Use of Aggregation Prediction in Protein Formulations for Excipient Design

Brock C. Roughton[1], Anthony I. Pokphanh[1], T. Steele Reynolds[1], J. Laurence[2], E.M. Topp[3], and K.V. Camarda[1]

[1] Department of Chemical and Petroleum Engineering, The University of Kansas [2] Department of Pharmaceutical Chemistry, The University of Kansas [3] Department of Industrial and Physical Pharmacy, Purdue University

A major concern in the development therapeutic protein formulations is protein aggregation. Proteins can interact to form bound groups of protein molecules or aggregates. Aggregates in protein formulations reduce effectiveness and can lead to severe immune responses in patients. Excipients are additive molecules that are not therapeutically active, but can increase the stability of protein formulations. An ideal excipient binds with aggregation prone regions on the protein to limit interaction of that region with another protein molecule. The goal of this project is to predict aggregation prone regions and design excipients to interact with these regions.

Several tools exist to predict which regions on a protein will be most likely to initiate aggregation. Aggrescan (http://bioinf.uab.es/aggrescan/) and SAP (Spatial Aggregation Potential) were used to predict aggregation prone regions on proteins and the results were compared. Aggrescan uses experimental data to assign each amino acid an aggregation propensity score. An aggregation prone region is identified by a sequence of amino acids with high propensities. The three-dimensional structure is not used in the aggregation prediction. SAP uses molecular simulation to determine regions that are hydrophobic and solvent accessible. Each residue is scored and the results are mapped to the three-dimensional protein structure. A successful prediction tool must use parameters that correlate with aggregation potential for a folded protein.

The aggregation prone regions predicted by Aggrescan and SAP were compared to experimental data on protein aggregation. Proteins with a high number of predicted regions or large predicted regions were found to have higher experimental percent aggregation. With the regions identified, molecular simulations were performed for protein-excipient systems. A protein and small molecule docking algorithm was used to determine which regions of the protein certain excipients interacted with. Trehalose, poly(vinylpyrrolidone), and guanadine hydrochloride were used. For an excipient to successfully stabilize a protein and prevent aggregation, the excipient should interact with the aggregation prone regions predicted by Aggrescan and SAP. The predicted regions were compared to the regions where the excipient docks in the molecular simulation. The simulation results were compared to experimental data on the percent aggregation observed in several protein-excipient formulations. The excipients that were

found to interact with the predicted aggregation prone regions in simulations should also experimentally prohibit aggregation, leading to lower percent aggregation. Hydrogen-deuterium swapping along with FTIR analysis will be performed experimentally to determine exposed regions on the protein. Proteins with a high number of exposed regions are less stable. The exposed regions will be compared to the aggregation prone regions predicted by Aggrescan and SAP.