Received: 1 July 2022 | Revised: 5 December 2022 | Accepted: 3 January 2023

DOI: 10.1002/prot.26463

RESEARCH ARTICLE

PROTEINS WILEY

Effects of the Jokela type of spinal muscular atrophy-related G66V mutation on the structural ensemble characteristics of CHCHD10

Hakan Alici[1] | Vladimir N. Uversky[2] | David E. Kang[3,4] | Junga Alexa Woo[3] | Orkid Coskuner-Weber[5]

[1]Department of Physics, Faculty of Arts and Sciences, Zonguldak Bulent Ecevit University, Zonguldak, Turkey

[2]Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida, USA

[3]Department of Pathology, School of Medicine, Case Western Reserve University, Cleveland, Ohio, USA

[4]Louis Stokes Cleveland VA Medical Center, Cleveland, Ohio, USA

[5]Molecular Biotechnology, Turkish-German University, Istanbul, Turkey

Correspondence
Hakan Alici, Department of Physics, Faculty of Arts and Sciences, Zonguldak Bulent Ecevit University, Zonguldak 67100, Turkey.
Email: hakanalici@beun.edu.tr

Orkid Coskuner-Weber, Molecular Biotechnology, Turkish-German University, Sahinkaya Caddesi, No. 106, Beykoz, Istanbul 34820, Turkey.
Email: weber@tau.edu.tr

Abstract

The G66V pathological variant of the coiled-coil-helix-coiled-coil-helix domain-containing protein 10 (CHCHD10), mitochondrial, plays a role in Jokela type spinal muscular atrophy. The wild-type and G66V mutant-type CHCHD10 proteins contain intrinsically disordered regions, and therefore, their structural ensemble studies have been experiencing difficulties using conventional tools. Here, we show our results regarding the first characterization of the structural ensemble characteristics of the G66V mutant form of CHCHD10 and the first comparison of these characteristics with the structural ensemble properties of wild-type CHCHD10. We find that the structural properties, potential of mean force surfaces, and principal component analysis show stark differences between these two proteins. These results are important for a better pathology, biochemistry and structural biology understanding of CHCHD10 and its G66V genetic variant and it is likely that these reported structural properties are important for designing more efficient treatments for the Jokela type of spinal muscular atrophy disease.

KEYWORDS

bioinformatics, CHCHD10, G66V, genetics, Jokela type spinal muscular atrophy, multiple run molecular dynamics simulations, structural properties

1 | INTRODUCTION

CHCHD10 is encoded by the *CHCHD10* gene and is enriched at cristae junctions in the intermembrane space region of mitochondria. This protein that has intrinsically disordered protein regions (IDPRs) is detected at increased levels in skeletal muscles, brain and heart.[1–3] Recently, the genetically mutated forms of CHCHD10 have been detected in patients suffering from different neurodegenerative diseases.[2] Even though clinically, it has been proven as a fact that CHCHD10 genetic mutations are linked to a series of neurodegenerative diseases, the physiological role of these mutant-type CHCHD10 proteins is poorly understood. The current literature does not involve the information on the structural ensemble characteristics of the G66V CHCHD10 mutant-type protein and on how G66V mutation impacts the structural ensemble characteristics of wild-type CHCHD10. Recently, we reported details of the effects of the S59L genetic mutant, which is at the center of amyotrophic lateral sclerosis-frontotemporal dementia.[1] In fact, we anticipated a higher aggregation propensity for the S59L mutant because of the increased number of the amino acid residues adopting β-sheet structure, which agrees with experiments.[1] Currently, existing studies focus on the definition of the CHCHD10 impacts in mitochondria by investigating the influences of genetic mutants in vivo (cell culture).[4] Such investigations showed that CHCHD10 impacts the stability of mitochondrial DNA, COX activity, mitochondrial cristae stability, mitochondrial fusion/fission, mitochondrial networks formation, as well as

apoptosis.[5–8] The phenotype of the G66V mutation in the Jokela type spinal muscular atrophy (SMAJ) patients, while being dominant and early onset, is mild and slow progressing.[4] SMAJ G66V carriers exhibit motor neuron abnormalities with little to no mitochondrial myopathy.[4]

CHCHD10 possesses intrinsically disordered regions.[1] Therefore, experiments applying conventional techniques face challenges in characterizing the dynamic structural ensembles properties of the G66V mutant-type of CHCHD10 in solution.[9] In parallel, the impact of G66V mutation on CHCHD10's internal deregulation tendency is poorly understood. The SMAJ-related G66V mutation of the CHCHD10 protein was first detected in 55 patients in 17 Finnish families.[4] One study showed that CHCHD10 is a regulator of mitochondrial respiration and can play a role in the transcriptional repression of the oxygen-responsive element-containing genes, and some genetic CHCHD10 variants (including G66V) are impaired in these activities.[10] As a result, CHCHD10 is recognized as an important Jokela type of spinal muscular atrophy-causative gene.[2] We should mention here that we provided the structural ensemble knowledge on wild-type CHCHD10 for the first time in the literature,[1] but our studies are based on computational investigations and experiments including nuclear magnetic resonance and mass spectrometry measurements are needed for determining its structural ensemble properties in solution.

Following our recent studies on CHCHD10,[1] we used bioinformatics and homology modeling linked to multiple run molecular dynamics (MD) simulations to investigate the structural ensemble properties of the G66V CHCHD10 mutant-type protein and the impacts of G66V mutation on the structural ensemble properties of wild-type CHCHD10 in water. Presented findings may be important for gaining more detailed insights into the structural ensemble properties of Jokela type spinal muscular atrophy-related G66V mutant CHCHD10 and may be useful for treatment design studies.

2 | MATERIALS AND METHODS

To be consistent with our most recent studies on CHCHD10 and its pathological variants, we conducted multiple run MD simulations of wild-type and G66V mutant-type CHCHD10 in aqueous solution utilizing the same simulation protocols and methods as explained in Ref. 1. The bioinformatics and analyses protocols and methodologies were reported in Ref. 1 as well. Specifically, we used the Iterative Threading ASSembly Refinement (I-TASSER) homology modeling method[11] for predicting the initial configuration of CHCHD10. I-TASSER was ranked as the best server for protein structure prediction in community-wide CASP7, CASP8, CASP9, CASP10, CASP11, CASP12, CASP13, and CASP14 experiments. However, we should note that the initial structure determination affects the simulation outcome. Currently, we are validating various homology modeling tools linked to multiple run MD simulations on IDPRs. The G66V mutant-type CHCHD10 initial configuration was produced by replacing the Gly66 residue with Val. The Gromacs 5.1.4 software package was utilized,[12] and the temperature was set to 310K with a time-step of 2 fs. The TIP3P parameters[13] and CHARMM36m[14] force field parameters were chosen for water and the proteins because these parameters yield findings in accordance with experiments for IDPs.[15] Each protein was solvated by 10 948 water molecules and periodic boundary conditions were applied.[16] For defining the long-range interactions, the Particle Mesh Ewald method was utilized.[17] For achieving charge neutrality, one Cl^- ion was used. The energies of solvated protein systems were minimized by utilizing the steepest descent method. NVT and NPT ensembles were used for two-stage equilibrations of 2 ns, respectively.[16] Triple independent multiple run MD simulations were conducted after equilibration for a total of 3.0 μs for each solvated protein. In these simulations, varying velocity distributions were used. Nose–Hover thermostat was used[18] for keeping the temperature constant while Parrinello–Rahman barostat was utilized for keeping the pressure constant[19] along with the application of the LINCS algorithm.[20] We saved the production run trajectories every 10 ps. The free energy change surface area was calculated based on end-to-end distance and radius of gyration values utilizing the following equation:

