COMPUTATIONAL STUDIES ON TGF-β LIGANDS AND THEIR INTERACTION WITH RECEPTOR I AND RECEPTOR II

by

MD. SHAHID NAYEEM

Department of Chemistry

Submitted

in fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

December, 2012

CERTIFICATE

This is to certify that the thesis entitled, "***Computational studies on TGF-β ligands and their interaction with receptor I and receptor II***", being submitted by Mr. **Md. Shahid Nayeem** 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 him. Mr. Md. Shahid Nayeem 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. Shashank Deep)**

Assoc. Professor, Department of Chemistry

Indian Institute of Technology Delhi

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere thanks to all the people who directly or indirectly have been instrumental to my success throughout my Ph. D., here at I.I.T., Delhi.

First and foremost, I wish to express my sincere gratitude and respect to my Research Supervisor, **Dr. Shashank Deep**, for having been so supportive, for inspiring me to stretch myself over my limits, and for his never ending enthusiasm. I am extremely appreciative for his guidance with constructive suggestions throughout my research work which made possible this thesis.

I am thankful to the SRC members for giving their valuable time and expert suggestions during comprehensive and synopsis presentations.

I wish to extend my thanks to Prof. B. Jayaram for providing me the computational facilities at SCFBIO, IIT, Delhi. I must thank Shashank Shekhar, at SCFBIO, for his technical support. I express my sincere thanks to Sumit Mookerjee, Bhushan Awasarmol and Sanjiv Chacko for providing me computational facility at IUAC, New-Delhi, where most of the work reported in this thesis has been carried out.

I take this opportunity to express my deep sense of respect and thanks to Prof. Barry Honig, Columbia University, New York and Dr. Emil Alexov, Clemson University, Atlanta, U.S.A for helping me in running Delphi programme and responding to each of my queries with efficient and effective suggestions. I must thank Prof. A. M. J. J. Bonvin, Utrecht University, Netherland for his support in running HADDOCK and his worthwhile suggestions for explaining the scoring functions of the docked structure. I

feel short of words to express my gratitude and thanks to Marc Baaden, Université Paris Diderot, Paris, France for providing me MMPBSA tool and his critical remarks for manuscript preparation and his support in providing me his own computational resources. I must thank the Gromacs User Group for providing me fundamental support at every problem in running Gromacs and special thanks to Justin Lemkul on this group for his quick and crucial suggestion.

Many thanks are due, to all my teachers who have taught me at any point of time and have laid the foundation and helped me in growing both at the academic as well as personal level.

I thank all my labmates Anurag, Komal, Unnati, Ashhar, Robin, Babu, Shakir, Vinay, Preeti, Shivnetra, Nidhi and M.Sc. students for respecting my seniority and being always supportive to me.

I would like to express my sincere thanks to Aligarh Muslim University, Aligarh for providing me study leave, and Prof. Shamsul Haque siddiqui and Dr. Syed Mashhood Ali as well as all my colleagues in the department at A.M.U. for their support and encouragement.

My next acknowledgments are due to my parents and family members who are always behind me with their prayers, blessings and support.

And finally and most of all I thank my wife Atiya for her unending love and support during these years, ensuring the success of this work.

(Md. Shahid Nayeem)

ABSTRACT

Signal transduction pathways are central to the functioning, regulation and maintenance of cells and living systems. TGF-β3 is one of the ligands of TGF-β signal transduction pathway which is known to be associated with growth, differentiation, angiogenesis and several other physiological roles. An understanding of TGF-β signalling pathway is of immense importance in dealing with physiological malfunctions and drug discovery. This thesis entitled "***Computational studies on TGF-β ligands and their interaction with receptor I and receptor II"*** is aimed at understanding the role of TGF-β3 ligand in TGF-β signaling pathway.

The thesis has been divided into seven chapters. Chapter-1 entitled "**Introduction**" is a brief introduction of TGF-β superfamily ligands, receptors and the signal transduction pathways. The Protein-Protein interaction phenomena, which is indispensable for such pathways are described with its salient features including hot residues and anchor residues. Some examples from literatures are described here to show the use of computational investigation and its validity in the study of proteins and protein-protein interactions. The Chapter-2 entitled "**Methods**" describes in brief the theory of Molecular Dynamic simulation and Molecular Docking. An introduction to MMPBSA binding energy calculation, Poisson-Boltzman solver for electrostatic analysis and computational alanine scanning are given in this chapter. A very brief notes about several softwares used are also given.

