Classifying Membrane Exposure of Transmembrane Helices

Jennifer Metzger, Sikander Hayat, and Volkhard Helms

Center for Bioinformatics, Saarland University, 66041 Saarbrücken, Germany

'Helical membrane proteins' (HMPs) play a crucial role in diverse physiological processes of all organisms, including energy generation, signal transduction, and transport of solutes across the membrane. As a result of the wide range of essential functions, they provide the major category of drug targets [1]. Given the difficulty in determining high resolution structures by experimental techniques, it is highly desirable to develop sequence-based computational methods for predicting structural characteristics of HMPs. In this context, the ability to predict the solvent accessibility of transmembrane residues directly from the sequence is a valuable tool. Here, accessibility refers to the extent to which a residue in the hydrophobic core region of helical transmembrane bundles is in contact to the surrounding phospholipid molecules.

In our studies, we applied a two-stage approach for predicting the solvent accessibility of transmembrane residues in HMPs. In the first stage of our method a positional score for each sequence position was calculated based on conservation indices, frequency profiles, and position specific scoring matrices obtained by AL2CO [2] and PSI-BLAST [3]. In the second stage these positional scores were then used to predict the solvent accessibility of transmembrane residues. Because real values of solvent accessibility directly reflect the degree to which the residues are in contact with the solvent molecules, 'support vector regression' (SVR) was applied to predict real valued 'relative solvent-accessible surface areas' (rSASA) of residues.

To gain a further step towards the determination of three dimensional structures of transmembrane protein, we also developed a method which predicts the membrane exposure of transmembrane helices based on the real valued rSASA determined by our best model in the phase before.

[1] A. Elofsson, G. von Heijne, *Annu Rev Biochem*, **2007**, *76*, 125-140.
[2] J. Pei, N.V. Grishin, *Bioinformatics*, **2001**, *17*, 700-712.
[3] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, *J Mol Biol*, **1990**, *215*, 403-410.