

REFERENCES

- Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F., and Wang, J. 1987. Protein data bank. In F.H. Allen ; G. Bergerhoff ; and R. Sievers(eds.), Crystallographic databases-information content, software systems, scientific applications, pp. 107-113. Bonn/Cambridge/Chester : Data Commission of the international Union of Crystallography.
- Aleksander, I., and Morton, H. 1990. An Introduction to Neural Computing. Great Britain : TJ Press.
- Andreas, *et al.* 1989. SNNS User Manual, Version 3.3. University of Stuttgart.
- Anfinsen, C.B., Haber, E., Sela, M., and White, F.H. 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Nat. Acad. Sci. 47 : 1309-1314.
- Argos, P. 1987. Analysis of sequence-similar pentapeptides in unrelated protein tertiary structures. J. Mol. Biol. 197 : 331-348.
- Bairoch, A., and Boeckmann, B. 1991. The SWISS-PROT protein sequence data bank. Nucleic Acids Res. 19 : 2247-2249.
- Bairoch, A., and Boeckmann, B. 1993. The SWISS-PROT protein sequence data bank, recent developments. Nucl. Acid Res. 21 : 3093-3096.
- Bazzo, R., *et al.* 1988. The structure of melittin using H-NMR study in methanol. Eur. J. Biochem. 173 : 139-146.

- Bernstein, F.C., et al. 1977. The protein Data Bank : a computer-based archival file for macromolecular structures. J. Mol. Biol. 112 : 535-542.
- Blout, E.R., de Loze, C., Bloom, S.M., and Fasman, G.D. 1960. The dependence of the conformations of synthetic polypeptides on amino acid composition. J. Amer. Chem. Soc. 82 : 3787-3789.
- Bohm, G. 1996. Review : New approaches in molecular structure prediction. Biophysical Chemistry. 59 : 1-32.
- Bryson, A.E., and Ho Y.-C. 1969. Applied Optimal Control. New York : Blaisdell.
- Chandonia, J., and Karplus, M. 1995. Neural networks for secondary structure and structural class predictions. Protein science. 4: 275-285.
- Chothia, C. 1992. One thousand families for the molecular biologist. Nature. 357 : 543-544.
- Chothia, C., and Finkelstein, AV. 1990. The classification and origins of protein folding patterns. Annu. Rev. Biochem. 59 : 1007-1039.
- Chou, P., and Fasman, G. 1974. Prediction of protein conformation. Biochemistry. 13 : 222-245.
- Creighton, T.E. 1983. Proteins : structures and Molecular Principles. New York : W.H. Freeman.
- Creighton, T.E. 1993. Protein. 2nd ed. New York : W.H. Freeman.
- Denhoff, J. 1990. Neural Network Architectures. New York : Van Nostrand Reinhold.
- Dubchak, I., Holbrook, and Kim, S.H. 1993. Prediction of folding class from amino acid composition. Proteins Struct. Funct. Genet. 16 : 79-91.

- Duchak, I., Muchnik, I., Holbrook, S., and Kim S.H. 1995. Prediction of protein folding class using global description of amino acid sequence. Pro. Natl. Acad. Sci. USA. 92 : 8700-8704.
- Eisenhaber, F., Persson, B., and Argos, P. 1995. Protein Structure Prediction : Recognition of Primary, Secondary, and Tertiary Structural Features from Amino Acid Sequence. Critical Reviews in Biochemistry and Molecular Biology. 30(1) : 1-94.
- Eisenhaber, F.; Persson, B.; and Argos, P. 1995. Protein Structure Prediction : Recognition of Primary, secondary, and Tertiary Structural Features from Amino Acid Sequence. In G.D. Fasman(ed.), Critical Reviews in Biochemistry and Molecular Biology, pp. 1-94. Heidelberg : CRC Press.
- Finkelstein, A.V., and Ptitsyn, O.B. 1987. Why do globular proteins fit the limited set of folding patterns?(review). Prog. Biophys. Mol. Biol. 50 : 171-190.
- Fu, L. 1994. Neural Networks in Computer Intelligence. New York : McGraw-Hill.
- Garnier, J., Osguthorpe, D., and Robson, B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120 : 97-120.
- Hertz, J., Krogh, A., and Palmer, G.R. 1991. Introduction to the theory of neural computation. Vol. 1. California : Addison-Wesley.
- Hirst, J.D., and Sternberg, M.J.E. 1992. Prediction of structural and functional features of protein and nucleic acid sequences by artificial neural networks. Biochemistry. 31 : 615-623.

