Conformational Dynamics and Molecular Characterization of Alsin MORN Monomer and Dimeric Assemblies

Proteins: Structure, Function, and Bioinformatics

WILEY

PROTEINS

RESEARCH ARTICLE

Conformational Dynamics and Molecular Characterization of Alsin MORN Monomer and Dimeric Assemblies

Marcello Miceli | Marco Cannariato | Riccardo Tortarolo | Lorenzo Pallante | Eric A. Zizzi | Marco A. Deriu

PolitoBIOMed Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy

Correspondence: Marco A. Deriu (marco.deriu@polito.it)

Received: 8 December 2023 | **Revised:** 5 June 2024 | **Accepted:** 27 June 2024

Funding: This work was supported by Fondazione Telethon (grant # GSP 20005_PAsIAHSP007) and Associazione Help Olly Onlus—Italy (https://helpolly.it/) within the framework of CRYSTAL (www.crystal.m3b.it).

Keywords: ALS | Alsin | IAHSP | molecular dynamics | MORN | rare disease

ABSTRACT

Despite the ubiquity of membrane occupation recognition nexus (MORN) motifs across diverse species in both eukaryotic and prokaryotic organisms, these protein domains remain poorly characterized. Their significance is underscored in the context of the Alsin protein, implicated in the debilitating condition known as infantile-onset ascending hereditary spastic paralysis (IAHSP). Recent investigations have proposed that mutations within the Alsin MORN domain disrupt proper protein assembly, precluding the formation of the requisite tetrameric configuration essential for the protein's inherent biological activity. However, a comprehensive understanding of the relationship between the biological functions of Alsin and its three-dimensional molecular structure is hindered by the lack of available experimental structures. In this study, we employed and compared several protein structure prediction algorithms to identify a three-dimensional structure for the putative MORN of Alsin. Furthermore, inspired by experimental pieces of evidence from previous studies, we employed the developed models to predict and investigate two homo-dimeric assemblies, characterizing their stability. This study's insights into the three-dimensional structure of the Alsin MORN domain and the stability dynamics of its homo-dimeric assemblies suggest an antiparallel linear configuration stabilized by a noncovalent interaction network.

1 | Introduction

Infantile-onset ascending hereditary spastic paralysis (IAHSP) is a rare genetic disease associated with improper tetramerization of a protein known as Alsin, encoded by the ALS2 gene. The Alsin protein consists of five structured domains [1], and the manifestation of IAHSP has been connected to mutations in these domains, ultimately leading to improper Alsin tetramerization [2–5]. Unfortunately, a thorough understanding of the mechanical and biological functions of Alsin in both physiological and pathological conditions is currently hindered by the absence of a three-dimensional experimental structure of the protein. As a matter of fact, three-dimensional protein structure models enable the exploration of structural characteristics and underlying molecular mechanisms, through the use of computational methods which facilitate an atomistic-level investigation of protein biophysics and behavior. This approach helps unravel the complex relationships between protein structures and their biological functions, which is particularly relevant in the investigation of proteins associated with rare diseases [6–9]. In this context, we recently developed and characterized molecular models for three of the five structured domains of Alsin, namely the RCC-like domain (RLD) [6], the Dbl homology and Pleckstrin homology (DH-PH) domain [7], and the vacuolar protein sorting (VPS9) domain [8].

In the present work, we focus on the membrane occupation and recognition nexus (MORN)-like domain, which has been

suggested to be located in the region between the DH-PH and VPS9 domains of Alsin [10, 11]. MORNs are protein domains formed by several structural repetitions and are found ubiquitously in both eukaryotes and prokaryotes [12]. The name derives from the originally proposed role of mediating plasma membrane association, which was hypothesized when this assembly was first discovered as a domain of eight repetitions in junctophilins [13]. The analysis of the sequences of MORN-containing proteins showed that the number of MORN repetitions varies greatly (from 2 to 20) and that MORN motifs can be found in combination with several other domains, making a complete understanding of their biological roles challenging. The limited structural and biological characterization of MORN domains has been for a long time the primary obstacle to comprehending these ubiquitous motifs and their roles in the biological processes they participate in [12]. Even though these domains are generally assumed to be lipid-binding sites [13, 14], there is no direct evidence of this biological function, and experimental data are often conflicting and ambiguous. [13, 15, 16].

Three MORN domains have been recently solved by x-ray crystallography, obtained from three different all-MORN proteins found in *Trypanosoma brucei* (*Tb*), *Toxoplasma gondii* (*Tg*), and *Plasmodium falciparum* (*Pf*). Following the comparison of these structures, the MORN unit was described as a 23–24 amino acid domain [14, 17, 18] with a highly conserved secondary structure, comprising an all-β-strand motif featuring a β-hairpin structure, consisting of a strand-loop-strand pattern. Furthermore, the crystallographic structures led to a new hypothesis regarding the biological function of MORN motifs, suggesting a potential role as mediators in homophilic interactions [19]. Most interestingly, the three MORN structures exhibited two potential homo-dimeric spatial arrangements: linear and V-shaped conformations. Specifically, the *Tb*-MORN structure displayed a preference for a linear assembly, the *Tg*-MORN structure showed both linear and V-shaped arrangements, and the *Pf*-MORN structure was characterized by a preferentially V-shaped arrangement [19]. Interestingly, the interaction between adjacent homo-dimers is coordinated by a Zn^{2+} ion in V-shaped configurations of the *Tg*- and *Pf*-MORNs: these structures conserve a cysteine (CYS), a glutamic acid (GLU), and an aspartate (ASP) at the same sequence position and the interface between adjacent monomers [19]. In the context of the Alsin protein, experiments have highlighted that its aggregation process may be facilitated through its MORN domain, specifically regions of amino acids between 1233–1351 and 1351–1454 [10, 11].

Based on this hypothesis that the MORN domains of Alsin play a role in the protein's tetramerization process and given that improper tetramerization is linked to IAHSP pathology, we conducted a computational modeling investigation to examine the MORN domain of Alsin along with two potential dimeric assemblies.

After characterizing the monomer, we examined the stability of the dimeric interface and identified peculiarities among different dimeric assemblies. This study lays the groundwork for understanding potential mechanisms of protein domain interactions associated with physiological Alsin aggregation processes. In particular, whereas the presence of specific amino acid repetitions in the Alsin MORN sequence might suggest that the Alsin could have a stable linear and V-shaped conformation, our computational investigation suggested that the Alsin could form an interface more compatible with a linear assembly than with a V-shaped one.

This study is situated within a broader context focused on characterizing and comprehending the molecular features that underlie the functions and pathological behavior of the Alsin protein. The employed approach constitutes a crucial step toward understanding the aggregation mechanism of this complex multidomain protein and how it is impaired by gene-related mutations, such as those associated with IAHSP pathology.

2 | Materials and Methods

2.1 | Alsin MORN Model Building

In this section, the detailed materials and methods employed to build the molecular structures of the Alsin MORN (A-MORN) and the linear/V-shaped Alsin dimeric assemblies (L-ALS/V-ALS) are described in detail.

2.2 | Alsin MORN (A-MORN)

Since no experimental structure of Alsin MORN was available on the Protein Data Bank (PDB) database, three different protein structure prediction techniques have been employed to generate a three-dimensional structure for the Alsin MORN starting from its primary sequence. In detail, we used (i) classical template-based homology modeling, (ii) I-TASSER [20, 21], and (iii) the Alphafold v2 algorithm [22].

