**CHOPnet**

CHOPnet is a de novo method that predicts structural domains in the absence of homology to known domains. The method is based on neural networks and relies exclusively on information available for all proteins. Multiple sequence alignments are obtained by searching with PSI-BLAST[1] against all known sequences contained in SWISS-PROT, TrEMBL[2] & PDB[3]. All hits below a PSI-BLAST E-value of 10-3 are filtered and included in the sequence profile. For these filtered alignments, amino acid composition, predicted secondary structure and solvent accessibility are calculated and used as an input to the neural network. The aforementioned neural network is a three layer feed forward artificial neural network using the standard back-propagation algorithm with momentum term[5][6] with a post processing step which removes the noisy data.

**References:**

[1]: Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 25, 3389–3402.

[2]: Bairoch,A. and Apweiler,R. (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. 28, 45–48.

[3]: Berman,H.M., Westbrook,J., Feng,Z., Gillliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. Nucleic Acids Res., 28, 235–242.

[4]: Rumelhart,D.E., Hinton,G.E. and Williams,R.J. (1986) In Rumelhart,D.E. and McClelland,J.L. (eds), Parallel Distributed Processing. MIT Press, Cambridge, UK, Vol. 1, pp. 318–362.

[5]: Rost,B. and Sander,C. (1993) Prediction of protein secondary structure at better than 70% accuracy. J. Mol. Biol., 232, 584–599.