ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

# Computational and Structural Analysis to Assess the Pathogenicity of Bardet-Biedl Syndrome Related Missense Variants Identified in Bardet-Biedl Syndrome 10 Gene (BBS10)

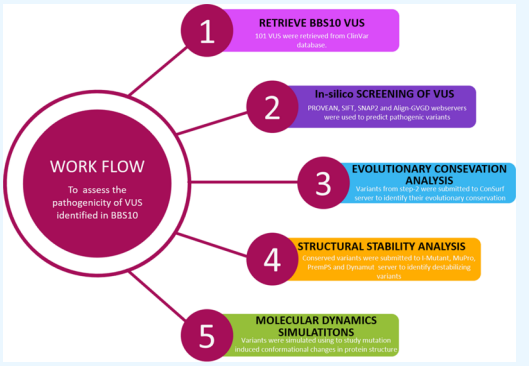
Neha Gupta,*[,⊥] Mudassar Ali Khan,[⊥] Giovambattista Capasso, and Miriam Zacchia

Cite This: https://doi.org/10.1021/acsomega.2c04522

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Bardet-Biedl Syndrome (BBS) is a rare inherited disorder resulting in multiple organ dysfunctions, whose cardinal clinical features include cognitive impairment, obesity, and renal dysfunction. Although it is highly heterogeneous at genetic levels, *BBS10* is one of the major causative genes worldwide. The BBS10 protein is part of a multiprotein complex localized at the basal body of the primary cilium. With the advancement of sequencing technologies, novel missense mutations are regularly reported in *BBS10*. However, prioritizing missense variants and conducting further in-depth analysis are key challenges in addressing their pathogenic effect. This study aims to characterize the known missense mutations of *BBS10* by combining nine different *in silico* tools (SIFT, SNAP2, PROVEAN, Align-GVGD, ConSurf, I Mutant, MuPro, PremPS, and Dynamut) and molecular dynamics (MD) simulations. A total of 101 *BBS10* missense variants have been analyzed. Our results showed that six *BBS10* missense variants (Ser191Leu, Cys19Gly, Ile342Thr, Cys371Ser, Ala417Glu, and Tyr613Cys) were potentially deleterious. Overall, this study provides a comprehensive workflow for screening BBS10 missense mutations to identify pathogenic variants effectively.






## 1. INTRODUCTION

Bardet-Biedl syndrome (BBS; MIM #209900) is an autosomal recessive disorder resulting in multiple organ dysfunctions, including retinal degeneration, cognitive impairment, obesity, and renal dysfunction.[1] The genetic heterogeneity and clinical variability have been reported to be very high among BBS patients.[2] BBS is a rare genetic disorder, and its rate of incidence varies geographically. The estimated cases of BBS were found to be 1 in 160,000 in Switzerland and 1 in 36,000 in the mixed Arab population of Kuwait.[3−5] Interestingly, in small isolated populations, the rate of incidence of BBS is quite high, such as 1 in 18,000 in Newfoundland, Canada and 1 in 13,500 in Kuwaiti families of Bedouin ancestry.[5,6] To date, 24 genes have been discovered, but three of them, *BBS1*, *BBS2*, and *BBS10*, account for almost 50% of patients.[7,8]

BBS10 is among the major contributors of BBS, accounting for 20% of all of the cases, with few exceptions in the ethnically homogeneous group of Danish and Spanish BBS cohorts.[9−11] The BBS10 gene located on chromosome 12q21.2 encodes 723 amino acids and plays an essential role in the formation and function of the basal body and primary cilia.[12] BBS10 has a type II chaperonin-like structure and has been shown to assist BBSome assembly, a multiprotein complex involved in intraflagellar trafficking.[13−15]

Nonsense or frameshift mutations identified in *BBS10* are generally pathogenic as they result in a nonfunctional protein.[10,16] Nevertheless, missense mutations could also be deleterious; however, the impact of identified missense mutations is rarely studied.[17,18]

Consequently, it is crucial to develop *in silico* approaches to identify the significant functional mutations that might aid in the early detection and clinical management of BBS.

A total of 213 mutations, including 101 missense variants of uncertain significance (VUS), have already been reported in the *ClinVar* database (https://www.ncbi.nlm.nih.gov/clinvar). Elucidating their pathogenic effect has a pivotal relevance, but it is challenging.

In the past decade, various computational tools have been developed to predict the effect of missense variants on a protein's structure and, eventually, its function. The combination of multiple computational approaches and the consensus of their prediction outcomes could narrow down the candidate mutations for further validation.[19] We have utilized nine different servers, SIFT, SNAP2, PROVEAN, Align-GVGD, ConSurf, I Mutant, MuPro, PremPS, and Dynamut, to predict

Received: July 18, 2022
Accepted: September 28, 2022

the deleterious missense variants of BBS10.[20−28] Protein functions are not only related to the static structures determined by their amino acid sequence but are also highly dependent on protein dynamics. Therefore, we analyzed protein stability via molecular dynamics simulation to deeply analyze the structural changes arising in the potentially deleterious mutants of BBS10. This article highlights the workflow of computational screening and analysis of missense variants of *BBS10*. We conclude that our study will aid researchers in further investigating the roles of the BBS10 gene and its encoded protein in Bardet-Biedl syndrome.

## 2. MATERIALS AND METHODS

**2.1. Data Collection.** All information about the human BBS10 gene was retrieved from public web-based resources. The reported *BBS10* missense variants of uncertain significance (VUS) were collected from the *ClinVar* database (https://www.ncbi.nlm.nih.gov/clinvar).[29] The amino acid sequence (UniProt ID: Q8TAM1) encodes a BBS10 protein retrieved from the UniProt database (https://www.uniprot.org/uniprot/Q8TAM1).

**2.2. Prediction of Disease-Related Missense Mutations.** *2.2.1. Prediction of Functional Consequences of Missense Mutations.* The functional effects of missense mutations were predicted by SIFT (Sorting Intolerant from Tolerant), SNAP2 (screening of nonacceptable polymorphism 2), PROVEAN (Protein Variation Effect Analyzer), and the Align-GVGD tool, and missense mutations were assigned as deleterious mutations by consistent predictions of all four tools:

**PROVEAN** (Protein Variation Effect Analyzer; http://provean.jcvi.org/index.php) is a software tool that utilizes an alignment-based score approach to predict the functional effect of single or multiple amino acid substitution and indels. The server provides rapid analysis of protein variants from any organism and supports high-throughput analysis for human and mouse variants at genomic and protein levels (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528627/). We submitted the query protein sequence and amino acid variations to the PROVEAN server that performed a BLAST search to collect homologous sequences. The scores were calculated for each mutation. The threshold PROVEAN score was −2.5 to distinguish deleterious substitutions from neutral ones.[22,30]

**SIFT** (Sorting Intolerant from Tolerant; http://sift.bii.a-star.edu.sg) is a bioinformatics tool to predict if an amino acid substitution in a protein will affect function. This tool is based on the physiochemical properties of amino acids in the protein sequence and its sequence homologies. SIFT can be applied to naturally occurring nonsynonymous polymorphisms and laboratory-induced missense mutations. The predicted results of the SIFT program can be categorized into two classes: tolerated and deleterious. The variant is predicted to be deleterious if its substitution SIFT score is between 0 and 0.05, and 1 is marked as tolerable.[20]