For the template-based homology modeling, the MORN domain sequence was extracted from the complete human Alsin protein sequence (UniProtKB accession: Q96Q42) available in the UniProt database. The specific region employed encompassed amino acids 1049–1244 [23]. Sequence alignment was then performed against the highest resolution structures of MORN domains currently available on the RSCB database (PDB IDs: 6T4D [19], 6T6Q [19], 6T4R [19], 7JR9 [24], 1H3I [25], and 6JLE [24]) using the Molecular Operating Environment (MOE) software [26]. Homology modeling was performed using MOE's homology tool, using as a template the structure with the highest identity and similarity percentages. The obtained model will be referred to as A-MORN^{HO} in the following.

I-TASSER suite [20] [21] was also used to produce the molecular structure of the Alsin protein: we selected the same sequence employed for the classical homology modeling procedure (aa 1049–1244) as input, then the best model was selected according to the best *C*-score (1.81) and TM-score (0.97 ± 0.05). The I-TASSER model will be referred to as A-MORN^{IT} in the following.

The third model was extracted from the whole Alsin model downloaded from the Alphafold protein structure database (https://alphafold.ebi.ac.uk/entry/Q96Q42) and will be referred to as A-MORN^{AF}. From a visual inspection of the whole Alsin model predicted by Alphafold, the MORN β-sheet region involves amino acids ranging from 1031 to 1265. Thus, since the

amino acids of the C-terminal tail are involved in the dimeric interface, the residues comprising this area have been kept in the 3D structure. The quality of A-MORN^{AF} prediction was evaluated in terms of average per residue confidence score (mean pLDDt = 90.08).

The models were evaluated by assessing the stereochemical quality through the Ramachandran plots generated by the MOE software. They were also compared in terms of root mean square deviation (RMSD) and TM-Score, evaluated by the TM-Align tool [27]. The best structure among the three obtained A-MORN models, that is, A-MORN^{HO}, A-MORN^{IT}, and A-MORN^{AF}, was chosen based on the *Z*-score, obtained using the ProSA web server [19]. The *Z*-score of a model is the energy deviation of its structure from a set of random configurations and, therefore, can be considered as a measure of the overall quality of a model [17, 18].

2.3 | Linear and V-Shaped Alsin Dimeric Assemblies (L-ALS and V-ALS)

The best A-MORN model was considered for the following analysis and was employed to build the two dimeric assemblies conformations. Using the MOE software, we recreated the target spatial arrangements employing the superposition tool to align the model chains to the interfaces of the linear and V-shaped structures of the *Pf*-MORN and *Tb*-MORN (PDB ID: 6T4D and 6T4R [19]), respectively. As previously mentioned, the dimeric assembly of the V-shaped conformations is coordinated by a Zn^{2+} ion thanks to the presence of a conserved quadruplet sequence formed by residues CYS306-x-GLU308-ASP309 [19]. Since the Alsin MORN sequence presented a similar group of residues (SER1215-GLU1216-ASP1217-ASP1218) in a similar sequence position, the Zn^{2+} ion was preserved in the V-shaped super-assembly. Then, for both the linear and V-shaped Alsin MORN, energy minimizations were performed with MOE standard parameters, and the Ramachandran plots were analyzed to assess the stereochemical quality of the models. The linear and V-shaped models will be referred to as L-ALS and V-ALS, respectively, and are represented in Figure 1.

2.4 | Molecular Dynamics Investigation of Alsin MORN Monomer and Dimeric Assemblies

Molecular dynamics (MD) simulations were carried out for the A-MORN^{AF} monomer model and the two dimeric assemblies, that is, L-ALS and V-ALS, employing GROMACS 2020.4 [28]. The systems were inserted into a dodecahedral box with periodic boundary conditions imposing a minimum distance between the protein and the box edge of 1.5 nm. Each system was solvated by the explicit TIP3P water model [29] imposing physiological Na^+/Cl^- salt concentration conditions (0.15 M) to counterbalance the net charge of the systems. The steepest descent method was applied for 5000 steps during the energy minimization phase. Then, each system undergoes two equilibration steps with constant volume and temperature (NVT) and with constant pressure and temperature (NPT), respectively. The NVT and NPT equilibration steps were performed using position restraints on the protein-heavy atoms and with reference values for temperature and pressure of $T = 300$ K and $P = 1.0$ bar, respectively. Berendsen thermostat [30] was used

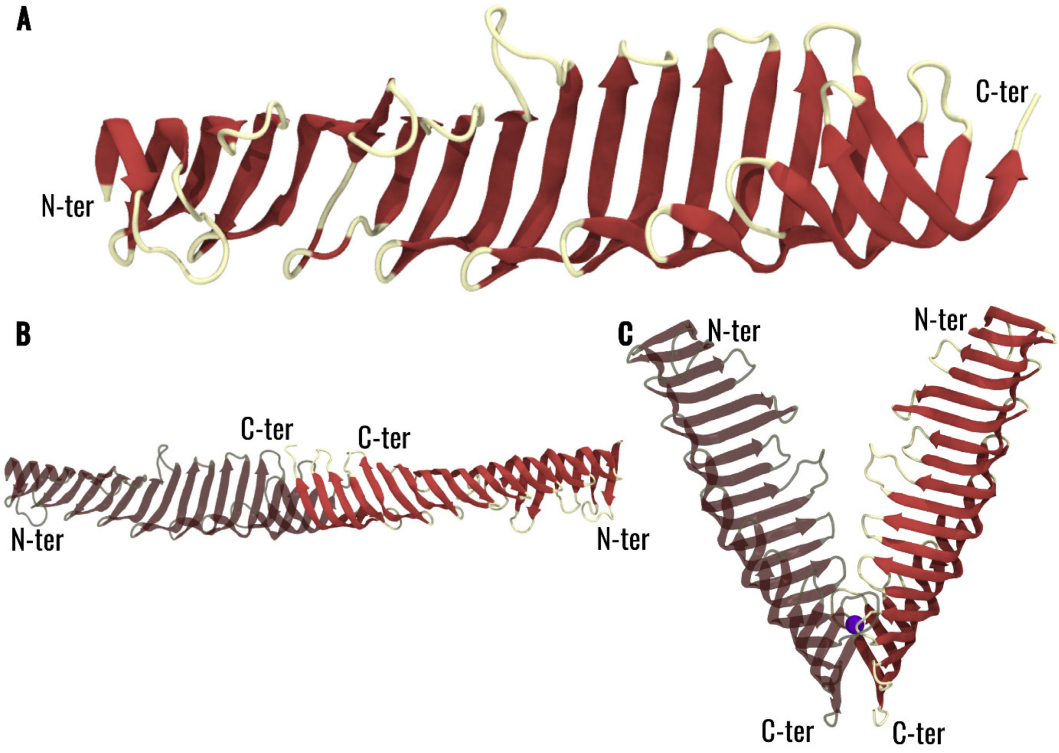

FIGURE 1 | Protein structures rendering of the Alsin MORN models, that is, (A) A-MORN^{AF}, (B) L-ALS, and (C) V-ALS. The structure is rendered as a new cartoon representation and colored according to the secondary structure (red for β-sheet, yellow for coil, and orange for α-helix). For the V-ALS the zinc ion (violet) is rendered as Van der Waals.

for NVT equilibration ($\tau = 0.1$), whereas Parrinello–Rahman [31] isotropic coupling was selected for NPT equilibration ($\tau = 2.0$). The total time for the NVT/NPT equilibration phase was 1 ns. After finishing the equilibration step and removing the position restraints, the production phase in an NPT ensemble was performed for 200 ns for the monomer and 1000 ns for the dimeric assemblies. The simulation time step was imposed to 2 fs, and the AMBER ff99SB-ILDN force field [32] was employed for the system parametrization. Electrostatic interactions were treated using the PME scheme [33] with a cutoff of 1 nm for both Van der Waals and Electrostatic interactions. Three replicas for each system were made to increase the statistics of the data and ensure the repeatability of the results. Therefore, a total of 600 ns for the monomer system and 3 μs for each dimeric system during the unrestrained simulation phase were performed.