$$\Delta G = -k_B T[lnP - lnP_{max.}]$$

where T is the temperature and k_B is the Boltzmann constant. Following our recent studies, P represents the probability distribution of the reaction coordinates (R) and P_{max} represents the maximum probability distribution of the molecular system selected to yield a value of $\Delta G = 0$.[1] The selection of R depends on the system properties under examination and R is defined usually by Rg, RMSD, the number of contacts and/or the number of hydrogen bonds.[21–23] We used a cut-off distance of 3.5 Å and an angle cut-off of 30 between the acceptor and donor atoms for predicting the intra-molecular hydrogen bonds. We defined a salt bridge using a cut-off distance of 7.0 Å for oppositely charged residues' centers of mass distance.[1] Furthermore, intramolecular interactions were determined using a distance between two residues centers of mass of 12.0 Å.[1] Principal component analysis was determined for investigating the major motion dynamics in wild-type and G66V mutant-type CHCHD10 in solution. The first two principal components that were obtained from covariance matrix diagonalization of atomic coordinates were utilized.[1]

3 | RESULTS AND DISCUSSION

Wild-type CHCHD10 and G66V mutant-type CHCHD10 proteins' average root mean square deviation (RMSD) values and average α-helix contents in aqueous solution are shown in the Supporting Information section. Following our earlier studies, these values were used for estimating the convergence of the simulations. In agreement with our previous studies,[1] the initial 200 ns of the simulation time is necessary for reaching convergency and conformations belonging to the last 800 ns—after convergence—are used in this study for interpreting the structural and thermodynamic characteristics. The obtained conformations from multiple run MD simulations of the

G66V mutant-type of CHCHD10 and those of wild-type CHCHD10 are depicted in Figure 1A. Figure 1B illustrates the average root mean square fluctuation (RMSF) values for each amino acid residue for the G66V mutant-type CHCHD10 and their comparison to those values of wild-type CHCHD10. Here, we note substantial differences especially in the N-terminal region of CHCHD10 (residues Pro16-Ser30). A comparison to our recent study involving the effect of S59L mutation on wild-type CHCHD10 reveals that S59L mutation impacts only the RMSF values of N-terminus and C-terminus[1] but G66V mutation has an impact on a larger region in the N-terminal region. Figure 1C depicts the difference in RMSF values between the G66V and S59L mutant forms of CHCHD10.[1] Stark differences occur in the N-terminal and C-terminal as well as mid-domain regions. We also show the average end-to-end distance and radius of gyration values with time and RMSF values from each multiple run MD simulation in the Supporting Information section.

3.1 | Secondary structure properties

Figure 2A shows the mean secondary structure element abundances per amino acid residue of G66V mutant-type CHCHD10 and its comparison to those of wild-type CHCHD10 in aqueous solution. Based on these findings, residues Ser5, Met65, Ser67, Thr70, Gln98, and Met99 form abundant β-sheet structure (probability > 10%) in the conformations of G66V mutant-type CHCHD10. On the other hand,

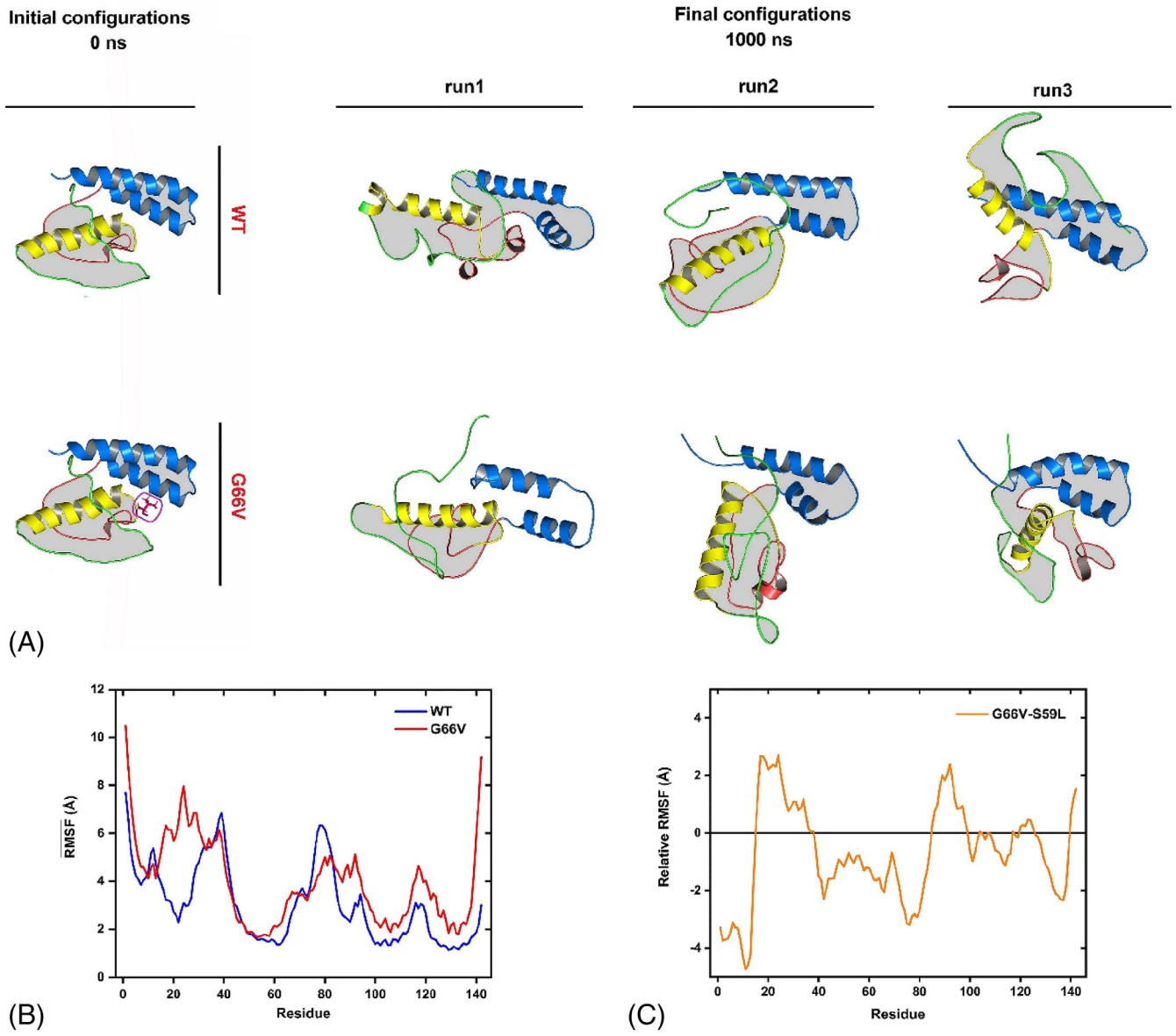

FIGURE 1 (A) Representative structures of wild-type and G66V mutant-type CHCHD10 in water obtained from homology modeling and multiple run MD simulations. (B) The mean RMSF values for wild-type and G66V mutant-type CHCHD10. (C) The difference in RMSF values between the G66V and S59L mutant-type CHCHD10. MD, molecular dynamics; RMSF, root mean square fluctuation.