- Holley, H.L., and Karplus, M. 1989. Protein secondary structure prediction with a neural network. Proc. Natl. Acad. Sci. USA. 86 : 152-156.
- Holly, H., and Karplus, M. 1991. Neural networks for protein structure prediction. Methods Enzymol. 202 : 204-224.
- Kabsch, W., and Sander, C. 1983. Dictionary of protein secondary structure : pattern recognition of hydrogen bonded and geometrical features. Biopolymers. 22 : 2577-2637.
- Kabsch, W., and Sander, C. 1984. On the use of sequence homologies to predict protein structure : Identical pentapeptides can have completely different conformations. Proc. Natl. Acad. Sci USA. 81 : 1075-1078.
- Klein, P., and DeLisi, C. 1986. Prediction of protein structural class by discriminant analysis. Biochim. Biophys. 874 : 205-215.
- Kneller, D.G., Cohen, F.E., and Langridge, R. 1990. Improvements in protein secondary structure prediction by an enhanced neural network. J. Mol. Biol. 214 : 171-182.
- Kneller, G.R., and Geiger, A. 1989. A method to calculate the g-coefficients of the molecular pair correlation function from molecular dynamics simulations. Mol. Simul. 3 : 283-300.
- Kyte, J, and Doolittle, R.F. 1982. A simple method for displaying the hydrophatic character of a protein. J. Mol. Bio. 157(1) : 105-132.
- Le Cun, Y. 1985. Une procedure d'apprentissage pour reseau a seuil assymetrique [A learning procedure for asymmetric threshold network]. Proceedings of Cognitiva. 85 : 599-604.

- Lehninger, A.L., Nelson, D.L., and Cox, M.M. 1993. Principles of Biochemistry. 2nd ed. New York : Worth Publishers.
- Levin, J.M., Pascarella, S., Argos, P., and Garnier, J. 1993. Quantification of secondary structure prediction improvement using multiple alignments. Prot. Eng. 6 : 849-854.
- Levitt, M., and Chothia, C. 1976. Structure pattern in globular protein. Nature. 202 : 552-558.
- Lim, V.I. 1974. Structural principles of globular organization of protein chains. A stereochemical theory of globular protein secondary structure. J. Mol. Biol. 88 : 873-894.
- Marion, D., and Wathrith, K. 1983. Application of phase, sensitive 2-D correlated spectroscopy (cosy) for measurement of H-H spin coupling constant in protein. Biochemical and Biophysical research communication. 113(3) : 967-974.
- Marqusee, S., and Baldwin, R.L. 1987. Helix stabilization by Glu⁻Lys⁺ salt bridges in short peptides of *de novo* design. Proc. Nat. Acad. Sci. 84 : 8898-8902.
- Marqusee, S., Robbins, V.J., and Baldwin, R.L. 1989. Unusually stable helix formation in short alanine-based peptides. Proc. Nat. Acad. Sci. 86 : 5286-5290.
- McGregor, M.J., Flores, T.P., and Sternberg, M.J.E. 1989. Prediction of beta-turns in proteins using neural networks. Prot Eng. 2(7) : 521-526.
- Muggleton, S., King, R.D., and Sternberg, M.J.E. 1992. Protein secondary structure prediction using logic based machine learning. Prot. Eng. 5 : 647-657.

