Preface

The central goal in designing a novel protein or peptide structure is to devise an amino acid sequence that will adopt a unique and stable 3-dimensional structure. Although nature has provided an enormous number of natural proteins, which fold in to different variety of structures and carry out innumerable functions, there are two main motivations to pursue the protein and peptide design: (i) design is ultimate test of our understanding of natural proteins; and (ii) it represents the essential first step towards a new generation of novel structures that will have practical applications in industry and biomedicine. Among several approaches that have been tried, the one that emerged is *De Novo* design - design of peptides and proteins from scratch - address the problem in steps, *i.e.*, first build modules of secondary and super secondary structures, and then assemble them in to 3-D structures of desired conformation and function. The available tools to design are not just limited to conventional 20 genetically-coded amino acids, but comprises various mutated, constrained or restrained, flexible amino acids or even templates. In building the modules of secondary structures, the first problem lies in finding the conformational preferences of these non-standard amino acids (also called non-coded amino acids), and their potency in dictating the conformation of peptides in presence of neighboring residues. In our work, we have made an attempt to address these issues, focusing primarily on conformational considerations (Chapter 1).

Conformationally constrained mutants offer a good means to design specific folding motifs, and to develop new pharmaceutical agents. Our investigations were focussed on the potential bioactive analogs of inflammation response inducing N-formyl peptide, N-formyl-L-Methionyl-L-Leucyl-L-Phenylalanyl-methyl ester (fMLF), which stimulates a wide range of functions from chemotaxis and lysosomal enzyme release to super oxide generations. In view of the current model on structure- function relationship of chemotactic factors, which suggest two widely different conformation -folded & extended - to be biologically

active, three peptides of fMLF family were investigated (Chapter 2): N-formyl-L-Methionyl-1-Amino 1-cyclooctane carbonyl (Ac_8c)-L-Phenylalanyl-methyl ester (fMACF); N-formyl - L-Methionyl- Dipropylglycyl-L-Phenylalanyl-methy ester (fMDF); & N-formyl-L-Methionyl-L-Leucyl-L-*paraiodo*Phenylalanyl- OH (fMLIF), using local cyclic and linear constraints, or modifications, at position 2 and 3. Conformational studies of peptides indicate that fMACF and fMLIF have folded conformation, while fMDF assumes extended backbone conformation, both in crystalline state and solution. Peptides in these conformations demonstrate significant chemotactic and secretagogue (lysozyme secretion) activity. The high biological activity of peptides with widely different conformations may possibly be explained from *induced-fit* theory *i.e.*, there is a possible change in the conformation of peptides on receptor binding. The molecular dynamics study performed on fMDF with a view to observe other conformation accessible to peptide, does indeed demonstrate that folded conformation is accessible to molecule at room temperature. Dipropylglycyl residue undergoes a conformational transition from fully-extended to relatively high energy C^7 (axial and equatorial)-state. The induction of a particular conformation in a residue by the molecule as a whole, which would not be obtained in the isolated residue, it can be inferred that receptor would induce otherwise unfavorable conformational states in a single residue or parts of peptide. The induction would take place by minimizing the free energy of the peptide-receptor complex as a whole.

Other conformationally important peptides from linear- and cyclic- disubstituted glycine have been studied in Chapter 3. In the *de novo* design, long peptides have been studied for stability of regular secondary structures, while small peptides rich in non-coded amino acids are examined to address such issues as conformational variability, sequence dependent conformational preferences, design and stereochemistry of linker segments. These results are useful in developing molecular guidelines for peptide design. Boc-Aib-Ac_8c-Aib-OMe (bUACU) is one such peptide, which incorporates 1-Amino 1-cyclooctane carboxylic acid (Ac_8c). It forms type -III(III') β-turns in crystal. Examining the conformational preference of Aib-Xaa-Aib sequence revel that similar to homo-oligomeric sequence of Aib, it also has a strong tendency to make type-I or -III β -turn. Another molecule, Boc-L-Val-Aib-Gly-L-Leu-OMe (bVUGL), which has achiral sequence Aib-Gly at the middle that has strong tendency to make β-turn, adopts consecutive type II-I' β-turn conformation.