**SNAP2** (https://rostlab.org/services/snap/) is a neural-network-based prediction mainframe that identifies the functional effect of amino acid sequence variants or predicts the impact of single amino acid acid substitution on protein function. The SNAP2 prediction score ranges from −100 (strong neutral prediction) to 100 (strong effect prediction), reflecting the likelihood of the specific mutation altering the native protein function.[21,31]

**Align-GVGD** (http://agvgd.iarc.fr) is a web-based server that combines the biophysical characteristics of amino acids and protein multiple sequence alignments to predict where missense substitutions in genes of interest are positioned in a spectrum from enriched neutral to enriched deleterious. The A-GVGD software program is an extension of the Grantham difference to multiple sequence alignments and accurate concurrent multiple comparisons.[23]

*2.2.2. Estimation of Evolutionary Conservation of Missense Mutations.* The ConSurf bioinformatics (http://consurf.tau.ac.il/) tool estimates the evolutionary conservation of amino acids in a protein sequence based on the phylogenetic relations between homologous sequences.[24] This tool determines the degree to which an amino acid position is evolutionarily conserved and is strongly dependent on its structural and functional importance. The level of evolutionary conservation of each sequence position corresponds to the evolutionary rate, i.e., not constant among all amino acids in a protein. The amino acid positions that evolve gradually are commonly considered conserved sites that are thought to be an essential factor for protein structure and function. We submitted the amino acid sequence of BBS10 to the ConSurf server, which calculates the conservation scores categorized into a discrete scale of nine bins. The position with bin 9 indicates the most conserved sites, while the position with bin 1 indicates the most variable sites.

*2.2.3. Prediction of Protein Change Stability of Missense Mutations.* Accurate prediction of protein stability changes upon single point mutations is essential for understanding protein structure and function. In the present study, we used MuPro, I Mutant 3.0, PremPS, and Dynamut Web servers to predict protein stability changes for the missense variants. The MuPro (http://mupro.proteomics.ics.uci.edu/) and I Mutant 3.0 Web servers (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) are vector machine-based tools to predict protein stability changes for single amino acid mutations based on the protein sequence. The protein sequence, the position of mutation, and the mutant residue were uploaded, and the protein stability was analyzed at default temperatures and pH.[25,26]

The PremPS (https://lilab.jysw.suda.edu.cn/research/PremPS) and Dynamut web servers (http://biosig.unimelb.edu.au/dynamut/) take protein structure as an input to predict the effect of single amino acid substitution on the stability of the protein structure. PremPS predicts the change in stability by calculating the changes in unfolding Gibbs free energy ($\Delta\Delta G$).[27] The Dynamut web server predicts the impact on protein stability using the Normal Mode Analysis (NMA) approach. It combines the score of changes in protein stability and vibrational entropy between the wild-type and mutant protein structures.[28]

**2.3. Molecular Dynamics Simulations.** Molecular Dynamics Simulation (MDS) studies investigate the structural stability and dynamics of *wild-type* and mutant protein structures.[32] MDS was carried out using GROMACS 2018.1, implementing the OPLS-AA/L force field for a time scale of 100 ns.[33,34] The system was solvated using the TIP3P water model in a cubic box with periodic boundary conditions, and counterions $Na^+$ and $Cl^-$ were added to neutralize the system. Energy minimization was performed using the steepest descent algorithm with a tolerance of 1000 kJ/mol/nm. The system was equilibrated by applying positional restraints on the structure using NVT (constant Number of particles, Volume,


https://doi.org/10.1021/acsomega.2c04522
ACS Omega XXXX, XXX, XXX−XXX

and Temperature) followed by the NPT (constant Number of particles, Pressure, and Temperature) ensemble for 100 ps each. The temperature of 300 K was coupled by the Berendsen thermostat with a pressure of one bar.[35] The equilibrated system was subjected to 100 ns of the production run with time-step integration of 2 fs. The trajectories were saved at every 2 ps and analyzed using Gromacs 2020.4. Furthermore, RMSD, RMSF, the radius of gyration (Rg), SASA, and the free energy landscape were analyzed using Gromacs inbuilt functions.

## 3. RESULTS

**3.1. Modeling of BBS10 Three-Dimensional Structure.** The crystal structure of BBS10 has not been determined yet, so the three-dimensional structure was modeled using the Robetta protein structure prediction service (https://robetta.bakerlab.org).[36] The model with a maximum number of residues in the allowed region was refined with the ModLoop server (https://modbase.compbio.ucsf.edu/modloop/).[37] The modeled structure was further refined using GalaxyWEB (seoklab.org) server. The refined models obtained from GalaxyWEB were subjected to Ramachandran plot analysis. The model with the highest residues in the core and the allowed region was selected as the final model (Figure 1). The

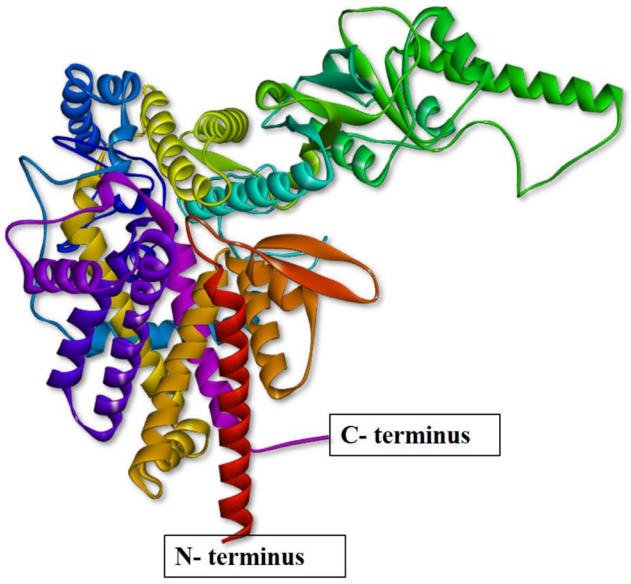



**Figure 1.** Three-dimensional modeled structure of BBS10 protein.

model was also validated using SAVESv6.0 - Structure Validation Server (ucla.edu), SWISS model workspace, and Protein Structure Analysis (ProSA) Web.[38,39] Ramachandran plot statistics of the BBS10 model obtained from the Procheck package on SAVESv6.0 - Structure Validation Server shows 94.9% residues at the core region, 4.1% in the allowed region, and 0.0% in the disallowed region (Supplementary Figure S1).

The modeled structure of BBS10 was further compared with the structure of BBS10 submitted in the AlphaFold Protein Structure Database.[40] The Ramachandran plot statistics of the modeled BBS10 structure were found to be better as compared to the AlphaFold structure. Secondary structure analysis of the two PDB structures indicated a close correlation in the high confidence regions predicted by AlphaFold. The major difference between the two structures was observed in the low confidence region predicted by AlphaFold (from residue: 460−588). The two structures were also compared using the SWISS model workspace, and the modeled structure was observed to have better stereochemistry than the AlphaFold structure (Supplementary Figure S2). Therefore, the modeled structure of BBS10 has been used for structure-based mutational analysis and MD simulation studies.

