

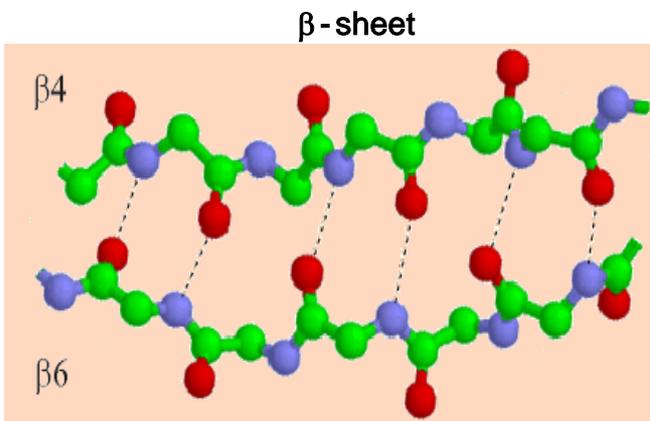
(secondary structure)

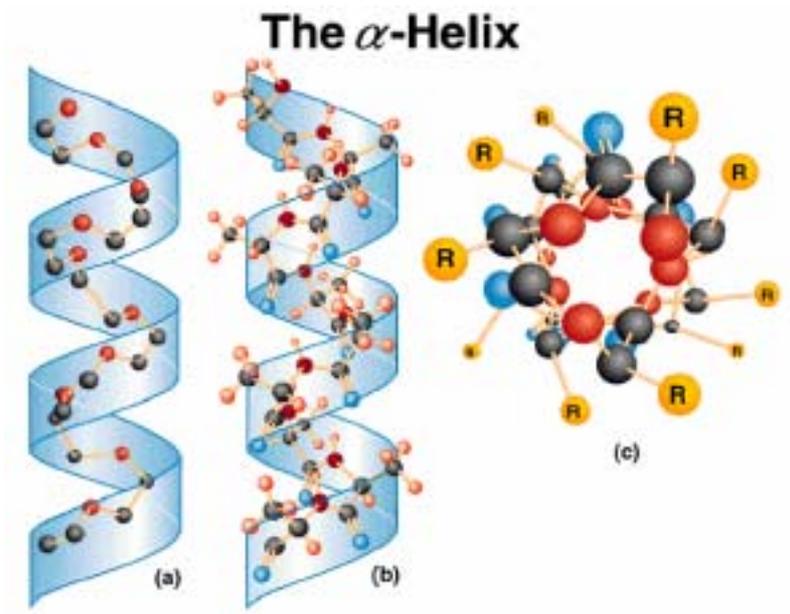
1. DSSP

(primary structure; DNA structure) (secondary structure) (tertiary structure)

DSSP(dictionary of protein secondary structure) [1]. DSSP (solvent accessibility)

DSSP DSSP
 “<http://www.cmbi.kun.nl/gv/dssp/>”
 가 DSSP 가 (G: 3₁₀ helix, H: alpha helix, I: pi helix, B: residue in isolated beta-bridge, E: extended strand, participates in beta ladder, T: hydrogen-bonded turn, S: bend, X: otherwise)
 가 가 (H: helix, E: extended, C: coil)
 (G, H, and I → H; B and E → E; T, S, and X → C).





2. PSIPRED

가
 PSIPRED [2]. PSIPRED (prediction accuracy) 80%
 . PSIPRED Protein Structure Prediction
 Server(<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>)
 [3].

PSIPRED ()
 . PSIPRED
 “<ftp://bioinf.cs.ucl.ac.uk/pub/psipred/>” . PSIPRED
 PSI-BLAST PSIPRED 가 . PSI-
 BLAST [4].

PSIPRED PSI-BLAST 가 “ ()
 ”(position-specific scoring matrix;
 PSSM) . PSSM $n \times 20$. n
 , 20 20 가
 . PSIPRED PSSM
 15 \times 20 , (input vector)
 15 (window size) , 15 .

PSIPRED (David T. Jones)가 15 ,
15 가 . PSIPRED
(neural network) , (pattern recognition)
(machine learning) 가 .

(training)
(parameters) .
가 . 15x20 가
, 15 가
(DSSP)

가

가

PSIPRED (PSI-BLAST
) PSSM . PSSM

PSIPRED PSSM .

PSIPRED .

3.

(support vector machine)

(nearest neighbor) .

SVMpsi [5]. SVMpsi PSIPRED .

PREDICT PSIPRED SVMpsi [6,7].
가 가 가
가 가 가
가 100 가 가
가 helix 가 20 , extended 가 50 , coil 30 가
extended extended 15x20

PSI-BLAST PSSM

(PSSM)

(<http://www.cheric.org>), PSSM

[]

- [1] W. Kabsch and C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers* **22** (1983) 2577-2637.
- [2] D. T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, *J. Mol. Biol.* **292** (1999) 195-202.
- [3] L. J. McGuffin, K. Bryson, and D. T. Jones, The PSIPRED protein structure prediction server, *Bioinformatics* **16** (2000) 404-405.

- [4] S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* **25** (1997) 3389-3402.
- [5] H. Kim and H. Park, Protein secondary structure prediction based on an improved support vector machines approach, *Protein Eng.* **16** (2003) 553-560.
- [6] K. Joo, J. Lee, S.-Y. Kim, I. Kim, J. Lee, and S. J. Lee, Profile-based nearest neighbor method for pattern recognition, *J. Korean Phys. Soc.* **44** (2004) 599-604.
- [7] K. Joo, I. Kim, S.-Y. Kim, J. Lee, J. Lee, and S. J. Lee, Prediction of the secondary structures of proteins by using PREDICT, a nearest neighbor method on pattern space, *J. Korean Phys. Soc.* **45** (2004) 1441-1449.