

Molecular Dynamics Simulation and Structural Insights into The SMN1 Protein Involved in the Pathogenesis of Spinal Muscular Atrophy

Pankaj Bagga¹, Sudhakar Singh^{2*}, Gobind Ram³, Subham Kapil⁴, Ramesh Eerapagula⁵, Punit Puri^{6*}, Jitender Kumar⁷, Imran Sheikh⁸

¹School of Bioengineering and Biosciences, Lovely Professional University, Phagwara, Punjab - 144411, India.

Email ID: bagga021@gmail.com

²School of Bioengineering and Biosciences, Lovely Professional University, Phagwara, Punjab - 144411, India.

Email ID: sudhakarsingh86@gmail.com

³PG Department of Biotechnology, Lyallpur Khalsa College, Jalandhar, Punjab - 144001, India.

Email ID: ramgobind4@gmail.com

⁴Division of Zoology & Centre for Green Energy Research, Career Point University, Bhoranj (Tikka-Kharwarian) Hamirpur, Himachal Pradesh (India) - 176041.

Email ID: shubhamkapil143@gmail.com

⁵Department of Genome Informatics, Centre for DNA Fingerprinting and Diagnostics (CFDFD), Hyderabad – 500001, India.

Email ID: ramesheerapagula76@gmail.com

⁶Department of Zoology, DAV College Jalandhar, Mahatma Hans Raj Marg, GT Road, Jalandhar, Punjab – 144008, India.

Email ID: puripunit@gmail.com

⁷Department of Plant Sciences, School of Life Sciences, Central University of Himachal Pradesh, Kangra 176206, India.

Email ID: jitenderkumarcuhp@gmail.com

⁸University Centre for Research and Development (UCRD), Chandigarh University, Gharuan, Mohali-140413, Punjab, India.

Email ID: imransheikh485@gmail.com

*Corresponding Author:

Sudhakar Singh, Punit Puri

Email ID: sudhakarsingh86@gmail.com , puripunit@gmail.com

Cite this paper as: Pankaj Bagga, Sudhakar Singh, Gobind Ram, Subham Kapil, Ramesh Eerapagula, Punit Puri, Jitender Kumar, Imran Sheikh, (2025) Molecular Dynamics Simulation And Structural Insights Into The SMN1 Protein Involved In The Pathogenesis Of Spinal Muscular Atrophy. *Journal of Neonatal Surgery*, 14 (13s), 1066-1077.

ABSTRACT:

Spinal muscular atrophy (SMA) is an autosomal recessive inherited neuromuscular condition distinguished by the deterioration of alpha motor neurons within the spinal cord. This degeneration leads to a gradual onset of muscle weakness and paralysis primarily affecting muscles close to the body's center. The SMA is categorized into four severity grades (SMA I, SMA II, SMA III and SMA IV) determined by the age of onset and the level of motor function attained. This condition arises from homozygous mutations in the survival motor neuron 1 (SMN1) gene, with diagnostic tests typically revealing homozygous deletion of SMN1 exon 7 in the majority of patients. Herein, we have applied bioinformatics approaches to predict the structure of SMN1 protein (using AlphaFold, I-TASSER and RoseTTAFold), structure validation (PSVS v1.5, PROCHECK and ProSA-web) and molecular dynamics (MD) simulation using GROMACS 2022.3 at 100 ns (nanoseconds) to analyze the Root Mean Square Deviation, Root Mean Square Fluctuation, and Radius of Gyration. MD results clearly indicate that RoseTTAFold predicted structure of SMN1 is highly stable and consistent.

Keywords: Spinal Muscular Atrophy; Survival Motor Neuron1; Molecular Dynamics Simulation; Root Mean Square Deviation; Root Mean Square Fluctuation; Radius of Gyration

1. INTRODUCTION

Spinal muscular atrophy (SMA) is a genetic disorder that affects the nerves responsible for muscle movement. Over the time, these nerves deteriorate, leading to the muscle weakness and loss of function, results in significant and gradual decrease in muscle tone and strength (Younger and Mendell (2023). With a prevalence rate of about 1 in 10,000 live births, SMA is considered the second most common autosomal recessive disorder (Nilay *et al.*, 2021). Most cases of SMA are a result of

genetic mutations or deletions in the *SMN1* gene, present on chromosome 5q13.2 (Savad *et al.*, 2023). The *SMN2*, which is nearly identical to *SMN1*, has the greatest impact on disease modification. With just seven nucleotide differences, *SMN2* showcases an alternative splicing of exon 7, leading to the creation of a mostly non-functional protein. As a result, it only encodes a small portion of the complete SMN protein (Ramos *et al.*, 2019; Nishio *et al.*, 2023). There is variation among individuals in the number of copies of *SMN2*. The severity of the phenotype is heavily impacted by the quantity of *SMN2* copies (Babić *et al.*, 2023).

SMA is categorised into five clinical types (0–4), based on the highest motor milestone achieved and the age of onset. The most severe form of SMA is Type 0, which usually begins before or shortly after birth and often leads to death at birth (Nicolau *et al.*, 2021). The most severe and prevalent type of SMA is SMA type 1, also known as Werdnig-Hoffmann disease. It accounts for approximately 50% of all SMA diagnoses. Infants with SMA type I usually show clinical signs before they turn 6 months old and they are unable to sit without support and intervention and sadly, their life expectancy is generally less than 2 years (Jędrzejowska 2020; Aragon-Gawinska *et al.*, 2023). Symptoms of type 2 SMA known as Dubowitz disease usually manifest in individuals around 6 to 18 months of age. Although individuals may develop the capability to sit or stand with assistance but they never attain the ability to walk independently (Oskoui *et al.*, 2023). The SMA type III, also known as Kugelberg-Welander disease, typically manifests its symptoms after the child reaches 18 months of age. Patients are able to walk on their own, but those who experience symptoms before the age of 3 (Type 3a) generally lose their ability to walk earlier than those who develop symptoms after the age of 3 (Type 3b) (Al-Taie and Köseoğlu 2023). Type 4 SMA typically appears in adulthood and is considered the least severe variation of the condition (Farrar and Kiernan 2015; Belter *et al.*, 2023). The start of medication development for SMA patients, which marks a turning point in the treatment of SMA. The first FDA-approved medication for the treatment of SMA was Nusinersen, which was approved in December 2016. This intrathecal antisense oligonucleotide acts as a modulator of *SMN2* pre-mRNA splicing (Angilletta *et al.*, 2023). Next came Risdiplam (August 2020), an oral *SMN2* splicing modifier used to treat spinal muscular atrophy, an intravenous one-time gene transfer therapy that aims to deliver a functional copy of the *SMN1* gene (Ratni *et al.*, 2018; Chiriboga *et al.*, 2023; Mendell *et al.*, 2023).

