

Development of a High Performance Prediction Method for Single Spanning Membrane Proteins

T. Tsuji, S. Mitaku

Nagoya University, Japan

Membrane proteins constitute 20-25% of open reading frames in a biological genome [1]. Previously we developed a membrane protein predictor SOSUI [2] and a signal peptide predictor SOSUIsignal [3] whose web site is visited by many researchers in the world. However, this system is not good at prediction of single spanning (TM1) membrane proteins. It is a common problem to all membrane protein prediction tools. TM1 membrane proteins occupy 30-35% of membrane proteins in a genome and have various important functions.

In this study, we prepared a non redundant dataset of membrane and soluble proteins from Swissprot for developing a method for discriminate TM1 membrane proteins with the signal peptide (SP) from other types of membrane and soluble proteins.

First, we classified the dataset into soluble, single spanning and multi spanning protein by the number of transmembrane helices predicted by SOSUI ver.3. The relationship between the position of predicted transmembrane helix and the physicochemical properties around amino terminus was investigated, leading to the fact that TM1 membrane proteins with SP tend to have a transmembrane helix around carboxyl terminus and hydrophobic amino terminus.

Using this result, we can predict 85% TM1 membrane protein with SP and 72% TM1 membrane protein without SP with the high accuracy. This result improves the accuracy of past prediction tools by 15% in prediction of TM1 with SP.

TM1 membrane proteins with SP contain a very important family of receptors which bind with proteins such as growth factors. High performance prediction system for TM1 membrane proteins with SP is the first step for the prediction of receptors which will be useful for the medical application.

**Proceedings of The Joint 2nd Pacific Rim International
Conference on Protein Science and 4th Asian-Oceania
Human Proteome Organization, Cairns- Australia, 22-26
June 2008**

1. T. Hirokawa et al., Bioinformatics, 14, 378-379 (1998).
2. M. Gomi et al., CBIJ, 4, 142-147 (2004).
3. S. Mitaku et al., Biophys. Chem., 82, 165-171 (1999).

**Proceedings of The Joint 2nd Pacific Rim International
Conference on Protein Science and 4th Asian-Oceania
Human Proteome Organization, Cairns- Australia, 22-26
June 2008**