Sampling the Conformationional Space of the Parkinson's Disease Associated Protein Alpha-Synuclein

Michael Dean[a,b,c], **Aram Davtyan**[c], **Peter G. Wolynes**[c]

a. Department of Biology and Biochemistry, University of Houston b. Houston Community College

c. Center for Theoretical Biological Physics, Rice University

Alpha-synuclein and Parkinson's Disease

Aggregation of the protein Alpha-synuclein into amyloid fibrils is linked to the neurodegenerative condition known as Parkinson's Disease. This aggregation pathway is not fully known, and the intrinsically disordered (IDP) nature of the protein adds to its complexity. The present study aims to explore the conformation space of an Alpha-synuclein tetramer using a series of replica exchange molecular dynamics simulations.

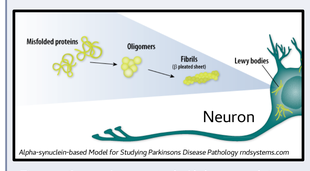


Figure 1 *- Aggregated Alpha-synuclein is found in Neuron inclusions known as "Lewy Bodies" in patients with Parkinson's Disease*

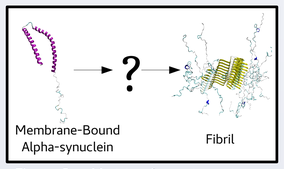


Figure 2 *– Models of membrane-bound (1XQ8) and fibrillic (2N0A) Alpha-synuclein from the Protein Data Bank. Pre-fibrillic structures are not well known.*

Molecular Dynamics Simulations

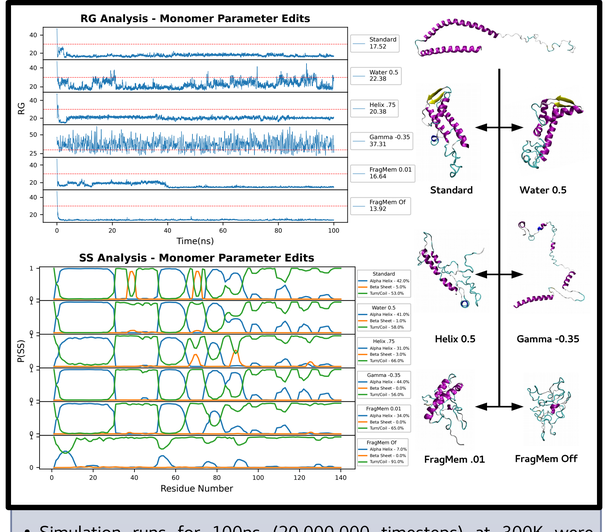
- Alpha-synuclein was investigated using Molecular Dynamics (MD) simulations, where the dynamics of the system is modeled based on forces calculated from a chosen energy function (force field).
- Simulations were performed using the LAMMPS molecular dynamics code **[1]** with the Associative Memory, Water Mediated, Structure and Energy Model (AWSEM-MD) coarse-grained force field **[2].**
- Reductions to term weights for the AWSEM-MD potential energy function (***Figure 3***) were tested in monomer simulations, as the standard settings may not be ideal for IDP simulation.

$$V_{total} = V_{backbone} + V_{contact} + V_{burial} + V_{HBBeta} + V_{HBHelix} + V_{FM}$$

$$V_{backbone} = V_{con} + V_{chain} + V_{\chi} + V_{rama} + V_{excl} \qquad V_{FM} = -\lambda_{FM} \sum_{m} \sum_{ij} \exp\left[-\frac{(r_{ij} - r_{ij}^{m})^2}{2\sigma_{ij}^2}\right]$$

Figure 3 *- Potential energy function for the AWSEM-MD force field. Each term describes different aspects of protein dynamics. V_{FM} is a bioinformatic term that pulls experimental information on a given 'fragment' from the Protein Data Bank. r_{ij} is the distance between $C_{a/b}$ atoms in the simulation and r_{ij}^{m} is the corresponding distance in the memory.*

Monomer Parameter Reduction Simulations



- Simulation runs for 100ns (20,000,000 timesteps) at 300K were analyzed for radius of gyration (RG) and secondary structure (SS).
- Reduction of Water(0.5), Helix(0.5), and Fragment Memory(0.1) terms reduced helical propensity, increased beta sheet propensity in the region around residue 90, and produced fluctuation in RG.

