Verifying Whether Mutations M76V and A80Y of 3SRP Influence Antibody Binding Regions

Jin Jun Li

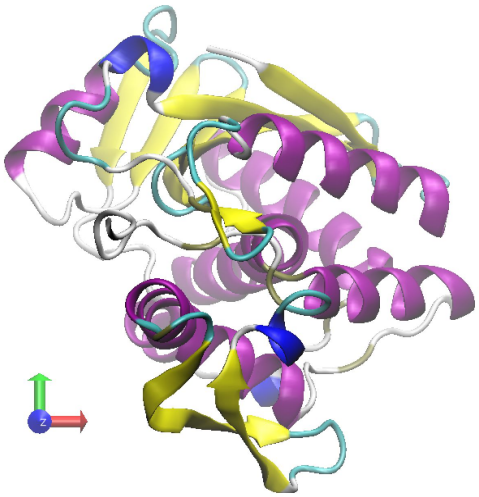
Introduction

Molecular dynamics (MD) simulations have allowed scientists to model and understand interactions between molecules in ways that physical experiments can't. By using computers to replicate Newtonian forces and force fields in a system of molecules, scientists are able to analyze the structure and functions of macromolecules such as proteins. For example, molecular dynamics has allowed scientists to predict secondary structures of proteins, understand how and why proteins fold, visualize where hydrogen bonds and disulfide bridges are formed, analyze how mutations affect proteins, and search for viable medical drugs to cure diseases.[3]

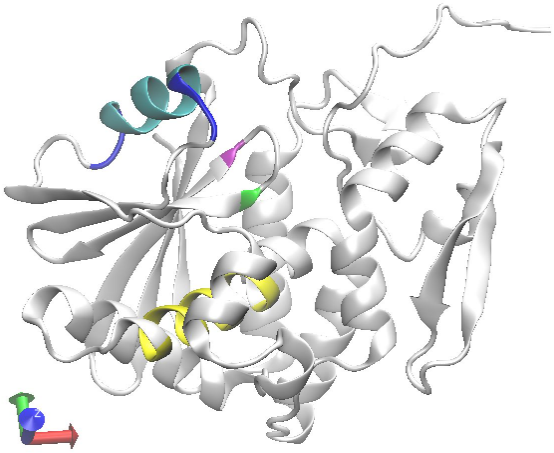
In this paper, we use MD simulations to help further understand the protein RiVax, (PDB entry 3SRP), which has the potential to be a vaccine for ricin, one of the deadliest poisons. Scientists hope to inject RiVax into patients, eliciting the necessary antibodies from the immune system to combat actual ricin. RiVax is composed of the ricin A-chain (RTA) residues 1–268 but with two mutations V76M (to ensure the removal of any trace of vascular leak syndrome (VLS) activity from the immunogen) and Y80A (to inactivate enzyme catalysis). Both mutations help reduce toxicity of RTA. After the mutation, the structure of the protein was experimentally verified using X-ray diffraction with a resolution of 2.14 Angstrom.[7]

This protein has multiple areas of interest that we wish to consider during the MD simulation. The loop–helix–loop segment anchored by two β-strands between residues ASN97 and PHE108 is a binding region for UNIVAX 70/138, a neutralizing antibody produced against RTA. In addition, the conformation of the structural region GLU99–HSD106 is an important epitope for more toxin-neutralizing antibodies.[5] Furthermore, region THR163–MET174 is known to be a binding site for antibody GD-12.

In this paper, we will compare 3SRP with its reversed mutation M76V and A80Y to confirm the author's claim that these mutations do not alter overall protein fold or change the conformation of side chain residues corresponding to known neutralizing epitopes. We also want to see whether the properties of regions around residue 76 and 80 changed to verify that its toxicity changes. It is crucial to make sure that the overall properties of the protein doesn't change, otherwise elicited antibodies will not be able to bind and respond appropriately to this protein, invalidating the usefulness of RiVax as a vaccine. Furthermore, it is important to check whether those mutations actually made a difference, elsewise 3SRP would not be any less toxic, making the protein a poison rather than a vaccine. In this paper, we will call 3SRP the original (or control) protein, and the resulting protein after the reversed mutation the mutated protein.



3SRP Colored By Secondary Structure



3SRP Colored By Special Regions.
Purple: MET76
Green: ALA80
Blue: ASN97-PHE108
Cyan: GLU99–HSD106
Yellow: THR163–MET174

Methods

To compare the control protein and the mutated protein, we used VMD to create an MD trajectory for each protein. Our first step was to prepare the protein and the MD system. We first checked the PDB file for 3SRP for any missing residues. We found that the first 6 residues are missing, which would not influence our simulation because the missing residues are at the beginning of the protein. Next, we checked if the PDB file had any duplicate subunits, and we found that it did not. Afterward, we removed all molecules that do not pertain to the protein, such

as extra water molecules and ligands. Then we solvated the system with water molecules (with 10 Angstrom padding around the protein) to model a real cell as closely as possible. We also rotated the molecule to minimize volume. Furthermore, we ionized the system by first neutralizing the entire charge of the system and then adding 0.1mol/L for NaCl concentration. Ionization allows us to further mimic the cell, as proteins are usually surrounded by ions. Next, we centered the system for a more symmetrical analysis. Finally, we restrained the protein with simple harmonic springs to make sure that the backbone does not move for our next step.

After preparing the protein, we then minimize the energy of the system. We do this by artificially moving the atoms in the system to minimize the potential energy created by the force fields in our simulation. Minimization allows us to start with a stable state that would prevent thermodynamically unlikely and energetically unfavorable scenarios from occurring. We chose the cutoff distance for forces to be 12A.

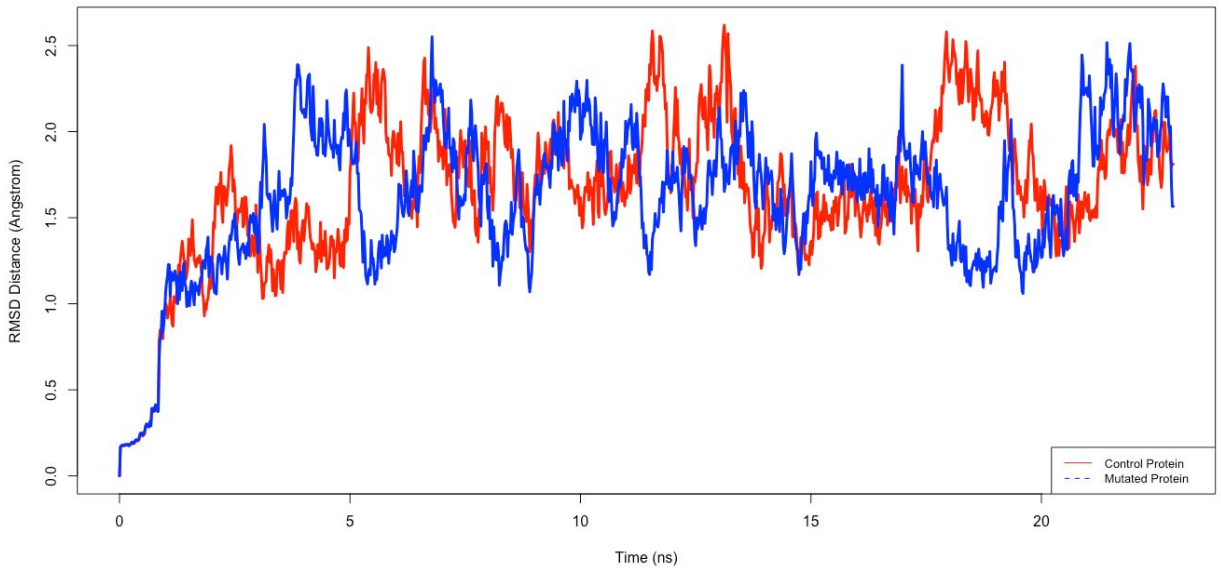
Next, we heat up and equilibrate the protein for 2.85 nanoseconds to model a more realistic protein at room temperature. We warm up the system gradually from 240K to 300K and apply atmospheric pressure to the system. Then we gradually remove the restraints on the protein backbone, as we do not want to immediately release the restrains to reduce the possibility of unrealistic scenarios. Finally, we continue the simulation for another 20 ns with 300K as the temperature. During the heat up, we turned on Langevin dynamics with a damping coefficient of 5/ps, target pressure of 1.01325 bar, oscillation period of 100fs, a damping time scale of 50fs. During the rest of the simulation, the Langevin parameters were the same, except the damping coefficient was changed to 1/ps.