**3.2. Screening of Missense VUS Based on Functional Analysis.** A total of 101 missense variants of *BBS10* were analyzed for their functional effect using sequence-based *in silico* servers such as PROVEAN, SIFT, SNAP2, and Align-GVGD. Out of 101 variants, PROVEAN predicted 15 variants as "deleterious" with a cutoff score less than or equal to −2.5, and 86 variants were predicted as "neutral" with a score greater than −2.5. SIFT server predicted 46 variants as "damaging" with a cutoff score less than or equal to 0.05, and the remaining 55 variants were predicted as "tolerated". Out of 101 variants, the SNAP2 server predicted 55 variants as "effect" with scores > 0, and the remaining 46 variants were predicted as "neutral". Moreover, all of the variants were also analyzed by the Align-GVGD server, which predicted 52 variants as "class C65" (Supplementary Table S1). The "class C65" indicates that the variants are most likely to interfere with the protein's function. Overall, 13 variants were predicted to be deleterious by all four servers and, hence, were chosen for further analysis (Table 1).

**3.3. Screening of Deleterious VUS Using Evolutionary Conservation of Sequences.** The evolutionary conservation of a protein provides essential information regarding its structure and function. Conversely, the evolutionary conservation of an amino acid in a protein affects its structure and function. Therefore, it is essential to analyze the evolutionary conservation of amino acids to understand variants which could be deleterious to the function of its protein. The ConSurf server was utilized to test the extent of evolutionary conservation of potentially deleterious variants predicted by the PROVEAN, SIFT, SNAP2, and Align-GVGD tools. ConSurf server calculates the conservation scores on a scale of 1 to 9 grades. The degree of conservation in the range of 1 to 3 is considered variable, 4 to 6 is considered average, and the scale from 7 to 9 is considered conserved. Based on the results obtained from ConsSurf, variants having a conservation score of 7 or above were selected (Figure 2). Out of 13 variants, 10 were found to be highly conserved (Table 2). These 10 conserved variants were further investigated to examine their deleterious effect on the stability of the protein.

**3.4. Screening of Potentially Deleterious VUS based on the Stability Analysis.** Changes in protein's stability upon single point mutation could interfere with its function. The stability of 10 evolutionary conserved BBS10 variants was analyzed by I Mutant, MuPro, PremPS, and Dynamut Web servers. The I Mutant 3.0 server predicts the stability of variants by computing the free energy value (DDG) at pH 7 and a temperature of 298 K. I Mutant 3.0 predicted seven variants to decrease the stability of the protein with a DDG score less than 0 (Table 3). The MuPro server, on the other hand, predicted nine variants to decrease the stability of the protein out of a total of 10 variants (Table 3). Structure-based stability analysis of variants was carried out using PremPS and Dyanmut Web servers by uploading the structural coordinates of BBS10. PremPS predicted all 10 variants to have a destabilizing effect on the protein structure with unfolding Gibbs free energy change ($\Delta\Delta G$ > 0 kcal/mol). The Dynamut server predicted eight variants to be potentially destabilizing


https://doi.org/10.1021/acsomega.2c04522
ACS Omega XXXX, XXX, XXX−XXX

**Table 1. BBS10 VUS Predicted to Be Deleterious by PROVEAN, SIFT, SNAP2, and Align-GVGD**

| | | PROVEAN | | SIFT | | SNAP2 | | AGVGD | | |
|---|---|---|---|---|---|---|---|---|---|---|
| sr. no. | variant | score | prediction | score | prediction | score | prediction | GV | GD | prediction |
| 1 | V11G | −3.02 | deleterious | 0.002 | damaging | 79 | effect | 0 | 108.79 | class C65 |
| 2 | P31L | −9.12 | deleterious | 0 | damaging | 74 | effect | 0 | 97.78 | class C65 |
| 3 | I68T | −3.04 | deleterious | 0.007 | damaging | 53 | effect | 0 | 89.28 | class C65 |
| 4 | S191L | −3.37 | deleterious | 0.01 | damaging | 56 | effect | 0 | 144.08 | class C65 |
| 5 | C195G | −3.6 | deleterious | 0.03 | damaging | 51 | effect | 0 | 158.23 | class C65 |
| 6 | S311F | −3.26 | deleterious | 0.001 | damaging | 58 | effect | 0 | 154.81 | class C65 |
| 7 | I342T | −3.78 | deleterious | 0.005 | damaging | 41 | effect | 0 | 89.28 | class C65 |
| 8 | C371S | −4.65 | deleterious | 0 | damaging | 41 | effect | 0 | 111.67 | class C65 |
| 9 | S378F | −3.54 | deleterious | 0.004 | damaging | 12 | effect | 0 | 154.81 | class C65 |
| 10 | A417E | −4.22 | deleterious | 0.002 | damaging | 83 | effect | 0 | 106.71 | class C65 |
| 11 | R422W | −5.38 | deleterious | 0.002 | damaging | 77 | effect | 0 | 101.29 | class C65 |
| 12 | Y613C | −2.56 | deleterious | 0.001 | damaging | 62 | effect | 0 | 193.72 | class C65 |
| 13 | C694R | −3.97 | deleterious | 0 | damaging | 73 | effect | 0 | 179.53 | class C65 |