To understand more about protein function and regulation, the stability of the structure is critical. As there is no complete structure for the *SMN1* protein available on PDB, there is a need for prediction of the stable structure of the protein. In the present study, multiple software programs were utilized to predict the highly stable structure of the *SMN1* protein and the predicted structure was further validated using various tools.

2. MATERIALS AND METHODS

Primary sequence and secondary structure Prediction

The primary sequence of the *SMN1* protein was retrieved from the National Center for Biotechnology Information (NCBI) GenBank (<https://www.ncbi.nlm.nih.gov/>) had Accession number NP_000335.1 and the UniProt Knowledgebase (UniProtKB) database (<https://www.uniprot.org/>) (Lussi *et al.*, 2023). The Cytogenetic location of *SMN1* is validated using Online Mendelian Inheritance in Man (OMIM) (<https://www.omim.org/>) (Cacheiro *et al.*, 2024) and the NCBI Gene (<https://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve>) (Macnee *et al.* 2023). Further, the motifs and domain of *SMN1* were identified using PROSITE (<https://prosite.expasy.org/>) (Subbiah *et al.*, 2023; Derry and Altman (2023)), SMART (<https://smart.embl-heidelberg.de/>) (Kavas *et al.*, 2023) and InterProScan (<https://www.ebi.ac.uk/interpro/search/sequence/>) (Gumerov *et al.*, 2024). In addition, Position-Specific Iterated - BLAST (PSIPRED) was used to assess the proportion of alpha-helices in the targeted *SMN1* protein sequence, the corresponding amino acid residues, and disrupted amino acids in the entire protein sequence (<http://bioinf.cs.ucl.ac.uk/psipred/>) (Pahlavan *et al.*, 2024). The secondary structure elements were predicted by the Self-Optimized Prediction method with Alignment (SOPMA) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Sawal *et al.*, 2023).

Model building and validation

A fundamental bioinformatics method for determining a target protein's three-dimensional structure using the target primary sequence information and the template protein sequence alignment is the structure modeling strategy. The 3-D structure of *SMN1* was predicted using AlphaFold (<https://alphafold.ebi.ac.uk/>) (Fasimoye *et al.*, 2022), I-TASSER (<https://zhanggroup.org/I-TASSER/>) (Sharmin *et al.*, 2023), and RoseTTAFold (<https://rosetta.bakerlab.org/>) (Baek *et al.*, 2024). We also compared the structures modeled by I-TASSER, AlphaFold, and RoseTTAFold in order to determine which the best was. The next stage of identifying a protein's three-dimensional structure is structural validation. The Protein Structure Validation Software (PSVS) suite v1.5 (<https://montelionelab.chem.rpi.edu/PSVS/PSVS/>) received the PDB file containing the predicted structure (Zeindl *et al.*, 2023). To evaluate the stereochemical quality of the protein, the Ramachandran plot of the target protein (*SMN1*) was constructed using PROCHECK (Alam *et al.*, 2023). The Z-score is calculated using the ProSA-web program in order to evaluate the entire protein structure (<https://prosa.services.came.sbg.ac.at/prosa.php>) (Khan *et al.*, 2023).

Molecular dynamics (MD) simulation

The stability and consistency of the predicted structure of SMN1 were assessed using a molecular dynamics (MD) simulation running at 100 ns (100000 ps). The structure predicted by RoseTTAFold was used in (Setlur *et al.*, 2023) study on the GROMACS 2022.3 (Groningen Machine for Chemical Simulations) software (Spoel *et al.*,2005; Praveena *et al.*,2022). Newton's motion equation was solved by the simulation of every atom in order to gain insight. Using the Optimized Potentials for Liquid Simulations (OPLS-AA/M) all-atom force field, GROMACS' pdb2gmx modules were used to construct the protein file and protein topology (Hosseini and Van der Spoel,2023). A triclinic box was employed for protein solvation, and the protein was positioned at least 2.0 nm from the box edge. The steepest descent algorithm was used to minimize energy (Jaidhan *et al.*,2014; Lu *et al.*,2023). The V-rescale temperature coupling method was then used to equilibrate the system in an isothermal-isochoric ensemble (NVT ensemble with a constant particle number, volume, and temperature) at 300 K. Using the isotropic Parrinello-Rahman pressure coupling strategy, the system was once more equilibrated at 1.0 bar of atmospheric pressure under isobaric-isothermal conditions (NPT ensemble with constant particle number, pressure, and temperature). The Leap-Frog integrator, periodic boundary conditions, valet-cut-off for unbound parameters, LINCS constraints for bound parameters, and particle mesh Ewald for long-range electrostatics were used in both equilibration processes. For 100ns, an unrestricted production simulation of the equilibrated system was run. Following the addition of the necessary charges to the system, the simulation was run (Spoel *et al.*,2005; Ashiru *et al.*, 2023). The entire duration of the simulation was 100 nanoseconds. Trajectory analysis and the calculation of several parameters, including Radius of Gyration, root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), temperature, pressure, and potential energy, were then conducted using the software's integrated tool (Romero-Rivera *et al.*,2017).

3. RESULTS

Primary sequence and secondary structure Prediction

The SMN1 protein, which is implicated in the pathophysiology of spinal muscular atrophy, formed the basis for the current study's structural analysis and molecular simulation. The cytogenetic location of the SMN1 protein was 5q13.2, its length was 294 aa, and its accession number was NP_000335.1 (Figure 1).

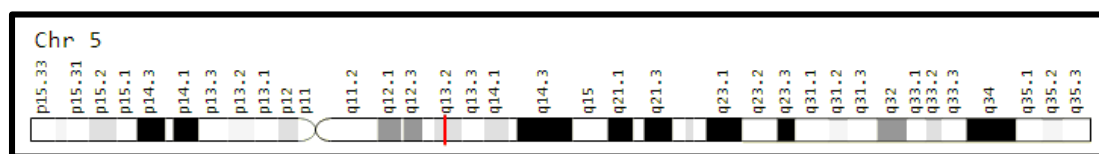


Figure 1: Cytogenetic location of SMN1 protein

Our target sequence's UniProt ID was Q16637. Name and taxonomy, subcellular location, disease and variations, post-translational modifications (PTM)/Processing, expression, interaction, structure, family and domains, sequence and isoforms, and comparable proteins related to or target protein SMN1 were all studied using UniProt (Table 1).

Table1: Basic information about the SMN1 protein

Gene Name	Survival of motor neuron 1, telomeric [<i>Homo sapiens</i> (human)]
NCBI Gene ID	6606
HGNC(HUGO Gene Nomenclature Committee) Approved Symbol	SMN1
HGNC Approved Name	Survival of motor neuron 1, telomeric
HGNC ID	11117
GenBank/Refseq Accession Number	NM_000344.4
Number of Nucleotides	1482 bp
Protein Accession Number	NP_000335.1
Number of Amino Acids	294 aa
UniProt ID	Q16637
Cytogenetic Location	5q13.2
Locus type	Gene with protein product(protein-coding genes (the protein may be

	predicted and of unknown function)
OMIM ID	600354
UCSC ID	uc003kak.4

The PROSITE, SMART, and InterProScan databases were used to identify the TUDOR domain. According to Figures. 2a and 2b, the domain in the query sequence begins at positions 90 to 151 aa.

