Computergestützte Strukturbiologie
(Strukturelle Bioinformatik)

Membrane proteins

Sommersemester 2009

Peter Güntert

Membrane proteins

- About 25% of all human proteins are membrane proteins.
- The majority of proteins that are drug targets are membrane proteins.
- However, less than 1% of the protein structures in the PDB are membrane proteins.
- Membrane proteins are difficult to crystallize.
- Membrane proteins are insoluble in water and thus difficult to study by NMR spectroscopy.

Membrane with an embedded protein

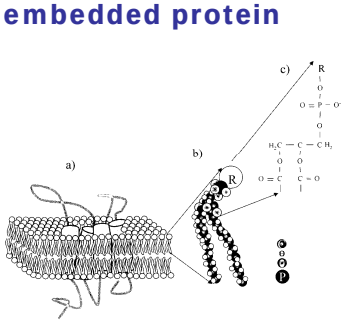


Figure 8.1 (a) Schematic diagram of a membrane with an embedded protein. (b) The structure of a phospholipid. (c) The chemical structure of its hydrophilic head. The phospholipid is formed by a glycerol molecule linked to two fatty acids, long linear carbon chains usually 16 or 18 carbon atoms in length, and a hydrophilic groups indicated by R in the figure. Glycerol is an alcohol with three carbon atoms, each bearing a hydroxyl group. Two of the hydroxyl groups of glycerol are bound to the fatty acids, which can be saturated or unsaturated, i.e. containing none or up to three double bonds, respectively. Two saturated hydrocarbon tails can pack tightly together An unsaturated fatty acid will have a kink in its shape wherever a double bond occurs; this results in looser packing and is important for the fluidity of the cell membrane. One of the hydroxyl groups of glycerol is joined to a phosphate group, which is negatively charged. Additional small molecules, usually charged or polar, can be linked to the phosphate group to form a variety of phospholipids. Because of the presence of a hydrophilic and hydrophobic region, phospholipids arrange in a double layer as shown in (a).

Two membrane proteins

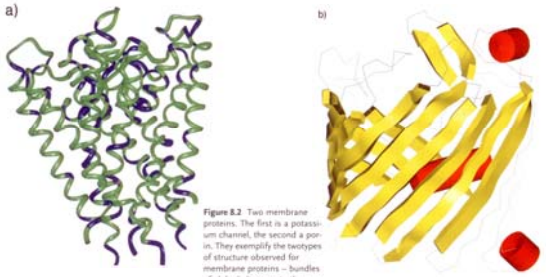


Figure 8.2 Two membrane proteins. The first is a potassium channel, the second a porin. They exemplify the twotypes of structure observed for membrane proteins – bundles of alpha helices, as in the channel, and beta barrels, as in the porin. In both examples main chain hydrogen-bonds are saturated.

Membrane protein structure and function

- Membrane proteins are either α-helix bundles or β-barrels
- Most of the important membrane proteins are α-helical. β-barrel membrane proteins are found only in the outer membrane.
- Functions:
  - Transport of ions/molecules through the membrane
  - Signal transduction (e.g. G-protein coupled receptors; GPCR)

Hydrogen bonding in membrane proteins

- In soluble proteins hydrogen bonds can be formed between protein atoms and to the solvent (water).
- In the hydrophobic environment of membrane proteins only intra-protein hydrogen bonds can be formed.
- The energy cost of burying a pair of unsatisfied hydrogen bond donors/acceptors is high: ~ 4 kcal/mol.
- Transmembrane segments that are α-helices or β-strands involve all main chain polar atoms in hydrogen bonds.
- Random coil structures within the membrane bilayer is energetically unfavorable and rare.

Membrane protein structure prediction

- The present data base of known membrane protein structures is small.
- Less structural variability than for soluble proteins
- Prediction of transmembrane helices based on secondary structure prediction and hydrophobicity

Which features of membrane protein structure can be predicted?

- Discrimination between globular and membrane proteins
- Identification of individual transmembrane helices
- Location of loops on the inside/outside of the membrane
- Topology

Chou-Fasman propensities for α, β, and coil secondary structure

Table 7.1 The Chou and Fasman parameters for secondary structure prediction.

Amino acid	*Pa*	*Pb*	*Pc*	Amino acid	*Pa*	*Pb*	*Pc*
Alanine	1.42	0.83	0.66	Leucine	1.21	1.30	0.59
Arginine	0.98	0.93	0.95	Lysine	1.14	0.74	1.01
Aspartic acid	1.01	0.54	1.46	Methionine	1.45	1.05	0.60
Asparagine	0.67	0.89	1.56	Phenylalanine	1.13	1.38	0.60
Cysteine	0.70	1.19	1.19	Proline	0.57	0.55	1.52
Glutamic acid	1.39	1.17	0.74	Serine	0.77	0.75	1.43
Glutamine	1.11	1.10	0.98	Threonine	0.83	1.19	0.96
Glycine	0.57	0.75	1.56	Tryptophan	1.08	1.37	0.96
Histidine	1.00	0.87	0.95	Tyrosine	0.69	1.47	1.14
Isoleucine	1.08	1.60	0.47	Valine	1.06	1.70	0.50

Chou-Fasman secondary structure prediction

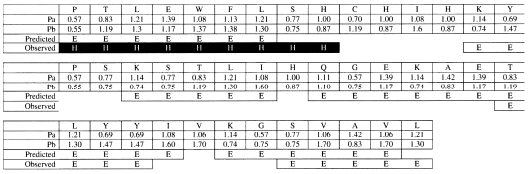


Figure 7.3 An example of the application of the Chou and Fasman method to a protein. It is apparent that the predicted and experimental structures only have limited overlap.