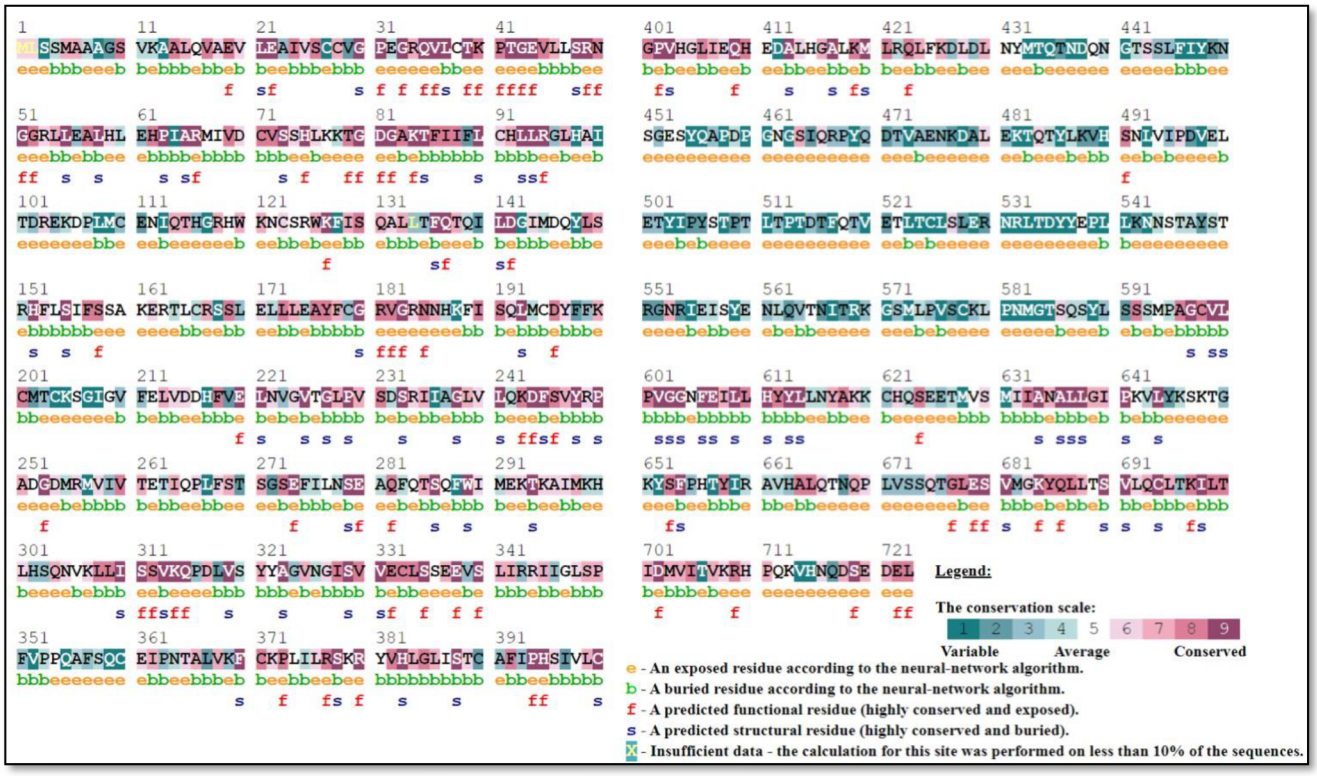



**Figure 2.** Evolutionarily conserved regions of BBS10 protein predicted by the ConSurf server.

based on the variation in Gibbs free energy ($\Delta\Delta G$ < 0.0 kcal/mol). Together, all four web servers predicted six variants to be potentially pathogenic based on their destabilizing effect on the protein's structure.

**3.5. Potentially Pathogenic Variants Affect the Conformational Stability of BBS10 Protein.** The stability of a protein structure could be determined by analyzing the deviations produced in its structure during the course of the simulation. The conformational stability of BBS10 wild-type and missense mutants was accessed by plotting the RMSD values throughout the 100-ns-long MD simulation. RMSD is used for measuring the difference between the backbones of a protein from its initial structural conformation to its final conformation. Smaller deviations in RMSD values indicate a more stable protein structure. Notable differences were observed between the RMSD profile of the wild-type and mutants (Figure 3). The RMSD of wild-type BBS10 got stabilized after 10 ns with slight deviations observed at 55 and 88 ns, whereas higher fluctuations were observed for mutants. Three mutants, Ser191Leu, Ile342Thr, and Tyr613Cys, showed higher RMSD than the wild-type (Figure 3a, c, and f). Higher fluctuations in the RMSD profile of mutants indicate lesser stability of BBS10 mutants' protein structure conformations.

**3.6. Potentially Pathogenic Variants Affect the Dynamic Behavior of BBS10 Protein.** The dynamic characteristic of native BBS10 and mutant protein structures were evaluated by calculating the residue-based root-mean-square fluctuations (RMSF). As observed in the RMSF profile, the mutations do not seem to affect the overall fluctuation of protein residues as compared to the wild-type protein (Figure 4). Interestingly, Ser191Leu, Ile342Thr, Cys371Ser, and


https://doi.org/10.1021/acsomega.2c04522
ACS Omega XXXX, XXX, XXX−XXX

**Table 2. Evolutionary Conservation Status of BBS10 Variants As Predicted by ConSurf (Conserved Variants Highlighted in Bold)**

| sr. no. | variant | ConSurf scale |
|---|---|---|
| 1 | V11G | 3 |
| 2 | **P31L** | **9** |
| 3 | I68T | 4 |
| 4 | **S191L** | **9** |
| 5 | **C195G** | **8** |
| 6 | **S311F** | **9** |
| 7 | **I342T** | **8** |
| 8 | **C371S** | **9** |
| 9 | **S378F** | **9** |
| 10 | **A417E** | **9** |
| 11 | R422W | 5 |
| 12 | **Y613C** | **9** |
| 13 | **C694R** | **9** |

Tyr613Cys showed high fluctuations in the amino acid region from 435 to 465 residues, suggesting the global effect of mutations on native protein conformation (Figure 4a, c, d, and f). The BBS10 435−465 residue region is made up of loops and contains charged amino acids at the protein surface. Loops are known to perform critical molecular functions and recognition of interactor molecules. An increase in the flexibility of the loop region may affect the physiological role of BBS10 protein. Therefore, from RMSF profile analysis, it could be concluded that potentially pathogenic variants affect the dynamics of BBS10 protein by particularly affecting the loop region.

**3.7. Potentially Pathogenic Variants Decrease the Overall Compactness of BBS10 Protein.** The radius of gyration (Rg) provides insights into the compactness of a protein structure throughout the simulation. The Rg value of the wild type showed a sharp change up to 40 ns with a steep dip observed at 28 ns; the Rg value later got stabilized at 3 nm. The change in Rg value was found to be significant for all of the variants (Figure 5). All of the variants showed a zigzag pattern of Rg values throughout the simulation, indicating mutation-induced alterations in structural compactness. The variants, Cys195Gly, Ile342Thr, and Tyr613Cys, showed much higher values of Rg as compared to the wild-type BBS10 (Figure 5b, c, and f). Changes in Rg values suggest the disruption of the hydrophobic core of mutant proteins leading to the partial unfolding transition and, therefore, a decrease in overall compactness. Surprisingly, the solvent accessible surface (SASA) values were found to be decreasing throughout the simulation, and no significant change was observed between the wild-type and mutant protein (Supplementary Figure S3).

The free energy landscape was plotted to obtain the minimum energy conformation ensemble of the structures. The Wild type BBS10 showed the highest number of minimum energy structures compared to all six variants indicating it to be more stable (Supplementary Figure S4). Together, the MD simulation results conclude that potentially pathogenic mutations alter the overall stability of BBS10, which could adversely affect the native protein function.

## 4. DISCUSSION

BBS is a rare genetic disorder which affects multiple body systems. BBS is characterized by progressive retinal dystrophy, postaxial polydactyly, obesity, cognitive impairment, and renal and urogenital anomalies as primary diagnostic features.[1,2] Therefore, BBS has also been considered a model disease to study the biology of the primary cilium. Diagnosis of BBS is often tricky during early age due to clinical heterogeneity and the common incomplete presentation of the disease. Therefore, genetic testing has become an important diagnostic method for disease confirmation, risk assessment, genetic counselling, and clinical management. *BBS10* is one of the most common hotspots for mutations, along with *BBS1* and *BBS2*. With the advancement of genetic sequencing, a large number of novel missense mutations are identified in *BBS10*; hence, it becomes crucial to characterize the identified mutations into pathogenic or benign categories. Clinically identified missense variants with family history can be classified as pathogenic, such as E679K mutation in BBS10.[18] However, the classification of a large number of missense variants with no obvious clinical features or family history remains a challenging task.