the wild-type CHCHD10 protein possesses with high probability β-sheet structure formation at Ala21, Met65, Gly75, Gly76, Thr90, and Ala93 (probability > 10%). Even though the residues that adopt β-sheet structure differ from each other in these two IDPRs, the oligomerization may be affected by its location but not in its kinetics due to the same number of residues adopting β-sheet structure with high probability. This is an observation that needs to be validated by experiments. A comparison to S59L mutation effect on CHCHD10 reveals that S59L mutant is expected to be more susceptible toward oligomerization than the wild-type and G66V mutant-type CHCHD10 proteins due to increased number of residues adopting abundant β-sheet conformation.[1] Prominent α-helical structure formation (abundance > 10%) is detected in four regions, namely, at Pro42-His63, Ser81-Val85, Cys102-Thr115, and Leu119-Ser140 while wild-type CHCHD10 has abundant α-helix formation in three regions located at Pro42-His63, Cys102-Thr115, and Leu119-Ser140. This comparison illustrates that an additional region in the mid-domain region (Ser81-Val85) forms α-helical structure upon G66V mutation. 3_{10}-helix formation (probability > 10%) is noted at Pro23-Ala25 and Gln82-Val85 in the structural ensemble of mutant G66V CHCHD10 while Pro23-Ala25, Gln82-Val85 and Thr90-Ala92 form 3_{10}-helix in the structures of wild-type CHCHD10. This comparison indicates that 3_{10}-helix formation is diminished upon G66V mutation. Finally, abundant turn structure formation (probability > 10%) is detected in five regions located at Ser5-Gln41, Met65-Pro96, Met99-Ala103, Ser113-Leu121, and Gly137-Ser140 in the structures of G66V mutant CHCHD10 protein. Wild-type CHCHD10 possesses turn structure adaptation at Pro2-Ala33, Met65-Pro96, Gly100, Pro101, and Thr113-Leu119 and the residues that form turn structure are altered by the G66V mutation of wild-type CHCHD10. Figure 2B shows the formed average differences in secondary structure element abundances between the G66V and S59L mutant forms of CHCHD10. Based on these results, The β-sheet content is reduced in the structures of G66V mutant form in comparison to S59L mutant of CHCHD10 (especially in the N-terminal and mid-domain regions of CHCHD10). Therefore, we expect a higher likelihood toward oligomerization for the mutant S59L CHCHD10 but not for the mutant G66V CHCHD10 protein in solution.[1] The secondary structure element probabilities obtained from each multiple run MD simulation are depicted in the Supporting Information section (Figure S5).

3.2 | Conformational-free energy surface areas

As reported recently,[1] the mean end-to-end distance (R_{EE}) value of wild-type CHCHD10 in aqueous medium is 12.96 ± 3.32 Å, but the R_{EE} value is 17.74 ± 2.62 Å upon G66V mutation. This result demonstrates a significant increase (by 36.88%) of the value in the G66V mutant-type CHCHD10 protein when we do not consider standard deviation values. In fact, we detected such an increase of this value also for S59L mutant-type CHCHD10.[1] In fact, the average of this value for S59L mutant-type CHCHD10 is 18.73 ± 5.78 Å. However, the compactness of the protein by means of the radius of gyration (R_g) value is not affected significantly especially when we consider standard deviation values. Specifically, the G66V mutant-type CHCHD10 has a mean R_g value of 15.51 ± 0.28 Å and wild-type CHCHD10 has a mean R_g value of 15.86 ± 0.26 Å. When we do not consider the standard deviation, we note that G66V mutant CHCHD10 is only 2.25% more compact than wild-type CHCHD10 in aqueous solution. This is an ignorable deviation. In fact, we note that we recently found a mean

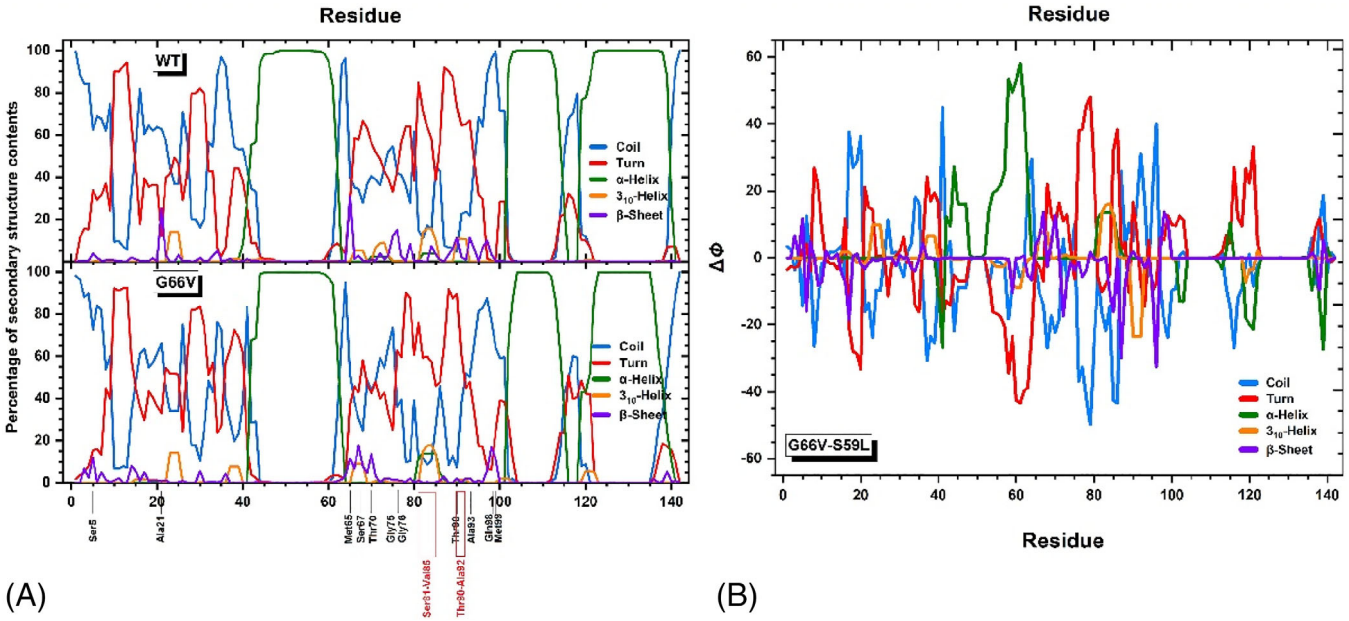

FIGURE 2 (A) The mean secondary structure element abundances per amino acid for wild-type and G66V mutant-type CHCHD10 in an aqueous medium. Coil (blue), turn (red), α-helix (green), 3_{10}-helix (orange), and β-sheet (purple). (B) The differences between the mean secondary structure element abundances of the G66V and S59L mutant forms of CHCHD10.

