Published in final edited form as:

Curr Opin Virol. 2015 April ; 11: 98–102. doi:10.1016/j.coviro.2015.03.012.

Antibody specific epitope prediction – emergence of a new paradigm

Inbal Sela-Culang¹, Yanay Ofra¹, and Bjoern Peters²

¹The Goodman Faculty of Life Sciences, Nanotechnology Building, Bar-Ilan University, Ramat Gan, 52900, Israel

²La Jolla Institute for Allergy and Immunology, Division of Vaccine Discovery, 9420 Athena Circle, La Jolla, CA 92037, USA

Abstract

The development of accurate tools for predicting B-cell epitopes is important but difficult. Traditional methods have examined which regions in an antigen are likely binding sites of an antibody. However, it is becoming increasingly clear that most antigen surface residues will be able to bind one or more of the myriad of possible antibodies. In recent years, new approaches have emerged for predicting an epitope for a specific antibody, utilizing information encoded in antibody sequence or structure. Applying such antibody-specific predictions to groups of antibodies in combination with easily obtainable experimental data improves the performance of epitope predictions. We expect that further advances of such tools will be possible with the integration of immunoglobulin repertoire sequencing data.

Introduction

The epitope of an antibody can be defined as the minimal structural determinant that it recognizes. Correct identification of an antibody's epitope is crucial for understanding the molecular basis of immunity and autoimmunity. It may also allow for the design of immunogens that elicit similar antibodies in a vaccine or therapeutic setting. Moreover, characterizing the epitope of an antibody helps understand and predict possible cross-reactivity, which is particularly important when the antibody is used as a drug, as a diagnostic tool or as a reagent. Multiple experimental methods have been successfully applied to the identification of antibody epitopes such as X-ray crystallography, NMR spectroscopy, peptide ELISAs, phage display, expressed fragments, partial proteolysis, mass spectrometry, and mutagenesis analysis. However, such experimental methods can be expensive, time consuming and no single method will consistently succeed in identifying epitopes for all antibodies [1]. Moreover, the rapid and inexpensive methods, such as peptide ELISA, typically identify linear epitopes, rather than conformational ones although

© 2015 Published by Elsevier B.V.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. 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.

the latter are assumed to constitute about 90% of all epitopes [2,3]. Therefore, computational methods are a desirable alternative to identify antibody epitopes [4].

Traditional B-cell epitope prediction

The first epitope prediction methods were published in the 1980s and were fairly simple. They were based on a single propensity scale such as flexibility, amino-acid composition or solvent accessibility [5–10]. A new generation of methods that combined multiple physicochemical properties was introduced in the 1990s [11–13]. However, the predictive quality of these approaches was questioned in 2005 in a study by Blythe and Flower [14] which showed that almost 500 propensity scales performed only slightly better than random. Since then, the field has moved away from simple propensity scales towards the development of more sophisticated knowledge-based methods [15]. Those with the better performance are usually structure-based [15], relying on antigen structure to identify patches on the surface of the antigen as putative epitopes. Whether sequence- or structure-based, all these traditional tools predict which residues in an antigen could be recognized by *some* antibody. We refer to these methods as traditional- or antibody-independent predictors in the following.

The performance of antibody-independent predictors has incrementally increased over the years, but their practical usefulness is limited [16–18]. Several reviews of such tools and studies evaluating their performance are available [1,15,18–23]. What could be the reasons for this difficulty in differentiating between epitopic and non-epitopic residues of an antigen? As more epitopes are discovered, it is becoming apparent that essentially any surface accessible region of an antigen can be the target of some antibody [16,24–28]. This phenomenon may explain the fact that epitopic and other surface residues are almost indistinguishable in their amino-acid composition, as was shown recently by several studies [29–31]. Figure 1 exemplifies this phenomenon using the hemagglutinin antigen of the Influenza virus. In this example, a specific antibody (purple ribbon representation) binds to its epitope (orange space-fill representation), but multiple other epitopes exist (cyan space-fill representation). A traditional antibody epitope prediction method would be considered correct if it identified all epitope residues, which here cover a large part of the hemagglutinin surface, and therefore would provide information that is not very useful.

