PLOS ONE

RESEARCH ARTICLE

Comprehensive in silico analysis and molecular dynamics of the superoxide dismutase 1 (SOD1) variants related to amyotrophic lateral sclerosis

Gabriel Rodrigues Coutinho Pereira[1], Bárbara de Azevedo Abrahim Vieira[2], Joelma Freire De Mesquita[1]*

1 Department of Genetics and Molecular Biology, Bioinformatics and Computational Biology Laboratory, Federal University of the State of Rio de Janeiro (UNIRIO), Rio de Janeiro, Rio de Janeiro, Brazil, **2** Faculty of Pharmacy, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro, Brazil

* joelma.mesquita@unirio.br

Abstract

Amyotrophic Lateral Sclerosis (ALS) is the most frequent motor neuron disorder, with a significant social and economic burden. ALS remains incurable, and the only drugs approved for its treatments confers a survival benefit of a few months for the patients. Missense mutations in superoxide dismutase 1 (SOD1), a major cytoplasmic antioxidant enzyme, has been associated with ALS development, accounting for 23% of its familial cases and 7% of all sporadic cases. This work aims to characterize *in silico* the structural and functional effects of SOD1 protein variants. Missense mutations in SOD1 were compiled from the literature and databases. Twelve algorithms were used to predict the functional and stability effects of these mutations. ConSurf was used to estimate the evolutionary conservation of SOD1 amino-acids. GROMACS was used to perform molecular dynamics (MD) simulations of SOD1 wild-type and variants A4V, D90A, H46R, and I113T, which account for approximately half of all ALS-SOD1 cases in the United States, Europe, Japan, and United Kingdom, respectively. 233 missense mutations in SOD1 protein were compiled from the databases and literature consulted. The predictive analyses pointed to an elevated rate of deleterious and destabilizing predictions for the analyzed variants, indicating their harmful effects. The ConSurf analysis suggested that mutations in SOD1 mainly affect conserved and possibly functionally essential amino acids. The MD analyses pointed to flexibility and essential dynamics alterations at the electrostatic and metal-binding loops of variants A4V, D90A, H46R, and I113T that could lead to aberrant interactions triggering toxic protein aggregation. These alterations may have harmful implications for SOD1 and explain their association with ALS. Understanding the effects of SOD1 mutations on protein structure and function facilitates the design of further experiments and provides relevant information on the molecular mechanism of pathology, which may contribute to improvements in existing treatments for ALS.

Introduction

Amyotrophic Lateral Sclerosis (ALS) is a highly disabling, progressive, and fatal neurodegenerative disorder characterized by injury and death of upper motor neurons in the cerebral cortex and lower motor neurons in the brain stem and spinal cord [1]. ALS is characterized by the progressive loss of voluntary motor activity, which impairs the patient's ability to work and perform daily activities [2], usually resulting in the patient's death due to respiratory paralysis within three to four years after the symptoms' onset [3]. The patient's illness-related absence from work, caregiver needs, and the medical costs of ALS represent a significant economic burden, which has been estimated at over one billion dollars a year only in the United States [4]. ALS is the most frequent motor neurodegenerative disorder in adults [5], affecting more than 220,000 people worldwide. Due to population aging, the number of people affected by ALS is projected to increase by 69% in the next 20 years, particularly among developing countries [6].

Despite its importance, ALS remains uncurable [7]. The only drugs approved for the treatment of ALS, Riluzole, and Edaravone, confers a survival benefit of a few months for the patients [8]. Thus, there is an urgent need to develop more effective treatments for diseases such as ALS, but these will only come with a deep understanding of their causes and associated mechanisms [9]. Due to its heterogeneous and complex nature, the exact pathway that causes ALS degeneration is not yet fully understood [10]. However, glutamate cytotoxicity, inflammatory pathway, oxidative stress, and protein aggregation are among the main mechanisms associated with ALS development [11].

ALS is classified as sporadic (sALS) and familial (fALS). Most ALS cases are sporadic (95–90%), while 5–10% are familial [12]. Mutations in the *SOD1 gene* are frequent and important causes of ALS, accounting for 23% of the fALS cases, and approximately 7% of all sALS cases [13]. The SOD1 gene codes superoxide dismutase 1 (SOD1), a major cytoplasmic antioxidant enzyme. SOD1 is a Cu/Zn metalloenzyme that catalyzes the dismutation of superoxide radicals (O_2-.) to hydrogen peroxide (H_2O_2) and molecular oxygen (O_2) [14]. This mechanism protects cells from the harmful effects of superoxide radicals and, consequently, from oxidative stress [15]. More than 185 SOD1 mutations have already been associated with ALS [13]. Although these mutations affect different locations within the SOD1 structure [16], many of them lead to protein-aggregation and increased oxidative damage, which are central events in ALS pathogenicity [17, 18].

The computational approach has become widely used to characterize the effects of mutations in proteins and to identify potentially deleterious mutations, particularly because of its relatively high cost/efficiency and accuracy [19]. We applied computational predictions to the study of SOD1 protein variants following the methodology previously established by our group [16, 19–22]. This work aims to characterize the effects of these variants on the SOD1 structure, which could assist the design of future experiments and provide relevant information on the molecular mechanism of pathology that may lead to improvements in existing treatments for ALS [22–24].

Materials and methods

Dataset

The structure, sequence, and natural variants of SOD1 protein were retrieved from the Protein Data Bank (PDB) [25], UniProt [26], OMIM [27], ALSoD [28], dbSNP [29], ClinVar [30], and PubMed [27].

Functional and stability prediction analysis

The functional and stability effects of the SOD1 protein variants were predicted using the following algorithms: SNAP2, SNPs&GO, PolyPhen2, PMUT, PhD-SNP, MutPred2 [19], Panther [31], Provean [32], MAPP [33], I-Mutant3.0 [34], SIFT and SNPEffect4.0 [35, 36].

Evolutionary conservation analysis

The evolutionary conservation analysis of human SOD1 was performed using the ConSurf server, which estimated the evolutionary conservation degree of each SOD1 amino acid [37]. The following parameters were selected for this analysis: PDB ID: 2C9V; Chain identifier: A; homologous search algorithm: PSI-BLAST; number of iterations: 3; E-value cut-off: 0.0001; protein database: UniProt; reference sequence: closest; number of reference sequences selected: 150; maximum sequence identity: 95%; minimum identity for counterparts: 35%; alignment method: MAFFT-L-INS-i; calculation method: Bayesian; evolutionary substitution model: best model (standard).

Molecular dynamics simulations

Molecular Dynamics (MD) simulations of the wild-type SOD1 protein and its variants A4V, D90A, H46R, and I113T were performed using the GROMACS 2018.8 package [38]. The software Visual Molecular Dynamics 1.9.2 was used to induce the mutations A4V, D90A, H46R, and I113T on the crystallographic structure of wild-type SOD1 (PDB ID: 2C9V) [25]. The variants A4V, D90A, H46R, and I113T were selected for the MD simulations because they account for approximately half of all ALS-SOD1 cases in the United States [39], Europe, Japan, and United Kingdom [15], respectively.

The MD simulations were performed in triplicates according to the methodology described by Pereira, Tellini, and De Mesquita, 2019 [22]. The force-field AMBER99SB-ILDN was selected for the simulations. The molecules were solvated in a dodecahedral box with TIP3P water molecules, neutralized by adding Na^+Cl^- ions, and minimized using the steepest descent method. After system minimization, an NVT ensemble (constant number of particles, volume, and temperature) followed by an NPT ensemble (constant number of particles, pressure, and temperature) were performed for 100ps at a pressure of 1atm and temperature of 300K. The V-rescale thermostat and Parrinello-Rahman barostat were selected for the NVT and NPT ensembles. The production simulations were further performed at a temperature of 300K for a duration of 300ns using the LINCS (linear constraint solver) and PME (particle mesh Ewald) algorithms.

The MD trajectories were analyzed using the following GROMACS distribution programs: *gmx rms*, *gmx rmsf*, *gmx gyrate*, *gmx dssp*, *gmx sasa*, and *gmx hbond*. These MD analyses generated parameters values for root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), B-factor, radius of gyration (Rg), solvent accessible surface area (SASA), and secondary structure (SS). The MD trajectories were also submitted to the Bio3D library in R software [40], which was used to perform principal component analysis (PCA) of the wild-type SOD1 and its variants. PCA was performed for the Cα atoms of protein structure, and their Cartesian coordinates were used to generate the covariance matrices.

Results

Dataset

The complete crystallographic structure of wild-type SOD1 retrieved from the Protein Data Bank [PDB ID: 2C9V] is shown in Fig 1A. The 2C9V structure was selected because it has the
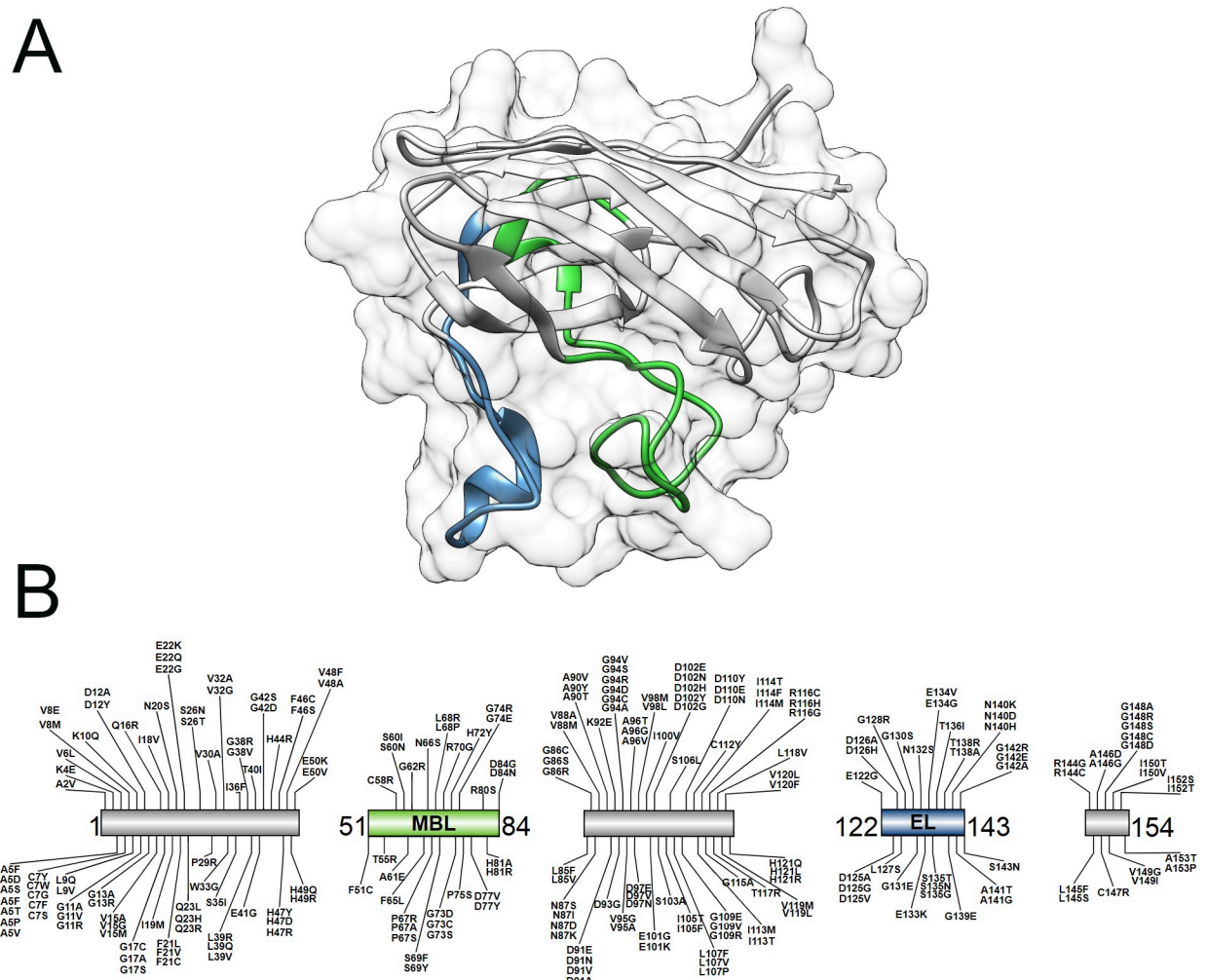
Fig 1. Crystallographic structure and schematic representation of SOD1 protein and its variants. The metal-binding (amino-acids 51–84) and electrostatic loops (amino-acids 122–143) of SOD1 are represented in green and blue, respectively. (A) Secondary structure representation of the crystallographic structure of SOD1 (PDB ID: 2C9V). The protein's surface is also shown. (B) Schematic representation of the SOD1 protein sequence and its 233 variants compiled from databases and literature.

https://doi.org/10.1371/journal.pone.0247841.g001

highest atomic resolution (1.07Å) among the SOD1 structures available in PDB [25]. The complete amino-acid sequence of the SOD1 protein was retrieved from the UniProt database [UniProt ID: P00441] [26]. SOD1 contains an 8-stranded beta-barrel motif and several loops. Two loop elements termed metal-binding (amino-acids 51–84) and electrostatic loops (amino-acids 122–143) project from the beta-barrel (Fig 1A) and are important for the SOD1 enzymatic activity [39]. The electrostatic loop is composed of positively charged amino-acids that guide superoxide radicals into the SOD1 active site where the copper ion is located [41]. The metal-binding loop coordinates the binding of Zinc and Copper ions [42, 43], which is structurally important for SOD1 [44]. These loops together shape the active site pocket [43].

