Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Epitope-Analyzer: A structure-based webtool to analyze broadly neutralizing epitopes

Daniel Montiel-Garcia, Oscar Rojas-Labra, Nelly Santoyo-Rivera, Vijay S. Reddy

PII: S1047-8477(22)00009-0
DOI: https://doi.org/10.1016/j.jsb.2022.107839
Reference: YJSBI 107839

To appear in: Journal of Structural Biology

Received Date: 18 September 2021
Revised Date: 27 January 2022
Accepted Date: 1 February 2022

Please cite this article as: Montiel-Garcia, D., Rojas-Labra, O., Santoyo-Rivera, N., Reddy, V.S., Epitope-Analyzer: A structure-based webtool to analyze broadly neutralizing epitopes, Journal of Structural Biology (2022), doi: https://doi.org/10.1016/j.jsb.2022.107839

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier Inc.
Epitope-Analyzer: A structure-based webtool to analyze broadly neutralizing epitopes

Daniel Montiel-García¹*, Oscar Rojas-Labra², Nelly Santoyo-Rivera² and Vijay S. Reddy¹*
¹Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.
²Departments of Computer Systems and Information Technologies, Tecnologico Nacional de Mexico & Instituto Tecnológico Superior de Irapuato, Irapuato, Guanajuato, Mexico.

*Corresponding authors:
Vijay S. Reddy, e-mail: reddyv@scripps.edu
Daniel Montiel-Garcia, e-mail: dmontiel@scripps.edu
Abstract

The antigenic epitope regions of pathogens (e.g., viruses) are recognized by antibodies (Abs) and subsequently cleared by the host immune system, thereby protecting us from disease. Some of these epitopes are conserved among different variants or subgroups of pathogens (e.g., Influenza (FLU) viruses, Coronaviruses), hence can be targeted for potential broad-neutralization. Here we report a web-based tool, Epitope Analyzer (EA), that rapidly identifies conformational epitope and paratope residues in an antigen-antibody complex structure. Furthermore, the tool provides the ways and means to analyze broadly neutralizing epitopes by comparing the equivalent epitope residues in similar antigen structures. The similarity in the epitope residues between (multiple) pairs of similar antigen molecules suggest the presence of conserved epitopes that can be targeted by broadly neutralizing antibodies. These details can be used as a guide in developing effective treatments, such as the design of novel vaccines and formulation of cocktail of broadly neutralizing antibodies, against multiple variants or subgroups of viruses. The web application can be freely accessed from the URL, http://viperdb.scripps.edu/ea.php.

Keywords: Epitopes, Paratopes, Epitope conservation, broadly neutralizing antibodies, Vaccine design

1. Introduction

Human and animal viruses undergo continuous evolution to counter host immune pressures particularly directed against the surface exposed regions of viruses (Smith et al., 2004). While these regions belong to the surface glycoproteins (e.g., Envelope (Env), Hemagglutinin (HA) and Spike glycoproteins) in the case of enveloped viruses (e.g., HIV, FLU and Coronaviruses, respectively), they correspond to viral coats in the case of non-enveloped or naked viruses (e.g., Picornaviruses). Interestingly, some of these regions involve receptor binding domains as well as those important for other
aspects of viral function (e.g., fusion machinery of enveloped viruses). Remarkably, however, these functionally important regions evade immune surveillance and the critical residues responsible for the specific viral function are often preserved. Furthermore, these functionally important regions are conserved in terms of their sequence and local structure across different subtypes of viruses (e.g., FLU type A). Such regions are of particular interest for developing the broad-spectrum antivirals and structure-based vaccine design (Corti et al., 2011; Dreyfus et al., 2012; Laursen and Wilson, 2013; Whittle et al., 2011; Wu and Wilson, 2020; Yoshida et al., 2009). The antibodies that target the conserved regions across different subtypes of viruses are known as broadly neutralizing antibodies (bNAbs) (Barba-Spaeth et al., 2016; Cockburn et al., 2012; Laursen and Wilson, 2013; Wu and Wilson, 2020). Acquisition of “passive immunity” through the infusion of bNAbs effectively protects the recipients against the corresponding viral infections (Morris and Mkhize, 2017).

