CHE622A - Introduction To Molecular Simulation

Structure and Internal Dynamics of the Bovine Pancreatic Trypsin Inhibitor Protein in Aqueous solution from the long-time molecular dynamic simulations

Instructor: Dr. Jayant Singh

Submitted By:

Vaibhav Anand (18807843)

Saumya Sehgal (190779)

Date Of Submission - 2 May, 2022

Abstract -

We have studied the structural and dynamic properties of BPTI protein using two molecular dynamic simulations of each 5ns. The first simulation was a normal simulation of molecules; the second was with atom-atom distance restraints. For each simulation, we have analyzed properties such as potential energy, coulombic energies, the radius of gyration, and rmsd values. To study protein dynamism, we tried to study the temporal variation of dihedral angles. It is known from the literature that dihedral angles (phi and psi) of consecutive residues result in such transitions. For each simulation, we get conformation at which backbone angle transition occurs. These conformers are further used for the comparative study of protein dynamics.

Introduction -

We aim to study the structural and dynamic properties of BPTI protein. Bovine pancreatic trypsin inhibitor (BPTI) is one of the smallest and simplest globular proteins, a mere 58 residues long. Due to its small size, BPTI has a relatively small hydrophobic core and needs some extra stability. This is provided by the protein's conserved cysteine residues that form disulfide bonds and stabilize its 3D structure. The presence of three disulfide bonds makes BPTI one of the most stable globular proteins.

Of the two molecular dynamic simulations we performed, the first was a standard simulation. In the second simulation, we chose to include atom-atom distance restraints to improve the quality of trajectory and obtain results closer to those expected for the crystalline structure of the protein.

Tools Used -

We made use of

- The 2022 version of Gromacs for running simulations.
- Pymol/Chimera for protein visualization and sequence analysis.
- xmgrace for visualizing graphs out of generated files such as xvg.

Methodology/Simulation Details -

For Simulation A, an extended simple point charge water model was used, a parameterized version of the simple point charge model of water. The dimensions of the cubic boxes used for simulating were one nm, and the system was equilibrated to a temperature and pressure of 277K and 1 Bar, respectively. The SHAKE method was used to keep bonds rigid. A cutoff radius of 0.8 nm was used. The OPLS-AA force field was used, and the protein database ID was 1 PIT.

For Simulation B, similar steps were taken, but we equilibrated the system with the use of atom-atom distance restraints until 277K and 1 Bar for greater accuracy with NMR structures. The atom-atom distance constraint files for our protein were unavailable, so we used a similar protein that differed from our protein by a single mutation. The PDB ID for this protein was 1JV8. To use its distance constraints, we first deleted the restraints for the atoms corresponding to mutated residues. As atom numbers may vary from amino acid to amino acid, we changed atom numbers to make sure that the remaining atoms map with constraints.

Later, crystallographic B-factor crystal structure II of BPTI protein was used for comparison of properties with conformers we observed based on backbone angle transitions.
PDB ID of Crystal Structure used for comparison: 5PTI

Results With Discussion -

Result 1: Stability of Total PE and LJ Protein-solvent energy for Simulation A:

Stability of Total Potential energy and Lennard-Jones energy for protein-solvent interactions are used to verify the equilibration of simulation A. In the first image, the red part indicates the time-averaged values of energy (averaged over 10ps). In the second image, we observe some fluctuations, but the energy values still look stable and less scattered around their mean value.

