

A  
MAJOR PROJECT REPORT  
ON  
**“Protein Structure Alignment Algorithm Studies”**

*Submitted to*

**DELHI TECHNOLOGICAL UNIVERSITY, DELHI**

*In Partial Fulfillment of the Requirement For  
The Award of Degree of*

**MASTER OF TECHNOLOGY  
IN  
BIOINFORMATICS**

**2010-2012**

By

**Ankita Mehta (01/BINF/2010)**

*Under The Guidance Of*

***Proff. Jooyoung Lee***

***CIPS, KIAS, Seoul, South korea***

*and*

***Dr. Yasha Hasija***

***Assistant Professor, DTU, Delhi, India***




**Department of Biotechnology**


**Delhi Technological University, Delhi (Formerly known as DCE)**

## CERTIFICATE

This is to certify that the project entitled "Protein Structure Alignment Algorithm Studies" submitted by Ankita Mehta to the Delhi Technological University, Delhi in partial fulfillment of the requirement for the major project as a part of curriculum of fourth semester of Master of Technology degree in Bioinformatics (Science) is a record of bona fide research work carried out by her under my supervision and guidance at the Korea Institute for Advanced Study (KIAS) in Seoul, South Korea and no part of this work has been submitted for any other degree or diploma.

University Guide  
(Dr. Yasha Hasija)  
DTU  
Delhi, India

Professor July 9<sup>th</sup>, 2012  
(Jooyoung Lee)   
Center for In Silico Protein Science  
KIAS, Seoul, South Korea



# DECLARATION

I hereby declare that the work presented in this thesis has been carried out by me under the supervision of Proff. Jooyoung Lee CIPS, KIAS, Seoul, South Korea and Dr. Yasha Hasija, Department of Biotechnology Delhi Technological University, Bawana Road, Delhi and has not been submitted for a degree or diploma of any other university. All assistance and help received during the course of the investigation has been duly acknowledged.

Ankita Mehta

## *ACKNOWLEDGEMENT*

*“To speak gratitude is courteous and pleasant, to enact gratitude is generous and noble, but to live gratitude is to touch Heaven”*

*I would like to express my heartfelt gratitude to all those who have contributed directly or indirectly towards obtaining my master degree. I am highly indebted to my esteemed supervisor, Proff. Jooyoung Lee and Dr. Yasha Hasija who has guided me through thick and thin. Many people have contributed to this project in a variety of ways over the past few months. I also acknowledge the many helpful comments received from Dr. Bala Chandran Manavalan, research fellow at KIAS, Seoul, South Korea. I am indebted to all those who provided reviews & suggestion for improving the result and the topics covered in my project.*

Ankita Mehta

Contents	Page number
1. Abstract	1
2. Introduction	2-3
3. Definitions	3
3.1 Units of Structure Descriptions	3
3.2 Distance Matrices	5
3.3 Torsion Angles	7
3.4. Helices	8
3.5 Strands and Sheets	8
3.6 Identifying the SSEs	11
3.7 Protein Domains	12
3.8 Folds	17
3.9 Structure Comparison	18
3.10 Superposition and Dynamic Programming	20
3.11 Protein Structure Classification	22
3.12 Databases for Structure classification	23
4 Literature survey	24
4.1 DALI	25
4.2 Maxsub	29
4.3 RMS/Coverage Graphs	30
4.4 Protein structure alignment by combinatorial extension (CE)	31
4.5 TM align	33
5 Methodology	35
5.1 Idea behind TMscore	35

5.2 Modification of TM-score	36
5.3 Source Code	37
6. Results and Discussion	69
6.1 Original TMscore	69
6.2 Modified TMscore	69
6.3 Discussion, Conclusion and Future Work	70
7. References	71
8. Appendix	