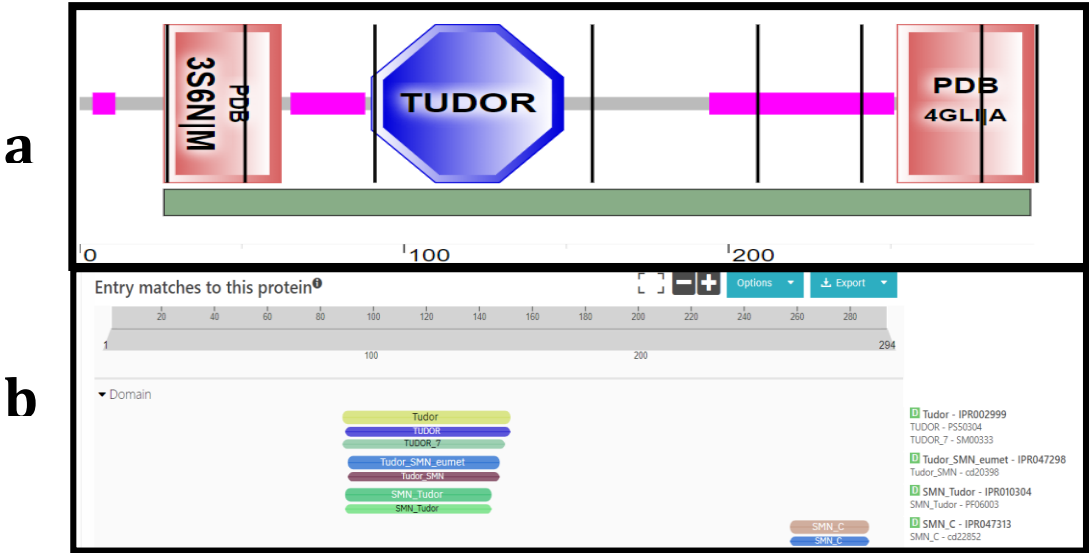


Figure 2: Predicted results of the SMN1 Tudor domain. a. SMART result b. InterProScan result.

Using specific color codes—yellow for strands, pink for helices, and gray for coils—the PSIPRED result depicts the secondary structural elements (Figure 3a). Furthermore, out of 294 amino acids, the SOPMA analysis shows the amounts of several structural components in the SMN1 protein, such as 58- α helices, 39-extended strands, 15- β twists, and 182-random coils (Figure 3b).



Figure 3: Secondary Structure Prediction a. PSIPRED Result. b. SOPMA result.

Model building and validation

AlfaFold (Figures 4a), I-TASSER (Figures 1b), and the RoseTTAFold server (Figures 5a and 5b) were used to predict the

structure of SMN1.

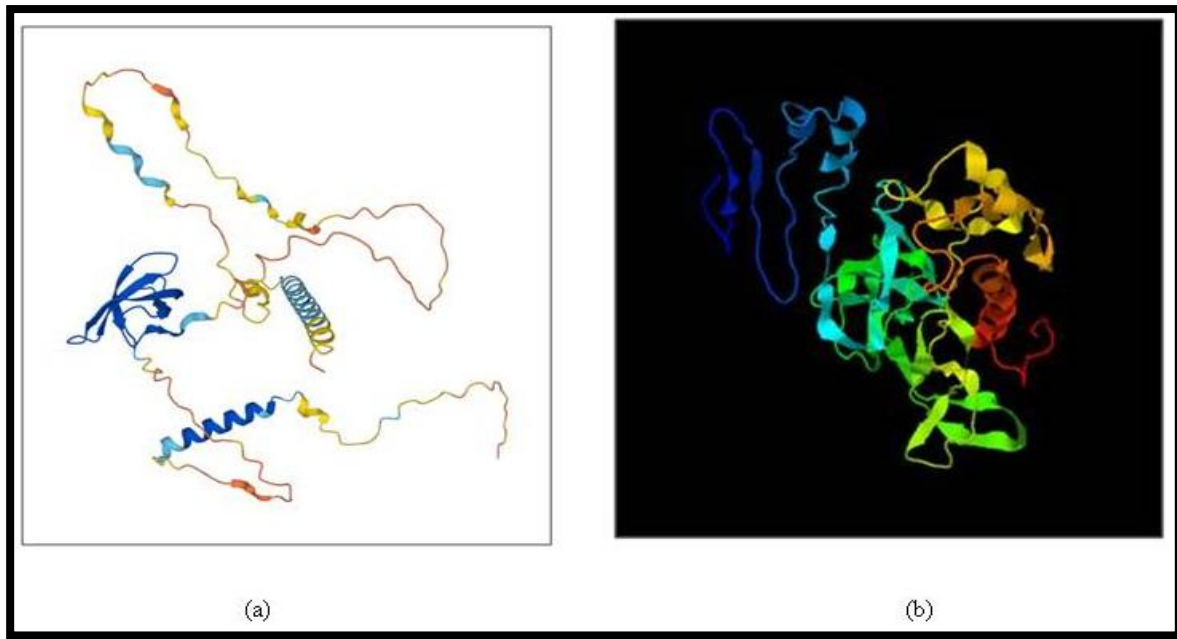


Figure 4: SMN1 structure predicted by a. AlphaFold and b. I-TASSER

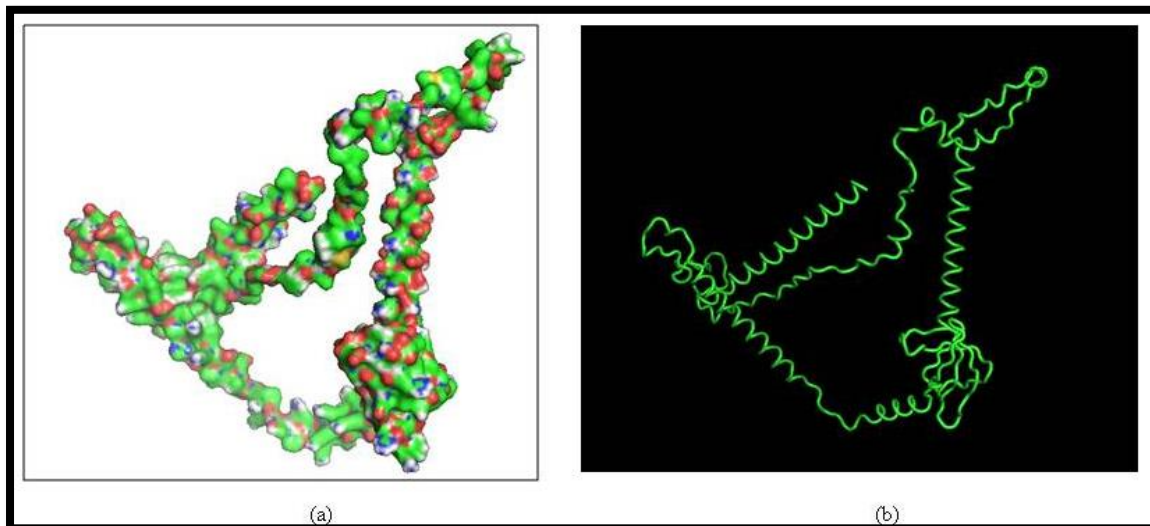


Figure 5. Structure predicted by RoseTTAFold a. surface view of SMN1 b. ribbon structure of SMN1.