Two hundred and thirty-three (233) mutations in SOD1 protein were compiled from the databases and literature consulted (S1 Table). As shown in Fig 1B, SOD1 mutations occur all over the protein structure. Thirty-one (31) mutations were found occurring at the metal-binding loop, while 30 mutations were found at the electrostatic loop of SOD1.
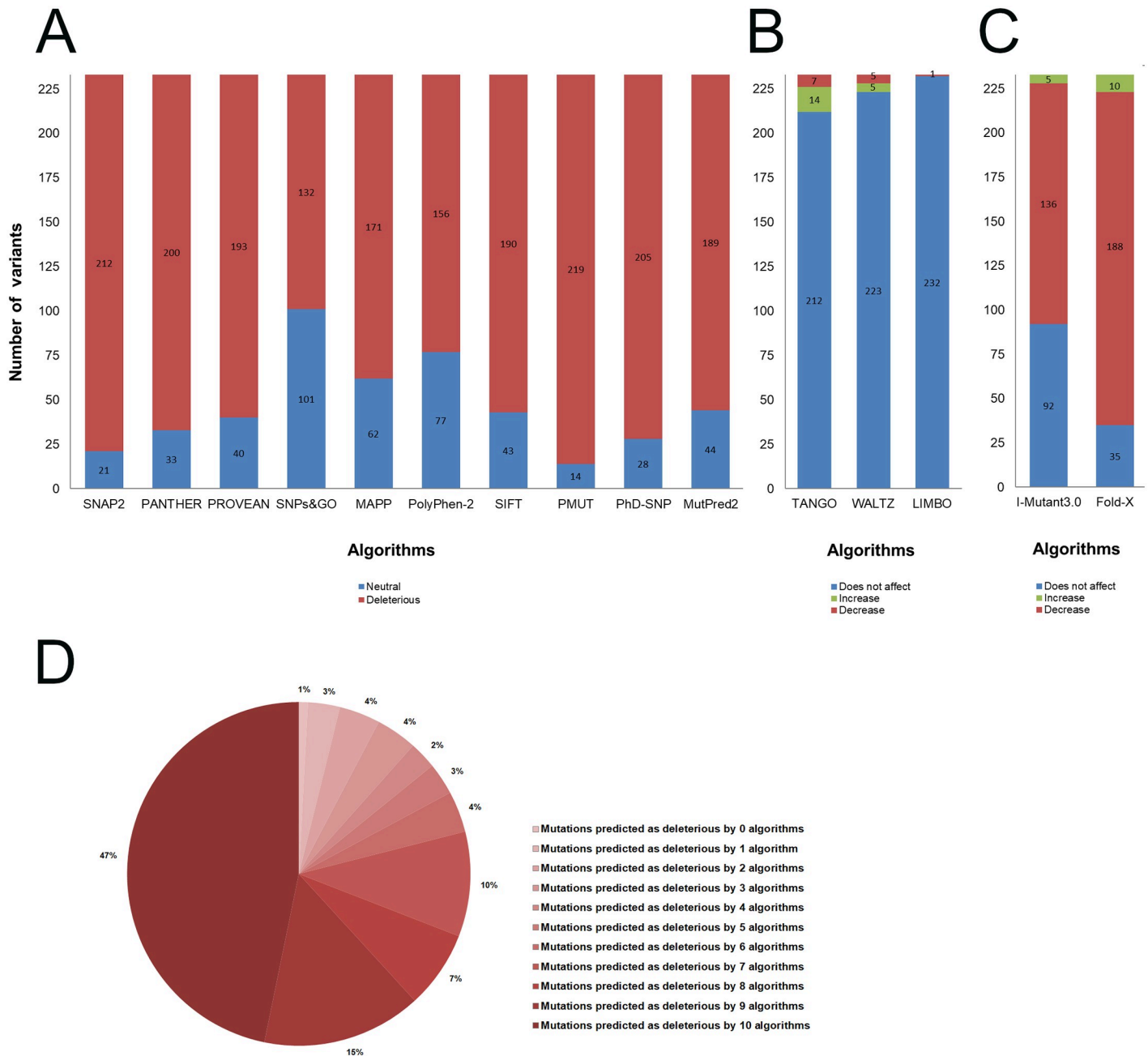
Fig 2. Functional and stability predictions of SOD1 protein variants. The SOD1 variants compiled from the literature and databases (233) were analyzed using functional and stability prediction algorithms. (A) Functional prediction using ten different algorithms. The bar plot indicates the number of neutral (blue) and deleterious (red) predictions for each algorithm. (B) Prediction of aggregation propensity (TANGO), amyloid propensity (WALTZ), and chaperone binding (LIMBO) using the SNPEffect4.0 algorithm. The bar plot indicates the number of variants that increase (green), decrease (red), and do not affect (blue) the parameter analyzed. (C) Stability prediction using the FoldX and I-Mutant3.0 algorithms. The bar plot indicates the number of variants that increase (green), decrease (red), and do not affect (blue) protein stability. (D) The circle chart indicates the percentage of variants predicted as deleterious by zero to ten algorithms.

https://doi.org/10.1371/journal.pone.0247841.g002

Functional and stability prediction analysis

As shown in Fig 2A, all the algorithms used predicted more than 50% of the SOD1 mutations as deleterious. SNPs&GO was the algorithm with the lowest number of deleterious predictions, while PMUT was the algorithm with the highest number of deleterious predictions (Fig 2A).

As shown in Fig 2D, 47% of the analyzed variants were predicted as deleterious by all the ten functional prediction algorithms used, while only 1% of them were predicted as neutral by all the algorithms. Moreover, the SOD1 protein variants presented an elevated rate of deleterious predictions (Fig 2D). This analysis also showed that the mutations with the highest rates of deleterious predictions were concentrated in the metal-binding (amino-acids 51–84) and electrostatic (amino-acids 122–143) loops of SOD1 (S2 Table).

According to the SNPEffect4.0, twenty-one variants were predicted to affect the SOD1 aggregation tendency, ten variants were predicted to affect amyloid propensity, and one variant was predicted to affect chaperone binding (Fig 2B). The variants predicted to affect protein aggregation mostly occur in the N and C-terminal portions of SOD1, while the variants predicted to affect amyloid propensity mostly occur in its N-terminal portion (S3 Table).

As shown in Fig 2C, each algorithm used predicted more than 50% of the SOD1 variants as decreasing protein stability. This analysis also showed that 49% of these variants were predicted to decrease protein stability by I-Mutant and Fold-X, while less than 6% were predicted by both algorithms to increase or not affect this feature (S3 Table). The I-Mutant and Fold-X predictions were concordant for 54.5% of the analyzed variants, but only 2.2% of all predictions were opposed (i.e. increase and decrease for the same variant).

The variants A4V, D90A, H46R, and I113T, which are the most frequent ALS-related mutations in SOD1 [15, 39], were predicted as deleterious by 90%, 50%, 100%, and 100% of the algorithms used, respectively (S1 Fig). The variants A4V, D90A, and I113T were also predicted to decrease protein stability by Fold-X and I-Mutant, while the variant H46R was only predicted as destabilizing by Fold-X. Furthermore, the A4V variant was predicted to increase protein aggregation. None of these variants were predicted to affect amyloid propensity and chaperone binding (S3 Table).

Evolutionary conservation analysis

The crystallographic structure of wild-type SOD1 (PDB ID: 2C9V) was submitted to the ConSurf server, which estimated the evolutionary conservation score (ConSurf-score) of each protein's amino acid. ConSurf performs multiple-sequence alignments to compare a given amino acid sequence to its homologous sequences and estimate their evolutionary conservation. ConSurf is useful for identifying functionally important regions on proteins, as these are usually conserved throughout evolution [37]. The computed ConSurf-score was projected on the protein's surface (Fig 3A). The SOD1 amino-acids were colored according to their ConSurf-score following a coloring-code scheme, which varies from (1) cyan and variable to (9) maroon and conserved.

As shown in Fig 3B, 59% of the analyzed SOD1 variants occur in conserved positions (i.e. ConSurf-score ≥ 7), while only 17% of them occur in variable positions (i.e. ConSurf-score ≤ 3). This analysis also showed that variants affecting highly conserved positions (i.e. ConSurf-score = 9) mainly occur at the metal-binding (amino-acids 51–84) and electrostatic (amino-acids 122–143) loops of SOD1. Among the most frequent ALS-related mutations in SOD1 [15, 39], the variant I113T was the only variant predicted to affect a variable region. The variants A4V and H46R were predicted to affect conserved amino-acids of SOD1, while the variant D90A was predicted to affect an average conserved position (S2 Fig).

Molecular dynamics simulations

MD is a computational method that predicts the time-dependent motion of an atomic system by solving Newton's equations of motion [45]. The MD simulations can be used to reproduce the real behavior of a protein in its environment, which allows the study of relevant
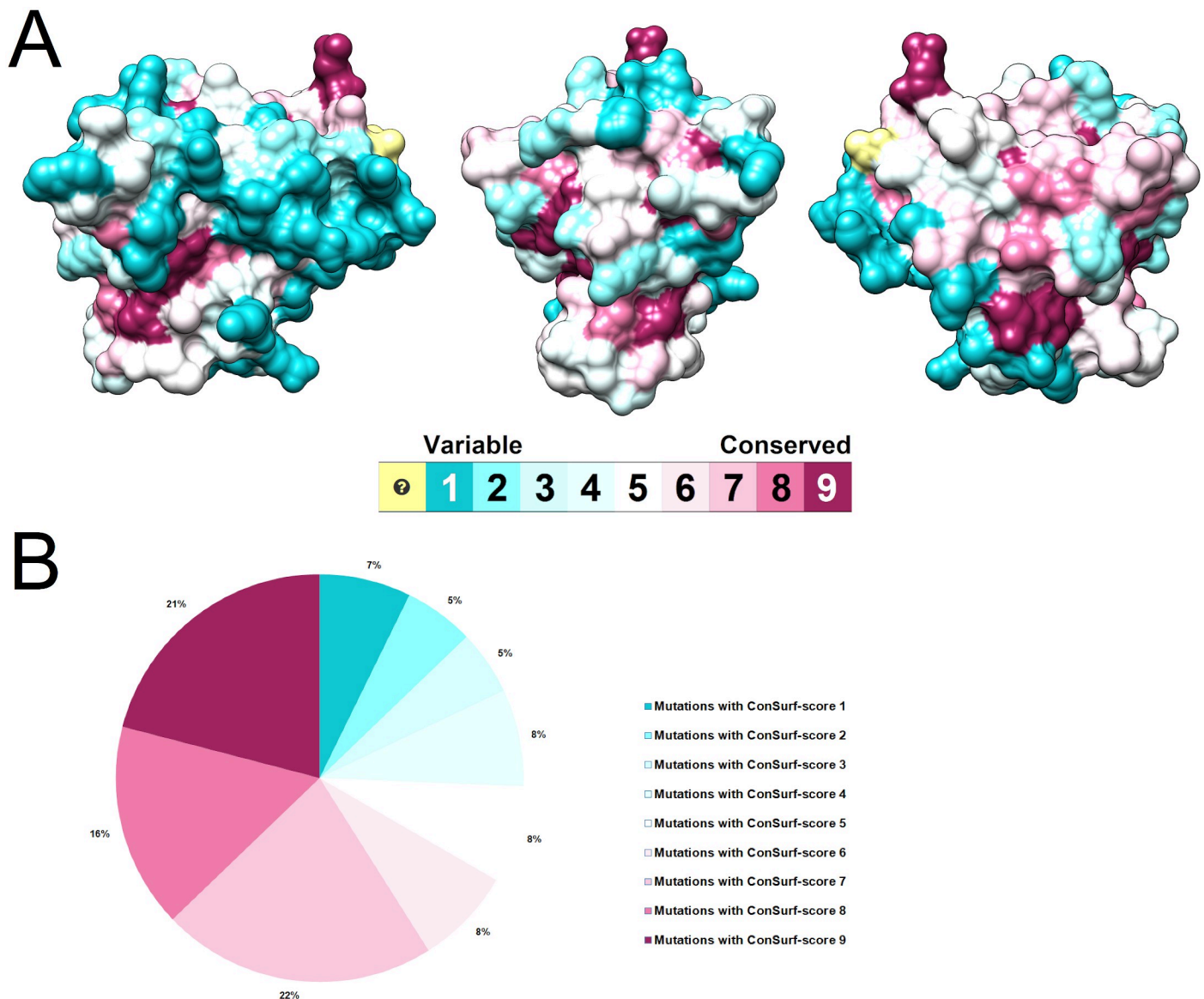
Fig 3. Evolutionary conservation analysis of SOD1 protein. (A) The evolutionary conservation profile of SOD1 is shown in three different angles, displaying different sides of the protein. Each SOD1 amino-acid is represented as a surface model and colored according to its ConSurf conservation score. The color-coding bar shows the ConSurf coloring scheme, which varies from (1) cyan and variable to (9) maroon and conserved. The amino-acids colored yellow did not receive conservation scores due to insufficient data. (B) The circle chart indicates the percentage of variants affecting amino-acids predicted as variable to conserved. The circle chart is also colored according to the ConSurf coloring scheme.

biomolecular processes, including conformational change, protein folding, and ligand binding. These simulations can also be used to predict the effects of perturbations such as mutation, phosphorylation, and protonation on the protein's structure [46]. During an MD, the inter-atomic interactions are calculated at every simulation step and the resulting atomic positions are registered in a trajectory file, which provided detailed information on changes in protein conformation that can be used to assess several structural parameters [19].

We performed MD simulations of the wild-type SOD1 (PDB ID: 2C9V) and its most frequent ALS-related variants, i.e. A4V, D90A, H46R, and I113T [15, 39], to better understand the impact of these amino-acid substitutions on protein structure. The following structural

parameters were analyzed from the MD trajectories: RMSD, RMSF, B-factor, Rg, SASA, SS, and PCA.

RMSD is commonly used in MD simulations to measure the spatial differences between a starting structure and its subsequent coordinates computed over time [47]. This parameter is useful to analyze the time-dependent motion of protein structures and to determine their structural convergence throughout the simulation [19]. RMSD values were calculated from the total number of protein conformations computed in the MD trajectories. As shown in Fig 4, triplicates for wild-type SOD1 and its variants presented a similar behavior throughout the simulation, given their respective means and confidence intervals. A sudden increase in RMSD values was observed at the beginning of all simulations. This was followed by the establishment of a plateau in RMSD values after approximately 150ns, which suggests that the protein structures float around average stable conformations [16]. The initial effects of the trajectories, i.e. those occurring until 150ns, were discounted in further analyses for meaningful comparison [19, 34].

RMSF is a measure of the structural displacement of an amino-acid from its average position throughout the simulation [16]. RMSF is useful to assess local flexibility in protein structures, allowing the identification of flexible and rigid regions [48]. The RMSF analysis pointed to a flexibility increase at the metal-binding and electrostatic loops of the variants A4V and D90A, in addition to a flexibility decrease in the same regions of H46R (Fig 5).

