Divya Kappor (2009). *Short-circuiting natural evolution through selected transplantations of residues to engineer the physical behaviour of protein surface.* Ph.D. Thesis. CSIR-IMTECH, Chandigarh/ Jawaharlal Nehru University, New Delhi: India.
Supervisor: Dr. P. Guptasarma **Acc. No. : Th-202**

SUMMARY OF THE THESIS

For the past many decades, protein engineering has been put to great use for designing new proteins and enzymes with improved, novel properties with potential application in biotechnology. Virtually all efforts made till date have been directed towards altering the chemical properties of enzymes, like substrate affinity, substrate specificity, catalytic turnover etc. To the best of our knowledge, no efforts have been made to change any physical parameters of protein function, such as the temperature of optimal enzyme activity, except some protein structural domain-swapping experiments. Therefore, in the present thesis, we have described the invention and application of a novel rational engineering strategy likely to be applicable to all beta sheet-based proteins, for the mixing and matching of traits of homologous proteins taken from different domains of life. In nature, mesophile organisms produce meso-active and meso-stable enzymes while thermophile or hyperthermophile organism produce only thermo-active and thermo-stable enzymes. Nature normally does not produce enzymes which are meso-active but extremely thermo-stable.

It is well known that enzymes/proteins performing similar/same type of chemical reaction tend to have similar polypeptide backbone structures, and also similar amino acid sequences. The physical properties of such 'homologous' enzymes/proteins are determined, however, by their outer shape characteristics, or surface features, since it is the surface (consisting mostly of sidechain atoms) that interacts with other molecules, while the polypeptide backbone remains largely hidden. Thus, two proteins that are evolutionarily unrelated can potentially have almost identical polypeptide backbone structures, with almost insignificant sequence identity/similarity, and yet have very different surface features, due to the differential decoration of their backbones by interacting groups of side-chains of amino acid residues that differ in details amongst homologous proteins. This consideration prompted us to explore the feasibility of transplanting/grafting surface features of one protein [illegible] the structural scaffold of another homologous protein/enzyme with a similar backbone [illegible] and this novel rational strategy that we proposed for the engineering of beta-sheet [illegible] proteins/enzymes may be referred to as "Surface Remodeling" or "Surface Grafting" [illegible] transplantation".

[illegible] the literature, the concept of the enzyme "active site" is quite well known. We [illegible] introduced (in particular, for enzymes that bind long polymeric chain substrates) a new

term called "active surface". The main difference between the two terms is that the "active site" defines only the key catalytic residues while the "active surface" refers to not only residues involved in catalysis, but also all residues involved in enzyme-substrate contacts. It is our hypothesis that the nature, and temperature-dependence of microconformational flexibility of the interactions of the substrate with the entire active surface (and not just with the active site) determine enzyme functional features such as the temperature of optimal function. Therefore, it is our hypothesis that the transplantation/grafting of the active surface, or the whole surface, between homologous enzymes in functional form may result in the creation of novel enzymes with the structural core, and stability characteristics of one enzyme/protein progenitor, and the functional characteristics of a different enzyme/protein progenitor sharing good polypeptide backbone structural homology. This technique of surface grafting was introduced by us successfully into the field of enzyme engineering, in the work presented in this thesis, and it has provided us with new insights and also opened up new avenues for designing novel enzymes/proteins of commercial significance with desired features. This approach has great implications both in basic science, from the viewpoint of gaining new insights and in applied science, from the viewpoint of designing novel enzymes with desired features required by the industry.

The entire thesis is divided into seven chapters with references mentioned at the end of each chapter. **Chapter 1** provides a very brief introduction to the field of protein engineering, by summarizing all the techniques/strategies used, till date, to design enzymes with improved properties. **Chapter 2** describes in detail the materials and methods commonly applicable to all the aspects of the present study while "materials and methods" unique to each part of the study have been detailed individually in a section "Experimental strategies and design" in each chapter.

Chapter 3 titled: **Design of a meso-active thermo-stable Cel12A through "Active Surface Grafting"** describes in detail our new strategy, applicable to beta-sheet based protein structures, for the rational design of novel enzymes with desired stability ($T_{m)}$ and activity (T_{opt}) characteristics. The chapter demonstrates the successful creation of the first such enzyme made, called "MT Cel12A" (for meso-active thermo-stable Cel12a) with structural stability features of one enzyme and activity features of another homologous enzyme. The significance of this is as follows: The temperatures of optimal activity (T_{opt})

and structural melting (T_m) of any enzyme are invariably very close, differing by only a few degrees Centrigrade. Dogma in enzymology holds, therefore, that enzyme activity must double with every 10 °C rise in temperature whilst global 3-D structure is retained. Here, however, we demonstrate the complete uncoupling of T_{opt} from T_m by our novel technique (functionally akin to skin-grafting in cosmetic surgery). Using several tens of rationally-selected mutations of non-contiguous residues, we have remodeled the entire solvent-exposed face of the substrate-binding and catalytically-active beta sheet-based groove of a thermophile cellulase, *Rhodothermus marinus* Cel12A, and caused it to resemble instead the groove of *Trichoderma reesei* Cel12A, a known mesophile structural homolog. X-ray crystallographic structure determination and stability/activity tests establish incontrovertibly that the enzyme bearing the skin-graft displays the functionality of the *T.reesei* enzyme and the stability of the *R.marinus* enzyme. The T_{opt} is effectively lowered by 35 °C, from 90 °C to 55 °C, whereas the T_m remains $>$ 90 °C. We believe that this successful first application of a protein skin-grafting technique to dissect and independently assort the T_{opt} and T_m properties of two different enzymes establishes the scope for future 'mix-and-match' creations of many novel enzymes, using proteins sourced from different domains of life.

