COMPUTATIONAL STUDIES ON THE DYNAMICS, THERMODYNAMICS AND STRUCTURE OF DNA- MINOR GROOVE BINDER COMPLEXES AND DEVELOPMENT OF A PROTOCOL TO COMPUTE DNA-DRUG BINDING AFFINITY

by

SAHER AFSHAN SHAIKH
DEPARTMENT OF CHEMISTRY

SUBMITTED
IN FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

TO THE

INDIAN INSTITUTE OF TECHNOLOGY, DELHI

INDIA

MARCH 2007

To

My Parents

Deeba Shaikh and Nisar Ahmed Shaikh

CERTIFICATE

This is to certify that the thesis entitled, "COMPUTATIONAL STUDIES ON THE DYNAMICS, THERMODYNAMICS AND STRUCTURE OF DNA-MINOR GROOVE BINDER COMPLEXES AND DEVELOPMENT OF A PROTOCOL TO COMPUTE DNA-DRUG BINDING AFFINITY", being submitted by Ms. **SAHER AFSHAN SHAIKH** 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. Saher Afshan Shaikh has worked under my guidance and supervision, and has fulfilled the requirements for 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: 19/7/2007

B. Jayaram
Professor
Department of Chemistry
Indian Institute of Technology, Delhi

ACKNOWLEDGEMENTS

I express my sincere gratitude to Prof. B. Jayaram for his constant support and guidance with which I steadily traveled my way in the world of scientific research. He continually encouraged independence and innovation and has been highly patient while allowing me to evolve and progress as a researcher. I also thank him for always showing confidence in my ability, and assigning me different responsibilities which also helped me hone my teaching and managerial skills.

I am grateful to the Indian Institute of Technology, Delhi and its Deptt. of Chemistry for its inspiring and motivating environment. I am grateful to all my teachers and the administrative staff for their kind cooperation. I thank the Council of Scientific and Industrial Research, India, for financial support.

I cannot express my gratitude in words to my parents, for bringing me to this stage today. My mother, Deeba Shaikh, always trying to understand what work I did with 'DNA', and ever proud about my pursuing PhD. My father, Nisar Ahmed Shaikh, always a silent inspiration with his honest and sincere attitude towards work and life. My parents ensured that I imbibed their mental strength and dedication, without which I could not have done justice to my work. Their unfaltering love and faith in me has been a constant driving force. I dedicate this thesis to them.

I thank my brothers, Raashid and Nazar Shaikh for their strong support and their pride in me; and my sisters-in-law, Manasi Shaikh and Zeenat Rehman for their interest in my progress and their pleasant presence in my life.

I am extremely glad to have spent a long time with exceptionally warm-hearted people at my lab and sharing many fun-filled moments with them. I thank Pooja for her sweet companionship and for the numerous times she cheerfully accommodated me at her hostel room when I got late; Pankaj for his unique presence

and sense of humor, always making the lab environment positive and pleasant; Surojit for his caring nature and being a person one can always turn to for help; Praveen for his cheer and the several times he made sure I reached home safe when late; Gandhimathi for her smiling support and the rapport we shared; Gurvisha for her care and enjoyable company; Vidhu for his quiet help in sorting out my programming woes; Shashank for his concern; Anuj and Shailesh for their help in setting up the webserver; Tarun, Kumkum and Poonam for being highly cooperative co-PhD students, and always making sure I did not miss out on any departmental formality deadlines. I also thank Dr. Sandhya, Lipi, Neelam and Gaurav for their amiability and co-operation. I am grateful to earlier lab members, Sajeedha, Priyanka and Sivasankari for their contribution to this study in its initial stages, to Dr. Parul and Dr. Latha for their presence in my initial years in the lab, and to Srinivas for carrying on the baton now. I also thank Preeti for efficiently handling the accounts and paperwork and Manohar for making the innumerable trips to the coffee shop to get my constant supply of caffeine.

Throughout this research work, I derived strength from my friends, while they stood by me, always believing in my ability. I would especially like to thank Amit, the pillar in my life, for being happy in my happiness, being the punching bag when I was agitated, constantly forgiving me for my difficult behavior and always bringing back my smile; Namit, calming me when I would lose focus, supporting me despite my faults, always refreshing with his unique humor and most importantly, for always being there; Kalyani and Ruchi, for being friends I could always trust and count on. My friends, with their total and unconditional love, made me believe I am good and able, and I will always be grateful to them for this feeling.

Saher Afshan Shaikh

ABSTRACT

Design of small molecules that can bind to DNA in a sequence-specific manner with affinities comparable to regulatory proteins, thus having the potential to modulate DNA replication or transcription, is of considerable interest in anticancer and antibiotic therapy. A reliable and theoretically rigorous computational protocol for studying the structure-energy relationships in DNA-drug complexes and estimation of DNA-drug binding affinity would greatly facilitate rapid design and screening of lead-like molecules targeted to DNA as well as enhance the present understanding of DNA recognition. Development of such a protocol entails an exhaustive study of the structure, dynamics and thermodynamics of DNA-drug complexes to discern factors important to binding affinity and specificity. DNA-minor groove binder complexes form a suitable series for a systematic investigation since many crystal structures are available for these and their thermodynamics has been well characterized experimentally. Chapter I discusses the current status of computer-aided drug design in general, the significance of DNA as a macromolecular target, small molecules targeting DNA and experimental as well as theoretical studies on DNA-drug interactions.

