58. Marcos, P. & Coveñas, R. Involvement of the orexinergic system in cancer: Antitumor strategies and future perspectives. *Appl. Sci.* **13**, 7596 (2023).
59. Couvineau, A., Nicole, P., Gratio, V. & Voisin, T. The Orexin receptors: Structural and anti-tumoral properties. *Front. Endocrinol.* **13**, 931970 (2022).
60. Koehl, A. *et al.* Structure of the μ -opioid receptor-Gi protein complex. *Nature* **558**, 547–552 (2018).
61. Heifetz, A. *et al.* Study of human Orexin-1 and-2 G-protein-coupled receptors with novel and published antagonists by modeling, molecular dynamics simulations, and site-directed mutagenesis. *Biochemistry* **51**, 3178–3197 (2012).
62. Nicole, P., Couvineau, P., Jamin, N., Voisin, T. & Couvineau, A. Crucial role of the orexin-BC-terminus in the induction of OX1 receptor-mediated apoptosis: Analysis by alanine scanning, molecular modelling and site-directed mutagenesis. *Br. J. Pharmacol.* **172**, 5211–5223 (2015).
63. Sherry, S. T. *et al.* dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
64. Consortium U. The universal protein resource (UniProt) in 2010. *Nucleic Acids Res.* **38**, D142–D148 (2010).
65. Kim, D. E., Chivian, D. & Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **32**, W526–W531 (2004).
66. Colovos, C. & Yeates, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **2**, 1511–1519 (1993).
67. Lüthy, R., Bowie, J. U. & Eisenberg, D. Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85 (1992).
68. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
69. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
70. Wiederstein, M. & Sippl, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410 (2007).
71. Shihab, H. A. *et al.* Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum. Mutat.* **34**, 57–65 (2013).
72. Sim, N.-L. *et al.* SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **40**, W452–W457 (2012).
73. Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* **39**, e118–e118 (2011).
74. Hecht, M., Bromberg, Y. & Rost, B. Better prediction of functional effects for sequence variants. *BMC Genomics* **16**, 1–12 (2015).
75. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
76. Tang, H. & Thomas, P. D. PANTHER-PSEP: Predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics* **32**, 2230–2232 (2016).
77. Niroula, A., Urolagin, S. & Vihinen, M. PON-P2: Prediction method for fast and reliable identification of harmful variants. *PLoS ONE* **10**, e0117380 (2015).
78. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019).
79. Tavtigian, S. V., Byrnes, G. B., Goldgar, D. E. & Thomas, A. Classification of rare missense substitutions, using risk surfaces, with genetic-and molecular-epidemiology applications. *Hum. Mutat.* **29**, 1342–1354 (2008).

80. Carter, H., Douville, C., Stenson, P. D., Cooper, D. N. & Karchin, R. Identifying Mendelian disease genes with the variant effect scoring tool. *BMC Genomics* **14**, 1–16 (2013).
81. Ferrer-Costa, C. *et al.* PMUT: A web-based tool for the annotation of pathological mutations on proteins. *Bioinformatics* **21**, 3176–3178 (2005).
82. Yates, C. M., Filippis, I., Kelley, L. A. & Sternberg, M. J. E. SuSPect: Enhanced prediction of single amino acid variant (SAV) phenotype using network features. *J. Mol. Biol.* **426**, 2692–2701 (2014).
83. Capriotti, E., Calabrese, R. & Casadio, R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics* **22**, 2729–2734 (2006).
84. Capriotti, E. *et al.* WS-SNPs&GO: A web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics* **14**, 1–7 (2013).
85. Takeda, J. *et al.* InMeRF: Prediction of pathogenicity of missense variants by individual modeling for each amino acid substitution. *NAR Genomics Bioinforma.* **2**, 038 (2020).
86. Hunter, S. *et al.* InterPro in 2011: New developments in the family and domain prediction database. *Nucleic Acids Res.* **40**, D306–D312 (2012).
87. Hulo, N. *et al.* The PROSITE database. *Nucleic Acids Res.* **34**, D227–D230 (2006).
88. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
89. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–D226 (2015).
90. Capriotti, E., Fariselli, P. & Casadio, R. I-Mu[1] E. Capriotti, P. Fariselli, R. Casadio, I-Mutant20: Predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res.* **33**, W306–W310 (2005).
91. Savojardo, C., Fariselli, P., Martelli, P. L. & Casadio, R. INPS-MD: A web server to predict stability of protein variants from sequence and structure. *Bioinformatics* **32**, 2542–2544 (2016).
92. Cheng, J., Randall, A. & Baldi, P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins Struct. Funct. Bioinform.* **62**, 1125–1132 (2006).
93. Pires, D. E. V., Ascher, D. B. & Blundell, T. L. mCSM: Predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **30**, 335–342 (2014).
94. Pandurangan, A. P., Ochoa-Montaño, B., Ascher, D. B. & Blundell, T. L. SDM: A server for predicting effects of mutations on protein stability. *Nucleic Acids Res.* **45**, W229–W235 (2017).
95. Pires, D. E. V., Ascher, D. B. & Blundell, T. L. DUET: A server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* **42**, W314–W319 (2014).
96. Rodrigues, C. H. M., Pires, D. E. V. & Ascher, D. B. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* **46**, W350–W355 (2018).
97. Berezin, C. *et al.* ConSeq: The identification of functionally and structurally important residues in protein sequences. *Bioinformatics* **20**, 1322–1324 (2004).
98. Ashkenazy, H. *et al.* ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
99. Venselaar, H., Te Beek, T. A. H., Kuipers, R. K. P., Hekkelman, M. L. & Vriend, G. Protein structure analysis of mutations causing inheritable diseases: An e-Science approach with life scientist friendly interfaces. *BMC Bioinform.* **11**, 1–10 (2010).
100. Geourjon, C. & Deleage, G. SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* **11**, 681–684 (1995).
101. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
102. Dapkūnas, J. *et al.* The PPI3D web server for searching, analyzing and modeling protein–protein interactions in the context of 3D structures. *Bioinformatics* **33**, 935–937 (2017).
103. Laskowski, R. A. & Swindells, M. B. *LigPlot+*: Multiple ligand–protein interaction diagrams for drug discovery (Springer, 2011).
104. Rodrigues, C. H. M., Myung, Y., Pires, D. E. V. & Ascher, D. B. mCSM-PPI2: Predicting the effects of mutations on protein–protein interactions. *Nucleic Acids Res.* **47**, W338–W344 (2019).

Author contributions

Conceptualization, M.F.-D., L.M., H.H.-Y., H.P. and B.R.; methodology, M.F.-D., L.M., and A.H.; formal analysis, M.F.-D., L.M., and A.H.; writing—original draft preparation; M.F.-D. and L.M.; writing—review and editing, L.M., and B.R.; supervision, B.R. All authors reviewed and approved the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-49809-4>.

Correspondence and requests for materials should be addressed to B.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024