To identify and analyze conserved epitopes across a group of similar antigens, which can be targeted by bNAbs, we developed a web-based tool, namely Epitope Analyzer (EA) (http://viperdb.scripps.edu/ea.php). This application locates the conformational (non-linear) epitope residues from a chosen Antigen (Ag) and Antibody (Ab) complex structure from PDB or a user uploaded pdb/cif file and compares them with other user specified antigen structures.

2. Results

Firstly, for any selected or uploaded Ag-Ab complex structure (PDB-1), the EA-utility identifies and displays the epitope and paratope residues (Fig. 1). Users can select a preferred distance cutoff (default: 4Å) to identify epitope and paratope residues from the interacting antigen and antibody molecules, respectively. This information on interactions along with the details of molecules involved (e.g., PDB-ID, antigen, and antibody) are stored in a behind-the-scenes MySQL database for reuse when
the same structure (PDB-ID) is chosen again. Subsequently, the conserved conformational epitope residues are identified by superimposing the $2^{nd}$ chosen or uploaded antigen structure (in PDB-2) onto the $1^{st}$ Ag molecule (in PDB-1) using Kpax program (Ritchie, 2016; Ritchie et al., 2012) and comparing the equivalent epitope residues in terms of their sequence and structural conservation. The equivalent epitope residues in the $2^{nd}$ antigen structure (PDB-2) are identified by rigid docking (translocation) of the antibody molecule in the reference structure (PDB-1) onto the $2^{nd}$ structure after superposition of the two antigen molecules. It is not required that PDB-2 contain coordinates for an antibody molecule. If PDB-2 does contain coordinates for an antibody molecule, they will not be considered by the EA-utility. If the user desires the antibody coordinates in PDB-2 to be considered, a separate run must be performed with this coordinate file designated as PDB-1. It is important to note that the reference structure (PDB-1) should have coordinates for an antigen and antibody, while the second structure (PDB-2) only needs to have coordinates for an antigen.

Additionally, while the structural similarities are calculated as RMSDs between the aligned C-alpha atoms, the sequence similarities (not identities) are estimated based on the similar properties of the structurally aligned a.a. residues. Of note, the 20 a.a. residues are categorized into 7 groups based on the similarity in their properties: 1) small hydrophobics (A, G, P, V); 2) large hydrophobics (I, L, M); 3) positively (+) charged (H, K, R); 4) negatively (-) charged (D, E); 5) small polar (C, S, T); 6) large polar (N, Q) and 7) aromatics (F, Y, W). The aligned a.a. residues that belong to the same group are considered equivalent (similar). It is notable that one can obtain two sets of RMSDs: 1) using all the aligned C-alpha atoms in the two superimposed antigen molecules from the [Aligned residue pairs] tab in the EA-tool and 2) using only the aligned epitope residues from the [Aligned epitope residues] tab. The RMSDs of the aligned epitope residues are listed in Table 1 for the case study described below.
Despite the differences in the surface exposed loops, the overall structures of similar antigens are highly conserved (e.g., HAs of FLU, Envs of HIV and protomers of Picornaviridae). This allows one to obtain proper structure-based sequence alignments by superimposing the overall antigen structures from the same virus family, even when the overall sequence identity between different antigens is small (<20%) (Montiel-Garcia et al., 2016). One of the known limitations of the EA-tool is that currently there is no direct way to consider the molecular symmetry defined in the BIOMT matrices within the EA-tool. A workaround is to generate the full molecule outside of the tool and upload the expanded pdb file to analyze the interactions.

