Different approaches to solvation in molecular dynamics simulations

E.Pepeliaeva, Tutors: J.Belyaeva, A.Zlobin, A.Golovin

INTRODUCTION

Water plays an important role in maintaining protein stability and mediating enzymatic reactions. The most popular software for performing MD simulations (Gromacs) solvates MD system through a single command. Due to its implementation, water molecules can be placed into the thermodynamically irrelevant places. Therefore, it is important to place them accurately while performing *in silico* simulations of molecular systems.

AIM

Here we aim to check existing protocoles of solvating molecular systems for MD simulations, while inspecting the impact of water molecules positioning on the protein conformational stability.

METHODS

We selected 3 proteins for our work: ubiquitin (1UBQ), lysozyme (1IEE), and pancreatic trypsin inhibitor (4PTI). All simulations were performed with Gromacs toolkit. We used numpy, pandas, and matplotlib Python3 libraries for analysis of the obtained results.

REFERENCES

1. Mark Abraham et al. (2023). GROMACS 2023 Manual (Version 2023). Zenodo. https://doi.org/10.5281/zenodo.7588711
2. Yin H, Feng G, Clore GM, Hummer G, Rasaiah JC. Water in the polar and nonpolar cavities of the protein interleukin-1β. J Phys Chem B. 2010 Dec 16;114(49):16290-7. doi: 10.1021/jp108731r. Epub 2010 Nov 3. PMID: 21047091; PMCID: PMC3005849.
3. Lemkul, Justin. (2018). From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0]. Living Journal of Computational Molecular Science. 1. 10.33011/livecoms.1.1.5068.

