

## RESEARCH ARTICLE

# Molecular dynamics and protein frustration analysis of human fused in Sarcoma protein variants in Amyotrophic Lateral Sclerosis type 6: An *In Silico* approach

L. F. S. Bonet , J. P. Loureiro, G. R. C. Pereira, A. N. R. Da Silva , J. F. De Mesquita \*

Department of Genetics and Molecular Biology, Laboratory of Bioinformatics and Computational Biology, Federal University of the State of Rio de Janeiro, Rio de Janeiro, Brazil

\* [joelma.mesquita@unirio.br](mailto:joelma.mesquita@unirio.br)



## OPEN ACCESS

**Citation:** Bonet LFS, Loureiro JP, Pereira GRC, Da Silva ANR, De Mesquita JF (2021) Molecular dynamics and protein frustration analysis of human fused in Sarcoma protein variants in Amyotrophic Lateral Sclerosis type 6: An *In Silico* approach. PLoS ONE 16(9): e0258061. <https://doi.org/10.1371/journal.pone.0258061>

**Editor:** Maria Gasset, Consejo Superior de Investigaciones Cientificas, SPAIN

**Received:** July 14, 2021

**Accepted:** September 16, 2021

**Published:** September 29, 2021

**Copyright:** © 2021 Bonet et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting information files](#).

**Funding:** This study was supported in the form of funding by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (<http://www.faperj.br/>) awarded to LFSB, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (<http://www.capes.gov.br/>) awarded to LFSB, Financiadora de Estudos e

## Abstract

Amyotrophic lateral sclerosis (ALS) is the most frequent adult-onset motor neuron disorder. The disease is characterized by degeneration of upper and lower motor neurons, leading to death usually within five years after the onset of symptoms. While most cases are sporadic, 5%-10% of cases can be associated with familial inheritance, including ALS type 6, which is associated with mutations in the Fused in Sarcoma (*FUS*) gene. This work aimed to evaluate how the most frequent ALS-related mutations in *FUS*, R521C, R521H, and P525L affect the protein structure and function. We used prediction algorithms to analyze the effects of the non-synonymous single nucleotide polymorphisms and performed evolutionary conservation analysis, protein frustration analysis, and molecular dynamics simulations. Most of the prediction algorithms classified the three mutations as deleterious. All three mutations were predicted to reduce protein stability, especially the mutation R521C, which was also predicted to increase chaperone binding tendency. The protein frustration analysis showed an increase in frustration in the interactions involving the mutated residue 521C. Evolutionary conservation analysis showed that residues 521 and 525 of human *FUS* are highly conserved sites. The molecular dynamics results indicate that protein stability could be compromised in all three mutations. They also affected the exposed surface area and protein compactness. The analyzed mutations also displayed high flexibility in most residues in all variants, most notably in the interaction site with the nuclear import protein of *FUS*.

## Introduction

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disease characterized by damage to upper and lower motor neurons, resulting in progressive muscle atrophy, paralysis, and usually death due to respiratory failure within 1 to 5 years after the onset of symptoms [1]. With an annual incidence of 2 per 100000 people worldwide, ALS is the most common adult-onset motor neuron disorder [2]. Despite the vast literature on the disease, its physiological basis is

Projetos (FINEP) (<http://www.finep.gov.br/>) awarded to LFSB, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (<http://cnpq.br/>) awarded to LFSB, Universidade Federal do Estado do Rio de Janeiro to LFSB, and in the form of material support by NVIDIA Corporation awarded to LFSB. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have read the journal's policy and the authors of this manuscript have the following competing interests: LFSB received material support from NVIDIA for this study. This does not alter our adherence to PLOS ONE policies on sharing data and materials. There are no patents, products in development or marketed products associated with this research to declare.

still poorly understood, being considered a multifactorial condition in about 90–95% of cases [3]. The remaining 5–10% can be associated with familial inheritance, termed familial ALS (fALS), with several known genes relating to ALS either causatively or as susceptibility factors [4].

Among these genes, the Fused in Sarcoma (*FUS*) is associated with approximately 5% of fALS cases [5]. *FUS* is a DNA- and RNA-binding protein associated with splicing regulation, stress granule formation, mRNA transport, and DNA repair [6]. *FUS* is a nuclear protein composed of 526 amino acid residues and seven domains [7]. Notably, most of the *FUS* mutations associated with ALS development occur in the protein's C-terminal region [8], specifically in its nuclear localization signal (NLS) domain. Among these non-synonymous single nucleotide polymorphisms (nsSNP), variant R521C is the most frequent mutation in ALS type-6 patients [5,9,10]. Variants R521H and P525L follow closely as second and third most common [6] mutations in ALS type 6 cases. Additionally, P525L has been associated with a more aggressive disease phenotype [6].

There are very few treatment options for ALS patients, with merely two approved drugs, both of which increase patient lifetime by 3–5 months [11]. The lack of knowledge regarding the pathogenesis of the disease contributes to the absence of effective treatments [12]. Nevertheless, advances in computational technologies and applied bioinformatics have allowed scientists a new perspective on the molecular basis of diseases, which is an essential piece in rational drug design and precision medicine [13]. Here, we performed molecular dynamics to better understand how ALS-related mutations in *FUS* behave inside a simulated cellular environment and how this behavior can be associated with fALS development. Furthermore, we used several computational algorithms to predict the effects of mutations R521C, R521H, and P525L and analyze the evolutionary conservation of amino acids in the human *FUS* protein.

## Materials and methods

We used the methodology previously described by our group [13–15] to analyze the effects that mutations R521C, R521H, and P525L have over the human *FUS* protein.

## Ethics statement

This research does not involve human participants or human tissues. All data used as a starting point for our study were obtained from freely available databases, such as Protein Data Bank and UniProt. No participant consent or ethics committee approval is needed in this case.

## Sequence and structure retrieval

The wild-type sequence of the human *FUS* protein, as well as a list of its naturally occurring variants, were obtained from the UNIPROT [16] database [ID: P35637]. The structural models for the human *FUS* protein were obtained from the Protein Data Bank [17]. As there is no model for the complete protein structure, two fragments containing the nuclear localization signal of *FUS* wild-type [ID: 4FDD] [18] and variant P525L [ID: 7CYL] [19] were selected for computational analysis after meticulous research showed that molecular dynamics simulations (MD) involving protein fragments were viable [20,21]. Both fragments ranged from position 507 to 526 of the human *FUS* protein.

## nsSNP analysis

Ten algorithms were used to predict whether the variants could cause deleterious functional alterations on the protein: PredictSNP [22], MAPP [23], PhD-SNP [24], Polyphen-1 [25],

Polyphen-2 [26], SIFT [27], SNAP [28], SNPs&GO [29], PMut [30] and PROVEAN [31]. Additional factors, such as aggregation tendency (TANGO), amyloid propensity (WALTZ), chaperone binding tendency (LIMBO), and protein stability were analyzed using the SNPeffect [32] and FoldX [33] algorithms.

### Protein frustration analysis

To investigate if the studied mutations caused changes in protein frustration levels, the wild-type structure and the variant structures were analyzed in the Protein Frustratometer 2 server [34]. The variant structures (R521C, R521H, and P525L) were generated using the model obtained from the PDB and the *in silico* mutagenesis extension of the Visual Molecular Dynamics software [35], version 1.9.3.

### Evolutionary conservation analysis

The CONSURF algorithm was used to evaluate the evolutionary conservation of amino acid residues in the FUS protein [36]. UNIPROT was selected as the database, and the homolog search algorithm used was HMMER, in 1 iteration and an E-value cutoff of 0.0001. Maximum and minimum identities for homologs were set as 95% and 35%, respectively. The alignment method used to build the Multiple Sequence Alignment was MAFFT-L-INS-i. The calculation method used was the Bayesian method, and the evolutionary substitution model was set to default.

### Molecular dynamics

The wild-type and variant models of the FUS protein were submitted to molecular dynamics using the Linux GROMACS [37] package, version 5.0.7. Using the Amber99SB-ILDN [38] force field, the protein fragment was solvated in TIP3P [39] water molecules inside a triclinic box with dimensions of 56Å, 34Å, and 31Å. The dynamics were neutralized by the addition of Na<sup>+</sup> and Cl<sup>-</sup> ions. The systems underwent energy minimization using the steepest descent method. After minimization, NVT (constant number of particles, volume, and temperature) and NPT (constant number of particles, pressure, and temperature) ensembles were performed, both at a pressure of 1atm and temperature of 300K for 100ps. The Linear Constraint Solver (LINCS) algorithm [40] was utilized to contain covalent bonds, while the electrostatic interactions were computed using the Particle Mesh Ewald (PME) method [41]. Production simulations were performed at 300K and 1atm for 300ns for the wild-type FUS and variants.

