Case Study for CS 379a *Julia Salas* 3-22-06

McGovern, S.L.; Shoichet, B.K. Information Decay in Molecular Docking Screens against Holo, Apo, and Modeled Conformations of Enzymes. *J. Med. Chem.* **2003,** *46*, 2895-2907.

Introduction

What is the influence of protein structure on docking success? How good does a receptor structure have to be for successful docking? These two questions lie at the heart of McGovern and Shoichet's paper. Due to cost concerns, docking is often accomplished by modeling the target (protein) site as rigid. As such, when performing a docking experiment, researchers must select a single protein structure to screen an entire library of small molecules. For well-known drug targets, many crystal structures are often available including both holo (target bound to a small molecule) and apo (target in the absence of a binder) forms. For most sequenced proteins, however, no crystal structures are available and thus homology modeling is used to determine a modeled structure.

Methods

For this comparative study, researches used the program DOCK3.5 (accommodating ligand flexibility) to dock the MDL Drug Data Report Database into target protein structures for ten proteins that are common drug targets. Known ligands composed about 0.03%-1% of the total 95,000 compound database. For each protein, the database was docked into a holo and an apo structure as well as a homology modeled structure of a related but non-structurally characterized protein. Enrichment factors were measured and compared to probe the effect of each starting structure.

Results

In general, and as expected, success of docking was best when using the holo structure of the protein followed by the apo structure and lastly the modeled target structure. There were some notable exceptions, however, which elucidate the advantages and the drawbacks of each of the three starting structure types. For seven of the 10 protein targets, the holo starting structure led to the best enrichment. For 2 targets, Thrombin and Thymidylate Synthase (TS), the apo structure led to the most successful docking and for one target, Purine Nucleoside Phosphorylase (PNP), the starting modeled structure led to the best enrichment. For PNP, however, the holo starting structure resulted in the same enrichment but with a lower concentration of hits compared to the modeled structure. For nearly all starting structures, successful enrichment was observed and for 8 docked, 2 apo, and 3 modeled starting structures, enrichment of at least 20-fold was observed.

Conclusions

This study confirms the theory that in general, holo starting structures lead to the best enrichment in docking. The exceptions highlighted in this paper bring new insight into specific cases where a holo-starting structure is inappropriate, however. For TS and Thrombin, the apo structures performed better than the holo structures. The authors conclude that these two holo structures are overspecialized, that is, they can only accommodate structures where the target molecule binds in a similar mode to the bound-ligand. As such, small molecules with different binding modes are missed. For the cases of Thrombin and TS, the apo structure is more open, allowing for a greater variety of binding modes of the ligands and thus more successful hits. For PNP, the modeled structure performed the best, but its similarity to the holo structure likely accounts for this. In general, modeled structures which were not successful docking targets had poorly placed side chains which hindered docking. Apo structures that were not successful generally had very different geometries than the bound structure. Thus, in general holo structures are the best starting point for a docking experiment unless the structure is known to be overspecialized. In

this case, an apo structure would be more appropriate. Apo structures are not likely to be successful if the protein undergoes a large structural change upon ligand binding. Successful apo starting structures will have "promiscuous" binding sites, that is, the sites will be able to accommodate many different ligand binding confirmations.

Analysis

This paper addressed important issues relevant to rigid docking. A notable study that is lacking here, however, is how enrichment for these three structures (holo, apo, and modeled) docked using rigid docking methods compares to a docking study including protein flexibility. Cost analysis and overall benefit could have been explored to determine the advantages of each type of study.