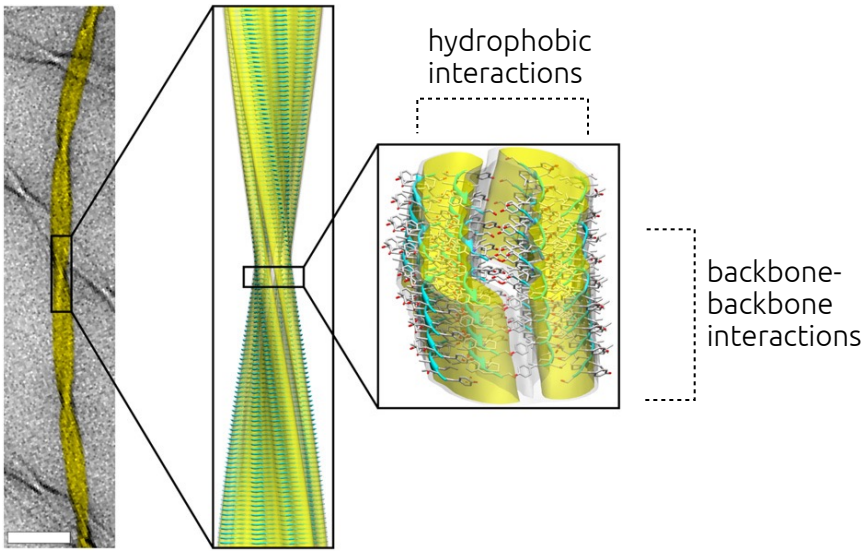
Introduction

Amyloid structures are fibrous protein structures that play a key role in a huge number of debilitating diseases, including Alzheimer's disease, Huntington's disease, Parkinson's disease, type 2 diabetes, and rheumatoid arthritis. Many amyloid structures are formed from the cooperative aggregation of polypeptides, which results in an array of β-sheets running antiparallel to the main fibrils. The β-sheets are stacked on top of each other via backbone-backbone interaction, and hydrophobic residues interact in the middle of the two sheets to create the highly-ordered amyloid fibrils:



Fitzpatrick et al. (2013) Atomic structure and hierarchical assembly of a cross-amyloid fibril. Proc. Natl. Acad. Sci. U.S.A. 110, 5468-5473

Method

Experimentally, it is very difficult to calculate the exact conformations of multimeric amyloid-like systems because of their complex structure. However, new high speed computational techniques allow us to easily simulate an array of peptides to help understand whether certain sequences will aggregate, and if they do, their specific conformations. I used Desmond, a molecular dynamics software package from D.E Shaw Research, to perform these simulations on the following sequences:

Protein	Species	Sequence (AFR)
Als3	*Candida albicans*	IVIVA
Aβ	*Homo sapiens*	LVFFA
N/A	N/A	AAAAA
FLO1	*Saccharomyces cerevisiae*	TVIVA

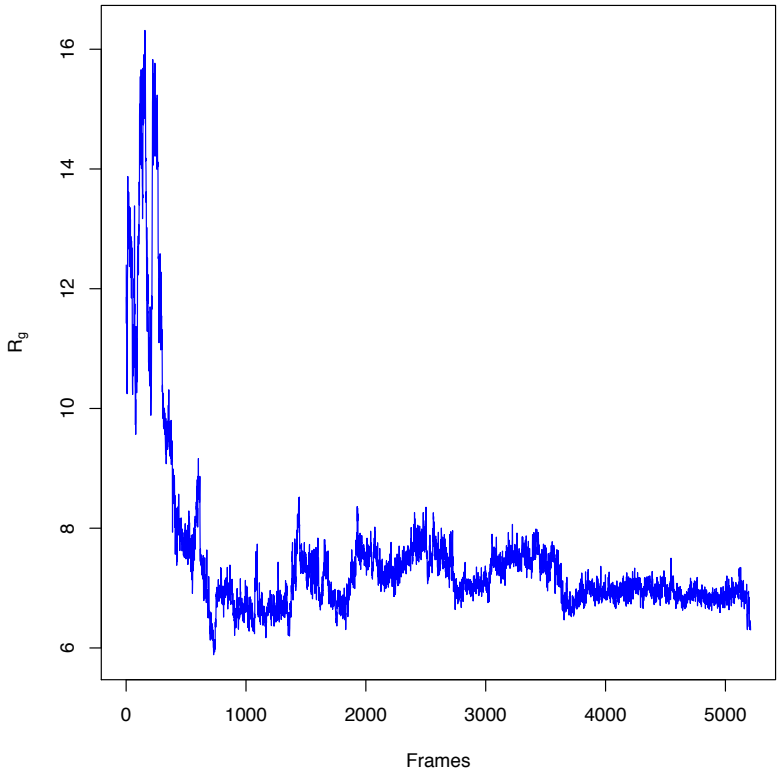
Once complete, I wanted to analyze the degree of amyloid-like association for each mutation. After reading through some of the literature on structure and computational analysis of

amyloid formation, I couldn't find any computational techniques employable in an entire trajectory. Thus, the only option was to create a tool myself. First, I needed to find an objective metric to measure the degree of association. We used a measurement called the radius of gyration (R_g), which measures an average distance between a set of coordinates and their mean coordinate (or center of mass) in a given structure, calculated by this equation:

$$R_g = \sqrt{\frac{\sum_{k=1}^{N} \left(r_k - r_{mean} \right)^2}{N}}$$

Variable	Definition
R_g	radius of gyration
N	coordinates
r_{mean}	mean coordinate

This is a very good metric to use because amyloidosis requires closely aggregated peptides in a relatively unchanging structure. By taking the R_g of each structure individually, we can graph each data to analyze the change in R_g throughout the trajectory:



With a graphical representation, it is easy to see the changes in R_g, which tell us that in this case, the dimer system associates within the first 6 nanoseconds of the simulation, and does not dissociate within the rest of the 50 nanosecond simulation.

Computation

I then wanted to use this function to automate the process of determining amyloid-like association in Maestro. To do this, I first wrote three classes in Java: class 'Point' that stores

the data for a single coordinate in a structure and includes a function that compares the distance between two points, class 'ROG' that computes R_g values by reading a pdb file (exported from the Trajectory tool in Maestro), and class 'MultiROG' that initializes objects and starts the calculation process. Then, I implemented these classes into a user-friendly Bash shell, that receives user input for the file to read from and the range of residues to calculate R_g for, computes R_g for each structure, and finally writes the results to a text file, which can be used to create graphs using software such as R.

Discussion

Although my technique is a step towards quantitative analysis of amyloid aggregation, there is still a lot of work to be done in terms of subjectively assessing whether an amyloid has formed. This would require advanced data analysis in terms of the specific conformation of the secondary structure elements, and the resistance of the structure to conformational changes, proving a resilient, sustained amyloid-like fibril structure. Additionally, tools are required to quantitatively analyze the interactions between different secondary structure elements, like the inter-β-sheet backbone-backbone interaction and hydrophobic interaction between peptides.