

Chapter 9

Protein Docking and Drug Design

Aditi Gangopadhyay
Jhargram Raj College, India

Hirak Jyoti Chakraborty
Central Inland Fisheries Research Institute, India

Abhijit Datta
Department of Botany, Jhargram Raj College, India

ABSTRACT

Protein docking is integral to structure-based drug design and molecular biology. The recent surge of big data in biology, the demand for personalised medicines, evolving pathogens and increasing lifestyle-associated risks, asks for smart, robust, low-cost and high-throughput drug design. Computer-aided drug design techniques allow rapid screening of ultra-large chemical libraries within minutes. This is immensely necessary to the drug discovery pipeline, which is presently burdened with high attrition rates, failures, huge capital and time investment. With increasing drug resistance and difficult druggable targets, there is a growing need for novel drug scaffolds which is partly satisfied by fragment based drug design and de novo methods. The chapter discusses various aspects of protein docking and emphasises on its application in drug design.

INTRODUCTION

Docking the Big Data Surge to Drug Design

With decreasing costs in human genome sequencing, advances in exome sequencing, high-throughput peptide sequencing and a growing network of scientific collaborations, the golden age of Molecular Biology is presently sensing a massive big data surge. The cost of sequencing a single human genome dropped steeply after 2007 with the advent of next generation sequencers (Wetterstrand, 2016). Genome-wide studies are now possible, which feed information to the human diseasome (Hirschhorn & Daly,

DOI: 10.4018/978-1-5225-2607-0.ch009

2005; McCarthy et al., 2008). Colossal advances in genome informatics, personal genomes, make way for the field of personalised medicines, which hold the future of medicine and drug discovery (Agyeman & Ofori-Asenso, 2015; Ginsburg & McCarthy, 2001; Stein, 2010). However, the discovery of a single drug is presently a matter of years and billions of US dollars (Avorn, 2015). With personalised medicines out on the field, the future demands high-throughput drug design to stay at par with the big data explosion in biology and medicinal chemistry (Lusher, McGuire, van Schaik, Nicholson, & de Vlieg, 2014). As if this wasn't enough, the human race is constantly at battle with evolving viral strains, bacterial multidrug resistance, undruggable targets, epidemics and accelerated lifestyle-associated risks. Future drug design will demand smart and robust technologies capable of handling the five V's of biological big data: volume, velocity, variety, veracity and value. The discovery of computational power has been a blessing to Science. Computer-aided drug design (CADD) is presently able to screen chemical libraries in the order of millions, in minutes. An important technique in CADD is molecular docking, used in structure-based drug design. CADD is evolving rapidly and sharply, and probably by the time this book is published, there will be newer tools and techniques in the field. Molecular docking serves one of the most important objectives in drug design and molecular biology: to model and comprehend molecular interactions.

This chapter attempts to discuss the various aspects of protein docking: the kinds, purpose, algorithms, scoring functions, tools and some practical facets such as the docking tools, file formats, visualisation and computational time. The chapter places emphasis on the application of high-throughput protein docking for handling big data in CADD, and illustrates with case studies.

Background

The docking technique was first applied to biology in 1975 by Levinthal et al. to determine the interactions of sickle haemoglobin. The earliest mention of “molecular docking” in ScienceDirect is in a work by Luskey et al. in 1981. Today, “molecular docking” returns over 5,000 articles for the year 2016 alone in ScienceDirect, of which over 3,000 are linked to drug design. Docking is applied in CADD to model protein interactions to small molecules, fragments and peptides and to examine the structure of protein-protein and protein-nucleic complexes. The most commonly encountered docking computations in CADD are protein-small molecule and protein-fragment, usually on a high-throughput scale.

Molecular docking involves two steps: pose generation and scoring. One molecule can bind to the other molecule in n number of ways or poses. In pose generation, the n poses of one molecule with respect to the other are generated by the docking algorithm. The poses are scored by a scoring function. The optimum solution is selected on the basis of the score and other parameters (Lengauer, 2008). One of the main difficulties in any docking computation is identifying false positives and selecting the natural binding mode, which is a test of the power of the scoring function. Generally, the use of consensus scoring using different scoring functions and cross-platform docking with different algorithms is suggested to eliminate false positives and identify the correct binding mode.

Protein Docking: Getting Them Sorted

The individuality and dimensionality of biological problems makes every docking computation unique and each case requires an exclusive solution. This section will attempt to integrate protein docking com-

33 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/protein-docking-and-drug-design/182948

Related Content

Brain Interface: Nano-Scaled Device as an Improvement in the Process of Learning

Rinat Galiautdinov (2021). *International Journal of Applied Research in Bioinformatics* (pp. 23-34).

www.irma-international.org/article/brain-interface/278749

Linking Interactome to Disease: A Network-Based Analysis of Metastatic Relapse in Breast Cancer

Maxime Garcia, Olivier Stahl, Pascal Finetti, Daniel Birnbaum, François Bertucci and Ghislain Bidaut (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 151-172).

www.irma-international.org/chapter/linking-interactome-disease/76061

Statistical Analysis for Radiologists' Interpretations Variability in Mammograms

Ahmad Taher Azar (2012). *International Journal of Systems Biology and Biomedical Technologies* (pp. 28-46).

www.irma-international.org/article/statistical-analysis-radiologists-interpretations-variability/75152

Animal Actin Phylogeny and RNA Secondary Structure Study

Bibhuti Prasad Barik (2015). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 46-61).

www.irma-international.org/article/animal-actin-phylogeny-and-rna-secondary-structure-study/165549

Introduction to Classification Error Estimation

(2011). *Feature Selection and Ensemble Methods for Bioinformatics: Algorithmic Classification and Implementations* (pp. 334-340).

www.irma-international.org/chapter/introduction-classification-error-estimation/53912