In the present study, we have (1) modeled the three-dimensional structure of BBS10 and (2) used a combination of

**Table 3. Evolutionary Conservation and Mutation-Induced Changes in Protein Stability of BBS10 VUS As Predicted by ConSurf, I-Mutant, MuPro, PremPS, and Dynamut Web Servers (Variants Highlighted in Red Chosen for Further Analysis)**

| | I-Mutant | | MuPro | | PremPS | | Dynamut | |
|---|---|---|---|---|---|---|---|---|
| Variants | DDG | Stability | DDG | Stability | ΔΔG (kcal/mol) | Effect | ΔΔG (kcal/mol) | Effect |
| P31L | -0.53 | Decrease | 0.225 | Increase | 0.19 | Destabilizing | 1.653 | Stabilizing |
| **S191L** | **-0.19** | **Decrease** | **-0.04** | **Decrease** | **0.69** | **Destabilizing** | **-1.61** | **Destabilizing** |
| **C195G** | **-1.4** | **Decrease** | **-1.75** | **Decrease** | **2.07** | **Destabilizing** | **-3.03** | **Destabilizing** |
| S311F | 0.1 | Increase | -0.264 | Decrease | 0.19 | Destabilizing | 2.065 | Stabilizing |
| **I342T** | **-2.27** | **Decrease** | **-2.416** | **Decrease** | **2.89** | **Destabilizing** | **-0.16** | **Destabilizing** |
| **C371S** | **-0.84** | **Decrease** | **-1.55** | **Decrease** | **2.22** | **Destabilizing** | **-1.15** | **Destabilizing** |
| S378F | 0.21 | Increase | -0.675 | Decrease | 0.4 | Destabilizing | -0.692 | Destabilizing |
| **A417E** | **-0.3** | **Decrease** | **-0.935** | **Decrease** | **2.03** | **Destabilizing** | **-0.78** | **Destabilizing** |
| **Y613C** | **-1** | **Decrease** | **-1.5** | **Decrease** | **2.43** | **Destabilizing** | **-1.18** | **Destabilizing** |
| C694R | 0.04 | Increase | -0.799 | Decrease | 1.72 | Destabilizing | -0.119 | Destabilizing |


https://doi.org/10.1021/acsomega.2c04522
ACS Omega XXXX, XXX, XXX−XXX

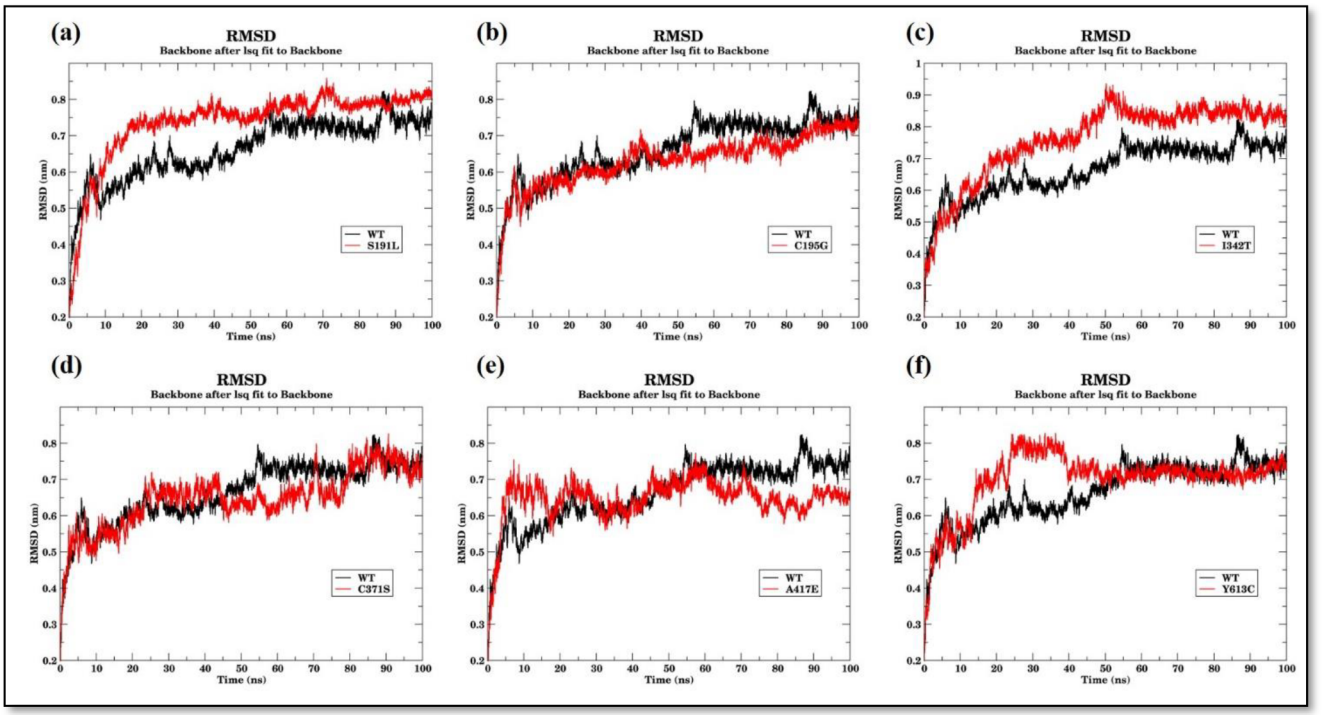



**Figure 3.** Comparative backbone root-mean-square deviation (RMSD) for wild-type BBS10 (WT) and potentially pathogenic missense variants. (a) Ser191Leu, (b) Cys19Gly, (c) Ile342Thr, (d) Cys371Ser, (e) Ala417Glu, (f) Tyr613Cys.

