Chapter 2
Virtual Screening:
An Overview on Methods and Applications

Khaled H. Barakat
University of Alberta, Canada

Jonathan Y. Mane
University of Alberta, Canada

Jack A. Tuszynski
University of Alberta, Canada

ABSTRACT

Virtual screening, or VS, is emerging as a valuable tool in discovering new candidate inhibitors for many biologically relevant targets including the many chemotherapeutic targets that play key roles in cell signaling pathways. However, despite the great advances made in the field thus far, VS is still in constant development with a relatively low success rate that needs to be improved by parallel experimental validation methods. This chapter reviews the recent advances in VS, focusing on the range and type of computational methods and their successful applications in drug discovery. The chapter also discusses both the advantages and limitations of the various techniques used in VS and outlines a number of future directions in which the field may progress.

INTRODUCTION

Once, a US General summarized his philosophy on warfare in just four concise statements, "The art of war is simple enough. Find out where your enemy is. Get at him as soon as you can. Strike him as hard as you can, and keep moving." Although these overarching statements formed the basic premise of modern war strategies, the same concepts have been applied in designing new drugs aimed at combating a broad range of diseases. In this context, rational drug design (Mandal et al., 2009) has been established as an exciting research approach aimed at developing safer and more efficacious drugs. The ultimate goal of this research is to design small organic non-peptidic compounds that bind to a specific molecular target, and result in the inhibition (or less frequently, activation) of a particular protein or enzyme involved in a given cellular pathway.

DOI: 10.4018/978-1-60960-491-2.ch002

The development of such drugs has been recognized early on by the pharmaceutical industry as a principal foundation that provides it with the necessary return on investment to fuel further research and development (Szymkowski, 2005) leading to a discovery and development cycle.

Our understanding of cell mechanisms and pathways becomes deeper and clearer every day. This is largely due to the great efforts and hard work of genomic and proteomic research groups who add novel targets for drug intervention on a regular basis (Drews, 2000; Fishman & Porter, 2005; Hopkins & Groom, 2002). Thus far, several hundred proteins have been synthetically reproduced and many of them are currently evaluated for their druggability (Hopkins & Groom, 2002). These targets involve several families comprised of G-protein coupled receptors (GPCRs), ligand-gated ion channels (LGICs), cytoskeleton proteins, phosphatases, kinases, nuclear receptors (NRs) and DNA repair proteins. The growing list of potential drug targets encourages a bold question if it is in principle possible to restore any diseased cell to a healthy state by uncovering a drug for every potential druggable target? Certainly, if this dream is ever realized, many diseases will be cured and relegated to the dustbin of history in a manner similar to the effect of the discovery of vaccines in the 19th and early 20th century.

Without a doubt, developing a new drug is a highly structured and expensive route that begins with the identification of the target and concludes with a phase III clinical trial followed by marketing (Fishman & Porter, 2005). A candidate drug may never materialize into a safe and efficacious medicine due to its failure to comply with stringent requirements at any stage of the drug discovery process. The further a potential drug progresses in the development process, the more costly its failure becomes. Accordingly, it is important to reduce the probability of late-stage expensive failures by identifying a diversity of lead compounds that are suitable for structural optimizations. Throughout the last two decades, experimental high throughput screening (HTS) and combinatorial chemistry formed the principal source for lead identification. However, as these approaches are particularly expensive and require considerable resources in terms of equipment and skills of the highly qualified personnel (Lahana, 1999), it is vital to search for an alternative or a complementary low-cost technique that aids in the discovery of new bioactive compounds while maintaining the high yield and rapidity of HTS. More recently, a new trend was born which has been named computational virtual screening (VS) or *in silico* screening (Stahura & Bajorath, 2004; Zoete, Grosdidier, & Michielin, 2009).

While the fundamentals of VS have been accumulated from a few studies in the early nineties (DesJarlais et al., 1990; Kuntz, 1992; Kuntz et al., 1982), the term "virtual screening" was first used by Horvath in 1997 in his study that led to the discovery of new Trypanothione Reductase inhibitors (Horvath, 1997). These pioneering efforts defined the overall concept of a typical VS computational protocol as "searching for bioactive molecules within large compound databases". These molecules are predicted to complement a specific binding site of a particular molecular target in terms of parameters such as shape, charge, the number of hydrogen-bond donors/acceptors, and several additional biochemical characteristics. Over the past decade, the method has undergone immense improvements and gained popularity as a result of an exponential increase in the performance of computer hardware, more efficient algorithms and methods as well as vastly enhanced human expertise. Driven by the combined efforts of many research groups these advances have been directed toward increasing the accuracy in selecting active compounds and amplifying hit (active) rates while keeping the computational cost as low as possible (Abagyan & Totrov, 2001; Schneider & Bohm, 2002; Shoichet et al., 1993). Currently, VS is considered to be a valuable prototype within the rational drug design tool box, helping in prioritizing compounds for