ESPript/ENDscript: Sequence and 3D Information from Protein Structures

Patrice Gouet[a], Xavier Robert[a], Emmanuel Courcelle[b], [a]*Laboratoire de BioCristallographie, IBCP IFR128, France.* [b]*Laboratoire de Biologie Moléculaire et des Relations Plantes Microorganismes, Castanet Tolosan, France.* E-mail: p.gouet@ibcp.fr

The fortran program ESPript has been created to display on a single PostScript figure, multiple sequence alignments adorned with secondary structure elements [1]. A web server is available at http://espript.ibcp.fr/ESPript/ESPript. It has been linked to three web tools: ProDom which identifies protein domains, PredictProtein which predicts secondary structure elements and NPS@ which runs sequence alignment programs.

An extension of ESPript named ENDscript has been made available at the same electronic address [2]. It enables the creation from a single Protein Data Bank identifier, of a multiple sequence alignment figure adorned with secondary structures of each sequence of known structure. ENDscript uses programs such as BLAST, CLUSTAL and PHYLODENDRON to work on protein sequences and such as DSSP and CNS to work on protein coordinates. Similar structures are superimposed in turn with the program PROFIT. Final 3D figures are drawn with MOLSCRIPT, BOBSCRIPT and DINO, so as to show sequence conservation as well as structure conservation.

[1] Gouet P., Courcelle E., Stuart D. I., Metoz F., *Bioinformatics*, 1999, **15**, 305. [2] Gouet P., Courcelle E. , *Bioinformatics*, 2002, **18**, 767.

Keywords: bioinformatics, sequence homology, protein structure comparison