Flexibility can also be assessed during an MD simulation by analyzing the B-factor, a temperature-displacement factor. B-factor is a measure of the structural displacement of an amino-acid around its average position due to thermal vibrations [49]. The B-factor values computed for each SOD1 amino-acid were projected on the protein surface, providing an interesting three-dimensional representation of structural flexibility [19]. The B-factor distribution over a protein structure also provides an important indicator of its dynamics [50]. As shown in Fig 6, the variants A4V and D90A presented increased flexibility at their metal-binding and electrostatic loops, while the variant H46R presented decreased flexibility at those regions when compared to the wild-type SOD1. The B-factor analysis suggested flexibility alterations in the variants A4V, D90A, and H46R at regions similar to those found altered in the RMSF analysis.

PCA is also known as essential dynamics (ED) when applied to analyze MD simulations [19]. ED linear transforms the high-dimensional and complex data contained in a molecular trajectory into a low-dimensional space in which the large-scale protein motions occur [51], reducing the number of dimensions needed to describe protein dynamics. This statistical technique systematically filters the observed motions from the largest to smallest spatial-scale in a molecular trajectory using a covariance matrix constructed with the Cartesian coordinates representing atomic positions [52]. ED allows separating the essential from the remaining protein motions. The essential motions (i.e. the largest-scales motions) are usually biologically relevant movements such as opening, closing, and flexing, while the remaining motions describe small irrelevant local fluctuations [20]. The essential motions are often confined into the first two PCA modes (principal components) [53].

The projections for the MD trajectories of wild-type SOD1 and its variants into the subspace spanned by PC1 and PC2 are shown in Fig 7. The ED analysis suggested that the first two principal components (PC1 and PC2) capture the dominant motions, accounting for 62.8%, 73.3%, 70.7%, 35.3%, and 62.1% of the total variance for wild-type SOD1 and its variants A4V, D90A, H46R, and I113T, respectively. As shown in Fig 7, the variants A4V, D90A, and I113T occupied a higher area in conformational space, while the variant H46R occupied a smaller area. Changes in cluster shape were also observed in the conformational space of all

Fig 4. RMSD analysis of wild-type SOD1 and its variants. The RMSD values computed for the backbone atoms of wild-type SOD1 and its variants at 300K are shown over time. The means (solid lines) and confidence intervals (smooth lines) are displayed for the triplicates. (A) The wild-type is represented in black, while variant A4V is represented in red. (B) The wild-type is represented in black, while variant D90A is represented in blue. (C) The wild-type is represented in black, while variant H46R is represented in green. (D) The wild-type is represented in black, while variant I113T is represented in yellow.
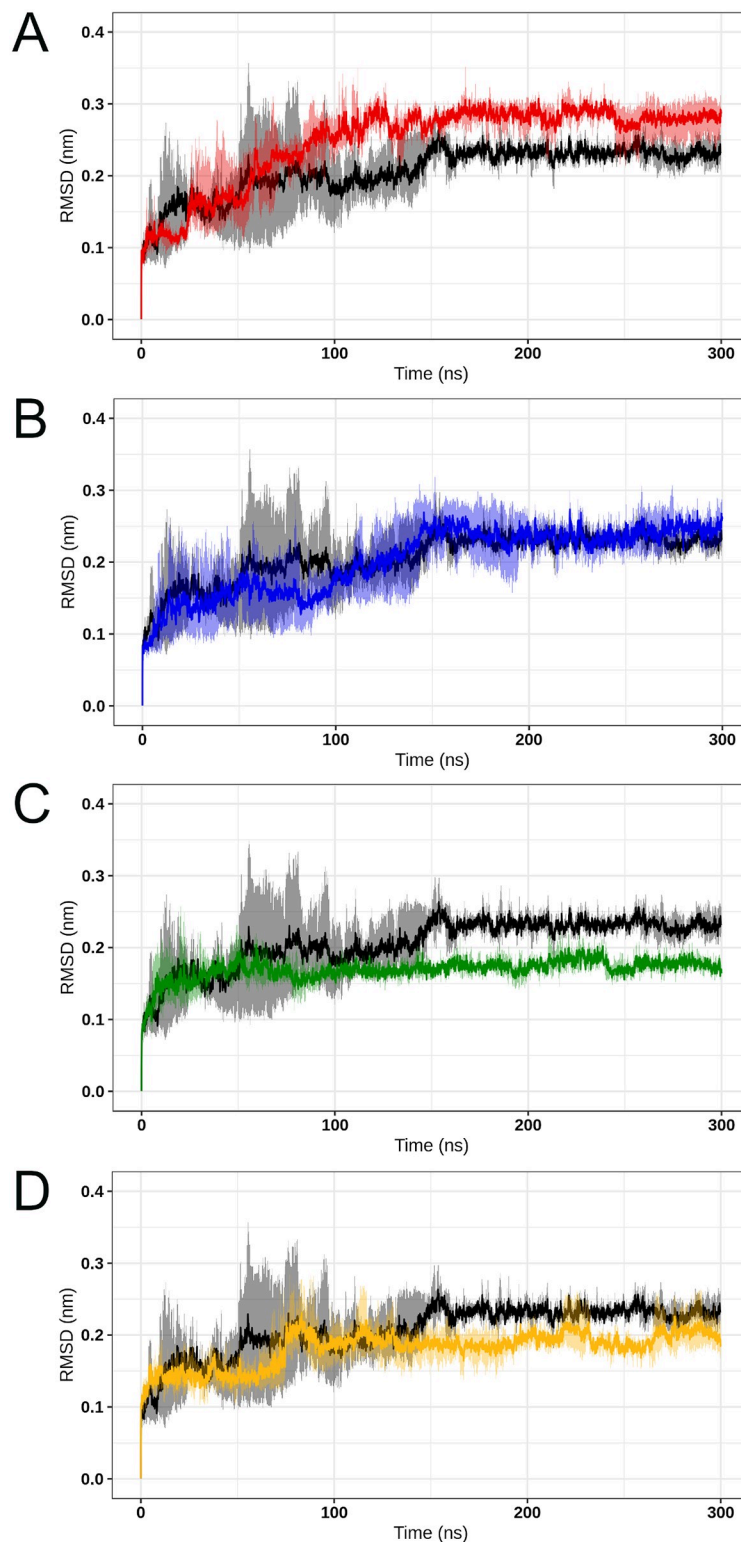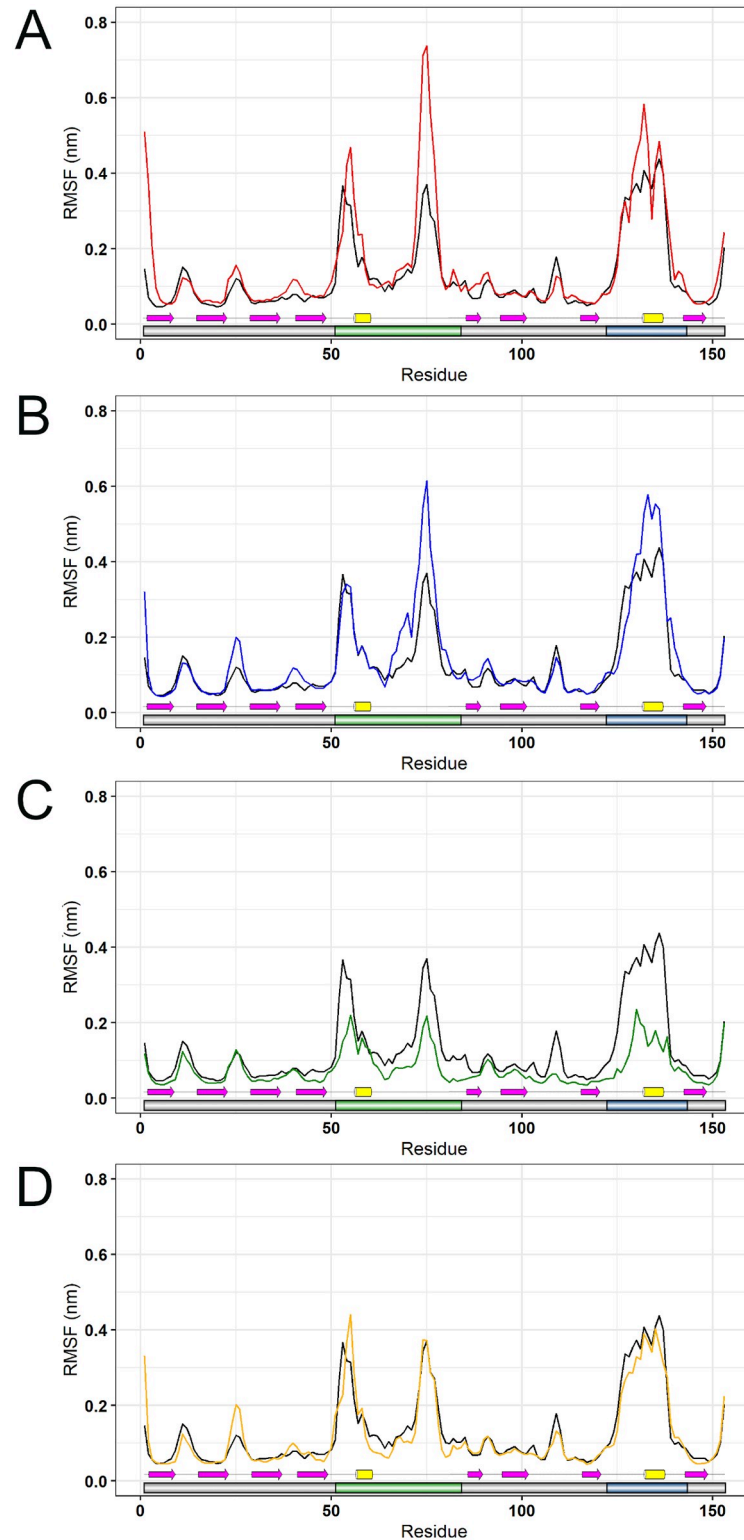
Fig 5. RMSF analysis of wild-type SOD1 and its variants. The RMSF values for each residue of wild-type SOD1 and its variants at 300K are shown as a line plot. Schematic representations of the SOD1 functional regions and secondary structure are shown to further comparison. The metal-binding and electrostatic loops of SOD1 are highlighted in green and blue, respectively. β-strands are represented by yellow barrels, α-helices are represented by magenta arrows, and the coils are represented by the thin black lines. (A) The wild-type is represented in black, while variant A4V is

represented in red. (B) The wild-type is represented in black, while variant D90A is represented in blue. (C) The wild-type is represented in black, while variant H46R is represented in green. (D) The wild-type is represented in black, while variant I113T is represented in yellow.

variants when compared to the wild-type. The ED analysis thus pointed to alterations in the overall essential dynamics of all variants.

We also analyzed the RMSF contribution of each protein amino-acid for the first two principal components. RMSF values for the PC1 and PC2 projections are shown in Figs 8 and 9, respectively. The RMSF contribution to PC1 pointed to increased essential mobility at the metal-binding loops of the A4V, D90A, and I113T (Fig 8), while the RMSF contribution for the PC2 pointed to essential mobility alterations at the metal-binding and electrostatic loop of the A4V and H46R variants (Fig 9). As shown in Fig 8, these alterations are particularly high at the metal-binding loop of A4V and D90A variants.

