Structural bioinformatics

HELIQUEST: a web server to screen sequences with specific \( \alpha \)-helical properties

Romain Gautier*, Dominique Douguet, Bruno Antonny and Guillaume Drin

Institut de Pharmacologie Moléculaire et Cellulaire, Université de Nice Sophia Antipolis and CNRS, 660 route des Lucioles, 06560 Valbonne, France

Received on April 17, 2008; revised on July 16, 2008; accepted on July 25, 2008

Advance Access publication July 28, 2008

Associate Editor: Anna Tramontano

ABSTRACT

Summary: HELIQUEST calculates the physicochemical properties and amino acid composition of an \( \alpha \)-helix and screens databank to identify protein segments possessing similar features. This server is also dedicated to mutating helices manually or automatically by genetic algorithm to design analogues of defined features.

Availability: http://helquest.ipmc.cnrs.fr

Contact: gautier@ipmc.cnrs.fr

1 INTRODUCTION

Several proteins bind to membranes via a small amphipathic helix with one face made of hydrophobic residues that insert between the lipid acyl chains and the other one containing polar residues that interact with the lipid polar heads and solvent (Cornell and Taneva, 2006). This structural motif and its mode of membrane binding could appear simplistic, compared to well-structured domains (PH, PX, FYVE) that recognize specific lipids (Lemmon, 2008). However, depending on their sequence, amphipathic helices have different properties: some form holes in membrane as certain antimicrobial peptides (Dathe and Wieprecht, 1999); others permit proteins of vesicular transport to modify or detect membrane shape (Drin et al, 2007). This result illustrated how ALPS-like helices (amphipathic, transmembrane, etc.) and use the results as a starting point to extract putative equivalent helices in unexpected proteins. An additional module was developed to manually mutate a helix or to automatically design analogues by genetic algorithm.

2 METHODS

2.1 Sequence analysis and database screening

A sequence submitted by the user, considered as helical, is analyzed by a sliding window (14-54 aa, i.e. up to three repeats of a complete helical wheel of 18 aa). The analysis module displays for each generated segment a table reporting its net charge and hydrophobicity and hydrophobic moment and calculated with a standard hydrophobic scale (Fauchere and Pliska, 1983), as well as statistics on its composition (percentage or enumeration of specific residues). A helical wheel representation of each segment with its hydrophobic moment vector is downloadable.

If user analyzes sequences with an 18 aa window, each table displays a link to the screening module. Values calculated on the selected segment and to specify if the protein segment to be extracted must contain a minimal number of polar residues (E, D, K, R, S, T, N, H, Q+G), of charged residues (E, D, K, R) and of specific polar residues (S, T, N, Q, H). Protein segments containing proline at their ends, cysteine or both can be either accepted or excluded. The screening module is directly accessible if the parameters are known. For example, we identify ALPS-like motifs by limiting the number of charged residues and a sum of serine, threonine and glycine superior to 6; cysteine was excluded.

User screens either personal or annotated SWISSPROT databases (Boeckmann et al., 2003); in this case, one can discard poorly...
Any characterized helix can be mutated either manually (and reanalyzed to (Fig. 1). A procedure exists to refine the screening of amphipathic helices sequence of the databank and save segments fulfilling the required criteria defined proteins. The algorithm slides an 18 aa window along each protein. This algorithm detects the existence of an uninterrupted hydrophobic face of at least five residues adjacent on a helical wheel. If such a face exists, it verifies whether the facing residues are polar.

An output text file lists the proteins containing at least one positive segment, whose sequence, localization, physicochemical features and content in amino acid are reported. If several segments in one protein overlap or are adjacent, they are merged into a unique sequence, reported in a second file. Corresponding PDF files contain helical wheel representations of all sequences. A decision tree, integrating results from TMHMM (Krogh et al., 2001), PSIPRED (Jones, 1999) and a discriminant analysis performed on lipid-binding helices, order sequences in six classes. A sequence is classified for example as a helix, a lipid-binding helix or as a non-relevant sequence with a high propensity to form a β-sheet. With appropriate screening parameters, our procedure was found to identify transmembrane segments from the MPtopo database (Jayasinghe et al., 2001) and highly amphipathic helices (with a <μH> superior to 0.6) from a subset of non-redundant PDB with a positive predictive value of 95 and 86%, respectively. Finally, we noted that the screening procedure was able to identify a majority of lipid-binding helices extracted from a small dataset of known perimembrane proteins.

2.2 Helix mutation

Any characterized helix can be mutated either manually (and reanalyzed to examine how the mutation changes its features) or automatically by genetic algorithm (GA). As <H>-, <μH> and z are interdependent properties, changing manually one property without modifying others is difficult (Dathe and Wieprecht, 1999). The GA-based module allows modifying independently or simultaneously these parameters with a minimal number of mutations. Alternatively, user can impose an amino acid composition within an α-helix under defined <H> and <μH> constraints. Generating sequences de novo with precise features is possible.

2.3 Implementation

HELIQUEST, written in Python 2.5, is organized as interconnected CGI programs and use R language to draw helical wheel.

Funding: Agence Nationale pour la Recherche; the CNRS (RISC).

Conflict of Interest: none declared.

REFERENCES


