Positional Biases of the Experimentally Characterized T-cell Epitopes

Mirjana D. Pavlović[1], Ana M. Jelović[2], Davorka R. Jandrlić[3], and Nenad S. Mitić[4]

[1]Institute of General and Physical Chemistry, University of Belgrade, Studentski trg 12, 11000 Belgrade, Serbia
[2]Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia
[3]Faculty of Mechanical Engineering, University of Belgrade, Kraljice Marije 16, Belgrade, Serbia
[4]Faculty of Mathematics, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia
nenad@matf.bg.ac.rs

Positional Biases of the Experimentally Characterized T-cell Epitopes

INTRODUCTION

Binding to Major Histocompatibility Complex (MHC) class I and class II molecules is the most restrictive step in T-cell epitope presentation. In humans these molecules are named Human Leukocyte Antigen (HLA) class I and II. However, only a minor part of MHC-binding epitopes are immunogenic. Thus, the most challenging part of T-cell prediction algorithms is to distinguish between immunogenic and nonimmunogenic epitopes. Using HLA-binding data and T-cell functional assay data we have analyzed positional biases and hydropaty of T-cell epitopes in predicted structured and unstructured regions of protein antigen(Ag), in an attempt to find out differences between T-cell immunogenic and nonimmunogenic epitopes.

MATERIAL AND METHODS

Data bases and epitope prediction tools: The Immune Epitope Database (IEDB, http://www.iedb.org) hosts the largest set of experimentally characterized T-cell immune epitope data, as well as data on peptides tested for MHC binding, 'MHC binders'. Both positive and negative experimental results were captured. Epitopes were divided in 2 different datasets: a) epitopes tested for binding to HLA class-I or class II molecules (positive and negative) and b) T-cell functional assay positive/negative (immunogenic/non-immunogenic) epitopes (with known HLA specificity). First dataset contained HLA-binding epitopes from 696 proteins. Second dataset contained 77663 epitopes from 5555 proteins. **Disorder prediction methods**: In order to eliminate the impact of the individual disorder prediction methods on the positions of epitopes, we have used the consensus of 7 different disorder predictors with the total of 9 variants. The set of predictors includes: IsUnstruct, VSL2b, DisEMBL_Remark465, DisEMBL_Hot_loops, IUPred-L, IUPred-S, and RONN. Consensuses of protein regions were defined by 7 disorder predictors as: disorder (D), and order (O). **Peptide epitope hydropathy** (Average number of hydrophobic AA/epitope) was determined using Kyte–Doolittle (KD) hydrophobicity scale, by counting the majority of hydrophobic/hydrophilic amino acids (AA) in each epitope.

RESULTS AND DISCUSSION:

The frequency of epitopes/region length, which bind to either HLA-I class or class HLA-II molecules is much higher in the consensus of structurally stable (O) regions, defined by different disorder predictors (or majority of these predictors), than in the consensus of structurally unstable (D) regions. HLA-I binding epitopes are more frequent than HLA-II binders in O than in D regions. Epitopes, negative for binding to HLA-I molecules (HLA-I non-binders), were also concentrated in the consensus of O regions, being several times more frequent than HLA-I binders, possibly due to overestimation in epitope prediction. The frequency of HLA-II binders is significantly higher than non-binders in the all types of predicted structural regions, **Fig. 1**. The majority of AA per epitope is hydrophilic in epitopes in both D an O protein regions, but HLA binding-positive epitopes of both classes are more hydrophobic than negative epitopes. HLA-class I binders are the most hydrophobic of all analyzed epitopes, **Fig. 2**.

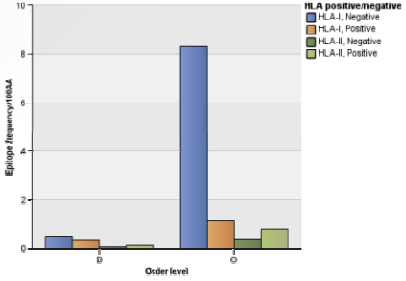

Figure 1. Frequency of exper. validated HLA class I and II negative and positive epitopes in the consensus regions

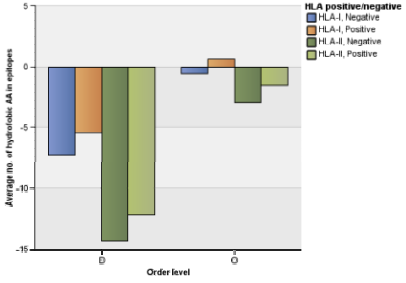

Figure 2. Average number of hydrophobic AA in HLA class I and II negative/positive epitopes in the consensus regions