Rg is commonly used to measure the structural displacement of protein atoms from their common center of mass throughout the simulation, thus, providing comprehensive information on protein compactness over time [19]. The Rg values computed for the wild-type SOD1 and its variants pointed to an initial moment of structural instability, which is observed in the first half of all simulations (Fig 10). After this time, all the analyzed protein structures presented steady Rg values, which further suggests stable protein folding [54]. This analysis also
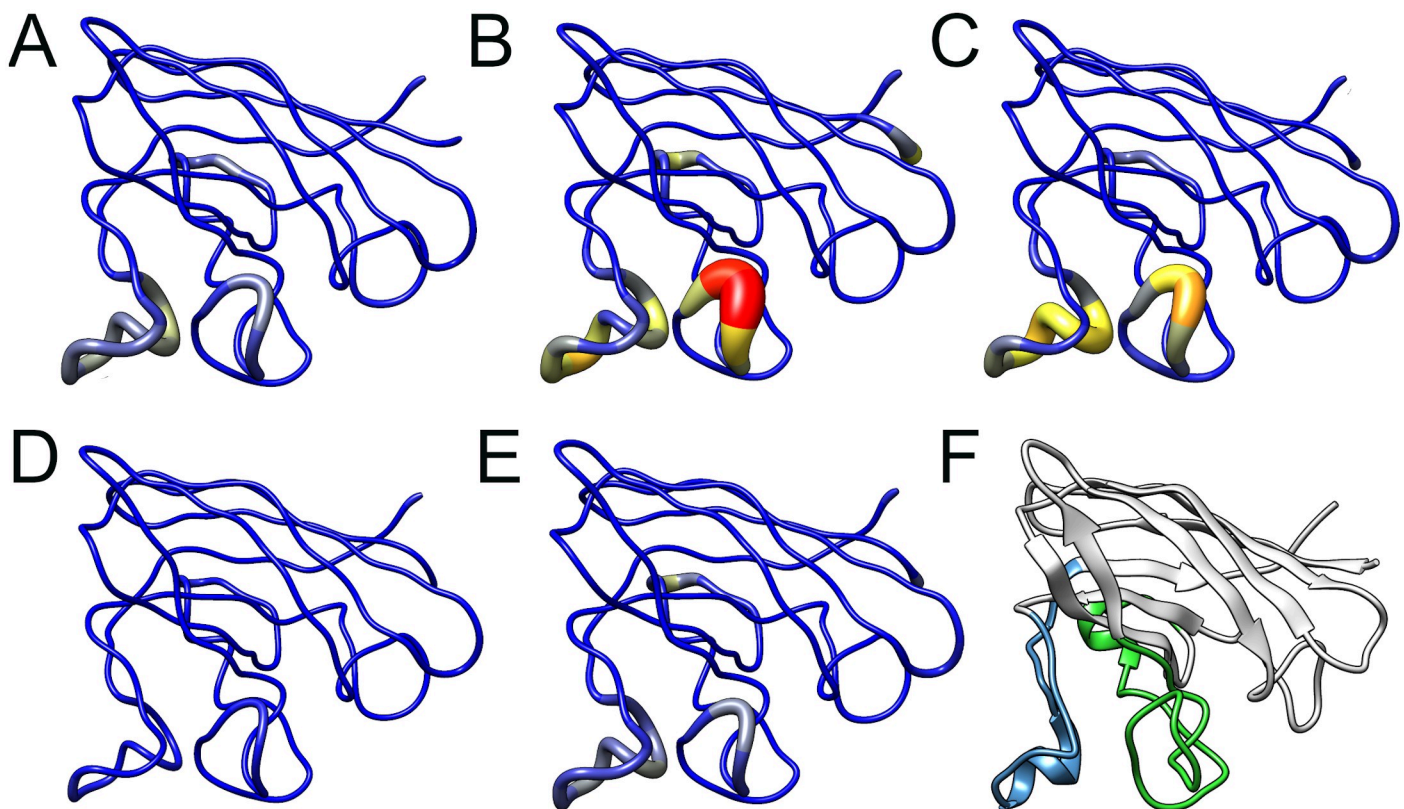
Fig 6. B-Factor analysis of wild-type SOD1 and its variants. The amino-acids residues of wild-type SOD1 (A) and its variants A4V (B), D90A (C), H46R (D), and I113T (E) are sized and colored according to their B-factor values, following a coloring-thickness scale that varies from blue and thin (rigid residues) to red and thick (flexible residues). The three-dimensional representation of SOD1 and its functional regions (F) are also shown to further comparison. The metal-binding and electrostatic loops are colored in green and blue, respectively.
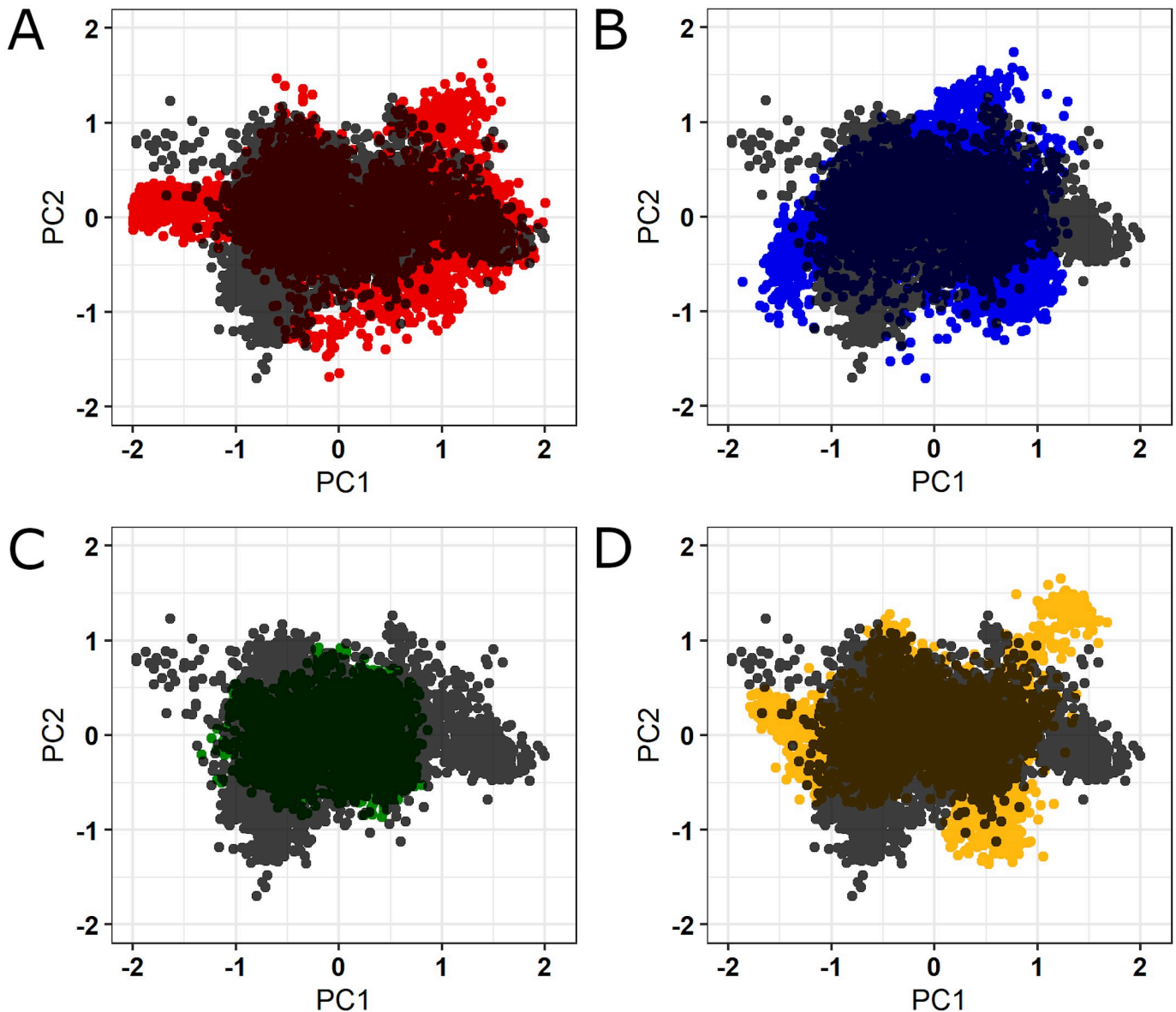
Fig 7. PCA for the wild-type SOD1 and its variants. Projections for the first two principal components extracted from the MD trajectories. (A) Comparison between the PCA projections for the wild-type SOD1 (black) and its variant A4V (red). (B) Comparison between the PCA projections for the wild-type SOD1 (black) and its variant D90A (blue). (C) Comparison between the PCA projections for the wild-type SOD1 (black) and its variant H46R (green). (D) Comparison between the PCA projections for the wild-type SOD1 (black) and its variant I113T (yellow).

https://doi.org/10.1371/journal.pone.0247841.g007

indicated that the average Rg values for the wild-type SOD1 (1.422±0.016) are similar to that of variants A4V (1.428±0.015), D90A (1.433±0.01), H46R (1.405±0.008), and I113T (1.421±0.01). This result suggested no compactness alterations in the analyzed variants.

SASA is a measure of the exposed surface in protein structures that can be accessible to solvent molecules [55]. SASA analysis thus provides relevant information into the protein's degree of exposure to its environment over time [56]. The SASA values computed throughout the simulations (Fig 11) pointed to an unstable behavior at the beginning of all simulations. After approximately 150ns, the SASA values assumed a steady behavior until the end of the MD trajectories. This analysis indicated that the average SASA values for the wild-type SOD1

