

Research Interests: Protein stability, folding and design

Research goals

1. Understanding and quantitating protein stability.
2. Engineering proteins for increased stability.
3. Engineering proteins for decreased stability.
4. Mechanism of action of chaperone SecB

Experimental systems: RNase A, thioredoxin, SecB, MBP, CcdB, Gal4, HIV-1 gp120.

Techniques used : Molecular biology, calorimetry, X-ray crystallography, NMR spectroscopy, computational analysis

Proteins are biopolymers composed of 20 different types of naturally occurring monomer units called amino acids.

Proteins vary in length from about 50-500 amino acids.

A given protein is characterized by a specific amino acid sequence and a specific three dimensional shape.

The sequence of the protein contains sufficient information to specify the shape. However given the sequence of a protein, it is not currently possible to predict the shape from first principles.

My laboratory works on understanding the relation between protein sequence, protein shape and protein stability. We also apply this knowledge to design proteins for specific biological applications.

Protein sequences can be deduced from corresponding gene sequences. With the advent of genome sequencing, protein sequences for complete genome for several organisms, including humans are now known.

It is obviously important to be able to determine the structure and function of these proteins.

Gene knockouts are a powerful tool to study protein function. However this technique is difficult to use for essential genes.

In such cases it is very useful to make conditional mutants (mutants which function like the wild type protein under some conditions, but which are non-functional under other conditions where the wild type protein is fully functional).

Ts mutants are typically generated by random mutagenesis of the gene or entire genome followed by laborious screening procedures.

We have developed methods for prediction of possible Ts mutants solely from amino acid sequence (Varadarajan et al, PNAS 93, 13908-13 (1996), Chakshusmathi et al, PNAS, 101:7925-7930 (2004)).

The method indicates a small number of residue locations that are likely to be highly buried inside the protein and a suggested list of substitutions that will generate Ts mutants at a reasonable frequency.

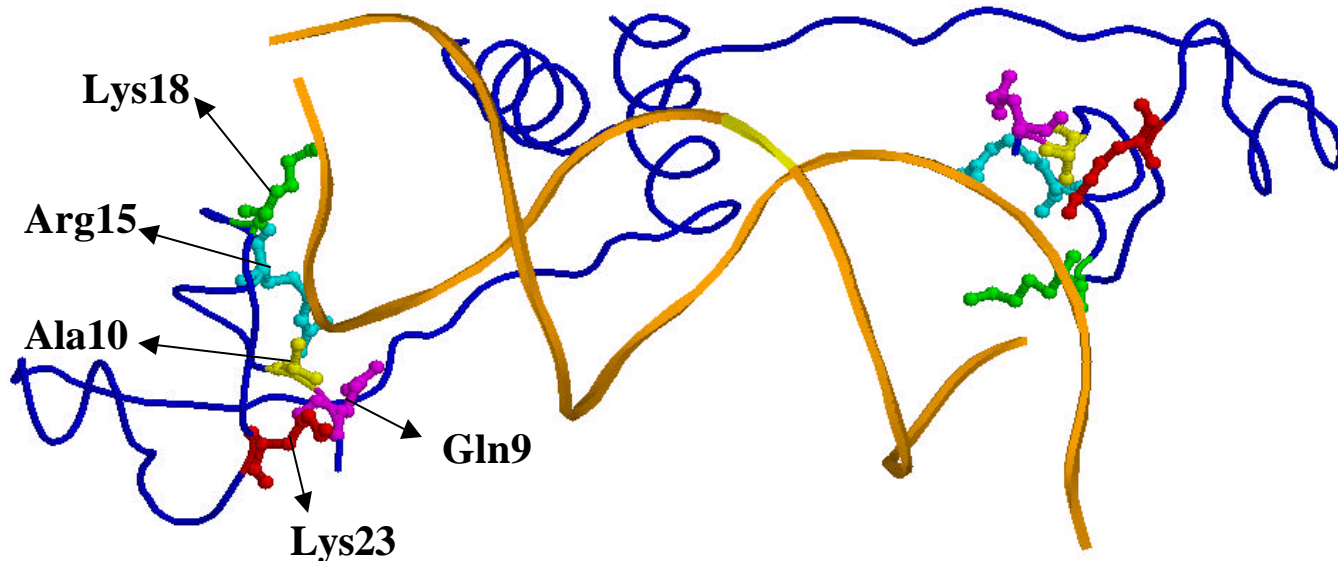
Is the Ts phenotype transferable between organisms?

