JOURNAL OF
EMERGING INVESTIGATORS

Article

Exploring Interactions between PFAS (Per- and Polyfluoroalkyl Substances) and proteins

Hera Lu-Yang[1], Xiaoyu Wang[2], Min Zang[2], Hao Dong[2]
[1] BASIS Independent McLean, McLean, Virginia
[2] Nanjing University, Nanjing, China

SUMMARY

PFAS, or per- and poly-fluoroalkyl substances, while not as infamous or widely discussed as microplastics, are among the biggest threats to human health and development. Like microplastics, PFAS take an extremely long time to degrade, earning them the title of the "forever chemicals." We investigated the potential dangers of PFAS by focusing on how perfluorooctanoic acid (PFOA), one of the most studied PFAS, binds to bovine serum albumin (BSA), and how this interaction could impact protein function and human health. We used BSA because of its structural similarity to human serum albumin (HSA), the most abundant protein in adult plasma, making it integral to bodily functions. We hypothesized that PFOA would bind with high affinity to BSA, causing structural changes that might affect the protein's functions. Using several software programs, we examined two distinct configurations of the PFOA-BSA complex. We identified two configurations with differing properties, one having a greater stability than the other. However, despite having less stability, the second configuration had a higher binding energy, implying a stronger and more likely binding scenario in the human body. This may contribute to long-term health issues due to persistent PFAS binding, potentially impairing protein functions and leading to various health problems. Our study highlights the need for further research to understand these interactions, which could lead to strategies to address the health risks posed by PFAS exposure.

INTRODUCTION

According to the EPA, per- and poly-fluoroalkyl substances (PFAS) are a class of synthetic chemicals that have been used in various industrial and consumer products since the 1940s, due to their water- and oil-repellent properties (1). We interact with PFAS often throughout our lives, as these chemicals are found in a variety of materials, including non-stick cookware, water-resistant clothing, and cleaning products, (2). Though PFAS are less widely known by the public, their widespread use has led to their accumulation in the environment. PFAS have been detected in human bodies worldwide and are often called "forever chemicals" because they resist degradation in biologically relevant timeframes; their extremely slow degradation rates in natural environments and human systems lead to persistent bioaccumulation (3). For example, a study found that the half-life of perfluorooctanoic acid (PFOA), a type of PFAS, in the human body is around 2.36 years (4). PFAS have already been linked to development issues, liver and kidney damage, immune system suppression, and an increased risk of certain cancers (5).

In this study, we modeled the molecular interactions between PFAS and proteins in the body to help researchers understand how PFAS may contribute to these health concerns. Human serum albumin (HSA) is critical in the human body, providing oncotic pressure within capillaries, transporting fatty acids, bilirubin, minerals, and hormones, and functioning as both an anticoagulant and an antioxidant (6). Because HSA is integral to human health, we wanted to use HSA in our simulations. However, because its structure is so complex, we decided to use bovine serum albumin (BSA) instead, as BSA is a homologous protein with a simpler structure and is often used as a model for other homologous serum albumin proteins, including HSA (7). We aimed to investigate the interaction between PFOA, one of the most prevalent and studied PFAS, and BSA using a combination of molecular dynamic simulations, docking software, and energy minimization techniques.

We hypothesized that PFOA would bind with a high affinity to BSA, leading to structural changes that may affect the protein's function. To simulate this interaction, we first used a molecular docking program to mimic the interaction between the ligand and the protein before using a molecular dynamics simulation program to minimize the energy of the ligand-protein complexes and calculate their binding free energies. Our results supported our hypothesis. By doing this, we aimed to better understand the causes of potential health risks associated with PFAS exposure, particularly the long-term effects of PFAS-protein binding. The unusually high binding energy of the PFOA-BSA complex suggests a potential for disruption of protein function, which could contribute to the negative health effects observed in populations exposed to PFAS.