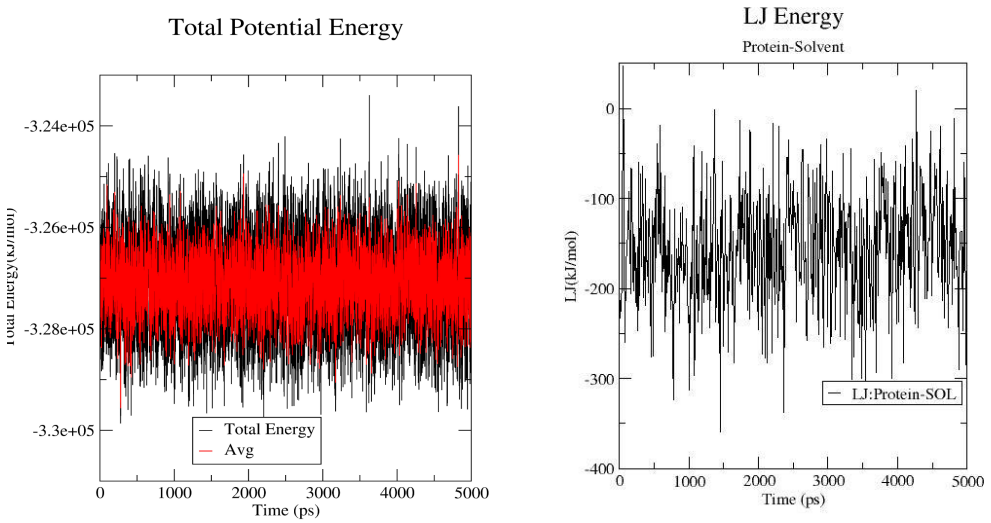


Figure 1a (left) and 1b(right)

Result 2: Coulomb energy of protein-solvent and protein-protein for simulation B -

Similar to results in 1, we generated and analyzed the Coulombic energy values for protein-protein and protein-solvent to verify equilibration. Here, we get oscillatory data points similar to those in the literature. It can be observed that oscillations are large until 2ns; after that values come close to the mean value.

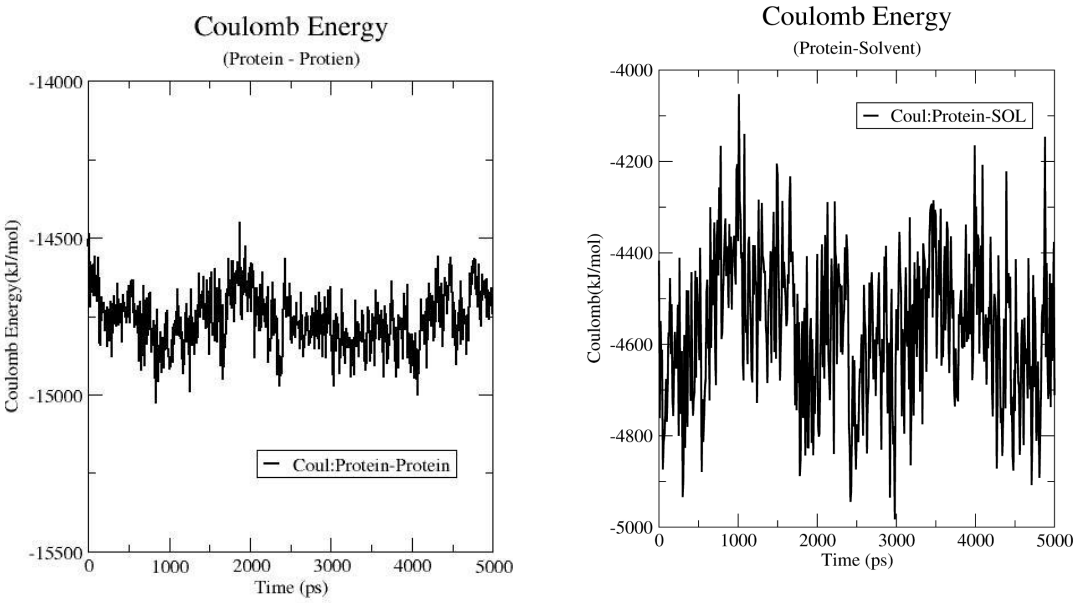


Figure 2a(left) and 2b(right)

Result 3: Variation of 10ps average of RG with time -

We have generated the radius of gyration plots for both simulations. The values have been averaged over 10ps, to get clear results. The values are stable and close to their mean value for both simulations.