In a case study, we calculated the similarity between the epitope residues, which occur in the stem region of HA from multiple strains of FLU viruses that bind to different (known) bNAbs, CR6261, CR8020, and CR9114 (Laursen and Wilson, 2013; Wu and Wilson, 2020). Table 1 shows the sequence and structural similarities between the epitope regions of selected pairs of HA-Ab complexes. Of note, the upper half and lower half of the comparisons in Table 1 correspond to same pair of antigens. However, the statistics are not identical for the equivalent pairs (e.g., 3gbm:5c0s vs. 5c0s:3gbm) because the epitope residues being compared will be different depending on the chosen reference antigen molecule (PDB-1). Four of the Ag-Ab coordinate sets in the case study involve bNAb CR6261 in complex with HA of various group-1 FLU viruses. It is notable that when these four coordinate sets are compared with each other, the identified epitope RMSDs (0.38 to 1.05 Å) are relatively small, and the epitope sequence similarities that range from 63.2% to 95.0% are relatively high. Thus, if it were not already known that a bNAb exists for this region of HA, it might be concluded from an EA analysis with one Ag-Ab coordinate set and multiple antigen-only coordinate sets that this region would be a good target for a bNAb.

However, not surprisingly the epitope RMSDs based on group-2 specific bNAb:CR8020 are relatively higher (1.48 to 2.81 Å), and the epitope sequence similarities (48.0% to 67.0%) are lower.
between group-1 and group-2 FLU viruses. Interestingly, the group-1 epitope residues associated with CR9114 antibody, which supposedly neutralizes both group-1 and group-2 viruses (Laursen and Wilson, 2013), correlate well with those of group-1 viruses but poorly with the one group-2 virus in the case study. As shown in Table 1, when the HA - CR9114 coordinates (PDB-ID: 4fqi) are designated as PDB-1, the other group-1 HA antigens (PDB-IDs: 3gbm, 3gbn, 5c0s, 6uyn) are found to have relatively low epitope RMSDs (0.42 to 1.09 Å) and high epitope sequence similarities (58.0% to 100.0%). These EA results are consistent with our knowledge that the CR9114 antibody targets a similar epitope as targeted by the bNAb CR6261. In contrast, the same reference structure (PDB-ID: 4fqi) compared to one group-2 HA antigen (PDB-ID: 3sdy) results in a higher epitope RMSD (2.54 Å) and lower epitope sequence similarity (37.0%). These results suggest that while CR9114 may neutralize some group-2 viruses, perhaps it does not neutralize HK68/H3. As expected, when both the PDB-1 and PDB-2 antigens are from the same FLU virus (e.g., H5N1), the EA-utility finds a low epitope RMSD (0.41-0.42 Å) and 100% sequence similarity.

3. Conclusions and outlook

The EA-tool is a web-based application (http://viperdb.scripps.edu/ea.php) that aims to assist researchers to analyze the conserved epitope residues targeted by bNAbs. Even though it was originally designed to analyze the Ag-Ab interactions, the EA-tool can also be used to evaluate protein-protein interactions between any two set(s) of chains in a pdb file. Moreover, user can upload and analyze the coordinates that are not yet deposited in PDB. Going forward, we plan to analyze all the available antigen-antibody complex structures in PDB and store this information in a relational database so that researchers can search for the antigens or antibodies of interest and associated broadly neutralizing epitopes.

Acknowledgements
We thank Dr. Robyn Stanfield for the helpful comments on the manuscript. We also would like to thank Dalton Karlinsey for testing and providing the feedback on the EA-tool.

**Funding**

This work was supported by the NIH grant R21 AI137580 to VSR.

**Author contributions:**

D.M.G and V.S.R. designed the research. D.M.G, O.R.L and N.S.R. developed and implemented the tool. Everyone was involved in testing the tool. D.M.G. and V.S.R. wrote the manuscript. All authors were asked to comment on the manuscript.

