14 Structure Prediction of Membrane Proteins

Xiche Hu

14.1 Introduction

Membrane proteins play a central role in many cellular and physiological processes. It is estimated that integral membrane proteins make up about 20–30% of the proteome (Krogh et al., 2001b; Stevens and Arkin, 2000; von Heijne, 1999). They are essential mediators of material and information transfer across cell membranes. Their functions include active and passive transport of molecules into and out of cells and organelles; transduction of energy among various forms (light, electrical, and chemical energy); as well as reception and transduction of chemical and electrical signals across membranes (Avdonin, 2005; Bockaert et al., 2002; Pahl, 1999; Rehling et al., 2004; Stack et al., 1995). Identifying these transmembrane (TM) proteins and deciphering their molecular mechanisms, then, is of great importance, particularly as applied to biomedicine. Membrane proteins are the targets of a large number of pharmacologically and toxicologically active substances, and are directly involved in their uptake, metabolism, and clearance (Bettler et al., 1998; Cohen, 2002; Heusser and Jardieu, 1997; Tibes et al., 2005; Xu et al., 2005). Despite the importance of membrane proteins, the knowledge of their high-resolution structures and mechanisms of action has lagged far behind in comparison to that of water-soluble proteins: less than 1% of all three-dimensional structures deposited in the Protein Data Bank are of membrane proteins. This unfortunate disparity stems from difficulties in overexpression and the crystallization of membrane proteins (Grisshammer and Tate, 1995; Michel, 1991).

Due to the aforementioned difficulties in solving their 3D structures by X-ray or NMR, theoretical prediction is important for revealing structures and functions of TM proteins. Membranes are complex structures containing bilayers of amphiphilic phospholipids with proteins either loosely associated on the surface (peripheral membrane proteins, e.g., phospholipase A_2) or embedded (integral membrane protein, e.g., bacteriorhodopsin). This chapter deals with the structure prediction of integral membrane proteins, defined as proteins having peptide chains with substantial tertiary structure within the nonpolar region of the lipid bilayer.

Integral membrane proteins reside in the lipid bilayer that is a highly asymmetric environment consisting of hydrated ionic head groups and the hydrophobic interior region (see Fig. 14.1). Due to the constraint of the lipid bilayer, membrane proteins must adopt secondary structure conformations with closed hydrogen bonds that pair the amide hydrogen bond donor and the carboxyl hydrogen bond acceptor
of the protein main chain in the interior region of the membrane. There are only two secondary structural architectures of the protein backbone that satisfy this energetic requirement, i.e., α-helix and β-sheet. As a matter of fact, all integral membrane proteins with known high-resolution structures belong to one of the two classes: bundles of hydrophobic α-helices or β-barrels (von Heijne, 1999; White and Wimley, 1999) (see Fig. 14.2).

**Fig. 14.2** Classes of integral membrane proteins: α-helix bundle and β-barrel. (a) The structure of the light harvesting complex II (LH-II) from the purple bacterium *Rhodospirillum molischianum* (Koepke et al., 1996); (b) the structure of porin from *Rhodobacter capsulatus* (Weiss et al., 1991).