Three independent experiments were performed for the wild-type protein and its variants. The resulting trajectories were analyzed using the following GROMACS distribution programs: *gmx rms*, *gmx rmsf*, *gmx gyrate* and *gmx sasa*, which generated parameter values for Root-mean-Square Deviation (RMSD), Root-mean-Square Fluctuation (RMSF), radius of gyration (Rg), and solvent-accessible surface area (SASA). The ggplot2 package from the R software was used to generate the MD simulation graphs and calculate the mean values for the wild-type FUS and variants among the triplicates.

## Results and discussion

### nsSNP analysis

Most of the algorithms accurately predicted the mutations R521C, R521H, and P525L to be deleterious, as shown in Table 1. PredictSNP, MAPP, Polyphen-1, Polyphen-2, SIFT, SNAP, and Pmut were the most accurate predictors, classifying all three mutations as deleterious,

**Table 1. Functional effect prediction of FUS mutations by ten nsSNP prediction algorithms.**

Prediction algorithms	Protein Variants		
	R521C	R521H	P525L
PredictSNP	Deleterious [72%] <sup>1</sup>	Deleterious [72%] <sup>1</sup>	Deleterious [65%] <sup>1</sup>
MAPP	Deleterious [77%] <sup>1</sup>	Deleterious [56%] <sup>1</sup>	Deleterious [86%] <sup>1</sup>
PHD-SNP	Neutral [78%] <sup>1</sup>	Neutral [72%] <sup>1</sup>	Neutral [78%] <sup>1</sup>
Polyphen-1	Deleterious [74%] <sup>1</sup>	Deleterious [59%] <sup>1</sup>	Deleterious [59%] <sup>1</sup>
Polyphen-2	Deleterious [55%] <sup>1</sup>	Deleterious [55%] <sup>1</sup>	Deleterious [56%] <sup>1</sup>
SIFT	Deleterious [79%] <sup>1</sup>	Deleterious [79%] <sup>1</sup>	Deleterious [79%] <sup>1</sup>
SNAP	Deleterious [89%] <sup>1</sup>	Deleterious [81%] <sup>1</sup>	Deleterious [56%] <sup>1</sup>
SNPs&GO	Neutral [0.114] <sup>2</sup>	Neutral [0.072] <sup>2</sup>	Neutral [0.039] <sup>2</sup>
PMut	Deleterious [87%] <sup>3</sup>	Deleterious [86%] <sup>3</sup>	Deleterious [91%] <sup>3</sup>
PROVEAN	Deleterious [-4.37] <sup>4</sup>	Neutral [-2.37] <sup>4</sup>	Deleterious [-6.54] <sup>4</sup>

<sup>1</sup>The numbers in brackets represent the expected accuracy of the classification of each algorithm, as provided by PredictSNP.

<sup>2</sup>The numbers in brackets represent the disease probability according to SNPs&GO.

<sup>3</sup>The numbers in brackets represent the reliability of the classification of PMut.

<sup>4</sup>The numbers in brackets represent the score behind PROVEAN's classification.

<https://doi.org/10.1371/journal.pone.0258061.t001>

while PHD-SNP and SNPs&GO classified all three as neutral. Interestingly, PROVEAN classified variants R521C and P525L as deleterious but R521H as neutral.

Despite affecting the same amino acid position, the R521C mutation exchanges a positively charged and basic amino acid for a non-charged and polar amino acid. On the other hand, variant R521H swaps a positively charged and basic amino acid for another positively charged and basic amino acid, which would possibly have less impact on protein structure and function. Particularly regarding that alteration in the protein net charge can affect its ability to interact with other molecules [42].

Moreover, two algorithms classified as neutral three mutations that are known to be disease-related, reinforcing the importance of using several prediction algorithms. Since these algorithms use different parameters to make predictions, using several algorithms can yield more reliable results, as previously shown by our group [43].

Table 2 shows that none of the mutations were predicted to affect the aggregation tendency (TANGO) or amyloid propensity (WALTZ) of FUS. The human FUS protein is composed of seven domains (Fig 1), a prion-like region in its N-terminus that is rich in low-complexity (LC) amino acids, three regions with Arg-Gly-Gly motifs separating an RNA recognition

**Table 2. SNPeffect and FoldX effect predictions of FUS mutations.**

Predictive algorithms	Protein Variants		
	R521C	R521H	P525L
TANGO	No effect [0.00] <sup>1</sup>	No effect [0.00] <sup>1</sup>	No effect [0.00] <sup>1</sup>
WALTZ	No effect [0.00] <sup>2</sup>	No effect [0.00] <sup>2</sup>	No effect [0.00] <sup>2</sup>
LIMBO	Increases [151.11] <sup>3</sup>	No effect [28.67] <sup>3</sup>	No effect [47.73] <sup>3</sup>
FoldX	Greatly reduces [-6.03 kcal/mol] <sup>4</sup>	Reduces [-3.9 kcal/mol] <sup>4</sup>	Reduces [-3.78 kcal/mol] <sup>4</sup>

<sup>1</sup>TANGO evaluates aggregation tendency, the number in brackets represents the numerical dTANGO value behind the algorithm's classification.

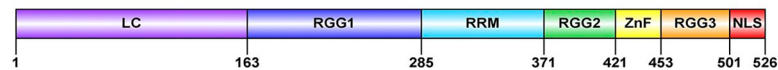
<sup>2</sup>WALTZ evaluates amyloid propensity, the number in brackets represents the numerical dWALTZ value behind the algorithm's classification.

<sup>3</sup>LIMBO evaluates chaperone binding tendency, the number in brackets represents the numerical dLIMBO value behind the algorithm's classification.

<sup>4</sup>FoldX evaluates protein stability, the number in brackets represents the numerical ddG value behind the algorithm's classification.

<https://doi.org/10.1371/journal.pone.0258061.t002>

## FUS Protein Domains



**Fig 1. Graphic representation of the domains that compose the FUS protein.** LC represents the low complexity, prion-like domain; RGG1 represents the first R-G-G pattern-rich domain; RRM represents the RNA-recognition motif; RGG2 represents the second R-G-G pattern-rich domain; ZnF represents the zinc-finger domain; RGG3 represents the third R-G-G pattern-rich domain, and NLS represents the nuclear localization signal.

<https://doi.org/10.1371/journal.pone.0258061.g001>

motif, and a RanBP2-type zinc-finger domain [44], and a Pro-Tyr nuclear localization signal (PY-NLS) in its C-terminus [7]. Studies have shown that both the prion-like LC region [45] and the RGG motif regions [46] contribute to the self-association tendencies of the protein. These SNPeffect results corroborate this idea; since these mutations occur in the NLS, they would likely have no direct effect on aggregation tendency, according to the algorithm. Similarly, amyloid fibrils have been observed only in the LC region [47] and are also unlikely to be directly affected by the studied mutations.

