

# Contact prediction for transmembrane proteins

Hennerdal A. et al.

April 4, 2007

Due to difficulties in the experimental determination of the structure of transmembrane (TM) proteins, fewer than 250 such structures have been solved since 1985. Clearly, other methods are needed to increase the rate of structure discovery.

The alpha-helical class of TM proteins is the most common and the most interesting, not least since it contains many novel drug targets. Structure prediction for this class is in its infancy and many strategies that have been proven useful for soluble proteins have yet to be implemented.

We have developed the first predictor for residue-residue contacts in alpha-helical TM proteins, using a methodology similar to that previously used on soluble proteins. It thus utilizes data on sequence space separation and correlated mutations of residues as tried for soluble proteins, but our aim has been to incorporate additional features unique to alpha-helical TM regions.

Studies of residue-residue contact data in TM regions reveals a tendency for intra-membrane contacts to occur every 37 residues along the TM chain and another such tendency, albeit weaker, for contacts every 4 residues. This is interpreted as two alpha-helices in contact with each other in tilted and parallel fashion respectively.

Our final prediction method achieves an average of 23% accuracy on predicted contacts, which is equivalent to a 10-fold improvement over a random predictor. This is the same level of accuracy as is seen in other contact prediction methods optimized for soluble proteins. The predictor's accuracy is highest for protein chains with 5-7 TM helices, such as those in the rhodopsin family.