Chapter-3 entitled "**Rationalization of poor solubility of TGF-β3 using MD simulation**" describes the Molecular dynamic simulation of TGF-β3 and TGF-β1 in explicit water at different temperature and in the presence of urea solutions. The TGF-β3 ligand is most aggregation prone among the homologue despite sequence and structure similarity and hence a comparative molecular dynamic simulation study was done for TGF-β1 and TGF-β3. H3-helix of TGF-β3 is found to be relatively unstable The role of this unstable H3 α-helix of TGF-β3 for causing aggregation was validated by MD simulation of the three of small peptide fragments (containing residues of H3-helix from residue number 50 to 75) from TGF-β3 and TGF-β1 in a box. These small fragments of TGF-β3 and TGF-β1 were also simulated as a function of pH.

In Chapter-4 entitled **"Conformational preference of Transforming Growth Factor β1 and β3 for its open versus closed conformation",** the role of H3-Helix is investigated in deciding the conformational preference of TGF-β3 and TGF-β1 using HADDOCK and Rosetta dock. In contrast to TGF-β1, TGF-β3 is believed to exist in an open conformation, in which both its monomers are loosely packed against each other. To understand the difference in its conformational preference for the open versus the closed structure, two monomers of TGF-β3 with intact and altered H3 α-helix, respectively, were docked against each other using HADDOCK and Rosetta servers. In order to understand the differences in stability of open vs. closed conformation in TGF-β3 and TGF-β1, the calculation of binding energy between two monomers using the MMPBSA approach and anchor residues analysis were done. Computational Alanine

scanning studies using FoldX and PRICE server was also carried out to find hot residues at the interface of open and closed conformations in TGF-β3 and TGF-β1. It was observed that altered H3-Helix is important for the existence of open conformation. In Chapter 5 entitled **"Specificity in the interaction of the TGF-β1 and TGF-β3 with receptor I and receptor II",** computational analysis of the crystal structure of ternary complex of TGF-β1 and TGF-β3 aimed at understanding their specificity is described. Although the structural aspects of complex assembly of TGF-β3 (TGF-β3:TβR2:TβR1) as well as TGF-β1 (TGF-β1:TβR2:TβR1) are known, there is very little information about their conformational preferences and specificity, since both the ligands have similar interfaces with receptors. All important interactions were studied at the interface using NACCESS and Ligplot. Anchor and Hot residues at each interface in TGF-β3 and TGF-β1 ternary complex were determined. Binding energies with receptors were calculated using MMPBSA. Electrostatic contribution to binding free energy, as calculated by Delphi at different pH, indicates that the specificity in complex formation by two ligands may be decided by the pH environment of the cell. Chapter 6 entitled **"BMP2 ligand binding with its receptor II and receptor I: A comparative analysis with TGF-β3 and TGF-β1"** describes the comparative study of the two ternary complex of BMP-2 and ternary complex of TGF-β ligands. This chapter aims to provide an understanding of specificity and promiscuity in ligand-receptor interaction of BMP subfamily. The differential preference of site of binding of Type II receptor in TGF-β and BMP-2 has been attributed to the buried surface area. Hot and anchor residues were

calculated at all interfaces of the complex. MMPBSA and Delphi binding energy of receptors I and II with BMP-2 is also carried out for both the ternary complexes. The electrostatic analysis shows that the ligand receptor binding is influenced by the pH of the environment and probably the influence of electrostatic environment on binding might answer the larger question of specificity and promiscuity in TGF-β signalling complexes. Chapter 7 entitled "**Summary and Future Perspectives**" describes the salient features of these works. In nutshell we have shown that the relatively unstable H3-α-Helix in TGF-β3 as compared to TGF-β1 is responsible for its high aggregation behavior and its preference for open conformation. We have also shown that the pH environment of the cell could be the driving force of the specificity in TGF-β ligand receptor binding and the promiscuous behavior of the ligand.

Further study may result in controlling aggregation during expression and understanding the binding of TGF-β ligands with receptors eventually leading to a thorough understanding of TGF-β signalling pathway.

CONTENTS