Haider Tarar

HOMETOWN
Islamabad, Pakistan

MAJORS
Chemical Engineering

ACADEMIC LEVEL
Senior

RESEARCH MENTOR
Kyle Camarda, *Associate Professor of Chemical Engineering*

Q&A

How did you become involved in doing research?
Dr. Camarda was my academic adviser. I asked him about the research he was involved in and what he did sounded interesting, so I decided to start working with him.

How is the research process different from what you expected?
It involves a lot more initiative from your end. There are no set directions or standards for what you are doing and you have to find your way forward.

What is your favorite part of doing research?
It is satisfying when experimental results match your hypothesis. Also, you get to see the application of a lot of theory.

Optimum protein-excipient interactions using molecular docking simulations

Haider Sulaiman Tarar

THE PROBLEM

Protein drugs have a tendency to aggregate, which adversely affects their shelf life and delivery as a stable drug formulation. Of the top 100 drugs by U.S. sales in the fourth quarter of 2012, 28 were protein drugs or other biologics (source: IMS Health, http://www.imshealth.com). More than 100 genuine, and similar numbers of modified, therapeutic proteins are approved for clinical use in the European Union and the USA, with 2010 sales of 108 billion (US$) (Dimitrov, 2012). Thus, there is a high demand for protein drugs, but any positive consequences of protein drugs are inaccessible if these drugs cannot be stored and delivered in a stable form. Therapeutic effects of the protein drugs are dependent on the specific primary structure of the proteins. During Aggregation, the protein irreversibly loses its unique shape, and thus its function. The aggregation process can occur due to physical interactions between protein surfaces or can result from chemical interactions of the amino acids, forming covalent bonds between proteins (Wang, 2005).

Lyophilization is the primary method used in the industry to improve the stability of protein drug formulations. Lyophilization removes water from the formulation by sublimation via a freezing step and then by evaporation through primary and secondary drying steps (Cleland, Langer et al. 1994; Costantino and Pikal 2004). The idea is to improve stability of the protein

by removing water to decrease its mobility and make it less vulnerable to water-based reactions. Even after this process, proteins may lose their native structure due to aggregation, leaving lyophilization alone as a dubious solution.

PROPOSED SOLUTION

The aim of this project is to find an effective, cheap and efficient method to develop stable formulations for protein drugs which are not vulnerable to aggregation and in turn have stable shelf life and effective delivery. A method using a combination of computer-aided molecular design and molecular docking simulations is proposed to find a specific optimum formulation for a specific protein, minimizing aggregation probability.

The choice of excipient in the protein drug formulation can lower the probability of protein aggregation through interactions between the excipient and hot spots on the protein. Excipients are inactive ingredients in the drug product that help hold a dose of the active pharmaceutical ingredient (API) together and keep it stable for a long shelf life (Ritter, 2008). Hot spots are aggregation-prone regions on a protein. In this project, a pool of candidate sugar and surfactant excipients is chosen based on specific physical and chemical properties. For each candidate excipient, molecular docking simulation is conducted with the protein. The interactions of the excipients with the hot spots on the protein are fed into Computer Aided Molecular Design (CAMD) which determines the optimum excipient formulation for that specific protein.

METHOD

The pool of excipient candidates are chosen based on their physical and chemical properties. An excipient having a high glass transition temperature is preferred. In the case of sugar excipients, glass transition temperature refers to phase change. Higher glass transition temperature means less mobility and more stability. Another aspect looked at is the water content in a lyophilized excipient-protein solid. Lower water content is preferred because removal of water slows down water-based reactions and also reduces protein mobility. Furthermore, the lower the protein loss after lyophilization, the better the excipient-protein combination will function. Loss of monomeric protein may indicate protein aggregation.

