AUTOMATING LEAD MOLECULE DISCOVERY FOR PROTEIN TARGETS VIA *SANJEEVINI* SERVER

TANYA SINGH

DEPARTMENT OF CHEMISTRY
INDIAN INSTITUTE OF TECHNOLOGY DELHI
APRIL 2015

AUTOMATING LEAD MOLECULE DISCOVERY FOR PROTEIN TARGETS VIA *SANJEEVINI* SERVER

by

TANYA SINGH

DEPARTMENT OF CHEMISTRY

Submitted

in fulfillment of the requirements of the degree of

Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 2015

Certificate

This is to certify that the thesis entitled, "**Automating Lead Molecule Discovery for Protein Targets via Sanjeevini Server**", being submitted by **Ms. Tanya Singh** to the Indian Institute of Technology Delhi for the award of the degree of **Doctor of Philosophy** in Chemistry is a record of bonafide research work carried out by her. Tanya Singh has worked under my guidance and supervision and has fulfilled the requirements for the submission of this thesis, which to my knowledge has reached the requisite standard.

The results contained in this dissertation have not been submitted in part or full to any other University or Institute for the award of any degree or diploma.

Dr. B. Jayaram
Professor
Department of Chemistry
Indian Institute of Technology Delhi
INDIA

Dated

Acknowledgements

I express my sincere gratitude to Prof. B. Jayaram. I appreciate all his contributions which have helped me to travel in the world of Science. He has always supported me through his ideas, funding, encouragements and his patience which allowed me to progress as a researcher. The joy and passion he has for Science was always motivational for me. I am also thankful to him for showing confidence in my abilities and providing with different opportunities in the field of Science as well as Teaching. I would like to thank him for encouraging my research and allowing me to grow as researcher.

I am thankful to all the faculty and staff members of Chemistry Department Indian Institute of Technology Delhi for their help and support during this project. I thank Department of Biotechnology, India and the Council of Scientific and Industrial Research, India, for financial support.

I am grateful to all my lab members of the Supercomputing Facility for Bioinformatics and Computational Biology and the Chemistry Department. The lab members have always been a source of friendships and good scientific advises. I am extremely thankful to my friend Priyanka for her support, her companionship, her encouragements and her scientific advises throughout the PhD duration. I thank her for her unique presence through thick and thin.

I specially thank my parents and my elder brother, for bringing me to this stage today. I thank my Mom, to be a patient listener throughout my PhD work and for her motivational words which always kept me sincere towards my work. I thank

them all for giving me all the liberties in life which helped me constantly to progress as a researcher. I thank my sister-in-law for her strong support and her pride in me.

I greatly appreciate the love, affection, patience and encouragement received from my husband, Nishant Kumar. I thank him for his support, for forgiving me for my difficult behavior, for staying calm, hopeful and patient which made me to finish the dissertation work.

Finally, I dedicate this thesis to my parents, for whom words are not just enough to describe.

Abstract

This thesis focuses on developing atomic level computational algorithms for binding site detection and applying them in improving protein-ligand docking in an automated mode. The protocols have been validated on diverse datasets and a good agreement/correlation has been achieved with experiment. The algorithm developed as part of this thesis has been web enabled and integrated as part of the *in silico* drug design software christened *Sanjeevini* which is freely available at http://www.scfbio-iitd.res.in/sanjeevini/sanjeevini.jsp.

The thesis is divided into six chapters. Chapter I discusses the current status of computer-aided drug design in general and its role in drug discovery process. A brief overview of protein binding site detection algorithm is reported. Various virtual screening methodologies including docking and scoring approaches and some of the software known for them is also presented.

Given a three dimensional structure of a protein, some of the questions like, if the protein is a suitable drug target, how one can detect possible binding sites where a small molecule can bind, or if the binding sites have suitable properties for a high affinity binding with small molecules, are of immediate value in understanding the protein-ligand binding mechanism. Chapter 2 of this thesis explains various physico chemical properties involved in binding of a small molecule against protein. An algorithm for binding site detection in a protein target, its validation on more than 600 proteins with known binding sites and the web enabling of the algorithm is described.

In Chapter 3, development of a docking and scoring algorithm and its automation with the binding site detection is presented. The results of the automated binding site detection followed by docking and scoring on 170 protein-ligand complexes are presented and discussed. This chapter describes in detail, the computational protocol, preparation of the target protein and the ligand molecule, various model validations used in this work and the web enabling of the algorithm developed. Chapter 4 describes the development of a flexible docking and scoring algorithm considering flexibility of a ligand molecule, its validation, and results obtained on 170 protein-ligand complexes and the web enabling of the tool developed.

Design of novel molecules being a direct application of the computational methodologies discussed in this thesis, the methodology was extended to an *in silico* design and study of molecules with high affinity and specificity against an important class of drug targets, the matrix metalloproteinases, is discussed in Chapter V. The computational algorithm helped in formulating important design principles to come up with specific molecules against MMP-2, MMP-3, MMP-8, MMP-9 and MMP-13. Finally in Chapter VI, a summary and some perspectives emerging from this thesis work on drug design *in silico* are provided.

CONTENTS

List of Figures

List of Tables