Name of the Research Scholar: **SYED AUSAF ALI**
Name of the Supervisor: **PROF. FAIZAN AHMAD**
Name of the Co-supervisor: **DR. ASIMUL ISLAM**
Name of the Centre: **CENTRE FOR INTERDISCIPLINARY RESEARCH IN BASIC SCIENCES**

Title of the thesis: **PREDICTING PROTEIN STABILITY FROM ITS ATOMIC COORDINATES IN PROTEIN DATA BANK (PDB): AN EMPIRICAL APPROACH**

ASTRACT OF THESIS

Proteins are movers and shakers of the cell. Proteins and enzymes are the backbone of many important commercial products like food stuffs, cosmetics, soaps, detergents, drugs, medicines, fabrics, so on and so forth. In view of that, predicting protein stability is of immense use and applicability in applied scientific and industrial research. In principle, the conformational protein stability can be defined as the change in Gibbs free energy (ΔG_D) between its folded and unfolded states during denaturation process. ΔG_D can be inferred from the folding-unfolding equilibrium of the protein. The fundamental algorithm, on which our proposed prediction model is structured, is based on Tanford's thermodynamic model. According to this model, the overall change in Gibbs free energy (ΔG_D) of a protein unfolding reaction is a function of transfer free energies of the constituent groups of amino acid residues constituting that particular protein. Both the overall ΔG_D of the reactions and transfer free energies of protein residues are dependent on denaturant concentration. The correlation of transfer free energy of residues with the denaturant concentration is not linear. In order to make it linear, this model was a bit modified by implying some empirical parameters. These empirical parameters for particular denaturants (urea and GdnHCl) were calculated by fitting the values of transfer free energy for residues into a power series of differing denaturant concentrations using least square method in MATLAB®. At mid-point

of denaturation, where the number of denatured molecules are half maximal, overall ΔG_D becomes zero and change in Gibbs free energy in 0M concentration of denaturant (ΔG_D^0), can be estimated as a function of C_m (denaturant concentration at mid-point of denaturation), residue transfer free energy ($\Delta g_{tr,i}$) and fractional exposure of the residues in moving from water to denaturant solution ($\Delta \alpha_i$).

Thus, from the observations of our research project, we can conclude that

(i) Our online tool "ProStab" is able to predict the conformational stability of the globular proteins (provided their C_m value and X-ray coordinates are known).

(ii) The accuracy of our proposed model is better for the protein candidates undergoing urea denaturation than those undergoing GdnHCl denaturation.

(iii) Re-scaling of data improves both the correlation coefficients and slope of the linear fit (an indicator of one-to-one correspondence).

(iv) There is better prediction for the proteins which strictly follow a two-state process without having ambiguous data.

(v) Our tool "ProStab" based on thermodynamic model is better than other online tools (like FoldX or I-Mutant2.0) for stability prediction in terms of accuracy, robustness, universality and portability.

(vi) It is the only online tool predicting conformational stability of wild type proteins based on transfer free energy values for the constituents groups (*m* value predictor is fully erroneous).

(vii) The SASA of amino acid residues is linearly and perfectly correlated with their average residual mass, monoisotopic mass, van der Waals volume and average volume of buries residues.