Chapter 41
Protein Ligand Interaction Fingerprints

Ali HajiEbrahimi
Shiraz University of Medical Sciences, Iran

Mohsen Ranjbar
Shiraz University of Medical Sciences, Iran

Hamidreza Ghafouri
Shiraz University of Medical Sciences, Iran

Amirhossein Sakhteman
Shiraz University of Medical Sciences, Iran

ABSTRACT

A most challenging part in docking-based virtual screening is the scoring functions implemented in various docking programs in order to evaluate different poses of the ligands inside the binding cavity of the receptor. Precise and trustable measurement of ligand−protein affinity for Structure-Based Virtual Screening (SB-VS) is therefore, an outstanding problem in docking studies. Empirical post-docking filters can be helpful as a way to provide various types of structure-activity information. Different types of interaction have been presented between the ligands and the receptor so far. Based on the diversity and importance of PLIF methods, this chapter will focus on the comparison of different protocols. The advantages and disadvantages of all methods will be discussed explicitly in this chapter as well as future sights for further progress in this field. Different classifications approaches for the protein-ligand interaction fingerprints were also discussed in this chapter.

INTRODUCTION

The number of protein-ligand complex structures from both in silico and experimental methods have highly increased during the past decade. This was due to advances in structure-based drug design in pharmaceutical industry, high-throughput NMR technology, crystallography and structural genomics. In the public domain, availability of over 5000 small molecule complexes within databases is the result of many effective studies performed in this field (Deng, Chuaqui, & Singh, 2004).

Utilizing this experimental and in silico information depends highly on the ability to organize, analyze and mine the structural data, based on virtual chemical library screening (Deng et al., 2004).

DOI: 10.4018/978-1-5225-1762-7.ch041

The most important merit of virtual screening as a complementary approach to the traditional high-throughput screening is to give insights into molecular recognition in biological systems as well as to facilitate the design of novel therapeutics (Kelly & Mancera, 2004). However, some pitfalls are observed in virtual screening studies, which can be summarized as bellow (Scior et al., 2012):

1. **Detecting High Affinity Molecules by Virtual Screening:** Virtual screening is not considered as a successful tool for identifying high affinity substances. Since reaching highly attracted compounds to particular target is possible during the optimization process.
2. **Certainty of Inquiries:** The setting chosen to select potent materials from a large number of substances are very arbitrary and may depend on the performance of screening.
3. **An Expectation the False Binding Site:** Possibility of binding to exactly correct poses is low in fact. Actually, a few studies revealed exactly true prediction of binding modes.
4. **Predicting Water Mediated Interactions:** Possibly there is many water mediated hydrogen bonds between ligand and receptor. Prediction of count, situation and orientation of these interactions is very ambitious.
5. **Allosteric or Multiple Binding Sites vs. Single Ones:** Many receptors possess allosteric or multiple binding sites, but VS approaches are unable to predict the affinity to an allosteric binding site.

Retrieving useful and crucial information for rational drug discovery is not facile due to an overwhelming number of 3D X-ray and NMR structures. In order to perform virtual screening, these prospective ligands need to be filtered in order to reduce their number for more precise and intensive studies. (Kelly & Mancera, 2004).

Up to now, a variety of approaches have been developed to capture structural knowledge ranging from 2D ligand filters and pharmacophores to 3D interaction constraints and pharmacophoric features (Nandigam, Kim, Singh, & Chuaqui, 2009). All these methods can be used to limit poses generated from ligand receptor docking. Although analyzing a few structures by in silico methods is now considered as a usual practice, mining and making comparisons between a huge number of protein-ligand complexes needs special care such as simplification of the 3D information. Among the most useful simplification processes for analyzing protein-ligand interactions is the conversion of atomic coordinates into simpler 1D or 2D fingerprints (Desaphy, Raimbaud, Ducrot, & Rognan, 2013). Moreover, because binding information is encoded as 1D fingerprint strings, pattern identification of the fingerprints can be done using machine learning and clustering methods. This approach is beneficial in enhancing the ability to make useful implications that are not apparent by looking at individual structures (Nandigam et al., 2009).

BACKGROUND

Fingerprints are easy to generate and can manipulate, compare and, therefore, enable a systematic analysis of large data sets. They are largely used to describe and compare molecular objects (small molecular weight ligands, pharmacophores, proteins and protein-ligand binding sites) and can be used as descriptors. This feature of fingerprints makes it possible to use them by drug design programs as well as in silico screening tools. These fingerprints are frequently manipulated in either ligand based, structure