

Understanding the essential requirements for success in structure-based design

Gregory L. Warren

OpenEye Scientific Software, USA

Abstract

We know that for any prediction the quality of the data used to build a model has a direct impact on the quality of the predictions from that model. Developing and validating software for modeling protein-ligand interactions requires protein-ligand structure data; developing and validating high quality software for modeling protein-ligand interactions therefore requires high quality structure data. If predictions from a molecular docking program fail is the failure the result of poorly developed software, using poor data for the prediction, or the use of poor data invalidation? This presentation will discuss methods to measure the suitability of protein-ligand structure data for docking software validation. Using the recently published Iridium data set, it will be shown that a high proportion of structures used for validating docking software in the past are, in fact, entirely unsuitable for this task, but heuristics will be presented that will allow assembly of datasets appropriate for this purpose.

Biography

Gregory Warren received his PhD in Biochemistry from Massachusetts Institute of Technology, Cambridge, MA. He worked as a Postdoctoral Fellow at Yale University in the laboratory of Axel Brunger as part of the development team for the Crystallography & NMR System (CNS) refinement suite. Dr. Warren worked for 8 years as a molecular modeler at GlaxoSmithKline Pharmaceuticals before moving to OpenEye Scientific Software, Inc., Santa Fe, NM where his work currently includes structure based design and X-ray crystallography applications.