Fig 8. RMSF contribution to PC1 of wild-type SOD1 and its variants. The RMSF contribution of each amino-acid to PC1 is shown as a line plot and projected on the corresponding structure. Schematic representations of the SOD1 functional regions and are shown to further comparison. The metal-binding and electrostatic loops of SOD1 are highlighted in green and blue, respectively. Each amino-acid of SOD1 wild-type and variants was colored and sized according to its RMSF contribution, following a coloring-thickness scale that varies from blue and thin (low fluctuations) to red and thick (high fluctuations). (A) RMSF contribution of wild-type SOD1 to PC1. (B) RMSF contribution of variant A4V to PC1. (C) RMSF contribution of variant D90A to PC1. (D) RMSF contribution of variant H46R to PC1. (E) RMSF contribution of variant I113T to PC1.
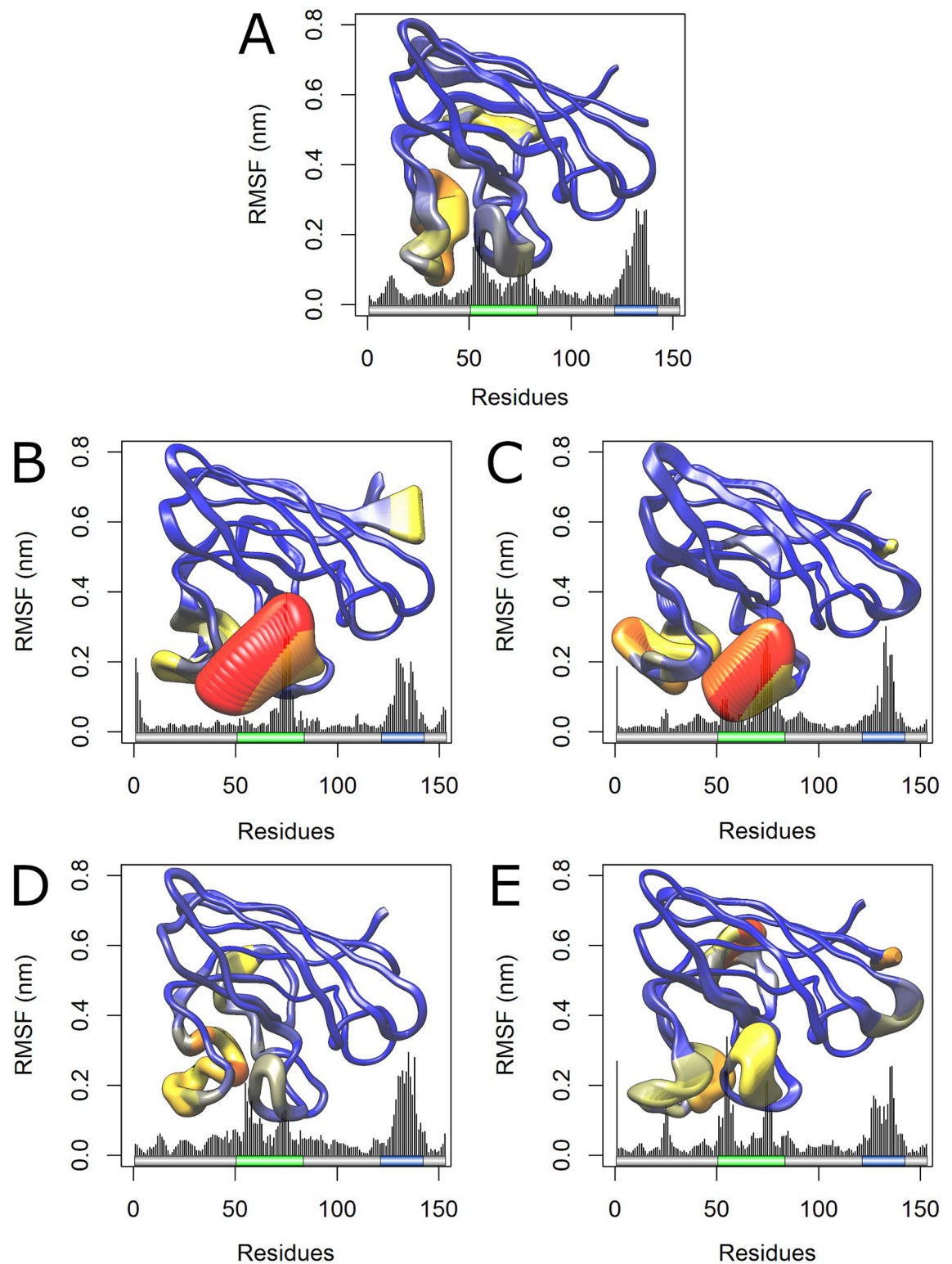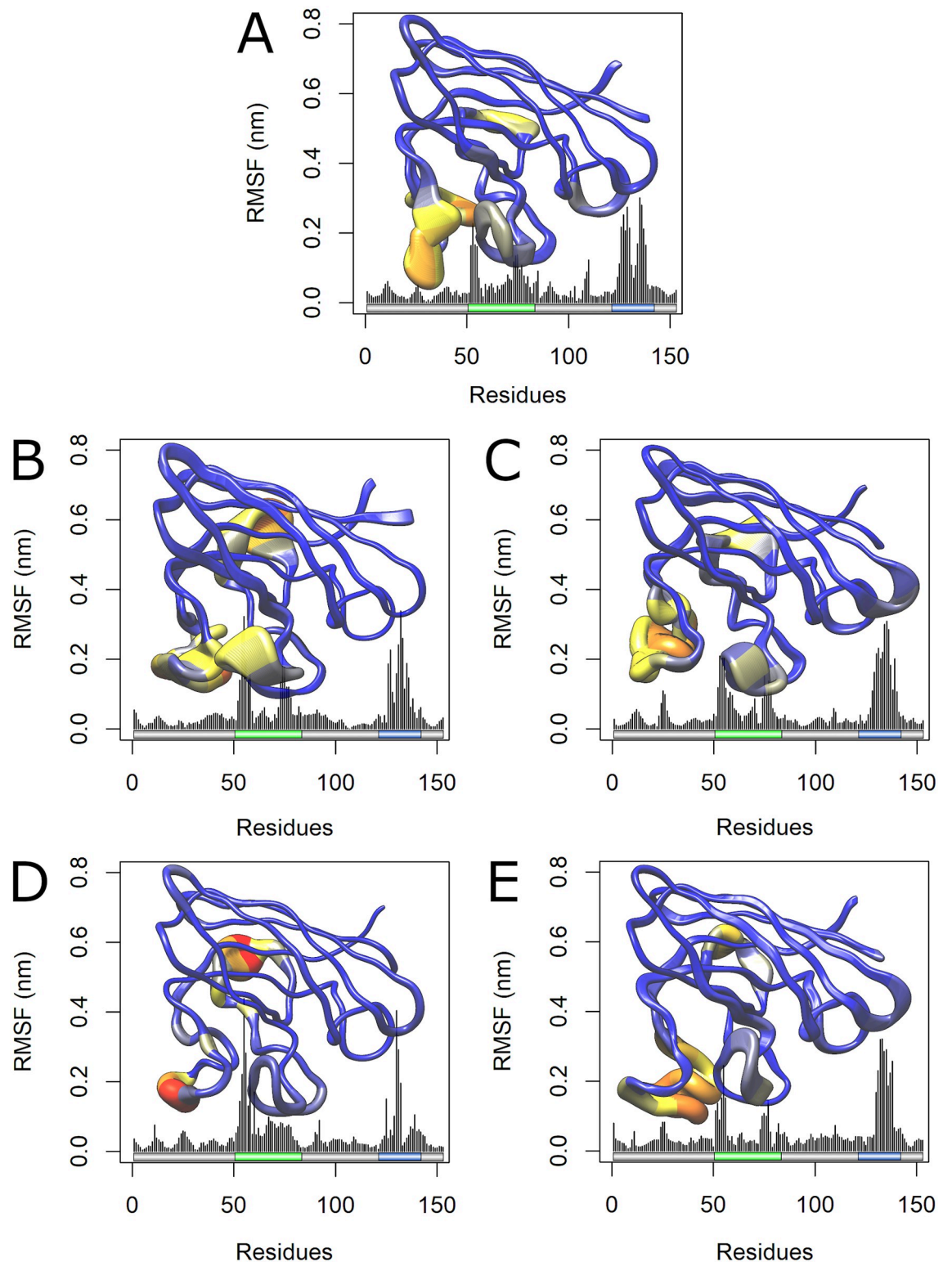
Fig 9. RMSF contribution to PC2 of wild-type SOD1 and its variants. The RMSF contribution of each amino-acid to PC2 is shown as a line plot and projected on the corresponding protein structure. Schematic representations of the SOD1 functional regions are also shown to further comparison. The metal-binding and electrostatic loops of SOD1 are highlighted in green and blue, respectively. Each amino-acid of SOD1 wild-type and variants was colored and sized according to its RMSF contribution, following a coloring-thickness scale that varies from blue and thin (low fluctuations) to red and thick (high fluctuations). (A) RMSF contribution of wild-type SOD1 to PC2. (B) RMSF contribution of variant A4V to PC2. (C) RMSF contribution of variant D90A to PC2. (D) RMSF contribution of variant H46R to PC2. (E) RMSF contribution of variant I113T to PC2.
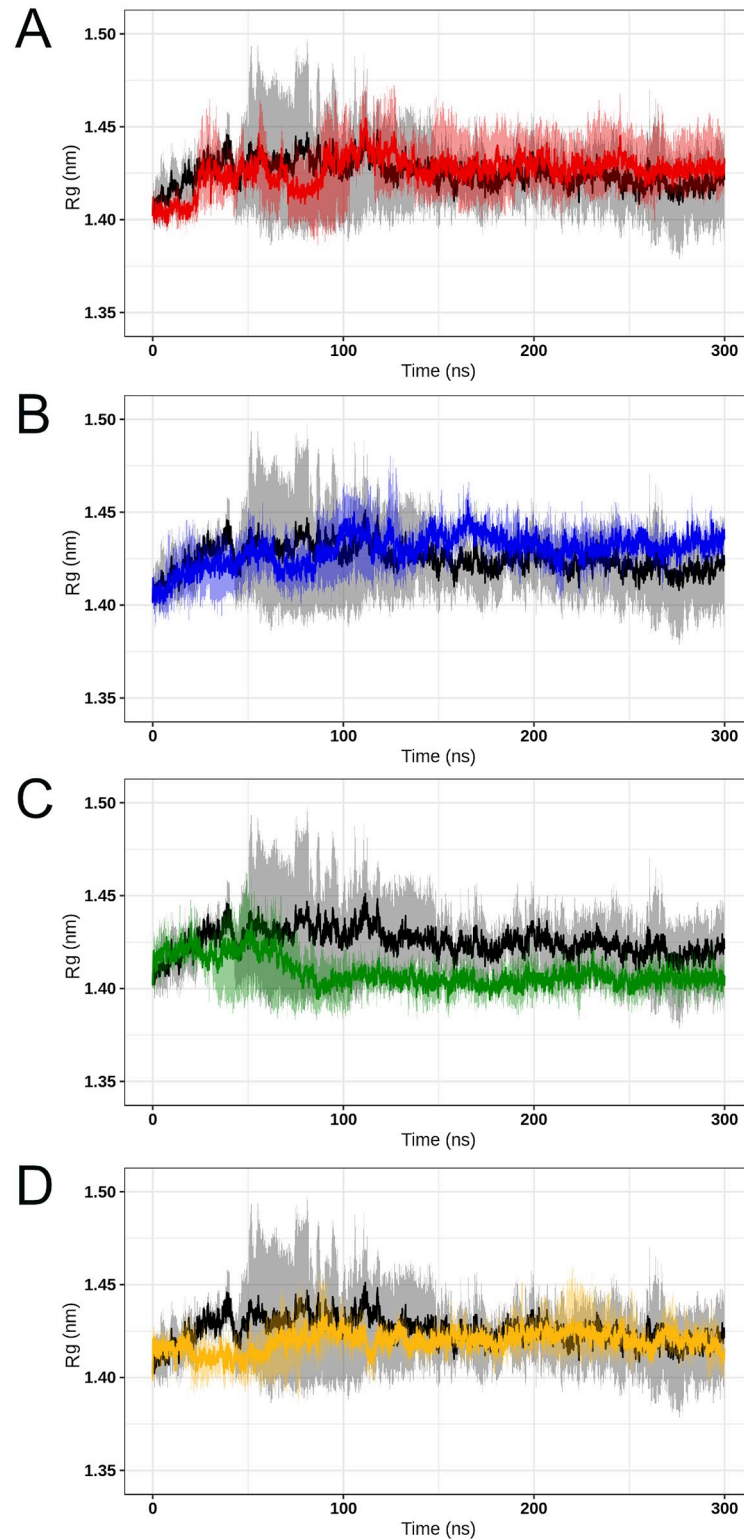
Fig 10. Rg analysis of wild-type SOD1 and its variants. The Rg values computed for the wild-type SOD1 and its variants at 300K are shown over time. The means (solid lines) and confidence intervals (smooth lines) are displayed for the triplicates. (A) The wild-type is represented in black, while variant A4V is represented in red. (B) The wild-type is represented in black, while variant D90A is represented in blue. (C) The wild-type is represented in black, while variant H46R is represented in green. (D) The wild-type is represented in black, while variant I113T is represented in yellow.
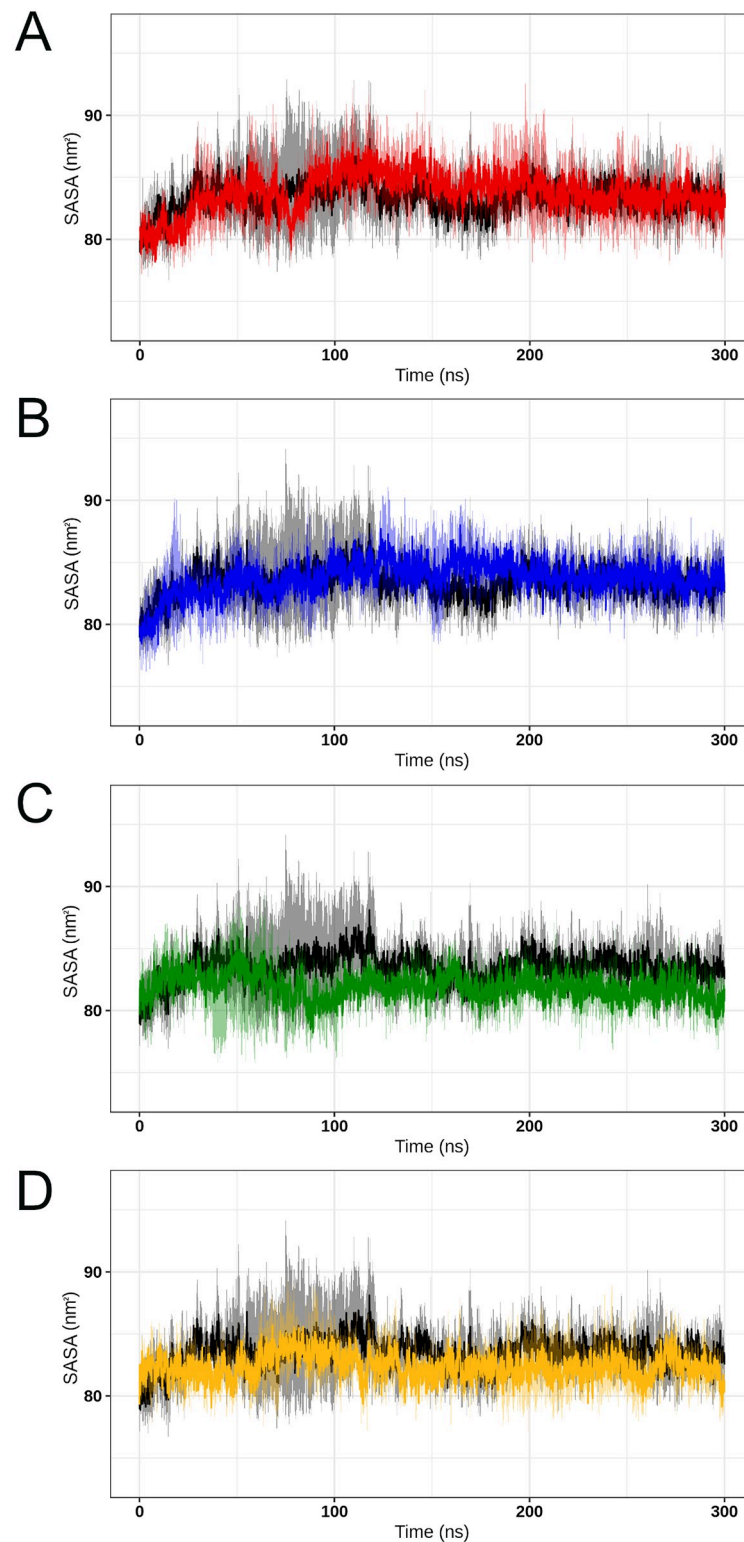
Fig 11. SASA analysis of wild-type SOD1 and its variants. The SASA values computed for the wild-type SOD1 and its variants at 300K are shown over time. The means (solid lines) and confidence intervals (smooth lines) are displayed for the triplicates. (A) The wild-type is represented in black, while variant A4V is represented in red. (B) The wild-type is represented in black, while variant D90A is represented in blue. (C) The wild-type is represented in black, while variant H46R is represented in green. (D) The wild-type is represented in black, while variant I113T is represented in yellow.
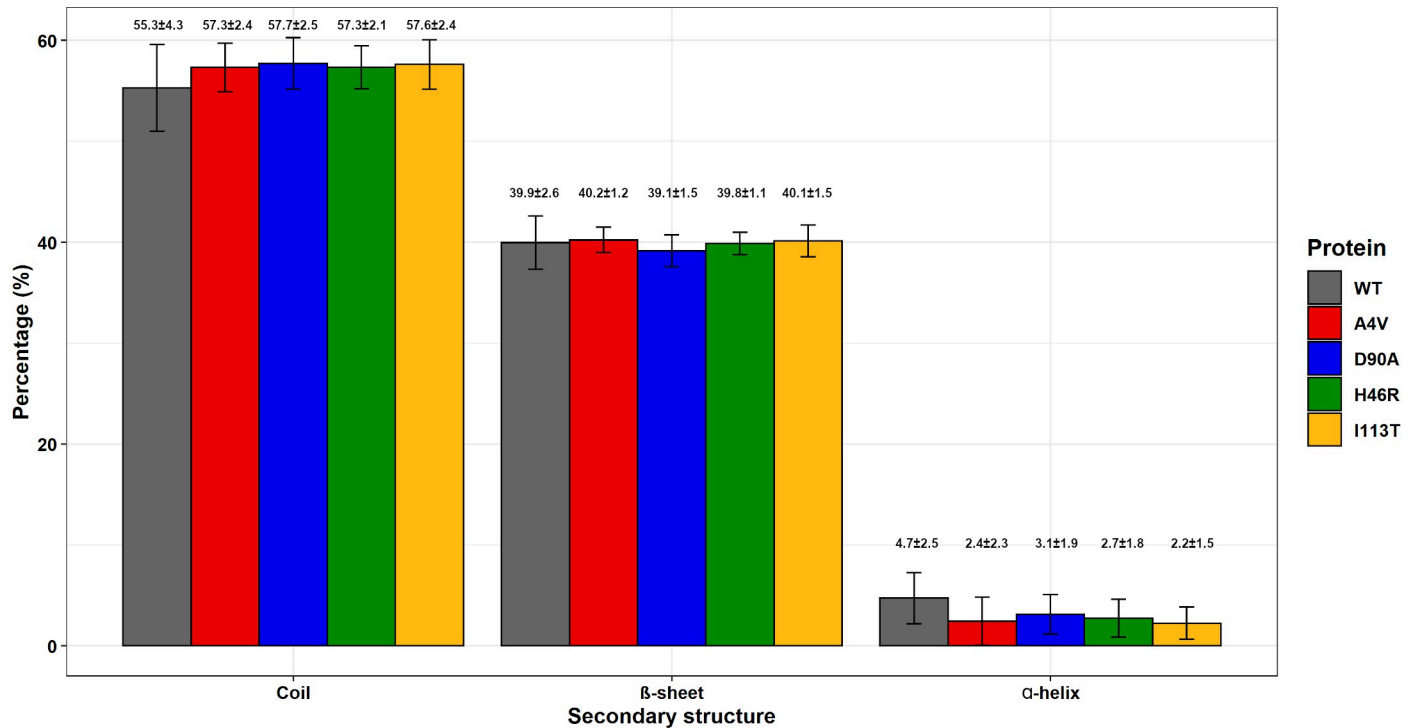
Fig 12. Secondary structure analysis of wild-type SOD1 and its variants. The average number (in percentages) of alpha-helices, beta-sheets, and coils formed throughout the simulations of wild-type SOD1 and its variants are shown as a bar plot. The error bar represents the standard deviation. The bar plot was designed using the ggplot2 library in R software. The average numbers and their respective standard deviations are also shown in the figure.

https://doi.org/10.1371/journal.pone.0247841.g012

(83.3±1.9) are similar to that of variants A4V (83.7±2.2), D90A (84.0±2.0), H46R (81.7±1.6), and I113T (82.2±1.8). This result suggested no SASA alterations in the analyzed variants.

We also performed an SS analysis to further investigate the structural impact of SOD1 protein variants. The average number (in percentage) of coils, beta-sheets, and alpha-helices formed throughout the simulations was computed and displayed in barplot graphs. As shown in Fig 12, no alterations in the average number of SS were observed for the analyzed variants when compared to the wild-type SOD1.