RESULTS

Using three software programs, Visual Molecular Dynamics (VMD), AutoDock Vina, and GROMACS, we simulated the interactions between PFOA and BSA, allowing us to observe the ligand-protein complex in an environment that resembled the human body (8-11). After using VMD, a visual molecular dynamics program, to select one of the two identical monomers of BSA that we wanted to study, we uploaded both the protein structure and the ligand (PFOA) structure to AutoDock Vina, a docking software. We also observed the ligand structure in VMD before being uploaded to AutoDock Vina (**Figure 1**). We selected only half of the protein, as BSA is a homodimer, containing two identical subdomains. We did this to reduce the time needed for our simulation, as well as to help us visualize the results better.

https://doi.org/10.59720/25-038

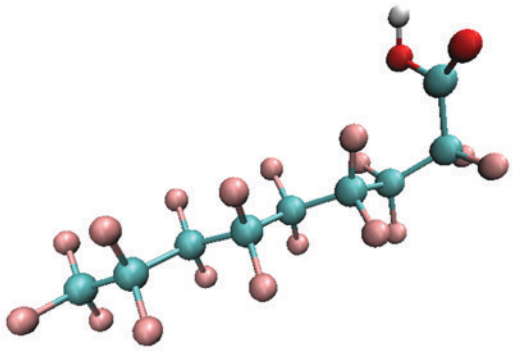

Figure 1: Molecular structure of perfluorooctanoic acid (PFOA). The structure was generated in PyMol and is loaded on Visual Molecular Dynamics (VMD) for visualization purposes. The atoms are color-coded: carbon (blue), fluorine (pink), oxygen (red), and hydrogen (white).

AutoDock Vina yielded nine possible configurations of the ligand-protein complex. We would only continue to analyze two of the nine possible arrangements, choosing one in each distinct protein pocket (**Figure 2**) with the strongest binding energies for the most contrasting results, complexes 1 and 3. However, once we re-uploaded each of the complexes' structure files back into VMD, we had less than ten frames of the complex to observe, making it difficult to see details within the interaction.

To solve this issue of having minimal snapshots to analyze, we used an energy minimization program to not only give us more than a thousand frames of the interaction for each complex, but we also were able to produce root mean square deviation (RMSD), and root mean square fluctuation (RMSF), plots for each configuration, allowing us to visualize the stability of each complex. RMSD is a numerical measurement of the difference between two structures, or in our case, how much our structure moved from its original unbound conformation. A lower RMSD means smaller average displacement and higher stability. RMSF shows us the specific regions of the structure that had an unusually high distance from their original positions. Those amino acids with high fluctuation were also able to be highlighted in VMD, showing that the parts of the structure with unusual movement were all close to the ligand binding site (**Figure 3**).

The residues with the most movement are clustered near the ligand binding site, suggesting localized fluctuations in response to PFOA binding (**Figure 3A**). The spatial orientation of the ligand relative to nearby residues also allows for several potential interactions (**Figure 3B**). However, it is important to note that the hydrogen bonds shown were generated using an automated distance-based selection tool in VMD and have not been validated experimentally. Because true hydrogen bonds between amino acid residues and fluorine atoms rarely exceed 4 Å in length, any interactions shown beyond this threshold were considered hypothetical and may instead reflect transient spatial proximity rather than stable bonding (12). Complex 1 had an average RMSD of around 2.5 Å, while complex 3 had an average RMSD of around 3 Å (**Figure 4**). Moreover, complex 1 had two areas of unusually high fluctuation, while complex 3 had four areas of high fluctuation (**Figure 5**).

Finally, using GROMACS, we calculated the final binding energy of each complex. We found that complex 1 had an average binding energy of -13.66 ± 4.07 kJ/mol (mean ± STD), while complex 3 had an average of -19.75 ± 3.23 kJ/mol, suggesting that complex 3's configuration is more thermodynamically favorable.

