WinGene/WinPep: User-Friendly Software for the Analysis of Amino Acid Sequences

Lars Hennig
University of Freiburg, Freiburg, Germany

ABSTRACT

WinGene1.0/WinPep1.2 is a pair of Microsoft® Windows® programs designed to read nucleotide or amino acid sequence data. These versatile programs have the following capabilities: (i) searches for open reading frames and their translation, (ii) assisting the design of primers for PCR and (iii) calculation of molecular weight, isoelectric point and molar absorption coefficients of polypeptides. Furthermore, hydropathic plots and helical wheel displays are easily produced. The programs run with an intuitive Windows interface, contain a comprehensive help file and enable data exchange with other applications by means of the Copy&Paste command. The software is free for academic and noncommercial users.

INTRODUCTION

The powerful tools of molecular biology and genome sequencing projects yield a rapidly growing number of new nucleic acid sequence data (2). To elucidate the mostly unknown functions of the encoded proteins, many experimental approaches start with the careful analysis of the deduced amino acid sequence. Although a multitude of World Wide Web (WWW)-based services for sequence analysis is available, a single stand-alone, PC-based program package is needed by many investigators. Some of the commonly experienced problems include slow network connections, difficulties localizing an appropriate service or the need to perform analysis with computers not connected to the internet.

The WinGene/WinPep programs were developed to circumvent these problems. They are also intended as a basic freeware alternative to expensive commercial software packages. The programs were written in C++ primarily for Windows® 95/98 (Microsoft, Redmond, WA, USA); they can also be used under Windows NT™ (Microsoft). WinGene/WinPep were designed to have a convenient, standard Windows-interface and to carry out a number of different (often required) tasks from a single sequence input. An extensive help system documents all calculations. An additional main goal was to enable easy export of data into other applications. Therefore, WinGene/WinPep is not only intended for analysis but also for data display. To ensure maximal flexibility, the programs utilize initialization (INI) files, containing all the parameters needed for calculations; moreover, these files can be modified with any text editor. Further important advantages of this program are that amino acid’s pKa values, hydropathic indices and other parameters can be easily changed within appropriate dialog boxes.

MATERIALS AND METHODS

Data input into WinGene/WinPep can be achieved manually in a text editor window, with the Copy&Paste command or by import from text files. Any numbers, spaces or special characters are deleted by the program. Moreover, WinGene performs the translation of nucleotide sequences into amino acid sequences and the transfer of the latter into WinPep. WinGene also contains the ability to reverse-complement a nucleotide sequence. A “primer-check” function calculates base composition and melting temperature of oligonucleotides as described by Baldino et al. (1).

WinPep offers several possibilities of amino acid sequence analysis. First, a set of basic physicochemical properties can be derived. This set includes (i) sequence length, (ii) frequencies of amino acid occurrences, (iii) isoelectric point [pKa values of Creighton (3)], (iv) molecular weight and (v) molar absorption coefficients for both native (7) and denatured (4) proteins. For more detailed analysis, a sequence-specific cleavage of the protein can be simulated. After selecting one of several preset proteases and chemical agents or providing a recognition sequence of the user’s choice, all possible
resulting peptide fragments are displayed with sequence position, length and molecular weight. This allows easy identification of fragments after HPLC and mass spectrometry. In addition, the sequence can be searched for the occurrence of any sequence motif.

The visualization of amphipathic helices, often involved in protein-protein interactions (3), can be achieved by the display of a (sub)sequence as a helical wheel in WinPep. To aid the identification of potential helices with strong amphipathy, a value referred to as “amphipathic moment” is calculated. This value is the gradient of the hydropathy of individual residues across an ideal helix. Consequently, values close to zero result from an even distribution of hydrophilic and hydrophobic amino acids around a hypothetical helix, while larger values indicate an asymmetric distribution. A plot of the amphipathic moment along the entire sequence can be displayed. While the amphipathic moment does not predict any structural elements, it constitutes a helpful tool for the design of biochemical experiments.

Hydropathic plots are commonly used to identify putative transmembrane regions of proteins (6). With WinPep, it is possible to display a hydropathic plot based on a running window average with variable window size. The scales of Kyte and Doolittle (6) or Sweet and Eisenberg (8) are predefined. Alternatively, individual scales can be provided. Thus, other tasks (e.g., the prediction of immunogenic epitopes) can individually be set up. Based on the currently active scale, the main hydropathy is calculated. Similarly, for the other graphic displays, hydropathic plots can be transferred into other applications through the clipboard (Copy&Paste). Figure 1 displays a screen of a hydropathic plot of cyclic nucleotide gated channel 1 (CNGC1). This protein of an as-yet-unknown function but with homology to animal cyclic nucleotide gated channels is encoded by a recently identified gene in the model plant *Arabidopsis thaliana* (5). Based on the display in Figure 1, the topology of six transmembrane helices similar to the channels in animals could be predicted. In contrast to the animal proteins, however, the fifth and sixth putative transmembrane helices are much further apart in CNGC1; this indicates that the plant protein could belong to a new, distinct family of channel proteins.

To summarize, WinGene/WinPep carry out a variety of often required tasks preceding and paralleling biochemical experiments. The intuitive user interface and the extensive export possibilities based on Copy&Paste make this program applicable to many needs.
AVAILABILITY

For academic and noncommercial usage, this program is available from the author free of charge. It is available at: (http://www.biologie.uni-freiburg.de/data/schaefer/lhennig/winpep.html). It is also available from the Software Library on the BioTechniques Web site (http://www.BioTechniques.com).

ACKNOWLEDGMENTS

This work was partly supported by the Graduiertenkolleg “Molekulare Mechanismen pflanzlicher Differenzierung”. The author thanks Claudia Köhler for discussions and extensive testing of the software.

REFERENCES


Received 16 February 1999; accepted 5 April 1999.

Address correspondence to:
Lars Hennig
Department of Botany
University of Freiburg
Schänzlestr. 1
79104 Freiburg, Germany
Internet: lhennig@sun2.ruf.uni-freiburg.de