Computational studies on DNA–drug complexes traversed from the initial energy minimization-based modeling era through a period marked with rapid developments in the theory of DNA simulation, improvements in force fields for nucleic acids and small molecules and other methodological advancements to the present age where advanced explicit solvent simulations and free energy calculations can be performed routinely. However, computational complexities persist in the

treatment of DNA-drug complexes, thus limiting these studies to only a few systems and the information obtained from these studies, though valuable, is insufficient to enable the extraction of a set of general principles for DNA-drug binding. Thus, the development of a state-of-the-art computational protocol for accurate modeling of DNA-drug and estimation of binding affinity is of immediate value in understanding DNA recognition and aiding drug design. Chapter II discusses the computational methodology developed for this purpose and the theory behind the binding affinity estimation methods developed and employed to study DNA-drug binding energetics.

Thermodynamics of DNA-drug binding and the various energetic and structural factors involved in DNA minor groove recognition are investigated via a computational study on energy-minimized structures of 25 DNA-drug complexes and the results are presented in Chapter III. Energetics of the 25 DNA-minor groove binder complexes is developed using the MMGBSA methodology, the nature and contribution of the diverse binding free energy components including the significance of van der Waals, electrostatics and hydrogen bonding interactions and possible threshold limits for some of these properties in terms of optimal binding are analyzed. The successful application of this information to modify a known ligand for improving its binding affinity is demonstrated. Molecular dynamics simulations performed on two representative systems corroborate the conclusions derived from the thermodynamic study. Chapter III discusses the molecular thermodynamics of the 25 DNA-minor groove binder complexes and the availability of a computationally viable free energy methodology that could be of value in drug design.

CONTENTS

Structural and thermodynamic studies in a dynamic environment afford an enhanced view of binding and recognition, hence extensive molecular dynamics simulations (10 ns each) on 25 DNA-minor groove binder complexes were carried out and the analyses are reported in Chapter IV. The molecular dynamics simulations also introduce theoretical rigor into binding free energy estimates and eliminate some of the limitations inherent in the single point free energy calculations. Free energy component analyses reveal improved binding free energy estimates, reflecting the effects of thermal averaging, dynamics, solvent and counterion effects. Structural studies on the DNA and the surrounding solvent and counterions are also carried out. Also, a novel atomic level model for a static and dynamic view of DNA recognition derived from a new classification of contacts based on hydrogen bonding and steric complementarity is developed. A detailed discussion of these observations and results is presented in Chapter IV.

Development of a computational tool that can provide quick and accurate estimates of binding affinities for DNA-drug complexes can assist in speedy and efficient design of new drugs for DNA. Though several energy / scoring functions have been reported in literature for protein-drug, protein-DNA and protein-protein binding affinity estimations, and some presented as computational tools, parallel efforts for DNA-drug systems have been lacking and existing methods such as MMGBSA and MMPBSA have their limitations. PreDDICTA, discussed in Chapter V, is a computational protocol developed to calculate DNA-drug binding affinity, and represents a step towards filling this lacuna. The PreDDICTA energy function is validated on 50 DNA-minor groove binder complexes including crystal and model-

built structures and the protocol is translated into a swift, web-enabled computational tool and made freely available at http://www.scfbio-iitd.res.in/preddicta for ΔG° and ΔT_m predictions of DNA-minor groove binder complexes. Chapter V discusses the development of this energy function, its validation and details about the web-based computational tool.

Design of novel molecules and the assessment of their stability and energetics being a direct application of the computational methodologies and binding affinity prediction methods discussed in the thesis, the methodology was extended to an *in silico* design and study of novel modified nucleotides and is discussed in Chapter VI. The computational protocol adopted to study DNA-drug interaction and the energy function in PreDDICTA being phenomenon-based, are extendable to other biomolecular interactions such as DNA-protein, protein-drug, protein-protein as well as DNA interstrand interactions, as demonstrated in this study. Five novel modified nucleotides were designed keeping in view that structural asymmetry (5' → 3') associated with each strand of the DNA is due to the C3' to C2' bond in each nucleotide unit and the absence of this bond results in symmetric nucleic acids of potential synthetic and therapeutic interest. Detailed structural and energetic analyses were carried out on double helices comprising these acyclonucleotides and the results are reported in Chapter VI.

In summary, this thesis addresses dynamic, thermodynamic and structural studies of a series of DNA-minor groove binder complexes via computer modeling and simulations, discusses structure-energy relationships and molecular attributes of DNA recognition emerging from these studies, presents the subsequent development

of a web-enabled computational protocol to elicit reliable estimates of DNA-drug binding free energies and ΔT_m and finally discusses the application of the methodology to study some novel designed oligoacyclonucleotides.