- Muller, B., and Reinhardt, J. 1990. Neural Networks. In E. Domany ; J.L.van Hemmen ; and K. Schulten(eds.), Physics of Neural Networks, pp. 1-261. Berlin : Springer-Verlag.
- Nishikawa, K, and Ooi, T. 1982. Correlation of the amino acid composition of a protein to its structural and biological characters. J. Bjo. Chem. 91 : 1821-1824.
- Nishikawa, K., Kubota, Y., and Ooi, T. 1983. Classification of proteins into groups based on amino acid composition and other characters. I. Angular distribution. J. Biochem. 94 : 997-1007.
- Oian, N., and Sejnonowski, T.J. 1988. Predicting of the secondary structure of globular proteins using neural network models. J. Mol. Biol. 262 : 865-884.
- Ose, T.G., and Kim, P.S. 1988. A peptide model of a protein folding intermediate. Nature. 336 : 42-48.
- Parker, D.B. Learning Logic. Technical Report TR-47. Center for Computational Reserch in Economics and Management Science, Massachusetts Institute of Technology, Cambridge, MA.
- Pascarella, S., and Argos, P. 1992. A databank merging related protein structures and sequences. Protein engineering. 5(2) : 121-137.
- Presnell, SR., Cohen, B.I., and Cohen, F.E. 1992. A segment-based approach to protein secondary structure prediction. Biochemistry. 31 : 983-993.
- Roder, H., Elove, G.A., and Englander, S.N. 1988. Structural characterization of folding intermediates in cytochrom C by H-exchange labeling and protein NMR. Nature. 335 : 700-704.

- Roman, M., and Wodak, S.J. 1988. Identification of predictive sequence motifs limited by protein structure data base size. Nature. 335 : 45-49.
- Rosenblatt, F. 1962. Principles of neurodynamics. New York : Spartan.
- Rost, B., and Sander, C. 1993 Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Prot. Natl. Acad. Sci. USA. 90 : 7558- 7562.
- Rost, B., and Sander, C. 1994. Combining evolutionary information and neural networks to predict protein secondary structure. Proteins Struct Funct Genet. 19 : 55-72.
- Rost, B., and Sander, C. 1995. Protein structure prediction by Neural networks. The handbook of brain theory and Neural networks. Boston : Bradford Book/The MIT Press.
- Rumelhart, D.E., and McClelland J.L. 1986. Parallel Distributed Processing. Vols 1 & 2. Cambridge, Mass : MIT Press.
- Sasagawa, F., and Tajima, K. 1993. Prediction of protein secondary structures by a neural network. CABIOS. 9(2) : 147-152.
- Stryer, L. 1988. Biochemistry. 3rd ed. New York : W.H. Freeman.
- Tayer, W.R., and Thornton, J.M. 1984. Recognition of super-secondary structure in proteins. J. Mol. Biol. 173 : 487-514.
- Udgaonkar, J.B., and Baldwin, R.L. 1988. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. Nature. 335 : 694-699.

- Urabe, K., and Sakaguchi, H. Characteristics of current induced potential oscillations of a triolein impregnated membrane placed between identical salt solutions. Biophysical Chemistry. 59 : 33-39.
- Voet, D., Voet, J. 1990. Biochemistry. 2nd ed. New York : John Wiley & Sons.
- Werbos, P.J. 1974. Beyond regression : New tools for prediction and analysis in the behavioral sciences. Thesis, Harvard University.
- Williamson, M.P., Hall, M. J., and Hand, B.K. 1986. H-NMR assignment and secondary structure of Herpes simplex virus glycoprotein D-1 antigenic domain. J. Biochem. 158 : 527-536.
- Zeidenberg, M. 1990. Neural network models in artificial intelligence. London : Ellis horwood.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

APPENDIX A

List of train pattern files, test pattern file and network files in CD-ROM

Prediction of the existence of helix, sheet and turn in the same network

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr3op.pat	tehy2gr3op.pat	hy2gr3opn*.net
hydropathy (7 groups)	thy7gr3op.pat	tehy2gr3op.pat	hy2gr3opn*.net
amino acid side chain properties (8 groups)	taa8gr3op.pat	teaa8gr3op.pat	aa8gr3opn*.net
hydrophobicity (3 groups)	thydrogr3op.pat	thydrogr3op.pat	thydrogr3opn*.net