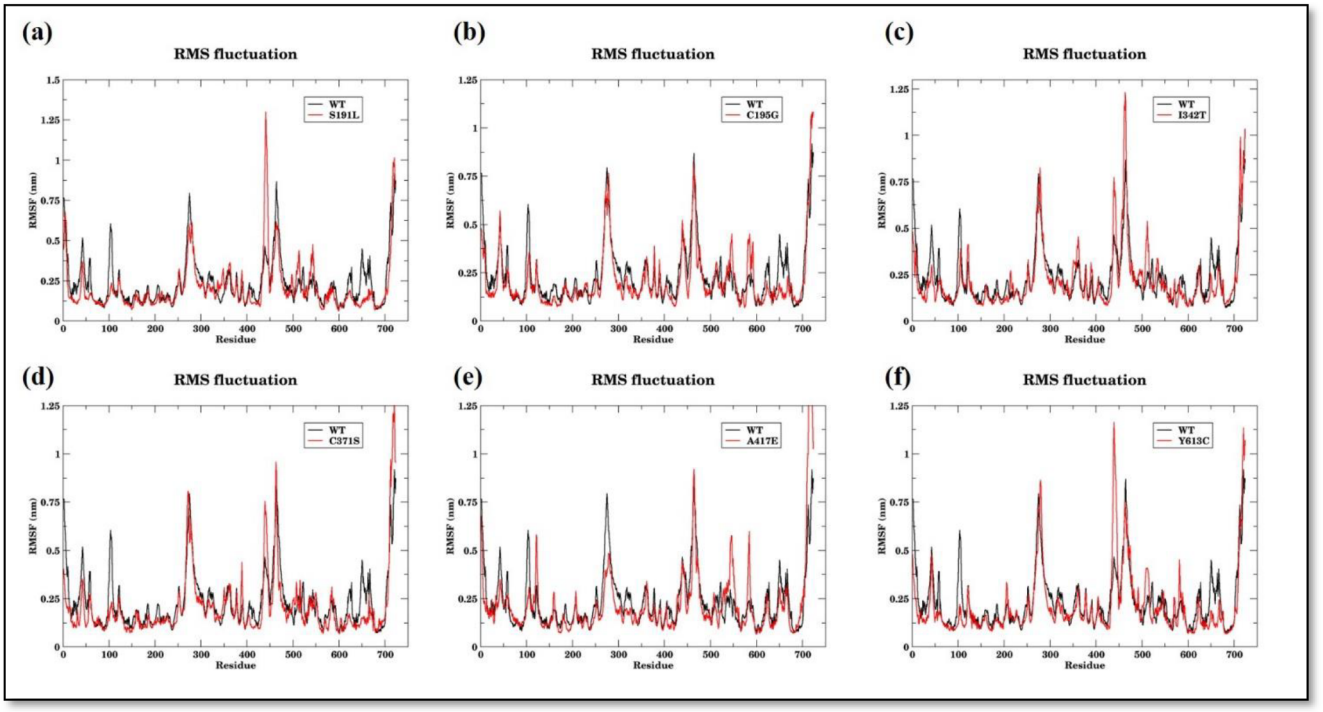



**Figure 4.** Comparative C$\alpha$-root-mean-square fluctuations (RMSF) for wild-type BBS10 (WT) and potentially pathogenic missense variants. (a) Ser191Leu, (b) Cys19Gly, (c) Ile342Thr, (d) Cys371Ser, (e) Ala417Glu, (f) Tyr613Cys.

*in silico* tools and a structure-based approach to identify pathogenic missense variants.

Extensive sequence analysis of *BBS10* and its homologues identified in vertebrates relate *BBS10* to the group II chaperonins.[10] Morpholino-based knockdown of BBS10 reveals developmental defects in zebrafish embryos, suggesting its role in early development. *BBS10* along with *BBS6* and *BBS12* are thought to function as a chaperonin-like protein and play a vital role in the assembly of the BBSome complex.[10,13] Recently, it has been shown that the inner medullary collecting duct (IMCD3) cells lacking BBS10 showed an increased expression of glycolytic enzymes, suggesting a process similar to the Warburg effect.[41] Analyzing, the three-dimensional structure of BBS10 could enhance our current understanding


https://doi.org/10.1021/acsomega.2c04522
ACS Omega XXXX, XXX, XXX−XXX

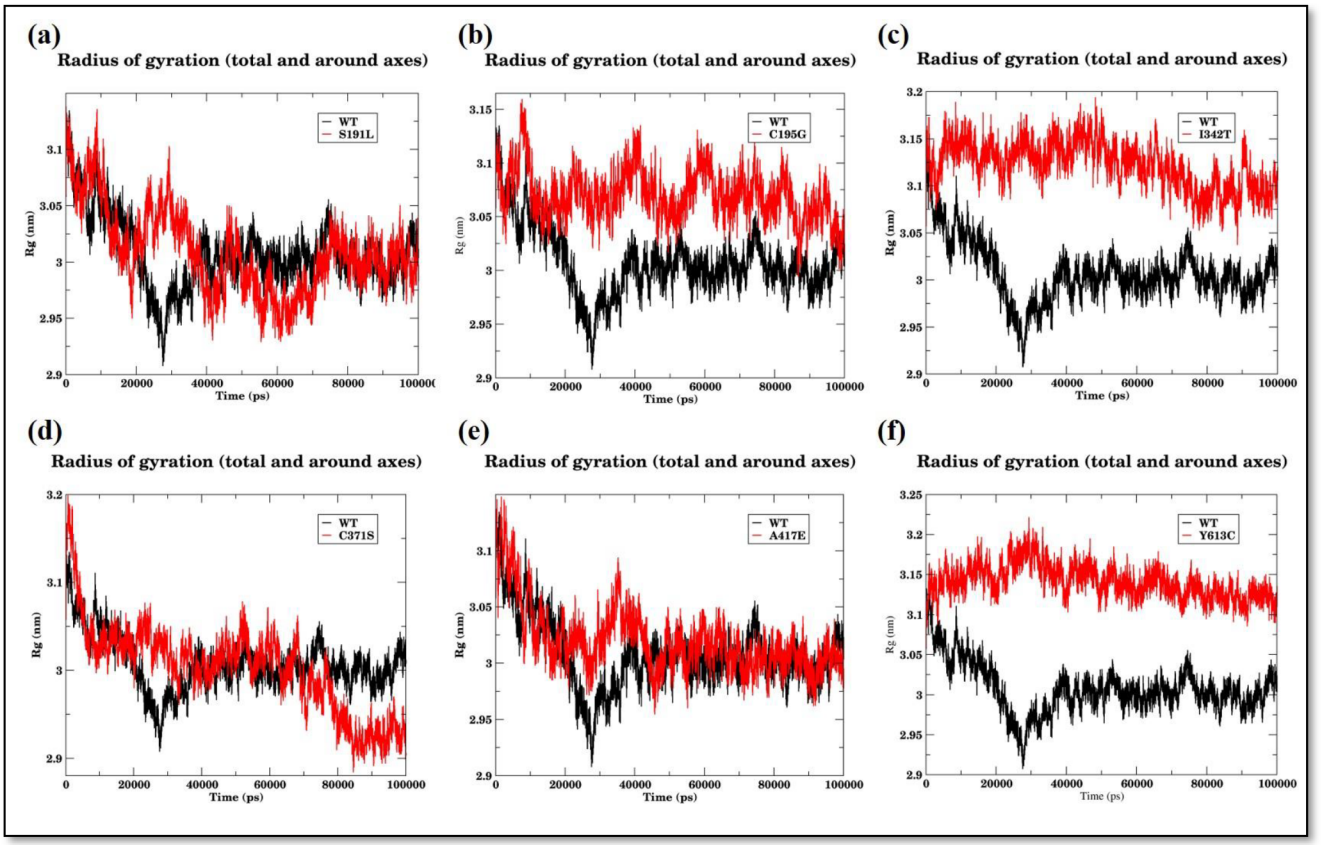



**Figure 5.** Comparative radius of gyration (Rg) profile for wild-type BBS10 (WT) and potentially pathogenic missense variants. (a) Ser191Leu, (b) Cys19Gly, (c) Ile342Thr, (d) Cys371Ser, (e) Ala417Glu, (f) Tyr613Cys.

of protein−protein interactions. Therefore, in the absence of the crystal structure of BBS10, we modeled the structure for the present study. Our results showed that stereochemical details of the BBS10 model used for the analysis were found to be better than the model structure predicted by the AlphaFold (Supplementary Figure 2). Nine publicly available web servers were utilized to determine the evolutionarily conserved pathogenic missense variants. Results obtained from *in silico* Web servers alone are not considered as independent evidence of pathogenicity; therefore, we applied the molecular dynamics approach to analyze whether detected mutations have an effect on protein structure and/or dynamics.

