BIOINFORMATICS APPLICATIONS NOTE Vol. 20 no. 17 2004, pages 3258–3260
doi:10.1093/bioinformatics/bth358

TMRPres2D: high quality visual representation of transmembrane protein models

Ioannis C. Spyropoulos, Theodore D. Liakopoulos, Pantelis G. Bagos and Stavros J. Hamodrakas*

Department of Cell Biology and Biophysics, Faculty of Biology, University of Athens, Panepistimiopolis, Athens 157 01, Greece

Received on April 5, 2004; revised on May 25, 2004; accepted on June 4, 2004
Advance Access publication June 17, 2004

ABSTRACT

Summary: The 'TransMembrane protein Re-Presentation in 2-Dimensions' (TMRPres2D) tool, automates the creation of uniform, two-dimensional, high analysis graphical images/models of alpha-helical or beta-barrel transmembrane proteins. Protein sequence data and structural information may be acquired from public protein knowledge bases, emanate from prediction algorithms, or even be defined by the user. Several important biological and physical sequence attributes can be embedded in the graphical representation.
Availability: http://bioinformatics.biol.uoa.gr/TMRPres2D
Contact: shamodr@cc.uoa.gr

INTRODUCTION

Transmembrane proteins are a major target of current research due to their vital importance in cell functions and disease. This tool was designed to assist molecular biologists with the creation of schematic drawings of transmembrane protein models for research, supervisory or educational purposes. Although several tools, which apparently satisfy the same needs, are already available (Beitz, 2000; Campagne and Weinstein, 1999; Lin and Hwang, 1998; Michalopoulos *et al*., 2004; Rice *et al*., 2000; Skrabanek *et al*., 2003), our method offers several characteristic unique features:

(1) Visualization of both alpha-helical and beta-barrel transmembrane proteins.
(2) 'One-step' data input procedure, including easy querying of the public databases and incorporation of output from various popular prediction tools available from the Web.
(3) Accurate dimensions of the molecular elements, with adjustable scaling.
(4) Representation of any available biological and physical attributes, using chromatic and symbolic variations (e.g. disulfide bridges, signal peptide).
(5) Ability to easily incorporate custom, user-defined annotations for any amino acid residue.
(6) Wide variety of supported vector and bitmap-based image formats.
(7) Integrated graphical user interface.
(8) Intelligible and elegantly styled depiction.

DATA INPUT

The data required for creation of an image are the exact amino acid sequence of the protein chain and the boundaries of one or more transmembrane segments. However, additional information may be utilized. Data input may be given utilizing various sources: (1) Public protein knowledge bases, namely SWISS-PROT (Boeckmann *et al*., 2003), PIR (Wu *et al*., 2003) and UniProt (Apweiler *et al*., 2004), (2) the output of prediction algorithms like HMMTOP (Tusnády and Simon, 2001) TMHMM (Krogh *et al*., 2001), PRED-TMR (Pasquier *et al*., 1999) and PRED-TMBB (Bagos *et al*., 2004) or (3) data provided manually by the user, via a special input form. Entries from external database sources may be obtained from the Web, by querying the corresponding hosting server with an entry identifier from within the application environment, or alternatively loaded from the local disc.

OUTPUT IMAGE

Image creation is made with the assumption that each transmembrane segment consists of one of the two possible types of secondary structure, namely an alpha-helix or a beta-strand that enters the membrane from one side and emerges from the other. Non-transmembrane regions are displayed as two-dimensional loose loops (Fig. 1).

A concrete scale of distances (pixels/Ångstrom) is applied to the attributes of every depicted element, namely lipid size, inter-lipid distance, helix radius and step, number of residues per turn, distance between beta-strands and atoms distance. The above values comply with theoretically acceptable atom distances from observations of

*To whom correspondence should be addressed.