Having completed the simulation for the original protein, we next use VMD to make two point mutations M76V and A80Y. We repeat the entire procedure again for this mutation.

Results and Discussion

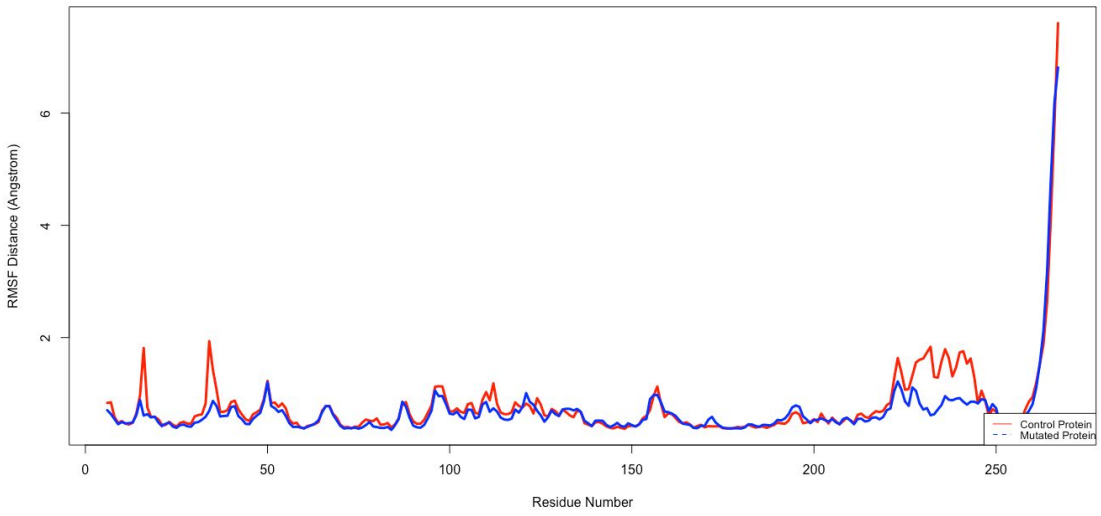
After setting up and running the simulations, we analyze the properties of both the original 3SRP and mutated 3SRP. We first compare the general dynamic behavior of both proteins, then we will focus on the mutated regions and binding sites of the proteins.

We first found the RMSD for each protein, which measures the root mean square distance of all atoms in the XYZ planes of the system from their respective starting locations. We will only measure RMSD for the backbone atoms because the backbone atoms determine the structure that we want to analyze. We aligned the trajectories to the last frame of the minimization to make sure that RMSD starts from 0. Although the graphs do not align exactly, the RMSD plots for both proteins are rather similar since both increases in the beginning during the heat up, then oscillates around 1.75 Angstrom. This aligns with what occurs in the simulation, as both proteins move increasingly faster during the heat up, then the change of movements becomes constant. This also confirms that the movements of both proteins are rather similar.

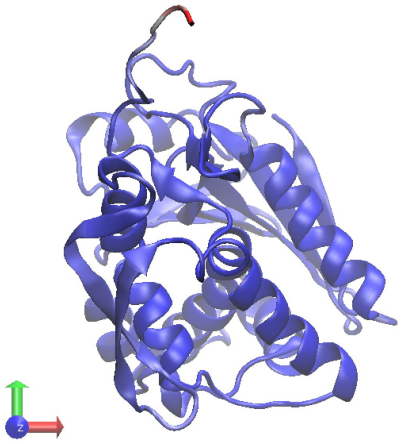


RMSD of 3SRP and Mutated 3SRP

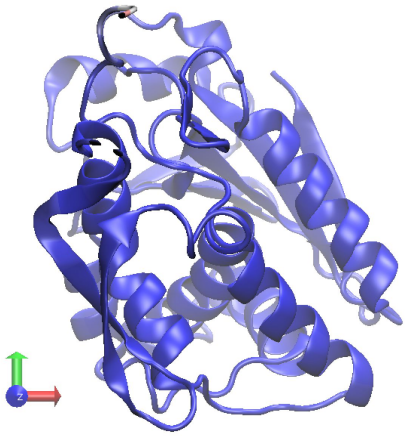
Next, we consider the RMSF of the stable regions of the trajectory (after 5 nanoseconds of runtime, as seen from the RMSD plot). RMSF calculates the average deviation (in Angstrom) for each amino acid's alpha-carbon in the protein against the amino acid number. In this plot, we see that the two proteins behave rather similarly since for most regions, the RMSF is the same. However, it seems that the mutations have decreased the flexibility of the region from residue 13 to 20, 30 to 35, and 225 to 250, as at those regions, the RMSF is lower in the mutated 3SRP. This statement is confirmed in the simulation, as those regions fluctutates less in the mutated protein, The tail of both proteins has a very high RMSF, which corresponds to the rapid movement of the tail in the simulation. Other regions with high flexibility in both proteins (as verified in the trajectory) include residues 45-55, 95-100, and 160-170.