Antibody-specific B-cell epitope prediction

Here we focus on a new approach to B cell epitope prediction that is based on reformulating the question being asked. Rather than attempting to predict which residues on an antigen can be recognized by some antibody, this approach attempts to predict where on the antigen a specific antibody will bind. Such predictions would be very valuable for monoclonal antibodies (mAbs) that are intended to be used as reagents, therapeutics or diagnostics. In all these applications, knowing the epitope is crucial for understanding possible cross-reactivity. Also, understanding how a specific antibody (and variants thereof) will recognize epitopes (and epitope variants) can serve as an input to optimize antibodies e.g. to ensure that they do or do not bind certain antigen-isoforms. Notably, such analyses are not possible with antibody-independent predictions.

To our knowledge, the first epitope prediction method taking into account antibody structure was suggested in 2007 by Rapberger et al [32]. Appreciating that the antigen epitope should geometrically and electrostatically match the antibody structure, they have generated a virtual library of 11,951 antibody models (based on computational mutations of known antibody structures). The epitope of a given antigen is then predicted by screening the antigen structure against this virtual library, trying to identify the best matching theoretical antibody model, based on shape complementarity and contact energies of the interacting residues. However, this method is not antibody-specific, in the sense that the antibody information used is not of the specific antibody of interest, but of some theoretical antibody. Therefore, similar to traditional epitope prediction methods, a single epitope would be predicted for a given antigen, instead of different epitopes for different antibodies, as in antibody-specific epitope prediction methods.

Truly antibody specific epitope predictions were first attempted using computational docking experiments of antibody and antigen structures. Until recently, such antibody-antigen docking used to be done with general protein-protein docking algorithms, filtering out solutions in which the CDRs are not in interaction with the antigen [33,34]. However, the application of these generic algorithms to antibody-antigen predictions has been disappointing [18,35]. Recently, a few docking approaches specific for antibody-antigen interactions have been published: SnugDock [36] is a docking algorithm that is based on RosettaDock [37], which simultaneously structurally optimizes the antibody-antigen rigid-body positions, the relative orientation of the antibody light and heavy chains, and the CDRs, to allow for antibody flexibility upon antigen binding. While the amino-acid composition of epitopic and non-epitopic surface residues is essentially the same [29–31], it appears that the set of contact preferences between amino-acids in the epitope and paratope (i.e. the preference of a certain aminoacid in the epitope to bind certain amino-acids in the paratope) is unique for antibody-antigen interfaces, and specific for each of the six CDRs [29,38]. Such an asymmetric set of contact preferences was implemented recently into an antibody-specific mode of ClusPro [39], and into EpiPred, a novel antibody-specific epitope prediction method [40], which scores candidate epitope patches by a combination of geometric fitting and antibody-antigen contact preferences. This algorithm was successfully used in re-scoring docking solution of other docking algorithms.

Docking based methods provide not only the location of the epitope, but also the specific interactions between epitope and paratope residues, which may assist in antibody design. On the other hand, a drawback of this approach and of other methods that require the antibody structure is that the structure of the antibody is many times not available. Moreover, for many antibodies, there is quite some flexibility of the antibody conformation upon antigen binding [41], which is currently taken into account only by SnugDock. A possible solution to this drawback is using a computational model of the antibody. Several methods that are specific for modeling antibodies are available, including RosettaAntibody [42] and PIGS [43]. While using homology models instead of X-ray structures for epitope prediction is even more challenging as some degree of inaccuracy originating from the modelling is inevitable, it was shown for the EpiPred method that the performance of a homology-modelled antibody dataset was not statistically significant different than that observed for a set of antibodies with 3D structure determined by X-ray crystallography [40].

Antibody-antigen contact preferences have been applied recently to antibody-specific epitope prediction methods that do not require antibody structure: ASEP [38] is an index, computed based on antibody-antigen contact preferences, used to narrow down candidate residues predicted by conventional methods. Bepar [44] and ABepar [45] are sequence-based methods that are based on association patterns of antibody-antigen residues (Bepar) and on antibody-antigen preferences of individual residues and residue pairs (ABepar).

