ATOMIC-LEVEL CHARACTERIZATION OF THE HUMAN IMMUNODEFICIENCY VIRUS (HIV) TYPE 1 ENVELOPE GLYCOPROTEINS USING MOLECULAR DYNAMIC SIMULATIONS

Stephanie Sauve[1], Dr. Maolin Lu[2], Dr. Mahmoud Moradi[1]
Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR, USA.[1]
Cellular and Molecular Biology, The University of Texas at Tyler, Tyler, Tx, USA.[2]

INTRODUCTION

- According to the World Health Organization human immunodeficiency virus (HIV), a virus that attacks the human body's immune system, affects more than 38 million people worldwide as of 2021.
- HIV type 1 (HIV-1), enters host cells after the envelope (Env) glycoprotein trimer [$(gp120/41)_3$], binds the CD4 receptor and a coreceptor, either CCR5 or CXCR4.
- Binding induces conformational changes that allow mature (cleaved) Env to transition from a closed pre-triggered state (state-1) to open states that are needed for the binding and fusion processes (state-2 known as the transition state and state-3 which is the bound state). However, the closed conformational state of the Env trimer is not yet known.

1. The HIV-1 Env trimer has three conformational states

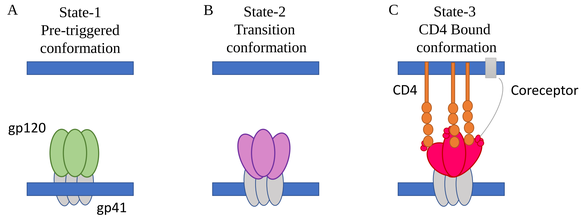

Fig 1. Schematic showing the three conformational states of the HIV-1 Env trimer previously determined through smFRET. The HIV-1 Env trimer transitions from state-1 to state-2 to state-3. A) State-1 represents the closed pre-triggered conformation. B) State-2 represents the open transitional conformation. C) State-3 represents the open CD4 and coreceptor bound conformation.

- This study aims to characterize local and global conformational changes of the closed and open states of the Env trimer and to determine the dynamic behavior of the glycoprotein trimer through a combination of equilibrium and nonequilibrium molecular dynamic (MD) simulations.
- The MD simulations were done using previously determined uncleaved structures in state-2. The structures used (PDB IDs: 7n6u (state U1) and 7n6w (state U2))[1] were previously characterized through electron microscopy (EM) and are able occupy all three states of the mature Env determined previously by single molecule fluorescence resonance energy transfer (smFRET)[2]. Compared to state U1, state U2 has unequal distances between the protomers when aligned and more conformational variation leading to greater asymmetry overall in the structure.
- Steered molecular dynamics (SMD) will be used to characterize the closed pre-triggered conformational state of HIV-1.

2. State U2 has greater degree of asymmetry than state U1 regarding rotation of the protomers respective to the trimer axis despite both structures being in the state-2 conformation.[3]

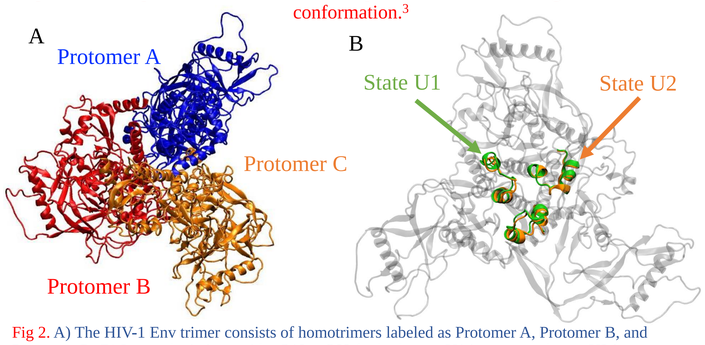

Fig 2. A) The HIV-1 Env trimer consists of homotrimers labeled as Protomer A, Protomer B, and Protomer C. B) Image depicting resides 546 to 568 for both state U1 in green and state U2 in orange that comprise the interprotomer interface.