RMSF of 3SRP and Mutated 3SRP

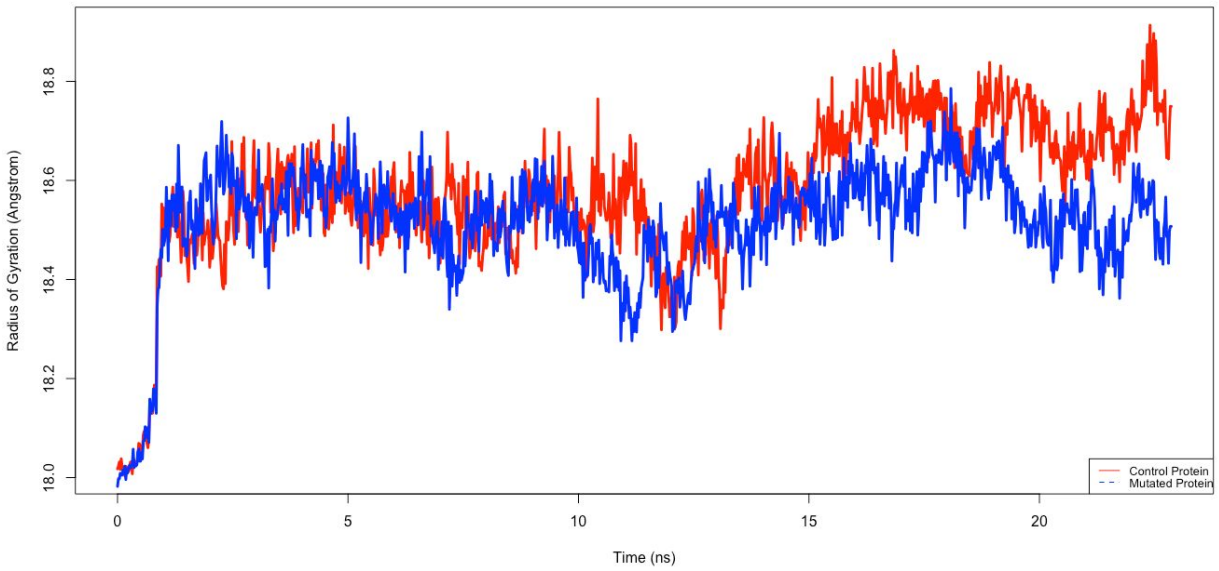


Regions of High RMSF for 3SRP



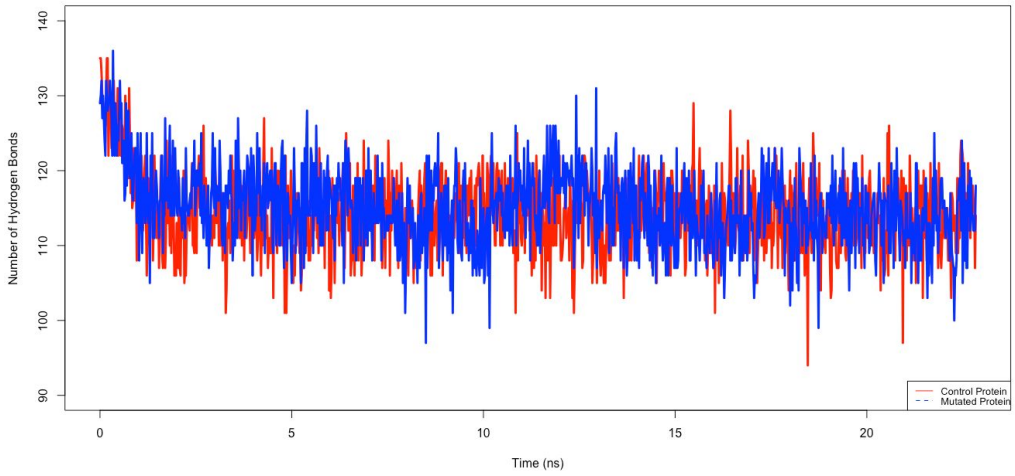
Regions of High RMSF for 3SRP Mutated

Next, we analyze the radius of gyration (RG), which measures the average squared distance of all atoms of the protein from the center of mass of the protein. In both simulations, the RG jumped at the beginning of the simulation, indicating that both proteins unraveled during the beginning of the heat up. After that initial jump, both RG values of the proteins stayed rather constant while oscillating, showing that the compactness of the protein is not changing drastically. However, the control protein 3SRP ended up with a higher RG value towards the end of the simulation, indicating that in the simulation, the initial 3SRP unraveled more than the mutated 3SRP.



RG of 3SRP and Mutated 3SRP

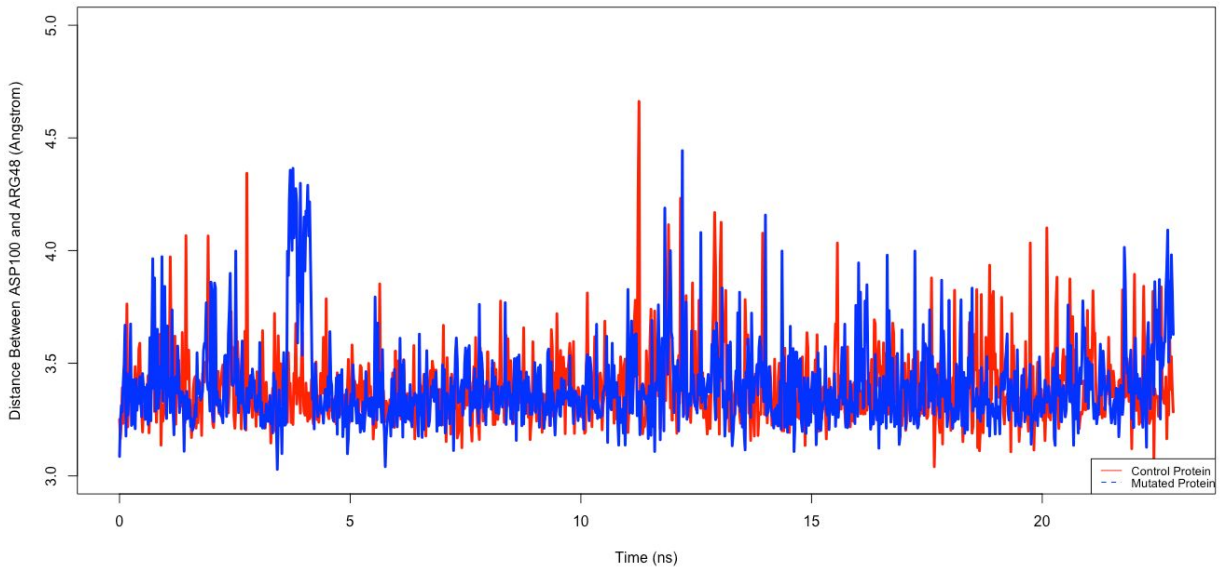
Next, we calculate the number of hydrogen bonds that exist in the protein backbones. During calculation, we used a donor-acceptance distance of 3.5 Angstrom and an angle cut off of 35 degrees. In both simulations, the number of hydrogen bonds decrease during the heat up process but stays rather constant afterward, as the hydrogen bonds oscillate around 115 over the simulation. This is another confirmation that the mutations did not change the overall property of the protein.



Number of Hydrogen Bonds of 3SRP and Mutated 3SRP

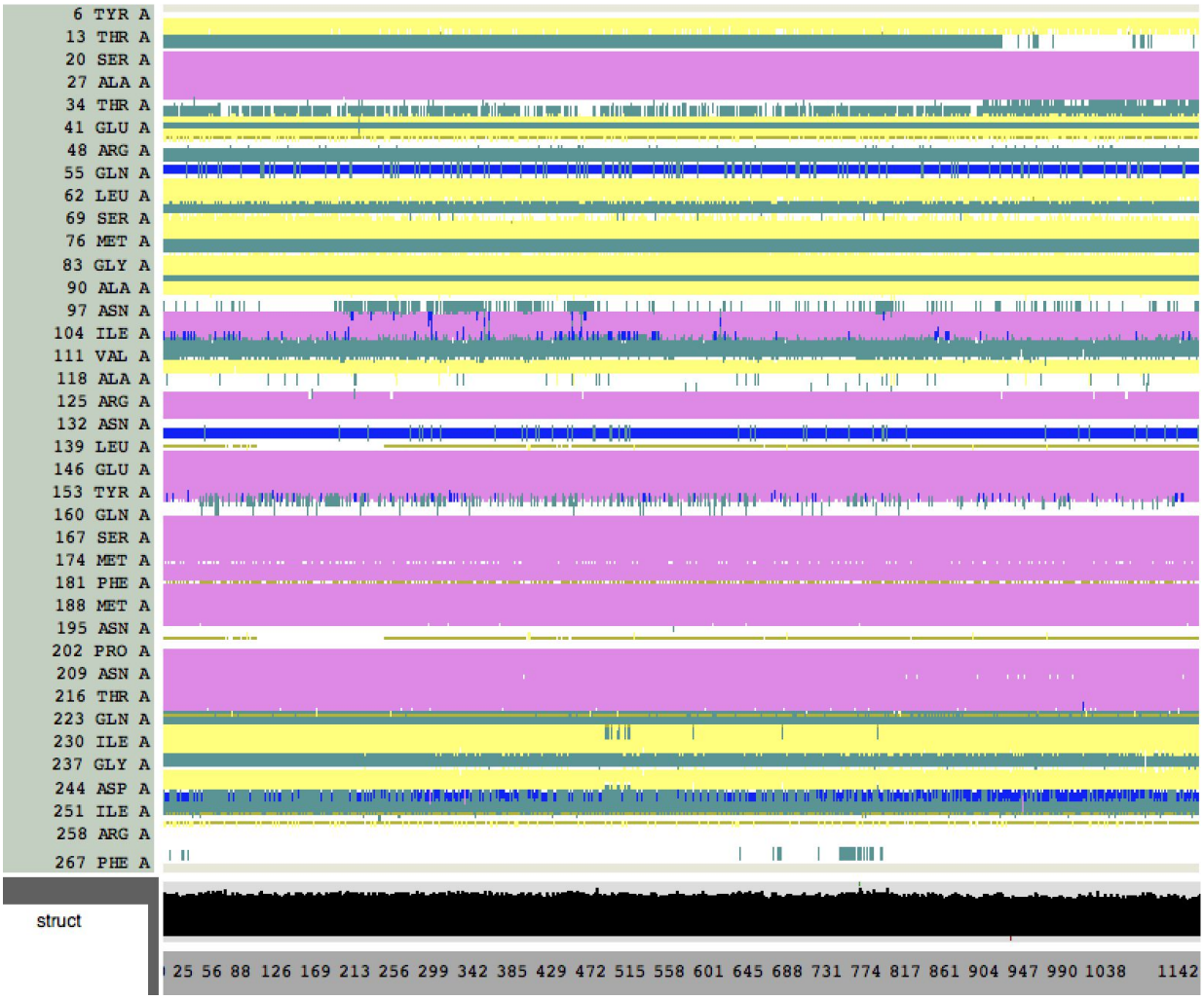
Next we consider salt bridges with O-N bond distance cut off of 3.2A. One stable salt bridge of interest is ASP100-ARG48, as ASP100 is in one of the binding regions that we are interested in observing. The salt bridge for ASP100-ARG48 stay stable for both proteins during

the simulation (as for the majority of the time, the distance stays below 3.2 Angstrom). And in both proteins, only occasionally does the salt bridge break and reform again. Since the salt bridge behaves very similarly in both proteins, this is another confirmation that the mutation did not affect the important binding areas.