2.5 | Conformational Analysis

For the dimeric assemblies, the two chains were named and analyzed separately: the first Alsin MORN monomer was defined as Chain A and the second monomer was defined as Chain B.

After the MD simulations, analyses to assess the stability of the systems were performed by RMSD and root-mean-square fluctuations (RMSF) calculations. In the case of the two superassemblies, RMSD was evaluated (i) separately on each chain after removing translational and rotational artifact fitting the coordinated on the position of the α-Carbons (Cα), (ii) for the complex after fitting the coordinates of the dimer on the Cα and (iii) on the Cα atoms in the C-terminal regions (aa 1216–1267) of both chains. RMSF analysis was performed on the Cα of each chain independently to avoid artifact generation due to global structure movements. The probability of an amino acid being involved in a secondary structure (i.e., α-helix, β-sheet, turn, or coil) was assessed using the MDTraj python library that utilized the Dictionary of Protein Secondary Structure (DSSP) program [34] as the backend. The secondary structure probability for each amino acid was calculated by dividing the total count of occurrences of a specific secondary structure by the total number of frames in the concatenated trajectories, as done previously [6, 9]. The probability of a specific interaction between the two chains of the dimer (i.e. hydrogen bonds, hydrophobic interactions, salt bridges, metal complexes, etc.) has been evaluated using the PLIP python library [35] on each frame and then averaging the number of the occurrences on the total number of frames. Interactions with a probability higher than 50% were selected, according to previous studies [6, 9].

To further investigate the hypothesis that the triplet of residues formed by the SER1214, GLU1215, and ASP1216 could act as a zinc coordination motif, as reported in previous literature for other MORN motifs [19], the minimum distance between the oxygens of the side chains of the putative amino acids and the Zn^{2+} ion was evaluated throughout the simulation. Analyses were performed employing GROMACS tools [28], specific Python packages, that is, MDAnalysis [36] and MDTraj [37], and in-house scripts. Analysis plots were obtained with the Matplotlib library [38], whereas molecular system rendering and visual inspections were made with the Visual Molecular Dynamics (VMD) software [39].

3 | Results

3.1 | Analysis of Alsin MORN Sequence and Model Building

The Alsin MORN sequence was compared with other six MORN protein domains (PDBID: 6T4D [19], 6T6Q [19], 6T4R [19], 7JR9 [24], 1H3I [25], and 6JLE [24]). Results revealed that the higher identity (34.7%) and similarity (49.5%) were achieved by the *Tg*-MORN sequence (PDBID: 6T6Q) followed by the *Pf*-MORN (PDBID: 6T4D) sequence (see also Table S1). The Alsin MORN sequences, together with the *Tg*-MORN and *Pf*-MORN sequences, were split to carry out a comparative analysis across MORN family, similarly to a previous consensus study [19]. The most conserved residues, observed in each repetition of MORNs, were the glycine amino acids in the starting position and the glycines in position 10 (Figure 2). Regarding the presence of a YEGEW motif in positions 13–17, in the Alsin sequence compared with the other MORNs, the TRP17 is partially replaced with another apolar amino acid (PHE) and the conservation of GLU16 is lost (Figure 2).

All the 3D-developed models (i.e., A-MORN^{HO}, A-MORN^{IT}, and A-MORN^{AF}) were then compared with select the best one for further investigation. All models showed negative *Z*-score (A-MORN^{AF}: −5.46; A-MORN^{IT}: −4.81; A-MORN^{HO}: −4.7) in the

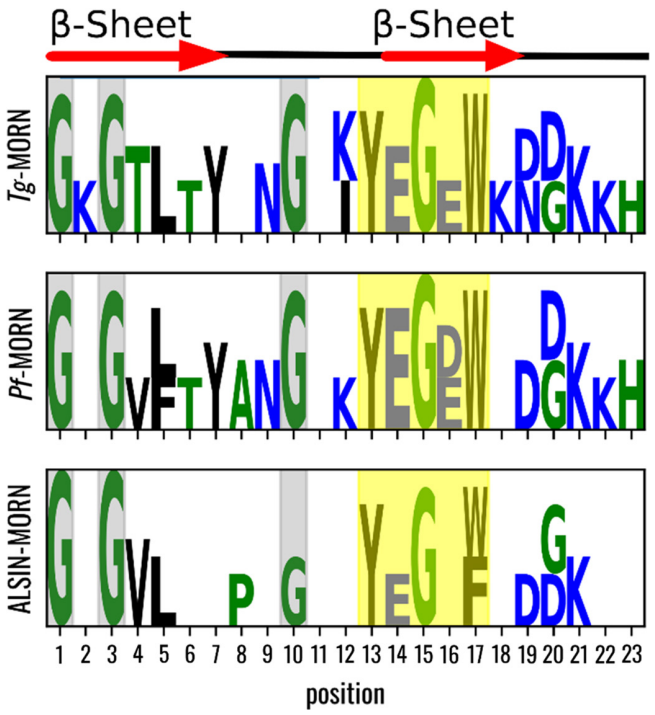

FIGURE 2 | Consensus analysis of the repetitions of the three MORN protein structures, that is, *Pf*-MORN, *Tg*-MORN, and Alsin MORN. Residues are colored according to hydrophobicity, with black for hydrophobic, green for slightly hydrophobic, and blue for hydrophilic. Key motifs reported in previous literature [19] are highlighted by gray and yellow shades.

region of experimentally resolved structures (see also Figure S1). The Ramachandran plot showed a good stereochemical quality of the models with no amino acids in the disallowed region for the A-MORN^{AF} and only two residues in the disallowed region for A-MORN^{IT} and A-MORN^{HO} (see also Figure S2). Then the structural similarities across the models were evaluated in terms of RMSD and TM-Score (see also Table S2). The RMSDs between the models were below 4 Å for each couple with a TM-Score above 0.5, indicating that the models correspond to the protein pairs of similar fold [20]. Thus, given the structural similarity between the three models (see also Figure S3), we selected the A-MORN^{AF} model, for the following reasons: the lowest *Z*-score, the lowest number of amino acids in the disallowed region and since it includes terminal regions, which could be involved in the dimerization. The model selected has been employed for further MD investigations and to build the dimeric assembly (Figure 2).

The selected model for the Alsin MORN domain, namely the A-MORN^{AF} model, is structurally characterized by an all-β-secondary structure, formed by nine repetitions, with each β-strand connected by flexible loops (Figure 1A).