Chapter 4 titled: **"Using DNA sequencing electrophoresis compression artifacts as reporters of stable mRNA structures affecting gene expression"** describes how compression artifacts in DNA sequencing can act as important clues for determination of translation-suppressing secondary structural elements in mRNA, especially in DNA/RNA encoding thermostable enzymes. It is a well known fact that the formation of secondary structure in DNA leads to 'compression artifacts', while secondary structure in mRNA can suppress translation. In the present study, we show how formation of secondary structure within the transcript of an engineered version of a thermophile gene sourced from *Rhodothermus marinus* introduced into *E. coli* manifests as a marked compression artifact in sequencing, thereby resulting in extremely poor gene expression, and how introduction of silent mutations at selective locations in the gene sequence destroys the compression effect, resulting in a tremendous enhancement in gene expression. Moreover, it is known in literature that these compression artifacts are dependent on the temperature at which electrophoresis is conducted. Therefore, we have proposed that identification of problem

regions of sequence affecting gene expression at the level of translation may be routinely aided by DNA sequencing electrophoresis at lower temperatures.

Chapter 5 titled: **"Active site transplantation in TET aminopeptidase"** describes how transplantation of the active site alone (rather than the entire active surface involved in substrate-binding and catalysis) can affect chemical parameters associated with the enzyme activity of the progenitor proteins, such as the rate of catalysis, the affinity for the substrate, and the nature of substrate-based inhibition. The TET aminopeptidase of *Pyrococcus horikoshii,* PhoAP, is a well characterized enzyme of known structure, whereas only the 3D structure of the homologous *Bacillus subtilis* enzyme, BsuAP, is known. We produced and characterized BsuAP and a PhoAP-homolog from *Pyrococcus furiosus*, PfuAP. The temperature/pH of optimal function of BsuAP and PfuAP were determined to be 70 °C/7.5 and 80 °C/8.0, respectively. BsuAP was twice as active as PfuAP, and displayed a 4-fold higher Km, in addition to severe substrate-based inhibition above a concentration of >5 mM substrate (unlike PfuAP). We made 9 mutations in PfuAP's active site to transform it into BsuAP's active site. The mutant, MutAP, showed a ~185-fold reduction in K_{cat}, a reduced temperature of optimal function (60 °C), a Km (3.8 mM) similar to that of PfuAP (5.0 mM), and substrate-based inhibition like BsuAP, suggesting that active site transplantation can help recombine enzyme functional characteristics in interesting ways.

Chapter 6 titled: **"Creation of a new eye lens crystallin through "Whole Surface Transplantation"** demonstrates the successful creation of a novel crystallin "Gambeta" through rational transplantation of the entire surface of βB2-crystallin upon the structural [illegible] of γB-crystallin, without altering the latter's interior. This work is an extension from grafting of "active surfaces" to grafting of "whole protein surfaces". This new protein, Gambeta, consists of: (a) 61 residues possessing the same identity at structurally-equivalent positions in βB2- and γB-crystallin; (b) 91 surface residues unique to βB2-crystallin; and (c) [illegible]ior residues unique to γB-crystallin. The protein incorporating these alterations, [illegible] was found to fold and display : (a) some characteristics unique to βB2-crystallin, [illegible]imeric quaternary structure, an identical near-UV CD spectrum, and maximal [illegible] emission at 334 nm; (b) other characteristics unique to γB-crystallin, such as an [illegible]-UV CD spectrum, precipitation upon cooling, aggregation upon heating, and [illegible] to chemical denaturation; (c) yet other characteristics common to both, such as

binding of the dye Stains-all, and solubility at concentrations approaching ~300 mg/ml, with persistence of transparency and high refractivity; and (d) characteristics found in neither, such as precipitation at low concentrations of guanidium hydrochloride. The folding of a protein with such a 'patchwork' residue ancestry suggests that interior/surface transplants involving all-beta proteins are a feasible engineering strategy.

Chapter 7 titled: **"Engineering design of a triple enzyme fusion cellulase" describes** an attempt to design a novel three enzyme fusion protein called "Fusion Cellulase" by fusing only the catalytic domains of three different cellulolytic enzymes (endoglucanase, exoglucanase and beta-glucosidase), sourced from different organisms, with the capability to degrade cellulose in a synergistic and cooperative manner with high efficiency. Cellulose forms the most abundantly available renewable biomass on earth and the degradation of cellulose to sugars is the rate limiting step in conversion of this biomass into utilizable forms of energy like ethanol. Ethanol can act as biofuel, thus, making this attempt of engineering "fusion cellulase" a worthwhile objective. The chapter describes the production of this fusion cellulase in *E.coli*.