Ideal Ligand Substituent Prediction Using HINT for the Purpose of De Novo Drug Design

J. Andrew Surface

Hampden-Sydney College / Virginia Commonwealth University

Introduction

Water is a well-known affector of ligand binding in active sites of target proteins. Klebe notes that about two-thirds of all liganded proteins have at least one water molecule mediating the ligand binding phenomenon. Furthermore, as demonstrated by Amadasi *et al.* in a recent publication [2] HINT, a computational protocol that scores Hydropathic Interactions proportionally to free energy [3] based on the LogP coefficient and other geometrical considerations [4], can be used in combination with another validated geometrical protocol, RANK, to accurately identify water molecules that will play a significant role in the binding of the ligand. These significant water molecules can be considered "relevant" in protein active sites and are important in docking and binding considerations. Nevertheless, the water molecules that are replaced by the ligand also have a particular Score and Rank. Given that HINT score and Rank give statistical indication via a double, weighted probability model that the water molecule will or will not play a role in the binding of the ligand, it can also be said that the score and Rank combination of the water molecules also gives indication of which water molecules are *not* relevant. Therefore presumably the Rank and HINT Score provide a hydropathic description of the active site of the protein and give a three-dimensional predictive map of the active site that can be used to provide information about the sort of ligand needed to best fill that active site.

This research proposes to expound and provide a proof-of-concept correlation of water molecule HINT Score with the purposes of providing a computational model and database upon which ideal ligand substituents can be dynamically chosen on the basis of the water HINT scores in protein active sites.

Methods

This project will expound on the data set created in the Summer of 2007 during the research which established a statisically valid water relevance calculation protocol. Using RCSB .pdb files (all of which have been established using X-Ray Crystallography at different resolutions) of the same protein before a ligand binds and after a ligand binds will have its non-relevant water scores in its active site re-evaluated and compiled into a database along with the replacing substituent identification. In addition, carefully mounted and restricted substituent groups will be computationally docked using GOLD [4] and HINT Scores will be calculated. Using simple statistical evaluation, a correlation between Score and specific substituent group will be sought. If there is a correlation, a database of substituents and relatives HINT scores will be built and tested.

Potential Results

Providing a correlation between the HINT Score and the replacing substituent group type is found, a statistically validated method will have been developed to successfully computationally build the outlying substituent groups of a ligand that would ideally bind in that protein active site. The result would be a rudimentary *de novo* "sketch" of a potentially synthesiable ligand that would bind to the given target protein. This would be the first computational layer to the HINT program that would enable it to complete *de novo* ligand prediction.

References:

[1] Klebe, G. (2003). Virtual Ligand Screening: Strategies, Perspectives, and Limitations. *Drug Discovery Today.* **11,** 580-594.

[2] Amadasi, A.; Surface, J.A.; Spyrakis, F.; Cozzini, P.; Mozzarelli, A.; Kellogg, G.E. (2008). Robust Classification of "Relevant" Water Molecules in Putative Protein Binding Sites. *J. Med. Chem.* **51,** 1063-1067.

[3] Spyrakis, F.; Amadasi, A.; Fornabaio, M.; Abraham, D.J.; Mozzarelli, A.; Kellogg, G.E.; Cozzini, P. (2007) The Consequences of Scoring Docked Ligand Conformations Using Free Energy Correlations. *Eur. J. Med. Chem.* **42,** 921-933.

[4] Verdonk, M.L.; Chessari, G.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Nissink, J.W.M.; Taylor, R.D.; Taylor, R. (2005). Modeling Water Molecules in Protein-Ligand Docking Using GOLD. *J. Med. Chem.,* **48,** 6504-6515.