R_g value for S59L mutant form of 15.70 ± 0.26 Å. This value did not deviate much from the one reported for wild-type CHCHD10 either. Overall, these findings indicate that the compactness of CHCHD10 is not altered starkly by S59L and G66V genetic mutations but the dynamics is.[1]

The free energy change surface area[22] of the G66V mutant-type CHCHD10 protein in water by means of average R_g and mean R_{EE} values and its comparison to wild-type CHCHD10 is demonstrated in Figure 3A. Wild-type CHCHD10 possesses the most preferred structures at mean R_g values ranging from 15.5 to 16.0 Å and R_{EE} values ranging from 10.2 to 12.3 Å. Despite, G66V mutant-type shows favorable structures at mean R_g values ranging from 14.7 to 16.3 Å but the mean R_{EE} values range from 15.1 to 28.3 Å. Therefore, we can state that the G66V mutation does not alter the compactness of wild-type CHCHD10 significantly, however, end-to-end distances values range from 12.1 to 29.3 Å on the free energy surface area. A similar trend was found for the S59L mutant-type CHCHD10. Again, these results further indicate that the structural ensemble

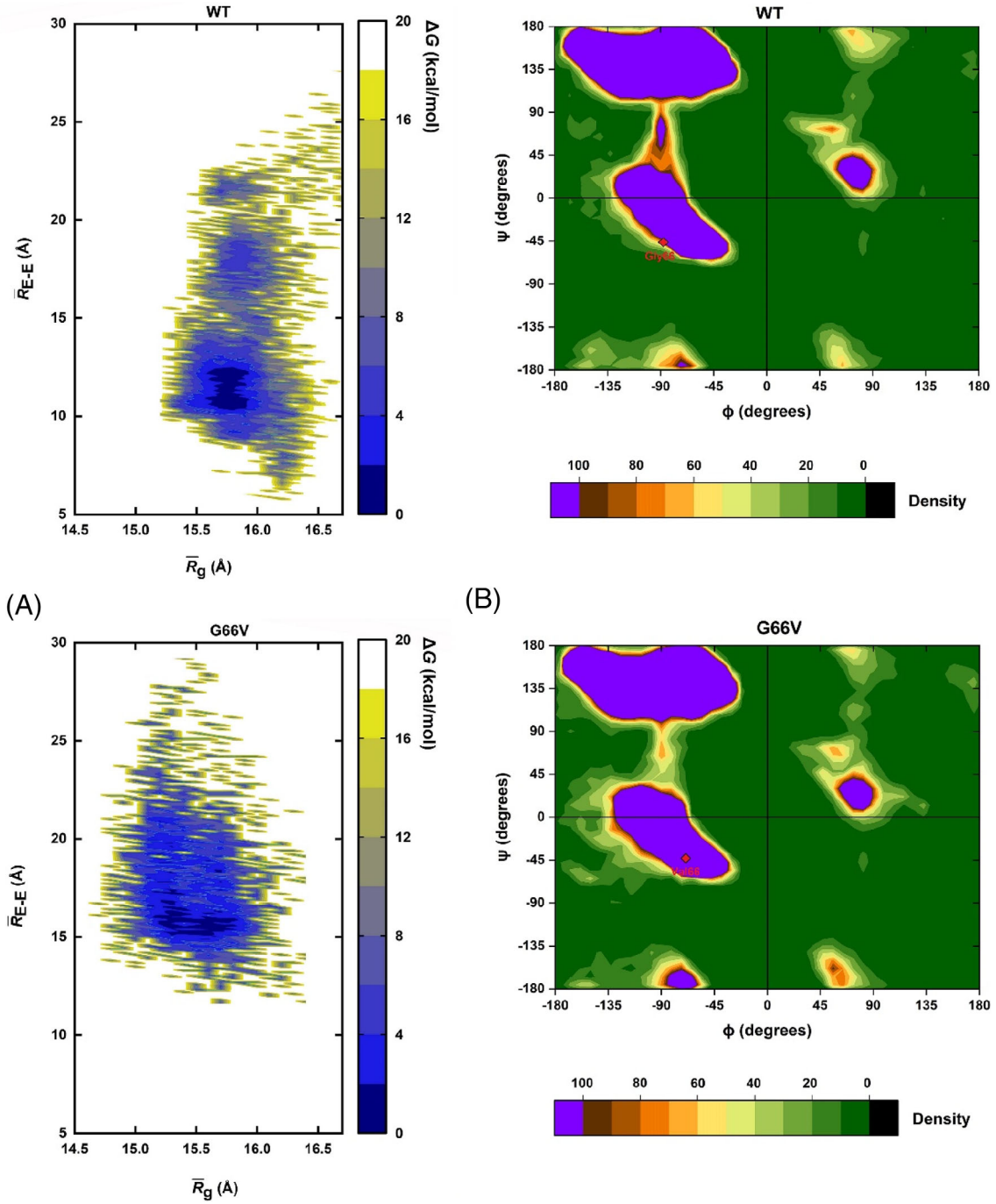

FIGURE 3 (A) Free energy surface areas of wild-type and G66V mutant-type CHCHD10 based on average R_g and R_{EE} values. (B) Ramachandran density plots demonstrating the specific torsional angles for wild-type and G66V mutant-type CHCHD10.

characteristics of G66V mutant-type CHCHD10 are different than that of its wild-type form. The obtained free energy surface areas from each multiple run MD simulation are illustrated in the Supporting Information section (Figure S6).

Figure 3B further demonstrates the structural property impacts of G66V mutation on neighboring amino acid residues by means of Ramachandran plots. For understanding the effect of G66V mutation on the neighboring residues in detail, we also computed the Ramachandran plots for residues located between Met45 and Pro80 (see Supporting Information section). It is noted that the conformations of Gly66 in wild-type CHCHD10 are scattered on very different dominant clusters, while Val66 residue of G66V mutant-type CHCHD10 possesses only two different main clusters. This decrease in the number of dominant clusters may not be an indicator to explain the effect of the genetic point mutation on the 66th amino acid residue. In fact, this trend may be foreseen for Gly because Gly possesses one hydrogen atom as a side chain, which leads to less steric hindrance compared to the side chains of other residues. Gly residue, therefore, has a higher flexibility and thus the polypeptide backbone makes turns that are not possible with other residues. When the dihedral angle distributions of the adjacent residues (Met45-Met65 residue region) are evaluated, the residues in Met45-Met65 possess a quasi-conformational overlap and these are significantly clustered on a dominant region representing α-helix. On the other hand, for His63, Val64, and Met65 residues adjacent to the point mutation, angular conformational clusters of wild-type CHCHD10 are found in dominant and mostly β-sheet domains while these amino acids in the G66V mutant have two dominant clusters, one found in the β-sheet areas and the additional one detected in left-hand α-helix for Val64 and α-helix regions for His63 and Met65. Another adjacent residue, Ser67, demonstrates that wild-type CHCHD10 exhibits two dominant clusters located in the α-helix and β-sheet regions, while the G66V mutant-type is detected to be located in α-helix, β-sheet, and left-hand α-helix clusters. Results indicate that the G66V genetic mutation could have a stark influence on adjacent His63, Val64, Met65, and Ser67 residues' angular conformations. When the neighbor residues at Ser67-Pro80 are investigated, it is noted that the abundant configurations of the wild-and G66V mutant-type are scattered in at least two dominant regions mostly placed in the α-helix and β-sheet areas and these relevant dominant clusters have conformational overlaps. Figure S10 in the Supporting Information section depicts the Ramachandran plots for the mutation site and neighboring residues.