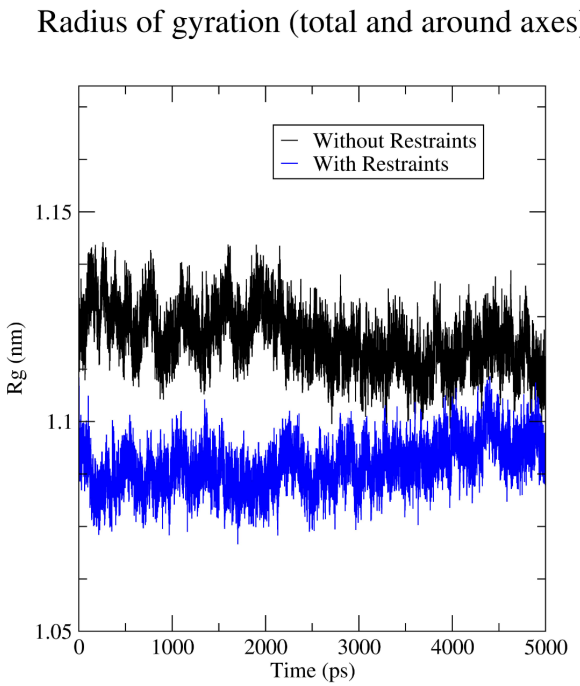


Figure 3

Result 4: rmsd of atoms averaged over 10ps -

We have also generated the plot for rmsd values for both simulations. Values are similar to what we observe in literature. Also, it can be observed that rmsd values for restraint simulation are lower than that without restraints. This can be attributed to the restrained moments of atoms in simulation B, as we are giving lower bound and upper bound for atom-atom movements.

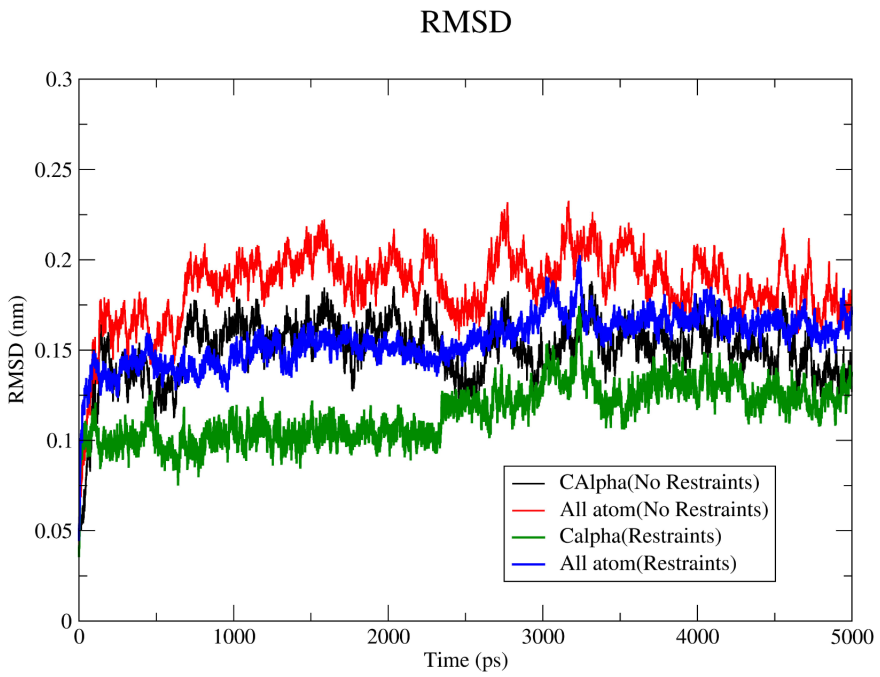


Figure 4

Result 5: Conformation from Trajectories -

To study the dynamism of the BPTI protein, we choose to analyze dihedral angles. It's been known from the literature that backbone angle transition occurs when phi and psi angles of consecutive residues fluctuate/change larger than 30 degrees. The following figures are from the literature we chose.

We can see 4 such fluctuations in (5a) and 3 in (5b). But, as we chose different force fields and slightly different parameters in mdp files, we didn't get the exact results. Instead, we get results which are shown in fig 6. Although we randomly chose three timestamps assuming transitions have occurred there.