Little is known about the transferability of the Ts phenotype from one organism to another. To this end, several Ts mutants of the yeast transcriptional activator Gal4 were isolated in yeast.

Gal4 was chosen because it is widely used for ectopic expression of transgenes in higher eukaryotes which lack endogenous Gal4.

Ts mutants screened in yeast have been cloned into Drosophila P element vectors. Transgenic flies were obtained and are currently being screened

- **The yeast protein Gal4 activates gene expression from genes downstream of specific DNA sequences.**
- **The DNA sequences (abbreviated UAS for upstream activating sequence) recognized by Gal4 are 17 base pairs (bp) in length, and each site binds a dimer of protein.**
- **Two different approaches were taken to generate Ts mutants of Gal4.**
 - a) **Residues involved in DNA binding were randomized and mutants screened for a Ts phenotype.**
 - b) **Putative buried hydrophobics were mutated as described above for CcdB and mutants screened for a Ts phenotype**
- **Yeast reporter strains used are deleted for *gal4* and *gal80*. Genes for histidine and adenine biosynthesis and for lacZ are placed downstream of tandem UAS sites**



Structure of DNA binding domain (amino acids 1-65) of Gal4 bound to its target upstream activating sequence (UAS) (Marmorstein et al, 1992). Residues, which have been mutated, are labeled.

Mutation based definitions and probes of residue burial in proteins

Methods

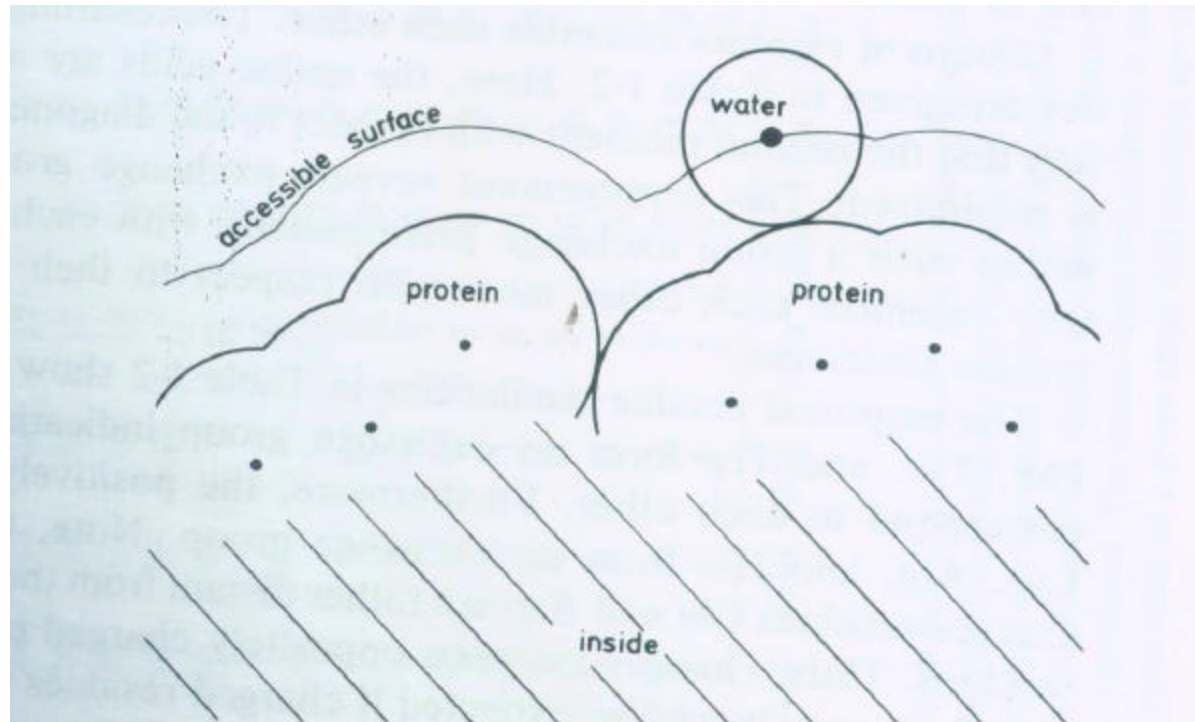
A high throughput 96 well based mega primer method was used to replace each of 101 residues of CcdB with Ala, Asp, Glu, Lys, Arg and Pro.