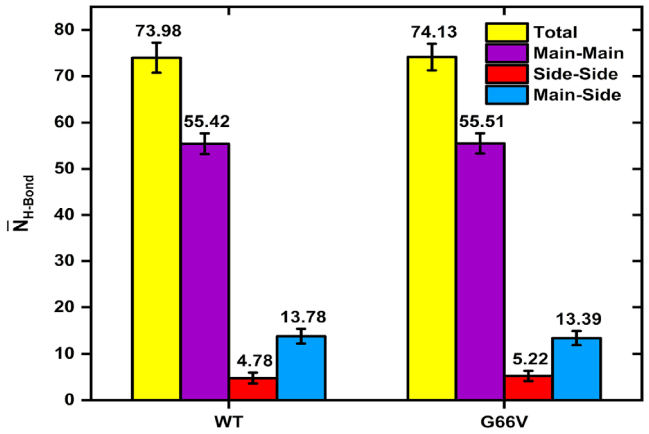

FIGURE 4 The average hydrogen bond numbers formed in wild-type and G66V mutant-type CHCHD10

3.3 | Hydrogen bond numbers

The hydrogen bond numbers of G66V mutant and wild-type CHCHD10 proteins in water are demonstrated in Figure 4. The side–side, main-–side, and main–main hydrogen bond formations are not affected by the G66V mutation. While we detected recently that these were somewhat reduced upon S59L mutation. salt bridges are formed between the residues Glu79 and Arg107, Glu105 and Arg3 or Arg6 when compared to its wild-type CHCHD10 form upon G66V mutation.

Using the SPSS version 23 software, we conducted independent samples *t* tests to assess whether the differences between the reported hydrogen bond number differences (Figure 4) for the wild-type and mutant G66V CHCHD10 possess meaningful significance (Table 1). Results indicate that the difference between the

TABLE 1 Independent samples *t* test of hydrogen bond numbers for the wild-type and G66V mutant CHCHD10; *t* values are the computed test statistics and here, df, the degrees of freedom, is 160 000. Sig (two-tailed) is the *p* value corresponding to the given test statistic and degrees of freedom. The Sig. (two-tailed) value indicates whether there is a difference between the groups. If this value is less than .05, it is decided that there is a difference between the groups. If this value is greater than .05, it is decided that there is no significant difference between the groups compared.

			Mean	Std. error	95% Confidence interval of the difference	
Type	t	Sig. (two-tailed)	Difference	Difference	Lower	Upper
Total	1.138	0.255	0.0785	0.0690	−0.0567	0.2136
Side–side	16.972	0.000	0.4366	0.0257	0.3861	0.4870
Main–side	−11.677	0.000	−0.4024	0.0345	−0.4700	−0.3348
Main–main	0.878	0.380	0.0044	0.0494	−0.0535	.01402

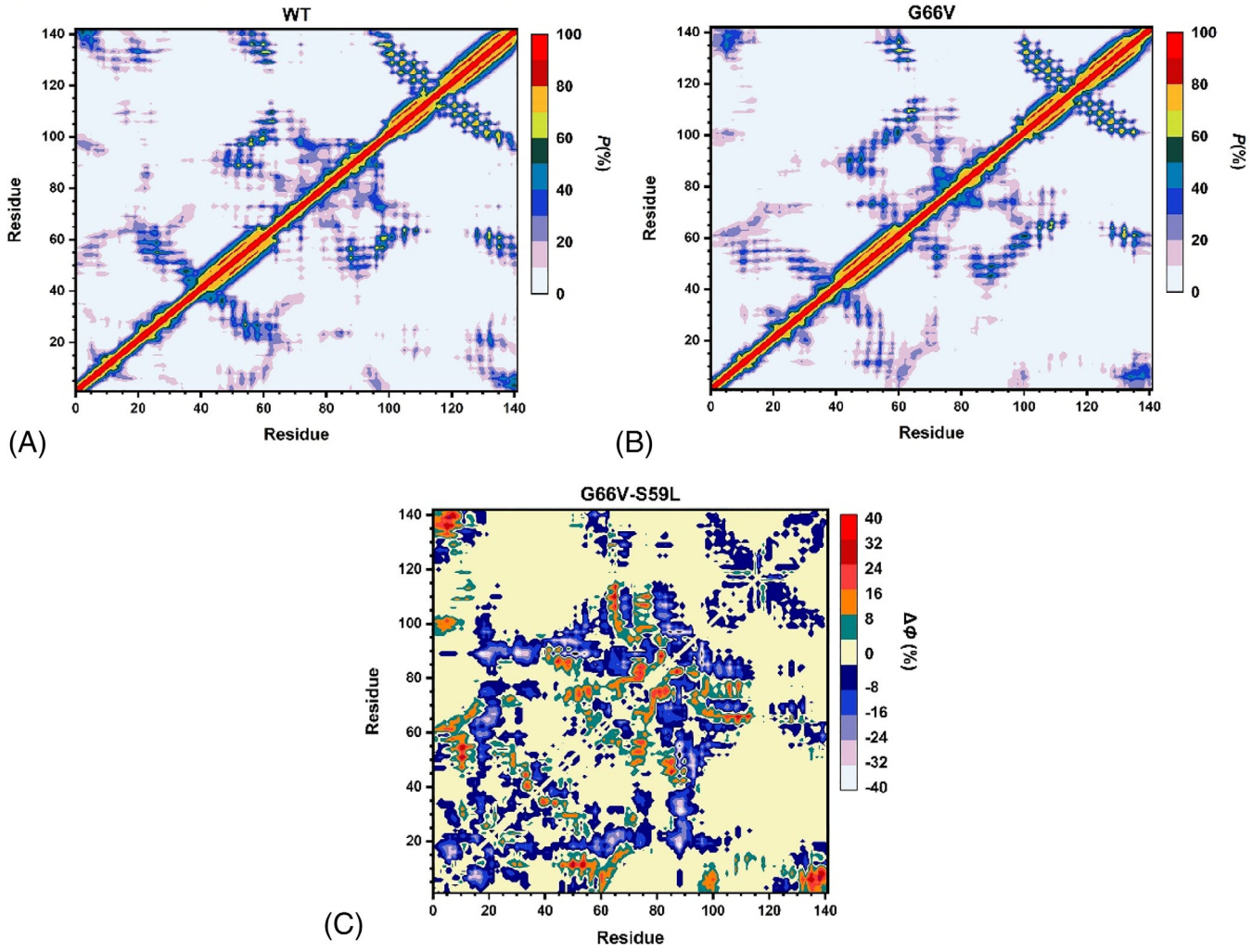

FIGURE 5 (A) The contact maps of wild-type and G66V mutant-type CHCHD10. (B) The difference between the tertiary structure abundances between the G66V and S59L mutant-type CHCHD10. The probability is defined as in Ref. 1.

average numbers of total hydrogen bond formations for the wild-type and G66V mutant CHCHD10 ($t = 1.138$, $p = .255$) and the average main–main hydrogen bond formations for the wild-type and G66V mutant CHCHD10 ($t = 0.878$, $p = .380$) are not statistically significant. Despite, the difference between the average side–side hydrogen bond formations for the wild-type and G66V mutant CHCHD10 ($t = 16.972$, $p < .001$) and the average main–side hydrogen bond formations for the wild-type and G66V mutant CHCHD10 ($t = -11.677$, $p < .001$) possess statistical significance. Thus, we may conclude that G66V mutant-type CHCHD10 (mean = 5.2244; standard deviation = 1.1117) shows more side–side hydrogen bond formations than the wild-type CHCHD10 protein (mean = 4.7819; standard deviation = 1.1881) whereas G66V mutant-type CHCHD10 has fewer main–side hydrogen bond formations (mean = 13.3889; standard deviation = 1.5192) than wild-type CHCHD10 (mean = 13.7913; standard deviation = 1.5630).