REFERENCES

1. Zhang, S., Wang, K., Wang, W. et al. Asymmetric structure and conformational plasticity of the uncleaved full-length human immunodeficiency virus envelope glycoprotein trimer. J. Virol 95(24), e00529-21 (2021) doi:10.1128/JVI.00529-21
2. Lu, M., Ma, X., Castillo-Menendez, L. et al. Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET. Nature 568, 415-419 (2019) doi:10.1038/s41586-019-1101-y
3. Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", J. Molec. Graphics 14, 33-38 (1996)
4. Eswar, N., Eramian, D., Webb, B. et al. Structural Proteomics. Methods Mol Biol 426, 145-159 (2008)
5. Jo, S., Kim, T., Iyer, V.G. et al. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. J. Comp. Chem. 29, 1859-1865 (2008)
6. Stanzione, D., West, J., Evans, R. et al. Frontera: the evolution of leadership computing at the national science foundation. In Practice and Experience in Advanced Research Computing (PEARC '20), 106-111 (2020) doi:10.1145/3311790.3396656

METHODS

Our simulations were based on electron microscopy structures of the Env trimer in the transitional conformation (state-2) (PDB ID: 7n6u and 7n6w)[1]. Missing residues were generated using Modeller.[4] The CHARMM-GUI webserver was used to prepare the proteins for MD simulations.[5] The proteins were solvated in a box of TIP3P waters and were neutralized by Na^+ and Cl^- ions with a salt concentration of 0.15 M.[5] All simulations were performed using the NAMD 2.13 simulation package with the CHARMM36m all-atom additive force field.[5] Approximate system sizes were 357,363 and 351,477 atoms for state U1 and state U2 respectively. The input files for energy minimization and production were generated using CHARMM-GUI.[5] MD simulations were energy-minimized and relaxed for 10,000 steps using the standard CHARMM-GUI protocol.[5] The initial relaxation was performed in an NVT ensemble while the rest of the simulations were performed in an NPT ensemble. Simulations were carried out using a 2-fs time step at 310 K using a Langevin integrator with a damping coefficient of 1.0 ps-1. The pressure was maintained at 1 atm using the Nose-Hoover Langevin piston method. The production runs were run for 500ns thus far. The simulations were repeated two additional times for 500 nanoseconds each for both systems (n=3). All simulations were run on TACC Frontera.[6] For smFRET data collection, CY3 and CY5 dyes were used as the donor and acceptor dyes respectively. The dyes were attached at residues 135 and 533 on all protomers. The distance between the dyes was converted to FRET efficiency using the FRET efficiency equation (efficiency = $R_0^6 / (R_0^6 + r^6)$ where r is the distance between donor and acceptor chromophores and R_0 is the distance with a 50% energy transfer efficiency. Then, FRET efficiency was converted to probability out of 100. Here we have the R_0 value as 82.53 which is the average distance of all center of distances of the U1 state crystal structure.

