

**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- β 3 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- β 3 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- β 3 using MD simulation**” describes the Molecular dynamic simulation of TGF- β 3 and TGF- β 1 in explicit water at different temperature and in the presence of urea solutions. The TGF- β 3 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- β 1 and TGF- β 3. H3-helix of TGF- β 3 is found to be relatively unstable. The role of this unstable H3 α -helix of TGF- β 3 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- β 3 and TGF- β 1 in a box. These small fragments of TGF- β 3 and TGF- β 1 were also simulated as a function of pH.

In Chapter-4 entitled “**Conformational preference of Transforming Growth Factor β 1 and β 3 for its open versus closed conformation**”, the role of H3-Helix is investigated in deciding the conformational preference of TGF- β 3 and TGF- β 1 using HADDOCK and Rosetta dock. In contrast to TGF- β 1, TGF- β 3 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- β 3 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- β 3 and TGF- β 1, 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- β 3 and TGF- β 1. 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- β 1 and TGF- β 3 with receptor I and receptor II”**, computational analysis of the crystal structure of ternary complex of TGF- β 1 and TGF- β 3 aimed at understanding their specificity is described. Although the structural aspects of complex assembly of TGF- β 3 (TGF- β 3:T β R2:T β R1) as well as TGF- β 1 (TGF- β 1:T β R2:T β R1) 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- β 3 and TGF- β 1 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- β 3 and TGF- β 1”** 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- β 3 as compared to TGF- β 1 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

	Page No.
Certificate	I
Acknowledgment	III
Abstract	V
Contents	IX
List of Figures	XV
List of Tables	XXIII
List of Abbreviation	XXV
<i>Chapter-1 Introduction</i>	1
<i>1.1 TGF-β Superfamily</i>	3
<i>1.1.1 Ligands of TGF-β superfamily</i>	4
<i>1.1.2 Receptors of TGF-β superfamily</i>	6
<i>1.1.3 Ligands of TGF-β subfamily</i>	8
<i>1.1.4 Receptors of TGF-β subfamily</i>	9
<i>1.2 Transforming Growth factor beta III</i>	11
<i>1.3 Complex of TGF- β3 and BMP-2 with receptor II and receptor I</i>	12
<i>1.4 Signal transduction pathways</i>	15
<i>1.4.1 Signaling pathway: An overview</i>	15
<i>1.4.2 TGF-β Signalling Pathway</i>	17
<i>1.5 Specificity and promiscuity in the TGF-β signalling pathway</i>	19
<i>1.5.1 Specificity and promiscuity: An overview</i>	19
<i>1.5.2 Specificity and promiscuity: TGF-β signaling pathway</i>	21
<i>1.6 Protein-protein interaction</i>	23
<i>1.6.1 Significant roles of protein-protein interaction in biological systems</i>	23
<i>1.6.2 Protein interface and interfacial residues</i>	24
<i>1.6.3 General Characteristics of protein interface</i>	25
<i>1.6.3.1 Size</i>	25
<i>1.6.3.2 Shape complementarity</i>	26
<i>1.6.3.3 Amino Acid Composition</i>	26
<i>1.6.3.4 Residue contact at the interface</i>	27
<i>1.6.3.5 Forces involved at protein-protein interface</i>	27
<i>1.7 Protein-protein interaction and protein aggregation</i>	30
<i>1.8 Binding Hot Spots and Anchor residues</i>	31
<i>1.9 Computational studies on Bio-molecular Systems</i>	32