**Figure Legends:**

**Figure 1.** Flowchart description of Epitope Analyzer (EA) tool (http://viperdb.scripps.edu/ea.php). There are 4 main components in the EA tool that overlap with each other: 1) Ag-Ab interaction analysis, 2) analysis of epitope conservation, 3) network analysis and 4) graphical user interface (GUI), identified blue, orange, green and pink colored areas. PDB-1 is the reference structure file of an antigen-antibody complex that can either be downloaded directly from PDB by specifying PDB-ID (e.g., 3gbn) or uploaded by the user from his/her computer. The molecular entities of PDB-1 and their sequences are displayed in the Web-GUI window. The [interactions] module calculates the non-covalent interactions between the selected chains using a chosen distance cutoff. The identified epitope and paratope residues are listed in one of the side panels and displayed on demand in the Web-GUI window. The antigen structure to be compared (PDB-2) is uploaded in [Structural comparison] panel. The results of the structure comparison and similarity in epitope residues are displayed in one of side the panels. The [Interaction Network] panel calculates and displays the network of interactions between the selected chains in the reference structure.
(PDB-1). Finally, the lists of epitope residues and sequence alignments can be downloaded in “CSV” and “FASTA” formats, respectively from the hyperlinks provided. JSMol is an open-source HTML-5 viewer for 3D molecular structures (http://wiki.jmol.org/index.php/JSmol).

References:

Barba-Spaeth, G., Dejnirattisai, W., Rouvinski, A., Vaney, M.C., Medits, I., Sharma, A., Simon-Loriere, E., Sakuntabhai, A., Cao-Lormeau, V.M., Haouz, A., England, P., Stiasny, K., Mongkolsapaya, J., Heinz, F.X., Screaton, G.R., Rey, F.A., 2016. Structural basis of potent Zika-dengue virus antibody cross-neutralization. Nature 536, 48-53.

Cockburn, J.J., Navarro Sanchez, M.E., Fretes, N., Urvoas, A., Staropoli, I., Kikuti, C.M., Coffey, L.L., Arenzana Seisdedos, F., Bedouelle, H., Rey, F.A., 2012. Mechanism of dengue virus broad cross-neutralization by a monoclonal antibody. Structure 20, 303-314.

Corti, D., Voss, J., Gamblin, S.J., Codoni, G., Macagno, A., Jarrossay, D., Vachieri, S.G., Pinna, D., Minola, A., Vanzetta, F., Silacci, C., Fernandez-Rodriguez, B.M., Agatic, G., Bianchi, S., Giacchetto-Sasselli, I., Calder, L., Sallusto, F., Collins, P., Haire, L.F., Temperton, N., Langedijk, J.P., Skehel, J.J., Lanzavecchia, A., 2011. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science 333, 850-856.

Dreyfus, C., Laursen, N.S., Kwaks, T., Zuidgeest, D., Khayat, R., Ekiert, D.C., Lee, J.H., Metlagel, Z., Bujny, M.V., Jongeneelen, M., van der Vlugt, R., Lamrani, M., Korse, H.J., Geelen, E., Sahin, O., Sieuwerts, M., Brakenhoff, J.P., Vogels, R., Li, O.T., Poon, L.L., Peiris, M., Koudstaal, W., Ward, A.B., Wilson, I.A., Goudsmit, J., Friesen, R.H., 2012. Highly conserved protective epitopes on influenza B viruses. Science 337, 1343-1348.

Laursen, N.S., Wilson, I.A., 2013. Broadly neutralizing antibodies against influenza viruses. Antiviral Res 98, 476-483.

Montiel-Garcia, D.J., Mannige, R.V., Reddy, V.S., Carrillo-Tripp, M., 2016. Structure based sequence analysis of viral and cellular protein assemblies. J Struct Biol.

Morris, L., Mkhize, N.N., 2017. Prospects for passive immunity to prevent HIV infection. PLoS Med 14, e1002436.

Ritchie, D.W., 2016. Calculating and scoring high quality multiple flexible protein structure alignments. Bioinformatics 32, 2650-2658.