Goals

Examining the correlation between residue burial and activity.

Formulation of rules relating protein sequence, structure and function.

Schematic representation of accessible surface area



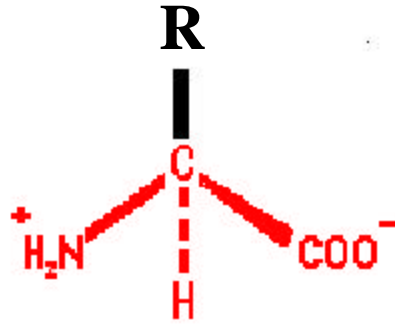
% Residue accessibility =

Lee and Richards; 1971

Accessible surface area of a residue in protein

Accessible surface area of residue in Gly-X-Gly conformation

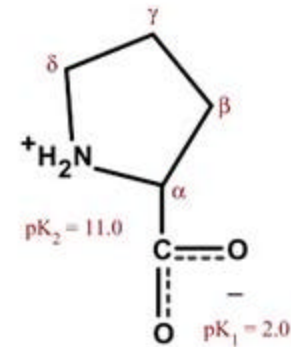
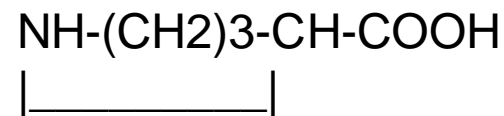
General Structure of an amino acid at pH 7.0



Where R =

- CH₃ for Ala,
- CH₂COO⁻ for Asp,
- CH₂CH₂COO⁻ for Glu,
- (CH₂)₄NH₃⁺ for Lys,
- (CH₂)₃-NH-(NH₂)C=NH for Arg

Proline:



Experimental test system: CcdB protein of F plasmid.

CcdB (control of cell death or division) is an inhibitor of DNA gyrase and is toxic to *E. coli*. It is a homodimer of 101 aa residue subunits.

Hence screening for Ts mutants of the protein is quite straightforward.

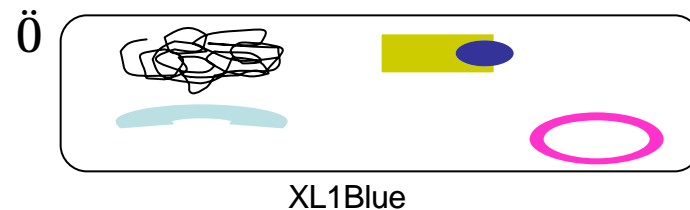
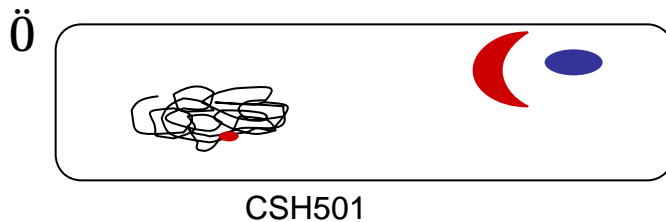
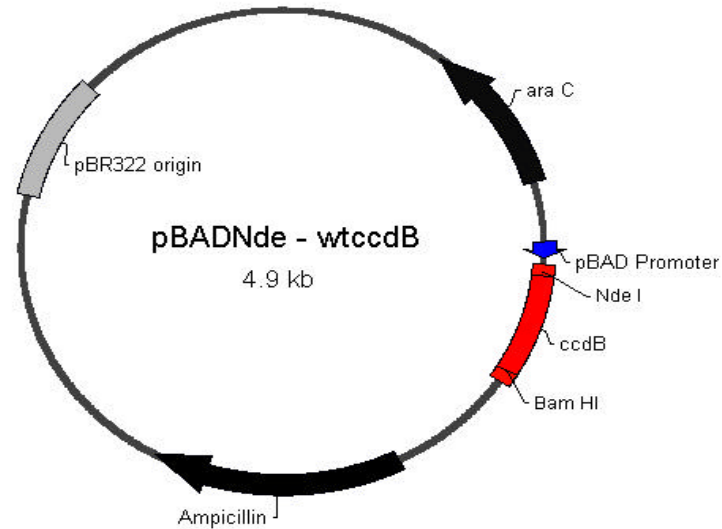
Active protein: Cells transformed with CcdB expressing plasmid die at all temperatures.

