ProPCalc –A TOOL FOR CALCULATING THE PROTEIN PROPERTIES

Boobalan C[1*] and Bharathi N [2]

[1]Department of Biotechnology, Indian Institute of Technology, Madras, Chennai 600036,INDIA.
[2]Department of Plant Molecular Biology and Biotechnology, Center for Plant Molecular Biology, Tamil Nadu Agricultural University, Coimbatore 641003, INDIA.

*- Corresponding Author

ABSTRACT:

Protein Property Calculator (*ProPCalc*) is a standalone tool developed by perl tk programming language which is used for calculating the physiochemical properties of a given protein deposited in Protein Data Bank (PDB). The parameters calculated by the tools include amino acid statistics, molecular weight, extension coefficient, optical density and GRAVY (Grand Average of Hydropathy). *ProPCalc* is available in http://code.google.com/p/tool-propcalc/.

Keywords: Protein properties, Protein property calculator (*ProPCalc*) ,Protein data bank (PDB), Practical extraction and report language (PERL).

INTRODUCTION

Protein Data Bank (PDB) is an internationally referred protein structure database, which comprises of the atomic coordinates of the three dimensional (3D) structure of proteins [1]. The coordinate file is organized in a common file format, called the PDB format [2]. Amino acid sequence and atom positions are organized systematically in the PDB file enabling machine learning [3].

Amino acid statistics

Calculating the statistics of individual amino acids based on the one letter code [4].

Molecular Weight

The molecular weight of the protein or peptide is shown in Daltons [5]. The formula used is:

MW = (nA x 71.07) + (nR x 156.18) + (nN x 114.08) + (nD x 115.08) + (nC x 103.10) + (nQ x 128.13) + (nE x 129.11) + (nG x 57.05) + (nH x 137.14) + (nI x 113.15) + (nL x 113.15) + (nK x 128.17) + (nM x 131.19) + (nF x 147.17) + (nP x 97.11) + (nS x 87.07) + (nT x 101.10) + (nW x 186.20) + (nY x 163.17) + (nV x 99.13) + 18.02

Extension Coefficient/Absorbance Factor

The absorbance factor is the conversion factor used to calculate the concentration of protein, using the measured absorbance at 280 nm (with a 1 cm light path) [6]. The following formula is used for calculation of extension coefficient/absorbance factor

Abs = (nW x 5500 + nY x 1490 + nC x 125) / MW

Optical density

The absorbance (optical density) is being calculated using the following formula:

Absorb(Prot) = E(Prot) / Molecular weight

Note-cysteine does not absorb appreciably at

wavelengths >260 nm, while cystine does

Aliphatic Index

The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine) [7]. The aliphatic index of a protein is calculated according to the following formula

Aliphatic index = X(Ala) + a * X(Val) + b * (X(Ile) + X(Leu))

Where X(Ala), X(Val), X(Ile), and X(Leu) are mole percent (100 X mole fraction) of alanine, valine, isoleucine, and leucine.

The coefficients a and b are the relative volume of valine side chain (a = 2.9) and of Leu/Ile side chains (b = 3.9) to the side chain of alanine.

GRAVY (Grand Average of Hydropathy)

The GRAVY value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence [8].

METHODS AND IMPLEMENTATION

ProPCalc was developed using Practical Extraction and Report Language (PERL) and Tool kit (Tk). PERL is used for scripting and Tk for designing graphical user interface (GUI). The user can select the desired protein structure via the browse button provided. The tool also provides an option to select interested protein properties for its calculation (Figure 1).

Here, we present the ProPCalc (**Pro**tein **P**roperty **C**alculator) standalone tool for calculating the protein properties of three dimensional structure of protein in Protein Data bank (PDB). The tool is user friendly and requires input file in PDB format for its calculation. ProPCalc is freely downloadable from http://code.google.com/ p/tool-propcalc/.

Figure 1. Flow chart illustrating the ProPCalc working process
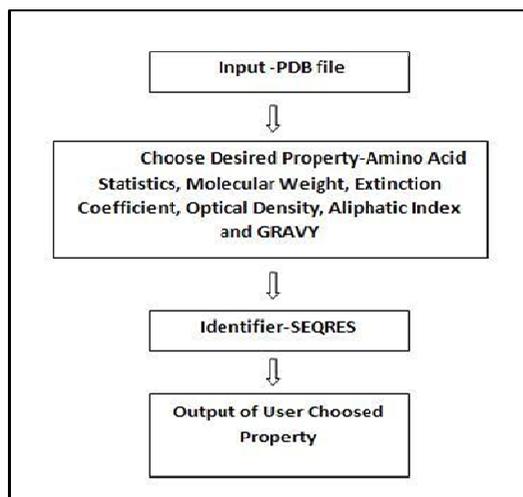
Figure 2. Home Page of ProPCalc with sample report

ACKNOWLEDGEMENTS

REFERENCES

1. HM. Berman., et al. (2000) The protein data bank. *Nucl. Acids Res.*, 28, 235-242 [PMID: 10592235]
2. K Henrick., et al.2008.Remediation of the protein data bank archive. *Nucl. Acids Res.*, (2008) 36:, D426 [PMID: 18073189]
3. S. R. Hall., et al. 1991.The STAR file: a new format for electronic data transfer and archiving. *J. Chem. Inf. Comp. Sci.*, (1991) 31: 326-333.

4. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.;*Protein Identification and Analysis Tools on the ExPASy Server;*(In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607.

5. Fasman, GD .1975. The Handbook of Biochemistry and Molecular Biology, CRC Press, 1:589.

6. Pace CN, Vajdos F, Fee L, Grimsley G, and Gray T., et al. 1995. How to measure and predict the molar absorption coefficient of a protein. Protein Sci, 4(10):2411-2423.

7. Ikai, A.J. ., et al. 1980. Thermostability and aliphatic index of globular proteins. J. Biochem. 88, 1895-1898. [PubMed: 7462208].

8. Kyte, J. and Doolittle, R.F. ., et al. 1982.A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132. [PubMed: 7108955]

Site References:
1. http://www.perl.org/
2. http://www.perltk.org/