Additionally, we examined RoseTTAFold's 100% sequence and structural alignment with AlfaFold and the I-TASSER structure after aligning the structure predicted by these tools. PSVS was utilized to validate the structure of the RoseTTAFold predicted structure (Figure 6).

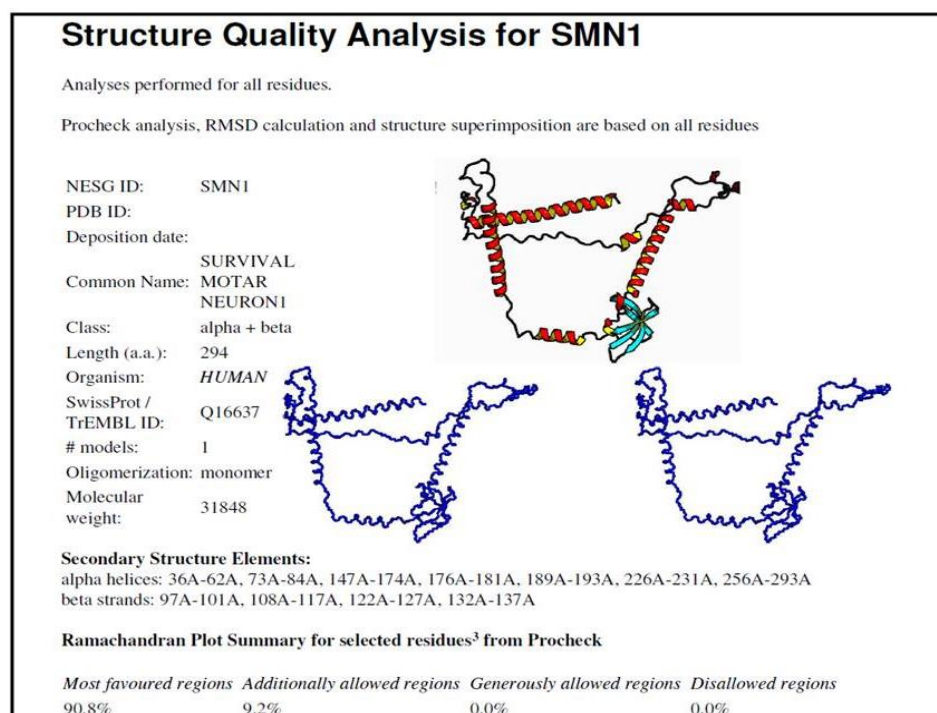


Figure 6: Structure validation using PSVS

The Ramachandran plot was made with PROCHECK, and it showed that there were 207 (90.8%) residues in the most favored regions, 21 (9.2%) in the additional allowed regions, 0 (0%) in the generously allowed regions, and 0 (%) in the disallowed regions. There were also 228 (100%) non-glycine and non-proline residues, 2 end residues (not including Gly and Pro), 25 glycine residues (shown in a triangle), 39 proline residues, and a total of 294 aa residues (Figure 7). RoseTTAFold's model had the best structure overall, according to the evaluation data, with all of its residues situated in the Ramachandran plot's most liked area.

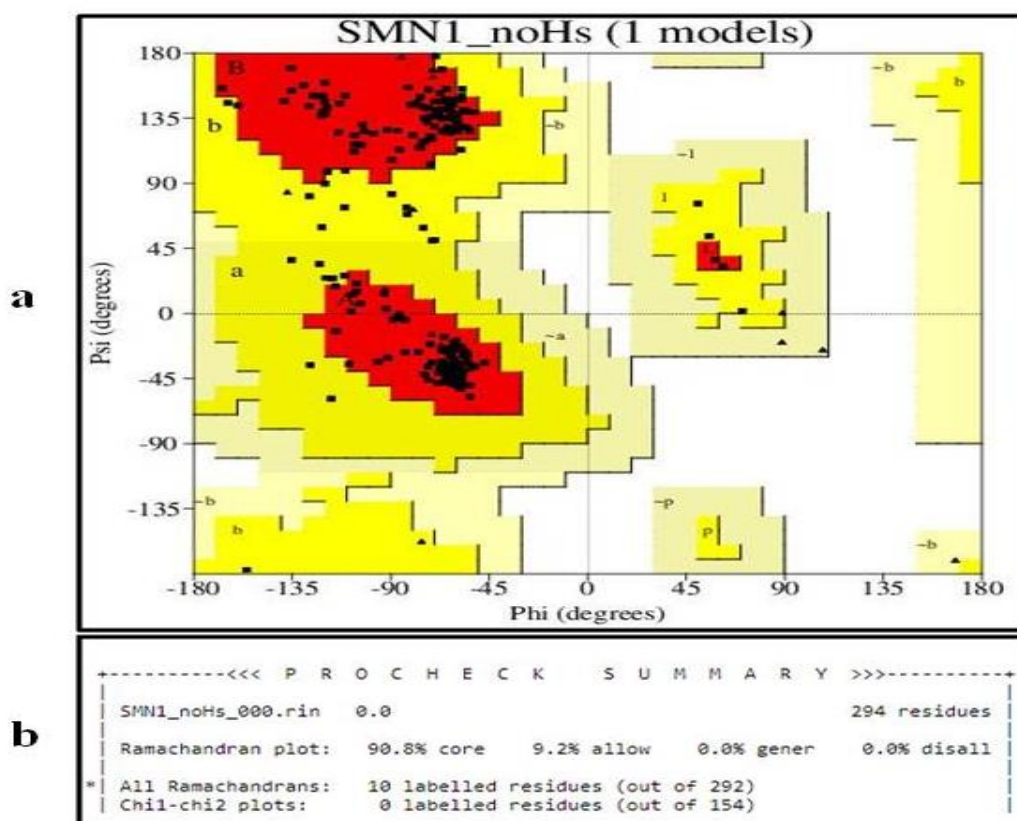


Fig. 7: Structural validation a. Ramachandran Plot b. Text summary of SMN1 protein

The Z-score acquired from ProSA-web servers was used to further validate the structure of the SMN1 protein, confirming its structural quality (a) The SMN1 structure is within the permissible range for X-ray and NMR investigations based on the curated data (the total score is -5.19) (b) Two distinct window sizes were used to depict the sequence position and knowledge-energy graph: light green denotes a window size of 10 and deep green denotes a window size of 40. This plot shows that all of the SMN1 protein's residues have energy values that are noticeably less than zero (Figure 8).