DISCUSSION

Using protein modeling software, we were able to identify and characterize interactions between PFOA and BSA at two distinct binding pockets. We found two complexes with negative average binding energies, supporting our hypothesis that PFOA would bind to BSA. RMSD is commonly used to assess the stability of molecular dynamics simulations, where smaller deviations from the starting structure often indicate

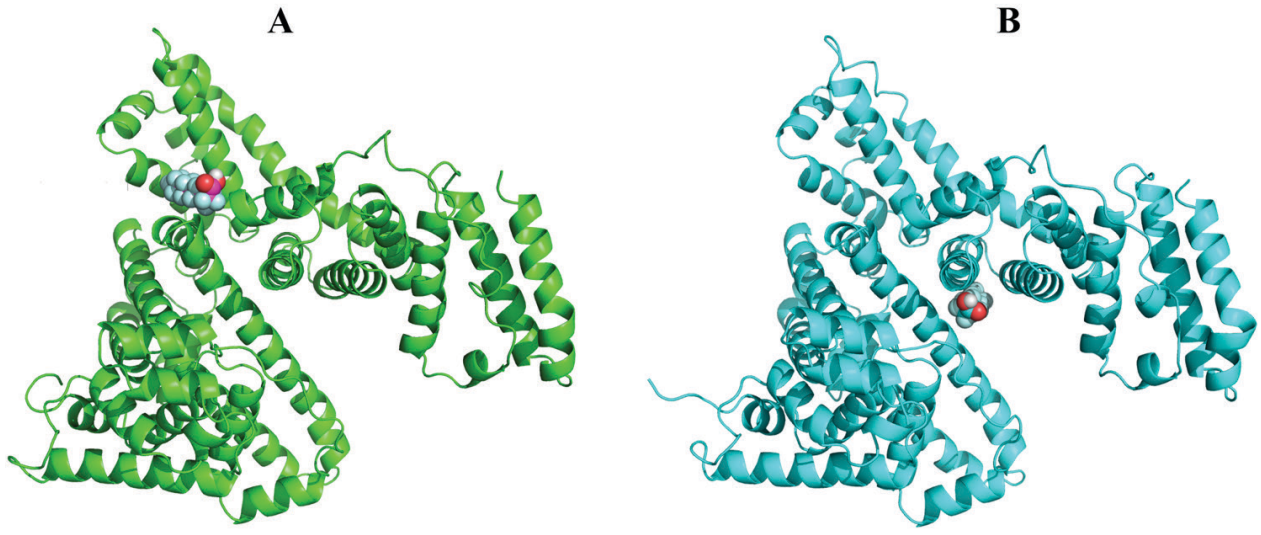

Figure 2: PFOA was docked on bovine serum albumin (BSA) using AutoDock Vina. A) Complex 1 and **B)** complex 3 as visualized in VMD. The BSA protein is displayed using a cartoon model (green and blue), while the ligand (PFOA) is displayed using a spheres model. The ligand is located in different pockets in both complexes. Both complexes are shown on the same scale.

