Membrane Protein Structure Prediction: Influence of Distance Constraints on the Efficiency of the Conformational Search with Monte Carlo Simulation

Claire Gervais, Zhong Chen*, Ying Xu
Computational Systems Biology Lab, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA
- Present address: 3M Corporate Research Materials Lab, 3M Center, St Paul, MN, USA

Solving the three-dimensional (3D) structure of integral membrane proteins still remains a challenge. Even though membrane proteins comprise about 20-30% of the genome [1], little is known about their 3D structure and the rules governing their packing. Up to now, the 3D structures of only 123 unique proteins have been solved [2].

For membrane proteins, computational methods aiming at predicting the 3D structure according to the protein sequence are still at their infancy. Indeed, homology-based approaches used traditionally with soluble proteins require a large database of known protein structures and thus cannot serve yet as a general purpose approach.

In this context, we present the general approach developed in our laboratory to predict the packing of transmembrane helices by using *ab initio* simulational techniques. In order to deal with the large conformational space, Monte Carlo simulations are carried out at two different levels of complexity: (i) Search for optimal packings is first performed at a coarse-grained (residue) level; (ii) Low-energy residue structures are then selected and used as input structures for simulations at the atomistic level. The main requirements for such an approach lie in a correct sampling of the conformational space [3] and in the development of energy functions correctly describing the system at both levels [4]. Along with preliminary results performed on Glycophorin A [5], we show how to include experimental data (e.g. in form of distance constraints) in the conformational search procedure in order to increase the overall efficiency of the algorithm.

[1] Wallin, E. and von Heijne, G., *Protein Sci.*, **1998**, 7, 1029-1038.
[2] White, S. H., *Protein Sci*., **2004**, 13, 1948-1949.
[3] Wang, F. and Landau, D.P., *Phys. Rev. Lett.*, **2001**, 86, 2050-2053.
[4] Chen, Z. and Xu, Y., *J. Bioinformatics Comp. Bio.*, **2006**, 4, 317-333.
[5] Chen, Z. and Xu, Y., *Proteins Struct. Funct. Bioinformatics*, **2006**, 62, 539-552.