Combined computational-experimental approaches

We have recently developed PEASE [46,47], a method that performs antibody-specific predictions in the absence of structural information for the antibody (and potentially even for the antigen). This method uses antibody-antigen contact preferences, as well as other properties computed from the antibody sequence and antigen structure or sequence. Similar to some of the most successful traditional methods (EPITOPIA [48], EPSVR [49] and CBTOPE [50]), PEASE is based on a machine-learning model that was trained on known antibody-antigen complexes. As a practical application, we applied this tool to predict epitopes for a panel of 12 mAbs targeting the D8 antigen of VACV. These 12 mAbs were part of a larger panel of hybridomas derived from VACV infected mice based and they were selected based on binding to the D8 antigen and having a unique CDR sequence. Furthermore, cross-blocking assays had been performed for these antibodies for a fast and cost-efficient categorization into groups of antibodies that bind different antigenic sites. It struck us that this is a typical starting point for mAb epitope mapping studies, and that we could directly utilize the cross-blocking information in the prediction scheme. Instead of predicting an epitope for each Ab, we predicted a single antigenic site for each group of Abs that cross-block each other, using the information from all Abs of the group. Then, we filtered out solutions that were not in agreement with the scheme of the relationships between the different groups, e.g. demanding that antibody groups that do not cross-block each other recognize different antigenic sites. When incorporating these conditions into the algorithm, we significantly increased the overall prediction performance [46].

A different kind of experimental data for antibody-specific epitope predictions was utilized by Chuang et. al. [51]. This method aims to identify epitope residues based on correlating the sequence variability of different viral strains with antibody neutralization potency. A limitation of this method is that it can only be applied to viral antigens for which large panels of neutralization potency data are available for different strains. At the same time, advantages of this method over other antibody-specific epitope prediction algorithms are that it does not require the antibody sequence, antibody structure or antigen structure. In cases where both neutralization and sequence and/or structure data are available, one could envision making integrated prediction approaches by e.g. combining likely epitope residues based on neutralization with those identified by PEASE.

Need for a comprehensive benchmark

Several antibody-specific epitope prediction methods exist, some of which require antibody sequence and/or structure, and some require different kinds of experimental data. In order to judge how promising the different approaches are, their prediction quality needs to be

compared. The currently available prediction evaluations for each algorithm are insufficient as the performance of each method is reported for a different dataset using slightly different definition of the epitope and using different metrics of performance success. Ideally, a non-biased independent study should be performed in which these algorithms are compared side-by-side on new datasets that were not included in the derivation of these algorithms. Several such studies were made for traditional epitope prediction methods [18,20]) in order to compare the relative performance of antibody-specific epitope prediction methods. Unfortunately, such an evaluation study of antibody-specific prediction methods is currently not feasible, as most of the antibody-specific epitope prediction methods are not available online or for download.

Conclusions

We propose that the emergence of antibody specific epitope predictions and the concept of making predictions for groups of antibodies have the potential to transform the epitope prediction field. Rapid advances in sequencing technologies have now made it possible to routinely characterize whole antibody repertoires. We hypothesize that families of clonally related antibodies are identifiable from antibody repertoire sequencing data, and that members of each such family will often bind to the same, or similar, antigenic sites. Thus, making predictions for a clonal family of antibodies identified by sequencing should improve prediction performance just as it did when we utilized cross-blocking data to identify antibody groups. This should be straightforward in a highly controlled model system such as immunization of inbred mice with a single antigen, but if successful could be scalable to more challenging application of antibody repertoires from humans immunized with a specific vaccine, or even repertoires from convalescent individuals recently infected with a specific agent.

Finally, a limiting factor on the current performance of most Ab-specific prediction approaches is the availability of high-resolution antibody-antigen complexes. These complexes are the basis for understanding Ab conformation, training Ab specific docking approaches, and learning the interface rules governing the mAb-antigen interaction. In particular, such new data will reduce the noise in the antibody-antigen contact potential matrix, which is the basis for most of the antibody-specific epitope prediction methods. For the 20X20 matrix of pairing potentials to be reliable, each cell (representing the propensity for an interaction between two types of amino-acids), should be based on large number of observations. In particular, it is desirable to have the expected number of contacts between each possible pair of residues to be at least in the order of magnitude of 100 contacts. Some types of amino acids constitute less than 3% of the interfaces [29]. To allow the expected number of contacts between two such amino-acids to be 100, we need over 110,000 residues in our set of interfaces. Given that each antibody-antigen interface covers about 45 residues [29], we will need over 2000 non-redundant complexes, whereas only about 120 non-redundant antibody-antigen structures were available in 2013 [46]. In the world of genomics it is already well accepted that we have to sequence many DNA and RNA sequences that are likely to teach us very little individually, if we ever want to understand genomic variations. Similarly, we will need to solve hundreds of antibody-antigen structures that may not

provide profound insights on their own in order to learn the general principles that govern antibody-antigen recognition.