Prediction of the existence of helix

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr1h.pat	tehy2gr1h.pat	hy2gr1hn*.net
hydropathy (7 groups)	thy7gr1h.pat	tehy2gr1h.pat	hy2gr1hn*.net
amino acid side chain properties (8 groups)	taa8gr1h.pat	teaa8gr1h.pat	aa8gr1hn*.net
hydrophobicity (3 groups)	thydrogr1h.pat	thydrogr1h.pat	thydrogr1hn*.pat

Prediction of the existence of sheet

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr1s.pat	tehy2gr1s.pat	hy2gr1sn*.net
hydropathy (7 groups)	thy7gr1s.pat	tehy2gr1s.pat	hy2gr1sn*.net
amino acid side chain properties (8 groups)	taa8gr1s.pat	teaa8gr1s.pat	aa8gr1sn*.net
hydrophobicity (3 groups)	thydrogr1s.pat	thydrogr1s.pat	thydrogr1sn*.net

Prediction of the existence of turn

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr1t.pat	tehy2gr1t.pat	hy2gr1tn*.net
hydropathy (7 groups)	thy7gr1t.pat	tehy2gr1t.pat	hy2gr1tn*.net
amino acid side chain properties (8 groups)	taa8gr1ht.pat	teaa8gr1t.pat	aa8gr1tn*.net
hydrophobicity (3 groups)	thydrogr1t.pat	thydrogr1t.pat	thydrogr1tn*.net

Prediction of percent helix (6 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%6h.pat	tehy2gr%6h.pat	hy2gr%6hn*.net
hydropathy (7 groups)	thy7gr%6h.pat	tehy2gr%6h.pat	hy2gr%6hn*.net
amino acid side chain properties (8 groups)	taa8gr%6h.pat	teaa8gr%6h.pat	aa8gr%6hn*.net
hydrophobicity (3 groups)	thydrogr%6h.pat	thydrogr%6h.pat	thydrogr%6hn*.net

Prediction of percent sheet (6 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%6s.pat	tehy2gr%6s.pat	hy2gr%6sn*.net
hydropathy (7 groups)	thy7gr%6s.pat	tehy2gr%6s.pat	hy2gr%6sn*.net
amino acid side chain properties (8 groups)	taa8gr%6s.pat	teaa8gr%6s.pat	aa8gr%6sn*.net
hydrophobicity (3 groups)	thydrogr%6s.pat	thydrogr%6s.pat	thydrogr%6sn*.net

Prediction of percent turn (6 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%6t.pat	tehy2gr%6t.pat	hy2gr%6tn*.net
hydropathy (7 groups)	thy7gr%6t.pat	tehy2gr%6t.pat	hy2gr%6tn*.net
amino acid side chain properties (8 groups)	taa8gr%6t.pat	teaa8gr%6t.pat	aa8gr%6tn*.net
hydrophobicity (3 groups)	thydrogr%6t.pat	thydrogr%6t.pat	thydrogr%6tn*.net

Prediction of percent helix (3 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%3h.pat	tehy2gr%3h.pat	hy2gr%3hn*.net
hydropathy (7 groups)	thy7gr%3h.pat	tehy2gr%3h.pat	hy2gr%3hn*.net
amino acid side chain properties (8 groups)	taa8gr%3h.pat	teaa8gr%3h.pat	aa8gr%3hn*.net
hydrophobicity (3 groups)	thydrogr%3h.pat	thydrogr%3h.pat	thydrogr%3hn*.net

Prediction of percent sheet (3 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%ss.pat	tehy2gr%3s.pat	hy2gr%3sn*.net
hydropathy (7 groups)	thy7gr%3s.pat	tehy2gr%3s.pat	hy2gr%3sn*.net
amino acid side chain properties (8 groups)	taa8gr%3s.pat	teaa8gr%3s.pat	aa8gr%3sn*.net
hydrophobicity (3 groups)	thydrogr%3s.pat	thydrogr%3s.pat	thydrogr%3sn*.net

Prediction of percent turn (3 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%3t.pat	tehy2gr%3t.pat	hy2gr%3tn*.net
hydropathy (7 groups)	thy7gr%3t.pat	tehy2gr%3t.pat	hy2gr%3tn*.net
amino acid side chain properties (8 groups)	taa8gr%3t.pat	teaa8gr%3t.pat	aa8gr%3tn*.net
hydrophobicity (3 groups)	thydrogr%3t.pat	thydrogr%3t.pat	thydrogr%3tn*.net