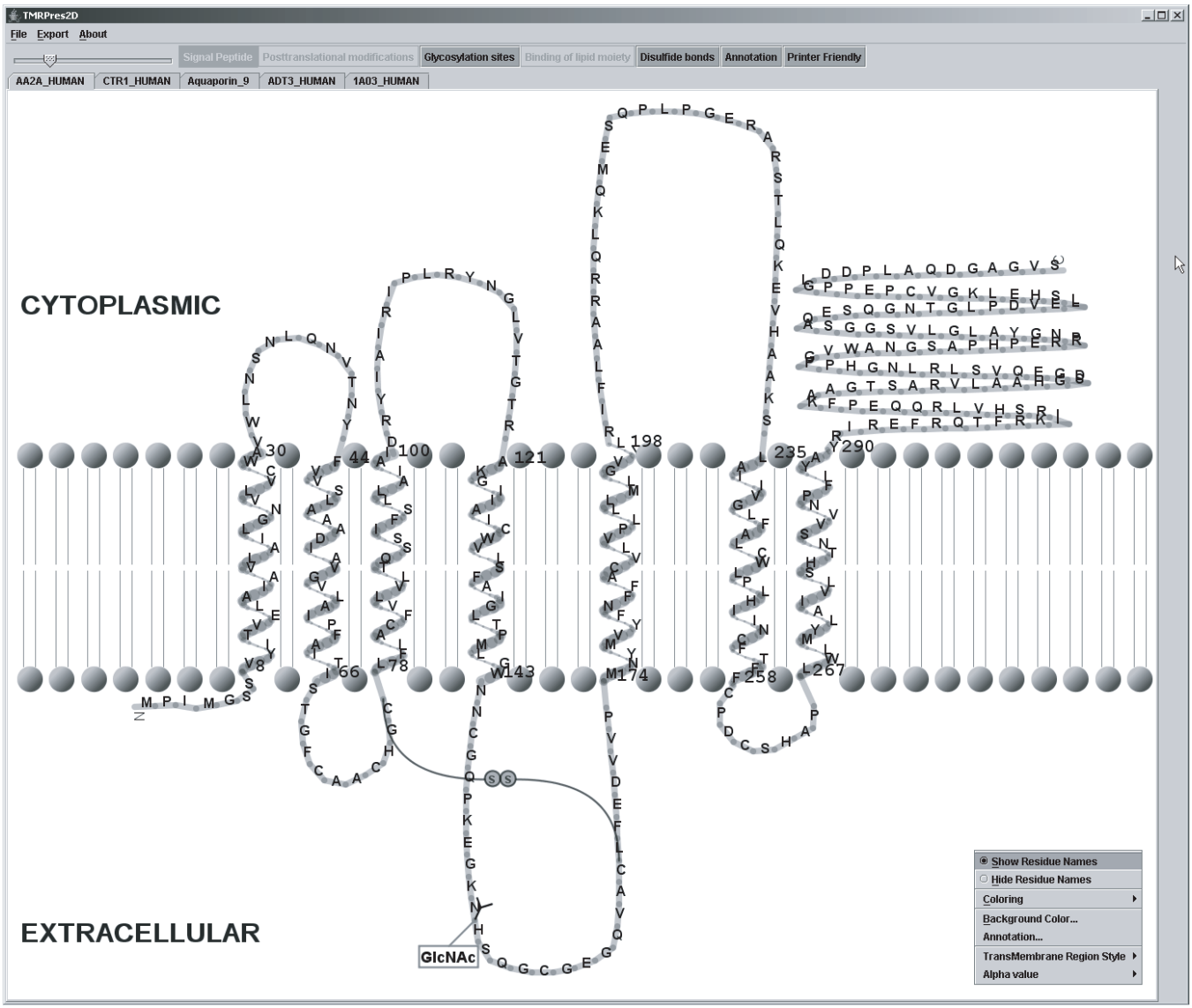

Fig. 1. A model of human adenosine A2a receptor (AA2A_HUMAN) utilizing TMRPres2D, in black and white ('printer friendly').

known structures (Andersen and Rost, 2003). The scale factor is adjustable and can be used for any scaling transformation.

Elementary data, such as transmembrane segment boundaries, cellular topology (if known) and sequence termini are always visible. However, additional image features may be customized to satisfy the user's special needs: amino acid residue types may be omitted from the depiction. The degree of transparency of the peptide chain may vary. Coloring may be based on: (1) an *ad hoc* hydrophobicity scale (Pasquier *et al.*, 1999), using a gradation from yellow to blue, (2) electrical charge (assuming pH = 7), using blue for positive, red for negative and gray for uncharged residues or (3) custom user-defined colors. In addition, an alternative 'printer-friendly' appearance option is available.