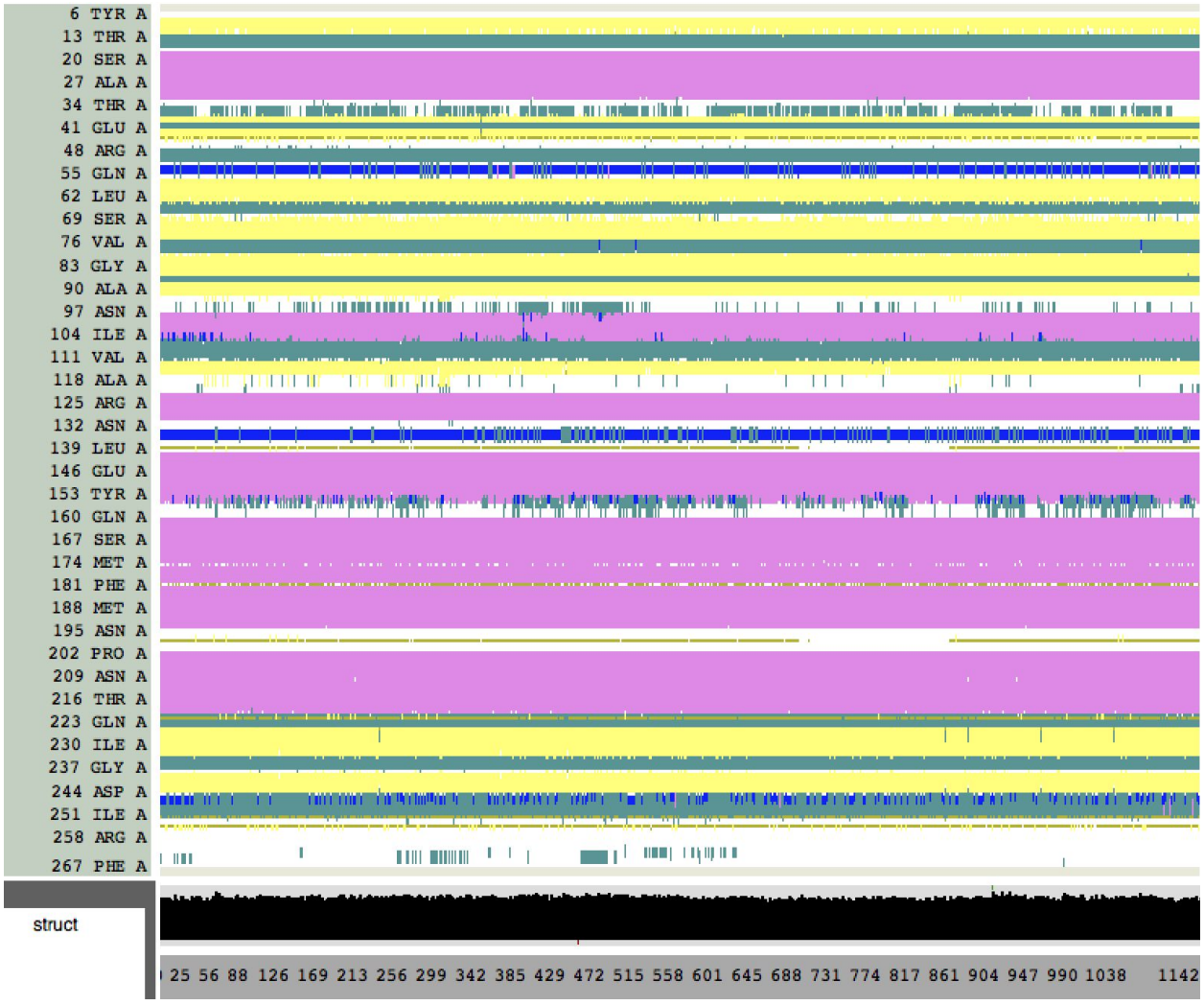


Salt Bridge Distance of ASP100-ARG48 of 3SRP and Mutated 3SRP

Finally, we analyze the change of secondary structures during the simulation for both of the proteins. The starting secondary structures are almost identical, including the mutated region 76, where it stayed as a turn. However, in the mutated region 80, the region changed from a coil to a beta sheet. This indicates that the mutations did not significantly change the overall secondary structures of the proteins, and in particular, the regions of binding sites. Furthermore, there are no radical and significant changes to the secondary structure during the trajectory. This again provides evidence that the mutation did not change the overall properties of the protein.

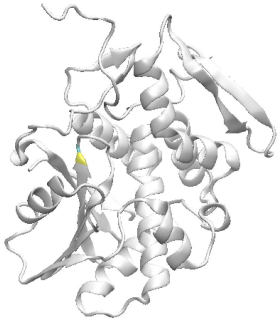


Secondary Structures of Control Protein

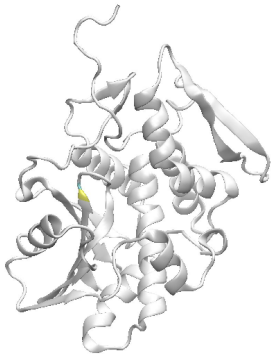


Secondary Structures of Mutated Protein

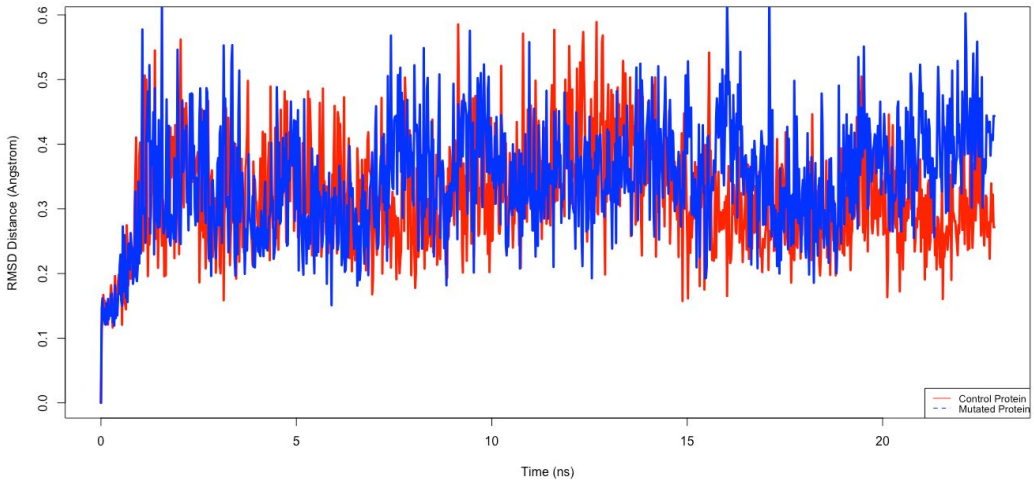
The first specific region that we want to consider is residue 76, which was mutated to ensure the removal of any trace VLS activity from the immunogen. We will consider the region of residue ALA73-ALA79, which surrounds that residue. We find that the RMSD of the region for both trajectories are rather similar. We also find that the RMSF at that region for both proteins to be very similar, indicating that the mutation did not influence the flexibility of that region. We also found that the number of hydrogen bonds nor the secondary at that region did not change in any significant ways. However, the RG value of the control protein was, on average, higher than the RG value of the mutated protein. This shows that the original mutation to reduce VLS activity worked, as this change in RG value shows a change of the binding action to cells that cause VLS.