As opposed to constrained residues, ω-amino acids are another important molecular tools, which relish high degree of flexibility, due to polymethylene spacer in between N and C^α

atoms, and are useful for generating several new kinds of sheets, turns or helices; sometimes as a linker in the design of super secondary structures. In Chapter 4, structure of Boc-β-Ala-L-Leu-Aib-L-Val-OMe (bBALUV) has been reported, consisting of β-Alanine, which is the first member of this class. The sequence bBALUV adopts an open β-turn conformation in solid state.

Limited success, however, were achieved with constrained residues, as they restrict the peptide backbone only in the helical region of conformational space (Ramachandran map). Needless to say, β-turn can also be enforced with shorter sequences. Extended structures, such as β-sheets, remain largely un-explored in the absence of any suitable candidate, which compulsorily give rise to β-sheet structures. They are poorly packed in crystal, and tend to aggregate therefore do not exist in isolation, which have limited their study in solid or solution state. We attempted to analyze those peptides which forms β-sheets in crystals (Chapter 5). With solid state studies on one of the peptide Z-L-Ala-L-Ala-L-Leu-pNA (ZAALN), four different conformers of which aggregate to form anti-parallel β-sheet in a crystallographic asymmetric unit, it was found that non-covalent π (aromatic) - interactions such as $\pi \cdots \pi$ & $\sigma \cdots \pi$ interactions, can play a very important role in stabilizing the packing of sheets in crystal. The high flexibility of peptide, as evidenced from four different conformers co-crystallized in a asymmetric unit, is also demonstrated in molecular dynamics study. In simulation experiments, which was carried out separately for individual conformers for 500 ps, peptide chain adopts several conformational states.

In conjunction with these studies, crystal structure of two peptide hydrates : L-Tryptophyl-Glycine Monohydrate (WGH); and Glycyl-L-Glutamine Monohydrate (GQH) were also examined and have been discussed in Appendix A. In the crystal structures of hydrates, water molecules are heavily involved in hydrogen bonding and form many three-center hydrogen bonds including *chelated bonds*. Water molecules, in WGH crystal, aggregate to form a channel along c-direction. Such water-rich crystals are interesting, which offer an atmosphere similar to solution around molecules.

Methodologies: We obtained the results using single crystal X-ray crystallographic and computational (molecular dynamics) techniques. Synthesis, solution and chemotactic activity studies of peptides were carried out elsewhere as a part of collaborative work. Peptides with standard amino acids alone, were commercial products. Diffraction Studies: Intensity data were collected on *Enraf-Nonius CAD-4* and *Rigaku MSC/AFC 7s* diffractometers. Structures were solved by the application of *direct* and *Patterson* methods, and

refined by *full-matrix least square* methods. Simulation Studies: molecules were simulated for 540 picoseconds (40ps for equilibration + 500ps for data collection), using *Consistent Valence Force Field (CVFF)* after minimization of refined crystal structure coordinates. Graphic illustration, associated computations, database analyses and preparation of materials were carried out on *VAX 785, VAX 8810 (VMS) IBM (RISC/6000), SGI's (Irix 5.3), Dec-ALPHA (OSF/1), HPC200 (HP-UX 10.20) and Sun (Solaris)* workstations in the X-ray crystallography laboratory of Physics department, Bioinformatics center, and at Supercomputer Education and Research Center (SERC) in the Institute, using available softwares *SIR97, MULTAN 87, SHELX 76/86/93/97, DISCOVER, INSIGHT-II, PARST95, PLUTO 78/PLATON, PLUTON, QUEST, ORTEP-76/3, XtalGX, SHOWCASE* and few other local programs for various applications in our use.