Acknowledgments

Funding was provided by National Institutes of Health contracts HHSN272201200010C and HHSN272200900048C.

References

1. XU X, SUN J, LIU Q, WANG X, XU T, ZHU R, WU D, CAO X. Evaluation of spatial epitope computational tools based on experimentally-confirmed dataset for protein antigens. *Chinese Science Bulletin*. 2010; 55:2169–2174.
2. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *Bmc Bioinformatics*. 2008; 9
3. Huang J, Honda W. CED: a conformational epitope database. *BMC Immunol*. 2006; 7:7. [PubMed: 16603068]
4. El-Manzalawy Y, Honavar V. Recent advances in B-cell epitope prediction methods. *Immunome Res*. 2010; 6 (Suppl 2):S2. [PubMed: 21067544]
5. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A*. 1981; 78:3824–3828. [PubMed: 6167991]
6. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990; 276:172–174. [PubMed: 1702393]
7. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*. 1986; 25:5425–5432. [PubMed: 2430611]
8. Pellequer JL, Westhof E, Van Regenmortel MH. Predicting location of continuous epitopes in proteins from their primary structures. *Methods Enzymol*. 1991; 203:176–201. [PubMed: 1722270]
9. Welling GW, Weijer WJ, van der Zee R, Welling-Wester S. Prediction of sequential antigenic regions in proteins. *FEBS Lett*. 1985; 188:215–218. [PubMed: 2411595]
10. Pellequer JL, Westhof E, Van Regenmortel MH. Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunol Lett*. 1993; 36:83–99. [PubMed: 7688347]
11. Pellequer JL, Westhof E. PREDITOP: a program for antigenicity prediction. *J Mol Graph*. 1993; 11:204–210. 191–202. [PubMed: 7509182]
12. Alix AJ. Predictive estimation of protein linear epitopes by using the program PEOPLE. *Vaccine*. 1999; 18:311–314. [PubMed: 10506656]
13. Odorico M, Pellequer JL. BEPITOPE: predicting the location of continuous epitopes and patterns in proteins. *J Mol Recognit*. 2003; 16:20–22. [PubMed: 12557235]
14. Blythe MJ, Flower DR. Benchmarking B cell epitope prediction: underperformance of existing methods. *Protein Sci*. 2005; 14:246–248. [PubMed: 15576553]
15. Gao J, Kurgan L. Computational prediction of B cell epitopes from antigen sequences. *Methods Mol Biol*. 2014; 1184:197–215. [PubMed: 25048126]
16. Greenbaum JA, Andersen PH, Blythe M, Bui HH, Cachau RE, Crowe J, Davies M, Kolaskar AS, Lund O, Morrison S, et al. Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *Journal of Molecular Recognition*. 2007; 20:75–82. [PubMed: 17205610]
17. Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*. 2012; 8:e1002829. [PubMed: 23300419]
18. Ponomarenko JV, Bourne PE. Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *Bmc Structural Biology*. 2007; 7