The next step is to model the interactions of this pool of excipient candidates with the protein. Molecular docking simulations of each protein-excipient combination are done in Autodock. The excipients looked at in this work are designed to provide stabilizing interactions for the protein myoglobin. In order to apply molecular docking, two commonly used excipients, the sugars sucrose and mannitol, have been investigated so far. Some of the surfactants which will be looked at as excipients in the future include Octyl glucoside, Tween 40 and Tween 80. Molecular docking simulations performed using Autodock predicted which residues were most likely to interact with the excipient in question. AutoDock employs a grid-based approach along with a semi-empirical free energy force field to identify interaction sites with minimal free energy.

Each simulation in Autodock provided ten docking conformations, which represent a local minimum free energy state. Ten docking simulations were performed for each excipient with myoglobin, providing a hundred docking conformations total for each excipient. Each conformation was evaluated to determine residues which were repeatedly involved in strong interactions (Van der Waal or hydrogen bonding) with the excipients.

The computational results from the above procedure were compared to experimental hydrogen-deuterium exchange (HDX) mass spectrometry experiments using lyophilized formulations of myoglobin with either sucrose or mannitol (Sophocleous, et al., 2012). The regions on the protein which exhibit free exchange between hydrogen and deuterium are uncovered and not protected by the protein-excipient interactions. Therefore, protein regions which show low deuterium uptake are protected by the excipient interactions covering those regions. These regions of low deuterium uptake are then correlated with protein-excipient interactions predicted by molecular docking simulations, to validate the simulation results.

Figure 1 below shows the concept by illustrating a 'perfect match':

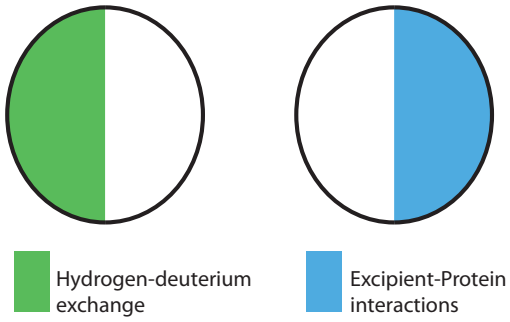

Fig. 1

RESULTS

Results were obtained from the execution of molecular docking simulations, providing 100 conformations for both myoglobin-mannitol and myoglobin-sucrose. The results were analyzed and compared with the results of HDX exchange experiments for the same

two pairs in the lyophilized state. Figure 2 shows the results for the HDX exchange experiments for those two excipient-protein pairs. Figure 2 is a histogram which presents the strength of molecular interactions at various residue locations, for both the simulation results and the HDX experimental results. Overall, the docking results compared favorably with the HDX results. Regions of high interaction in the docking simulations corresponded to regions that were protected from hydrogen-deuterium exchange. The regions which were protected by the excipients in the lyophilized state showed low deuterium uptake. These regions correspond to residue-excipient pairs, with sucrose and mannitol showing high frequency of interactions. Thus, the simulation predictions demonstrate utility as a formulation selection tool. The HDX experimental results show that the sucrose-myoglobin system always shows a lower deuterium uptake then the mannitol-myoglobin combination. For the simulations, the excipient (sucrose or mannitol) with higher interaction with a specific amino acid can be selected. An approximation made by the Autodock simulation algorithm is that the protein molecule remains rigid. While this simplification certainly suggests that the results of simulations alone should not guide formulation design, the use of Autodock in combination with the other software tools for screening of excipient candidates is not hampered.