https://doi.org/10.59720/25-038

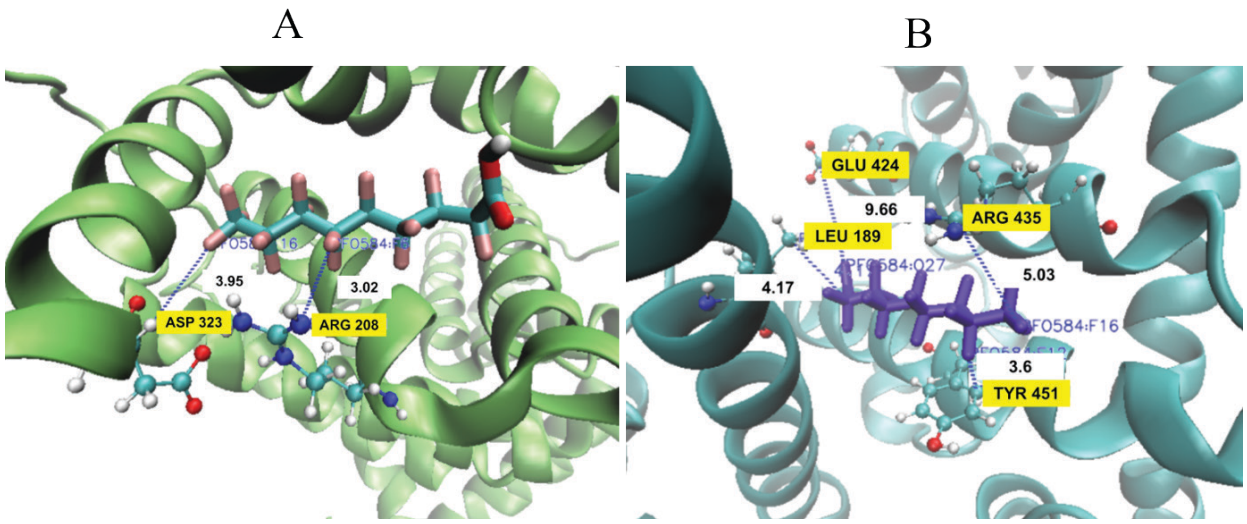

Figure 3: Close-up view of the binding sites. The BSA protein is represented by a cartoon model, and the substrate by stick model. Predicted interacting amino acids within the binding pocket are shown with a ball and stick model for **A)** complex 1 and **B)** complex 3. The blue dotted lines represent the intermolecular forces between the amino acids and the ligand, and the white boxes show the bond lengths in Å. They were found using the Non-Covalent Interaction (NCI) analysis method within VMD.

a more stable conformation over time (13). Since the graph showed consistent fluctuation around 2.5 Å, we can infer that the ligand-protein complex generally remained close to its unbound conformation during the simulation, indicating limited structural change. On the other hand, complex 3's RMSD graph was much more irregular, fluctuating around 3.0 Å, with larger and more extreme peaks. This suggests that complex 1 is more stable than complex 3, as it moved less due to the ligand's connection. This lines up with the RMSF plots, which show that complex 1 had just two unusually elevated fluctuation points, while complex 3 had four. Additionally, in both complexes, the amino acids with the highest fluctuation were all near the bound ligand, suggesting that the movement was because of the PFOA molecule.

Our molecular dynamics simulations revealed that PFOA binding to BSA is thermodynamically favorable, as evidenced by negative ΔG_0 values across all tested complexes. The more negative ΔG_0 values (particularly for complex 3) indicate stronger spontaneous binding affinity. While this metric provides a robust measure of overall binding favorability, we note that our current simulation approach cannot resolve the underlying energetic components (enthalpic versus entropic contributions) that collectively determine ΔG_0. One limitation of our experiment is that the binding energy calculations, which were based on the Newtonian laws and equations of motion that GROMACS provides, may not fully account for all factors influencing protein-ligand interactions in a biological environment. We used generalized force fields and solvent models that may not perfectly represent the true conditions within the human body. Additionally, subtle differences between BSA and HSA, including sequence variations and post-translational modifications, could influence the binding