Discussion

Here, we compiled 233 missense mutations in the SOD1 protein (S1 Table), including thirty mutations that have not yet been registered in the dbSNP database. Forty-seven percent of these variants were classified as deleterious by all the functional prediction algorithms used (Fig 2D), suggesting that they might be harmful to SOD1. Our findings also suggested that missense mutations in SOD1 are most likely to be harmful than neutral to protein function, given the elevated rates of deleterious predictions computed for the variants (Fig 2D). Furthermore, the SOD1 variants with the highest rates of deleterious predictions in the functional prediction analysis mainly affect its metal-binding and electrostatic loops (S1 Fig), which reinforces the functional importance of the amino-acids within these regions.

The functional prediction algorithms used (S2 Fig) presented high accuracy in detecting the known functional impact of the variants A4V, H46R, and I113T [15, 39]. However, these algorithms presented low accuracy in predicting the deleterious potential of variant D90A. Thus, as previously shown by our group, using a variety of algorithms is crucial for predicting the

functional impact of missense mutations. Especially because these algorithms apply different strategies to make predictions, and no gold standard method is currently established for this purpose [16].

Our findings also suggested that stability reduction is a common feature in SOD1 protein variants, given the elevated rate of destabilizing predictions (forty-nine percent) in the Mutant3.0 and Fold-X analyses (S3 Table). *In vitro* studies analyzing the stability of over 100 SOD1 protein variants have shown that most of these mutations are destabilizing [57, 58].

The divergent results observed at the stability prediction of some SOD1 protein variants (S3 Table) may occur due to the different strategies applied by FoldX and I-Mutant to make predictions. FoldX applies an empirical force field trained in a database of engineered proteins to estimate free-energy changes ($\Delta\Delta G$) upon mutations and evaluate the stability impact of missense mutations [59], while I-Mutant 2.0 uses Support Vector Machine and information from a database of experimentally determined structures for this purpose [60].

This analysis also showed that the variants A4V, D90A, and I113T were predicted to decrease protein stability by Fold-X and I-Mutant, which is in agreement with the literature consulted [61–65]. The variant H46R, in turn, was only predicted as destabilizing by Fold-X (S3 Table). The variant H46R is not destabilized relative to the wild-type SOD1 in the nascent state. After the post-translational modifications necessary for SOD1 activity [66], this variant becomes relatively destabilized [64, 67].

Four regions with high aggregation potential, comprised between the amino-acids 4–8, 100–106, 111–120, and 146–153, were predicted for SOD1 by Castillo and Ventura, 2009. Our findings showed that 14 variants (K3E, A4F, A4V, C6Y, C6W, C6F, D96V, R115G, R115C, H120L, C146R, G147C, G147R, G147A) were predicted to increase the SOD1 aggregation tendency, while 7 variants (A145D, G147D, V148G, I149T, I151T, I151S, and A152P) were predicted to decrease this feature (S3 Table). All these variants occur at known aggregation-prone regions of SOD1, except for variants D96V and H120L.

Although this association is not yet fully understood [68], stability reduction and protein aggregation are considered synergistic risk factors for ALS severity and short survival time [69, 70]. All the variants predicted to increase protein aggregation tendency, except for variants D96V and H120L, were also predicted to decrease protein stability by at least one functional prediction algorithm used (S2 Table and S3 Table). This result thereby suggested that these variants may be potentially harmful to SOD1 and, consequently, valuable targets for future investigation.

The SNPEffect analysis (S3 Table) also showed that five variants were predicted to increase SOD1 amyloid propensity (V14A, G16C, G16A, G16S, and A89Y), while 5 variants were predicted to decrease this feature (V14M, Q15R, I17V, I18M, and N19S). Four amyloid-prone segments were identified *in silico* at the SOD1 sequence by Ivanova *et al.*, 2014: 14–21, 30–38, 101–107, and 147–153. All the mutations predicted here to affect amyloid propensity, except for variant A90Y, occur in one of these regions, i.e. between amino-acids 14 and 21. Amyloid-like aggregates were observed *in vitro* for SOD1 variants occurring at these amyloid-prone regions [41, 71].

SOD1 folding and Cu/Zn acquisition are catalyzed by the copper chaperone for SOD1 (CCS). In the absence of CCS interaction, SOD1 is highly likely to remain inactive and unfolded, being degraded or accumulate in toxic aggregates [72]. The predictive analysis showed that the variant P66R decreases SOD1 chaperone binding (S3 Table). This mutation was previously found in an fALS patient, but the characterization of its effects has not yet been performed [73]. The amino-acid substitution P66R occurs at the metal-binding loop of SOD1 [74], which could hinder binding, preventing CCS action [75–77].

The variant A4V was predicted to increase protein aggregation (S3 Table), which is a known feature of this variant [78]. The variants D90A, H46R, and I113T were not predicted to affect protein aggregation, amyloid propensity, nor chaperone binding (S3 Table). However, studies have shown that these variants presented an increased tendency to form aggregates and amyloid-like fibrils [43, 67, 70, 79–84].

Functionally important amino-acids on proteins are usually conserved throughout evolution due to the high selective pressure [19]. Most of the analyzed SOD1 variants occur in conserved positions (Fig 3B), which might be harmful to SOD1 as they affect possibly important amino-acids to protein, especially the variants also predicted as deleterious by all the functional prediction algorithms used (S1 Fig). Mutations affecting conserved amino-acids within SOD1 can cause ALS, which suggests that even minor alterations in this protein may severely affect its structure and function [85].

The variants affecting highly conserved amino-acids mostly occur at the metal-binding and electrostatic loops (S2 Fig). The highly conserved amino-acids within the SOD1 structure are mainly located on these functionally important loops [43, 86], particularly at positions directly involved in metal-binding, superoxide guidance, and active site pocket formation [42, 43, 75].

As predicted by ConSurf (S2 Fig), the amino-acid substitution A4V and H46R occur at known conserved regions of SOD1 [21, 35, 87]. Although our results indicated that the variant I113T occurs in a variable position, this amino-acid is evolutionarily conserved in mammals [88]. The variant D90A, in turn, was predicted to affect an average conserved position. Unlike most ALS-related mutations in SOD1, D90A does not occur in a conserved amino-acid [65].

High fluctuations of RMSD values (Fig 4) were observed at the beginning of all simulations, which pointed to an initial moment of structural instability. This behavior is further confirmed by the high fluctuations of Rg (Fig 10), and SASA (Fig 11) observed in the first half of simulations. This might be attributed to the initial kinetic shock of molecular systems that occur in MD simulation processes [34]. The RMSD analysis also indicated that the protein structures fluctuate around average stable conformations after approximately 150ns (Fig 4), making sense to analyze local structural parameters, such as RMSF and B-factor [20]. The Rg analysis further suggested stable protein folding for SOD1 wild-type and its variants in the second half of trajectories (Fig 10). No alterations in the protein compactness (Fig 10), accessible surface (Fig 11), and secondary structure formation (Fig 12) were observed for the analyzed variants.

The essential mobility alterations observed in the ED analysis (Figs 7 and 8) occur at regions similar to those found altered in the RMSF (Fig 5) and B-factor (Fig 6) analyses, which mainly affect the metal-binding and electrostatic loops of the variants. Flexibility alterations can lead to strong and non-intuitive consequences for protein binding properties, as structural flexibility is determinant for binding affinity and specificity [89]. Moreover, the biological function of a protein is usually determined by its essential motions, especially those involved in protein interactions and binding to substrates [19]. Thus, the alterations observed during the MD simulations for the analyzed variants may affect SOD1 binding properties and, possibly, functional interactions, particularly at the metal-binding and electrostatic loops.

The ALS-SOD1 pathogenesis was initially believed to arise from the oxidative damage caused by reduced SOD1 activity, nevertheless, studies have further revealed that SOD1-null mice do not develop motor neuron disorder. On the other hand, transgenic mice expressing ALS-SOD1 mutant proteins become paralyzed despite possessing normal/elevated levels of SOD1 activity [90]. These pieces of evidence then suggest that SOD1 mutations lead to motor neuron death through toxic-gain of function rather than loss of enzymatic activity [91, 92].

The precise nature of ALS-SOD1 toxicity has not yet been fully elucidated. However, a substantial body of literature suggested that it arises from SOD1 misfolding and aggregation, which are common properties of ALS-SOD1 mutations and central events in ALS-SOD1

pathophysiology [85, 90, 93, 94]. More than 180 SOD1 coding mutations have already been described in the literature as occurring all over the protein and causing the same disease phenotype, ALS. Apparently, through similar toxic mechanisms, even mutations in SOD1 such as G93A, which cause changes as simple as the addition of a methyl group, can lead to the development of ALS [94, 95].

Many studies have been carried out to understand the molecular basis of ALS-SOD1 toxicity. The study of Elam *et.al.* 2003 assessed the structural impact of the S134N and H46R variants in SOD1 using X-ray crystallography [90]. The study of Banci *et.al.* 2005 evaluated the effect of S134N using nuclear magnet resonance methods (RMN) [35]. The study of Molnar *et.al.* 2009 analyzed the impact of 13 ALS-SOD1 mutations (A4V, L38V, G41S, H46R, G72Y, D76S, G85R, D90A, G93A, D124V, D125H, and S134N) on protein structure and dynamics using hydrogen/deuterium exchange mass spectrometry [95]. Despite the different methods used by the authors, all of them compared the ALS-related mutants with the wild-type protein and observed structural or dynamics perturbations at the electrostatic loop of the analyzed variants [35, 90, 95]. Molnar *et.al.* 2009 also observed alterations at the metal-binding loop of variants A4V, G93A, S134N, G85R, D124V, and D125H. No results were generated for the electrostatic and metal-binding loops of H46R in Molnar's hydrogen/deuterium exchange experiments [95]. Further studies comparing the reactivity of conformation-specific antibodies for wild-type SOD1 and a series of other variants confirmed that SOD1 mutations induce some degree of misfolding at the electrostatic and metal-binding loop [85]. Together with the wet-lab experiments previously described in this paragraph, the molecular dynamics results presented here for A4V, D90A, H46R, and I113T further suggest that dynamics and structural alterations at the metal-binding and electrostatic loops could be a common characteristic of ALS-SOD1 mutations [35, 90, 95].

The computational methods are important allies of the wet-lab experiments [96], as the *silico* analyses can complement and overcome some difficulties of *in vitro* and *in vivo* assays [97, 98]. The MD simulations performed in this study can contribute to a better understanding of the molecular basis of ALS-SOD1 since this method provides detailed information on protein conformation and fluctuations over time [20], which can complement the results of already performed wet-lab experiments and overcome some limitations such as the static images provided by X-ray crystallography and the limited size of peptic fragments that can be analyzed by hydrogen/deuterium exchange [95]. On the other hand, we did not analyze whether SOD1 mutations not related to ALS also cause dynamics and structural alterations at the electrostatic and metal-binding loops, which is a limitation of this study. The alterations observed here and in the wet-lab experiments previously mentioned [35, 85, 90, 95] for ALS-SOD1 mutations could be an effect of SOD1 mutations in general, not necessarily related to the development of ALS.

Although the biological consequences of alterations at the functional loops of ALS-SOD1 variants are not fully understood, a well-accepted hypothesis suggests that this common property could lead to ALS possibly through a similar toxic mechanism involving aberrant interactions with cellular constituents and protein aggregation [68, 85, 90, 93, 95]. The study of Elam *et.al.* 2003 indicated that well-ordered electrostatic and metal-binding loops of SOD1 are critical to prevent protein aggregation. This study suggested that the functional loops adopt non-native conformations upon H46R and S134N mutations, which deprotects a hydrophobic core of the β-barrel region of SOD1 that serves as a molecular interface for non-native protein interactions. Among the non-native interactions described by Elam *et.al.* 2003, extensive apolar and hydrogen-bond interactions were formed between the electrostatic loop of a mutant protein and the deprotected β-barrel region of a neighboring SOD1 [90]. Banci *et.al.* 2005 identified abnormal contacts at the electrostatic loop and increased aggregate formation for the

variant S134N in solution, which the authors attributed to the increased mobility and flexibility also observed at this functional loop. Protein surface regions characterized by high mobility and flexibility have been described as an important factor for molecular recognition [93]. Aberrant interactions with dynein and mitochondrial proteins, such as BCL-2, were also reported in the literature as a common feature of ALS-SOD1 mutations. The study of Zhang *et.al.* 2007 showed that the variants A4V, G85R, and G93A interact with the dynein, while the wild-type SOD1 does not. SOD1 mutant forms with disordered loop regions preferentially interact with the dynein complex, favoring aggregation [99]. Thus, together with these previous studies, our MD simulation results can further support this hypothesis since the flexibility and essential dynamics alterations observed at the functional loops of A4V, D90A, H46R, and I113T may favor or disfavor interactions, possibly leading to abnormal association with cellular constituents [19, 89].

Conclusion

Two hundred and thirty-three SOD1 variants were compiled from the databases and literature consulted, which underwent functional and stability predictions. The predictive analysis pointed to an elevated rate of deleterious and destabilizing predictions for these variants, indicating their harmful effects. Among the analyzed variants, those occurring at the electrostatic and metal-binding loops were noticeably more damaging to protein structure and function, which makes them valuable targets for future investigation. The ConSurf analysis indicated that mutations in SOD1 mainly affect conserved and possibly functionally important aminoacids. The MD analyses of variants A4V, D90A, H46R, and I113T pointed to flexibility and essential dynamics alterations at the electrostatic and metal-binding loops that could lead to aberrant interactions triggering toxic protein aggregation. These alterations may have harmful implications for SOD1 and explain their association with ALS. Understanding the effects of SOD1 mutations on protein structure and function facilitates the design of further experiments and provides relevant information on the molecular mechanism of pathology, which may contribute to improvements in existing treatments for ALS.