Further, these conformers are used for comparative analysis of positional rmsf of alpha and gamma carbon atoms.

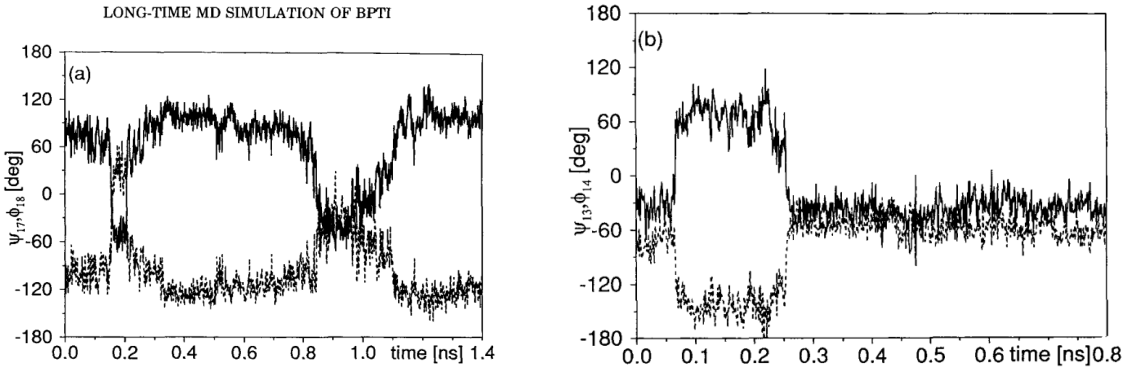


Figure 5a(left) and 5b(right)

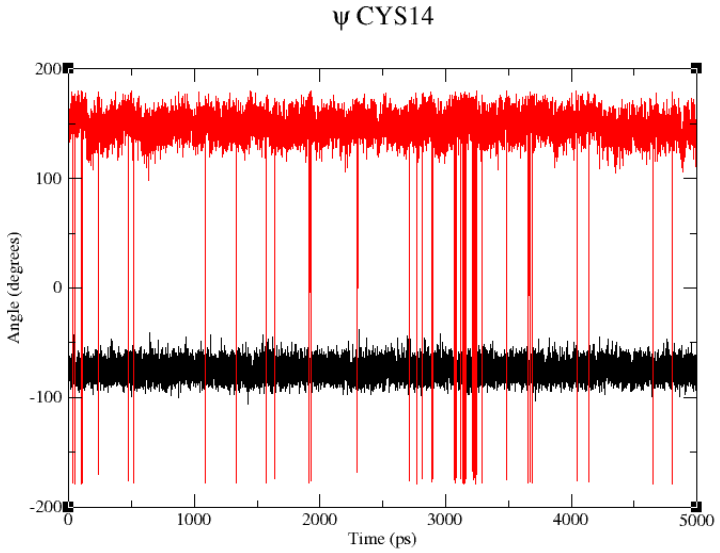


Figure 6: Variation of Backbone dihedral angle for residues Pro-13 phi and Cys-14 psi for simulation-A.

Result 6: Positional rmsf of alpha Carbon atom -

From our previous knowledge, we stated that we get three such conformers for each simulation. Further, these conformers are used for comparison of properties like backbone dihedral angles, rmsf of dihedral angles, positional rmsf of alpha carbon atoms, and positional rmsf of gamma carbon atoms with respect to crystal structure.

In the following fig 7a, we tried to compare rmsf values of 3 conformers for simulation A with respect to its crystal structure. It can be observed that conformer A1 is more close to experimental values, this is expected as conformers A2 and A3 will be different as we are selecting them from different time stamps. Similarly, we get such results for simulation B (with restraints) i.e. conformer B1 is more close to experimental values than B2 and B3.

But, as we are applying atom-atom distance restraints we expect the B1 to be more close to experimental values than A1. I think as we have used different force fields and different parameters while performing simulation we got varied results here.