A total of 101 missense VUS reported in the BBS10 protein were analyzed. Based on our results, six variants are found to be potentially pathogenic, Ser191Leu, Cys19Gly, Ile342Thr, Cys371Ser, Ala417Glu, and Tyr613Cys. The modeled structure of BBS10 reveals that all of the identified pathogenic variants were found to be located in the $\alpha$-helices, except Cys371 located in the $\beta$-sheet, signifying the importance of these variants in the native protein structure and protein folding.[42] An additional observation we made upon inspecting our model is that all of the identified pathogenic mutations were located in the buried region of protein, suggesting the role of these residues in the stability of the hydrophobic core of the native protein. Interestingly, except Tyr613, all of the remaining five potentially pathogenic variants were found to be located in the well conserved 60 kDa domain, *Cpn60/TCP-1* identified in HSP60 chaperone family and the TCP-1 (T-complex protein) family.[43,44] Identification of pathogenic mutations in the conserved *Cpn60/TCP-1* domain of BBS10 suggests the probable loss of chaperonin like function, resulting in impaired folding or stability of ciliary or basal body proteins.

In conclusion, our study provides a comprehensive scientific methodology to identify potentially pathogenic variants associated with BBS10. Our finding could prove valuable for the genetic counselling and early detection of BBS, which is often difficult due to the heterogeneity of the disease. The functional relevance of the identified potentially pathogenic BBS 10 variants could be studied in vitro using CRISPR-Cas9 based genome editing. From a broader perspective, our study yields new insights into a structurally less understood family of BBS proteins. It should be noted that the crystal structure of BBS10 is not available yet, and in the absence of any functional studies, it is not possible to validate our analysis. Additional studies will be required to fully understand the molecular mechanism associated with potentially pathogenic variants and their clinical relevance.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c04522.

Ramachandran plot of modeled BBS10 structure showing 94.9% residues in the allowed region and 5.1% in the additional allowed region; quality assessment of BBS10 structure modeled using Robetta server and AlphaFold; solvent accessible surface area (SASA) profile for wild-type BBS10 (WT) and potentially pathogenic missense variants; free energy landscape of BBS10 WT and mutant to identify minimum energy



https://doi.org/10.1021/acsomega.2c04522
ACS Omega XXXX, XXX, XXX−XXX

structures; BBS10 VUS predicted to be deleterious by PROVEAN, SIFT, SNAP2, and Align-GVGD (PDF)

## AUTHOR INFORMATION


### Corresponding Author

**Neha Gupta** − *Unit of Nephrology, Department of Translational Medical Sciences, University of Campania, 80131 Naples, Italy; BioGem S.C.A.R.L., Contrada Camporeale, 83031 Ariano Irpino AV, Italy;* orcid.org/0000-0001-8181-7378; Email: neha.gupta@biogem.it, neha.gupta@unicampania.it

### Authors

**Mudassar Ali Khan** − *Advanced Centre for Treatment, Research and Education in Cancer, Navi Mumbai, Maharashtra 410210, India; Homi Bhabha National Institute, Training School Complex, Anushaktinagar, Mumbai 400094, India;* orcid.org/0000-0001-5734-198X

**Giovambattista Capasso** − *Unit of Nephrology, Department of Translational Medical Sciences, University of Campania, 80131 Naples, Italy; BioGem S.C.A.R.L., Contrada Camporeale, 83031 Ariano Irpino AV, Italy*

**Miriam Zacchia** − *Unit of Nephrology, Department of Translational Medical Sciences, University of Campania, 80131 Naples, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c04522

### Author Contributions

⊥These authors contributed equally and share the first authorship


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors would like to thank the Bioinformatics Centre (BIC) at ACTREC for providing the necessary computational facility.

## REFERENCES


(1) Forsythe, E.; Beales, P. L. Bardet-Biedl syndrome. *Eur. J. Hum. Genet.* **2013**, *21*, 8−13.

(2) Beales, P. L.; Elcioglu, N.; Woolf, A. S.; Parker, D.; Flinter, F. A. New criteria for improved diagnosis of Bardet-Biedl syndrome: Results of a population survey. *J. Med. Genet.* **1999**, *36*, 437−446.

(3) Klein, D.; Ammann, F. The syndrome of Laurence-Moon-Bardet-Biedl and allied diseases in Switzerland. Clinical, genetic and epidemiological studies. *J. Neurol. Sci.* **1969**, *9*, 479−513.

(4) Farag, T. I.; TEEBI, A. S. Bardet-Biedl and Laurence-Moon syndromes in a mixed Arab population. *Clin. Genet.* **1988**, *33*, 78−82.

(5) Farag, T. I.; TEEBI, A. S. High incidence of Bardet Biedl syndrome among the Bedouin. *Clin. Genet.* **1989**, *36*, 463−4.

(6) Moore, S. J.; Green, J. S.; Fan, Y.; Bhogal, A. K.; Dicks, E.; Fernandez, B. A.; Stefanelli, M.; Murphy, C.; Cramer, B. C.; Dean, J. C.S.; Beales, P. L.; Katsanis, N.; Bassett, A. S.; Davidson, W. S.; Parfrey, P. S. Clinical and genetic epidemiology of Bardet-Biedl syndrome in Newfoundland: A 22-year prospective, population-based, cohort study. *Am. J. Med. Genet.* **2005**, *132A*, 352−360.

(7) Marchese, E.; Ruoppolo, M.; Perna, A.; Capasso, G.; Zacchia, M. Exploring Key Challenges of Understanding the Pathogenesis of Kidney Disease in Bardet-Biedl Syndrome. *Kidney Int. Reports* **2020**, *5*, 1403−1415.

(8) Gupta, N.; D'Acierno, M.; Zona, E.; Capasso, G.; Zacchia, M. Bardet-Biedl syndrome: The pleiotropic role of the chaperonin-like BBS6, 10, and 12 proteins. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **2022**, *190*, 9.

(9) Álvarez-Satta, M.; Castro-Sánchez, S.; Pereiro, I.; Piñeiro-Gallego, T.; Baiget, M.; Ayuso, C.; Valverde, D. Overview of Bardet-Biedl syndrome in Spain: Identification of novel mutations in BBS1, BBS10 and BBS12 genes. *Clin. Genet.* **2014**, *86*, 601−602.

(10) Stoetzel, C.; Laurier, V.; Davis, E. E.; Muller, J.; Rix, S.; Badano, J. L.; Leitch, C. C.; Salem, N.; Chouery, E.; Corbani, S.; et al. BBS10 encodes a vertebrate-specific chaperonin-like protein and is a major BBS locus. *Nat. Genet.* **2006**, *38*, 521−524.