Supporting information

S1 Table. SOD1 variants compiled from the literature (PubMed) and databases (ClinVar, UNIPROT, ALSod, dbSNP, and OMIM).
(DOCX)

S2 Table. Functional prediction for each SOD1 protein variant.
(DOCX)

S3 Table. Stability and SNPEffect4.0 predictions for each SOD1 protein variants.
(DOCX)

S1 Fig. Functional prediction of SOD1 protein variants. The SOD1 variants compiled from the literature and databases (233) were analyzed using ten different functional prediction algorithms. The bar plot indicates the number of neutral (gray) and deleterious (red) predictions for each SOD1 variant. The mutations predicted as deleterious by all the functional predictions algorithms used are marked by an asterisk. (A) Functional prediction from the variant A2V to L39V. (B) Functional prediction from the variant T40I to V88M. (C) Functional prediction from the variant A90V to V120L. (D) Functional prediction from the variant H121Q to A153P.
(TIF)

S2 Fig. Evolutionary conservation analysis of SOD1 protein variants. The bar plot shows the ConSurf score for each amino-acid of SOD1 protein affected by mutations. The bar plot was colored according to the ConSurf coloring-scheme, which varies from cyan and variable to maroon and conserved. The mutations affecting highly conserved amino-acids (i.e. ConSurf score = 9) are marked by an asterisk. (A) ConSurf prediction from the variant A2V to L39V. ConSurf did not predict the conservation score of the A2V variant due to the lack of evolutionary information. (B) ConSurf prediction from the variant T40I to A90Y. (C) ConSurf prediction from the variant A90V to V120L. (D) ConSurf prediction from the variant H121Q to A153P.
(TIF)

Author Contributions

Conceptualization: Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Data curation: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Formal analysis: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Funding acquisition: Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Investigation: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Methodology: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Project administration: Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Resources: Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Software: Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Supervision: Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Validation: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Visualization: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Writing – original draft: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

Writing – review & editing: Gabriel Rodrigues Coutinho Pereira, Bárbara de Azevedo Abrahim Vieira, Joelma Freire De Mesquita.

References

1. Tao QQ, Wu ZY. Amyotrophic lateral sclerosis: Precise diagnosis and individualized treatment. Chin Med J (Engl). 2017; 130: 2269–2272. PMID: 28937029

2. Zarei S, Carr K, Reiley L, Diaz K, Guerra O, Altamirano P, et al. A comprehensive review of amyotrophic lateral sclerosis. Surg Neurol Int. 2015; 6: 171. https://doi.org/10.4103/2152-7806.169561 PMID: 26629397

3. van Es MA, Hardiman O, Chio A, Al-Chalabi A, Pasterkamp RJ, Veldink JH, et al. Amyotrophic lateral sclerosis. Lancet. 2017; 390: 2084–2098. https://doi.org/10.1016/S0140-6736(17)31287-4 PMID: 28552366

4. Gladman M, Dharamshi C, Zinman L. Economic burden of amyotrophic lateral sclerosis: A Canadian study of out-of-pocket expenses. Amyotroph Lateral Scler Front Degener. 2014; 15: 426–432. https://doi.org/10.3109/21678421.2014.932382 PMID: 25025935

5. Traynor BJ, Codd MB, Corr B, Forde C, Frost E, Hardiman O. Amyotrophic Lateral Sclerosis Mimic Syndromes. Arch Neurol. 2000; 57: 109. https://doi.org/10.1001/archneur.57.1.109 PMID: 10634456

6. Arthur KC, Calvo A, Price TR, Geiger JT, Chiò A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. Nat Commun. 2016; 7: 1–6. https://doi.org/10.1038/ncomms12408 PMID: 27510634

7. Filipi T, Hermanova Z, Tureckova J, Vanatko O, Anderova M. Glial Cells—The Strategic Targets in Amyotrophic Lateral Sclerosis Treatment. J Clin Med. 2020; 9: 261. https://doi.org/10.3390/jcm9010261 PMID: 31963681

8. Jaiswal MK. Riluzole and edaravone: A tale of two amyotrophic lateral sclerosis drugs. Med Res Rev. 2018. https://doi.org/10.1002/med.21528 PMID: 30101496

9. Gitler AD, Dhillon P, Shorter J. Neurodegenerative disease: models, mechanisms, and a new hope. Dis Model Mech. 2017; 10: 499–502. https://doi.org/10.1242/dmm.030205 PMID: 28468935

10. Mancuso R, Navarro X. Amyotrophic lateral sclerosis: Current perspectives from basic research to the clinic. Prog Neurobiol. 2015; 133: 1–26. https://doi.org/10.1016/j.pneurobio.2015.07.004 PMID: 26253783

11. Lee J, Hyeon SJ, Im H, Ryu H, Kim Y, Ryu H. Astrocytes and microglia as non-cell autonomous players in the pathogenesis of ALS. Exp Neurobiol. 2016; 25: 233–240. https://doi.org/10.5607/en.2016.25.5.233 PMID: 27790057

12. Dekker AM, Seelen M, van Doormaal PTC, van Rheenen W, Bothof RJP, van Riessen T, et al. Large-scale screening in sporadic amyotrophic lateral sclerosis identifies genetic modifiers in C9orf72 repeat carriers. Neurobiol Aging. 2016; 39: 220.e9–220.e15. https://doi.org/10.1016/j.neurobiolaging.2015.12.012 PMID: 26777436

13. Tang L, Ma Y, Liu XL, Chen L, Fan DS. Better survival in female SOD1-mutant patients with ALS: A study of SOD1-related natural history. Transl Neurodegener. 2019; 8: 1–10.

14. Rodríguez-sureda V, Vilches Á, Sánchez O, Audí L, Domínguez C. Intracellular Oxidant Activity, Antioxidant Enzyme Defense System, and Cell Senescence in Fibroblasts with Trisomy 21. Oxid Med Cell Longev. 2015; 2015: 17. https://doi.org/10.1155/2015/509241 PMID: 25852816

15. Kaur SJ, McKeown SR, Rashid S. Mutant SOD1 mediated pathogenesis of Amyotrophic Lateral Sclerosis. Gene. 2016; 577: 109–118. https://doi.org/10.1016/j.gene.2015.11.049 PMID: 26657039

16. Pereira GRC, Da Silva ANR, Do Nascimento SS, De Mesquita JF. In silico analysis and molecular dynamics simulation of human superoxide dismutase 3 (SOD3) genetic variants. J Cell Biochem. 2019; 120: 3583–3598. https://doi.org/10.1002/jcb.27636 PMID: 30206983

17. de Araújo Brasil A, de Carvalho MDC, Gerhardt E, Queiroz DD, Pereira MD, Outeiro TF, et al. Characterization of the activity, aggregation, and toxicity of heterodimers of WT and ALS-associated mutant Sod1. Proc Natl Acad Sci U S A. 2019; 116: 25991–26000. https://doi.org/10.1073/pnas.1902483116 PMID: 31796595

18. Pansarasa O, Bordoni M, Diamanti L, Sproviero D, Gagliardi S, Cereda C. Sod1 in amyotrophic lateral sclerosis: "ambivalent" behavior connected to the disease. Int J Mol Sci. 2018; 19: 1–13. https://doi.org/10.3390/ijms19051345 PMID: 29751510

19. Pereira GRC, Tavares GDB, De Freitas MC, De Mesquita JF. In silico analysis of the tryptophan hydroxylase 2 (TPH2) protein variants related to psychiatric disorders. PLoS One. 2020; 15: 1–23. https://doi.org/10.1371/journal.pone.0229730 PMID: 32119710

20. De Oliveira CCS, Pereira GRC, De Alcantara JYS, Antunes D, Caffarena ER, De Mesquita JF. In silico analysis of the V66M variant of human BDNF in psychiatric disorders: An approach to precision medicine. PLoS One. 2019; 14: e0215508. https://doi.org/10.1371/journal.pone.0215508 PMID: 30998730

21. Da Silva ANR, Pereira GRC, Moreira LGA, Rocha CF, Mesquita JF. SOD1 in amyotrophic lateral sclerosis development–in silico analysis and molecular dynamics of A4F and A4V variants. J Cell Biochem. 2019; 1–9. https://doi.org/10.1002/jcb.29048 PMID: 31134679

22. Pereira GRC, Tellini GHAS, De Mesquita JF. In silico analysis of PFN1 related to amyotrophic lateral sclerosis. PLOS ONE. 2019; 14: e0215723. https://doi.org/10.1371/journal.pone.0215723 PMID: 31216283

23. Thusberg J, Vihinen M. Pathogenic or not? and if so, then how? Studying the effects of missense mutations using bioinformatics methods. Hum Mutat. 2009; 30: 703–714. https://doi.org/10.1002/humu.20938 PMID: 19267389

24. Roy Choudhury A, Cheng T, Phan L, Bryant SH, Wang Y. Supporting precision medicine by data mining across multi-disciplines: An integrative approach for generating comprehensive linkages between

single nucleotide variants (SNVs) and drug-binding sites. Bioinformatics. 2017; 33: 1621–1629. https://doi.org/10.1093/bioinformatics/btx031 PMID: 28158543

25. Rose PW, Prlić A, Altunkaya A, Bi C, Bradley AR, Christie CH, et al. The RCSB protein data bank: Integrative view of protein, gene and 3D structural information. Nucleic Acids Res. 2017; 45: D271–D281. https://doi.org/10.1093/nar/gkw1000 PMID: 27794042

26. Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R, et al. UniProt: The universal protein knowledgebase. Nucleic Acids Res. 2017; 45: D158–D169. https://doi.org/10.1093/nar/gkw1099 PMID: 27899622

27. Wheeler DL, Chappey C, Lash a E, Leipe DD, Madden TL, Schuler GD, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2000; 28: 10–4. https://doi.org/10.1093/nar/28.1.10 PMID: 10592169

28. Abel O, Powell JF, Andersen PM, Al-Chalabi A. ALSoD: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. Hum Mutat. 2012; 33: 1345–1351. https://doi.org/10.1002/humu.22157 PMID: 22753137

29. Sherry ST. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001; 29: 308–311. https://doi.org/10.1093/nar/29.1.308 PMID: 11125122

30. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018; 46: D1062–D1067. https://doi.org/10.1093/nar/gkx1153 PMID: 29165669

31. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. 2017; 45: D183–D189. https://doi.org/10.1093/nar/gkw1138 PMID: 27899595

32. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015; 31: 2745–2747. https://doi.org/10.1093/bioinformatics/btv195 PMID: 25851949

33. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. PLoS Comput Biol. 2014; 10: 1–11. https://doi.org/10.1371/journal.pcbi.1003440 PMID: 24453961

34. Agrahari AK, Sneha P, George Priya Doss C, Siva R, Zayed H. A profound computational study to prioritize the disease-causing mutations in PRPS1 gene. Metab Brain Dis. 2018; 33: 589–600. https://doi.org/10.1007/s11011-017-0121-2 PMID: 29047041

35. Moreira LGA, Pereira LC, Drummond PR, De Mesquita JF, Andersen P, Phukan J, et al. Structural and Functional Analysis of Human SOD1 in Amyotrophic Lateral Sclerosis. Le W, editor. PLoS One. 2013; 8: e81979. https://doi.org/10.1371/journal.pone.0081979 PMID: 24312616

36. De Carvalho MDC, De Mesquita JF. Structural Modeling and In Silico Analysis of Human Superoxide Dismutase 2. PLoS One. 2013; 8. https://doi.org/10.1371/journal.pone.0065558 PMID: 23785434

37. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 2016; 44: 344–350. https://doi.org/10.1093/nar/gkw408 PMID: 27166375

38. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. J Comput Chem. 2005; 26: 1701–1718. https://doi.org/10.1002/jcc.20291 PMID: 16211538

39. Battistini S, Benigni M, Ricci C, Rossi A. SOD1 mutations in amyotrophic lateral sclerosis. Eur Neurol J. 2010; 000: 1–11. Available: http://link.springer.com/article/10.1007/s00415-005-0742-y

40. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: An R package for the comparative analysis of protein structures. Bioinformatics. 2006; 22: 2695–2696. https://doi.org/10.1093/bioinformatics/btl461 PMID: 16940322

41. Keerthana SP, Kolandaivel P. Structural investigation on the electrostatic loop of native and mutated SOD1 and their interaction with therapeutic compounds. RSC Adv. 2015; 5: 34630–34644. https://doi.org/10.1039/C5RA00286A

42. Srinivasan E, Rajasekaran R. Computational Investigation on Electrostatic Loop Mutants Instigating Destabilization and Aggregation on Human SOD1 Protein Causing Amyotrophic Lateral Sclerosis. Protein J. 2019; 38: 37–49. https://doi.org/10.1007/s10930-018-09809-0 PMID: 30701485

43. Sirangelo I, Iannuzzi C. The role of metal binding in the amyotrophic lateral sclerosis-related aggregation of copper-zinc superoxide dismutase. Molecules. 2017; 22: 1–13. https://doi.org/10.3390/molecules22091429 PMID: 28850080

44. Srinivasan E, Sethumadhavan R, Rajasekaran R. A theoretical study on Zn binding loop mutants instigating destabilization and metal binding loss in human SOD1 protein. J Mol Model. 2017; 23: 1–7.

45. Knapp B, Frantal S, Cibena M, Schreiner W, Bauer P. Is an intuitive convergence definition of molecular dynamics simulations solely based on the root mean square deviation possible? J Comput Biol. 2011; 18: 997–1005. https://doi.org/10.1089/cmb.2010.0237 PMID: 21702691

46. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. Neuron. 2018; 99: 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011 PMID: 30236283

47. Martinez L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. Kleinjung J, editor. PLoS One. 2015; 10: e0119264. https://doi.org/10.1371/journal.pone.0119264 PMID: 25816325

48. Craveur P, Joseph AP, Esque J, Narwani TJ, Noël F, Shinada N, et al. Protein flexibility in the light of structural alphabets. Front Mol Biosci. 2015; 2. https://doi.org/10.3389/fmolb.2015.00020 PMID: 26075209

49. Kuzmanic A, Zagrovic B. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. Biophys J. 2010; 98: 861–871. https://doi.org/10.1016/j.bpj.2009.11.011 PMID: 20197040

50. Yuan Z, Bailey TL, Teasdale RD. Prediction of protein B-factor profiles. Proteins Struct Funct Genet. 2005; 58: 905–912. https://doi.org/10.1002/prot.20375 PMID: 15645415

51. Sang P, Hu W, Ye YJ, Li LH, Zhang C, Xie YH, et al. In silico screening, molecular docking, and molecular dynamics studies of SNP-derived human P5CR mutants. Journal of Biomolecular Structure and Dynamics. 2017. https://doi.org/10.1080/07391102.2016.1222967 PMID: 27677826

52. David CC, Jacobs DJ. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins Article. Proteins. 2014.

53. Stein SAM, Loccisano AE, Firestine SM, Evanseck JD. Principal Components Analysis: A Review of its Application on Molecular Dynamics Data. Annual Reports in Computational Chemistry. 2006. pp. 233–248.