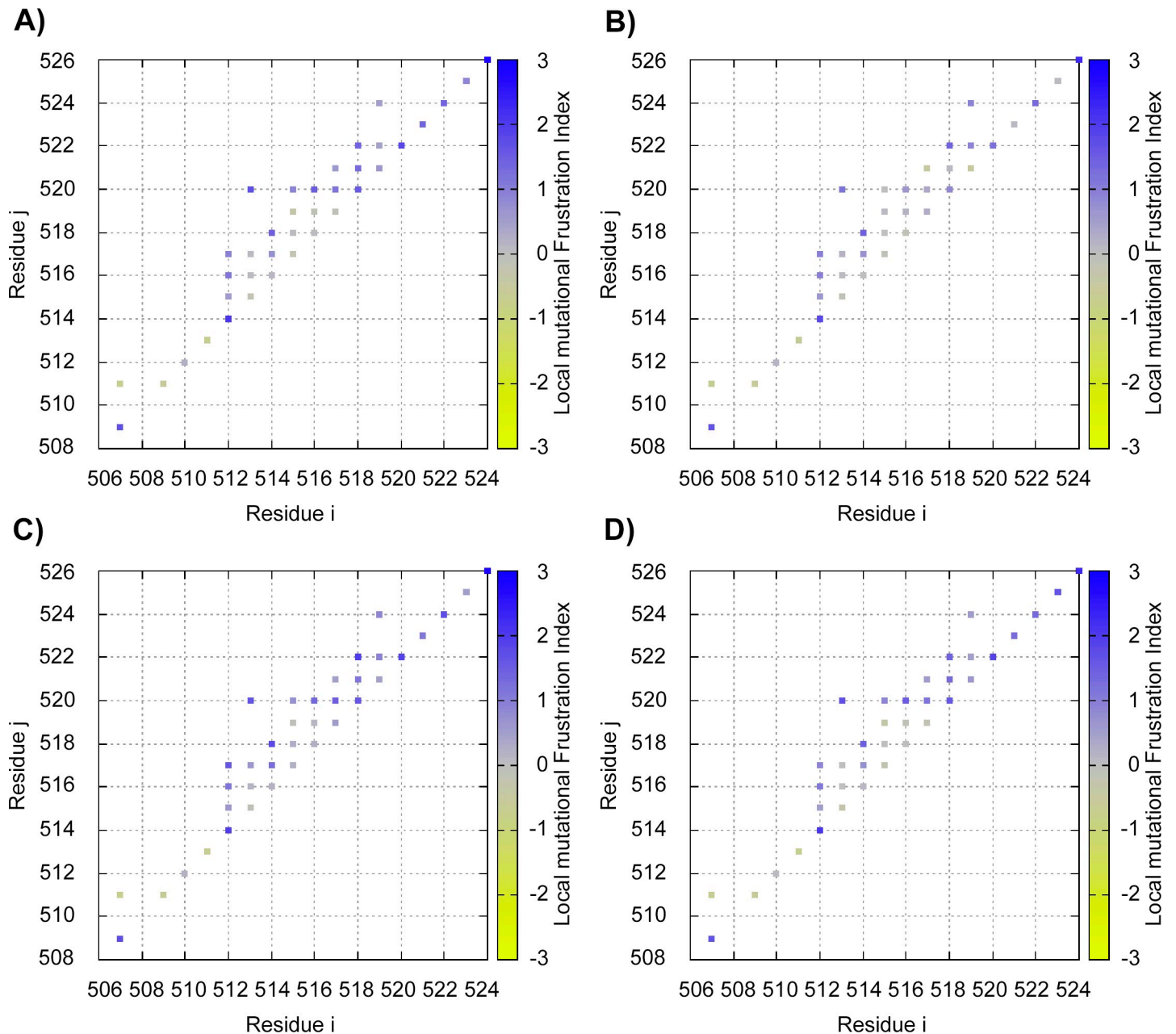
A hallmark of ALS is the formation of cytosolic aggregates. The composition of these aggregates depends mainly on the genetic mutation that is observed in the patient [48]. Aside from the mutated protein, other components can be found in these inclusions. For instance, misfolded wild-type TDP-43 is present in 97% of ALS cases [49]. Chaperones, specifically protein disulfide-isomerase (PDI), have also been observed as components in some FUS-positive inclusions [50], although the role of chaperones in ALS is still poorly understood. SNPeffect predicted an increase in chaperone binding tendency in mutation R521C, while R521H and P525L were unaffected.

## Protein frustration analysis

The results obtained from SNPeffect led us to investigate protein frustration in FUS since chaperones can mitigate protein frustration in the energy landscape of protein folding [51]. Protein frustration was best conceptualized by Ferreiro [52], where the author gives the example of a network of magnets that attract and repel each other according to their spins and the resulting energy state between pairs of magnets is termed frustration. Similarly, the amino acid residues that compose a protein have electrostatic charges and, in some cases, polarity, so the same concept can be applied.

As shown in Fig 2, there is little difference in mutational frustration in the variants compared to the wild-type protein, except for variant R521C. In this case, there is a noticeable increase in frustration in interactions involving the mutated residue. This indicates that other amino acids would be more favorable in that position than the mutated cysteine (C). This change in frustration could also corroborate the SNPeffect results, as the elevated level of protein frustration could increase the need for chaperones to mitigate the frustration [51]. Interestingly, Farg *et al.* (2012) study indicated that protein disulfide-isomerases (PDIs) colocalizes with inclusions of R521C mutant FUS. PDI proteins are general protein chaperones and are also responsible for disulfide bond formation. High levels of PDIs were observed in ALS patients inside TDP-43, SOD1, and VAPB inclusions, playing a protective role against toxicity and misfolding. Thus, it is more likely that PDIs act as a chaperone refolding misfolded FUS than a disulfide bond inducer, given that this type of interaction has not yet been described before for FUS protein. However, it is possible that FUS protein, which contains four cysteine residues, may form non-native disulfide bonds, especially after R521C mutation providing an additional cysteine. Non-native bonds were already observed in mutated forms of SOD1 and TDP-43 proteins [50].

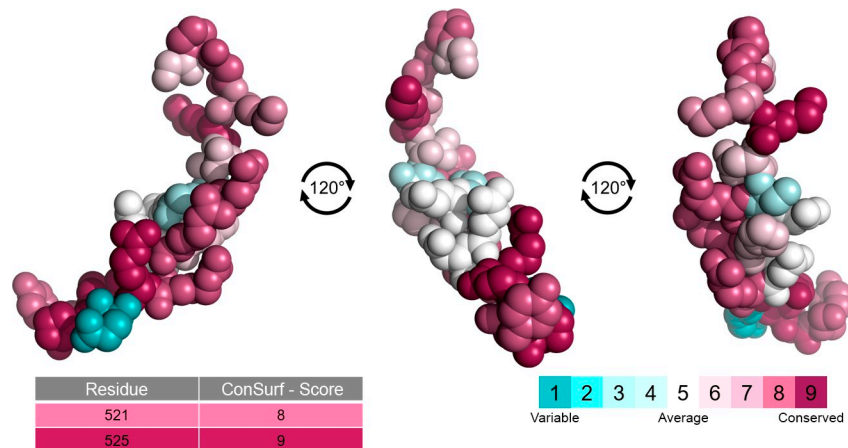




**Fig 2. Mutational frustration analysis.** Levels of mutational frustration by each residue-residue interaction. The colored scale represents the levels of frustration, in which blue represents minimal frustration, gray represents neutral frustration, and yellow represents high frustration. A) Wild-type FUS. B) R521C variant. C) R521H variant. D) P525L variant.

<https://doi.org/10.1371/journal.pone.0258061.g002>

Interestingly, these results show that residue 521 interacts with residue 517, but also with residues 518, 519, and 523, which are binding residues for Karyopherin- $\beta$ 2 (Kap $\beta$ 2), the importin that transports FUS, and other proteins, into the nucleus [20]. While it is unlikely that such electrostatic interactions would, by themselves, dramatically alter the structural stability of a protein [53], the significance of the affected residues could make them delicate to change. Especially since hydrophobicity is a hallmark of PY-NLSs [54] like the one present in FUS, and the amino acid cysteine is more hydrophobic.



**Fig 3. Evolutionary conservation analysis.** Amino acid conservation analysis of wild-type FUS protein fragment (residue 507 to 526) represented in a space-filling model and colored according to conservation scores. Conservation scores of the residues where R521C, R521H, and P525L mutations occur are highlighted in the table. CONSURF color scheme representing conservation levels. Cyan represents more variable positions, while bordeaux represents more conserved positions.

