# Relationship between protein surface and antibody binding

Virginie Lollier, Sandra Denery-Papini, Colette C. Larre, Dominique D. Tessier

# Relationship between protein surface and antibody binding

**Virginie Lollier, Sandra Denery-Papini, ColetteLarré and Dominique Tessier**

INRA UR Biopolymers, Interactions, Assemblies (Nantes, France). **Contact : vlollier@nantes.inra.fr**

Financé par ANR

## - Definitions -

**B-cell epitopes are small stretches of amino acids within antigens which interact specifically with antibodies, triggering immune reactions.**

**The B-cell epitopes are classified into 2 categories.**

**Continuous epitopes are composed of contiguous amino acids along the primary protein sequence**

**Discontinuous epitopes combine several short segments scattered along the sequence, brought together in spatial proximity when the protein is folded.**

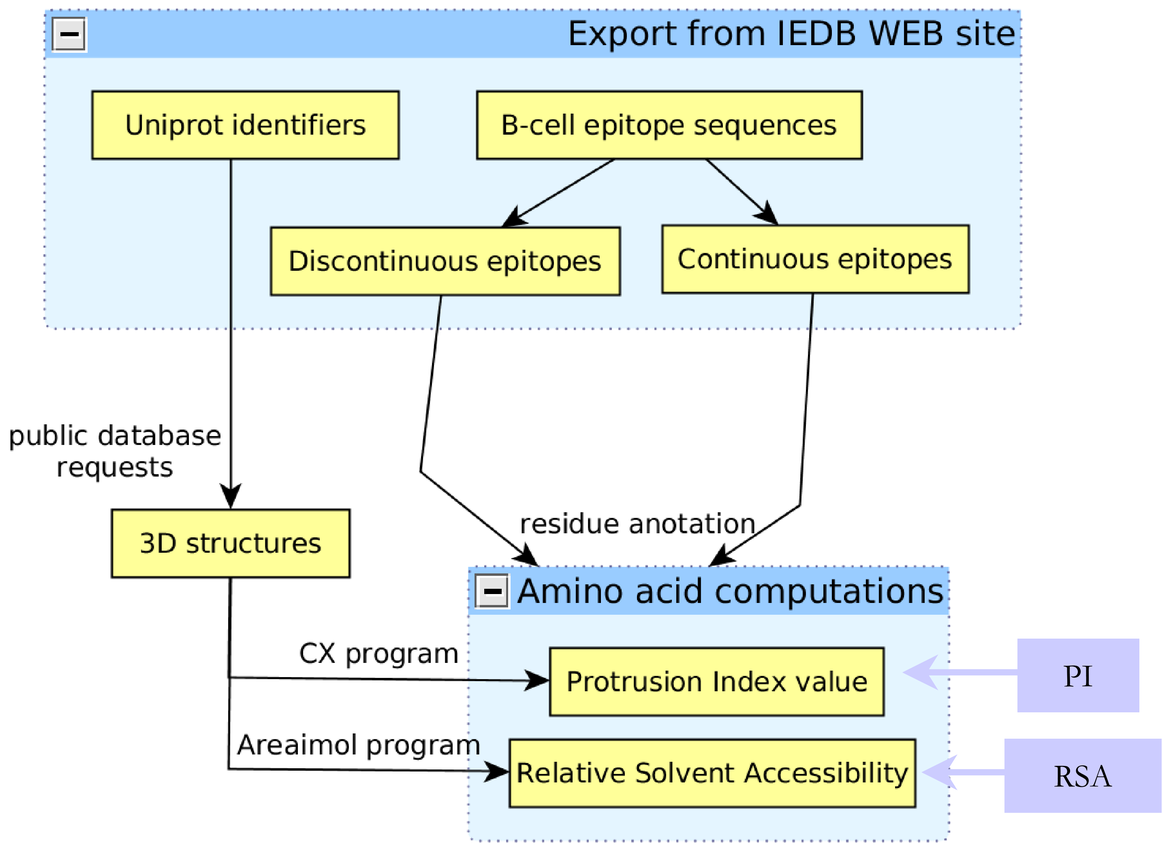
**Introduction** The assumption of finding B-cell epitopes on the molecule's surface is generally used as a criterion to locate them within the sequence of antigen. However, the current prediction systems, both from the protein 3D structures or from their sequences seem weakly efficient in view of benchmark tests. In order to assess how the surface accessibility feature is pertinent for epitope prediction, we studied how amino acids of many experimentally identified B-cell epitopes are surface exposed in comparison with the remain of the antigenic molecule.