Replica Exchange Simulation Method

- Multiple parallel simulations are run at different temperatures and conformations are exchanged periodically between simulations. (***Figure 4***)
- Lower temperature simulations are given access to conformations that would have been prevented by an energy barrier. (***Figure 5***)
- A more complete sampling of the conformation space is achieved.

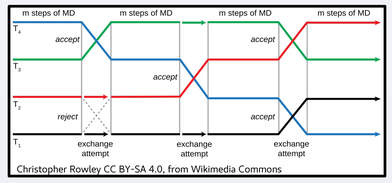


Figure 4 *– Diagram describing the exchange of conformations between replicas at different temperatures throughout a molecular dynamics simulation. Exchanges are either accepted or rejected based on Boltzmann-weigthed Metropolis criterion.*

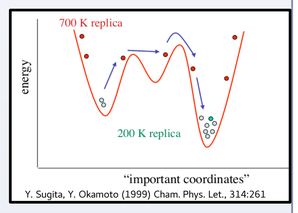


Figure 5 *– Lower energy replicas are allowed access to states previously prohibited by energy barriers.*

Tetramer Simulation Setup and Analysis Methods

• Replica Exchange Simulation Setup

- Four Alpha-synuclein chains at 1mM concentration
- Periodic Simulation Box
- 24 replicas ranging 280-470K
- 10 ns simulation time (2,000,000 timesteps)

• Simulation Groups

- **Group 1** – Standard AWSEM-MD Settings
- **Group 2** – Reduced Parameter Settings (Helix 0.5, Water 0.5, Fragment Memory 0.1)
- **Group 3** – E46K G51D Mutant with Reduced Parameters

• Analysis Methods

- Results from 300K conformations are combined into a discontinuous trajectory.
- RG and Average SS analysis is performed on the trajectory.
- 200 trajectory 'frames' are chosen and clustered. (***Figure 6***)
- Representative conformations are chosen from each cluster.

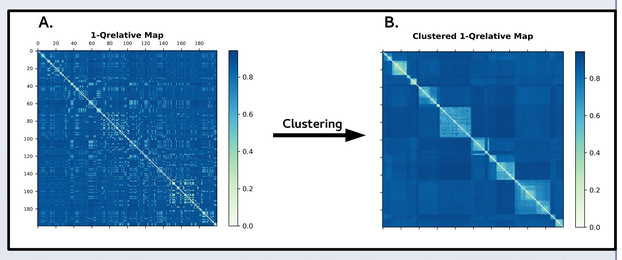


Figure 6 *– Clustering method applied to the combined 300K data from Simulation Group 1*

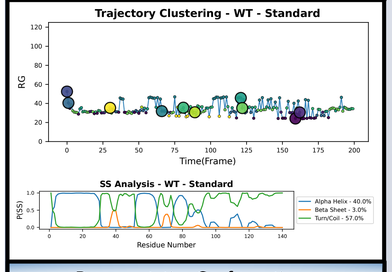
A *– 200 frames are chosen from the trajectory and compared using 1-Qrelative as a distance measure, where Qrelative is the fraction of contacts that are the same between frames. The 1-Qrelative value ranges from 0 (most similar) to 1 (least similar).*

B *- The trajectory frames are then clustered into similar groups using a hierarchical clustering algorithm and the 1-Qrelative similarity matrix.*

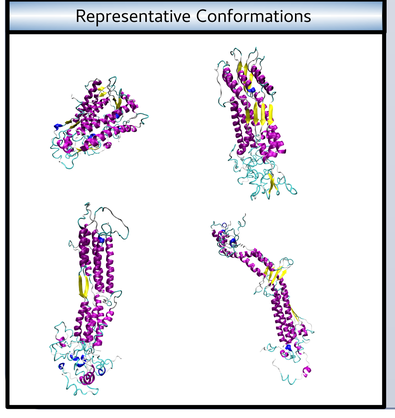
Simulation Group Results and Representative Structures