1.10	<i>Origin of the problem</i>	43
1.11	<i>Outline of the research work</i>	45
1.12	<i>References</i>	48
Chapter-2	<i>Methods</i>	59
2.1	<i>Introduction</i>	61
2.2	<i>Molecular Dynamic simulation</i>	61
2.2.1	<i>General steps of MD run</i>	62
2.2.2	<i>Equation of motion</i>	63
2.2.3	<i>Force Field</i>	64
2.2.4	<i>Constraints</i>	68
2.2.5	<i>Periodic Boundary condition</i>	68
2.2.6	<i>Neighbouring List</i>	70
2.2.7	<i>Electrostatics and long range interaction</i>	71
2.2.8	<i>Thermostat and Barostat for MD simulation</i>	74
2.2.9	<i>Treating Solvents: Explicit and implicit solvation</i>	75
2.2.9.1	<i>TIP Model</i>	76
2.2.9.2	<i>SPC Model</i>	77
2.2.9.3	<i>SPC/E Model: Extended SPC</i>	77
2.2.9.4	<i>Implicit Solvent Models</i>	78
2.3	<i>Molecular Docking</i>	79
2.3.1	<i>Representation of the system</i>	80
2.3.2	<i>Conformational space search</i>	81
2.3.3	<i>Scoring and ranking of potential solutions</i>	81
2.3.4	<i>Refinement of accepted solutions</i>	84
2.3.5	<i>Incorporating flexibility into protein-protein docking</i>	85
2.3.5.1	<i>Implicit treatment of flexibility in docking</i>	85
2.3.5.2	<i>Explicit treatment of flexibility in docking</i>	87
2.4	<i>Poisson-Boltzmann solver: Delphi</i>	88
2.5	<i>MMPBSA Calculation</i>	90
2.6	<i>Softwares used</i>	92
2.6.1	<i>Gromacs</i>	92
2.6.2	<i>HADDOCK</i>	93
2.6.3	<i>TINKER</i>	95
2.6.4	<i>MCCE</i>	95
2.6.5	<i>NACCESS</i>	95
2.6.6	<i>TANGO</i>	96
2.6.7	<i>PASTA Server</i>	96
2.6.8	<i>Computational Alanine Scanning</i>	96

2.6.8.1	<i>FoldX</i>	97
2.6.8.2	<i>Robetta</i>	98
2.7	<i>Limitations of Computational Methods</i>	98
2.7.1	<i>Limitations of MD Simulation</i>	98
2.7.2	<i>Computational Alanine Scanning for Hot Spot and Anchor Residue prediction</i>	101
2.7.3	<i>Docking: Protein-Protein complex prediction</i>	102
2.7.4	<i>MMPBSA Binding Energy</i>	103
2.7	<i>References</i>	104
Chapter-3	<i>Rationalization of poor solubility of TGF-β3 using MD Simulation</i>	111
3.1	<i>Introduction</i>	113
3.2	<i>Material and Methods</i>	113
3.2.1	<i>Preparation of 10M Urea solvation box</i>	115
3.3	<i>Results and Discussion</i>	116
3.3.1	<i>Standard MD simulations to characterize the intermediate during unfolding of TGF-β1 and TGF-β3</i>	116
3.3.2	<i>MD study of unfolding of TGF-β3 in 10 M Urea</i>	121
3.3.3	<i>MD simulation of the peptide fragment of H3-Helix</i>	123
3.3.4	<i>MD simulation of peptide fragment of H3-Helix at different pH</i>	125
3.3.5	<i>Presence of hydrophobic patch on the protein surface</i>	127
3.3.6	<i>Charge calculation</i>	128
3.3.7	<i>Hydrophobicity/hydrophilicity</i>	129
3.3.8	<i>Presence of hydrophobic residues on the protein surface</i>	129
3.3.9	<i>Aggregation Prediction Server</i>	130
3.4	<i>Conclusions</i>	133
3.5	<i>References</i>	134
Chapter-4	<i>Conformational preference of Transforming Growth Factor β1 and β3 for its open versus closed conformation</i>	137
4.1	<i>Introduction</i>	139
4.2	<i>Materials and Methods</i>	141
4.2.1	<i>Docking on Haddock</i>	141
4.2.2	<i>Docking using Rosetta</i>	143
4.2.3	<i>Free Energy of Binding using MMPBSA</i>	144
4.2.4	<i>Dynamics of Interfacial residues</i>	145
4.2.5	<i>Effect of Ala substitution on Binding Energy</i>	146
4.3	<i>Results and Discussion</i>	146
4.3.1	<i>Haddock docking of the TGF-β3 monomer</i>	146

