Improvement of the Transmembrane Helix Prediction System by Three-Stage Model

Takatsugu Hirokawa
hirokawa@cc.tuat.ac.jp

Seah Boon-Chieng
seah@cc.tuat.ac.jp

Shigeki Mitaku
mitaku@cc.tuat.ac.jp

Department of Biotechnology, Tokyo University of Agriculture and Technology
Nakacho 2-24-16, Koganei, Tokyo 184, Japan

Abstract

A new method to predict the transmembrane helices from amino acid sequences was developed, in which the effect of the stabilization of helices by interhelix binding was taken into account. It was assumed that there are three stages of transmembrane helix conformation: the binding to membrane surface, the formation of transmembrane core region, and the maturation of helix due to the tertiary structure formation in membrane. This method was applied to the amino acid sequences of membrane proteins whose number of transmembrane helix are given, and most transmembrane helices were truly predicted.

1 Introduction

The formation of transmembrane helices is the first process of the folding of membrane proteins. Transmembrane helices are usually very hydrophobic, because of the hydrophobic environment of membrane. However, the polar interactions between the side chains of transmembrane helices is also important for the stabilization of transmembrane helices as well as the tertiary structure formation, as revealed by the denaturation experiments of membrane proteins.[1] Therefore, the effect of the binding between transmembrane helices has to be taken into account for more accurate prediction of transmembrane helical regions according the physicochemical approach. We have devised several parameters for the secondary structure prediction of membrane proteins.

2 Methods

We have developed a new method under physicochemical consideration by assuming three-stage model of transmembrane helix formation: the binding to membrane surface, the formation of

transmembrane core region, and the maturation of helix due to the tertiary structure formation in membrane.
(1) Binding to membrane: As suggested from the amino acid sequences of signal peptides and venom peptides such as melittin, the existence of polar groups appears important for the binding to membrane.Therefore, the difference between the maximum and the minimum hydropathy in a window of nine residues was used as a parameter of the binding signal.
(2) Formation of transmembrane core region: The most important interaction for the formation of transmembrane helices is apparently the hydrophobic interaction. Therefore, among the segments that may bind to membrane we have selected segments with the hydrophobicity higher than some threshold value. Complementary assumptions were made in this stage. On the condition of the existence of very hydrophobic transmembrane helices, more hydrophilic segments may become transmembrane helices, if one of the following additional conditions is satisfied. Even when the hydrophobicity of the segments is not high enough, they may become transmembrane core, if they include a long noncharged segment. These assumptions seemed physically sound and actually improved the results of the analysis.
(3) Maturation of transmembrane helix: Some part of a transmembrane helix has amphiphilicity and become stable in membrane though the formation of the tertiary structure. In this stage, we have taken two factors into account for the determination of the ends of helices. First, the amino acid segments neighboring the core region and having periodicity of about 3.6 residues become helical in membrane. Second, the total length of a transmembrane helix including the core and the extended regions is not so different from the thickness of membrane. The power spectral density of helical periodicity was calculated by a maximum entropy method of Fourier transformation and used for the extension of the transmembrane core.

3 Results and Conclusions

This system was applied to the amino acid sequences of membrane proteins whose (a) atomic coordinates are given in Protein Data Bank : photoreaction center, bacteriorhodopsin, melittin and cytochrome c oxidase and (b) experimental topology is available from SWISS-PROT database. Amino acid sequences of soluble proteins were also analyzed as reference. As results, existence of transmembrane helices were predicted at high accuracy and the error in the ends of helices was about only two residues in average. False positive prediction was observed for several soluble proteins.

The present system is based on the physicochemical mechanism of transmembrane helix formation, i.e. the three-stage model, and may be applied to any amino acid sequence, providing the transmembrane helical regions with accuracy of about two residues.

References

[1] S. Mitaku, K. Suzuki, S. Odashima, K. Ikuta, M. Suwa, F. Kukita, M.Ishikawa, and H. Itoh, "Structure, Function, and Genetics," *PROTEINS*,22, pp.350-362,1995.