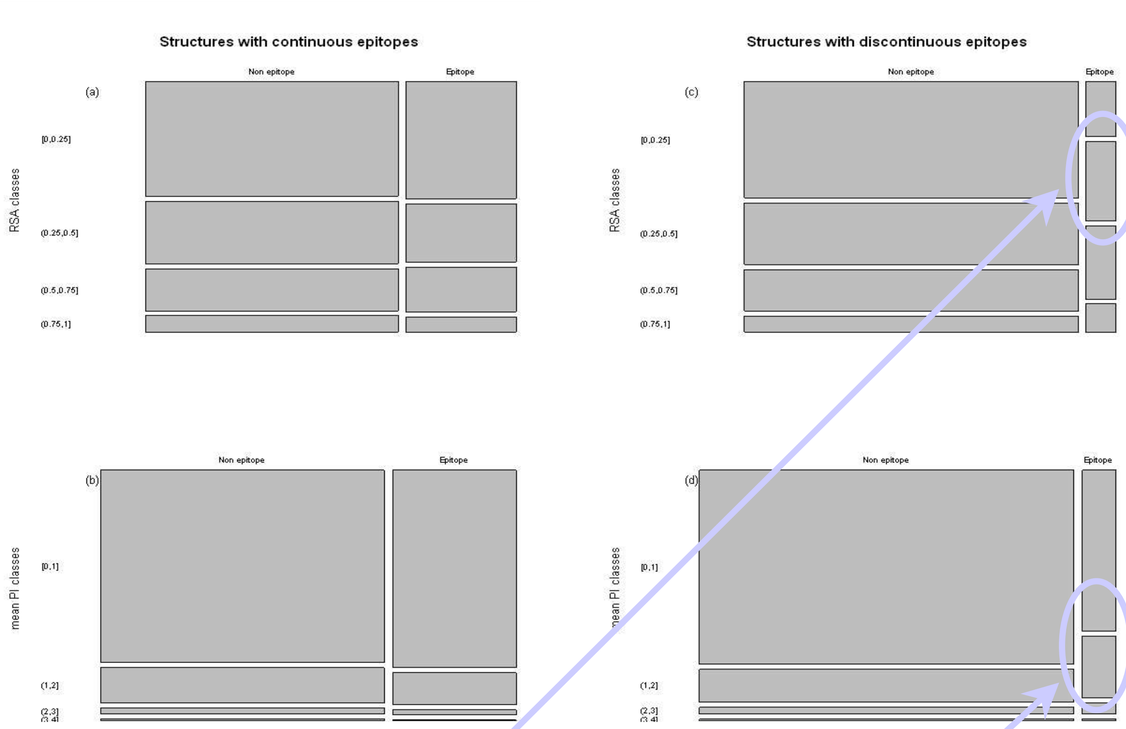
## Data collection

Epitope sequences and description is exported from Immune Epitope DataBase (IEDB). Epitopes are mapped to 61 3D structures, extracted from PDB site. The dataset is splitted into 2 subsets according to epitope types

A pipeline program computes relative surface exposure of amino acids for each PDB file.

Data are loaded into R environnement for statistical comparison between epitopic and no epitopic amino acids from antigens.



