MHC: Peptide Analysis: Implications on the Immunogenicity of Hantaviruses' N protein

Maurício Menegatti Rigo, Dinler Amaral Antunes, Gustavo Fioravanti Vieira, and José Artur Bogo Chies

Núcleo de Bioinformática, Department of Genetics, UFRGS, Porto Alegre, Brazil
{mauriciomr1985,dinler}@gmail.com, gusfioravanti@yahoo.com.br, jabchies@terra.com.br

Abstract. Hantaviruses, members of the *Bunyaviridae* family, are enveloped negative-stranded RNA viruses with tripartite genomes – S, M and L. The S genome codes for a nucleocapsid (N) protein, which is quite conserved among different species of the hantavirus genus and possess a recognized immunogenicity. In this work we analyzed the sequence of two regions in this protein ($N_{94\text{-}101}$ and $N_{180\text{-}188}$), which presents T cell epitopes for two species of hantaviruses – Sin Nombre and Puumala. Interestingly, the same region has no described epitopes for Hantaan virus, despite its similarity. A study using a bioinformatic approach for the construction of MHC:peptide complexes was performed to detect any variation on the cleft region that could explain such degrees of immunogenicity. Our results shown topological and charges differences among the constructed complexes**.**

Keywords: MHC-I, epitopes, Hantavirus, molecular docking.

1 Introduction

As members of the *Bunyaviridae* family, hantaviruses have a tripartite ssRNA(−) genome coding for a RNA-dependent RNA polymerase, two glycoproteins which are inserted into the viral envelope membrane, and the N protein associated with the viral genome [1]. The hantavirus nucleocapsid (N) protein fulfills several key roles in virus replication and assembly [1]. Also, it presents cross-reactivity among different members of the hantavirus genus [2]. Considering the recognized immunogenicity of this protein and that the final objective of our work is the development of a vaccine, two aspects should be here considered: the analysis of antigenic processing pathway and the presentation of epitopes to T cell on the MHC class I context. In a previous work (personal communication) we observed a concordance between conserved regions in the N protein and epitopes of hantaviruses described in literature [2]. Maeda *et al.* described epitopes for Puumala virus (PUU) and Sin Nombre virus (SNV) at region $N_{94\text{-}101}$ and $N_{180\text{-}188}$, however epitopes for Hantaan virus (HTN) in the same region were not observed, although the high similarity between these sequences (75% and 67%, respectively). Additionally, *in silico* simulations showed that epitopes of HTN

K.S. Guimarães, A. Panchenko, T.M. Przytycka (Eds.): BSB 2009, LNBI 5676, pp. 160–163, 2009.

in these regions can be generated by the antigen processing pathway in the same way as PUU and SNV. This data suggests a strong influence from some amino acids of the epitopes on the TCR recognition and on immune response induction. Therefore, the construction of a MHC:peptide complex for topological and charge analysis of these regions is quite important.

In the present work we analyzed two specific regions of the N protein (94-101/180-188) from three different species of hantaviruses – PUU, SNV and HTN. This analysis was made on modeled MHC:peptide complexes, mainly on the TCR surface contact, searching for charges and topological differences that could explain the different levels of immunogenicity observed among the epitopes from different hantaviruses.

2 Material and Methods

The epitope peptide sequences related to N_{94-101} and $N_{180-188}$ region from PUU, SNV and HTN were obtained from literature and written in the FASTA format. The crystal structure of murine MHC alleles, H2-K^b (1G7P) and H2-D^b (1WBY), were used to modeling the complexes of interest. Each peptide sequence was fitted on the parameters of the specific MHC allele epitope pattern using the SPDBV program [3]. Since we have the amino acids sequences within the parameters of three-dimensional shape, energy minimization was performed with the GROMACS 3.3.3 program [4] to stabilize the molecule, simulating an aqueous solution. The next step was a construction of the MHC:peptide complex using AutoDock 4.0 program [5]. After the construction of the first complex, a second energy minimization was performed aiming to a better interaction of the side chains of the MHC with the epitope. The MHC was separated from its epitope and a second docking was carried out for the construction of the final MHC:peptide complex. The surface of the resulting complex was visualized with the GRASP2 program [6] (Figure 1), where the electrostatic charge distribution and the shape were analyzed.

3 Results

All FASTA sequences were perfect fitted with PDB allele-specific peptides. Information about the positions of the anchor residues for H2-K^b and H2-D^b was obtained from SYFPEITHI [7], a MHC-ligand databank. Epitopes described in literature for SNV and PUU at the studied regions have the same sequence, therefore they were analyzed as a unique epitope. After two rounds of molecular docking and energy minimization we found good values of binding energy (BE) for each region (Table 1). At the topological and charge levels, categorical differences were found in MHC:peptide complex for N_{94-101} region, on the discordant residues. These differences could be important in TCR recognition. The $N_{180-188}$ region showed almost none differences at charge distribution, but a topological discrepancy was verified.