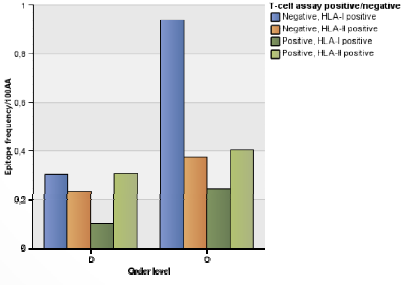

Figure 3. Frequency of exper. validated T-cell assay negative and positive epitopes in the consensus regions. All analyzed epitopes were positive for HLA class I or class II binding

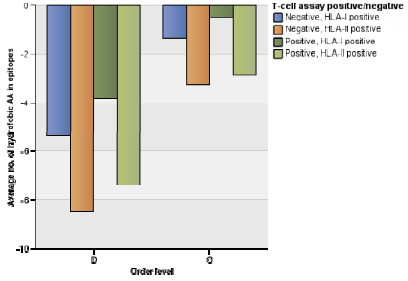

Figure 4. Average number of hydrophobic AA in T-cell assay negative/positive epitopes in the consensus regions. All analyzed epitopes were positive for HLA class I or class II binding

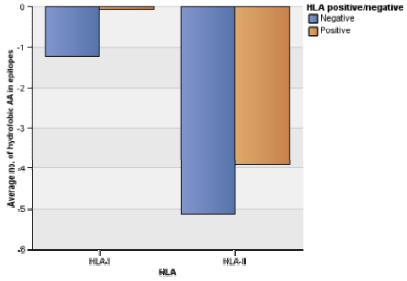

Figure 5. Average number of hydrophobic AA in epitopes in the dataset of HLA class I or class II negative/positive epitopes

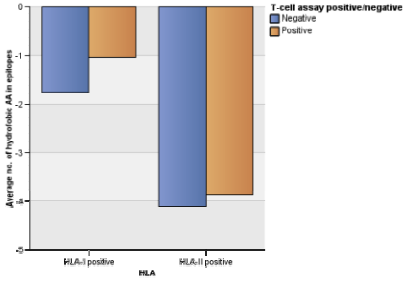

Figure 6. Average number of hydrophobic AA in epitopes in the dataset of T-cell assay negative/positive epitopes (HLA class I and class II-positive)

Regarding T cell assay negative or positive epitopes (all of them HLA binding positive), both HLA-I and HLA-II binders are preferentially concentrated in O regions and there is higher frequency of T cell assay negative than positive epitopes than in D regions. T cell assay negative/HLA-I binders are the most frequent in O regions. This result could possibly be attributed to the overestimated prediction of HLA-I binding epitopes towards structurally O regions, as noticed above. However, among HLA-II binders, the relation between T cell assay negative and positive epitopes is almost equal in O regions, and the percentage of T cell assay positive epitopes is slightly higher than those of negative epitopes in all types of structural regions**, Fig. 3**. The majority of AA in T-cell assay tested epitopes is hydrophilic in both D an O protein regions, but T-cell assay-positive epitopes of both classes are more hydrophobic than negative epitopes. T-cell assay positive/HLA-class I binding positive epitopes are the most hydrophobic of all analyzed epitopes, **Fig. 4.** They are, also, evidently, more hydrophobic than HLA I-binders from the first dataset, Fig.2

Although T cell epitopes have positional bias to O regions, which are more hydrophobic than D regions, all classes of analyzed epitopes have higher number of epitopes with majority of hydrophilic amino acids. However, HLA-I binding epitopes, which are positive in T-cell functional assays are more hydrophobic than those that are negative in T-cell functional assays, which is a significant difference between immunogenic and nonimmunogenic epitopes.

CONCLUSION

Analysis of positional biases of experimentally validated HLA-binding (both class I and class II) epitopes and T-cell functional assay positive epitopes in ordered and disordered regions of protein Ag reveals epitope concentration in the consensus of ordered protein regions. Although T cell epitopes have positional bias to ordered regions, which are more hydrophobic than disordered regions, all classes of analyzed epitopes have higher number of epitopes with majority of hydrophilic amino acids. HLA class-I binding epitopes, which are positive in T-cell functional assays are, however, much more hydrophobic than those that are negative in T-cell functional assays. This result could be useful in better prediction of immunogenic T-cell epitopes that bind to HLA class-I molecules. T-cell functional assay positive/HLA II binders are only slightly more positive than T-cell assay negative/HLA II binders.