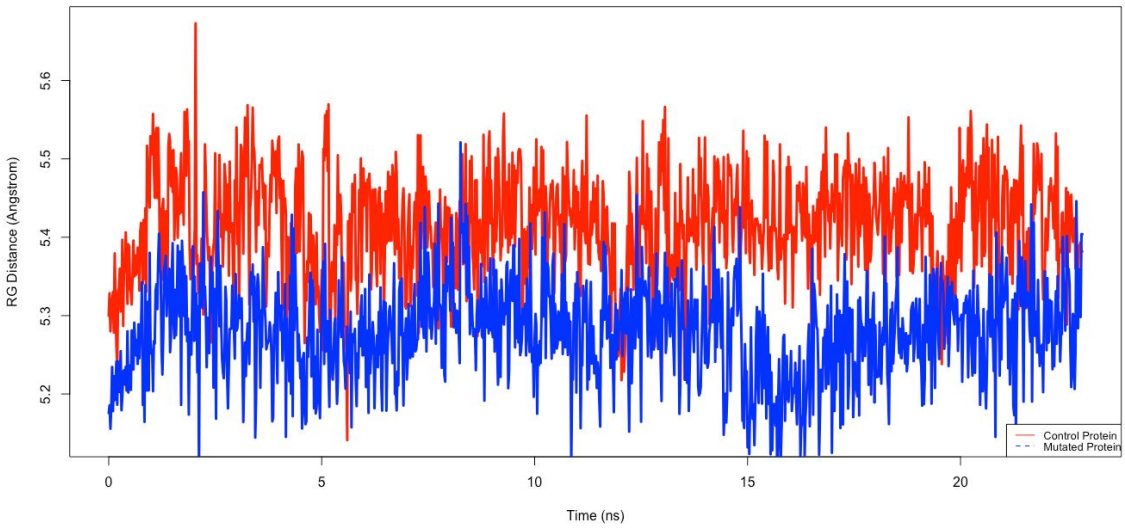
Residue MET76 (Original)



Residue VAL76 (Mutated)

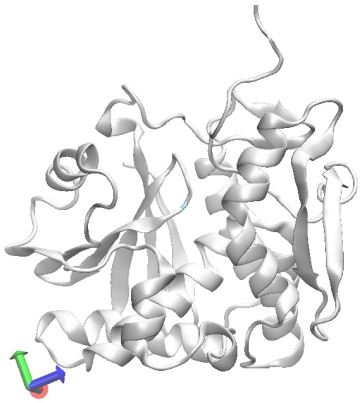


RMSD of 3SRP and Mutated 3SRP from Residues 73 to 79

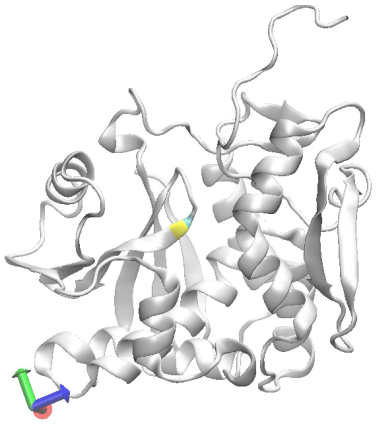


RG of 3SRP and Mutated 3SRP from Residues 73 to 79

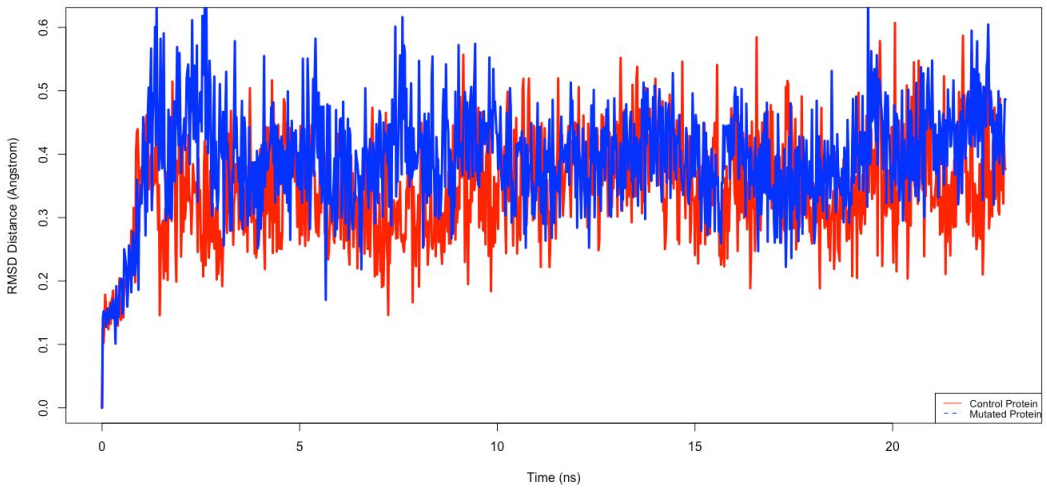
Next, we consider residue 80, which was mutated to inactivate enzyme catalysis. We will consider the region THR77-GLY83 that surrounds that residue. We found that there were no significant differences in the RMSD plot, RMSF plot, RG plot, and the hydrogen bonds plot. However, there was a change in the secondary structure at residue 80, which was changed from coil to a beta sheet. This was enough to change the interactions with enzyme catalysis, making 3SRP safer as a vaccine, as it will not interact destructively with nearby cells.



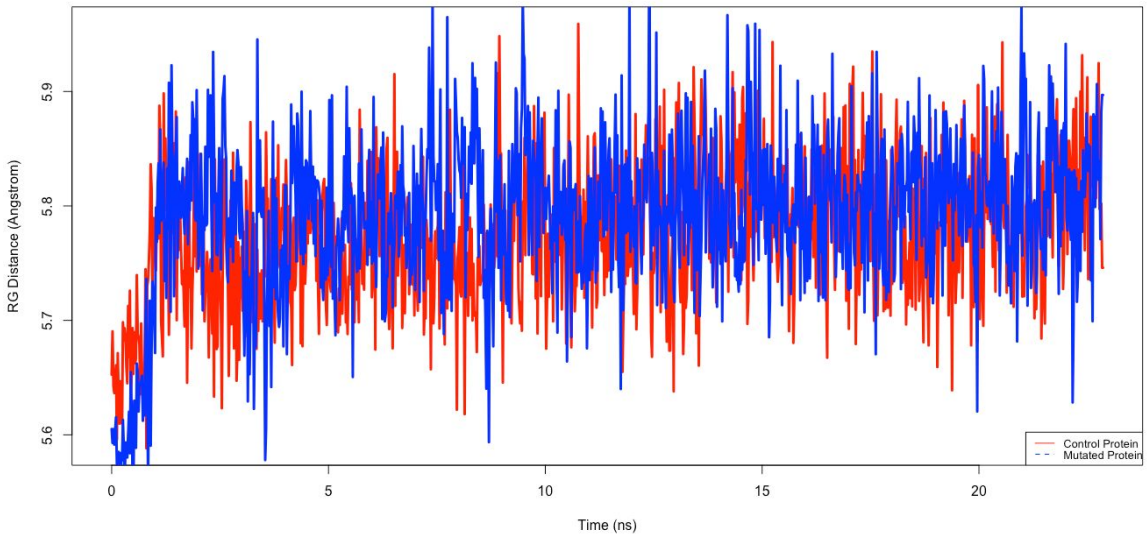
Residue ALA80 (Mutated)



Residue TYR80 (Mutated)

