Computational Analyses of the Binding Free Energies of Protease - Inhibitor Complexes: Implications to Drug Design

by

Parul Kalra
Department of Chemistry

Submitted
In fulfillment of the requirement of the degree of
Doctor of Philosophy

to the

Indian Institute of Technology, Delhi
May, 2002

CERTIFICATE

This is to certify that the thesis entitled "Computational Analyses of the Binding Free Energies of Protease-Inhibitor Complexes: Implications to Drug Design" being submitted by Ms. Parul Kalra 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. Ms. Parul Kalra 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.

Date: 17-5-2002

B. Jayaram

Professor

Department of Chemistry

Indian Institute of Technology, Delhi

ACKNOWLEDGEMENT

This thesis was prepared under the guidance of Prof. B. Jayaram. His constant encouragement, practical advice, inculcating an attitude of never to give up and confidence were essential for the completion of the thesis. He always supported me when I lost heart and infused fresh confidence in me that helped me to look at things from a different perspective. For all his attention, guidance, insight and support over the years, I wish to express my sincerest gratitude.

My thanks to the Head of the Department – Chemistry, IITD for providing me full access to the departmental facilities. Without his cooperation and willingness the research wouldn't have been possible.

I am also extremely thankful to all my teachers at the Department of Chemistry, IITD, for their cooperation at every stage. I would like to express special thanks to Prof. R. C. Anand and Dr. N. D Kurur.

My colleagues in the laboratory have been very supportive and helpful at every stage of my research. I would specially like to thank Dr. Achintya Das who was always there to help me get out of a computing problem. My special thanks also goes to Sonal Jain, Vasisht Reddy, Pooja Narang and Dr. Archana Milhotra for their timely support and encouragement. I would also like to thank Lata Narayanan, Preeti Misra, Tarun Jain, Arunima Acharya, Surojit Bose, Saher Afshan and Srikant Reddy.

I would like to take this opportunity to thank my husband Dhiraj Bhatia. without whose constant support, patience and his confidence in me, this thesis wouldn't have been possible. He was always there by my side and helped me in whatever way he could. He

had to tolerate my mood swings when the research wasn't going too well. I thank him for all his patience and dedicate my thesis to him.

My sincerest thanks to my mother in law, father in law, mother, brothers in law – Sulekh and Prashant, brother Puneet and bhabhi Shubi for their timely help and support.

I would also like to acknowledge the scholarship from the Council of Scientific and Industrial Research (CSIR) that supported me during the course of my research. The financial support received from the Indo-French Center for the Promotion of Advanced Research (IFCPAR) and the Department of Biotechnology (DBT) in the form of a grant to Prof B. Jayaram for creating research facilities used in this thesis work is greatly acknowledged.

Parul Kalra

ABSTRACT

The concept of specificity is central to the action of enzymes, which are the functional units of cell metabolism. The recognition of a specific substrate by an enzyme from a myriad of molecular species leads to the proper conduct of physiological activities and thus the sustenance of life. Certain diseases/disorders such as ulcers, molecular dystrophy, cancer, high blood pressure etc. are related to the functioning of certain enzymes. These can be controlled by designing a drug that inhibits the enzymes catalytic activity. In such cases, estimation of the binding free energies of the drug to its receptor is a prerequisite to determine the potency of the drug. This thesis is aimed at developing a computationally expedient and theoretically rigorous methodology for the estimation of binding free energies for protease-inhibitor complexes and formulating some general principles for drug design taking protease-inhibitor complexes as systems of study. The thesis is divided into seven chapters and the contents of each chapter are given underneath.

The introduction to the thesis in chapter 1 reviews the structure, function aspects of proteases, development of theoretical methods for estimating the thermodynamics of protein-drug interactions and related experimental studies on binding of inhibitors to the various proteases.

Estimation of the thermodynamic parameters of binding requires structures of the interacting molecules in their initial and final states, accurate description of the partial atomic charges on each of the interacting molecules, a potential of mean force between the interacting macromolecules in the given solvent environment which includes electrostatic

and nonelectrostatic contributions, methodologies to estimate translational, rotational, vibrational and configurational entropy losses upon binding, and methods to incorporate salt effects. An all atom treatment to arrive at the enzyme-inhibitor binding free energies, which considers all the above factors is beyond the capabilities of current generation supercomputers. There is a dire need in the field of computer aided drug design for system independent, computationally rapid and reliable protocols for free energy evaluations. A hierarchy of computer models with increasing levels of theoretical rigor are examined in chapter 2 with a special emphasis on speed and concord with experiment. This chapter gives a detailed account of the statistical mechanical theory of protein inhibitor binding in aqueous media, the master equation approach and other approximations. Details on the structural and energy analyses protocols developed and adopted in the study as pertinent to active site recognition are also provided.