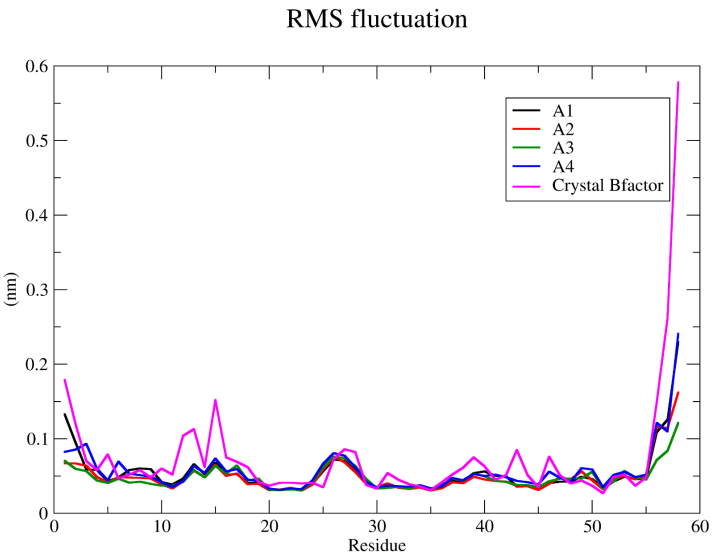


Figure 7a

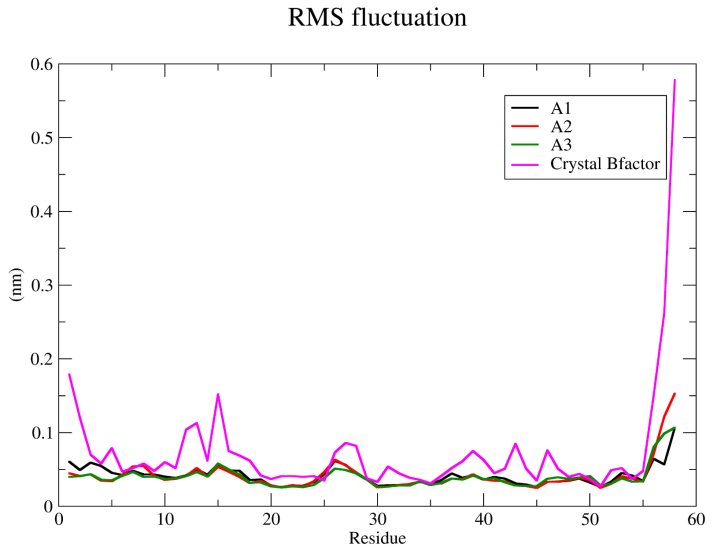
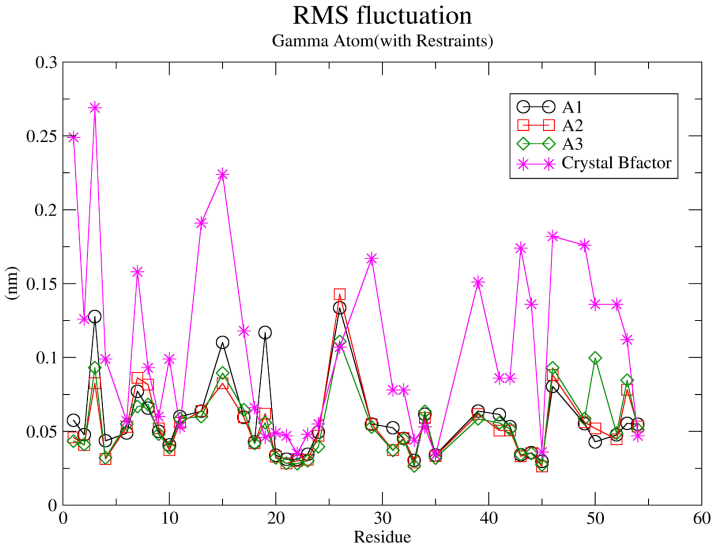


Figure 7b

Result 7: Positional rmsf of gamma-carbon atoms -

Similar to the result in 6, we can observe that A1 and B1 are much closer to experimental results. Again we didn't get the B1 close to the experimental structure than A1.