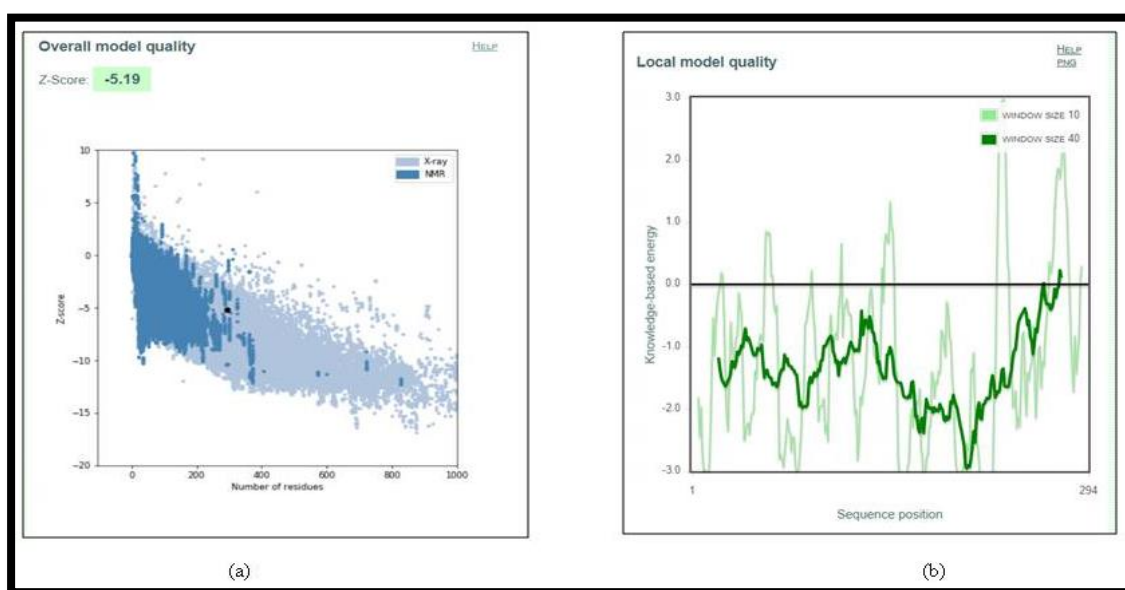


Figure 8: Structure validation by using ProSA-Web a. overall model quality and b. local model quality.

Molecular dynamics (MD) simulation

For predicting the stability of the protein structure, a molecular dynamics (MD) simulation was performed at 100 ns using the GROMACS 2022.3 software. At 50,000 steps, the steepest descent approach was used to minimize energy. For 1000 ps, the NVT and NPT ensembles were both run. Following the 1000 ps equilibration period, the temperature curve (Figure 9a) displayed an average temperature of 310 K. Throughout the 1000 ps equilibration phase, there were noticeable variations in the generated pressure graph. The graph (Figure 9b) shows that the pressure levels nearly stay constant (-50 ± 120 bar) over time during equilibration.

During the simulation, the radius of gyration (Rg) was calculated to examine the protein's level of compactness. The protein folded in its stable state and had a rather constant value for the predicted structure during the 100 ns MD simulation, as evidenced by the total radius of gyration of 30 Å (Figure 9c). A 100-ns MD simulation was then used to determine the protein backbone's root-mean-square deviation (RMSD), which showed minimal variability at 25 ns. The structure's general stability over the 100 ns run is demonstrated by the equilibration that appears at 50 ns and the stable RMSD value about 17 Å (Figure 9d). Additionally, the root-mean-square fluctuation (RMSF) was computed to assess the flexibility of residues. The RMSF plot was constant around 7 Å and exhibits the least amount of variability during the simulation (Figure 9e).