Pursuing the goal of developing and making available a software for assessing binding affinities, a computationally rapid, albeit elaborate, methodology is presented in chapter 3 to estimate and analyze the molecular thermodynamics of enzyme-inhibitor binding with crystal structures as the point of departure. The complexes of aspartic proteinases with seven inhibitors are adopted for this study. The standard free energy of complexation is considered in terms of a thermodynamic cycle of six distinct steps decomposed into a total of 18 well-defined components. The model employed involves explicit all atom accounts of the energetics of electrostatic interactions, solvent screening effects, van der Waals components and cavitation effects of solvation combined with a Debye-Huckel treatment of salt effects. The magnitudes and signs of the various components are estimated using the AMBER parm94 force field, generalized Born theory

and solvent accessibility measures. Estimates of translational and rotational entropy losses upon complexation as well as corresponding changes in the vibrational and configurational entropy are also included. The calculated standard free energies of binding at this stage are within an order of magnitude of the observed inhibition constants and necessitate further improvements in the computational protocols to enable quantitative predictions. The net binding free energies are a resultant of several competing contributions with 6 out of the 18 terms favoring complexation. In particular, the nonelectrostatic contributions, i.e. the net van der Waals interactions and the differential cavitation effects favor binding. Electrostatic contributions show considerable diversity and turn out to be favorable in a consensus view for the seven aspartic proteinase-inhibitor complexes examined here.

Previous studies on the aspartic proteinase-pepstatin complexes taught us that the binding free energy estimates are sensitive to the structure preparation protocols. To investigate this further and to improve the theoretical models, a detailed study on 30 protease-inhibitor complexes is carried out in the presence of explicit solvent in chapter 4. An additional goal of this study is to elucidate some structure and energy based principles for drug design. The analysis revealed that in all the 30 complexes packing, electrostatic and hydrophobic forces are found to be favorable to binding while the entropy and the salt effects effects are unfavorable to binding in a consensus view. A residue wise analysis is also carried out on all the systems to identify residues of each of the inhibitors important to binding in terms of their main chain and side chain contribution. An attempt is also made based on experiment and the present series of

computations to tabulate certain common properties that a good protease inhibitor should possess.

The theoretical rigor of the methods developed and applied in chapter 4 is taken one step further in chapter 5 by introducing Boltzmann averaging via molecular simulations. The method formulated involves developing molecular dynamics trajectories of the enzyme, the inhibitor and the complex, This is then followed by a free energy component analysis that conveys information on the physico-chemical forces driving the protein-ligand complex formation and enables an elucidation of drug design principles for a given receptor from a thermodynamic perspective. The complexes of HIV-1 protease with two peptidomimetic inhibitors are taken as illustrative cases. Four nanosecond level all atom molecular dynamics simulations using explicit solvent without any restraints are carried out on the protease-inhibitor complexes and the free proteases and the trajectories were analyzed via a thermodynamic cycle to calculate the binding free energies. The computed free energies are seen to be in good accord with the reported data. Net van der Waals and hydrophobic contributions are found to be favorable to binding while the net electrostatics, entropies and adaptation expense are unfavorable in these protease-inhibitor complexes. The hydrogen bond between the CH_2OH group of the inhibitor at the scissile position and the catalytic aspartate is found to be favorable to binding. Various implicit solvent models are also considered and their shortcomings discussed. In addition, some plausible modifications to the inhibitor residues are attempted which led to better binding affinities. The generality of the method and the transferability of the protocol with essentially no changes to any other protein-ligand system are emphasized.

An important question in drug design is under what conditions is a hydrogen bond between the enzyme and the inhibitor enhances binding. Chapter 6 addresses this question using the complexes of HIV-1 protease, carboxypeptidase A and thermolysin as case studies and the free energy methodology developed above in chapter 5. Mutations are also carried out on the inhibitors to delete or introduce hydrogen bonds in the complexes studied coupled with quantum mechanical charge derivations and molecular dynamics simulations on each of the mutants. In case of the HIV-1 protease inhibitor a hydroxyethylene group is introduced at the scissile position and it was found that such a mutation results in better binding affinity. For the carboxypeptidase A and thermolysin inhibitors the anionic group (phosphonamidate and mercapto respectively) were protonated and such a mutation causes a decrease in the magnitude of the binding free energy.

A summary of the series of theoretical studies undertaken on protease-inhibitor complexes, the lessons drawn for protease active site directed drug design are presented in chapter 7. The possibility of developing an automated software based on the methodology adopted in the thesis work is also explored.

CONTENTS