Neural network nodes

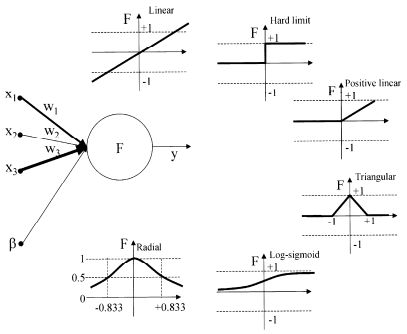


Figure 7.5 A node of a neural network is a computational unit that transforms input values x_i into an output value y. The input values can have different weights and the transformation can be different. A few examples of "transfer functions", used to compute the output of the node as a function of the weighted sum of its inputs, are shown in the figure.

Neural network layers

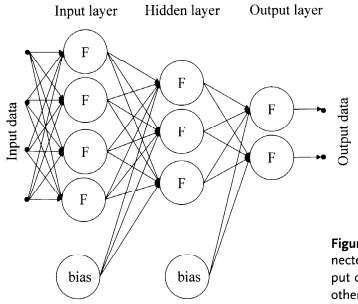


Figure 7.6 Several nodes can be connected in a network such that the output of some nodes is the input of others.

PHD secondary structure prediction algorithm

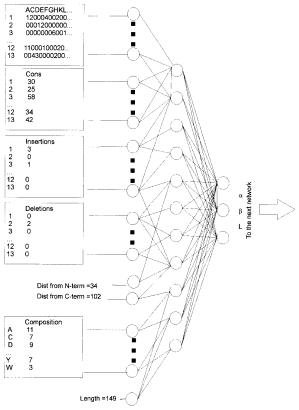
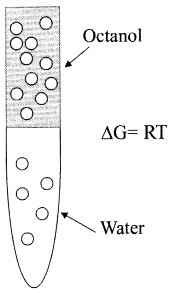


Figure 7.13 The secondary structure prediction method PHD uses several input nodes, containing information on the protein or on the thirteen-residue window under examination, to produce a prediction of the secondary structure of the central element of the window.

Free energy of transfer for a molecule between two media



$$\Delta G = RT \ln(c1/c2) = 2.53$$

Figure 8.3 The free energy of transfer of a molecule between two media can be derived from the equilibrium concentrations of the molecule in a two phase system.

Hydrophobicity plot

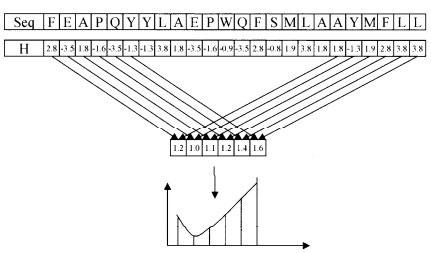


Figure 8.4 The method used to compute the hydrophobicity plot of a protein sequence. Each amino acid is assigned a hydrophobicity value. The average is computed over a sliding window of 19 residues.

Hydrophobicity plot for rhodopsin

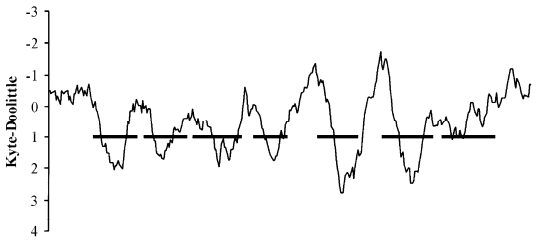


Figure 8.5 Hydrophobicity plot for the rhodopsin protein sequence. The black bars represent the experimentally determined position of the seven trans-membrane helices of this protein. Ideally, they should all correspond to large hydrophobicity values (minima in the plot).

Helical wheel and hydrophobic moment vectors

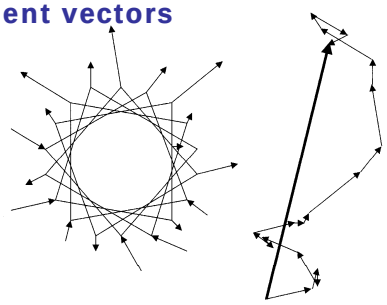


Figure 8.6 Helical wheel representation of a helix. The view is along the helix axis. The vectors represent the hydrophobic moment of each amino acid. The length of the vector is proportional to the hydrophobicity of the amino acid. The vectors connects the C alpha atom to the center of its side-chain. The hydrophobic moment of the helix is the sum of these vectors, indicated by the tick arrow on the left of the figure.

TMHMM: Hidden Markov model for transmembrane topology prediction

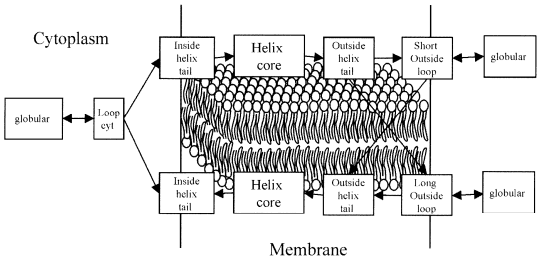


Figure 8.7 Possible scheme of a hidden Markov model for predicting the topology of a transmembrane protein. This is the layout used by the TMHMM server.

TMHMM performance

- **Discrimination globular/membrane: sensitivity and specificity > 98%**
- **Correct topology: 65-70%**
- **Single transmembrane helix identification: sensitivity 96%, specificity 98%**
- **Training set: 160 membrane proteins, 650 globular proteins**

http://www.cbs.dtu.dk/services/TMHMM/
Krogh et al., J. Mol. Biol. 305, 567-580 (2001)

Literatur

- Anna Tramontano: *Protein Structure Prediction, Wiley-VCH*, 2006.
- A. Krogh, B. Larsson, G. Von Heijne, E. L. L. Sonnhammer. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567-580 (2001)