BIOINFORMATICS APPLICATIONS NOTE Vol. 17 no. 12 2001
Pages 1236–1237

ProPred: prediction of HLA-DR binding sites

Harpreet Singh and G. P. S. Raghava*

Bioinformatics Centre, Institute of Microbial Technology, Sector 39A, Chandigarh-160036, India

Received on April 24, 2001; revised and accepted on June 21, 2001

ABSTRACT

Summary: ProPred is a graphical web tool for predicting MHC class II binding regions in antigenic protein sequences. The server implement matrix based prediction algorithm, employing amino-acid/position coefficient table deduced from literature. The predicted binders can be visualized either as peaks in graphical interface or as colored residues in HTML interface. This server might be a useful tool in locating the promiscuous binding regions that can bind to several HLA-DR alleles.

Availability: The server is available at http://www.imtech.res.in/raghava/propred/

Contact: raghava@imtech.res.in

Supplementary information: http://www.imtech.res.in/raghava/propred/page3.html

INTRODUCTION

Advances in the area of antigen processing and presentation has made subunit vaccines an integral part of vaccine design strategy. In the subunit vaccine, the vaccine candidates are immunogenic peptides/regions of protein instead of complete protein (Hagman, 2000). It is well established that binding of a peptide to an MHC molecule is a prerequisite for activation of antigen specific T-cells. As only certain peptides can bind to a given MHC molecule, the identification of these peptides is one of the bottlenecks in subunit vaccine design. In the past, a number of methods have been developed to predict MHC binding peptides. These include matrix-based methods (Hammer *et al.*, 1997), whose performance relies on the quality of matrix.

Recently, Sturniolo *et al.* (1999) described a novel approach called pocket profiles for constructing matrices. They generated matrices for a large number of human HLA-DR alleles from a small set of experiments and using these matrices developed a matrix based computer program called TEPITOPE, for windows. So far, the beta version of TEPITOPE is available only for 25 HLA-DR alleles. It has been used successfully in locating T-cell epitopes in the melanoma-associated glycoprotein and in MAGE-3 antigen (Manici *et al.*, 1999; Cochlovius *et al.*, 2000). The availability of such an important algorithm on the Internet will be useful for vaccine immunologists. With this objective, we developed this graphics based web server for 51 HLA-DR alleles. For this we extracted the matrices for 51 HLA-DR alleles from a pocket profile database described by Sturniolo *et al.* (1999). The threshold values for 25 alleles were obtained from TEPITOPE, while the threshold values corresponding to the remaining 26 alleles were generated from PIR database, as described in TEPITOPE.

The server uses ReadSeq program (developed by Dr Don Gilbert) to read the input sequence, thus it can accept most commonly used standard sequence formats (e.g. FASTA, PIR). The sequence can be uploaded, from a file or by using the cut and paste option. Users can customize the server by selecting single/multiple alleles, threshold and other parameters in order to achieve desirable results. The server analyzes sequence data and generates output as text or graphics (Figure 1). The text display has two options where the first option is similar to TEPITOPE in which the binding regions are shown by different colors. Although it is easy to detect binding regions in antigenic sequence using the above option, yet it is less expressive in presenting the overlapping binding regions. In the second option, the server presents overlapping regions individually on separate lines, making it easier to detect these regions from the display.

The graphical output presents the HLA-DR binding propensity along the primary structure of a protein. The server generates graphics in GIF format using GDPlot library (developed by Lincoln D. Stein). The graphics has an advantage over the text presentation that in addition to locating HLA-DR binders, it also presents their binding strength. Besides, the server also has option to plot the threshold profile (Threshold versus binding peptides), which can assist the users in selecting an appropriate threshold for locating promiscuous binders. The server also allows performing analysis on any selected sub-sequence. During subsequence analysis, in addition to analyzing the best scoring peptide frame (e.g. TEPITOPE), the server computes the binding strength of all peptide frames in the selected subsequence. The server

*To whom correspondence should be addressed.