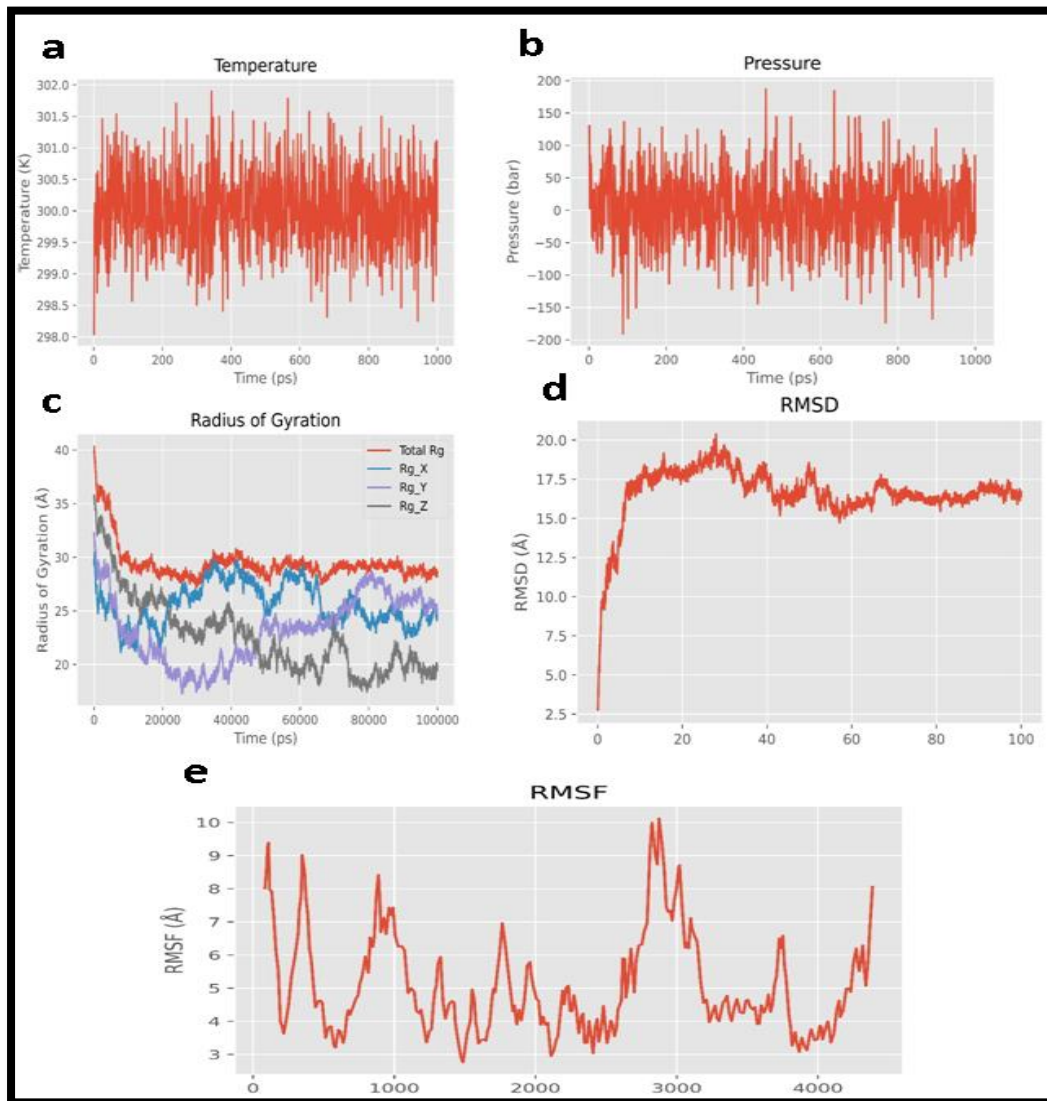


Fig. 9: Molecular dynamic simulation of SMN1 protein using GROMACS at 100 ns. a. Temperature Graph at 310K b. Pressure Graph c. Radius of Gyration d. RMSD e. RMSF

4. DISCUSSION

The target protein of this study is SMN 1 as SMA is mainly caused by mutation or deletion in this protein. The primary data for the SMN1 protein was evaluated in this study utilizing the NCBI and UniProt databases. According to previous research, the SMN1 protein has the accession number AAA66242 and 294 amino acids. The cytogenetic location of SMN1 has been

determined as 5q13.2, which is consistent with previous findings (Savad *et al.*,2023; Lefebvre *et al.*,1995; Li (2017)). In addition, we identified the target protein's domains and motifs using the SMART, InterProScan, and ScanProsite techniques. The Tudor domain has been discovered as a component of the SMN1 protein, which is corroborated by earlier studies (Li (2017); Chong, *et al.*,2021).

SOPMA and PSIPRED were used to collect secondary structural information on the target SMN1 protein. This made it simpler to count the amount of alpha, pi and beta helices, as well as bend areas, extended strands, and other structural elements. Using these approaches, the SMN1 protein exhibits the following structural features. The protein structure consists of 58 α helices, 39 extended strands, 15 β twists, and 182 random coils. After collecting primary and secondary data on the SMN1 protein, this study assessed its structure using the Protein Data Bank (PDB; <https://www.rcsb.org/>), however, the structure was not available. The SMN1 structure remained inaccessible as of April 23, 2024. Consequently, this study aimed to employ RoseTTAFold, AlfaFold and I-TASSER to simulate the structure of SMN1. One model was received from each of the top five model sets offered by AlfaFold, I-TASSER and RoseTTAFold. To determine the number of residues within the acceptable and banned zones, a Ramachandran plot was build for each model. The RoseTTAFold server provided us with the most favorable model because all residues were within the permitted range.

In addition, the Protein Structure Validation Software (PSVS) package v1.5, PROCHECK and ProSA-web server were utilized to validate the structure using the RoseTTAFold projected model. Initially, PSVS gave us the secondary structural components of SMN1: The protein structure is made up of beta strands at positions 97A-101A, 108A-117A, 122A-127A and 132A-137A. Alpha helices are also present at locations 36A-62A, 73A-84A, 147A-174A, 176A-181A, 189A-193A, 226A-231A and 256A-293A. Procheck was then used to construct a Ramachandran Plot for the SMN1 protein, resulting in the following plot statistics: The most popular places were 207 (90.8%), followed by 21 (9.2%), which were allowed. No regions were generously granted or denied. Furthermore, this study evaluated the structure using ProSA-web, which certifies its quality in two ways: (a) The Z-Score for the overall quality model, acquired by plotting the Z-Score on the y-axis against the number of residues (x-axis), was -5.19. This suggested that the structure of the SMN1 protein was found within a suitable region for conducting X-ray and NMR studies. (b) A graph depicting the quality of a local model was generated by graphing the energy derived from knowledge-based information on the y-axis and the location of the sequence on the x-axis. The graph displayed two clearly differentiated window sizes: 40 and 10, indicated by the dark and light green hues, respectively. This graph illustrates that all of the residue energy values of the SMN1 protein were negative.

In addition, RoseTTAFold predicted structure was used to conduct a 100 nanosecond molecular dynamics simulation using GROMACS 2022.3 software, aiming to forecast the stability of the structure. The energy minimization technique was executed over 50,000 iterations with the steepest descent method. Two NPT ensembles were executed for a duration of 1000 picoseconds. One ensemble maintained constant particle number, volume, and temperature, while the other maintained constant particle number, pressure, and temperature. According to the temperature graph, the system achieved an average temperature of 300 K after the 1000 ps equilibration phase. The pressure graph followed after the required temperature was attained. Throughout the 1000 picosecond equilibration phase, significant fluctuations were seen in the produced pressure graph. The graph depicts the consistent and relatively constant pressure measurements (-50 to 120 bar) throughout the equilibration process over time. Following the pressure graph, the subsequent step was determining the radius of gyration (Rg).

The Rg structural feature of the MD trajectory, which represents the events seen in the MD simulation, is important. The concept of protein axis refers to the arrangement of atoms in a protein. The length is a quantitative measurement of the distance between the pivot point and the point where the transfer of energy has the most significant effect, resulting in the emergence of Rg. The protein maintained a stable conformation throughout the 100 ns molecular dynamics (MD) simulation, as shown by the constant value of the system's total radius of gyration, which was measured to be 30Å.

To assess the conformational stability of the structure, we computed the root-mean-square deviation (RMSD) of the protein backbone using a 100-ns molecular dynamics (MD) simulation of the modeled protein. The system is exhibiting little variation at a time interval of 25 ns. The equilibration process takes around 50 nanoseconds, during which the root mean square deviation (RMSD) value stays consistently around 17 angstroms. This indicates that the structure maintains a stable state during the 100-nanosecond period. The simulation revealed that the expected protein did not undergo any structural alterations.

Furthermore, the root-mean-square fluctuation (RMSF) analysis was conducted to evaluate the flexibility of each residue. High RMSF values show significant deviation from the average position, indicating substantial structural mobility. Conversely, low RMSF values suggest less departure from the average position, indicating more rigidity during the simulation. The RMSF curve exhibited little volatility throughout the simulation and maintained a consistent value of around 7Å. It is worth noting that this study was the first study to predict the stable structure of SMN1 protein.

5. CONCLUSION AND RECOMMENDATIONS

SMN1 potentially plays a crucial role in the pathogenesis of spinal muscular atrophy. The complete structure of the SMN1

was predicted by a comparative study of AlphaFold, I-TASSER and RoseTTAFold prediction. Further structural analysis and validation reveal the RoseTTAFold Server model is of high quality. GROMACS 100 ns simulation revealed the overall stability of the structure based on the RMSD, RMSF and Radius of Gyration.

Funding: This research received no external funding.