(11) Hjortshøj, T. D.; Grønskov, K.; Philp, A. R.; Nishimura, D. Y.; Riise, R.; Sheffield, V. C.; Rosenberg, T.; Brøndum-Nielsen, K. Bardet-Biedl syndrome in Denmark: Report of 13 novel sequence variations in six genes. *Hum. Mutat.* **2010**, *31*, 429−436.

(12) Álvarez-Satta, M.; Castro-Sánchez, S.; Valverde, D. Bardet-biedl syndrome as a chaperonopathy: Dissecting the major role of chaperonin-like BBS proteins (BBS6-BBS10-BBS12). *Front. Mol. Biosci.* **2017**, *4*, 1−7.

(13) Seo, S.; Baye, L. M.; Schulz, N. P.; Beck, J. S.; Zhang, Q.; Slusarski, D. C.; Sheffield, V. C. BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1488−1493.

(14) Esposito, G.; Testa, F.; Zacchia, M.; Crispo, A. A.; Di Iorio, V.; Capolongo, G.; Rinaldi, L.; D'Antonio, M.; Fioretti, T.; Iadicicco, P.; Rossi, S.; Franze, A.; Marciano, E.; Capasso, G.; Simonelli, F.; Salvatore, F. Genetic characterization of Italian patients with Bardet-Biedl syndrome and correlation to ocular, renal and audio-vestibular phenotype: Identification of eleven novel pathogenic sequence variants. *BMC Med. Genet.* **2017**, *18*, 1−12.

(15) Ohto, T.; Enokizono, T.; Tanaka, R.; Tanaka, M.; Suzuki, H.; Sakai, A.; Imagawa, K.; Fukushima, H.; Fukushima, T.; Sumazaki, R.; Uehara, T.; Takenouchi, T.; Kosaki, K. A novel BBS10 mutation identified in a patient with Bardet-Biedl syndrome with a violent emotional outbreak. *Hum. Genome Var.* **2017**, *4*, 1−3.

(16) Putoux, A.; Mougou-Zerelli, S.; Thomas, S.; Elkhartoufi, N.; Audollent, S.; Le Merrer, M.; Lachmeijer, A.; Sigaudy, S.; Buenerd, A.; Fernandez, C.; Delezoide, A.-L.; Gubler, M.-C.; Salomon, R.; Saad, A.; Cordier, M.-P.; Vekemans, M.; Bouvier, R.; Attie-Bitach, T. BBS10 mutations are common in 'Meckel'-type cystic kidneys. *J. Med. Genet.* **2010**, *47*, 848−852.

(17) Pasińska, M.; Dudarewicz, L. Prenatal and Postnatal Diagnostics of a Child with Bardet-Biedl Syndrome: Case Study. *J. Mol. Genet. Med.* **2015**, *09*, DOI: 10.4172/1747-0862.1000189

(18) Dehani, M.; Zare-Abdollahi, D.; Bushehri, A.; Dehghani, A.; Effati, J.; Miratashi, S. A. M.; Khorshid, H. R. K. Identification of a Novel Homozygous Mutation in BBS10 Gene in an Iranian Family with Bardet-Biedl Syndrome. *Avicenna J. Med. Biotechnol.* **2021**, *13*, 230−233.

(19) Wang, Q.; Mehmood, A.; Wang, H.; Xu, Q.; Xiong, Y.; Wei, D. Q. Computational screening and analysis of lung cancer related non-synonymous single nucleotide polymorphisms on the human kirsten rat sarcoma gene. *Molecules* **2019**, *24*, 1951.

(20) Ng, P. C.; Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **2003**, *31*, 3812−3814.

(21) Hecht, M.; Bromberg, Y.; Rost, B. Better prediction of functional effects for sequence variants. *BMC Genomics* **2015**, *16*, S1.

(22) Choi, Y.; Chan, A. P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745−2747.

(23) Tavtigian, S. V.; Deffenbaugh, A. M.; Yin, L.; Judkins, T.; Scholl, T.; Samollow, P. B.; de Silva, D.; Zharkikh, A.; Thomas, A.; et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J. Med. Genet.* **2005**, *43*, 295−305.

(24) Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: an improved methodology to

estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **2016**, *44*, W344−W350.
(25) 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* **2006**, *22*, 2729−2734.
(26) Cheng, J.; Randall, A.; Baldi, P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins Struct. Funct. Genet.* **2006**, *62*, 1125−1132.
(27) Chen, Y.; Lu, H.; Zhang, N.; Zhu, Z.; Wang, S.; Li, M. PremPS: Predicting the impact of missense mutations on protein stability. *PLoS Comput. Biol.* **2020**, *16*, e1008543.
(28) 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.* **2018**, *46*, W350−W355.
(29) Landrum, M. J.; Lee, J. M.; Benson, M.; Brown, G.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Hoover, J.; et al. ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **2016**, *44*, D862−D868.
(30) Choi, Y.; Sims, G. E.; Murphy, S.; Miller, J. R.; Chan, A. P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* **2012**, *7*, e46688.
(31) Bromberg, Y.; Rost, B. SNAP: Predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res.* **2007**, *35*, 3823−3835.
(32) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Mol. Biol.* **2002**, *9*, 646−652.
(33) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(34) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(35) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(36) Yang, J.; Anishchenko, I.; Park, H.; Peng, Z.; Ovchinnikov, S.; Baker, D. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 1496−1503.
(37) Watanabe, Y. S.; Fukunishi, Y.; Nakamura, H. Modeling of Loops in Protein Structures. *Seibutsu Butsuri* **2004**, *44*, S41.
(38) Sippl, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins Struct. Funct. Bioinforma.* **1993**, *17*, 355−362.
(39) Wiederstein, M.; Sippl, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407−W410.
(40) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.
(41) Marchese, E.; Caterino, M.; Fedele, R.; Pirozzi, F.; Cevenini, A.; Gupta, N.; Ingrosso, D.; Perna, A.; Capasso, G.; Ruoppolo, M.; Zacchia, M. Multi-Omics Studies Unveil Extraciliary Functions of BBS10 and Show Metabolic Aberrations Underlying Renal Disease in Bardet - Biedl Syndrome. *Int. J. Mol. Sci.* **2022**, *23*, 9420.
(42) Ji, Y. Y.; Li, Y. Q. The role of secondary structure in protein structure selection. *Eur. Phys. J. E* **2010**, *32*, 103−107.
(43) Kubota, H.; Hynes, G.; Willison, K. The Chaperonin Containing t-complex polypeptide 1 (TCP-1): Multisubunit Machinery Assisting in Protein Folding and Assembly in the Eukaryotic Cytosol. *Eur. J. Biochem.* **1995**, *230*, 3−16.
(44) Gupta, R. S. Evolution of the chaperonin families (HSP60, HSP 10 and TCP-1) of proteins and the origin of eukaryotic cells. *Mol. Microbiol.* **1995**, *15*, 1−11.


https://doi.org/10.1021/acsomega.2c04522
*ACS Omega* XXXX, XXX, XXX−XXX