<https://doi.org/10.1371/journal.pone.0258061.g003>

### Evolutionary conservation analysis

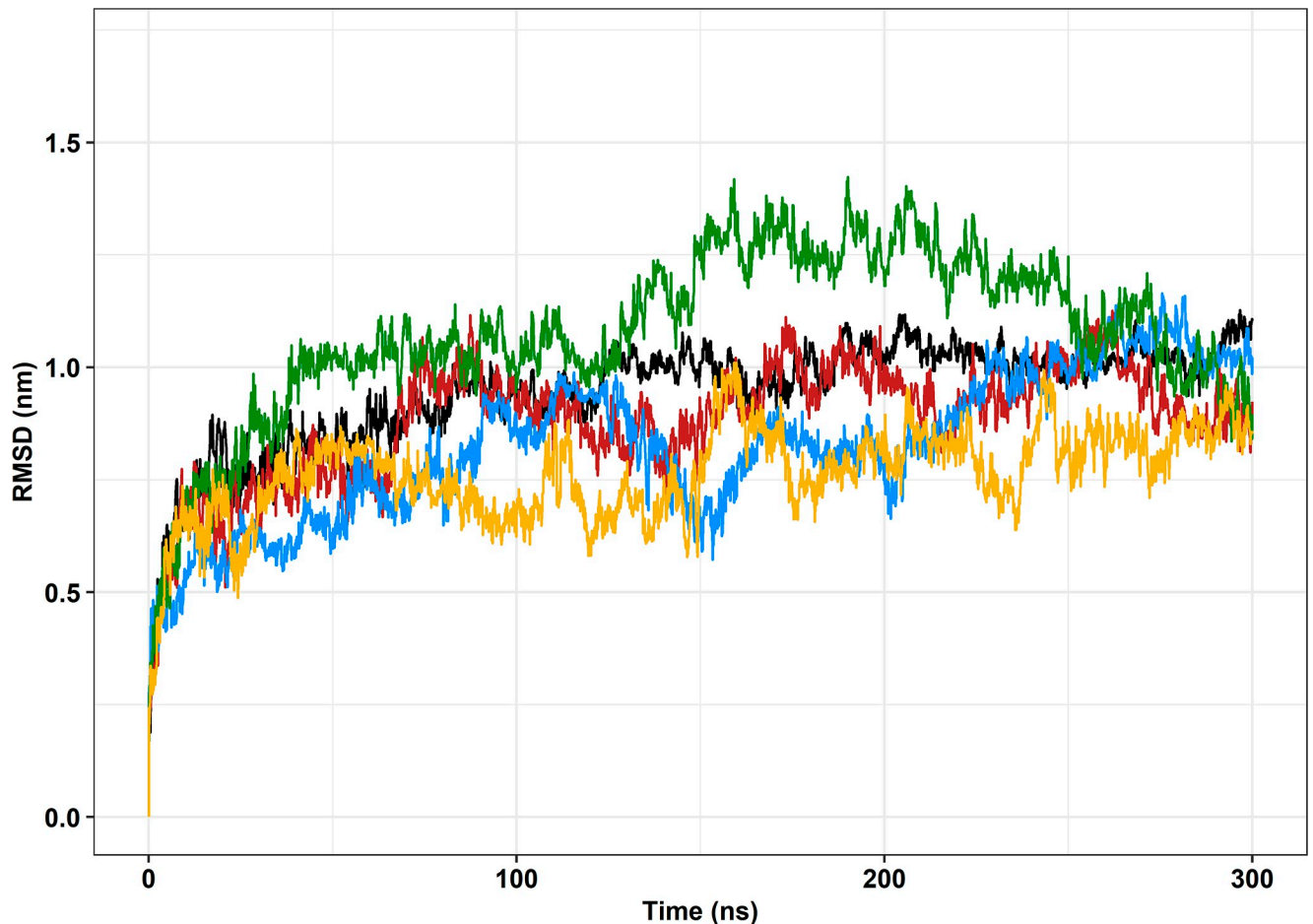
CONSURF server analyses the evolutionary dynamics of amino acid residues between homologous sequences and gives each residue a score ranging from 1, more variable, to 9, more conserved [36]. Results showed that the amino acid residues in the NLS of the protein tend to be more conserved than variable, as shown in Fig 3. This is expected due to the relevance of the domain since areas that have an important role in protein function or structure tend to be more conserved [36]. Indeed, seven out of the ten residues in the NLS that interact with Kap $\beta$ 2 were given conservations scores of seven or higher. Mutations in the NLS of FUS have been shown to weaken the affinity between FUS protein and the importin, and this loss of affinity causes the mislocalization of the protein to the cytosol, an effect that has been linked to ALS [55].

The amino acid residues where the studied mutations occur were also scored with high conservation levels by the algorithm. Residue 521 received a score of 8, while residue 525 received a score of 9. This corroborates that mutations occurring in more conserved sites tend to be more deleterious [56], seeing that variant P525L is frequently associated with a more aggressive disease phenotype [6]. Bulbar onset is more frequent in P525L than in other *FUS* mutations [6], which has been shown to have a smaller time frame between disease onset and death than spinal onset [57]. Additionally, the youngest case ever registered of ALS type 6 reported a carrier of the P525L *FUS* mutation with disease onset at merely 11 years of age [58], 28 years sooner than the median onset age for the condition [6].

### Molecular dynamics

During MD simulations, the trajectories of atoms and molecules are obtained by computational calculation of Newtonian equations of motion and application of different force fields [59]. Thusly, MDs can provide relevant information that would not be so easily discovered via experimental approaches [60], which tend to be much more costly. This type of information is also fundamental to improve efficiency and reduce the blindness of drug discovery [59].

RMSD is the average root-mean-square displacement of atoms between a reference structure and a dynamic structure, calculated at every frame of the simulation [37]. The reference



**Fig 4. Root mean square deviation (RMSD) of the backbone atoms of the protein as a function of time.** The wild-type FUS protein [PDB ID: 4FDD] is represented in black, variant R521C is represented in blue, variant R521H is represented in green, variant P525L is represented in red, and variant P525L-PDB [ID: 7CYL] is represented in yellow.

<https://doi.org/10.1371/journal.pone.0258061.g004>

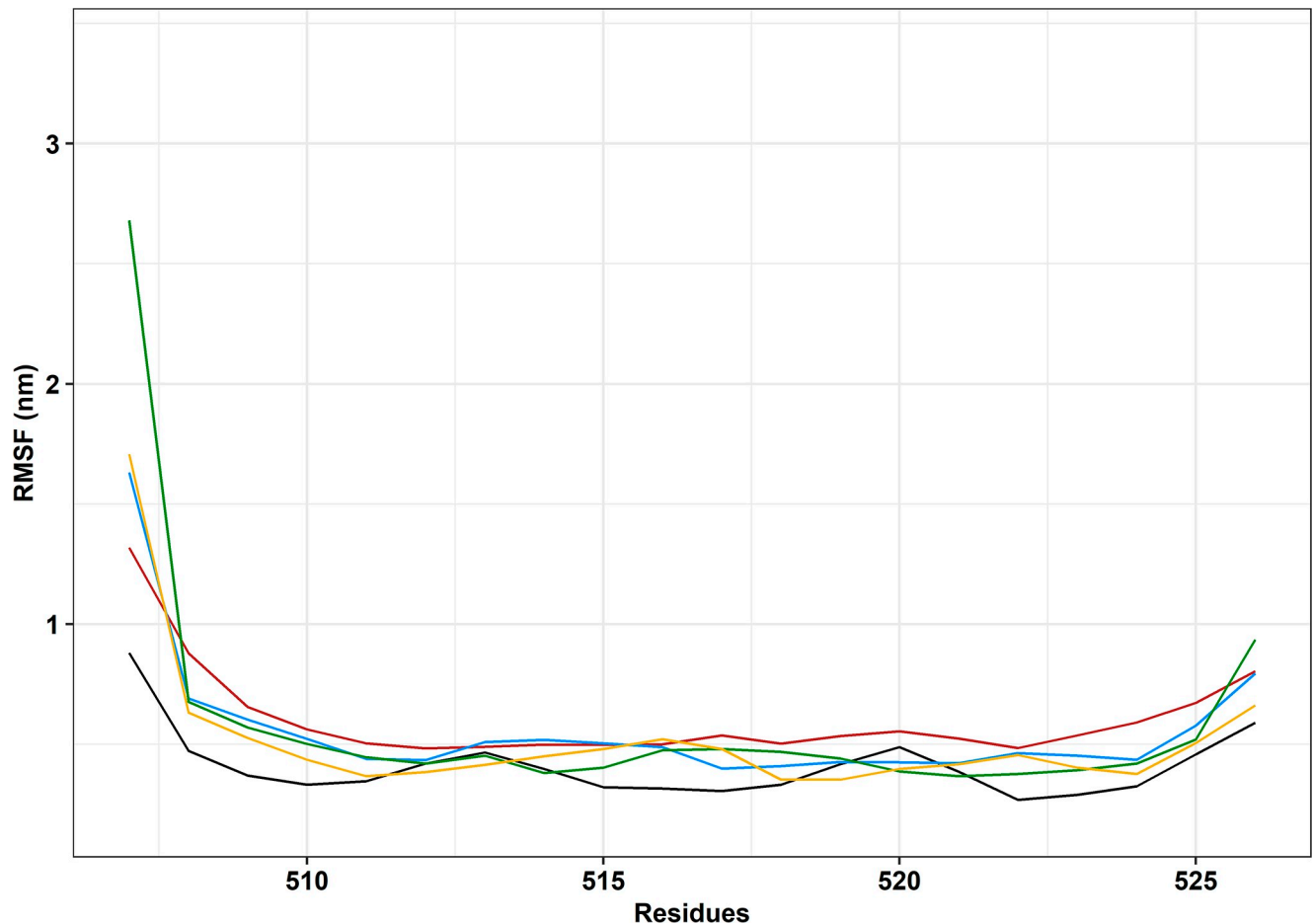
structure is usually the first frame of the structure in the simulation, which was the case in this study. RMSD, thus, is a good measure of structural similarity throughout the simulation [61]. When a molecular system approaches equilibrium, the protein structure generally floats around an average stable conformation, and consequently, the RMSD values tend to form a plateau [62].