3.2 | Alsin MORN (A-MORN) Conformational Dynamics

Structural integrity and dynamical behavior were assessed through the analysis along the three replicas of 200 ns trajectory produced by MD simulations of the A-MORN^{AF} model. The evaluation of the RMSD showed that the structure quickly reached equilibrium around the value of 0.40 ± 0.04 nm (see also Figure S4). All ensemble averages have been performed on the last 50 ns of each replica, extracting a sample every 100 ps. On average, the β-sheet repeat motif is maintained for the nine structural repetitions, with a slight tendency of the C-terminal β-strand to transit to a coil conformation (Figure 3A). The RMSF analysis showed four regions with high fluctuations: (i) the N-terminal tail, especially around SER1031 (0.35 ± 0.09 nm), (ii) the C-terminal tail, especially around LEU1268 (0.44 ± 0.03 nm), (iii) the coil region composed of residues LYS1092 (0.31 ± 0.04 nm) and ALA1093 (0.35 ± 0.04), and (iv) the loop formed by LEU1143 (0.35 ± 0.04 nm), THR1144 (0.37 ± 0.06 nm), SER1145 (0.35 ± 0.049 nm), and SER1146 (0.32 ± 0.09 nm) (Figure 3B,C).

3.3 | Conformational Dynamics of Alsin MORN Dimeric Assemblies (L-ALS and V-ALS)

Following the hypothesis that the MORN family could act as a dimerization domain and the evidence that the region comprising the Alsin MORN is fundamental for its proper tetramerization [10], two models of Alsin MORN homodimer were developed and simulated. The two models were built on the previously described experimental arrangements of the MORN dimeric assemblies, namely the linear and the V-shaped conformations [19]. The obtained linear and V-shaped models are referred to as L-ALS and V-ALS, respectively. Both models were simulated in three different replicas of 1000 ns each, and ensemble averages were extracted from the last 600 ns of each replica every 100 ps.

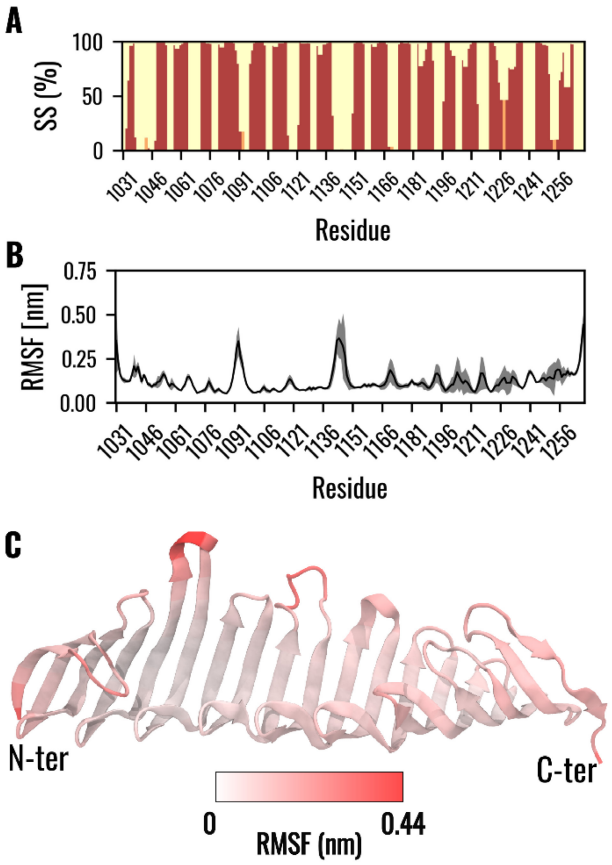

FIGURE 3 | (A) Probability distribution of secondary structures over time for the A-MORN^{AF} (red for β-sheet, yellow for coil, and orange for α-helix). (B) RMSF of the A-MORN^{AF} during the equilibrium reported as mean (black line) and confidence interval at 98% (±1.98σ) (gray shade). (C) RMSF values reported on the Alsin MORN 3D model and colored according to a reddish color scale.

To evaluate the structural integrity of the single chains of L-ALS and V-ALS, the RMSD on the C-alpha carbons with respect to the initial configurations was calculated for each chain separately after removing the rotational and translational artifacts.

The secondary structure probability and the RMSF were employed to assess the effect of dimerization on the dynamic behavior of the domain. An analysis of the amino acids most probably involved in the interface was performed to shed light on the residues fundamental for the stability of the complex.

3.3.1 | Structural Analysis and Interface Characterization of the Linear Alsin Dimer (L-ALS)

The RMSD evaluated on every single chain shows that each monomer in the complex does not undergo major conformational changes, and it is stabilized at 0.30 ± 0.05 nm for Chain A and 0.35 ± 0.05 for Chain B (see also Figure S5). On the other hand, the RMSD evaluated on the complex stabilized to higher values around 0.73 ± 0.21 nm (see also Figure S6). RMSD analysis of the heavy atoms of the C-tails, which are involved in the interface, showed a plateau across the three replicas with an average value between the last frames of each replica of 0.27 ± 0.03 nm (Figure S7).

The secondary structure of each chain in the linear conformation is stable during the simulation time and similar to the single monomer one, but with a greater stabilization for the C-terminal region (Figure 4A). In contrast to the monomeric Alsin, the C-terminal segment was stable in a β-sheet conformation created by the C-terminal regions of the two adjacent monomers forming an antiparallel dimeric interface (Figure 4C). Analysis of the RMSF showed several peaks corresponding to the loops connecting each β-strand and the ending termini. Interestingly, each chain reported a similar behavior to its counterpart and to the Alsin MORN monomer: the highest fluctuations are exhibited around residues SER1031 (0.43 ± 0.05 nm) belonging to the N-terminal, in the loop composed by residues LEU1143 (0.4 ± 0.02 nm), THR1144 (0.49 ± 0.08 nm), SER1145 (0.49 ± 0.14 nm), SER1146 (0.4 ± 0.09 nm), and around residues SER1267 (0.35 ± 0.43 nm) and LEU1268 (0.46 ± 0.08 nm) within the C-terminal (Figure 4B).

The most occurring interactions during MD simulations between the L-ALS monomers (Chain A and Chain B) are listed in Table 1. As also highlighted in Figure 4C, the dimeric interface was characterized by hydrophobic interactions, hydrogen bonds, and a salt bridge. In greater detail, a symmetric pattern of hydrophobic interactions between the amino acids ILE1259 and TYR1263 of the two chains was emphasized. Moreover, the results highlighted a symmetric pattern of hydrogen bonds formed by TYR1219-ILE1257 and ILE1259-GLY1261 between the two chains and a nonsymmetric hydrogen network of THR1219-LYS1265 and ASN1241-ASP1217 belonging to Chain A and Chain B, respectively. Finally, a symmetric salt bridge interaction between ASP1217-LYS1265 of chains A and B was observed.

3.3.2 | Structural Analysis and Interface Characterization of the V-Shaped Alsin Dimer (V-ALS)

The RMSD of the V-ALS stabilized at 0.33 ± 0.05 and 0.31 ± 0.02 nm for Chain A and Chain B, respectively, showing no major conformational changes on the single monomer (see also Figure S8). On the other hand, the RMSD on the whole complex stabilized on higher values around 0.63 ± 0.24 nm (see also Figure S9). RMSD of Cα of the C-tails interfacial residues deviated from the initial configuration, assuming an average value between the last frames of each replica of 0.43 ± 0.16 nm.