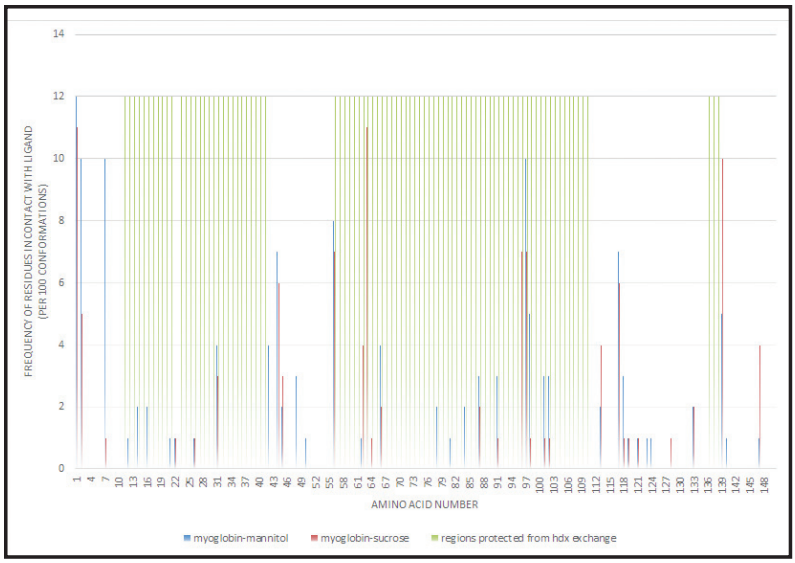

Fig. 2: Frequency of amino acid residues of Myoglobin in contact with sucrose and mannitol. Results are given for 100 docking conformations. The completely shaded regions in green indicate the amino acids that were protected by the excipient in the lyophilized state, as determined by HDX experiments.

Data provided by the docking simulations is used to optimally select excipients that have the highest number of interactions with aggregation-prone"hot spots."The hot spots were predicted using Aggrescan (Conchillo-Sole et al., 2007). For each hot spot, a sugar and a surfactant molecule are chosen to maximize protein-excipient interactions. For selection purposes, whichever excipient has the most interactions with a particular residue is chosen as the best candidate. It is assumed that the molecule with the most interactions with a residue would have the dominant effect and thus contribute the most to the protection of the residue of interest. The method may also be used to compare classes of excipients. For example, at each residue in the hot spot, the best sugar candidate and the best surfactant candidate can be compared. Whichever excipient had the most interactions is then selected and the process is repeated for the entire hot spot sequence. The sugar-surfactant pair providing the maximum number of interactions is selected as the optimal drug formulation. This part of the method aims to utilize Computer Aided Molecular Design (CAMD). Figure 3 summarizes the whole methodology:

A pool of excipient candidates, including surfactants and sugars, is selected based on certain physical and chemical properties.

Molecular docking simulations

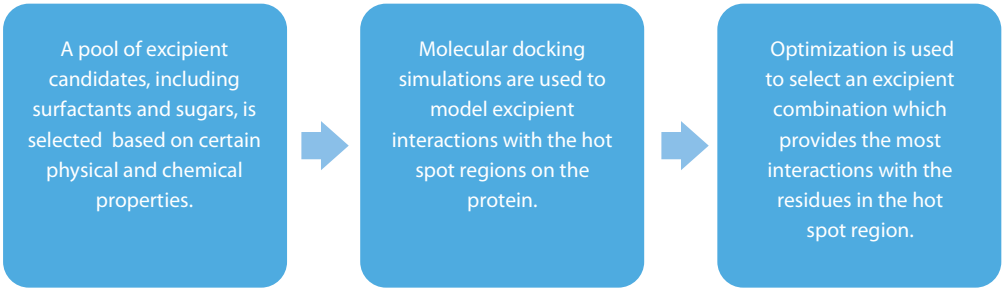

Fig. 3

are used to model excipient interactions with the hot spot regions on the protein.

Optimization is used to select an excipient combination which provides the most interactions with the residues in the hot spot region.

CONCLUSIONS AND FUTURE WORK

The hydrogen-deuterium exchange experiments were performed specifically to establish the utility of molecular docking simulations for the determination of the best excipients for a protein. For example, 25 excipient candidates may be generated for a particular protein using computer aided molecular design, based upon three properties of interest: high glass transition temperatures, low water content in the lyophilized state and minimum protein mass loss on lyophilization. The molecular docking simulation method can then be used to quickly screen this group into a short, ranked list of excipients for that protein. Formulations of different excipients which give optimal interactions with the hotspot regions of the proteins can also be determined. This procedure significantly decreases the time and resources expended in the search for novel excipients. Thus, this work aids in the ability to predict the optimal excipient for any provided protein, allowing the pharmaceutical industry to better develop safe, effective and financially viable protein drugs.

