A novel Arg H52 and Tyr H33 conservative binding motif in antibodies: a correlation between sequence of immunoglobulins and their binding properties

Artem Petrov, Vladimir Arzhanik

Kolmogorov Advanced Education Scientific Center of Lomonosov Moscow State University, Kremenchugskaya st., 11, Moscow, Russia, 121357, artempetrov@mail.ru

Lomonosov Moscow State University, Faculty of Bioengineering and Bioinformatics, LenGory, 1, b.73, Moscow, Russia, 119992

Oleg V. Koliasnikov

Kolmogorov Advanced Education Scientific Center of Lomonosov Moscow State University, Kremenchugskaya st., 11, Moscow, Russia, 121357, olkol@aesc.msu.ru

Introduction

Antibodies are the family of glycoproteins, which are responsible for antigen recognition and binding in vertebrates. In order to predict antibody-antigen interaction, several methods are used. When the crystallographic structure of a complex is unavailable, computational tools, such as homology modeling tools, are used for structure prediction. However, in the absence of additional biological information about the combining sites of immunoglobulins and without reliable methods of quality-control of modeled structures, it is challenging to model a correct high quality antibody-antigen complex basing on homological structures.

One of the ways to avoid errors tied to the use of homological structures in modeling of antibody-antigen interaction is to predict the exact interaction scheme without use of structural information, directly by the amino acid sequence of antibody. This method will help to increase the accuracy of computational modeling of the antibody-antigen complexes and to control the correctness of the modeled complexes. Our research presents the novel conservative binding motif and therefore the direct correlation between the amino acid sequence of antibodies and their specific binding properties.

The purpose of the investigation

The purpose of our research was to find a correlation between the sequence of antibodies and their specific interaction mechanism with negatively charged groups of antigens. This

criterion was chosen because the majority of antigens - proteins, peptides, many carbohydrates and haptens - contain negatively charged groups.

Methods of the investigation

We used multiple sequence alignment tools and Jalview software to find conservative motifs. Then we applied PyMOL software to analyze the structures and establish whether or not the residues of the motif are involved in interaction with negatively charged groups of antigens. We used molecular dynamics (GROMACS software, Lomonosov Supercomputer Center) and distance data analysis to establish interaction mechanism. Then we showed the conservatism of the motif using distance data and germline genes analysis.

Results of the research

Using this method we have found an amino acid motif containing four residues: arginine in the position H52, tyrosine in the position H33, threonine in the position H59 and glutamic acid in the position H61 (Arg H52, Tyr H33, Thr H59 and Glu H61). This motif is situated in heavy chain of antibodies, and three residues (Arg H52, Thr H59, Glu H61) are situated in one structural unit – CDR H2. This means that the residues of the motif have specific spatial orientation. The results of structural alignment are presented in the Figure 1.

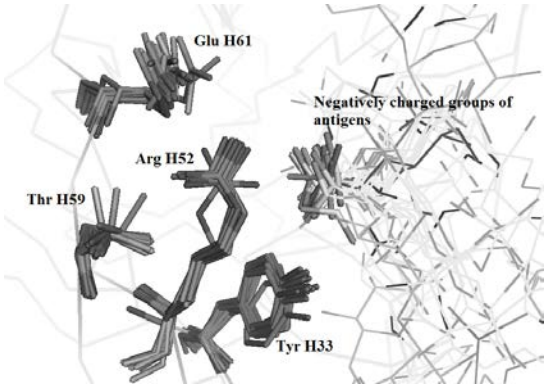

Figure 1 presents the structural alignment of the residues of the motif in complex with various antigens, which have negatively charged groups. It shows that the residues and antigens have specifically determined spatial orientation in the binding site.

These residues interact only with negatively charged groups of antigens according to a specific mechanism, which we have also established. Two residues are involved in antigen

binding: Arg H52 and Tyr H33. Arg H52 positively charged side chain forms salt bridges with negatively charged groups of antigens, forming a very energetically favorable complex, therefore contributing to the total antigen-antibody interaction energy. Tyr H33 forms a hydrogen bond with negatively charged group of antigen, coordinating the antigen spatially correctly towards Arg H52 side chain. Other two residues of the motif, Glu H61 and Thr H59, coordinate highly movable and charged side chain of Arg H52 residue by forming salt bridges and a hydrogen bond, respectively (Fig. 2).