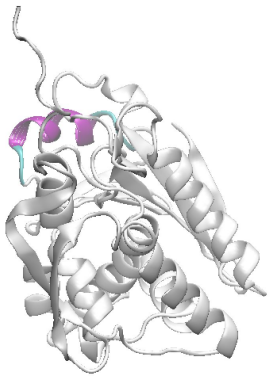


RMSD of 3SRP and Mutated 3SRP from Residues 77 to 83

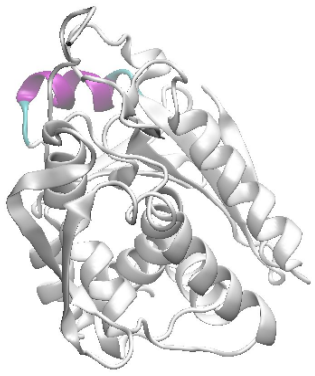


RG of 3SRP and Mutated 3SRP from Residues 77 to 83

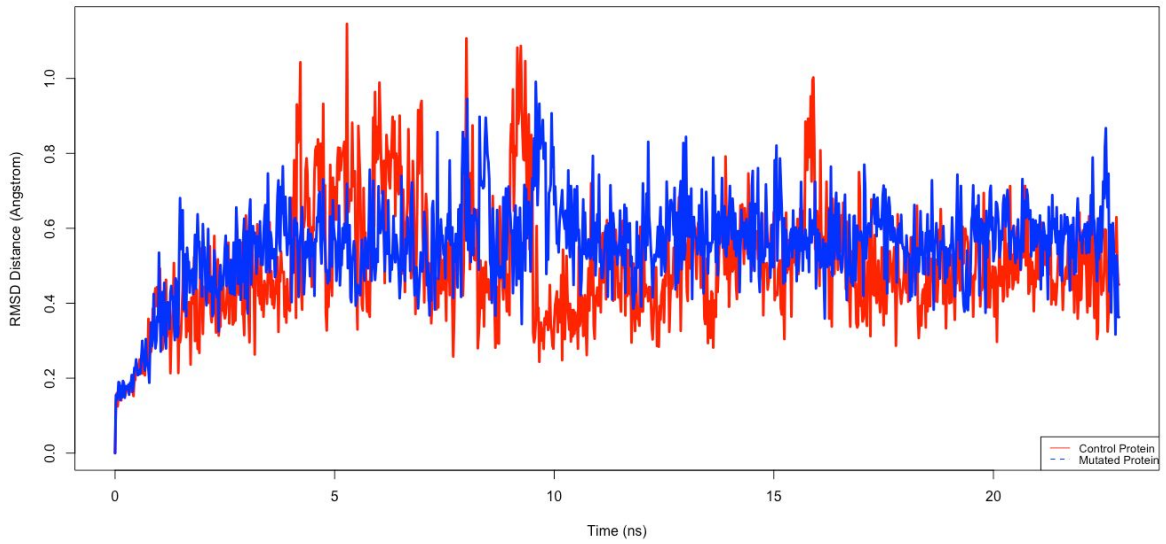
Next, we analyze regions of interest for residue 97 to 108, an important binding region for UNIVAX 70/138, a neutralizing antibody produced against RTA. This region is also important because it contains residue 99 to 106, an important epitope for generation of toxin-neutralizing antibodies. We want to ensure that the properties of this region did not drastically change from mutations in region 76 and 80, otherwise, antibodies will not bind properly to these regions, invalidating 3SRP's usefulness. In this region, RMSD, RMSF, and hydrogen bonds plots are very similar, and the secondary structure (an alpha helix contained between two turn structures) is the same. However, the RG plots are different, as the oscillation of RG values of the control protein is greater than that of the mutated protein. This difference may influence how antibodies may interact with that region.



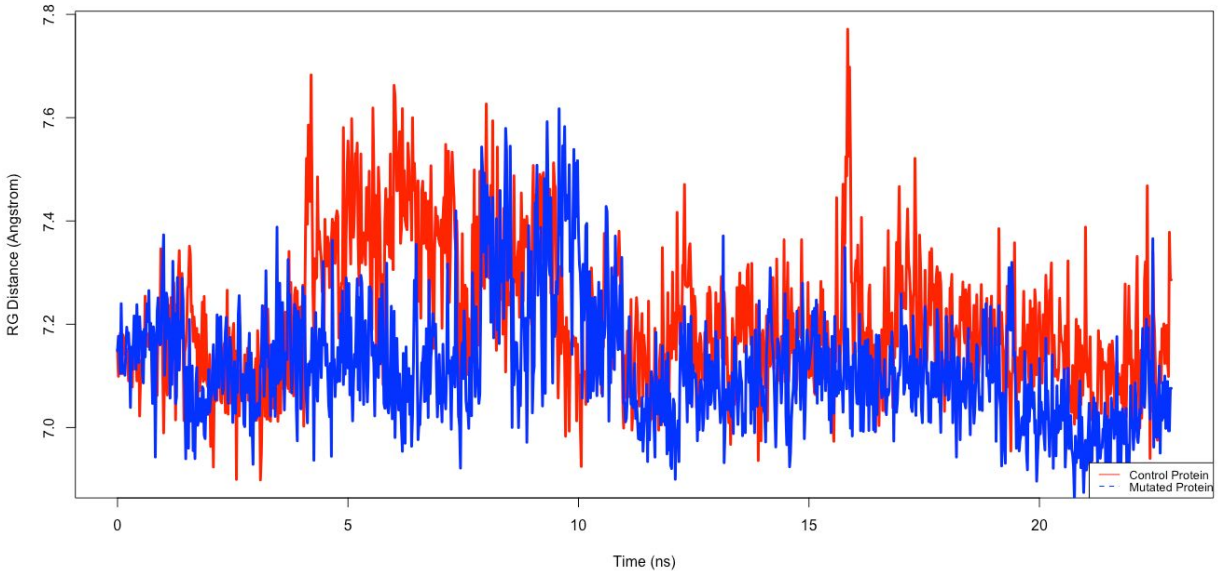
Residue ASN97 to PHE108 (Original)



Residue ASN97 to PHE108 (Mutated)

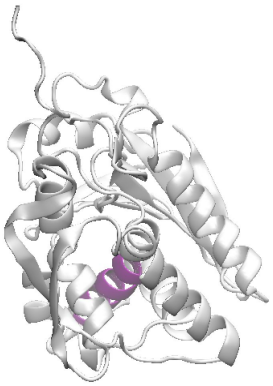


RMSD of 3SRP and Mutated 3SRP from Residues 97 to 108

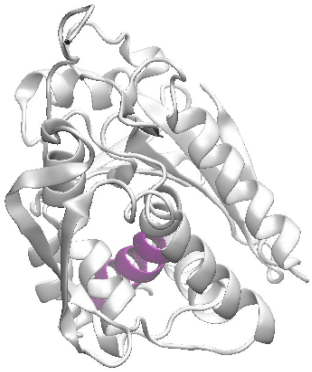


RG of 3SRP and Mutated 3SRP from Residues 97 to 108

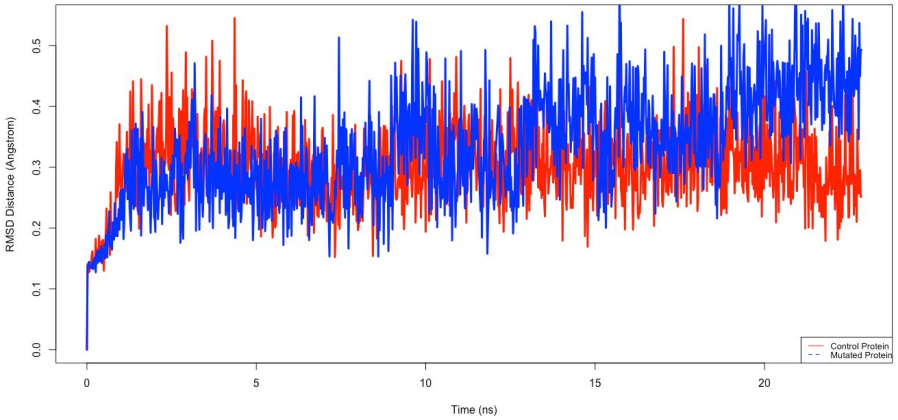
Next, we analyze regions of interest for residue THR163 to MET174, an important binding site for GD-12, another antibody. In this region, RMSF, RG, hydrogen bond plots are very similar, and in both proteins, the secondary structure is an alpha helix. The RMSD plots, on the otherhand, are slightly different, as the RMSD of the mutated protein is increasing, while it is staying constant in the control protein. However, the general properties seem to be the same, and thus verifies the author's claim that the mutations did not influence this binding region.