Ethical Statement: This study does not include any such human/animal studies.

Competing Interests: The authors declare no conflicts of interest

Author Contributions: Conceptualization, S.S; Methodology, S.S, G.R, R.E, and P.B; writing—original draft preparation, P.B; review and editing, S.K, P.P, J.K and I.S.

Acknowledgments: The authors are thankful to Lovely Professional University (LPU), Jalandhar, Punjab (India), Centre for DNA Fingerprinting and Diagnostics (CFDfD), Hyderabad for providing facilities and encouragement for the research and DAV College Jalandhar, Jalandhar, Punjab (India).

References

- [1] Alam, M. M., Saikat, A. S. M., & Uddin, M. E. (2023). Bioinformatics Approaches for Structural and Functional Annotation of an Uncharacterized Protein of *Helicobacter pylori*. *Engineering Proceedings*, 37(1), 61. <https://doi.org/10.3390/ECP2023-14671>
- [2] Al-Taie, A., & Köseoğlu, A. (2023). Evaluation of the therapeutic efficacy and tolerability of current drug treatments on the clinical outcomes of paediatric spinal muscular atrophy type 1: A systematic review. *Paediatric Respiratory Reviews*. <https://doi.org/10.1016/j.prrv.2023.06.004>
- [3] Angilletta, I., Ferrante, R., Giansante, R., Lombardi, L., Babore, A., Dell'Elice, A., ... & Rossi, C. (2023). Spinal muscular atrophy: an evolving scenario through new perspectives in diagnosis and advances in therapies. *International Journal of Molecular Sciences*, 24(19), 14873. <https://doi.org/10.3390/ijms241914873>
- [4] Aragon-Gawinska, K., Mouraux, C., Dangouloff, T., & Servais, L. (2023). Spinal muscular atrophy treatment in patients identified by newborn screening—a systematic review. *Genes*, 14(7), 1377. <https://doi.org/10.3390/genes14071377>
- [5] Ashiru, M. A., Ogunyemi, S. O., Temionu, O. R., Ajibare, A. C., Cicero-Mfon, N. C., Ihekuna, O. A., ... & Adeyemi, A. O. (2023). Identification of EGFR inhibitors as potential agents for cancer therapy: pharmacophore-based modeling, molecular docking, and molecular dynamics investigations. *Journal of Molecular Modeling*, 29(5), 128. <https://doi.org/10.1007/s00894-023-05531-6>
- [6] Babić, M., Banović, M., Berečić, I., Banić, T., Babić Leko, M., Ulamec, M., ... & Šimić, G. (2023). Molecular biomarkers for the diagnosis, prognosis, and pharmacodynamics of spinal muscular atrophy. *Journal of clinical medicine*, 12(15), 5060. <https://doi.org/10.3390/jcm12155060>
- [7] Baek, M., McHugh, R., Anishchenko, I., Jiang, H., Baker, D., & DiMaio, F. (2024). Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature methods*, 21(1), 117–121. <https://doi.org/10.1038/s41592-023-02086-5>
- [8] Belter, L., Peterson, I., & Jarecki, J. (2023). Evaluating perceived fatigue within an adult spinal muscular atrophy population. *Neurology and Therapy*, 12(6), 2161–2175. <https://doi.org/10.1007/s40120-023-00552-y>
- [9] Cacheiro, P., Lawson, S., Van den Veyver, I. B., Marengo, G., Zocche, D., Murray, S. A., ... & Smedley, D. (2024). Lethal phenotypes in Mendelian disorders. *Genetics in Medicine*, 26(7), 101141. <https://doi.org/10.1016/j.gim.2024.101141>
- [10] Chiriboga, C. A., Bruno, C., Duong, T., Fischer, D., Mercuri, E., Kirschner, J., ... & Muntoni, F. (2023). Risdiplam in patients previously treated with other therapies for spinal muscular atrophy: an interim analysis from the JEWELFISH study. *Neurology and therapy*, 12(2), 543–557. <https://doi.org/10.1007/s40120-023-00444-1>
- [11] Chong, L. C., Gandhi, G., Lee, J. M., Yeo, W. W. Y., & Choi, S. B. (2021). Drug discovery of spinal muscular atrophy (SMA) from the computational perspective: a comprehensive review. *International Journal of Molecular Sciences*, 22(16), 8962. <https://doi.org/10.3390/ijms22168962>
- [12] Derry, A., & Altman, R. B. (2023). COLLAPSE: A representation learning framework for identification and characterization of protein structural sites. *Protein Science*, 32(2), e4541. <https://doi.org/10.1002/pro.4541>
- [13] Farrar, M. A., & Kiernan, M. C. (2015). The genetics of spinal muscular atrophy: progress and challenges. *Neurotherapeutics*, 12(2), 290–302. <https://doi.org/10.1007/s13311-014-0314-x>
- [14] Fasimoye, R. Y., Spencer, R. E. B., Soto-Martin, E., Eijlers, P., Elmassoudi, H., Brivio, S., ... & Müller, B.