The secondary structure of the V-ALS was stable and comparable for both chains during simulation except for (i) the region comprising residues 1226–1235 of Chain B, which lost the secondary structure arrangement and (ii) the C-terminal part of both chains, which showed a transition to an unstructured conformation similar to the case of the single Alsin monomer (Figure 5A).

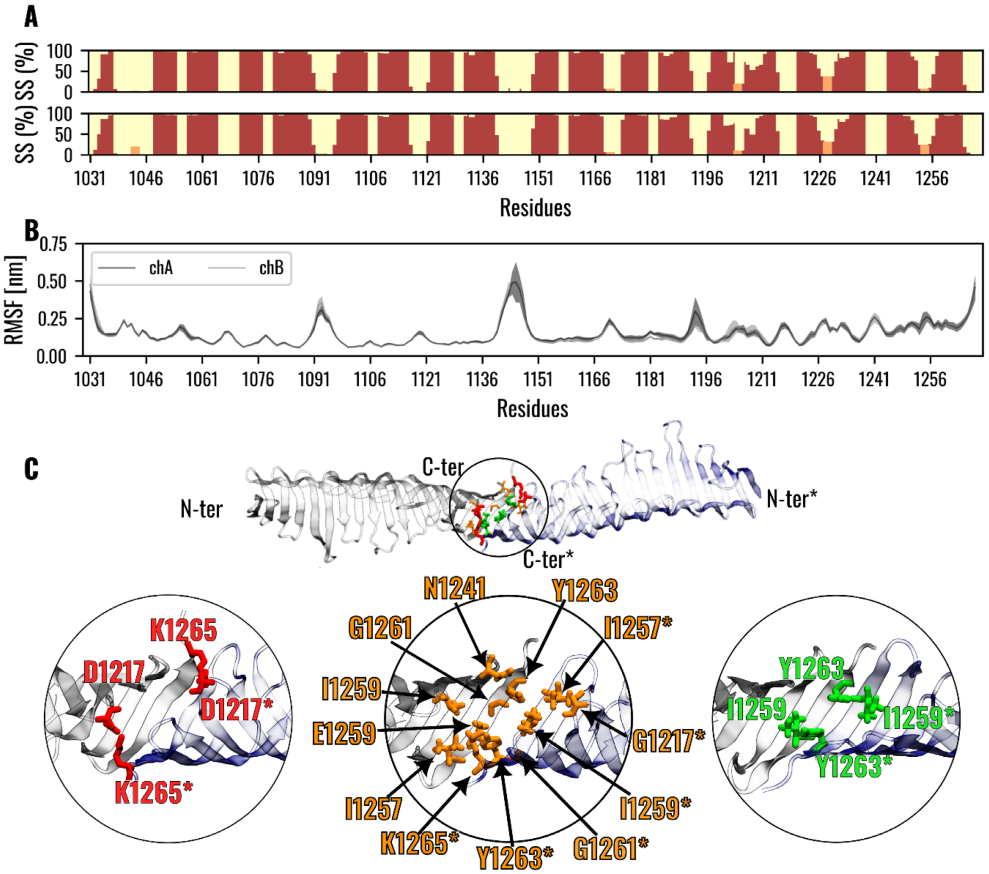

FIGURE 4 | (A) Probability distribution of secondary structure for Chain A (upper) and Chain B (lower) of the L-ALS structure (red for β-sheets, orange for α-helix, and yellow for unstructured). (B) RMSF evaluated on the single chain reported as mean (solid line) and standard deviation (shaded area). (C) Rendering of the L-ALS dimer with amino acids most involved in the interaction interface colored according to the specific type of interaction (red for salt bridges, orange for H-bonds, and green for hydrophobic interactions). The amino acid labels with the "*" character denote residues belonging to Chain B.

TABLE 1 | Contact probability of the most recurring amino acids involved in the interaction interface between L-ALS monomers (Chain A and Chain B).

Type	Chain A	Chain B	Probability
Hydrophobic interaction	ILE1259	TYR1263	70%
Hydrophobic interaction	TYR1263	ILE1259	70%
Hydrogen bond	ILE1257	TYR1263	67%
Hydrogen bond	ILE1259	GLY1261	67%
Hydrogen bond	GLY1261	ILE1259	67%
Hydrogen bond	TYR1263	ILE1257	67%
Hydrogen bond	THR1219	LYS1265	56%
Hydrogen bond	ASN1241	ASP1217	54%
Salt bridge	LYS1265	ASP1217	77%
Salt bridge	ASP1217	LYS1265	83%

Note: Types of interactions are colored according to the color code used in Figure 4, that is, red for salt bridges, orange for H-bonds, and green for hydrophobic interactions.

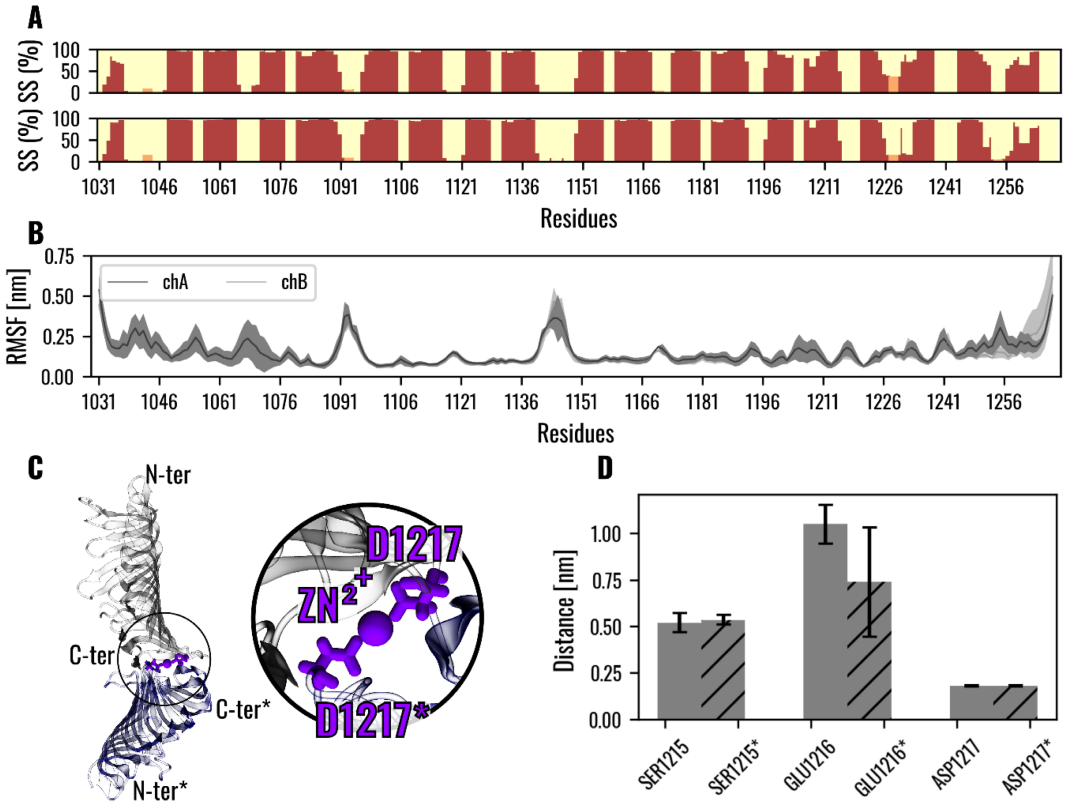

FIGURE 5 | (A) Probability distribution of the secondary structures for Chain A (upper) and Chain B (lower) of the V-ALS dimeric assembly (red for β-sheets, orange for α-helix, and yellow for unstructured regions). (B) RMSF evaluated on the single chain reported as mean (solid line) and standard deviation (shaded area). (C) Rendering of the V-ALS dimer in complex with the zinc ion. Amino acids most involved in the interactions are rendered in violet according to the specific type of interaction (metal complex). The amino acids labels with the "*" character denote residues belonging to Chain B. (D) Distribution of the distance between the putative coordination residues and the zinc ion, reported for each chain (solid bar for Chain A and hatched bar for Chain B) as median and interquartile range.