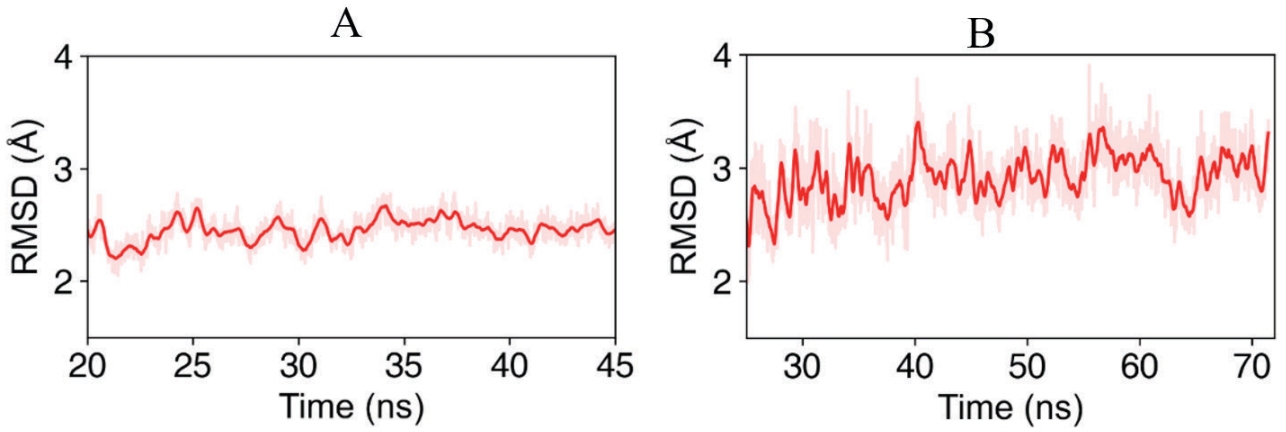

Figure 4: RMSD (Root Mean Square Deviation) plots. A) complex 1 and **B)** complex 3. These plots were made using the data that GROMACS produced and VMD's TK console. The light red line represents the raw data, while the dark red line is a smoothed version of the data, allowing us to see the overall trend. Both systems reached equilibrium at 2.5 Å and 3 Å, respectively.