Ritchie, D.W., Ghoorah, A.W., Mavridis, L., Venkatraman, V., 2012. Fast protein structure alignment using Gaussian overlap scoring of backbone peptide fragment similarity. Bioinformatics 28, 3274-3281.

Smith, D.J., Lapedes, A.S., de Jong, J.C., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D., Fouchier, R.A., 2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305, 371-376.

Whittle, J.R., Zhang, R., Khurana, S., King, L.R., Manischewitz, J., Golding, H., Dormitzer, P.R., Haynes, B.F., Walter, E.B., Moody, M.A., Kepler, T.B., Liao, H.X., Harrison, S.C., 2011. Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. Proc Natl Acad Sci U S A 108, 14216-14221.
Wu, N.C., Wilson, I.A., 2020. Influenza Hemagglutinin Structures and Antibody Recognition. Cold Spring Harb Perspect Med 10.

Yoshida, R., Igarashi, M., Ozaki, H., Kishida, N., Tomabechi, D., Kida, H., Ito, K., Takada, A., 2009. Cross-protective potential of a novel monoclonal antibody directed against antigenic site B of the hemagglutinin of influenza A viruses. PLoS Pathog 5, e1000350.
Fig. 1. Epitope Analyzer flowchart
Table 1. RMSDs and sequence similarities between the equivalent epitope residues in HA and bNAb complexes.

| Antigen-Antibody complex* | 3gbm† CR6261§ | 3gbn† CR6261§ | 5c0s† CR6261§ | 6uyn† CR6261§ | 4fqi† CR9114§ | 3sdy# CR8020§ |
|---------------------------|----------------|----------------|---------------|---------------|---------------|---------------|
| †HA (H5N1) - CR6261; 3gbm| 100 (0.0)‡    | 83.4 (0.93)    | 79.0 (0.97)   | 85.0 (0.78)   | 100.0 (0.42)  | 62.0 (2.44)   |
| †HA (H1N1) - CR6261; 3gbn| 84.2 (0.98)‡  | 100 (0.0)      | 94.7 (1.04)   | 95.0 (0.38)   | 74.0 (1.09)   | 62.0 (2.70)   |
| †HA (Gen4-HA-SS) – CR6261; 5c0s| 63.2 (0.92)    | 83.4 (1.05)    | 100 (0.0)    | 75.0 (0.92)   | 58.0 (0.75)   | 62.0 (1.48)   |
| †HA (A/Ohio/09/2015) – CR6261; 6uyn| 84.2 (0.81)      | 94.4 (0.38)    | 89.5 (1.04)   | 100 (0.0)    | 74.0 (0.91)   | 67.0 (2.61)   |
| †HA (H5N1) - CR9114; 4fqi| 100.0 (0.41)     | 83.3 (1.01)    | 79.0 (1.05)   | 85.0 (0.91)   | 100 (0.0)      | 48.0 (2.81)   |
| †HA (HK68/H3) – CR8020; 3sdy| 63.2 (2.12)     | 55.6 (2.51)    | 63.0 (1.06)   | 65.0 (2.41)   | 37.0 (2.54)    | 100 (0.0)     |

†Name of the antibody in the complex.
‡Group-1 FLU viruses, *Group-2 FLU virus
§Estimated sequence similarity and ‡RMSD (root mean square deviation) in Å between the equivalent pairs of epitope residues.

*The sequence similarities and RMSDs listed in the table were calculated by considering the structures (PDBs) indicated in the 1st row as PDB-1 and those in the 1st column as PDB-2 (see Fig. 1 for details).
Author contributions:

D.M.G and V.S.R. designed the research. D.M.G, O.R.L and N.S.R. developed and implemented the tool. Everyone was involved in testing the tool. D.M.G. and V.S.R. wrote the manuscript. All authors were asked to comment on the manuscript.

Conflict of Interest:

Authors declare that they have no conflicts of interest