SYNOPSIS

Integral membrane proteins (IMPs) are involved in a number of physiological and biochemical processes. The elucidation of their detailed structure would go a long way in understanding their mode of functioning. The conventional structure determination techniques, *viz.*, X-ray and nuclear magnetic resonance (NMR) spectroscopy have not been very successful in the case of IMPs. However, success has been achieved in a few cases, for *eg.*, X-ray studies have resulted in the atomic detailed structure of photosynthetic reaction center from bacteria and Porin from the outer membrane of gram negative bacteria, sustained efforts in developing and applying electron cryomicroscopy have resulted in the atomic resolution structure of bacteriorhodopsin (bR). However, a wealth of amino acid sequence information, biophysical and biochemical data are available on IMPs. A large number of IMPs are postulated to have helix bundle topology. In this thesis we are concerned only with this class of IMPs. In the case of bR, the reaction centers, rhodopsin (Rh) and a few other IMPs, experimental studies have clearly shown the existence of transmembrane helices (TMHs). Computer simulation and modeling can use the experimental data available on the IMPs to generate plausible model at atomic detail for these proteins.

A successful model can be built if accurate information obtained from experiments or generated by computation is provided as input at various stages of modeling. For instance, several steps are involved in the modeling of IMP helix bundles. It is desirable to know the detailed helical structures flexibility and rigidity at different segments and possible alternate conformations of the constituent helices. The inter-helical orientation is related to helix-helix packing and the protein folding problem. Special attention should be paid to the environment of functionally important residues in the IMP. Experimental data
on the structure and dynamics of the TMHs which constitute the IMPs are scarce. Hence, the first and the major part of this thesis is devoted to the characterization of the structure and dynamics of a large number of TMHs by molecular dynamics (MD) simulation. In the second part, the helix-helix packing in the IMPs are analysed from the coordinates of the known structures of bacterial reaction center from *Rhodobacter viridis* (PRC), *Rhodobacter Sphaeroides* (SPRC) and bR. The geometry of helix packing and amino acid substitutions of the residues buried at the helix-helix interface have been determined from this study. Preliminary analysis of modeling and minimization studies on two IMPs homologous to bR, Halorhodopsin (hR) and Sensoryrhodopsin I (sRI), are reported. A brief synopsis of each chapter is provided below.

Chapter 1 Introduction and review of literature

Chapter 2 Details of the methodology used in modeling molecular mechanics/dynamics simulations and analysis of structures are presented in this chapter.

Chapter 3 Bacteriorhodopsin, the only IMP with seven-helix bundle topology whose structure is known at atomic resolution has been used as a template to model other IMPs. In this chapter, MD studies on isolated helices of bR are reported. Three sets of helices were simulated:

(a) Each of the seven helices of bR as given by Henderson were simulated.

(b) The backbone geometry of each of the seven helices were kept as in (a) but the side chains were fixed from a backbone dependent rotamer library.

(c) To study the effect of the side chains' net charge on helix stability, helices III, IV and VI were simulated with all the residues being neutral.
The TMH simulations presented in this and other chapters were carried out in isolation for a period of 500 pico second using the AMBER 4.0 suite of programmes.

The results of the twenty simulations demonstrate the stability of the helices of bR in isolation. In particular, helix I, II, V and VII were relatively more stable than the helices III, IV and VI whose structures were influenced by the net charge on the side chains. The role of residues Gly, Ser/Thr along with Pro which results in a dynamic kink is reported. The backbone dependent rotamer library was found suitable to simulate the TMHs in IMPs.