4.3.2	<i>Rosetta Docking of TGF-β3 monomers</i>	156
4.3.3	<i>MMPBSA Calculation</i>	160
4.3.4	<i>Identification of Anchor Residues in the TGF-β Homo-dimers</i>	161
4.3.5	<i>Identification of Hot spot residues at the interface</i>	163
4.4	<i>Conclusions</i>	166
4.5	<i>References</i>	167
Chapter-5	<i>Specificity in the interaction of the TGF-β1 and TGF-β3 with receptor I and receptor II</i>	171
5.1	<i>Introduction</i>	173
5.2	<i>Materials and Methods</i>	174
5.2.1	<i>Electrostatic Interaction Energy using Delphi</i>	175
5.2.2	<i>Free Energy of Binding using MMPBSA</i>	176
5.2.3	<i>Computational alanine scanning</i>	177
5.3	<i>Results and Discussion</i>	177
5.3.1	<i>Interfacial residues of ligand and receptors in the ternary complex are almost similar except at the interface with TβR1:</i>	177
5.3.2	<i>There are subtle, yet important, differences in interaction of receptors with ligand or the other receptor at the interface in two ternary complexes</i>	180
5.3.3	<i>Anchor and hot residues of ligand and receptor II are similar in TGF-β3 and TGF-β1 complex</i>	187
5.3.4	<i>Anchor and hot residues of ligand and receptor I are similar at TGF:TβR1 interface in TGF-β3 and TGF-β1 ternary complexes</i>	198
5.3.5	<i>Anchor and hot residues of receptor I are markedly different at TβR1: TβR2 interface in TGF-β3 and TGF-β1 complex</i>	199
5.3.6	<i>pH modulates the affinity of TβR2 to the open conformation of TGF-β3</i>	202
5.3.7	<i>Formation of binary to ternary complex is less favored by electrostatic interaction in TGF-β3 at pH 7</i>	203
5.3.8	<i>pH modulates the affinity of TβR1 to binary complex in two ternary complexes</i>	204
5.3.9	<i>Comparative Binding affinity of TβR2 and TβR1 for TGF-β3 and TGF-β1 using MMPBSA</i>	207
5.4	<i>Conclusions</i>	209
5.5	<i>References</i>	211
Chapter-6	<i>BMP2 ligand binding with its receptor II and receptor I: A comparative analysis with TGF-β3 and TGF-β1</i>	213

6.1	<i>Introduction</i>	215
6.2	<i>Material and Methods</i>	217
6.3	<i>Results and Discussion</i>	217
6.3.1	<i>Few non-conserved residues of receptor II are present at the interface of ligand and receptor II in both complex of BMP-2</i>	217
6.3.2	<i>Similar residues are present at the interface of ligand and Receptor I</i>	218
6.3.3	<i>Different Hydrogen bonds and similar hydrophobic interactions at the interface are present in these ternary complexes</i>	220
6.3.3.1	<i>Interactions in 2H64</i>	220
6.3.3.2	<i>Interactions in 2GOO</i>	221
6.3.4	<i>Anchor residues analysis of BMP-2 are similar in the two ternary complexes</i>	225
6.3.5	<i>Computational alanine scanning mutagenesis for prediction of Hot Residues at the interface</i>	237
6.3.6	<i>The driving force of binding of receptor II at knuckle epitope in BMP-2 in contrast to TGF- β fingertips is the buried surface area</i>	238
6.3.7	<i>MMPBSA Binding Energy</i>	242
6.3.8	<i>Electrostatic Contribution of individual residues in binding in BMP-2 ternary complex</i>	245
6.4	<i>Conclusions</i>	249
6.5	<i>References</i>	250
Chapter-7	<i>Summary and Future perspectives</i>	253
7.1	<i>Summary and Future perspective</i>	255
	<i>Biodata</i>	