Prediction of percent helix (2 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%2h.pat	tehy2gr%2h.pat	hy2gr%2hn*.net
helical tendencies (5 groups)	thy7gr%2h.pat	tehy2gr%2h.pat	hy2gr%2hn*.net
amino acid side chain properties (8 groups)	taa8gr%2h.pat	teaa8gr%2h.pat	aa8gr%2hn*.net
hydrophobicity (3 groups)	thydrogr%2h.pat	thydrogr%2h.pat	thydrogr%2hn*.net

Prediction of percent sheet (2 groups)

Properties	Train pattern file	Test pattern file	Networks file
hydropathy (2 groups)	thy2gr%2s.pat	tehy2gr%2s.pat	hy2gr%2sn*.net
hydropathy (7 groups)	thy7gr%2s.pat	tehy2gr%2s.pat	hy2gr%2sn*.net
amino acid side chain properties (8 groups)	taa8gr%2s.pat	teaa8gr%2s.pat	aa8gr%2sn*.net
hydrophobicity (3 groups)	thydrogr%2s.pat	thydrogr%2s.pat	thydrogr%2sn*.net

n* is number of hidden units (7, 35, 70, 100, 120)

APPENDIX B

The networks files which gave the highest accuracies for secondary structure predictions

Prediction	Network files	Properties	Hidden units	% Accuracy
1. The existence of helix, sheet and turn in the same network	aa8gr3op70.net	amino acid side chain	70	52.63
	aa8gr3op100.net	amino acid side chain	100	
2. The existence of helix	aa8gr1h7.net	amino acid side chain	7	100
	hy2gr1h100.net	hydropathy (2 groups)	100	
3. The existence of sheet	aa8gr1s100.net	amino acid side chain	100	85.71
4. The existence of turn	aa8gr1t7.net	amino acid side chain	7	67.8
5. Percent helix (6 groups)	hy2gr%6h35.net	hydropathy (2 groups)	35	70.71
6. Percent sheet (6 groups)	hy2gr%6s35.net	hydropathy (2 groups)	35	50
	hy7gr%6s70.net	hydropathy (7 groups)	70	
7. Percent turn (6 groups)	aa8gr%6t35.net	amino acid side chain	35	46.43
	aa8gr%6t100.net	amino acid side chain	100	

Prediction	Network files	Properties	Hidden units	Percent accuracy
8. Percent helix (3 groups)	aa8gr%3h70.net	amino acid side chain	70	78.57
	aa8gr%3h100.net	amino acid side chain	100	
9. Percent sheet (3 groups)	hy7gr%3s7.net	hydropathy (7 groups)	7	78.57
	aa8gr%3s100.net	amino acid side chain	100	
10. Percent turn (3 groups)	hy2gr%3t70.net	hydropathy (2 groups)	70	64.29
	hy2gr%3t100.net	hydropathy (2 groups)	100	
11. Percent helix (2 groups)	hydrogr%2h120.net	hydrophobicity	120	78.57
	aa8gr%h100.net	amino acid side chain	100	
12. Percent sheet (2 groups)	hy2gr%2s35.net	hydropathy (2 groups)	35	75.00
	hy2gr%2s120.net	hydropathy (2 groups)	120	

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

BIOGRAPHY



NAME :

JURAI RAT PHROMJAI

DATE OF BIRTH :

MARCH 27, 1973

PLACE OF BIRTH :

CHIANGMAI, THAILAND

INSTITUTIONS ATTENDED :

CHULALONGKORN UNIVERSITY, 1991-1994 :

BACHELOR OF SCIENCE (MEDICAL TECHNOLOGY)

CHULALONGKORN UNIVERSITY, 1995-1998 :

MASTER OF SCIENCE (BIOCHEMISTRY)

RESEARCH GRANT :

FACULTY OF GRADUATE STUDIES,

CHULALONGKORN UNIVERSITY.