https://doi.org/10.59720/25-038

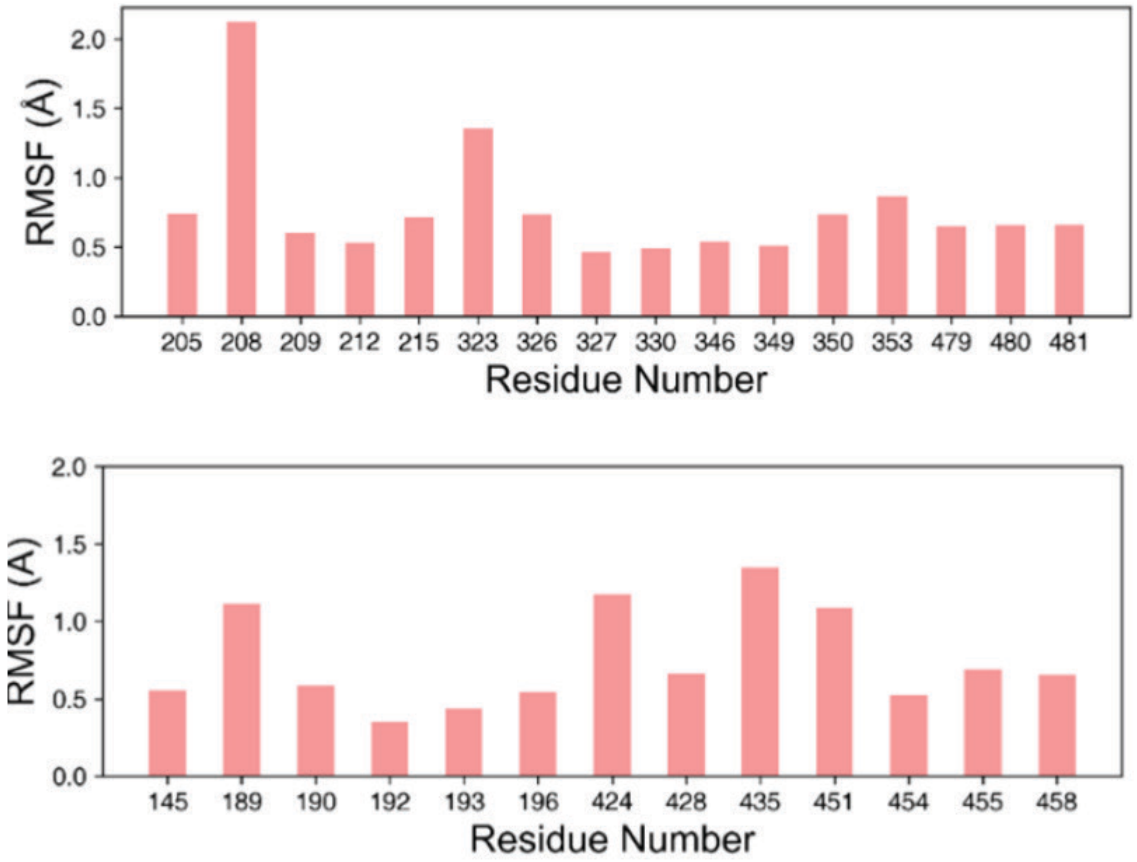

Figure 5: The RMSF (Root Mean Square Fluctuation) plot. Plots produced using GROMACS and VMD. **A)** Complex 1 has two unusually high peaks at amino acid residues 208 (Arg) and 323 (Asp). **B)** Complex 3 has four unusually high peaks at amino acid residues 189 (Leu), 424 (Glu), 435 (Arg), and 451 (Tyr).

behavior of PFOA. Despite these limitations, our simulations strongly suggest that PFOA would bind to BSA *in vivo*, supporting the hypothesis that PFOA (and potentially other PFAS) may interact with HSA in the human body.

Though complex 3 had a higher binding energy, both complexes demonstrated strong binding affinities as they had negative ΔG_0 values, which may suggest a molecular basis for how PFAS can persist in the human body. While our molecular dynamics simulations did not reveal significant conformational changes in BSA, as the RMSD remained relatively low and the RMSF fluctuations were limited to a few residues, the strong binding energies observed suggest that PFAS such as PFOA can stably associate with serum proteins.

Even though our data do not demonstrate major disruptions in protein structure, prior studies have proposed that PFAS binding may interfere with normal protein transport functions by competitively occupying key binding sites or altering ligand-binding kinetics (14). Molecular recognition is the highly specific interaction between a protein and its ligand. When a protein encounters a molecule like PFAS, this specificity may break down, and the protein may not be able to effectively distinguish between PFAS and its natural substrate. Therefore, the protein might undergo a conformational change to adapt to the structure of the PFAS, rather than its designated substrate, which could affect its function. Further research, particularly involving experimental approaches to observe real-life interactions rather than simulations, is needed to investigate whether PFAS binding induces allosteric effects or contributes to misfolding under physiological conditions.

Through this molecular simulation study, we demonstrated that PFAS compounds can bind to serum albumin with notable affinity and in multiple configurations. While no substantial structural destabilization was observed, our results support the hypothesis that the high binding affinity may explain why PFAS remain bound to proteins in the body, underscoring the need to further explore their potential biological impact. Future studies are warranted to study the interactions between HSA and PFAS to give more accurate information about the binding affinity and the effects of binding on human protein. Additionally, future research should investigate possible ways to combat the binding between the ligand and the protein. This could potentially offset the negative impacts of PFAS on the human body, as it would allow proteins to be protected from possible denaturation.

MATERIALS AND METHODS

First, the PDB file for BSA was downloaded (PDB: 4F5S). Since the protein structure is a homodimer (i.e., composed of

https://doi.org/10.59720/25-038

two identical monomers), chain A was used for our docking experiment. PyMol was used to create a 3D model of PFOA (15). VMD was then used to examine both the protein and the ligand structures separately. Since BSA is made up of two homologous chains, VMD's built-in program, TK console, was used to select one of the two chains to analyze further to save time and computational energy.

The protein structure file and ligand file were uploaded to AutoDock Vina, a program that simulates many different docking possibilities to give us the most probable orientation. Then, the complexes were sorted from the strongest to weakest binding energies. The two complexes with the strongest binding energies were kept to study.