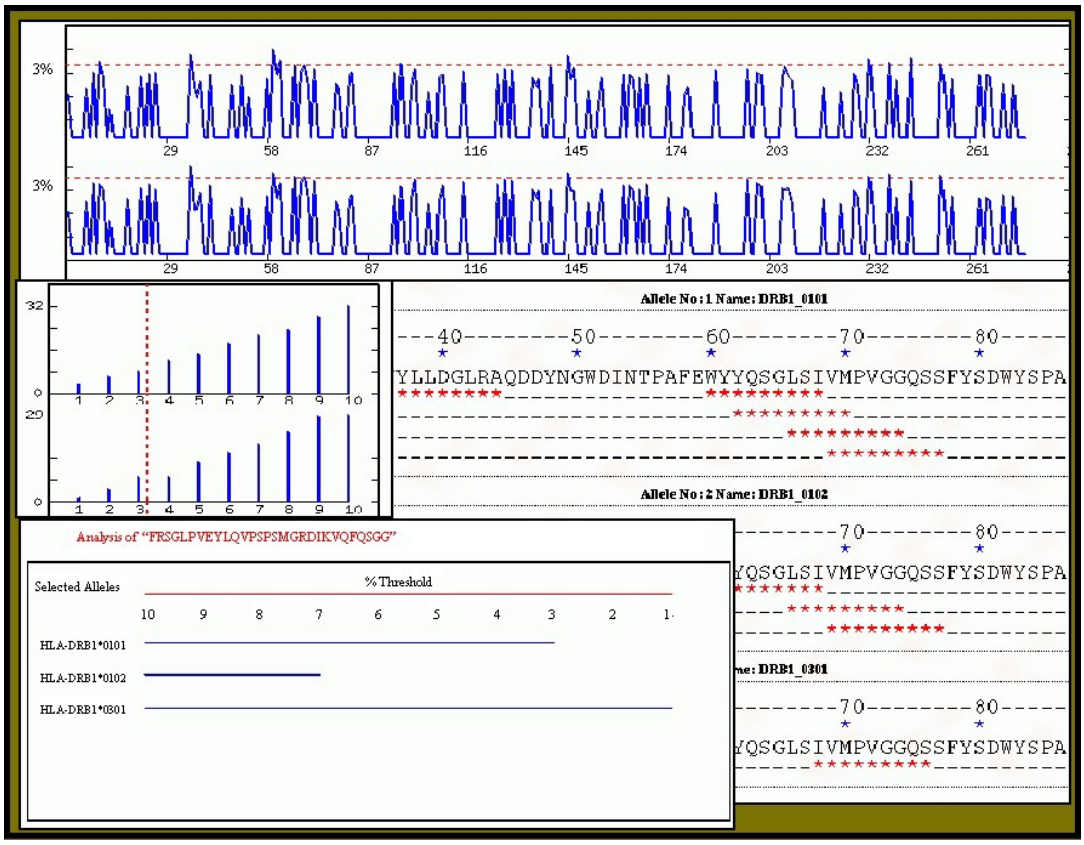

Fig. 1. Example screenshots from the ProPred server showing graphical and text output formats.

performs analysis for each HLA-DR allele independently and presents the output on a single screen, which will assist the user in rapid visualization of promiscuous binders. Promiscuous binders are those peptides that can bind to several HLA-DR molecules, and are important for subunit vaccine design. The server also provides an option to present the binding score calculated from matrix in a commonly used tabular format. In brief, the server is intended for a comprehensive analysis of an antigenic sequence, thereby assisting the vaccine immunologist in selection of a potential vaccine candidate.

ACKNOWLEDGEMENTS

We are grateful to Dr G.C.Varshney of our Institute for some useful discussions. The authors are also grateful to Council of Scientific and Industrial Research (CSIR) and Department of Biotechnology (DBT), Govt of India for financial assistance. This report has IMTECH communication No. 009/2001.

REFERENCES

Cochlovius,B., Stassar,M., Chriast,O., Raddrizzani,L., Hammer,J., Mytilineos,I. and Zoller,M. (2000) *In vitro* and *in vivo* induction of a Th cell response toward peptides of the melanoma-associated glycoprotein 100 protein selected by the TEPITOPE program. *J. Immunol.*, **165**, 4731–4741.

Hagman,M. (2000) Computer aided vaccine design. *Science*, **290**, 80–82.

Hammer,J., Sturniolo,T. and Sinigaglia,F. (1997) HLA class II peptide binding specificity and autoimmunity. *Adv. Immunol.*, **66**, 67–100.

Manici,S., Sturniolo,T., Imro,M.A., Hammer,J., Sinigaglia,F., Noppen,C., Spagnoli,G., Mazzi,B., Bellone,M., Dellabona,P. and Protti,M.A. (1999) Melanoma cells present a MAGE-3 epitope to CD4+ cytotoxic T-cells in association with HLA-DR11. *J. Exp. Med.*, **189**, 871–876.

Sturniolo,T., Bono,E., Ding,J., Raddrizzani,L., Tuercei,O., Sahin,U., Braxenthaler,M., Gallazzi,F., Protti,M.P., Sinigaglia,F. and Hammer,J. (1999) Generation of tissue-specific and promiscuous HLA ligand database using DNA microarrays and virtual HLA class II matrices. *Nature Biotechnol.*, **17**, 555–561.