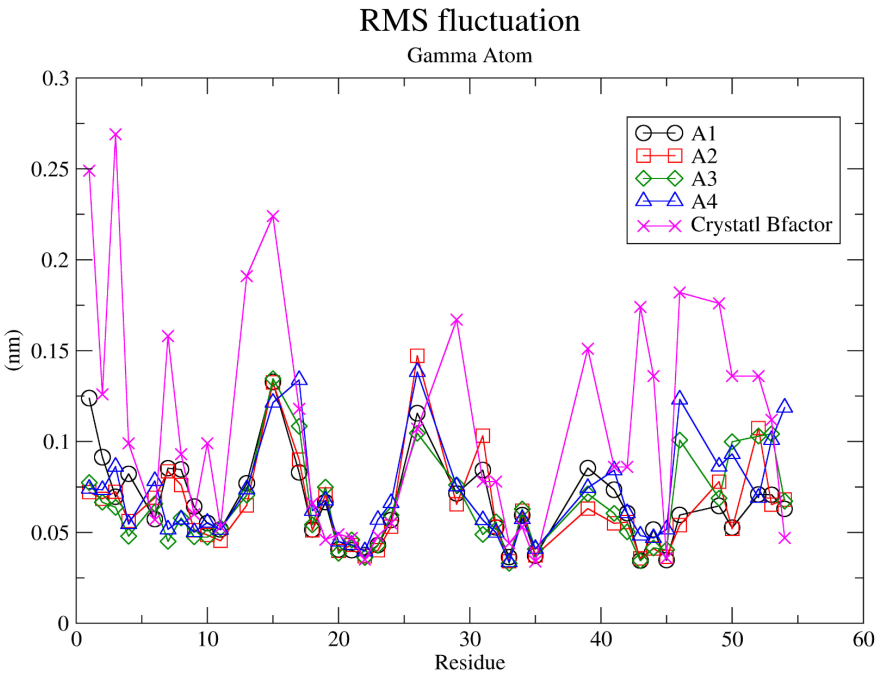


Figure 8a(above) and 8b(below)

Calculations -

- The following formula was used to calculate rmsf of alpha and gamma carbon atoms of the experimental structure. Please find the attached python code file for the same.

$$\langle(\Delta r)^2\rangle^{1/2} = [3B/(8\pi^2)]^{1/2}$$

Conclusions -

- Overall we get the simulation B results closer to the experimental values than from simulation A.
- In literature, they have performed the simulation of 1.4ns and 0.8ns. But as we have performed it for 5ns, we observe the stability of properties like the radius of gyration, coulomb energy (protein-protein), rmsd, and LJ potentials after some period of time. For example, It took 3.5ns for rmsd values to stabilize.
- While studying the dynamic behavior of protein we observed that backbone and side chains undergo transitions such that it forms new conformation. These regions where such transition occurs can be further studied to test their functional importance based on their high flexibility.
- It is observed that residues present at the end of secondary structures show high flexibility along with regions which flexible regions discussed in previous points. However it doesn't affect the overall structure and properties of the protein, but these variations can further be examined.
- Application of NOE derived atom-atom distance restraint data improved the quality of trajectory, this is shown using rmsf (C-alpha & C-gamma atoms) and rmsd values.

References -

- Kurt D. Berndt, Peter Güntert, Leonard P.M. Orbons, Kurt Wüthrich, Determination of a high-quality nuclear magnetic resonance solution structure of the bovine pancreatic trypsin inhibitor and comparison with three crystal structures, Journal of Molecular Biology.
- Brunne RM, Berndt KD, Güntert P, Wüthrich K, van Gunsteren WF. Structure and internal dynamics of the bovine pancreatic trypsin inhibitor in aqueous solution from long-time molecular dynamics simulations. Proteins.
- https://www.ks.uiuc.edu/Training/CaseStudies/pdfs/bpti.pdf