Wild-type FUS protein seemed to reach an equilibrium at around 125ns into the simulation when the RMSD value started fluctuating around 1nm, as shown in Fig 4. Variants P525L (induced *in silico*) and P525L-PDB [ID: 7CYL] were the only mutations to come close to an equilibrated state, which can be observed from the onwards of 160ns. These variants also presented similar behaviors throughout the simulations. R521C and R521H stand out by generally having lower and higher RMSD values, respectively, and seemingly did not achieve equilibrium during the simulations.

Root mean square fluctuation (RMSF) is similar to RMSD in that it measures the average displacement of atoms, but it does so relatively to individual residues throughout the simulation. Simply put, RMSF analyses flexibility differences for the amino acid residues [63].

Overall, variant P525L (induced *in silico*) was shown to have residues slightly more flexible than the other variants (Fig 5). Residue 507 was shown to be the most different between the





**Fig 5. Root mean square fluctuation (RMSF) of the backbone atoms of the protein as a function of each amino acid residue.** The wild-type FUS protein is represented in black, variant R521C is represented in blue, variant 521H is represented in green, variant P525L is represented in red, and variant P525L-PDB [ID: 7CYL] is represented in yellow.

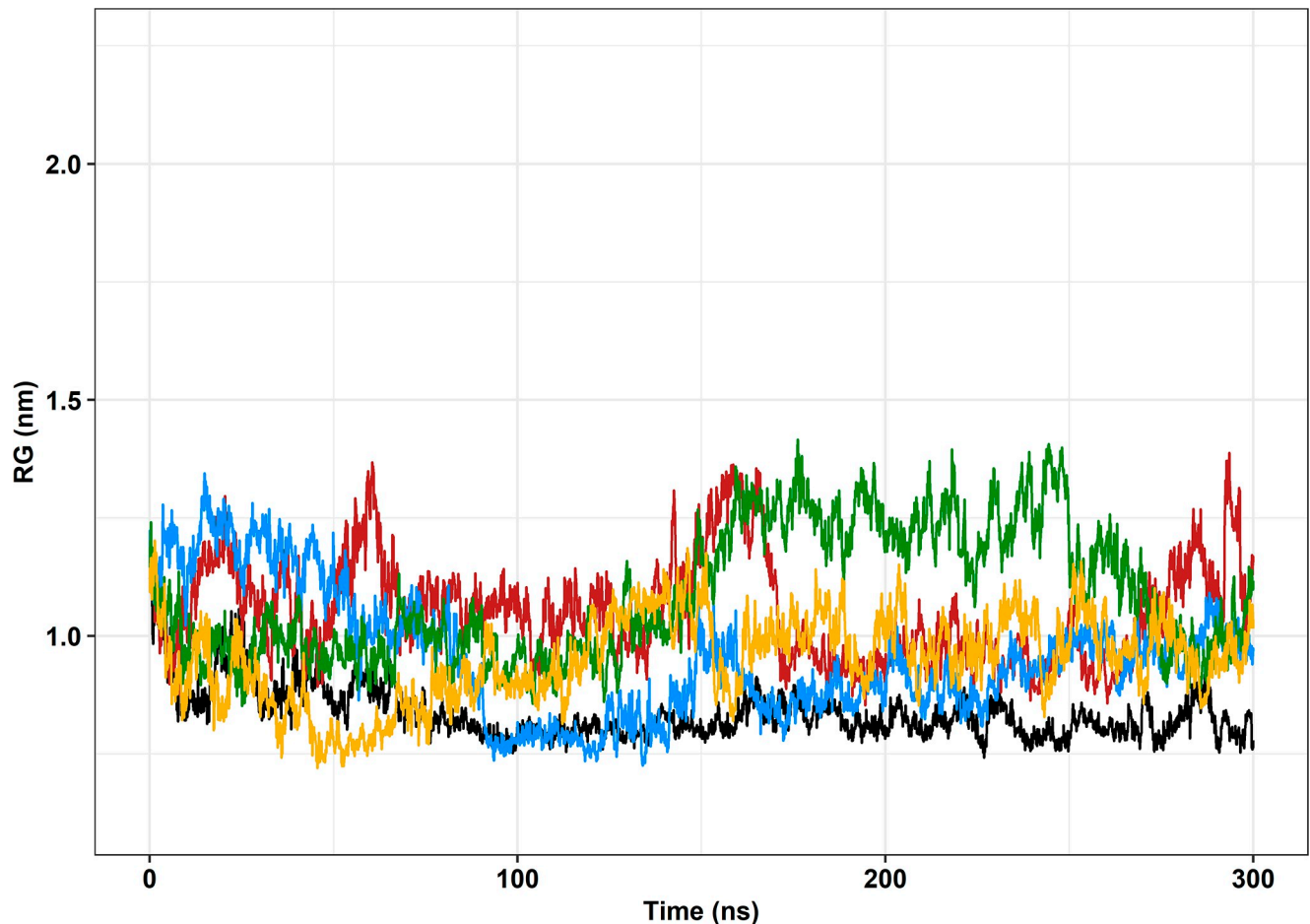
<https://doi.org/10.1371/journal.pone.0258061.g005>

variants and wild-type. While in the simulations of the P525L (induced *in silico*), P525L-PDB [ID: 7CYL], and R521C variants, residue 507 was approximately 33%, 60%, and 60% more flexible than the wild-type, respectively, this residue was nearly three times as flexible as the wild-type in the R521H simulation. Conversely, residue 513 showed the least variation of flexibility among all proteins, with all four receiving an RMSF value of approximately 0.5nm. Curiously, residue 507 is one of the sites that interact with Kap $\beta$ 2 [20].

Aside from residue 507, none of the other amino acid residues that interact with Kap $\beta$ 2 showed any distinct behavior other than that flexibility was generally higher in the variants than in the wild-type. Notably, alterations in flexibility can be very disruptive to proteins, putting at risk protein binding affinity and specificity [64], stability [65], and even misfolding [66], which is, by itself, linked to many diseases.

Radius of gyration (RG) is the mass-weighted root-mean-square distance of the atoms in a structure to their center of mass [20]. Thus, the radius of gyration can indicate the overall dimensions of a protein throughout an MD [63].