RESULTS

3. RMSD plots show that state U2 becomes stable faster than state U1.

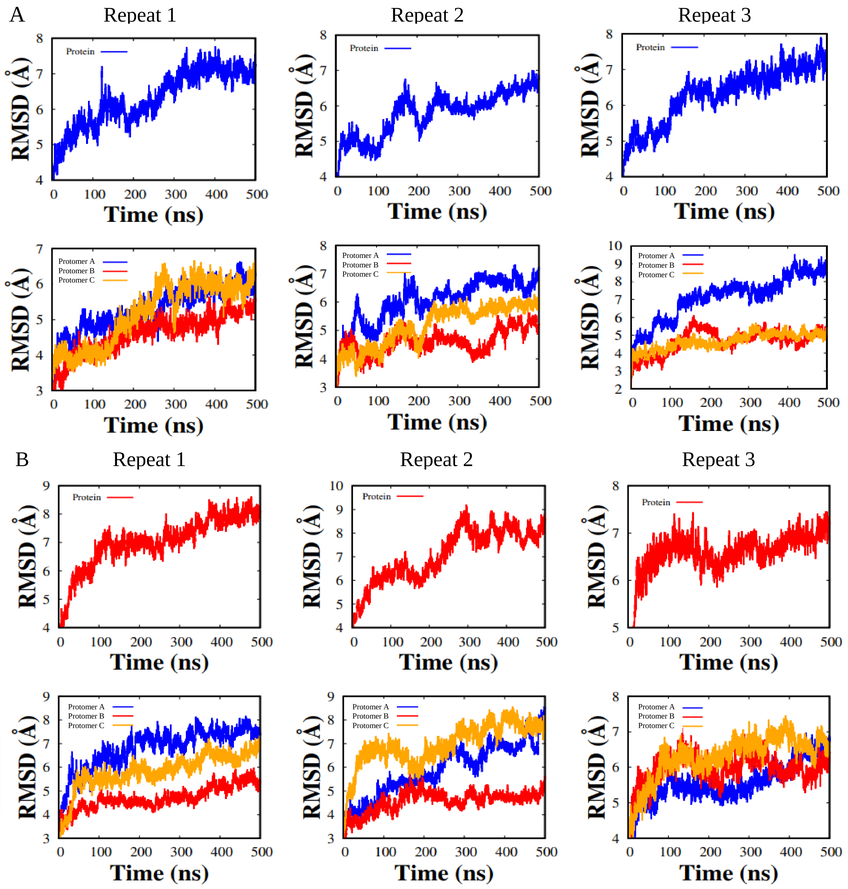

Fig 3. RMSD plots of both systems and their replicates U1 (A) and U2 (B) than have been simulated for 500ns with a plot of the whole protein on top and a plot of the individual protomers on the below. A) The RMSD appears to plateau at 400ns around 7 Å for Repeat 1 but has yet to plateau for Repeat 2 or Repeat 3 at 500ns for the overall protein. B) The RMSD appears to plateau at 400ns for all three replicates at 8 Å for Repeat 1 and Repeat 2 and at 7 Å for Repeat 3 for the overall protein. Overall, this shows that system U2 is becoming stable faster than system U1.

RESULTS CONTINUED

4. Principal Component Analysis (PCA) showing that despite being in the transitional open state, the systems behave differently.

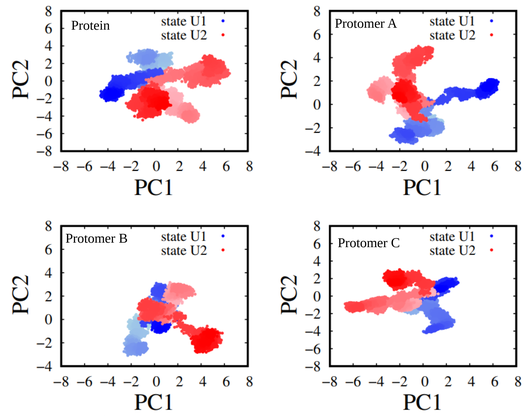

Fig 4. PCA plots of the whole Protein (top left), Protomer A (top right), Protomer B (bottom left), and Protomer C (bottom right) comparing states U1 and U2. The results show that despite being in the same transitional state-2 conformation, the systems behave differently.

5. State U1 and state U2 tend to have a high FRET with an R_0 value of 82.53. However, state U2 is more consistent.

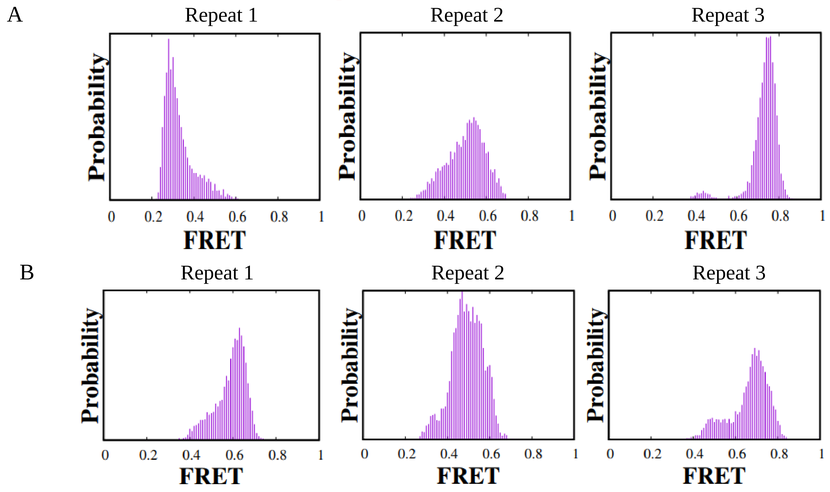

Fig 5. Probability versus FRET plots for A) system U1 with replicates and B) system U2 with replicates having an R_0 value of 82.53. A high FRET state is maintained for all plots except for U1 repeat 1. However, system U2 remains centered more consistently around 0.65 FRET.

CONCLUSION

- We have determined that despite both systems (state U1 for 7n6u and state U2 for 7n6w) being in state-2 conformation, there are differences in their stability and behavior.
- The RMSD for state U2 because stable faster than state U1 and in the smFRET data state U1 has a wider range of data while state U2 is centered around 0.65 FRET with an R_0 value of 83.
- The PCA shows that the two systems are behaving differently.
- In the future, steered molecular dynamic simulations need to be run to push the open conformation of state-2 into a closed state-1 conformation. After this state is obtained, we can compare the two states and the transition pathway.

Acknowledgements

Research reported in this work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award numbers R15GM139140 and R35GM147423. This research is also supported by the National Science Foundation grant CHE 1945465 and the Arkansas Biosciences Institute. This research is part of the Frontera computing project at the Texas Advanced Computing Center, made possible by National Science Foundation award OAC-1818253. This research is also supported by the Arkansas High Performance Computing Center which is funded through multiple National Science Foundation grants and the Arkansas Economic Development Commission.