54. Khan FI, Wei DQ, Gu KR, Hassan MI, Tabrez S. Current updates on computer aided protein modeling and designing. Int J Biol Macromol. 2016; 85: 48–62. https://doi.org/10.1016/j.ijbiomac.2015.12.072 PMID: 26730484

55. Sinha S, Verma S, Singh A, Somvanshi P, Grover A. Simulation Based Investigation of Deleterious nsSNPs in ATXN2 Gene and Its Structural Consequence Toward Spinocerebellar Ataxia. J Cell Biochem. 2018; 119: 499–510. https://doi.org/10.1002/jcb.26209 PMID: 28612427

56. Wang Q, Mehmood A, Wang H, Xu Q, Xiong Y, Wei DQ. Computational screening and analysis of lung cancer related non-synonymous single nucleotide polymorphisms on the human kirsten rat sarcoma gene. Molecules. 2019; 24. https://doi.org/10.3390/molecules24101951 PMID: 31117243

57. Synofzik M, Ronchi D, Keskin I, Basak AN, Wilhelm C, Gobbi C, et al. Mutant superoxide dismutase-1 indistinguishable from wild-type causes ALS. Hum Mol Genet. 2012; 21: 3568–3574. https://doi.org/10.1093/hmg/dds188 PMID: 22595972

58. Kumar V, Rahman S, Choudhry H, Zamzami MA, Sarwar Jamal M, Islam A, et al. Computing disease-linked SOD1 mutations: Deciphering protein stability and patient-phenotype relations article. Sci Rep. 2017; 7: 1–13. PMID: 28127051

59. Schymkowitz JWH, Rousseau F, Martins IC, Ferkinghoff-Borg J, Stricher F, Serrano L. Prediction of water and metal binding sites and their affinities by using the Fold-X force field. Proc Natl Acad Sci U S A. 2005; 102: 10147–10152. https://doi.org/10.1073/pnas.0501980102 PMID: 16006526

60. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 2005; 33: W306–W310. https://doi.org/10.1093/nar/gki375 PMID: 15980478

61. Shaw BF, Durazo A, Nersissian AM, Whitelegge JP, Faull KF, Valentine JS. Local unfolding in a destabilized, pathogenic variant of superoxide dismutase 1 observed with H/D exchange and mass spectrometry. J Biol Chem. 2006; 281: 18167–18176. https://doi.org/10.1074/jbc.M600623200 PMID: 16644738

62. Eisenberg DS. Perspective on SOD1 mediated toxicity in Amyotrophic Lateral Sclerosis*. 2013; 1: 362–369.

63. Hough MA, Grossmann JG, Antonyuk S V, Strange RW, Doucette PA, Rodriguez JA, et al. Dimer destabilization in superoxide dismutase may result in disease-causing properties: Structures of motor neuron disease mutants. PNAS. 2004; 101: 5976–5981. https://doi.org/10.1073/pnas.0305143101 PMID: 15056757

64. Rodriguez JA, Valentine JS, Eggers DK, Roe JA, Tiwari A, Brown RH, et al. Familial amyotrophic lateral sclerosis-associated mutations decrease the thermal stability of distinctly metallated species of human copper/zinc superoxide dismutase. J Biol Chem. 2002; 277: 15932–15937. https://doi.org/10.1074/jbc.M112088200 PMID: 11854285

65. Jonsson PA, Graffmo KS, Andersen PM, Marklund SL, Brännström T. Superoxide dismutase in amyotrophic lateral sclerosis patients homozygous for the D90A mutation. Neurobiol Dis. 2009; 36: 421–424. https://doi.org/10.1016/j.nbd.2009.08.006 PMID: 19703565

66. Winkler DD, Schuermann JP, Cao X, Holloway SP, Borchelt DR, Carroll MC, et al. Structural and Biophysical Properties of the Pathogenic SOD1 Variant H46R/H48Q. Biochemistry. 2009; 48: 3436–3447. https://doi.org/10.1021/bi8021735 PMID: 19227972

67. Antonyuk S, Elam JS, Hough MA, Strange RW, Doucette PA, Rodriguez JA, et al. Structural consequences of the familial amyotrophic lateral sclerosis SOD1 mutant His46Arg. Protein Sci. 2005; 14: 1201–1213. https://doi.org/10.1110/ps.041256705 PMID: 15840828

68. Schmitt ND, Agar JN. Parsing disease-relevant protein modifications from epiphenomena: perspective on the structural basis of SOD1-mediated ALS. J Mass Spectrom. 2017; 52: 480–491. https://doi.org/10.1002/jms.3953 PMID: 28558143

69. Auclair JR, Boggio KJ, Petsko GA, Ringe D, Agar JN. Strategies for stabilizing superoxide dismutase (SOD1), the protein destabilized in the most common form of familial amyotrophic lateral sclerosis. Proc Natl Acad Sci U S A. 2010; 107: 21394–21399. https://doi.org/10.1073/pnas.1015463107 PMID: 21098299

70. Wright GSA, Antonyuk S V., Kershaw NM, Strange RW, Samar Hasnain S. Ligand binding and aggregation of pathogenic SOD1. Nat Commun. 2013; 4: 1710–1758.

71. Ivanova MI, Sievers SA, Guenther EL, Johnson LM, Winkler DD, Galaleldeen A, et al. Aggregation-triggering segments of SOD1 fibril formation support a common pathway for familial and sporadic ALS. Proc Natl Acad Sci U S A. 2014; 111: 197–201. https://doi.org/10.1073/pnas.1320786110 PMID: 24344300

72. Sala FA, Wright GSA, Antonyuk S V., Garratt RC, Hasnain SS. Molecular recognition and maturation of sod1 by its evolutionarily destabilised cognate chaperone hCCS. PLoS Biol. 2019; 17: 1–22. https://doi.org/10.1371/journal.pbio.3000141 PMID: 30735496

73. Millecamps S, Salachas F, Cazeneuve C, Gordon P, Bricka B, Camuzat A, et al. SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: Genotype-phenotype correlations. J Med Genet. 2010; 47: 554–560. https://doi.org/10.1136/jmg.2010.077180 PMID: 20577002

74. Nogueira A, Da R, Rodrigues G, Pereira C. SOD1 in amyotrophic lateral sclerosis development–in silico analysis and molecular dynamics of A4F and A4V variants. 2019; 1–9. https://doi.org/10.1002/jcb.29048 PMID: 31134679

75. Rakhit R, Chakrabartty A. Structure, folding, and misfolding of Cu,Zn superoxide dismutase in amyotrophic lateral sclerosis. Biochim Biophys Acta—Mol Basis Dis. 2006; 1762: 1025–1037. https://doi.org/10.1016/j.bbadis.2006.05.004 PMID: 16814528

76. Das A, Plotkin SS. Mechanical probes of SOD1 predict systematic trends in metal and dimer affinity of ALS-associated mutants. J Mol Biol. 2013; 425: 850–874. https://doi.org/10.1016/j.jmb.2012.12.022 PMID: 23291526

77. Boyd SD, Ullrich MS, Calvo JS, Behnia F, Meloni G, Winkler DD. Mutations in Superoxide Dismutase 1 (Sod1) Linked to Familial Amyotrophic Lateral Sclerosis Can Disrupt High-Affinity Zinc-Binding Promoted by the Copper Chaperone for Sod1 (Ccs) Stefanie. Molecules. 2020; 25: 1–14.

78. Castillo V, Ventura S. Amyloidogenic regions and interaction surfaces overlap in globular proteins related to conformational diseases. PLoS Comput Biol. 2009; 5. https://doi.org/10.1371/journal.pcbi.1000476 PMID: 19696882

79. Münch C, Bertolotti A. Exposure of hydrophobic surfaces initiates aggregation of diverse ALS-causing superoxide dismutase-1 mutants. J Mol Biol. 2010; 399: 512–525. https://doi.org/10.1016/j.jmb.2010.04.019 PMID: 20399791

80. McAlary L, Aquilina JA, Yerbury JJ. Susceptibility of mutant SOD1 to form a destabilized monomer predicts cellular aggregation and toxicity but not in vitro aggregation propensity. Front Neurosci. 2016; 10: 1–16.

81. Pokrishevsky E, Hong RH, Mackenzie IR, Cashman NR. Spinal cord homogenates from SOD1 familial amyotrophic lateral sclerosis induce SOD1 aggregation in living cells. PLoS One. 2017; 12: 1–10. https://doi.org/10.1371/journal.pone.0184384 PMID: 28877271

82. Oztug Durer ZA, Cohlberg JA, Dinh P, Padua S, Ehrenclou K, Downes S, et al. Loss of metal ions, disulfide reduction and mutations related to familial ALS promote formation of amyloid-like aggregates from superoxide dismutase. PLoS One. 2009; 4. https://doi.org/10.1371/journal.pone.0005004 PMID: 19325915

83. Banci L, Bertini I, Boca M, Girotto S, Martinelli M, Valentine JS, et al. SOD1 and amyotrophic lateral sclerosis: Mutations and oligomerization. PLoS One. 2008; 3: 1–8. https://doi.org/10.1371/journal.pone.0001677 PMID: 18301754

84. Shaw BF, Valentine JS. How do ALS-associated mutations in superoxide dismutase 1 promote aggregation of the protein? Trends Biochem Sci. 2007; 32: 78–85. https://doi.org/10.1016/j.tibs.2006.12.005 PMID: 17208444

85. Rotunno MS, Bosco DA. An emerging role for misfolded wild-type SOD1 in sporadic ALS pathogenesis. Front Cell Neurosci. 2013; 7. https://doi.org/10.3389/fncel.2013.00253 PMID: 24379756

86. Antonyuk S V., Strange RW, Marklund SL, Hasnain SS. The Structure of Human Extracellular Copper-Zinc Superoxide Dismutase at 1.7 Å Resolution: Insights into Heparin and Collagen Binding. J Mol Biol. 2009; 388: 310–326. https://doi.org/10.1016/j.jmb.2009.03.026 PMID: 19289127

87. Alemasov NA, Ivanisenko N V., Ramachandran S, Ivanisenko VA. Molecular mechanisms underlying the impact of mutations in SOD1 on its conformational properties associated with amyotrophic lateral sclerosis as revealed with molecular modelling. BMC Struct Biol. 2018; 18: 1–14. https://doi.org/10.1186/s12900-018-0080-9 PMID: 29431095

88. Wei Q, Zhou Q, Chen Y, Ou R, Cao B, Xu Y, et al. Analysis of SOD1 mutations in a Chinese population with amyotrophic lateral sclerosis: A case-control study and literature review. Sci Rep. 2017; 7: 1–6.

89. Forrey C, Douglas JF, Gilson MK. The fundamental role of flexibility on the strength of molecular binding. Soft Matter. 2012; 8: 6385–6392. https://doi.org/10.1039/C2SM25160D PMID: 22707976

90. Elam JS, Taylor AB, Strange R, Antonyuk S, Doucette PA, Rodriguez JA, et al. Amyloid-like filaments and water-filled nanotubes formed by SOD1 mutant proteins linked to familial ALS. Nat Struct Biol. 2003; 10: 461–467. https://doi.org/10.1038/nsb935 PMID: 12754496

91. Hirayama K, Fujiwara Y, Terada T, Shimizu K, Wada K, Kabuta T. Virtual screening identification of novel chemical inhibitors for aberrant interactions between pathogenic mutant SOD1 and tubulin. Neurochem Int. 2019; 126: 19–26. https://doi.org/10.1016/j.neuint.2019.02.020 PMID: 30831216

92. Benkler C, O'Neil AL, Slepian S, Qian F, Weinreb PH, Rubin LL. Aggregated SOD1 causes selective death of cultured human motor neurons. Sci Rep. 2018; 8: 1–14.

93. Banci L, Bertini I, D'Amelio N, Gaggelli E, Libralesso E, Matecko I, et al. Fully metallated S134N Cu,Zn-superoxide dismutase displays abnormal mobility and intermolecular contacts in solution. J Biol Chem. 2005; 280: 35815–35821. https://doi.org/10.1074/jbc.M506637200 PMID: 16105836

94. Schmitt ND, Agar JN. Parsing disease-relevant protein modifications from epiphenomena: perspective on the structural basis of SOD1-mediated ALS. J Mass Spectrom. 2017; 52: 480–491. https://doi.org/10.1002/jms.3953 PMID: 28558143

95. Molnar KS, Karabacak NM, Johnson JL, Wang Q, Tiwari A, Hayward LJ, et al. A common property of amyotrophic lateral sclerosis-associated variants. Destabilization of the copper/zinc superoxide dismutase electrostatic loop. J Biol Chem. 2009; 284: 30965–30973. https://doi.org/10.1074/jbc.M109.023945 PMID: 19635794

96. Krebs BB, De Mesquita JF, Robberecht W, Philips T, Mancuso R, Navarro X, et al. Amyotrophic Lateral Sclerosis Type 20—In Silico Analysis and Molecular Dynamics Simulation of hnRNPA1. PLoS One. 2016. https://doi.org/10.1371/journal.pone.0158939 PMID: 27414033

97. Jean-Quartier C, Jeanquartier F, Jurisica I, Holzinger A. In silico cancer research towards 3R. BMC Cancer. 2018; 18: 1–12.

98. Raies AB, Bajic VB. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Comput Mol Sci. 2016; 6: 147–172. https://doi.org/10.1002/wcms.1240 PMID: 27066112

99. Zhang F, Ström AL, Fukada K, Lee S, Hayward LJ, Zhu H. Interaction between familial Amyotrophic Lateral Sclerosis (ALS)-linked SOD1 mutants and the dynein complex. J Biol Chem. 2007; 282: 16691–16699. https://doi.org/10.1074/jbc.M609743200 PMID: 17403682