GROMACS, a molecular dynamic package that simulates real biological environments (pH of 7 and 0.15 mM NaCl) was used to put the complex in an environment mimicking the human body Using the CHARMM36 force field, a topology file for both the protein and the ligand was produced. Next, to produce more frames of the interaction to examine, energy minimization techniques were used to allow the computer to produce them without exhausting itself.

Finally, the files were uploaded to a supercomputer in NPT equilibrium, meaning that the number of moles, the pressure, and the temperature were all held constant. Using the RMSD tool in VMD, RMSD and RMSF graphs were produced. The binding energies were also calculated using GROMACS.

Received: October 16, 2024
Accepted: September 1, 2025
Published: May 16, 2026

REFERENCES

1. "PFAS Explained | US EPA." *Environmental Protection Agency (EPA)*, 2023, https://www.epa.gov/pfas/pfas-explained. Accessed 13 August 2024.
2. "Chemicals: Perfluoroalkyl and Polyfluoroalkyl (PFAS) Substances." *Wisconsin Department of Health Services*, 2024, https://www.dhs.wisconsin.gov/chemical/pfas.htm. Accessed 13 August 2024.
3. "PFAS, what is it, why is it in our environment and why is it a problem?" *PFAS Free*, 2024, https://www.pfasfree.org.uk/about-pfas. Accessed 13 August 2024.
4. Batzella, Erich, et al. "Determinants of PFOA Serum Half-Life after End of Exposure: A Longitudinal Study on Highly Exposed Subjects in the Veneto Region." *Environmental Health Perspectives*, vol. 132, no. 2, 2024, p. 027002. https://doi.org/10.1289/EHP13152
5. "PFOA, PFOS, and Related PFAS Chemicals." *American Cancer Society*, 2024, https://www.cancer.org/cancer/risk-prevention/chemicals/teflon-and-perfluorooctanoic-acid-pfoa.html. Accessed 13 August 2024.
6. Chen, C. B., et al. "Overview of Albumin Physiology and Its Role in Pediatric Diseases." *Current Gastroenterology Reports*, vol. 23, no. 8, 2021, p. 11, https://doi.org/10.1007/s11894-021-00813-6.
7. Topală, Tamara, et al. "Bovine Serum Albumin Interactions with Metal Complexes." *Medicine and Pharmacy Reports*, vol. 87, no. 4, 12 Nov. 2014, pp. 215–219, https://doi.org/10.15386/cjmed-357.
8. Humphrey, W., et al. "VMD – Visual Molecular Dynamics." *Journal of Molecular Graphics*, vol. 14, 1996, pp. 33–38, https://doi.org/10.1016/0263-7855(96)00018-5.
9. Eberhardt, J., et al. "AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings." *Journal of Chemical Information and Modeling*. https://doi.org/10.1021/acs.jcim.1c00203.
10. Trott, O., and Olson, A. J. "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading." *Journal of Computational Chemistry*, 31(2), 455-461. https://doi.org/10.1002/jcc.21334.
11. Abraham, M., et al. GROMACS 2026.1 Manual. 2026.1, Zenodo, 6 Mar. 2026, https://doi.org/10.5281/zenodo.18886967.
12. Pietruś, W., et al. "Hydrogen Bonds with Fluorine in Ligand–Protein Complexes – The PDB Analysis and Energy Calculations." *Molecules,* vol. 27, no. 3, 2022, article 1005. https://doi.org/10.3390/molecules27031005
13. Hollingsworth, S. A., and Dror, R. O. "Molecular Dynamics Simulation for All." *Neuron*, vol. 99, no. 6, 2018, pp. 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011
14. Gallagher, A., et al. "Computational Modeling of Human Serum Albumin Binding of Per- and Polyfluoroalkyl Substances Employing QSAR, Read-Across, and Docking." *Molecules*, vol. 28, no. 14, 2023, article 5375. https://doi.org/10.3390/molecules28145375
15. Schrödinger, LLC. "The PyMOL Molecular Graphics System, Version 1.8." *Schrödinger, LLC*, 2015.