A wide range of biological attributes, if available in the input data, may be embedded and emphasized characteristically. These include disulfide bridges, signal peptides, posttranslational modifications, glycosylation sites and lipid moiety binding.

USER INTERACTION

User interaction, accomplished with the use of the mouse pointer, concerns directional movement and non-disfiguring scaling of the image. Besides, a pull-down menu serves for any possible combination of the available appearance choices. Custom annotations, in the form of 'labels', may be added to any amino acid residue. The produced picture may be saved using several major vector-based formats, namely

Encapsulated PostScript, Portable Document Format (PDF), PostSctipt and scalable vector graphics (SVG). Alternatively, pictures may be exported using the popular bitmap image formats JPEG, PNG.

The application has been implemented using the platform-independent Java language by Sun Microsystems. Consequently, the tool can be used as an applet too, and thus can serve to exhibit transmembrane protein schematic models in Web pages. For further information and downloading the application, please, visit http://bioinformatics.biol.uoa.gr/TMRPres2D.

ACKNOWLEDGEMENTS

We thank the anonymous referees and the editor for their useful comments.

REFERENCES

Andersen,C.A. and Rost,B. (2003) Secondary structure assignment. *Meth. Biochem. Anal.*, **44**, 341–363.

Apweiler,R., Bairoch,A., Wu,C.H., Barker,W.C., Boeckmann,B., Ferro,S., Gasteiger,E., Huang,H., Lopez,R., Magrane,M., *et al*. (2004) UniProt: the Universal protein knowledgebase. *Nucleic Acids Res.*, **32**, D115–D119.

Bagos,P.G., Liakopoulos,T.D. Spyropoulos,I.C., and Hamodrakas,S.J. (2004) A hidden Markov model capable of predicting and discriminating β-barrel outer membrane proteins. *BMC Bionformatics*, **5**, 29.

Beitz,E. (2000) TeXtopo: shaded membrane protein topology plots in LaTeX2e. *Bioinformatics*, **16**, 1050–1051.

Boeckmann,B.,Bairoch,A.,Apweiler,R.,Blatter,M.-C.,Estreicher,A., Gasteiger,E., Martin,M.J., Michoud,K., O'Donovan,C., Phan,I., Pilbout,S. and Schneider,M. (2003) The Swiss-Prot protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.*, **31**, 365–370.

Campagne,F. and Weinstein,H. (1999) Schematic representation of residue-based protein context-dependent data: an application to transmembrane proteins. *J. Mol. Graph. Model*. **17**, 207–213.

Krogh,A., Larsson,B., von Heijne,G. and Sonnhammer,E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.*, **305**, 567–580.

Lin,W.J. and Hwang,M.J. (1998) VHMPT: a graphical viewer and editor for helical membrane protein topologies. *Bioinformatics*, **14**, 866–868.

Michalopoulos,I., Torrance,G.M., Gilbert,D.R. and Westhead,D.R. (2004) TOPS: an enhanced database of protein structural topology. *Nucleic Acids Res.*, **32** (Database issue), D251–D254.

Pasquier,C., Promponas,V.J., Palaios,G.A., Hamodrakas,J.S. and Hamodrakas,S.J. (1999) A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. *Protein Eng.*, **12**, 381–385.

Rice,P., Longden,I. and Bleasby,A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.*, **16**, 276–277.

Skrabanek,L., Campagne,F. and Weinstein,H. (2003) Building protein diagrams on the web with the residue-based diagram editor RbDe. *Nucleic Acids Res.*, **31**, 3856–3858.

Tusnády,G.E. and Simon,I. (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics*, **17**, 849–850.

Wu,C.H., Yeh,L.L., Huang,H., Arvninski,L., Castro-Alvear,J., Chen,Y., Hu,Z.-Z., Ledley,R.S., Kourtesis,P., Suzek,B.E. *et al*. (2003) The protein information resource. *Nucleic Acids Res.*, **31**, 345–347.