An approximation made by the Autodock simulation algorithm is that the protein molecule remains rigid. Molecular docking simulations will be run in AutoDock Vina to see if that minimizes this approximation. Furthermore, additional proteins and excipients will be added in the dataset to obtain a more detailed and flexible tool for protein formulation design.

ACKNOWLEDGMENTS

Professor Kyle V. Camarda, Chemical and Petroleum Engineering, University of Kansas

Brock C. Roughton, Phd Chemical Engineering, University of Kansas, 2013.

Molecular Graphics and Modeling Lab, Molecular Structures Group, University of Kansas.

APPENDIX

The structure of the protein was obtained in the PDB file format from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). Structural information about the excipients was accessed via files in 3D-SDF format obtained from PubChem online (http://pubchem.ncbi.nlm.nih.gov/). Because Autodock cannot read sdf files, excipient files were converted to 3D-mol format using the CACTUS online toolkit (http://cactus.nci.nih.gov/translate/).

Docking Using ADT (Pokphanh, 2011).

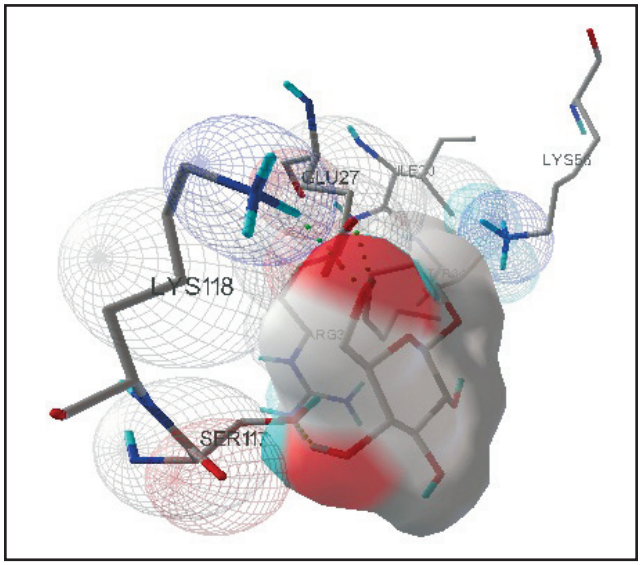

Fig. 4: A docking conformation of Octyl Glucoside with Myoglobin in Autodock.

References

Dimitrov Dimiter S., 2012, Therapeutic proteins, Methods Molecular Biology, 899:1-26.

Cleland, J. L., R. S. Langer and American Chemical Society. Division of Biochemical Technology. (1994). Formulation and delivery of proteins and peptides. Washington, DC, American Chemical Society.

Conchillo-Sole, O., de Groot, N., Aviles, F., Vendrell, J., Daura, X. and Ventura, S., 2007, AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics, 8, 1, 65.

Costantino, H. R. and M. J. Pikal (2004). Lyophilization of biopharmaceuticals. Arlington, VA, AAPS Press.

Pokphanh, Anthony, 2011, CPE 651 Research Report, 9-13.

Ritter, Stephen K., 2008, Excipients, ACS Publications, 86, 1, 25.

Roughton, Brock C., 2013, Development of Computer-Aided Molecular Design Methods for Bioengineering Applications, Dissertation, Department of Bioengineering, The University of Kansas.

Sophocleous, Andreas M., Zhang, Jun., and Topp, Elizabeth M., (2012), Localized Hydration in Lyophilized Myoglobin by Hydrogen-Deuterium Exchange Mass Spectrometry. 1. Exchange Mapping.

Wang Wei, 2005, Protein aggregation and its inhibition in biopharmaceutics, International Journal of Pharmaceutics, 289, 1-2, 1-30.