Analysis of the RMSF did not show any major effect on the residue fluctuations. Indeed, similarly to the single Alsin monomer and the L-ALS, several peaks were observed in the V-ALS in the positions corresponding to the loops connecting each β-strand and in the ending termini. Moreover, as for the L-ALS, the chains reported almost identical behavior in terms of residue fluctuations. The highest fluctuations for both V-ALS chains were demonstrated by the N-terminal SER1031 (0.54 ± 0.09 nm),

the loop formed by residues LYS1092 (0.37 ± 0.09 nm) and ALA1093 (0.39 ± 0.06 nm), a second loop comprising THR1144 (0.36 ± 0.07 nm), SER1145 (0.36 ± 0.15 nm), and SER1146 (0.34 ± 0.09 nm), and residues SER1267 (0.37 ± 0.04 nm) and LEU1268 (0.5 ± 0.03 nm) belonging to the C-terminal (Figure 5B).

Differently from the L-ALS, which forms a stable network of interactions, the V-ALS interactions are not stable during the simulation, meaning that there are no interactions with probability above the threshold of 50%. However, ASP1217 of both chains was able to form a stable metal complex interaction with the zinc ion for all the configurations sampled during simulations (Figure 5C). To further investigate the hypothesis that the triplet of residues formed by the serine (SER1215), glutamic acid (GLU1216), and aspartic acid (ASP1217) could act as a zinc coordination motif as for the other MORNs reported in the literature, each minimum distance between the oxygens of the side chains of the putative amino acids and the Zn^{2+} ion has been evaluated. Results reported as the median and interquartile range for each chain are not compatible with distances for metal site interactions [40] except for ASP1217 which reported values of 0.180 [0.177–0.183] nm for both chains, in line with the above-mentioned contact probability analysis (Figure 5D).

4 | Discussion

MORN is a class of poorly characterized protein domains found across diverse species. Recent studies suggested the possibility of these motifs forming dimeric structures, indicating two possible modes of interaction, namely the linear and V-shaped configurations [19]. Starting from these findings and from the hypothesis that the MORN domain can be a crucial domain for the oligomerization process in the Alsin protein, the present study proposed and characterized molecular models for the single Alsin MORN monomer and the relative dimeric structures in the linear and V-shaped configurations.

Due to the lack of an experimental structure of the human Alsin MORN domain, we tried different computational strategies (classical homology modeling, I-TASSER webserver, and Alphafold v2) to develop a reliable molecular model. The most promising model in terms of stereochemical quality and the *Z*-score was obtained with Alphafold v2. This model was therefore considered for the subsequent simulations and molecular refinements. Compared with the most similar MORN sequences available in the literature, that is, *Pf*- and *Tg*-MORNs, the most conserved residues in each repetition of MORNs were the glycine amino acids at the starting position and the glycines in position 10 (Figure 2). This analysis is in good agreement with previous pieces of literature which underlined, through a consensus sequence alignment among experimental MORNs, the presence of a highly conserved G × G motif at the start of the sequence and a conserved glycine at position 10 (GLY10) [19]. However, in the Alsin sequence, there was a partial replacement of TRP17 with another apolar amino acid (PHE), and the conservation of GLU16 was lost (Figure 2), in contrast to the conservation of the YEGEW motif in positions 13–17 highlighted in previous literature for *Tb*-, *Tg*-, and *Pf*-MORNs [19]. Moreover, the Alsin MORN model developed is formed by nine repetitions, differently from the previous literature that suggested an eight-repetition motif [10, 41]. Despite these small differences between the Alsin MORN and the available experimental MORN structures, the overall β-sheet repeated motif was preserved stable during the simulations, except for a slight tendency toward a coil structure in the C-terminal β-strand (Figure 3A). Moreover, the most fluctuating regions within the protein structure are located in the N- and C-terminal tail and two major loops at sequence positions 1092–1093 and 1143–1146 (Figure 3B,C).

Starting from the above-mentioned Alsin MORN model, we built two homo-dimeric assemblies: we recreated the desired spatial arrangements aligning the model chains with the interfaces of the linear and V-shaped structures of the *Pf*-MORN and *Tb*-MORN (PDB ID: 6T4R and 6T4D [8]), respectively. Therefore, the linear and the V-shaped conformation of Alsin have been obtained and referred to as L-ALS and V-ALS, respectively. Both dimers have an antiparallel shape, which involves the C-terminal region of the Alsin monomer. Interestingly, the L- and V-ALS showed an interface formed by amino acids ranging from 1216 to 1268 in agreement with regions indispensable for the proper Alsin tetramerization reported in previous experimental studies [10]. We then employed MD simulations to characterize the proposed dimeric structures, with particular focus on the dimeric interface. The L-ALS showed a symmetric pattern of interactions forming a continuous β-sheet-like structure (Figure 4A). Overall, the L-ALS converged to a stable interface, formed by the C-terminal tails, in all three analyzed replicas (Figure S7). As for the experimental *Tb*-MORN, the interface of the L-ALS monomers seemed to be stabilized by symmetric hydrophobic interactions, symmetric and asymmetric hydrogen bonds, and a symmetric salt bridge [19] (Figure 4C). On the other hand, the dimeric interface of the V-shaped conformation tends to considerably deviate from the imposed initial structure without forming stable interactions during simulations (Figure S10).

This is also confirmed by the observation that the V-shaped conformation, coordinated by a Zn^{2+} ion, showed less stable contact in the ensemble of configurations sampled during simulations. The only stable interaction was the symmetric metal complex between the Zn^{2+} ion and ASP1217 of both chains. Interestingly, this aspartic acid is in a position in the sequence close to a glutamic acid and a serine. However, both the contact probability (Figure 5C) and the distance analyses (Figure 5D) highlighted that the Alsin V-shaped conformation did not permit stable intermonomer contacts reducing the stability of the overall dimeric assembly. Previous work indicated that the V-shaped and linear conformation exhibited from the *Pf*- and *Tg*- MORN requires the presence of a triplet of aspartic acid, cysteine, and glutamic acid as crucial to stabilizing a V-shaped conformation [19]. In the *Tb*-MORN the cysteine residue is replaced with a leucine, and linear assembly is observed but not V-shaped. The results on the V-shaped conformation for the Alsin MORN observed in this work are in agreement with previously reported literature, suggesting that the presence of an aspartic acid conserved is not sufficient to maintain the V-shaped interface [19]. Indeed, not only is the presence of aspartic acid necessary for coordinating Zn^{2+}, but it also requires the presence of cysteine to stabilize the V-shaped conformation. Therefore, since the cysteine is not conserved in

the Alsin sequence, similar to *Tb*-MORN, it results in reduced stability of the V-shaped interface.