3.4 | Tertiary structure properties

Tertiary structure properties among with their probabilities for the wild-type and G66V mutant CHCHD10 are illustrated in Figure 5A,B. We notice differences in the intra-molecular contact maps upon G66V mutation. Specifically, fewer probable interactions are detected within N-terminus, N-terminal, and mid-domain regions, between the N-terminal and C-terminal regions of CHCHD10 upon G66V genetic mutation. Also, the C-terminus and mid-domain interactions disappear upon G66V mutation. Parallel to our R_{EE} and free energy surface area findings, G66V mutant-type is expected to be more flexible than its wild-type form. Such a general trend was also obtained for S59L mutant-type CHCHD10 in water. Figure 8B presents the differences in mean tertiary structure property abundances between G66V and S59L mutant-type CHCHD10. Significant differences between the tertiary structure property and abundances of G66V and S59L mutant-type CHCHD10 are detected in the mid-domain region and

C-terminus as demonstrated in Figure 5C.[1] Figure S9 in the Supporting Information section illustrates the contact maps obtained from each multiple run MD simulation.

The formed salt bridges in the conformations of G66V mutant and a comparison to those formed in wild-type CHCHD10 are presented in Table 2. We note that the G66V mutant-type possesses three additional salt bridges but with a rather longer distance.

TABLE 2 The formed average salt bridge distances including their standard deviations between specific amino acid residues of wild-type and G66V mutant-type CHCHD10

Residue	Residue	Wild-type (Å)	G66V mutant-type (Å)
Asp111	Arg107	6.48 (1.62)	6.50 (1.73)
Glu79	Arg107	–	14.46 (3.21)
Glu105	Arg3	–	14.51 (3.54)
Glu105	Arg6	–	14.38 (3.23)
Glu123	Lys130	8.69 (0.88)	8.96 (0.91)
Glu127	Lys130	4.34 (1.07)	4.83 (1.12)

3.5 | Principal component analysis

For gaining insights into the conformational equilibria of the G66V mutant and for providing a comparison with wild-type CHCHD10, we conducted principal component analysis (PCA).[1] This analysis helps for understanding their conformational dynamics onto two-dimensional space depending on the first two average principal coordinates and the conformational free energy was defined as an additional third coordinate. Figure 6A,B presents our results for the wild-type and G66V mutant CHCHD10 in water. We note—based on calculated first and second PCs—that the G66V mutant-type occupies more subspace than its wild-type form. Based on these findings, we can conclude that the G66V genetic mutation impacts the conformational ensembles of its wild-type form. These differences in dynamics

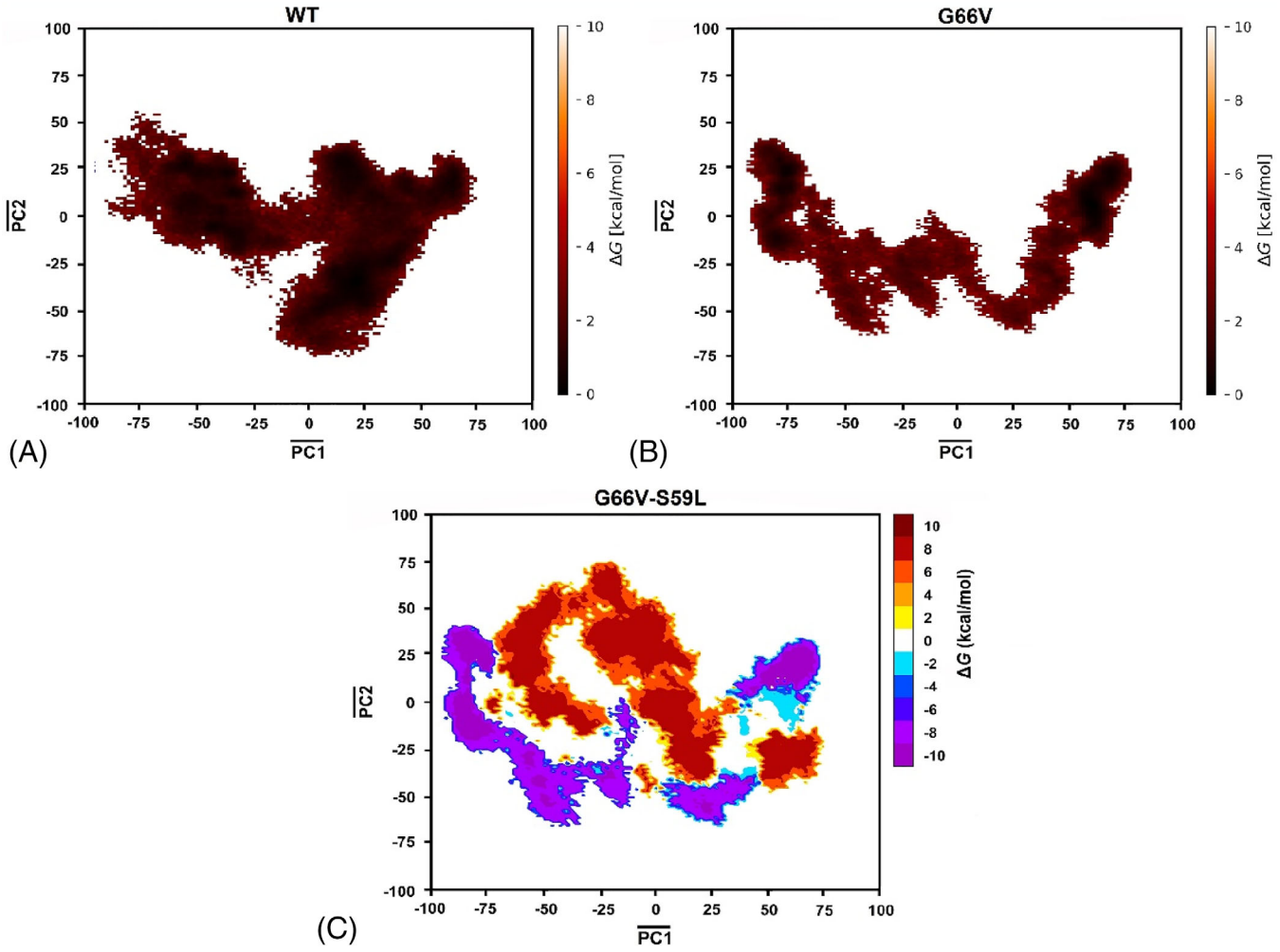

FIGURE 6 (A) Free energy surface areas of wild-type and (B) G66V mutant-type CHCHD10 from our PCA analysis. (C) The difference in free energy surface areas from PCA analysis between the G66V and S59L mutant-type CHCHD10. PC's in all three panels are calculated using all conformations from multiple run MD simulations for wild-type, G66V, and S59L mutant CHCHD10 proteins in water. The backbone atoms were used in these calculations. MD, molecular dynamics; PCA, principal component analysis.