Wild-type FUS shows steady RG values, especially after 50ns, where the RG values fluctuate around 0.6nm until the end of the trajectory (Fig 6). Variant R521C presented steady RG values around 150ns, with values similar to those of the wild-type. Variant P525L-PDB [ID:



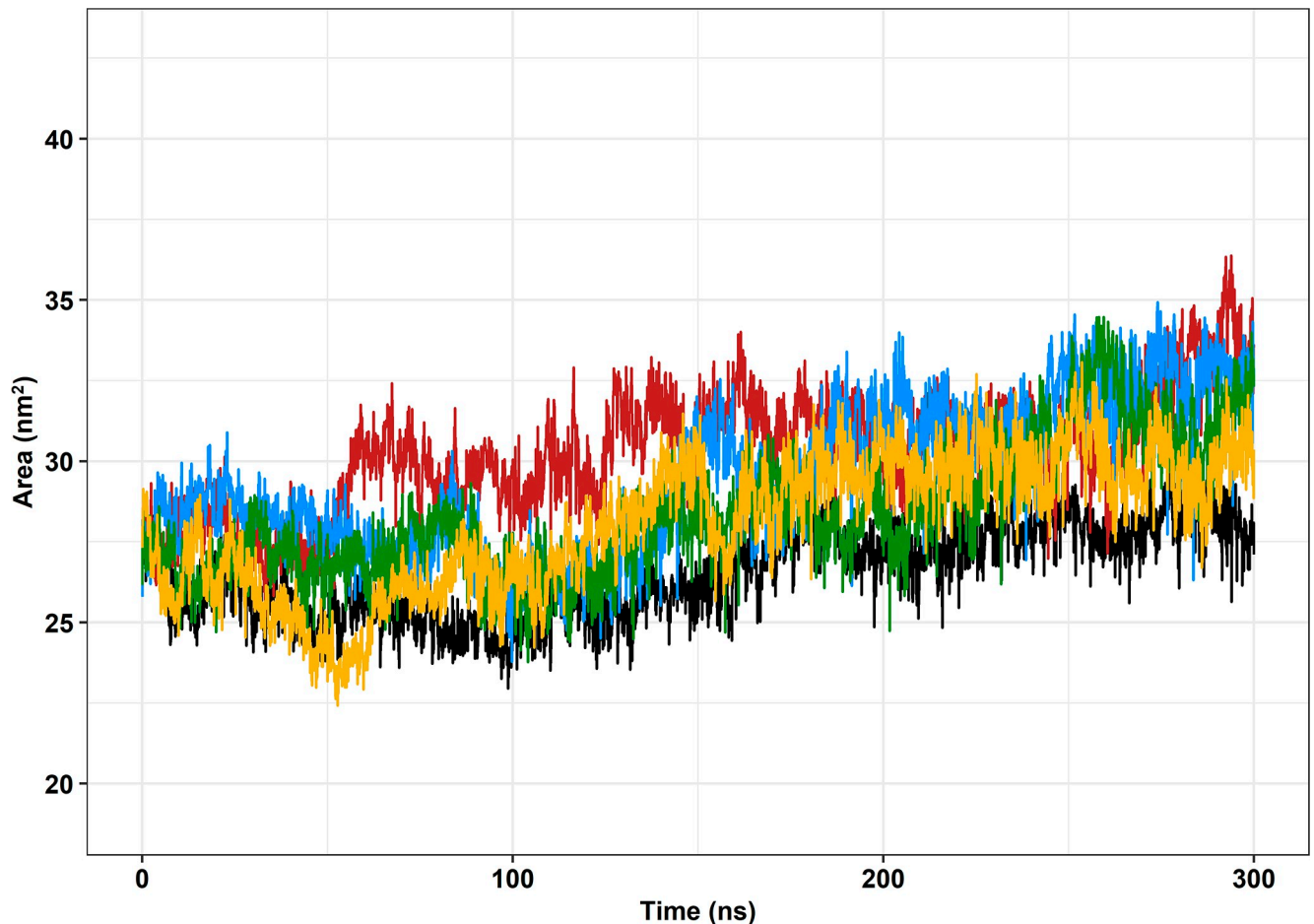
**Fig 6. Radius of gyration (RG) of the backbone atoms of the protein as a function of time.** The wild-type FUS protein is represented in black, variant R521C is represented in blue, variant 521H is represented in green, variant P525L is represented in red, and variant P525L-PDB [ID: 7CYL] is represented in yellow.

<https://doi.org/10.1371/journal.pone.0258061.g006>

7CYL] also presented steady RG values around 150ns, however, with values slightly higher than those of the wild-type. Interestingly, variants R521H and P525L (induced *in silico*) have seemingly opposite behavior between 160ns and 290ns, but they do not present a steady behavior throughout the simulation. In general, R521H, P525L (induced *in silico*), and P525L-PDB [ID: 7CYL] presented less compactness than the wild-type fragment over time. Variant R521C showed similarities to the wild-type at certain points but was generally less compact, as well.

Solvent-accessible surface area (SASA) is the area of the protein that is exposed to solvent [67]. Thus, SASA can provide information about the ability of the protein to interact [63]. Interestingly, FUS tends to interact with other FUS proteins and self-assemble, forming the cytosolic inclusions that are believed to contribute to neurodegeneration in ALS [45].

Overall, all variant fragments were shown to have higher surface areas, which corroborates the RG analysis results that showed that the mutant fragments were generally less compact than the wild-type protein fragment, as seen in Fig 7. Moreover, similarly to the observed in the RMSD, and RG analyses, the SASA values for the wild-type were considerably steadier than those of the analyzed variants, which could further suggest that the variants R521C, R521H, and P525L could reduce protein stability [13]. This could also corroborate the results obtained from SNPeffect, which showed that all three mutations decrease protein stability.



**Fig 7. Solvent-accessible surface area (SASA) of the protein as a function of time.** The wild-type FUS protein is represented in black, variant R521C is represented in blue, variant R521H is represented in green, variant P525L is represented in red, and variant P525L-PDB [ID: 7CYL] is represented in yellow.

<https://doi.org/10.1371/journal.pone.0258061.g007>

The results obtained in this study showed that mutations R521C, R521H, and P525L could alter stability and flexibility for the FUS protein, which are factors essential to the function of a protein [65]. Flexibility, as previously mentioned, is an important factor in binding affinity and specificity [64]. Interestingly, the increase in flexibility observed in the studied mutations occur on several residues that interact with Kap $\beta$ 2, the protein that carries transports FUS into the nucleus. Thusly, this increase in flexibility provides a possible link to the decrease in binding affinity between mutated FUS proteins and Kap $\beta$ 2 [20]. By lowering the binding affinity between these two proteins, mutated FUS would be more likely to remain in the cytoplasm, as is observed in type 6 ALS patients [48].

MD simulations and functional predictions have been proven to be valuable for uncovering the effects of single mutations in proteins [68]. In addition to the methods presented in this paper, other powerful *in silico* approaches for the study of missense mutations are also available and worth mentioning. For example, the combination of classical MD outputs with advanced energy calculation methods (*i.e.* free energy perturbation, energy decomposition, MM/PBSA, and MM/GBSA) can provide relevant mechanistic information on the effects of mutations, particularly regarding protein interactions and structural stability [69–72]. Another widely used approach consists of combining MD simulations with unsupervised machine learning techniques such as clustering and principal component analysis. This approach allows

researchers to assess representative conformations in proteins [69] and separate biologically relevant movements (*i.e.* opening, closing, and flexing) from small irrelevant local fluctuations, thus, providing detailed information on conformational dynamics [68].

## Conclusions

Through computational techniques, we were able to investigate how mutations R521C, R521H, and P525L affect the nuclear localization signal of FUS. Most of the functional prediction algorithms were capable of accurately predicting the three mutations as deleterious to the human organism. They also showed that the mutations decrease protein stability and, in the case of R521C, increase chaperone binding tendency. Relatedly, the protein frustration analysis showed an increase in frustration between the mutated R521C residue and the ones it interacts with. The evolutionary conservation analysis showed that the three mutations occur in highly conserved and probable important protein sites. The molecular dynamics results indicate that protein stability could be compromised in all three mutations. They also indicated that flexibility is elevated in most residues in all variants, most notably in the interaction site with the nuclear import protein of FUS. The analyzed mutations also affect the exposed surface area and protein compactness. These alterations in protein stability and flexibility could negatively impact the affinity between mutated FUS proteins and Kap $\beta$ 2, leading to the formation of the FUS-positive cytoplasmic aggregates associated with type 6 ALS.

## Supporting information

**S1 Dataset. Minimal dataset.** The tables contain the mean values used to create the molecular dynamics RMSD (Sheet 1), RMSF (Sheet 2), RG (Sheet 3), and SASA (Sheet 4) graphics, as well as the individual values obtained from each replicate. Sheet 5 contains the data output from ConSurf used in the evolutionary conservation analysis. (XLSX)

## Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation by donating the Titan X Pascal GPU used in this study.