5 | Conclusion

MORN motifs play roles in various physiological and pathological biological functions. In the Alsin protein, sequence variations in the MORN structures are suggested to be linked to incorrect protein aggregation, associated with the onset of IAHSP. However, the characterization of MORN motifs is still limited due to the lack of experimental molecular structures. Building upon the hypothesis that MORN motifs may act as interaction domains, this study, using in silico molecular methods, investigated the structure and interaction properties of Alsin MORN domains. The current investigation, through the comparison of various structure prediction approaches and the utilization of extensive MD simulations has produced an ensemble of configurations of the molecular model concerning the Alsin MORN domain. These ensemble of configurations, together with previously released models of Alsin domains by our group, serves as a starting point for a comprehensive understanding and characterization of the Alsin protein (https://crystal.m3b.it/results/). Furthermore, the creation of linear and V-shaped dimeric assemblies of the Alsin MORN provides insights into the stabilities of these aggregates, highlighting similarities and differences with similar experimental structures. Specifically, the linear conformation of Alsin MORN demonstrated stable interactions during simulations, suggesting a stable dimeric interface. On the contrary, the V-shaped assembly showed only a single residue involved in the intermonomer interaction, indicating a less stable dimeric interface.

These insights can lay the foundation for further in vitro and in silico investigations. For example, wet lab experiments, such as small x-ray light scattering or native mass spectroscopy, could confirm findings on the molecular shape of dimeric Alsin MORN. Moreover, alanine scans and aggregation experiments, aimed at investigating the role of the interfacial amino acids identified in this work, can provide further insights into the oligomerization process of Alsin. Furthermore, an enhanced sampling approach could be used to energetically characterize the Alsin MORN dimer and the effect of mutations on its stability. More specifically, techniques such as umbrella sampling can be used to characterize the free energy difference between the bound and unbound states of the dimeric conformations described here and to characterize the assemblies from a thermodynamic point of view. This approach could also be a future development of the current study to characterize the effect of mutation on the stability of the dimeric assembly.

This and future works are pivotal to unraveling the intricate relationships between Alsin's structural properties and its impact on biological processes, and thus understanding the mechanisms underlying IAHSP disease.

Author Contributions

Marcello Miceli: conceptualization, investigation, formal analysis, data curation, methodology, validation, writing – original draft, writing – review and editing. **Marco Cannariato:** investigation, data curation, formal analysis, writing – review and editing, conceptualization. **Riccardo Tortarolo:** conceptualization, writing – review and editing, investigation, data curation. **Lorenzo Pallante:** writing – review and editing, supervision. **Eric A. Zizzi:** supervision, writing – review and editing. **Marco A. Deriu:** conceptualization, funding acquisition, data curation, project administration, resources, validation, supervision, visualization, writing – review and editing.

Acknowledgments

This work was supported by Fondazione Telethon (Grant #GSP 20005_PAsIAHSP007) and Associazione Help Olly Onlus (https://helpolly.it/)—Italy within the framework of CRYSTAL (www.crystal.m3b.it). We acknowledge the CINECA award under the ISCRA initiative (Project CRYSTALC, Project code HP10CWRPPV), for the availability of high-performance computing resources and support.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All the files necessary to reproduce the simulations are accessible at the website https://github.com/M3B-Lab/Alsin-MORN.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/prot.26728.

References

1. K. Sato, A. Otomo, M. T. Ueda, et al., "Altered Oligomeric States in Pathogenic ALS2 Variants Associated With Juvenile Motor Neuron Diseases Cause Loss of ALS2-Mediated Endosomal Function," *Journal of Biological Chemistry* 293 (2018): 17135–17153, https://doi.org/10.1074/jbc.RA118.003849.

2. C. C. Verschuuren-Bemelmans, P. Winter, D. A. Sival, J.-W. Elting, O. F. Brouwer, and U. Müller, "Novel Homozygous ALS2 Nonsense Mutation (p.Gln715X) in Sibs With Infantile-Onset Ascending Spastic Paralysis: The First Cases From Northwestern Europe," *European Journal of Human Genetics* 16 (2008): 1407–1411, https://doi.org/10.1038/ejhg.2008.108.

3. C. Lai, C. Xie, S. G. Mccormack, et al., "Amyotrophic Lateral Sclerosis 2-Deficiency Leads to Neuronal Degeneration in Amyotrophic Lateral Sclerosis Through Altered AMPA Receptor Trafficking," *Journal of Neuroscience* 26 (2006): 11798–11806, https://doi.org/10.1523/JNEUROSCI.2084-06.2006.

4. S. M. Wakil, K. Ramzan, R. Abuthuraya, et al., "Infantile-Onset Ascending Hereditary Spastic Paraplegia With Bulbar Involvement due to the Novel ALS2 Mutation c.2761C>T," *Gene* 536 (2014): 217–220, https://doi.org/10.1016/j.gene.2013.11.043.

5. E. Eymard-Pierre, G. Lesca, S. Dollet, et al., "Infantile-Onset Ascending Hereditary Spastic Paralysis Is Associated With Mutations in the Alsin Gene," *American Journal of Human Genetics* 71 (2002): 518–527, https://doi.org/10.1086/342359.

6. M. Cannariato, M. Miceli, and M. A. Deriu, "In Silico Investigation of Alsin RLD Conformational Dynamics and Phosphoinositides Binding Mechanism," *PLoS One* 17 (2022): e0270955, https://doi.org/10.1371/journal.pone.0270955.

7. M. Cannariato, M. Miceli, M. Cavaglià, and M. A. Deriu, "Prediction of Protein–Protein Interactions Between Alsin DH/PH and Rac1 and Resulting Protein Dynamics," *Frontiers in Molecular Neuroscience* 14 (2022): 772122, https://doi.org/10.3389/fnmol.2021.772122.

8. M. Miceli, C. Exertier, B. Vallone, M. Cavaglià, and M. A. Deriu, "Elucidating Molecular Connection Between IAHSP Onset and Alsin Protein by Means of Homology Modelling and Molecular Dynamics," *Biomedical Science and Engineering* 4 (2021): 88–89, https://doi.org/10.4081/bse.2021.183.

9. M. Miceli, M. A. Deriu, and G. Grasso, "Toward the Design and Development of Peptidomimetic Inhibitors of the Ataxin-1 Aggregation Pathway," *Biophysical Journal* 121 (2022): 4679–4688, https://doi.org/10.1016/j.bpj.2022.10.021.

10. R. Kunita, A. Otomo, H. Mizumura, et al., "Homo-Oligomerization of ALS2 Through Its Unique Carboxyl-Terminal Regions Is Essential for the ALS2-Associated Rab5 Guanine Nucleotide Exchange Activity and Its Regulatory Function on Endosome Trafficking," *Journal of Biological Chemistry* 279 (2004): 38626–38635, https://doi.org/10.1074/jbc.M406120200.

11. M. Miceli, C. Exertier, M. Cavaglià, et al., "ALS2-Related Motor Neuron Diseases: From Symptoms to Molecules," *Biology* 11 (2022): 77, https://doi.org/10.3390/biology11010077.