may be linked to the reported cause of motoneuron disease by haploinsufficiency created by G66V mutation. Figure 6C shows the difference in PCA surface area between G66V and S59L mutant-type CHCHD10. As illustrated, different dynamics are obtained for G66V and S59L mutant-type of CHCHD10.[1] These dynamic differences may be related to their varying pathological roles in neurodegenerative diseases. Figure S8 in the Supporting Information section shows the PCA analysis obtained from each multiple run MD simulation.

Figure 7 shows the impact of the G66V mutation on the intrinsic disorder predisposition of human CHCHD10 protein.[1] According to Figure 7A, the N-terminal region of this protein (residues 1–98) preceding the CHCH domain is proposed to embrace increased intrinsic disorder. The average disorder score (which is calculated as a sum of the per-residues disorder probabilities normalized by the length of the considered region) for this region is 0.83 ± 0.11. Figure 7B indicates that the G66V causes some decrease in the disorder predisposition of this region (its mean disorder score decreases to 0.81 ± 0.13). This is an interesting observation considering the fact that we are looking at the impact of a genetic point mutation on the disorder predisposition of almost 100-residue-long region. To get a better representation of this effect on local disorder predisposition of human CHCHD10 protein, we looked at the "difference disorder spectra" generated a set of widely used disorder predictors, such as PONDR® VL3, PONDR® VLXT, PONDR® FIT, PONDR® VSL2, IUPred short, and IUPred long.[24,25] Data from this analysis are presented in Figure 7C that zooms into the region spanning residues Met45-Thr90. "Difference disorder spectra" are calculated via subtracting disorder profile generated for wild-type CHCHD10 utilizing a given predictor from the disorder profile of the G66V mutant-type generated by the similar predictor. Negative values of the resulting differences indicate that the local disorder propensity of CHCHD10 is decreased due to the G66V mutation, which agrees with other results of this study.

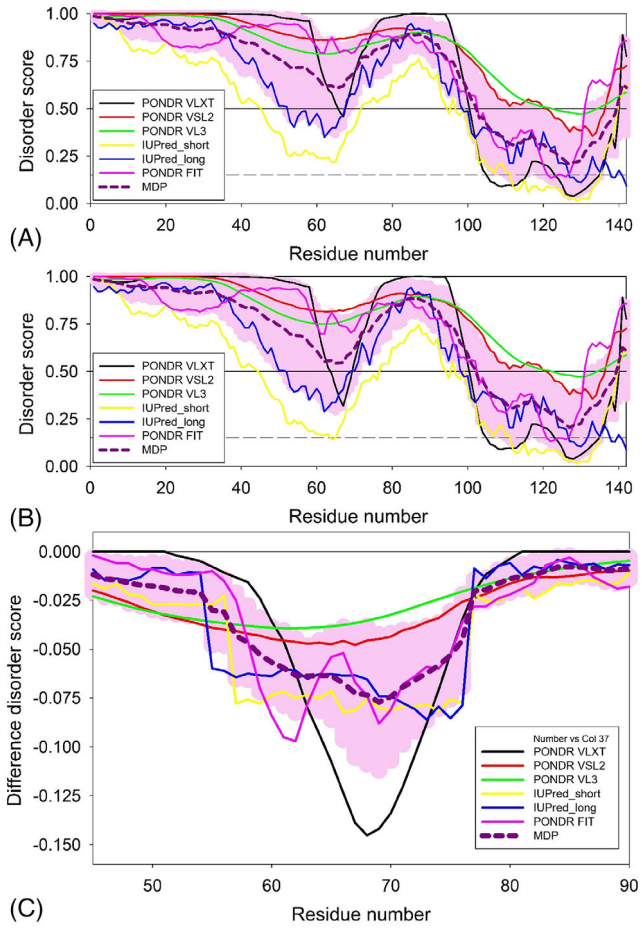

FIGURE 7 Impact of the G66V genetic mutation on intrinsic disorder predisposition of CHCHD10. (A). Intrinsic disorder profile computed for wild-type of CHCHD10 by a set of disorder predictors shown as sequence distributions of the corresponding per-residue intrinsic disorder propensities evaluated by PONDR® VLXT, PONDR® VSL2, PONDR® VL3, IUPred2A Long, IUPred 2A short, and PONDR® FIT. (B). Intrinsic disorder profile of the G66V mutant form calculated by the same predictors. (C). Visualization of the impact of the G66V mutation on local disorder propensity as "disorder spectra deviation," which were computed by subtracting the profile of wild-type CHCHD10 by PONDR® VLXT, PONDR® VSL2, PONDR® VL3, IUPred2A Long, IUPred 2A short, and PONDR® FIT, from the analogous disorder profile calculated for the G66V mutant-type CHCHD10. Average disorder profile along with standard deviations are presented as dashed bold dark pink line and light pink shade.

4 | CONCLUSION

Here, we report the impacts of G66V genetic mutation on the structures and dynamics of wild-type CHCHD10 in water. Results illustrate that α-helix and β-sheet formations of wild-type CHCHD10 are impacted by the G66V genetic mutation. Interestingly, radius of gyration values do not differ significantly and indicate that the genetic mutation does not influence the compactness of wild-type CHCHD10 strongly but an effect on its structural ensemble is expected. Despite, the end-to-end distance values demonstrate stark differences upon G66V mutation and the G66V mutant has a greater end-to-end distance value when we compare the results to those obtained for wild-type CHCHD10. Ramachandran plots were analyzed for further gaining insights into the structural properties caused by the G66V genetic mutation.

The number of hydrogen bonds is similar in the G66V mutant form of CHCHD10 and in wild-type CHCHD10. However, the tertiary structural property analyses reveal that the C- and N-terminus interactions occurring in the wild-type and some of the interactions that occur between the N-terminus and C-terminal regions, N-terminal and mid-domain region, and mid-domain and C-terminal region are weaker or disappear upon G66V mutation. There is a stark difference in the free energy surface areas depending on radius of gyration and end-to-end values upon G66V mutation. Moreover, a stark deviation in the conformational ensemble dynamics is also noted in our principal component analysis upon G66V mutation. Interestingly, intermolecular interactions of CHCHD10 with CHCHD2, COX6B, and CXXC5 are lost with the G66V mutation, indicating that this mutation induces broad changes in protein–protein interactions.[10] Our results indicate that these broad changes in protein–protein interactions may be a

result of different conformational ensemble characteristics upon G66V mutation. Furthermore, the slow progressing nature of G66V-related SMAJ may be due to the higher flexibility related enhanced flexible nature of the CHCHD10-G66V protein, which otherwise could produce a stronger phenotype and this may be related to the higher flexibility of G66V mutant form in comparison to wild-type CHCHD10. Such higher flexibility of CHCHD10-G66V may be related to its weaker intramolecular interactions and greater protein flexibility compared to wild-type CHCHD10, which could allow easier access by the degradation machinery. Unlike S59L and R15L mutations, the G66V mutation does not impair the MICOS complex, and no experimental study has shown that the G66V mutation alters CHCHD10 aggregation. This is in accord with our secondary structural property analysis since we do not expect an increased likelihood toward aggregation or oligomerization for the G66V mutant form of CHCHD10 (see above).