Group 1 – Standard AWSEM-MD

RG / Average SS Analysis

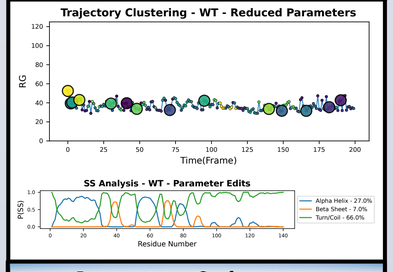


Representative Conformations

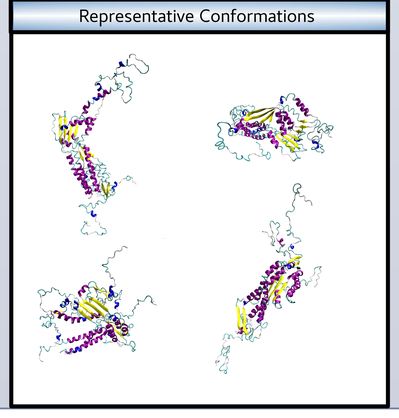


Group 2 – Reduced Parameters

RG / Average SS Analysis

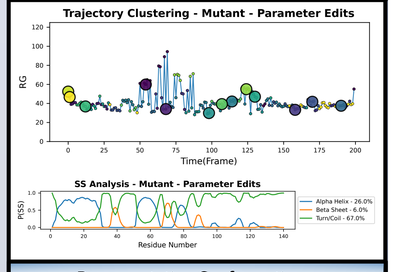


Representative Conformations

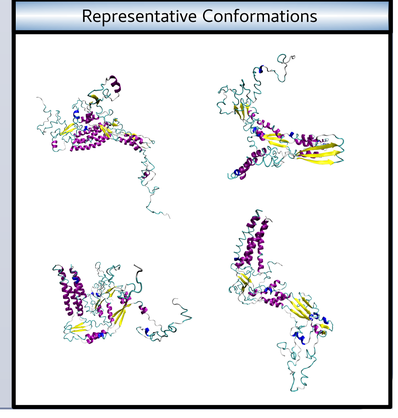


Group 3 – Mutant / Reduced

RG / Average SS Analysis



Representative Conformations



Simulation Group Comparison

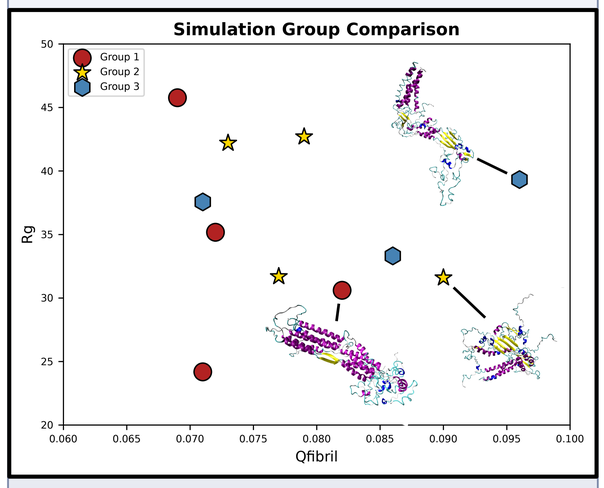


Figure 7 *- Representative configurations from the clusters of each simulation group are plotted along a Qfibril coordinate, which is the fraction of contacts that are the same between the structure and four chains from the fibril model (2N0A).*

Discussion and Future Work

• Observations

- The helical region around residue 55 was not affected by parameter reductions, which is interestingly were 4/5 of familial PD point mutations occur.
- Shifting of gamma levels resulted in no interaction between protein chains and was dropped from further testing.
- Interactions are seen between the N and C terminus, resulting in more compact folded tetramer conformations, which may be consistent with experimental reports of 'roughly spherical' tetramers that resist aggregation **[5]**.
- Mutant simulation group (Group 3) produced trimers as well as tetramers, along with the representative structure with greatest Qfibril value. (***Figure 7***)

• Future Work

- Free energy calculation to understand the energy landscape between the obtained conformations
- Longer simulations with increased number of protein chains
- Investigation into phosphorylation of C-terminus
- Collection of experimentally determined data to be used as reference

Acknowledgments

- 2018 FIS and REEMS programs
- NSF DMR-1460564
- NSF PHY-1427654
- NSF ACI-1531814
- XSEDE Comet Cluster sponsored by NSF
- Dr. Margaret Cheung – University of Houston / Rice
- Bartlett M. Sheinberg – Houston Community College
- Dr. Marion Alexander – Houston Community College

References

[1] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, J Comp Phys, 117, 1-19 (1995)

[2] Davtyan, et. al, "AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing", J. Phys. Chem. B 2012 116

[3] Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics", J. Molec. Graphics, 1996, vol. 14, pp. 33-38.

[4] Frishman,D & Argos,P., Knowledge-based secondary structure assignment. Proteins: structure, function and genetics, 23 (1995)

[5] Bartels, et. Al, "A-Synuclein Occurs Physiologically as a Helically Folded Tetramer That Resists Aggregation." Nature 477.7362 (2011): 107– 110. PMC. 2018.