

Inner Residues in the Transmembrane Helix Bundle are More Conservative

Sukenobu Konishi¹

gs02413@si.hirosaki-u.ac.jp

Takuhiro Nishio²

nishio@hama-med.ac.jp

Toshio Shimizu¹

slsimi@si.hirosaki-u.ac.jp

¹ Department of Electronic and Information System Engineering, Faculty of Science and Technology, Hirosaki University, Hirosaki 036-8561, Japan

² Department of Physics, Hamamatsu University School of Medicine, Hamamatsu 431-3192, Japan

Keywords: structure of membrane protein, transmembrane helix bundle, conservation of residues, prediction of structure

1 Introduction

Molecular structures of the membrane proteins are of great importance to the functions of the biomembranes. However, there is a small number of membrane proteins whose three-dimensional structures are determined. The properties of the conserved amino acid residues in the transmembrane helices (TMH) are valuable to investigate for the prediction of the protein structures. It is known that a residue at a helix contact in the TMH bundle frequently changes into another residue with similar volume. Furthermore, it is noticed that volume of the residue near the center of the protein scarcely changes.

In this study, the conservation tendency of the residues is analyzed in several transmembrane protein families. It is shown that the residues are more conservative at the inner positions in the TMH bundle. The analysis presented here is also valid for the prediction of the membrane protein structures.

2 Materials and Methods

2.1 Preparation of Dataset

One protein with known structure is chosen as a query from an objective family of the transmembrane proteins. Sequence similarity of each protein with the query is calculated within the protein family using BLAST [1] for the data from SWISS-PROT Release 41. Sequences whose degree of similarity is above a fixed value are collected and considered as a dataset for the present study. The value around 35% is applied as a critical value of sequence similarity. This value is effective in distinguishing between the inside and outside of the proteins. Fragmented sequences are excluded at this step.

2.2 Degree of Conservation

The degrees of conservation of the residues in the TMH bundles are evaluated from the dataset using ClustalW [2]. ClustalW classifies each residue into four types, i.e., fully conserved, strongly conserved, weakly conserved, and not conserved.

2.3 Conservation Index and Conservation Vector

Novel quantities, ‘conservation index’ and ‘conservation vector’, are introduced for each TMH in the bundle using the result from ClustalW in order to analyze the location of conserved residues.

Conservation index is a value which indicates the general degree of conservation of one TMH as a whole. The conservation index I_c is defined as follows;

$$I_c = f \cdot F + s \cdot S + w \cdot W \quad (f + s + w = 1, f \geq s \geq w),$$

where values F , S , and W mean the ratios of conservation on TMH, i.e., fully conserved, strongly conserved and weakly conserved, respectively. Parameters f , s , and w are used to regulate the index.

The conservation vector gives the synthetic direction of the conserved residues on TMH from the center axis of TMH. The conservation vector \mathbf{C} is defined as follows;

$$\mathbf{C} = f'\mathbf{F} + s'\mathbf{S} + w'\mathbf{W} \quad (f' + s' + w' = 1, f' \geq s' \geq w'),$$

where \mathbf{F} is a vector which indicates a direction of the fully conserved residues from a center axis of a TMH. This vector is made from the sum of all unit vectors for the fully conserved residues. Vectors \mathbf{S} and \mathbf{W} are synthesized for the strongly conserved and weakly conserved residues, respectively, in similar manner. Parameters f' , s' , w' and are for the regulation of the vector.

Angles between the vector \mathbf{C} and \mathbf{B} are also evaluated, where \mathbf{B} is a vector from the center axis of a TMH to the center of TMH bundle.

2.4 Prediction of Helix Arrangement

Applying the quantities I_c and \mathbf{C} , prediction of the arrangement of TMH becomes practical in the transmembrane proteins.

After amino acids of predicted helix are arranged on helical wheel, I_c and \mathbf{C} of the helix are calculated. Then, the helices in the bundle can be arranged on two concentric hypothetical circles, following the observed tendencies in the results.

3 Results and Discussion

The conservation index I_c of TMH is higher in the inner region than in the outer region of the bundle. The value of I_c decreases with the increase of the distance of TMH from the center of bundle. An example of the strong correlations from the data of sensory rhodopsin II (SWISS PROT entry: BACT_NATPH) is shown in Figure 1. In addition, most of the conservation vectors turn toward the inside of the bundle, although they are not exactly parallel to the vector \mathbf{B} .

The results obtained in this study indicate that the inner residues (especially, residues toward the inside) on TMH are more conservative than the others in the transmembrane proteins.

Method for the prediction of helix arrangement is examined for the structure of several proteins. Result of the trial is highly sufficient for sensory rhodopsin II (Figure 2). However, there are a few less sufficient cases in the other protein families. Further investigations and refinements are required.

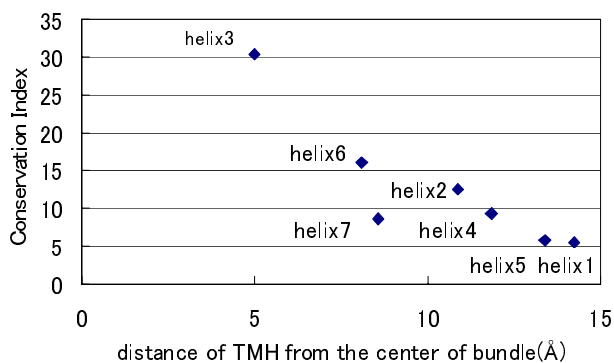


Figure 1: Predicted helix arrangement of sensory rhodopsin II. Large letters: TMH numbers, small letters: residue numbers.

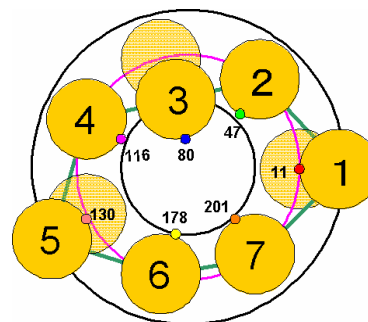


Figure 2: Correlation of I_c value with the distance of TMH from the center of bundle.

References

- [1] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J., Basic local alignment search tool, *J. Mol. Biol.*, 215(3):403–410, 1990.
- [2] <http://clustalw.genome.ad.jp/>