ProCoS: Protein composition server

Lavanya Rishishwar[1]*, Neha Mishra[1], Bhasker Pant[1], Kumud Pant[1], Kamal Raj Pardasani[2]

[1]Department of Bioinformatics, [2]Department of Mathematics, Maulana Azad National Institute of Technology, Bhopal - 462051, Madhya Pradesh, India; Lavanya Rishishwar - Email: lavanyarishishwar@gmail.com: Phone: +91 9981519470; *Corresponding author

Abstract:
ProCoS is a free online tool for computing different combinations of peptide compositions. It is developed as an applet and a server with a capability to handle multiple FASTA sequences. The generalized algorithm for computing poly-amino acid composition forms the core of ProCoS. It produces output in different formats for easy visualization of results. It also allows composition analysis of sequences in full or in specific parts. Thus, ProCoS is user-friendly, flexible and unique.

Background:
Sequences in databases like GenBank are usually not fully annotated with functional information. Therefore, it is important to annotate such sequences with useful functional information using prediction tools and techniques. A number of such tools are already available in the public domain. The parameters that have been widely used are amino acid composition [1][2], pseudo amino acid composition [3][4], amphiphilic pseudo amino acid composition [5][6], di-peptide composition [7], and even combination of few of these [8]. These parameters have been applied in the prediction of sub cellular localization of proteins, functionally characterizing proteins and identification of proteins from 2D gels [9]. Calculation of these features is usually non-trivial in the context of functional inference. Here, we describe ProCoS: PROtein COmposition Server (v 2.0) as a freely available tool at http://www.bifmanit.org/procos.

Figure 1: ProCoS screenshot

Input/Output:
ProCoS can be accessed both as an applet and as a server. Both the versions provide users with an extensive and interactive GUI enabling full command on the system. It requires input: (1) peptide sequence(s) in FASTA format; (2) calculation type (single or multiple); (3) output in four formats (JTable; text table; text list; feature value vector); (4) description and (5) composition degree; (6) break mode [automated; manual; disabled]

Features and caveats:
ProCoS eliminates user migration from one tool to another as per changes in his/her needs. The two versions are provided so as to suit user's need. The applet is designed in Java and the server works on Perl-PHP backbone. The applet version is best suited for small input data with less memory and computational processing. It can work in offline mode with a JRE installed in client machine**.** Its cons are covered by the server edition, which doesn't have any such restrictions. But it is less interactive than the former one.

Future developments:
Facility for the calculation of numerous plots will be added in the next version.

References:
[1] AG Garrow *et al. BMC Bioinformatics* **6**:56 [2005]   [PMID: 15769290]
[2] K Hiyama *et al. J Biochem* **78**:6 [1975] [PMID: 1225917]
[3] C Chen *et al. Journal of Theoretical Biology* **243**:3 [2006] [PMID: 16908032]
[4] KC Chou *et al. Protein Sci* **4**:3 [2005] [PMID: 15952744]
[5] P Mundra *et al. Pattern Recognition Letters* **28**:13 [2007]
[6] KC Chou. *Bioinformatics* **21**:1 [2005] [PMID: 15308540]
[7] P Petrilli. *Bioinformatics* **9**:2 [1992]
[8] K Hiyama *et al. J Biochem* **78:** 6 [1975]
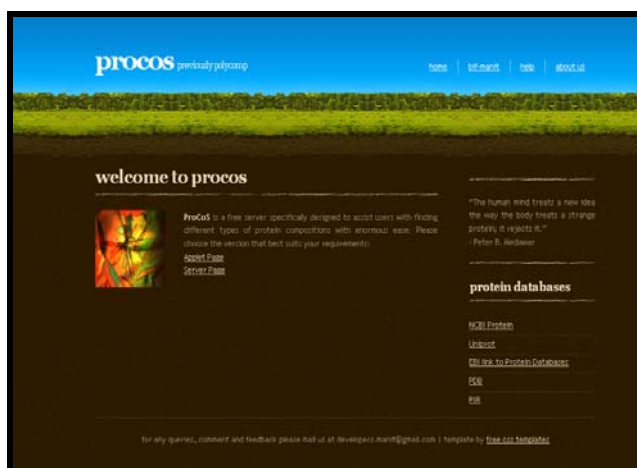[9] MR Wilkinson *et al. J Mol Biol* **278**:3 [1998] [PMID: 9600841]