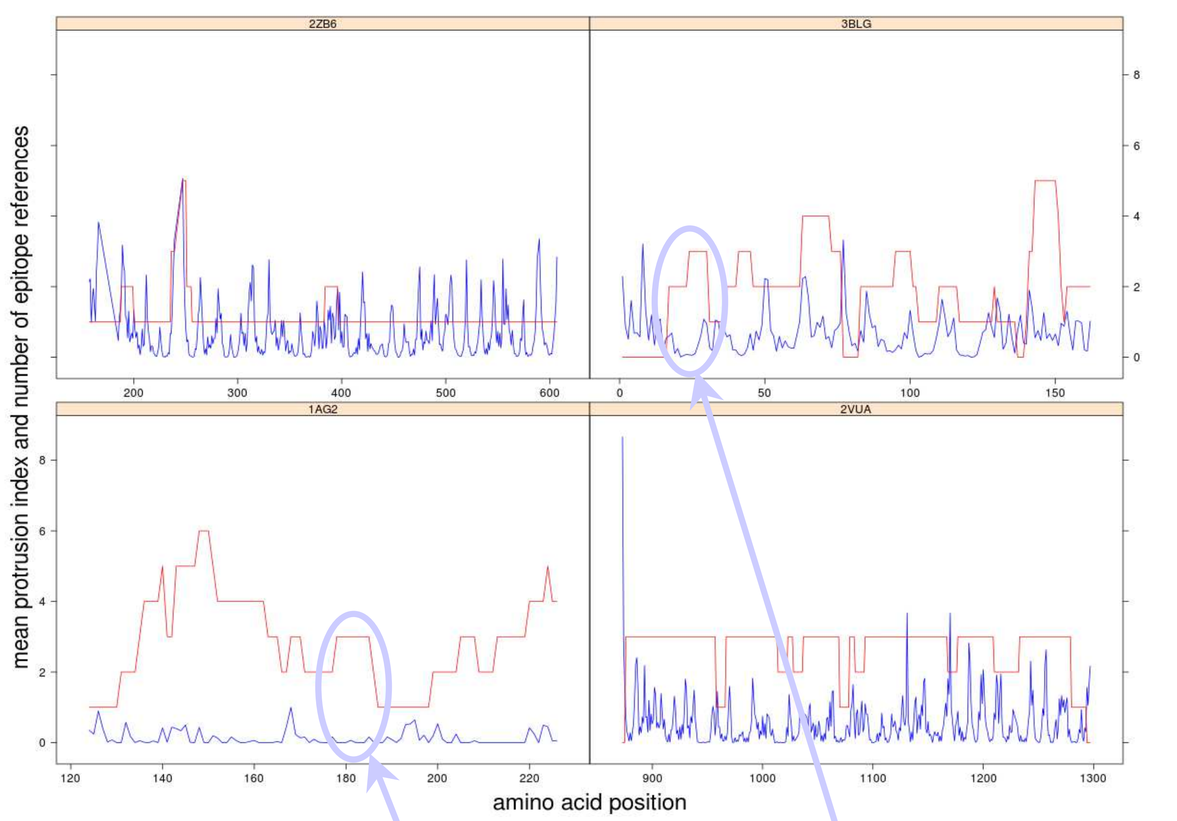

Different distributions but epitopic residues are not mostly surface-exposed

## Overall distributions

Regardless of the type of antigen (allergens, toxins, viral and autoimmune proteins) no threshold value could separate epitopic amino acids from the rest of the molecule.
Considering the continuous epitopes, the distribution of RSA and PI values for epitopic residues is identical to non epitopic.
Considering the discontinuous type, a statistical difference is observed, but, epitopic elements are not spread over higher classes of values.

## Individual analysis

When the epitopic values are mapped onto protein sequences, some protein sequences are entirely covered by continuous epitopes.
The relationship between the number of bibliographic references and the overlapping of the identified epitopes is examined on 4 structures.




Most often identified residues are not linked to high protrusion indexes

**Conclusion** Our work did not deny the requesite for B-cell epitopes to be exposed on surface. However, it indicates that the level of surface exposure has no discriminatory ability. As such, it would be difficult to include it as a parameter for any prediction system. Besides, the presence of epitopes all along several antigenic sequences appears more disconcerting. Apparently the epitope identification method depends on a well-defined context. Beyond the distinction of epitope types, it could be helpful to take this well defined context into account when compiling datasets through the use of a dedicated terminology. Such measures would enlarge the existing immune epitope ontology, which would be important before applying any generic approach to structural description.