## Author Contributions

**Conceptualization:** J. F. De Mesquita.

**Data curation:** G. R. C. Pereira, A. N. R. Da Silva, J. F. De Mesquita.

**Formal analysis:** L. F. S. Bonet, J. P. Loureiro, G. R. C. Pereira, A. N. R. Da Silva, J. F. De Mesquita.

**Funding acquisition:** J. F. De Mesquita.

**Investigation:** L. F. S. Bonet, J. P. Loureiro, G. R. C. Pereira, A. N. R. Da Silva, J. F. De Mesquita.

**Methodology:** L. F. S. Bonet, J. P. Loureiro, G. R. C. Pereira, A. N. R. Da Silva, J. F. De Mesquita.

**Project administration:** J. F. De Mesquita.

**Resources:** J. F. De Mesquita.

**Software:** J. F. De Mesquita.

**Supervision:** J. F. De Mesquita.

**Validation:** L. F. S. Bonet, J. F. De Mesquita.

**Visualization:** L. F. S. Bonet, J. F. De Mesquita.

**Writing – original draft:** L. F. S. Bonet, G. R. C. Pereira, A. N. R. Da Silva, J. F. De Mesquita.

**Writing – review & editing:** L. F. S. Bonet, G. R. C. Pereira, A. N. R. Da Silva, J. F. De Mesquita.

## References

1. Callister JB, Pickering-Brown SM. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. *Exp Neurol*. 2014; 262: 84–90. <https://doi.org/10.1016/j.expneurol.2014.06.001> PMID: 24915640
2. Redler RL, Dokholyan N V. The complex molecular biology of Amyotrophic Lateral Sclerosis (ALS). 1st ed. *Progress in Molecular Biology and Translational Science*. Elsevier Inc.; 2012. <https://doi.org/10.1016/B978-0-12-385883-2.00002-3> PMID: 22482452
3. Al-Chalabi A, Hardiman O. The epidemiology of ALS: A conspiracy of genes, environment and time. *Nature Reviews Neurology*. Nature Publishing Group; 2013. pp. 617–628. <https://doi.org/10.1038/nrneuro.2013.203> PMID: 24126629
4. Corcia P, Couratier P, Blasco H, Andres CR, Beltran S, Meininger V, et al. Genetics of amyotrophic lateral sclerosis. *Revue Neurologique*. Elsevier Masson SAS; 2017. pp. 254–262. <https://doi.org/10.1016/j.neurol.2017.03.030> PMID: 28449881
5. Kwiatkowski TJ, Bosco DA, LeClerc AL, Tamrazian E, Vanderburg CR, Russ C, et al. Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science* (80-). 2009; 323: 1205–1208. <https://doi.org/10.1126/science.1166066> PMID: 19251627
6. Naumann M, Peikert K, Günther R, van der Kooij AJ, Aronica E, Hübers A, et al. Phenotypes and malignancy risk of different FUS mutations in genetic amyotrophic lateral sclerosis. *Ann Clin Transl Neurol*. 2019; 6: 2384–2394. <https://doi.org/10.1002/acn3.50930> PMID: 31682085
7. Yoshizawa T, Ali R, Jiou J, Fung HYJ, Burke KA, Kim SJ, et al. Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell*. 2018; 173: 693–705.e22. <https://doi.org/10.1016/j.cell.2018.03.003> PMID: 29677513
8. Parakh S, Atkin JD. Protein Folding Alterations in Amyotrophic Lateral Sclerosis. *Brain Res*. 2016. <https://doi.org/10.1016/j.brainres.2016.04.010> PMID: 27064076
9. Vance C, Rogelj B, Hortobágyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* (80-). 2009; 323: 1208–1211. <https://doi.org/10.1126/science.1165942> PMID: 19251628
10. Yamashita S, Mori A, Sakaguchi H, Suga T, Ishihara D, Ueda A, et al. Sporadic juvenile amyotrophic lateral sclerosis caused by mutant FUS/TLS: possible association of mental retardation with this mutation. *J Neurol*. 2012; 259: 1039–44. <https://doi.org/10.1007/s00415-011-6292-6> PMID: 22057404
11. Breiner A, Zinman L, Bourque PR. Edaravone for amyotrophic lateral sclerosis: barriers to access and lifeboat ethics. *Can Med Assoc J*. 2020; 192: E319–E320. <https://doi.org/10.1503/cmaj.191236> PMID: 32392516
12. Petrov D, Mansfield C, Moussy A, Hermine O. ALS clinical trials review: 20 years of failure. Are we any closer to registering a new treatment? *Front Aging Neurosci*. 2017; 9: 68. <https://doi.org/10.3389/fnagi.2017.00068> PMID: 28382000
13. 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
14. Da Silva ANR, Pereira GRC, Moreira LGA, Rocha CF, De Mesquita JF. SOD1 in amyotrophic lateral sclerosis development—in silico analysis and molecular dynamics of A4F and A4V variants. *J Cell Biochem*. 2019; 120: 17822–17830. <https://doi.org/10.1002/jcb.29048> PMID: 31134679
15. Pereira GRC, De Azevedo Abraham Vieira B, De Mesquita JF. Comprehensive in silico analysis and molecular dynamics of the superoxide dismutase 1 (SOD1) variants related to amyotrophic lateral sclerosis. *PLoS One*. 2021; 16: 1–27. <https://doi.org/10.1371/journal.pone.0247841> PMID: 33630959
16. Consortium UniProt. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021; 49: D480–D489. <https://doi.org/10.1093/nar/gkaa1100> PMID: 33237286



17. 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
18. Zhang ZC, Chook YM. Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc Natl Acad Sci U S A.* 2012; 109: 12017–12021. <https://doi.org/10.1073/pnas.1207247109> PMID: 22778397
19. Gonzalez A, Mannen T, Çağatay T, Fujiwara A, Matsumura H, Niesman AB, et al. Mechanism of karyopherin-β2 binding and nuclear import of ALS variants FUS(P525L) and FUS(R495X). *Sci Rep.* 2021; 11: 1–15. <https://doi.org/10.1038/s41598-020-79139-8> PMID: 33414495
20. Swetha RG, Ramaiah S, Anbarasu A. R521C and R521H mutations in FUS result in weak binding with Karyopherinβ2 leading to Amyotrophic lateral sclerosis: a molecular docking and dynamics study. *J Biomol Struct Dyn.* 2017;35. <https://doi.org/10.1080/07391102.2016.1209130> PMID: 27381509
21. Sugihara M, Suwa M, Bondar A-N. Dynamics of bovine opsin bound to G-protein fragments. *J Struct Biol.* 2014; 188: 79–86. <https://doi.org/10.1016/j.jsb.2014.08.001> PMID: 25138898
22. 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
23. Stone EA, Sidow A. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res.* 2005. <https://doi.org/10.1101/gr.3804205> PMID: 15965030
24. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics.* 2006; 22: 2729–2734. <https://doi.org/10.1093/bioinformatics/btl423> PMID: 16895930
25. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* 2002; 30: 3894–3900. <https://doi.org/10.1093/nar/gkf493> PMID: 12202775
26. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;Chapter 7: Unit7.20. <https://doi.org/10.1002/0471142905.hg0720s76> PMID: 23315928
27. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res.* 2001; 11: 863–874. <https://doi.org/10.1101/gr.176601> PMID: 11337480
28. Bromberg Y, Yachdav G, Rost B. SNAP predicts effect of mutations on protein function. *Bioinformatics.* 2008; 24: 2397–2398. <https://doi.org/10.1093/bioinformatics/btn435> PMID: 18757876
29. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat.* 2009; 30: 1237–44. <https://doi.org/10.1002/humu.21047> PMID: 19514061
30. López-Ferrando V, Gazzo A, De La Cruz X, Orozco M, Gelpí JL. PMut: A web-based tool for the annotation of pathological variants on proteins, 2017 update. *Nucleic Acids Res.* 2017; 45: W222–W228. <https://doi.org/10.1093/nar/gkx313> PMID: 28453649
31. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics.* 2015. <https://doi.org/10.1093/bioinformatics/btv195> PMID: 25851949
32. De Baets G, Van Durme J, Reumers J, Maurer-Stroh S, Vanhee P, Dopazo J, et al. SNPEffect 4.0: On-line prediction of molecular and structural effects of protein-coding variants. *Nucleic Acids Res.* 2012; 40: D935–D939. <https://doi.org/10.1093/nar/gkr996> PMID: 22075996
33. 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
34. Parra RG, Schafer NP, Radusky LG, Tsai M-Y, Guzovsky AB, Wolynes PG, et al. Protein Frustrator 2: a tool to localize energetic frustration in protein molecules, now with electrostatics. *Nucleic Acids Res.* 2016; 44: W356–W360. <https://doi.org/10.1093/nar/gkw304> PMID: 27131359
35. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics.* 1996. pp. 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: 8744570
36. 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
37. 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
38. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct Funct Bioinforma.* 2010; 78: 1950–1958. <https://doi.org/10.1002/prot.22711> PMID: 20408171