12. S. El-Gebali, J. Mistry, A. Bateman, et al., "The Pfam Protein Families Database in 2019," *Nucleic Acids Research* 47 (2019): D427–D432, https://doi.org/10.1093/nar/gky995.

13. H. Takeshima, S. Komazaki, M. Nishi, M. Iino, and K. Kangawa, "Junctophilins: A Novel Family of Junctional Membrane Complex Proteins," *Molecular Cell* 6 (2000): 11–22, https://doi.org/10.1016/s1097-2765(05)00005-5.

14. J. I. Yang, A. J. Davis, I. Y. Perera, E. Johannes, N. S. Allen, and W. F. Boss, "The N-Terminal Membrane Occupation and Recognition Nexus Domain of Arabidopsis Phosphatidylinositol Phosphate Kinase 1 Regulates Enzyme Activity," *Journal of Biological Chemistry* 282 (2007): 5443–5452, https://doi.org/10.1074/jbc.M611342200.

15. H. J. Bennett, J. B. Davenport, R. F. Collins, A. W. Trafford, C. Pinali, and A. Kitmitto, "Human Junctophilin-2 Undergoes a Structural Rearrangement Upon Binding PtdIns(3,4,5)P3 and the S101R Mutation Identified in Hypertrophic Cardiomyopathy Obviates This Response," *Biochemical Journal* 456 (2013): 205–217, https://doi.org/10.1042/BJ20130591.

16. H. Ma, Y. Lou, W. H. Lin, and H. W. Xue, "MORN Motifs in Plant PIPKs Are Involved in the Regulation of Subcellular Localization and Phospholipid Binding," *Cell Research* 16 (2006): 466–478, https://doi.org/10.1038/sj.cr.7310058.

17. K. Hu, J. Johnson, L. Florens, et al., "Cytoskeletal Components of an Invasion Machine—The Apical Complex of *Toxoplasma gondii*," *PLoS Pathogens* 2 (2006): e13, https://doi.org/10.1371/journal.ppat.0020013.

18. M.-J. Gubbels, S. Vaishnava, N. Boot, J.-F. Dubremetz, and B. Striepen, "A MORN-Repeat Protein Is a Dynamic Component of the *Toxoplasma gondii* Cell Division Apparatus," *Journal of Cell Science* 119 (2006): 2236–2245, https://doi.org/10.1242/jcs.02949.

19. S. Sajko, I. Grishkovskaya, J. Kostan, et al., "Structures of Three MORN Repeat Proteins and a Re-Evaluation of the Proposed Lipid-Binding Properties of MORN Repeats," *PLoS One* 15 (2020): e0242677, https://doi.org/10.1371/journal.pone.0242677.

20. A. Roy, A. Kucukural, and Y. Zhang, "I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction," *Nature Protocols* 5 (2010): 725–738, https://doi.org/10.1038/nprot.2010.5.

21. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang, "The I-TASSER Suite: Protein Structure and Function Prediction," *Nature Methods* 12 (2015): 7–8, https://doi.org/10.1038/nmeth.3213.

22. J. Jumper, R. Evans, A. Pritzel, et al., "Highly Accurate Protein Structure Prediction With AlphaFold," *Nature* 596 (2021): 583–589, https://doi.org/10.1038/s41586-021-03819-2.

23. The UniProt Consortium, "UniProt: The Universal Protein Knowledgebase in 2023," *Nucleic Acids Research* 51, no. D1 (2023): D523–D531, https://doi.org/10.1093/nar/gkac1052.

24. I. Grossman-Haham, N. Coudray, Z. Yu, et al., "Structure of the Radial Spoke Head and Insights Into its Role in Mechanoregulation of Ciliary Beating," *Nature Structural & Molecular Biology* 28 (2021): 20–28, https://doi.org/10.1038/s41594-020-00519-9.

25. J. R. Wilson, C. Jing, P. A. Walker, et al., "Crystal Structure and Functional Analysis of the Histone Methyltransferase SET7/9," *Cell* 111 (2002): 105–115, https://doi.org/10.1016/S0092-8674(02)00964-9.

26. Molecular Operating Environment (MOE), (Montreal, Canada: Chemical Computing Group ULC, 2024).

27. Y. Zhang, "TM-Align: A Protein Structure Alignment Algorithm Based on the TM-Score," *Nucleic Acids Research* 33 (2005): 2302–2309, https://doi.org/10.1093/nar/gki524.

28. M. J. Abraham, T. Murtola, R. Schulz, et al., "GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism From Laptops to Supercomputers," *SoftwareX* 1–2 (2015): 19–25, https://doi.org/10.1016/j.softx.2015.06.001.

29. D. van der Spoel, P. J. van Maaren, and H. J. C. Berendsen, "A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for Use With a Reaction Field," *Journal of Chemical Physics* 108 (1998): 10220–10230, https://doi.org/10.1063/1.476482.

30. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular Dynamics With Coupling to an External Bath," *Journal of Chemical Physics* 81 (1984): 3684–3690, https://doi.org/10.1063/1.448118.

31. M. Parrinello and A. Rahman, "Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method," *Journal of Applied Physics* 52 (1981): 7182–7190, https://doi.org/10.1063/1.328693.

32. K. Lindorff-Larsen, S. Piana, K. Palmo, et al., "Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field," *Proteins: Structure, Function, and Bioinformatics* 78, no. 8 (2010): 1950–1958, https://doi.org/10.1002/prot.22711.

33. P. P. Ewald, "Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale," *Annalen der Physik* 369 (1921): 253–287, https://doi.org/10.1002/andp.19213690304.

34. W. Kabsch and C. Sander, "Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features," *Biopolymers* 22 (1983): 2577–2637, https://doi.org/10.1002/bip.360221211.

35. M. F. Adasme, K. L. Linnemann, S. N. Bolz, et al., "PLIP 2021: Expanding the Scope of the Protein–Ligand Interaction Profiler to DNA and RNA," *Nucleic Acids Research* 49 (2021): W530–W534, https://doi.org/10.1093/nar/gkab294.

36. R. Gowers, M. Linke, J. Barnoud, et al., "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations," in *Proceedings of the 15th Python in Science Conference*, Los Alamos National Laboratory (Los Alamos, NM (USA), 2016), 98–105.

37. R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, et al., "MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories," *Biophysical Journal* 109 (2015): 1528–1532, https://doi.org/10.1016/j.bpj.2015.08.015.

38. J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science & Engineering* 9 (2007): 90–95, https://doi.org/10.1109/MCSE.2007.55.

39. W. Umphrey, A. Dalke, and K. Schulten, "VMD-Visual Molecular Dynamics," *Journal of Molecular Graphics* 14 (1996): 33–38.

40. M. M. Harding, "Small Revisions to Predicted Distances Around Metal Sites in Proteins," *Acta Crystallographica Section D: Biological Crystallography* 62 (2006): 678–682, https://doi.org/10.1107/S0907444906014594.

41. J. Zhou, H. Liu, Y. Lin, and J. Zhao, "Membrane Occupation and Recognition Nexus (MORN) Motif Controls Protein Localization and Function," *FEBS Letters* 596, no. 14 (2022): 1839–1850, https://doi.org/10.1002/1873-3468.14378.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.