

From DNA to protein : comparative analyses to investigate structural relationships

Maria Luisa Chiusano and Giovanni Colonna

Centro di Ricerca di Scienze Computazionali e Biotechnologiche (CRISCEB), II Università di Napoli - via Costantinopoli 16, 80138 Napoli.
E-mail: marilu@crisceb.area.na.cnr.it

We developed a computational method to analyse a coding sequence considering all the information available for the sequence itself and its product. By a graphical analysis of the composition of the nucleic acid sequence, the corresponding amino acid sequence, its chemico-physical properties, the structural and functional information derived from the Swissprot database, the prediction of secondary structure derived from a consensus of five different predictive methods (3,4,5,6,8) and the three-dimensional information derived from the DSSP program (7) (when the experimental structure of the protein exists), it is possible to investigate and summarize the structural features of a protein from its coding sequence to its structure. The software allows the analysis of multiple alignments of sequences too. In this case, it is possible to perform a deeper analysis to infer on both functional and evolutionary information.

Integrating the software with a tool which calculates the substitution rate for both synonymous and nonsynonymous positions (1), we defined different compositional and substituting behaviour for the secondary structures of 34 mammal coding sequences (2).

References

- 1 Alvarez-Valin, F., Jabbari, K. and Bernardi, G., 1998. Synonymous and nonsynonymous substitutions in mammalian genes: intragenic correlations. *J. Mol. Evol.* 46, 37-44.
- 2 Chiusano M.L., D'Onofrio G., Alvarez-Valin F., Jabbari K., Colonna G., Bernardi G. Correlations of Nucleotide Substitution Rates and Base Composition of Mammalian Coding Sequences with Protein Structure. *Gene* 238, 23-31 (1999).
- 3 Deleage, G. and Roux, B., 1987. An algorithm for protein secondary structure prediction based on class prediction. *Prot. Eng.* 1, 289-294.
- 4 Geourjon, C. and Deleage, G., 1994. SOPM: a self optimised prediction method for protein secondary structure prediction. *Prot. Eng.* 7, 157-164.
- 5 Geourjon, C. and Deleage, G., 1995. SOPMA Significant improvements in protein secondary structure prediction by prediction from multiple alignments. *CABIOS* 11, 681-684.
- 6 Gibrat, JF., Garnier, J. and Robson, B., 1987. Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs. *J. Mol. Biol.* 198, 425-443.
- 7 Kabsch W. and Sander C., 1983. Dictionary of Protein Secondary Structure: Pattern recognition of Hydrogen-Bonded and Geometrical Features, *Biopolymers* 22, 2577-2637.
- 8 Levin, JM., Robson, B. and Garnier, J., 1986. An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS Lett.* 15, 303-308.