All in all, our data show that the structural ensemble properties of G66V mutant-type CHCHD10, which is at the center of the Jokela type spinal muscular atrophy disease, differ from those of wild-type CHCHD10 and those of S59L mutant form of CHCHD10.

AUTHOR CONTRIBUTIONS

Hakan Alici: Investigation; software; formal analysis. **Vladimir N. Uversky:** Conceptualization; formal analysis; writing – review and editing. **David E. Kang:** Writing – review and editing; conceptualization. **Junga Alexa Woo:** Conceptualization; writing – review and editing. **Orkid Coskuner-Weber:** Conceptualization; methodology; formal analysis; writing – original draft; writing – review and editing; software; validation; investigation; supervision.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Hakan Alici https://orcid.org/0000-0001-5105-8331
Orkid Coskuner-Weber https://orcid.org/0000-0002-0772-9350

REFERENCES

1. Alici H, Uversky VN, Kang DE, Woo JA, Coskuner-Weber O. Structures of the wild-type and S59L mutant CHCHD10 proteins important in amyotrophic lateral sclerosis-frontotemporal dementia. *ACS Chem Nerosci*. 2022;13(8):1273-1280. doi:10.1021/acschemneuro.2c00011
2. Ait-El-Mkadem Saadi S, Chaussenot A, Bannwarth S, Rouzier C, Paquis-Flucklinger V. CHCHD10-related disorders. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews®*. University of Washington, Seattle; 1993.
3. Ajroud-Driss S, Fecto F, Ajroud K, et al. Mutation in the novel nuclear-encoded mitochondrial protein CHCHD10 in a family with autosomal dominant mitochondrial myopathy. *Neurogenetics*. 2015;16(1):1-9. doi:10.1007/s10048-014-0421-1
4. Brockmann SJ, Freischmidt A, Oeckl P, et al. CHCHD10 mutations p.R15L and p.G66V cause motoneuron disease by haploinsufficiency. *Hum Mol Genet*. 2018;27(4):706-715. doi:10.1093/hmg/ddx436
5. Lehmer C, Schludi MH, Ransom L, et al. A novel CHCHD10 mutation implicates a Mia40-dependent mitochondrial import deficit in ALS. *EMBO Mol Med*. 2018;10(6):e8558. doi:10.15252/emmm.201708558
6. Auranen M, Ylikallio E, Shcherbii M, et al. CHCHD10 variant p.(Gly66-Val) causes axonal Charcot-Marie-Tooth disease. *Neurol Genet*. 2015;1(1):e1. doi:10.1212/NXG.0000000000000003
7. Burstein SR, Valsecchi F, Kawamata H, et al. In vitro and in vivo studies of the ALS-FTLD protein CHCHD10 reveal novel mitochondrial topology and protein interactions. *Hum Mol Genet*. 2018;27(1):160-177. doi:10.1093/hmg/ddx397
8. Genin EC, Bannwarth S, Lespinasse F, et al. Loss of MICOS complex integrity and mitochondrial damage, but not TDP-43 mitochondrial localisation, are likely associated with severity of CHCHD10-related diseases. *Neurobiol Dis*. 2018;119:159-171. doi:10.1016/j.nbd.2018.07.027
9. Akbayrak IY, Caglayan SI, Ozcan Z, Uversky VN, Coskuner-Weber O. Current challenges and limitations in the studies of intrinsically disordered proteins in neurodegenerative diseases by computer simulations. *Curr Alzheimer Res*. 2020;17(9):805-818. doi:10.2174/1567205017666201109094908
10. Purandare N, Somayajulu M, Hüttemann M, Grossman LI, Aras S. The cellular stress proteins CHCHD10 and MNRR1 (CHCHD2): partners in mitochondrial and nuclear function and dysfunction. *J Biol Chem*. 2018;293(17):6517-6529. doi:10.1074/jbc.RA117.001073
11. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res*. 2015;43(W1):W174-W181. doi:10.1093/nar/gkv342
12. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun*. 1995;91(1–3):43-56. doi:10.1016/0010-4655(95)00042-E
13. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935. doi:10.1063/1.445869
14. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem*. 2013;34(25):2135-2145. doi:10.1002/jcc.23354
15. Caliskan M, Mandaci SY, Uversky VN, Coskuner-Weber O. Secondary structure dependence of amyloid-β(1-40) on simulation techniques and force field parameters. *Chem Biol Drug des*. 2021;97(5):1100-1108. doi:10.1111/cbdd.13830
16. Systems M. In: Allison TC, Coskuner O, Gonzalez CA, eds. *A Quantum Chemist's Perspective*. 1st ed. CRC Press; 2011. doi:10.1201/b10835
17. Darden T, York D, Pedersen L. Particle mesh Ewald: an *N* ·log(*N*) method for Ewald sums in large Systems. *J Chem Phys*. 1993;98(12):10089-10092. doi:10.1063/1.464397
18. Evans DJ, Holian BL. The nose–hoover thermostat. *J Chem Phys*. 1985;83(8):4069-4074. doi:10.1063/1.449071
19. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 1981;52(12):7182-7190. doi:10.1063/1.328693
20. Hess B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J Chem Theory Comput*. 2008;4(1):116-122. doi:10.1021/ct700200b
21. Wise O, Coskuner O. New force field parameters for Metalloproteins I: divalent copper ion centers including three histidine residues and an oxygen-ligated amino acid residue. *J Comput Chem*. 2014;35(17):1278-1289. doi:10.1002/jcc.23622
22. Coskuner O, Uversky VN. Tyrosine regulates β-sheet structure formation in amyloid-β_{42}: a new clustering algorithm for disordered proteins. *J Chem Inf Model*. 2017;57(6):1342-1358. doi:10.1021/acs.jcim.6b00761

23. Coskuner-Weber O, Uversky VN. Alanine scanning effects on the biochemical and biophysical properties of intrinsically disordered proteins: a case study of the histidine to alanine mutations in amyloid-B42. *J Chem Inf Model*. 2019;59(2):871-884. doi:10.1021/acs.jcim.8b00926
24. Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta*. 2010;1804(4):996-1010. doi:10.1016/j.bbapap.2010.01.011
25. Dosztányi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*. 2005;21(16):3433-3434. doi:10.1093/bioinformatics/bti541

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Alici H, Uversky VN, Kang DE, Woo JA, Coskuner-Weber O. Effects of the Jokela type of spinal muscular atrophy-related G66V mutation on the structural ensemble characteristics of CHCHD10. *Proteins*. 2023;91(6):739-749. doi:10.1002/prot.26463