Table 1. Best binding energy values for each region of Hantavirus species

Hantavirus specie	Protein/Region	Peptide Sequence*	MHC allele	Best BE-1st docking (Kcal/mol)	Best BE - 2nd docking (Kcal/mol)
Hantaan	N_{94-101}	SMLSYGNV	H2-K^b	-6,11	-6,48
Sin Nombre & Puumala	N_{94-101}	SSLRYGNV	H2-K^b	-4,98	-7,49
Hantaan	$N_{180-188}$	SLPNAQSSM	H2-D^b	-11,07	-14,73
Sin Nombre & Puumala	$N_{180-188}$	SMPTAQSTM	H2-D^b	-12,01	-14,55

*Anchor amino acids are underlined.

4 Discussion

The described epitopes for N_{94-101} region are presented by H2-K^b allele, while $N_{180-188}$ are presented by H2-D^b allele [2]. There are a difference of only two amino acids at N_{94-101} region between SNV/PUU and HTN; at $N_{180-188}$ region, the difference is 3 amino acids. These substitutions seems to affect the recognition by the immune system, since there are no epitopes described for hantavirus in that region [2].

A first approach analyzed the binding energy value of molecular docking with peptides from these two regions of SNV/PUU and HTN. The computational program used was AutoDock. Values for binding energy were all negative. There are no data in literature about the best value for binding energy of peptide and MHC. However, this value is directly proportional to the entropy of the system. Thus, we admitted that

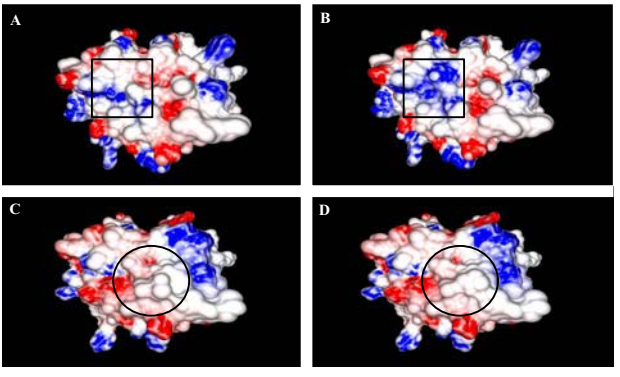

Fig. 1. Charge distribution of MHC:peptide complexes (Top-view) visualized at GRASP program. Peptide sequences from HTN (A) and SNV/PUU (B) on H2-Kb allele in N94-101 region; and from HTN (C) and SNV/PUU (D) on H2-Db allele in N180-188 region. Black squares show charge and topological differences between the first two complexes. Black circles show only topological difference on the remaining complexes.

the lowest value should be the best value. Accordingly to Binding energy data, all peptides have a good potential for attachment to the MHC. The conservation of anchor amino acids among the sequences for each MHC restriction could explain these good binding values.

It is known that changes in the distribution of charges interfere with the TCR recognition [8]. The best MHC:peptide models were visualized and analyzed at GRASP program, which provides a molecular surface with charge distribution, through the adjust of electric potential (Figure 1). The dark regions represent charged zones (both negative or positive charges). We could observe a topological and, especially, a charge difference between the MHC:peptide complexes of $N_{94\text{-}101}$ region, mainly on the fourth (Ser/Arg) residue. The $N_{180\text{-}188}$ region showed only topological changes, and the charges seems to remain equally distributed in both complexes. Our results showed that more than just charge differences, topological differences could explain the abrogation of an immune response.

References

1. Tischler, N.D., Rosemblatt, M., Valenzuela, P.D.T.: Characterization of Cross-reactive and Serotype-specific Epitopes on the Nucleocapsid Proteins of Hantaviruses. Virus Res. 135, 1–9 (2008)
2. Maeda, K., West, K., Toyosaki-Maeda, T., Rothman, A.L.: Identification and Analysis for Cross-reactivity Among Hantaviruses of H-2b-restricted Cytotoxic T-lymphocyte Epitopes in Sin Nombre Virus Nucleocapsid Protein. J. Gen. Virol. 85, 1909–1919 (2004)
3. Guex, N., Peitsch, M.C.: SWISS-MODEL and the Swiss-PdbViewer: an Environment for Ccomparative Protein Modeling. Electrophoresis 18, 2714–2723 (1997)
4. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.: GROMACS: Fast, Flexible, and Free. J. Comput. Chem. 26, 1701–1718 (2005)
5. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J.: Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. J. Comput. Chem. 19, 1639–1662 (1998)
6. Petrey, D., Honig, B.: GRASP2: Visualization, Surface Properties, and Electrosttics of Macromolecular Structures and Sequences. Meth. Enzymol. 374, 492–509 (2003)
7. Rammensee, H., Bachmann, J., Emmerich, N.N., Bachor, O.A., Stevanovic, S.: SYFPEITHI: Database for MHC Ligands and Peptide motifs. Immunogenetics 50, 213–219 (1999)
8. Kessels, H.W., de Visser, K.E., Tirion, F.H., Coccoris, M., Kruisbeek, A.M., Schumacher, T.N.: The Impact of Self-tolerance on the Polyclonal CD8+ T-cell Repertoire. J. Immunol. 172, 2324–2331 (2004)