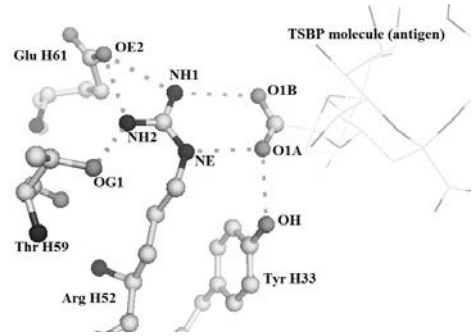

Figure 2 demonstrates the interaction of the residues in the conservative motif with antigen. All bonds are shown with dashed lines.

Then we have established the geometrical characteristics (distance length values) of the binding complex (Fig. 3a, 3b).

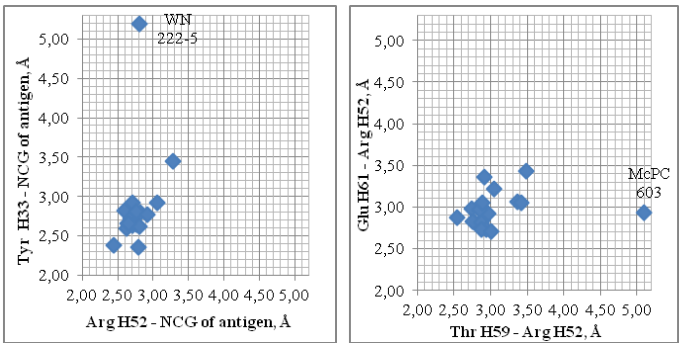

Figure 3a. Distribution of distances between interacting residues of the motif and negatively charged groups (NCG) of antigens. Figure 3b. Distribution of distances between the coordinating residues and Arg H52 side chain.

The plots show that the complex is very geometrically conservative. We have also established the statistical data regarding antibodies with the conservative motif: they comprise 4% of all

antibodies present in the X-Ray structures database (SAbDab, http://opig.stats.ox.ac.uk/webapps/sabdab). Germline genes analysis shows the genetic conservation of the motif: 4 out of 5 genes, coding CDR H2, which has 19 residues length, code CDR H2 with the conservative motif.

Conclusions

In this research we have found the conservative motif, which has specific antigen binding properties. The main result of this work is the finding of the new possibility to establish directly by the sequence, without structural analysis, how antibodies, which have the motif in their heavy chain sequences, interact with negatively charged groups of antigens. The results of this research can be used to facilitate the antibody-antigen interaction modeling (in case of antibodies containing the motif there is no more need to use homology modeling to predict the interaction of the antibodies with negatively charged groups of antigens), to increase the accuracy of computational antibody-antigen interaction modeling and for post-modeling quality control of the modeled structures (the determined geometrical parameters of the binding complex (Fig. 3a, 3b) can be applied in modeling software to build a correct model of antibody – antigen complex without use of homological structures or to control the accuracy of already modeled structures, therefore increasing the accuracy of the modeling). This work also demonstrates the possibility of studying sequence-dependent antibody binding properties.

1. C. Chothia, A.M. Lesk (1987) Canonical structures for the hypervariable regions of immunoglobulins, *J. Mol. Biol.,* **196(4):**901–917.

2. C.L. Brooks et al. (2008) Exploration of Specificity in Germline Monoclonal Antibody Recognition of a Range of Natural and Synthetic Epitopes, *J. Mol. Biol.,* **377:**450–468.

3. R.M. MacCallum et al. (1996) Antibody-antigen Interactions: Contact Analysis and Binding Site Topography, *J. Mol. Biol.,* **262:**732–745.

4. E.J. Sundberg (2009) Structural Basis of Antibody–Antigen Interactions, *Methods in Molecular Biology,* **524:**23–36.

5. D.J. Barlow, J.M. Thornton (1983) Ion-pairs in Proteins, *J. Mol. Biol.*, **168:**867–885.