Improving transmembrane protein prediction analysis using

Hassan Kazemian
Professor in Computer Science and Director of Intelligent Systems Research Centre,
Faculty of Life Sciences & Computing,
London Metropolitan University,
166-220 Holloway Road, London N7 8DB, UK.

Most membrane protein coats the surface of the cell. Some membrane of a cell is not only a border, it is also an interface. Many of these proteins span the distance from the outside to the inside of the cell. The part of the protein which makes contact with the cell interior is called the transmembrane protein. Long before the advent of genomics it was realised that the parts of proteins which have to contact membranes tend to be made of fat-loving amino acids, since the membranes of cells are largely made of fat. Membrane proteins constitute approximately 75% of possible targets for novel drugs. However, membrane proteins are one of the most understudied groups of proteins in biochemical research because of the technical difficulties of obtaining structural information about transmembrane regions.

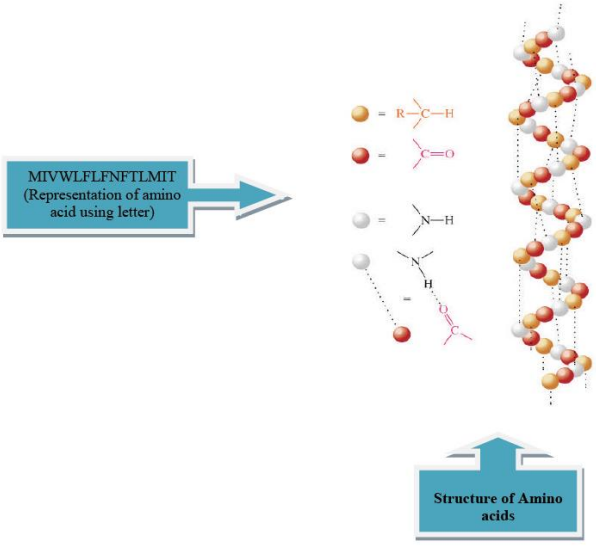

Figure 1 – Representation of amino acid using alphabetical letters.

This research is concerned with applications of artificial intelligence to amino acid sequences using the conventional alphabetical representation (Figure 1) to predict alpha-helices and beta barrels transmembrane regions. The research presents a novel evolutionary support vector machine (SVM) based alpha-helix transmembrane region prediction algorithm to solve the membrane helices in amino acid sequences. The SVM-genetic algorithm (GA) methodology is based on the optimisation of sliding window size, evolutionary encoding selection and SVM parameter optimisation. The computer simulation results demonstrate that the proposed SVM-GA methodology performs better than most traditional techniques producing an accuracy of 86.71% for cross-validation and 86.43% for jack-knife for randomly selected proteins containing single and multiple transmembrane regions [1].

The beta barrel research presents the applications of a dual Neural Network (NN) and Support Vector Machine (SVM) to prediction and analysis of beta barrel topology. The research applies a hybrid NN-SVM methodology to classify inter-class and intra-class transitions to predict the number and range of beta membrane spanning regions. The methodology uses sliding-windows for feature extraction to train two different class transitions entitled symmetric and asymmetric methods. For the NN-SVM to generate robust outcomes, the prediction methodology is analysed by Jack-knife tests and single protein tests. The computer simulation results demonstrate that NN-SVM tests with a 5 residue overlap for signal protein perform better than NN with and without redundant proteins for prediction of transmembrane beta barrel spanning regions.

[1] H. Kazemian, K. White and D. Palmer-Brown, S. A. Yusuf, Applications of Evolutionary SVM to Prediction of Membrane Alpha-Helices, Expert Systems with Applications, Elsevier, vol. 40, issue. 9, pp. 3412–3420, July 2013.

Presentation Method: Invited Oral 20 minutes. **Workshop:** Membrane for Pharmaceutical and Medical Applications