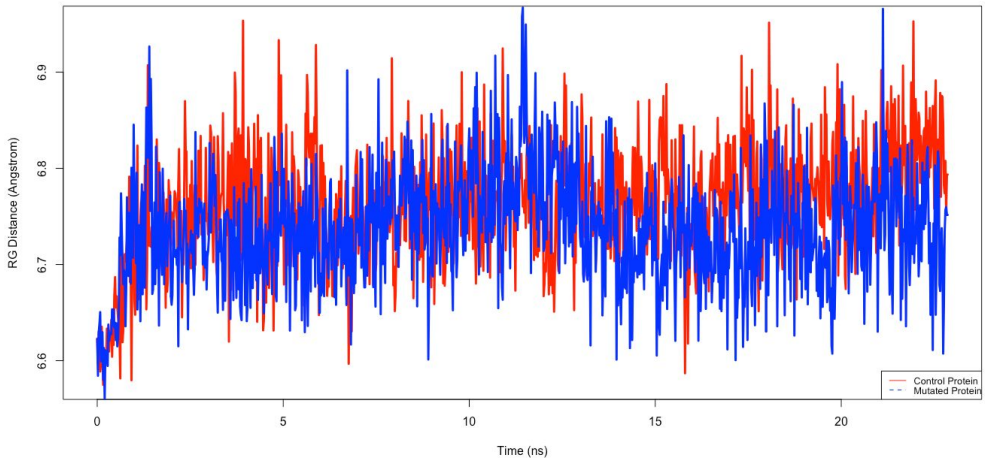
Residue THR163 to MET174 (Original)



Residue THR163 to MET174 (Mutated)

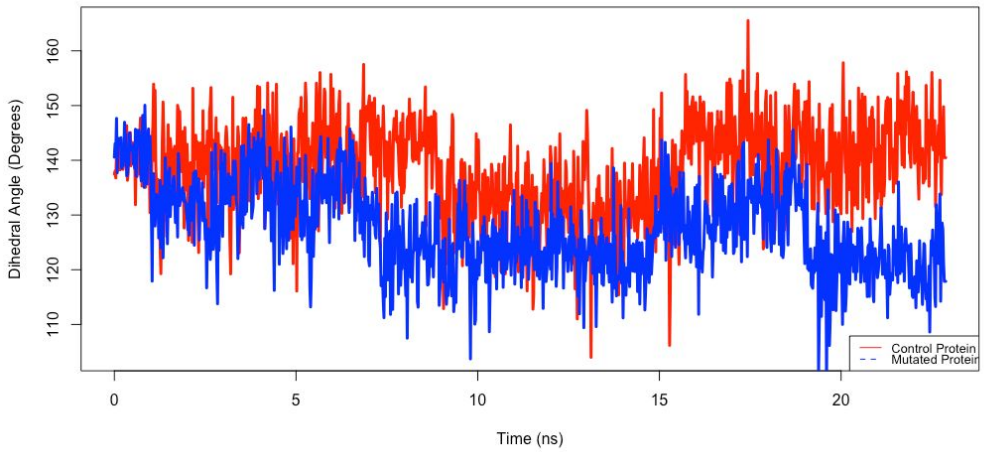


RMSD of 3SRP and Mutated 3SRP from Residues 163 to 174



RG of 3SRP and Mutated 3SRP from Residues 163 to 174

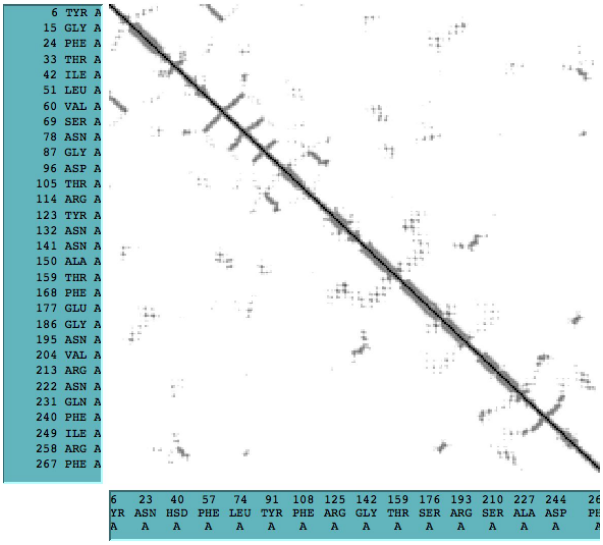
To further compare the two proteins, we calculate the change of torsion angles of four consecutive C-alpha atoms near residue 80 over the trajectory.[4] This is important to analyze because the author mutated residue 80 to inactivate possible enzyme catalyzes near that region. By analyzing the change of torsion angles near that region for both proteins, we are able to analyze and understand the region's receptibility for nearby possible enzymes. We find that the average dihedral angle is 138.81 degrees for the control protein and 127.78 degrees for the mutated protein. In addition, the standard deviation of dihedral angles for the control protein and mutated protein is 8.29 degrees and 8.20 degrees, respectively. This indicates that the oscillation of dihedral angles of both proteins is similar, meaning at the regions of both proteins, the flexibility are rather similar. However, whereas the dihedral angle of the control protein is generally staying constant, the dihedral angle of the mutated protein is decreasing. This decrease over time indicates that the mutation has changed the region's specificity to certain enzymes, confirming the authors claim that the mutation influenced the region's interactions with enzymes.



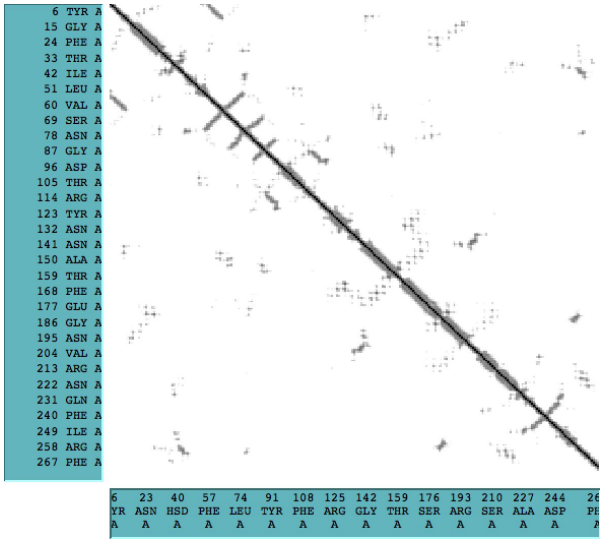
Dihedral Angle of C-Alpha for Resid 78-81

Next, we want to compare the contact maps between the proteins over the trajectory. We want to ensure that the contact maps of important binding regions between the two proteins are similar, as major differences would indicate that the properties of the protein dramatically changed, invalidating the author's claim. However, the contact map indicates that the regions of interest (MET76, ALA80, ASN97-PHE108, GLU99–HSD106, and THR163–MET174) of both proteins are identical.