39. Harrach MF, Drossel B. Structure and dynamics of TIP3P, TIP4P, and TIP5P water near smooth and atomistic walls of different hydroaffinity. *J Chem Phys*. 2014;140. <https://doi.org/10.1063/1.4872239> PMID: 24811640
40. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem*. 1997; 18: 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H)
41. Darden T, York D, Pedersen L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993. <https://doi.org/10.1063/1.464397>
42. Molinier-Frenkel V, Mestivier D, Castellano F. Alterations of the immunosuppressive IL411 enzyme activity induced by naturally occurring SNP/mutations. *Genes Immun*. 2016; 17: 148–52. <https://doi.org/10.1038/gene.2015.55> PMID: 26673964
43. 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
44. Nguyen CD, Mansfield RE, Leung W, Vaz PM, Loughlin FE, Grant RP, et al. Characterization of a Family of RanBP2-Type Zinc Fingers that Can Recognize Single-Stranded RNA. *J Mol Biol*. 2011; 407: 273–283. <https://doi.org/10.1016/j.jmb.2010.12.041> PMID: 21256132
45. Matsumoto T, Matsukawa K, Watanabe N, Kishino Y, Kunugi H, Ihara R, et al. Self-assembly of FUS through its low-complexity domain contributes to neurodegeneration. *Hum Mol Genet*. 2018; 27: 1353–1365. <https://doi.org/10.1093/hmg/ddy046> PMID: 29425337
46. Patel A, Lee HO, Jawerth L, Maharana S, Jahnel M, Hein MY, et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. 2015; 162: 1066–77. <https://doi.org/10.1016/j.cell.2015.07.047> PMID: 26317470
47. Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*. 2017; 171. <https://doi.org/10.1016/j.cell.2017.08.048> PMID: 28942918
48. Ciryam P, Lambert-Smith IA, Bean DM, Freer R, Cid F, Tartaglia GG, et al. Spinal motor neuron protein supersaturation patterns are associated with inclusion body formation in ALS. *Proc Natl Acad Sci U S A*. 2017; 114: E3935–E3943. <https://doi.org/10.1073/pnas.1613854114> PMID: 28396410
49. Maekawa S, Leigh PN, King A, Jones E, Steele JC, Bodi I, et al. TDP-43 is consistently co-localized with ubiquitinated inclusions in sporadic and Guam amyotrophic lateral sclerosis but not in familial amyotrophic lateral sclerosis with and without SOD1 mutations. *Neuropathology*. 2009; 29: 672–83. <https://doi.org/10.1111/j.1440-1789.2009.01029.x> PMID: 19496940
50. Farg MA, Soo KY, Walker AK, Pham H, Orian J, Horne MK, et al. Mutant FUS induces endoplasmic reticulum stress in amyotrophic lateral sclerosis and interacts with protein disulfide-isomerase. *Neurobiol Aging*. 2012; 33: 2855–68. <https://doi.org/10.1016/j.neurobiolaging.2012.02.009> PMID: 22459602
51. Bucci M. Minimizing frustration. *Nat Chem Biol*. 2017; 13: 1–1. <https://doi.org/10.1038/nchembio.2277> PMID: 27984573
52. Ferreira DU, Komives EA, Wolynes PG. Frustration in Biomolecules. *Q Rev Biophys*. 2014; 47: 285–363. <https://doi.org/10.1017/S0033583514000092> PMID: 25225856
53. Pace CN, Laurents D V., Thomson JA. pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry*. 1990; 29: 2564–2572. <https://doi.org/10.1021/bi00462a019> PMID: 2110472
54. Soniat M, Chook YM. Karyopherin- $\beta$ 2 Recognition of a PY-NLS Variant that Lacks the Proline-Tyrosine Motif. *Structure*. 2016; 24: 1802–1809. <https://doi.org/10.1016/j.str.2016.07.018> PMID: 27618664
55. Tyzack GE, Luisier R, Taha DM, Neeves J, Modic M, Mitchell JS, et al. Widespread FUS mislocalization is a molecular hallmark of amyotrophic lateral sclerosis. *Brain*. 2019; 142: 2572–2580. <https://doi.org/10.1093/brain/awz217> PMID: 31368485
56. Subramanian S. Fixation of deleterious mutations at critical positions in human proteins. *Mol Biol Evol*. 2011; 28: 2687–2693. <https://doi.org/10.1093/molbev/msr097> PMID: 21498603
57. Chiò A, Mora G, Leone M, Mazzini L, Cocito D, Giordana MT, et al. Early symptom progression rate is related to ALS outcome: a prospective population-based study. *Neurology*. 2002; 59: 99–103. <https://doi.org/10.1212/wnl.59.1.99> PMID: 12105314
58. Conte A, Lattante S, Zollino M, Marangi G, Luigetti M, Del Grande A, et al. P525L FUS mutation is consistently associated with a severe form of juvenile amyotrophic lateral sclerosis. *Neuromuscul Disord*. 2012; 22: 73–5. <https://doi.org/10.1016/j.nmd.2011.08.003> PMID: 21907581
59. Liu X, Shi D, Zhou S, Liu H, Liu H, Yao X. Molecular dynamics simulations and novel drug discovery. *Expert Opin Drug Discov*. 2018; 13: 23–37. <https://doi.org/10.1080/17460441.2018.1403419> PMID: 29139324

60. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol.* 2002; 9: 646–52. <https://doi.org/10.1038/nsb0902-646> PMID: 12198485
61. Martínez L. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS One.* 2015; 10: e0119264. <https://doi.org/10.1371/journal.pone.0119264> PMID: 25816325
62. Oliveira CCS De, Pereira GRC, Alcantara JYS De, Antunes D, Caffarena ER, Mesquita JF De. 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
63. Kumar CV, Swetha RG, Anbarasu A, Ramaiah S. Computational analysis reveals the association of threonine 118 methionine mutation in PMP22 resulting in CMT-1A. *Adv Bioinformatics.* 2014; 2014: 10. <https://doi.org/10.1155/2014/502618> PMID: 25400662
64. 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
65. Kumar S, Wolfson HJ, Nussinov R. Protein flexibility and electrostatic interactions. *IBM J Res Dev.* 2001; 45: 499–512. <https://doi.org/10.1147/rd.453.0499>
66. Teilum K, Olsen JG, Kragelund BB. Functional aspects of protein flexibility. *Cell Mol Life Sci.* 2009; 66: 2231–2247. <https://doi.org/10.1007/s00018-009-0014-6> PMID: 19308324
67. 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
68. Pereira GRC, Vieira B de AA, Mesquita JF De. Comprehensive in silico analysis and molecular dynamics of the superoxide dismutase 1 (SOD1) variants related to amyotrophic lateral sclerosis. *PLoS One.* 2021; 1: 1–27. <https://doi.org/10.1371/journal.pone.0247841> PMID: 33630959
69. Montefiori M, Pilotto S, Marabelli C, Moroni E, Ferraro M, Serapian SA, et al. Impact of Mutations on NPAC Structural Dynamics: Mechanistic Insights from MD Simulations. *J Chem Inf Model.* 2019; 59: 3927–3937. <https://doi.org/10.1021/acs.jcim.9b00588> PMID: 31408337
70. Serapian SA, Marchetti F, Triveri A, Morra G, Meli M, Moroni E, et al. The Answer Lies in the Energy: How Simple Atomistic Molecular Dynamics Simulations May Hold the Key to Epitope Prediction on the Fully Glycosylated SARS-CoV-2 Spike Protein. *J Phys Chem Lett.* 2020; 11: 8084–8093. <https://doi.org/10.1021/acs.jpcclett.0c02341> PMID: 32885971
71. Meli M, Morra G, Colombo G. Simple Model of Protein Energetics to Identify Ab Initio Folding Transitions from All-Atom MD Simulations of Proteins. *J Chem Theory Comput.* 2020; 16: 5960–5971. <https://doi.org/10.1021/acs.jctc.0c00524> PMID: 32693598
72. Pontiggia F, Colombo G, Micheletti C, Orland H. Anharmonicity and self-similarity of the free energy landscape of protein G. *Phys Rev Lett.* 2007; 98: 1–4. <https://doi.org/10.1103/PhysRevLett.98.048102> PMID: 17358816