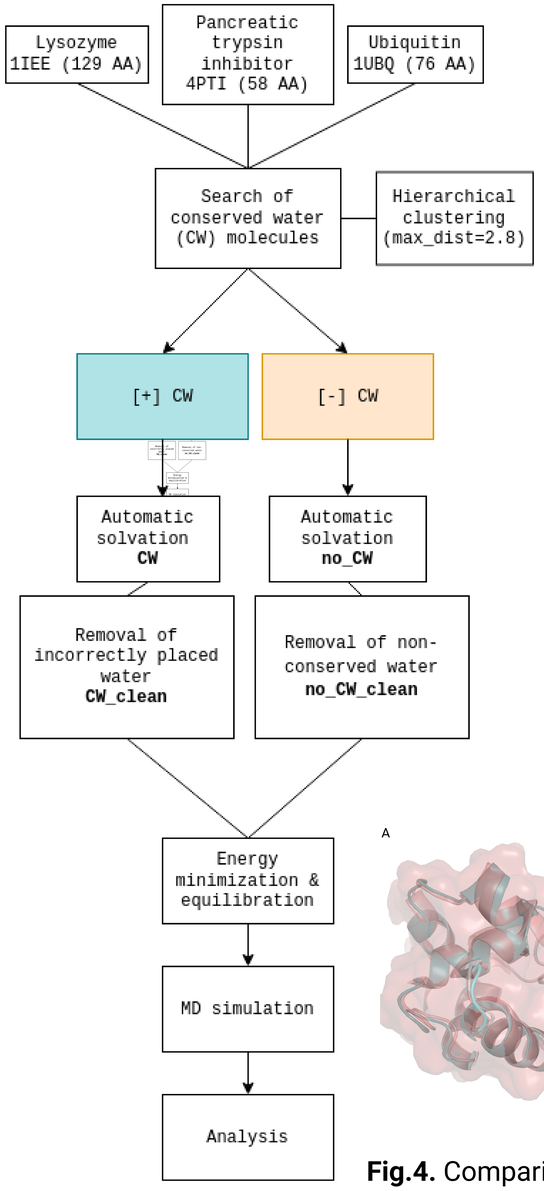

Fig.1. Workflow

RESULTS

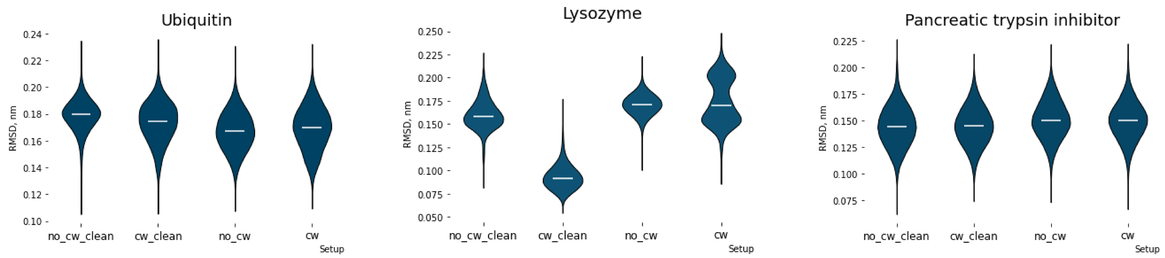

Fig.2. Distribution of RMSD values of secondary structure elements for different setups

We used root-mean-square deviation (RMSD) to measure secondary structure elements displacement during simulation. Respective X-ray structures were used as a reference. Thus, we were able to indirectly estimate protein stability.

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}$$

where N — the number of atoms used for RMSD calculation, δi — the distance between atom i and respective atom from a reference structure.

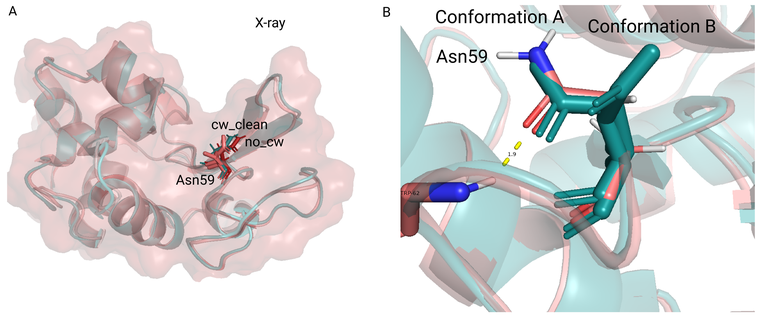

Fig.4. Comparison of lysozyme (A) average structures (B) Asn59 conformations in no_cw_clean system VS PDB structure.

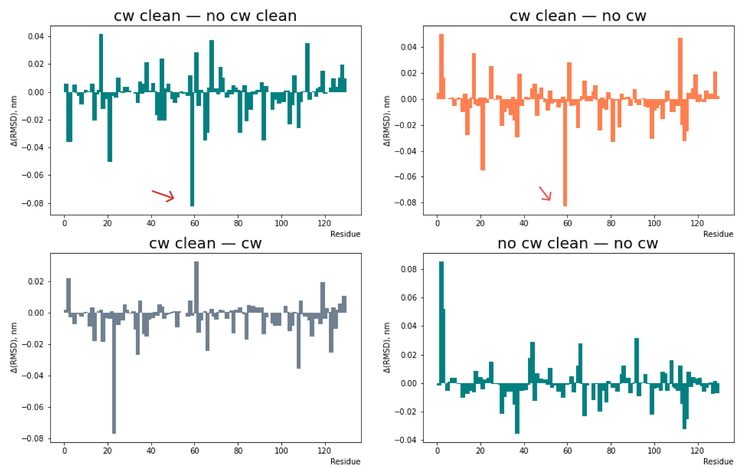

Fig.3. Difference in RMSD for each residue of lysozyme (1IEE). We noticed a sharp difference in RMSD values for ASN59 (arrow), which isn't present in other cases. The value is negative, indicating that cw setups have lower RMSD compared to no_cw ones.

CONCLUSION

We observed a significant difference in the performance of the lysozyme MD simulation with the **cw_clean** setup showing the least average RMSD and, hence, the greater stability of the protein. Other proteins didn't have drastic difference in performance, which might happen due to the lack of any inner crystallographic water (in the case of ubiquitin) or negligible amount of them (pancreatic trypsin inhibitor).