Chapter 4 Rhodopsin (Rh) is a well-studied member of the ever-increasing family of G Protein Coupled Receptors (GPCR) and has been shown to have a helix bundle topology similar to bR though the details of the structure are not known. An extensive characterization of the structure and dynamics of individual TMHs can aid in understanding the assembly of helices in the IMP. Hence, the Rh helices were subjected to MD simulation. Each helix was modeled as an ideal right-handed α-helix with the side chains fixed from the rotamer library. The role of different residues in the stability and dynamics are reported. The role of Pro which dominated the dynamics of the helix is important as it is conserved in many of the GPCR. Helices II and III are the most stable and behave like helices I, II, V and VII of bR. Helices I, IV and VII exhibit alternate conformations similar to helix IV of bR. Helices V and VI exhibit typical Pro-containing helix behaviour as do helices III and VI of bR. The results are compared with the behaviour of bR helices and placed in the perspective of the structures of GPCR family of proteins.
Chapter 5 Proline occurs in the middle of a number of TMHs. They dominate the behaviour of the TMHs containing them as was shown in bR and Rh in the last two chapters. Several of these Pro-containing helices also contain polar residues Ser/Thr. In this chapter, the role played by these residues in the stability and dynamics of helix II of bR (with Thr9, Thr10, Pro13, Thr18 and Ser22) is investigated. Four helices,

(a) helix II of bR
(b) helix II with Thr-> Val below Pro13
(c) helix II with Thr -> Val, Ser-> Ala above Pro13
(d) helix II with all Thr-> Val and Ser-> Ala

were studied by MD. The results are interpreted in terms of its conformational properties as a function of the position of the Ser/Thr with respect to Pro and their ability to hydrogen bond to the helix backbone.

Chapter 6 The high percentage of Proline residue present in the TMHs of transport proteins indicate that they are important in the functioning of these IMPs. Cis-trans isomerization has been suggested as a possible mechanism in many cases. The introduction of cis-Pro in an ideal α-helix results in a helix-turn-helix motif. Such a conformation would result in a drastic change in the protein conformation during cis-trans isomerization. Hence, it was considered interesting to investigate if and how a straight helix can accommodate a cis-Pro. As MD studies are ideally suited to explore the conformational space, especially, the allowed alternate conformations, this problem was studied by molecular dynamics simulation. The analysis of the conformations accessed during MD simulation showed that the residues near cis-Pro can adopt alternate conformations other than the
right handed helical conformation such that the helix direction is retained and the kink produced was around 20-30°. This value of the kink angle is also observed in the crystal structure analysis of helices in globular proteins and IMPs, and also seen in the MD simulation studies of helices with all the peptides in trans conformation. This may have implications in the involvement of cis-trans isomerization in folding and/or function of IMPs.

Chapter 7 The geometry of helix packing in PRC, SPRC and bR were evaluated in terms of the inter-helical distance and relative orientation. The buried residues at the helix-helix interface were determined in terms of the percentage surface accessible area loss (PSAAL) of the residues in each helix when two helices interact with each other. The amino acid sequences of the corresponding TMHs in proteins homologous to these were aligned. The amino acid substitutions at the positions of buried residues were determined for different amount of burial in terms of PSAAL. The substitutions were chiefly among the hydrophobic residues and it was seen that all substitutions need not be size-conserving. The substitutions between PRC and SPRC are such that the helix packing geometry was very similar in both of them.

Chapter 8 Bacteriorhodopsin's structure as given by Henderson et al. was used as a template to model hR and sRI, and these structures were energy minimized. The analysis of the intra- and inter-helical hydrogen bonds which are crucial for the structure and function of these proteins are reported. The neighbouring residues of the retinal in the three structures are also reported.
Chapter 9  This chapter summarizes the work reported in the present investigation. The simulation of isolated helices of bR with different starting conformations show that they are stable and the dynamics is dominated by residues like Gly, Ser/Thr and Pro. The charged residues also influence the stability of the helices. Similar results are obtained for the helices of Rh. The effect of polar residues Ser/Thr, at various positions with respect to the Pro, on the behaviour of such helices has been characterized. The "two stage model of integral membrane folding and oligomeration" by Engelman and Popot is discussed in light of these studies. There has been no conclusive evidence to accept or reject the possible involvement of Proline cis-trans isomerization in the folding and/or function of IMPs. Using MD simulation studies we have shown that, cis-Pro can indeed be accommodated in a TMH without disrupting the overall direction of propagation of the helix. The helix-helix packing in PRC, SPRC and bR have been analysed and the geometric packing parameters along with the residue substitutions found at the helix-helix interface are reported. Preliminary studies of modeling, minimization and analysis of hR and sRI have shown that they are likely to exhibit structural interactions similar to that of bR.

Chapter 10  The thesis ends with a discussion on the future lines of work that can emerge from results presented in this thesis.