- (2022). A novel, essential trans-splicing protein connects the nematode SL1 snRNP to the CBC-ARS2 complex. *Nucleic Acids Research*, 50(13), 7591-7607. <https://doi.org/10.1093/nar/gkac534>
- [15] Gumerov, V. M., Ulrich, L. E., & Zhulin, I. B. (2024). MiST 4.0: a new release of the microbial signal transduction database, now with a metagenomic component. *Nucleic Acids Research*, 52(D1), D647-D653. <https://doi.org/10.1093/nar/gkad847>
- [16] Hosseini, A. N., & Van der Spoel, D. (2023). Simulations of Amyloid-Forming Peptides in the Crystal State. *The Protein Journal*, 42(3), 192-204. <https://doi.org/10.1007/s10930-023-10119-3>
- [17] Jaidhan, B. J., Rao, P. S., & Apparao, A. (2014). Energy minimization and conformation analysis of molecules using steepest descent method. *Int. J. Comput. Sci. Inf. Technol*, 5(3), 3525-3528.
- [18] Jędrzejowska, M. (2020). Advances in newborn screening and presymptomatic diagnosis of spinal muscular atrophy. *Degenerative Neurological and Neuromuscular Disease*, 39-47.
- [19] Kavas, M., Mostafa, K., Seçgin, Z., Yerlikaya, B. A., Yıldırım, K., & Gökdemir, G. (2023). Genome-wide analysis of duf221 domain-containing gene family in common bean and identification of its role on abiotic and phytohormone stress response. *Genetic Resources and Crop Evolution*, 70(1), 169-188. <https://doi.org/10.1007/s10722-022-01421-7>
- [20] Khan, A., Adil, S., Qudisia, H. A., Waheed, Y., Alshabirmi, F. M., & Wei, D. Q. (2023). Structure-based design of promising natural products to inhibit thymidylate kinase from Monkeypox virus and validation using free energy calculations. *Computers in Biology and Medicine*, 158, 106797. <https://doi.org/10.1016/j.combiomed.2023.106797>
- [21] Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., ... & Melki, J. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, 80(1), 155-165.
- [22] Li, W. (2017). How do SMA-linked mutations of SMN1 lead to structural/functional deficiency of the SMA protein?. *PLoS One*, 12(6), e0178519. <https://doi.org/10.1371/journal.pone.0178519>
- [23] Lu, G., Ou, K., Zhang, Y., Zhang, H., Feng, S., Yang, Z., ... & Chen, Z. (2023). Structural analysis, multi-conformation virtual screening and molecular simulation to identify potential inhibitors targeting pS273R proteases of African swine fever virus. *Molecules*, 28(2), 570. <https://doi.org/10.3390/molecules28020570>
- [24] Lussi, Y. C., Magrane, M., Martin, M. J., Orchard, S., & UniProt Consortium. (2023). Searching and navigating UniProt databases. *Current Protocols*, 3(3), e700. <https://doi.org/10.1002/cpz1.700>
- [25] Macnee, M., Pérez-Palma, E., López-Rivera, J. A., Ivaniuk, A., May, P., Møller, R. S., & Lal, D. (2023). Data-driven historical characterization of epilepsy-associated genes. *European Journal of Paediatric Neurology*, 42, 82-87. <https://doi.org/10.1016/j.ejpn.2022.12.005>
- [26] Mendell, J., Wigderson, M., Alecu, I., Yang, L., Mehl, L., & Connolly, A. (2023). P223 Long-term follow-up of onasemnogene abeparvovec gene therapy in patients with spinal muscular atrophy (SMA) type 1. *Neuromuscular Disorders*, 33, S91. <https://doi.org/10.1016/j.nmd.2023.07.105>
- [27] Nicolau, S., Waldrop, M. A., Connolly, A. M., & Mendell, J. R. (2021, April). Spinal muscular atrophy. In *Seminars in pediatric neurology* (Vol. 37, p. 100878). WB Saunders. <https://doi.org/10.1016/j.spen.2021.100878>
- [28] Nilay, M., Moirangthem, A., Saxena, D., Mandal, K., & Phadke, S. R. (2021). Carrier frequency of SMN1-related spinal muscular atrophy in north Indian population: The need for population based screening program. *American Journal of Medical Genetics Part A*, 185(1), 274-277. <https://doi.org/10.1002/ajmg.a.61918>
- [29] Nishio, H., Niba, E. T. E., Saito, T., Okamoto, K., Takeshima, Y., & Awano, H. (2023). Spinal muscular atrophy: the past, present, and future of diagnosis and treatment. *International Journal of Molecular Sciences*, 24(15), 11939. <https://doi.org/10.3390/ijms241511939>
- [30] Oskoui, M., Day, J. W., Deconinck, N., Mazzone, E. S., Nascimento, A., Saito, K., ... & SUNFISH Working Group. (2023). Two-year efficacy and safety of risdiplam in patients with type 2 or non-ambulant type 3 spinal muscular atrophy (SMA). *Journal of Neurology*, 270(5), 2531-2546. <https://doi.org/10.1007/s00415-023-11658-6>
- [31] Pahlavan, Y., Yeganeh, O., Asghariazar, V., & Karami, C. (2024). Multi-epitope vaccine against SARS-CoV-2 targeting the spike RBD: an immunoinformatics approach. *Future Science OA*, 10(1), FSO939. <https://doi.org/10.2144/fsoa-2023-0081>
- [32] Praveena, R., Balasankar, A., Aruchamy, K., Oh, T., Poliseti, V., Ramasundaram, S., & Anbazhakan, K. (2022). Structural Activity and HAD Inhibition Efficiency of Pelargonidin and Its Glucoside—A Theoretical

Approach. *Molecules*, 27(22), 8016. <https://doi.org/10.3390/molecules27228016>

- [33] Ramos, D. M., d'Ydewalle, C., Gabbeta, V., Dakka, A., Klein, S. K., Norris, D. A., ... & Sumner, C. J. (2019). Age-dependent SMN expression in disease-relevant tissue and implications for SMA treatment. *The Journal of clinical investigation*, 129(11), 4817-4831. [10.1172/JCI124120](https://doi.org/10.1172/JCI124120)
- [34] Ratni, H., Ebeling, M., Baird, J., Bendels, S., Bylund, J., Chen, K. S., ... & Mueller, L. (2018). Discovery of risdiplam, a selective survival of motor neuron-2 (SMN2) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). <https://doi.org/10.1021/acs.jmedchem.8b00741>
- [35] Romero-Rivera, A., Garcia-Borras, M., & Osuna, S. (2017). Role of conformational dynamics in the evolution of retro-aldolase activity. *ACS catalysis*, 7(12), 8524-8532. <https://doi.org/10.1021/acscatal.7b02954>
- [36] Savad, S., Ashrafi, M. R., Samadaian, N., Heidari, M., Modarressi, M. H., Zamani, G., ... & Ghafouri-Fard, S. (2023). A comprehensive overview of SMN and NAIP copy numbers in Iranian SMA patients. *Scientific Reports*, 13(1), 3202. <https://doi.org/10.1038/s41598-023-30449-7>
- [37] Sawal, H. A., Nighat, S., Safdar, T., & Anees, L. (2023). Comparative in silico analysis and functional characterization of TANK-binding kinase 1-binding protein 1. *Bioinformatics and Biology Insights*, 17, 11779322231164828. <https://doi.org/10.1177/1177932223116482>
- [38] Sharmin, N., Yee, K. K., Dhaliwal, J. K., & Chow, A. K. (2023). Thinking outside the box: Using homology models and interactive PowerPoints for active learning. *Journal of Dental Education*, 88(Suppl 3), 1902. [10.1002/jdd.13421](https://doi.org/10.1002/jdd.13421)
- [39] Spoel, D. V. D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: fast, flexible, and free. *J Comput Chem*, 26(16), 1701-1718. <https://doi.org/10.1002/jcc.20291>
- [40] Subbiah, T., Maheswaran, G., & Raja, V. K. (2023). In Silico Protein Analysis of Phoenix dactylifera (Date Palm) by a Bioinformatics Approach. *ACS Agricultural Science & Technology*, 3(6), 528-534.
- [41] Younger, D. S., & Mendell, J. R. (2023). Childhood spinal muscular atrophy. *Handbook of clinical neurology*, 196, 43-58. <https://doi.org/10.1016/B978-0-323-98817-9.00030-2>
- [42] Zeindl, R., Unterhauser, J., Röck, M., Eidelpes, R., Führer, S., & Tollinger, M. (2023). Structural characterization of food allergens by nuclear magnetic resonance spectroscopy. In *Food Allergens: Methods and Protocols* (pp. 159-173). New York, NY: Springer US. https://doi.org/10.1007/978-1-0716-3453-0_10

..