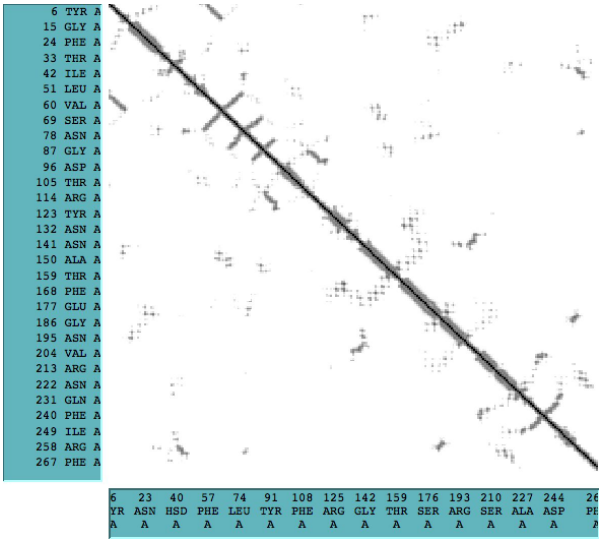
Beginning of Trajectory



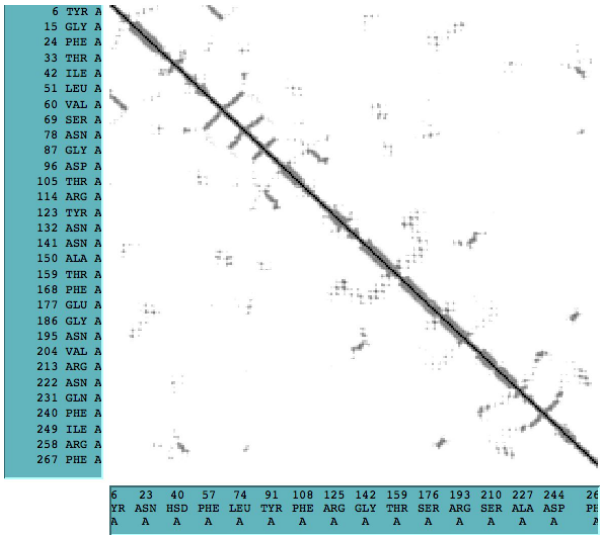
End of Trajectory



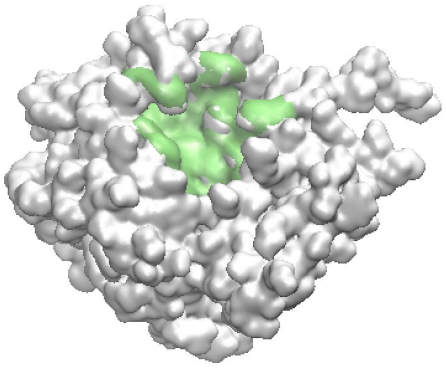
Contact Map of Original



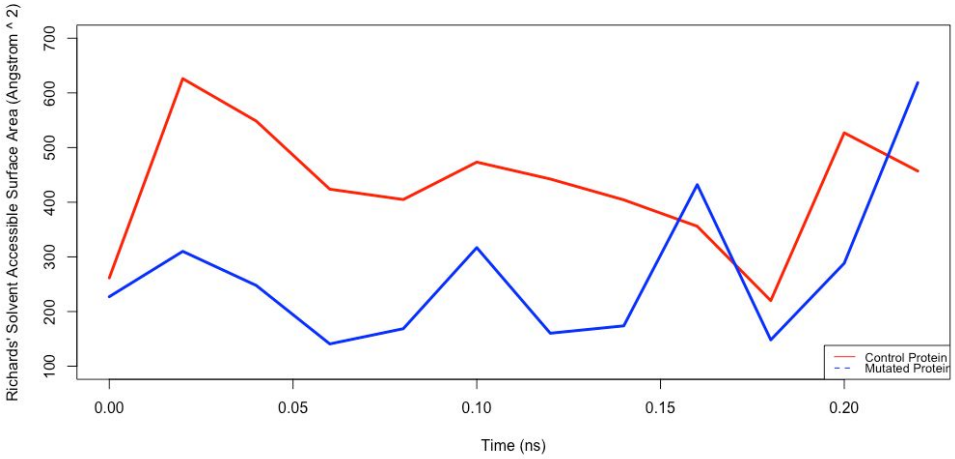
Contact Map of Mutated



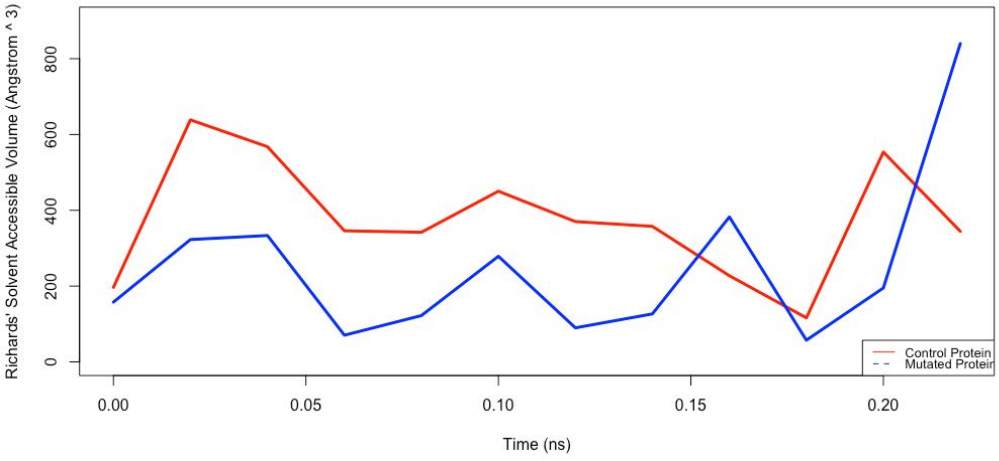
And lastly, we wish to analyze the cavity of the mutated region 76 for both proteins, as it will allow us to compare how the binding regions have changed from the mutation.[8] We used the CASTp program to find concave surface regions of 3SRP and its mutation. For every 2 nanoseconds, we calculated the Richards' Solvent Accessible Surface Area and Volume for the cavity at the region around residue 76 to see how the binding region changes. From the plot, we see that our mutation has decreased the size of the binding region of 3SRP, indicating that that region no longer fit with specific enzymes that would bind to that region. This verifies the claim that the mutation influenced the region's interactions with enzymes.



Cavity of Region Around Residue 76



Solvent Accessible Surface Area of 3SRP and Mutation



Although MD simulations are very valuable for analyzing interactions between molecules and gleaming insight that experiments can't, there are many pitfalls that we have to consider when using MD simulations to understand proteins.[2] For one, force field equations used to replicate reality (like spring equations to replicate bond stretching) are not necessarily accurate, and small errors can accumulate over long periods of simulations. The same force fields may introduce biases towards certain conditions. Equations can be modified to increase their accuracy, but that would be in expense of computational time. Another problem with MD simulations is that it takes an incredible amount of time to accurately simulate some important properties of proteins, such as hinge bending and elastic vibrations. Other actions of proteins, like rotations of buried side chains and local denaturations, are nearly impossible to replicate because they require seconds of time, and simulation steps are in femtoseconds. We can address this by increasing the time step and constraining bonds, but accuracy is drastically reduced. Furthermore, biological systems are incredibly complex, making important biological systems like membranes very computationally expensive to replicate. However, the increase in computational power may help address this problem.[1] In addition, simulating proteins in a simple system does not necessarily imply that the same protein will interact the same way with a more complex system. Furthermore, simulating a protein once does not imply that the protein will necessarily act the way it does in that one simulation. It is necessary to run the same simulation multiple times, and take averages over those simulations, which requires significantly more computation time.

And lastly, we acknowledge that removing the ligands from this simulation may reduce the stability of the protein and change the protein's interaction with the MD system.[6] However, we expect that the effects of removing ligands in our system is small.

Conclusion

In this paper, we used MD simulations to compare 3SRP and its reversed mutation. Using metrics such as RMSD, RMSF, RG, and the number of hydrogen bonds on both proteins helped confirm the claim that the mutations reduced the toxicity of RTA while not changing the overall properties of the important binding regions. Comparing salt bridges, secondary structures, contact maps, and dihedral angles help reached the same conclusion.

Although this MD simulation was able to confirm the author's claim that mutations did not affect the binding regions, there is still more work to be done. In the next steps, it would help to simulate 3SRP's interaction with antibodies that would bind to certain regions of 3SRP. For example, we could perform a simulation with 3SRP and multiple UNIVAX 70/138, and verify whether the binding action is the same for 3SRP and the mutation. Furthermore, we could also test if the mutation of V76M actually decreased its toxicity by doing an MD simulation with both variants of the protein with a membrane protein. We would analyze how that region binds to the membrane, and how it would break up regions of the membrane. And lastly, we would do another MD simulation to test whether the mutation Y80A actually inactivates catalysis by interacting both variants with known enzymes in the human body.

References

1. Adam Hospital, et al. 2015. Molecular dynamics simulations: advances and applications. *Adv Appl Bioinform Chem*. 8:37–47.
2. Daan Frenkel. 2012. Simulations: the dark side. *International School of Physics*.
3. Dimitrios Vlachakis, et al. 2014. Current State-of-the-Art Molecular Dynamics Methods and Applications. *Elsevier.* 94:269-313.
4. Fredrick Robin Devadoss, et al. 2014. Analysis and visual summarization of molecular dynamics simulation. *Journal of Cheminformatics.* 6 (1):16.
5. Lebeda F.J., et al. 1999. Prediction of a conserved, neutralizing epitope in ribosome-inactivating proteins. *Int J Biol Macromol.* 24(1):19-26.
6. M.R. Housaindokht, et al. 2008. Analysis of ligand binding to proteins using molecular dynamics simulations. *Journal of Theoretical Biology.* 254(2):294-300.
7. Patricia M. Legler, et al. 2011. Structure of RiVax: a recombinant ricin vaccine. *Acta Crystallogr D Biol Crystallogr*. 67(9): 826–830.
8. Wei Tian, et al. 2018. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Research*. 46(1):363–367.
9. William Humphrey, et al. 1996. VMD: Visual molecular dynamics. *Journal of Molecular Graphics.* 14(1):33-38.