Inactive protein: Cells survive upto 42°C

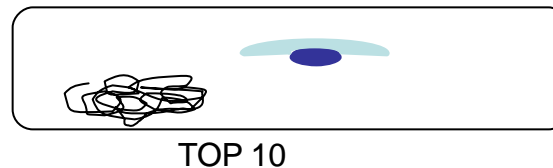
Ts mutants: Cells grow at higher temperatures (42°C), but not at a lower temperature (say 30°C).

Plasmids and host strains

CcdB, Controller of cell division or death, is 101 residue protein existing as a symmetric homodimer at physiological pH. (Loris *et al.*) Each CcdB monomer has a β -pleated sheet structure with single helix at the C-terminus. It is an *E.coli* gyrase inhibitor



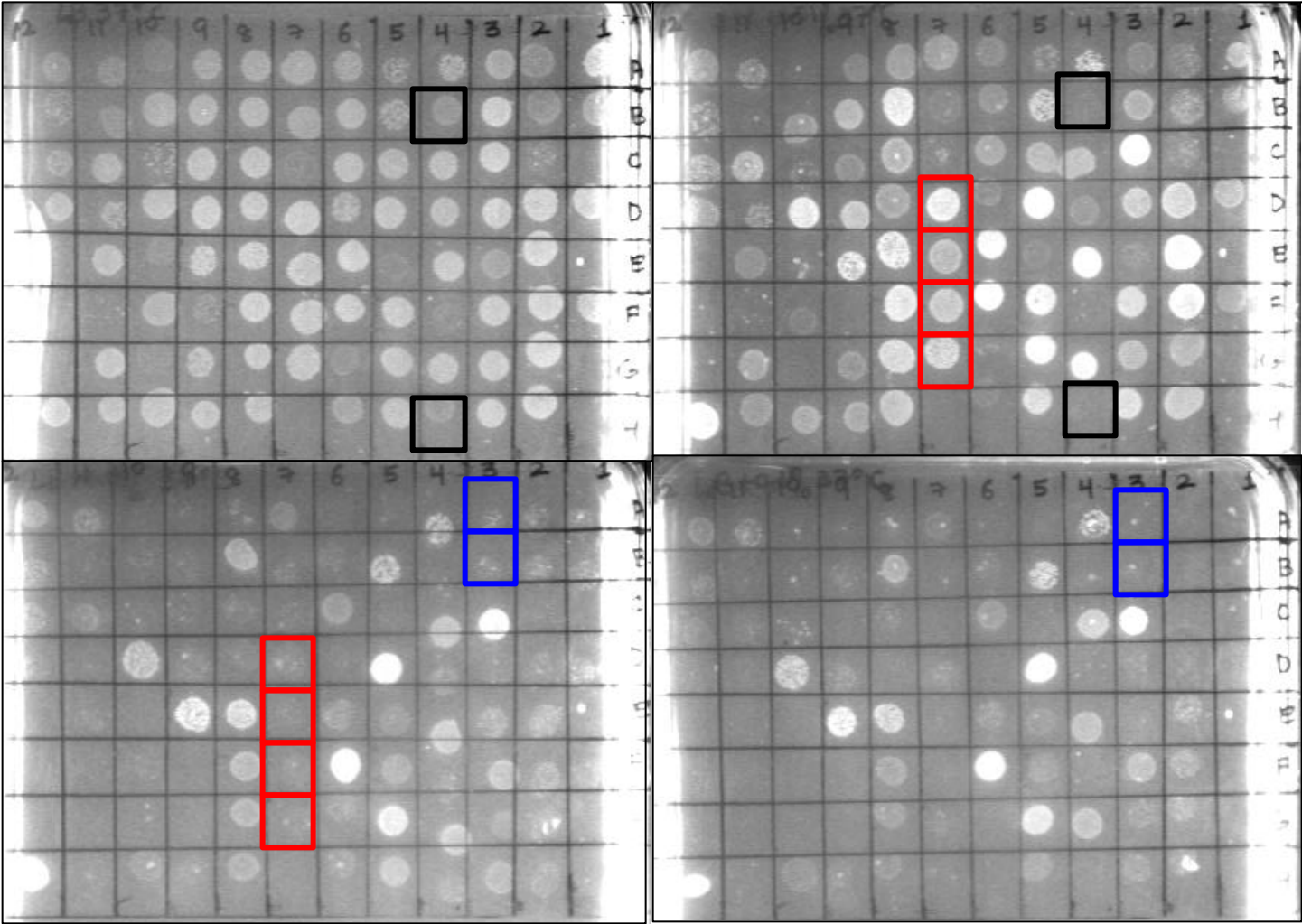
X



Phenotypic screening of CcdB scanning mutants as a function of arabinose

0% ara

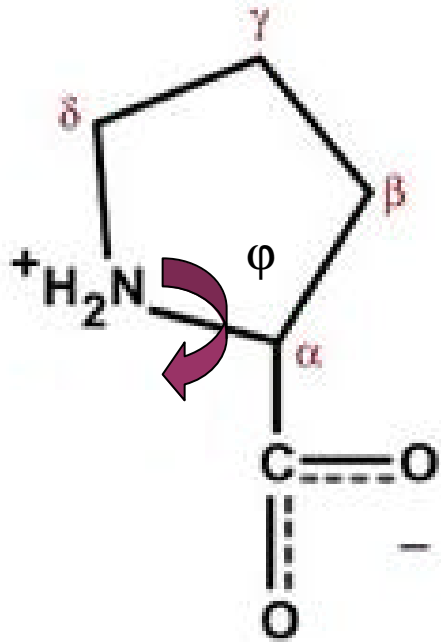
0.001% ara



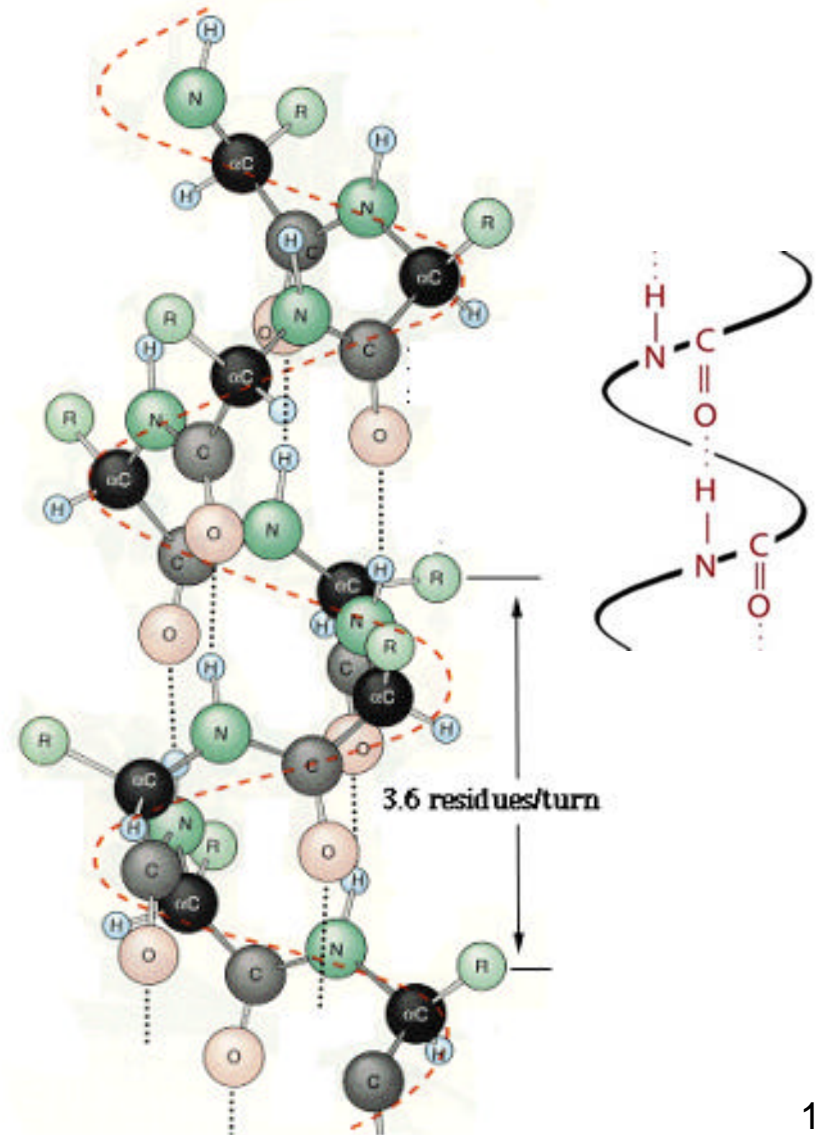
0.01% ara

0.1% ara

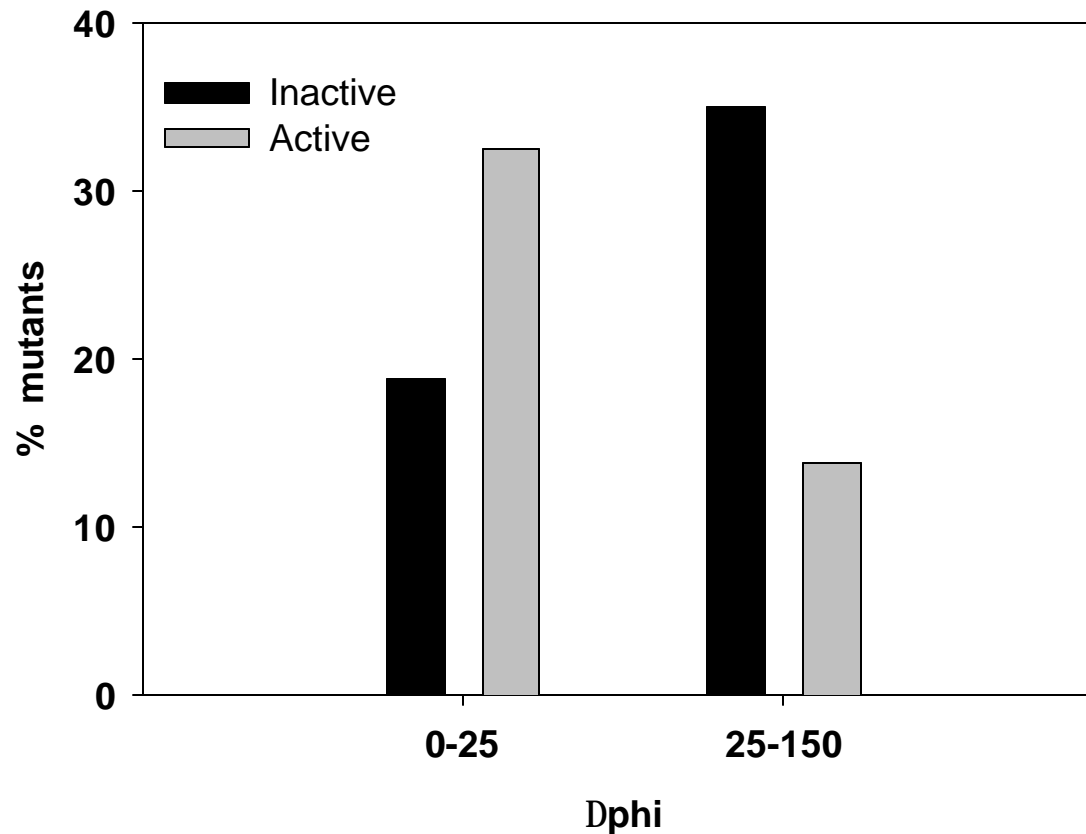
Proline as helix breaker?



where, $\phi(\phi) = -63 \pm 15$



Distribution of inactive and active proline substitutions as a function of $\Delta \phi$



$\Delta \phi = \text{Phi value of a residue} - \text{av. phi value of Pro } (-63 \pm 15)$

Acknowledgements

Kanika Bajaj
Kajari Mondal

Antara Ghosh Dastidar
Purbani Chakrabarti
Guramrit Singh

K. VijayRaghavan (NCBS)