19. Borley DW, Mahapatra M, Paton DJ, Esnouf RM, Stuart DI, Fry EE. Evaluation and use of in-silico structure-based epitope prediction with foot-and-mouth disease virus. *PLoS One*. 2013; 8:e61122. [PubMed: 23667434]
20. Yao B, Zheng D, Liang S, Zhang C. Conformational B-Cell Epitope Prediction on Antigen Protein Structures: A Review of Current Algorithms and Comparison with Common Binding Site Prediction Methods. *Plos One*. 2013; 8:e62249. [PubMed: 23620816]
21. Sun P, Ju H, Liu Z, Ning Q, Zhang J, Zhao X, Huang Y, Ma Z, Li Y. Bioinformatics resources and tools for conformational B-cell epitope prediction. *Comput Math Methods Med*. 2013; 2013:943636. [PubMed: 23970944]
22. Yang XD, Yu XL. An introduction to epitope prediction methods and software. *Reviews in Medical Virology*. 2009; 19:77–96. [PubMed: 19101924]
23. Zhang W, Niu Y, Xiong Y, Ke M. Prediction of conformational B-cell epitopes. *Methods Mol Biol*. 2014; 1184:185–196. [PubMed: 25048125]
24. Benjamin DC, Berzofsky JA, East IJ, Gurd FR, Hannum C, Leach SJ, Margoliash E, Michael JG, Miller A, Prager EM. The antigenic structure of proteins: a reappraisal. *Annu Rev Immunol*. 1984; 2:67–101. [PubMed: 6085753]
25. Novotný J, Handschumacher M, Haber E, Bruccoleri RE, Carlson WB, Fanning DW, Smith JA, Rose GD. Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). *Proc Natl Acad Sci U S A*. 1986; 83:226–230. [PubMed: 2417241]
26. Thornton JM, Edwards MS, Taylor WR, Barlow DJ. Location of ‘continuous’ antigenic determinants in the protruding regions of proteins. *EMBO J*. 1986; 5:409–413. [PubMed: 2423325]
27. Sela-Culang I, Kunik V, Ofra Y. The Structural Basis of Antibody-Antigen Recognition. *Front Immunol*. 2013; 4:302. [PubMed: 24115948]
28. Van Regenmortel MH. Specificity, polyspecificity, and heterospecificity of antibody-antigen recognition. *J Mol Recognit*. 2014; 27:627–639. [PubMed: 25277087]
29. Kunik V, Ofra Y. The indistinguishability of epitopes from protein surface is explained by the distinct binding preferences of each of the six antigen-binding loops. *Protein Eng Des Sel*. 2013; 26:599–609. [PubMed: 23754530]
30. Janin J, Chothia C. The structure of protein-protein recognition sites. *J Biol Chem*. 1990; 265:16027–16030. [PubMed: 2204619]
31. Kringelum JV, Nielsen M, Padkjær SB, Lund O. Structural analysis of B-cell epitopes in antibody:protein complexes. *Mol Immunol*. 2013; 53:24–34. [PubMed: 22784991]
32. Rapberger R, Lukas A, Mayer B. Identification of discontinuous antigenic determinants on proteins based on shape complementarities. *J Mol Recognit*. 2007; 20:113–121. [PubMed: 17421048]
33. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics*. 2014; 30:1771–1773. [PubMed: 24532726]
34. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*. 2005; 33:W363–367. [PubMed: 15980490]
35. Vajda S. Classification of protein complexes based on docking difficulty. *Proteins*. 2005; 60:176–180. [PubMed: 15981248]
- 36*. Sircar A, Gray JJ. SnugDock: Paratope Structural Optimization during Antibody-Antigen Docking Compensates for Errors in Antibody Homology Models. *Plos Computational Biology*. 2010; 6 Advanced antibody specific epitope predictions using docking - this algorithm does not assume rigid docking.
37. Lyskov S, Gray JJ. The RosettaDock server for local protein-protein docking. *Nucleic Acids Res*. 2008; 36:W233–238. [PubMed: 18442991]
38. Soga S, Kuroda D, Shirai H, Kobori M, Hirayama N. Use of amino acid composition to predict epitope residues of individual antibodies. *Protein Engineering Design & Selection*. 2010; 23:441–448.

39. Brenke R, Hall DR, Chuang GY, Comeau SR, Bohnuud T, Beglov D, Schueler-Furman O, Vajda S, Kozakov D. Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics*. 2012; 28:2608–2614. [PubMed: 23053206]
40. Krawczyk K, Liu X, Baker T, Shi J, Deane CM. Improving B-cell epitope prediction and its application to global antibody-antigen docking. *Bioinformatics*. 2014; 30:2288–2294. [PubMed: 24753488]
41. Sela-Culang I, Alon S, Ofra Y. A systematic comparison of free and bound antibodies reveals binding-related conformational changes. *J Immunol*. 2012; 189:4890–4899. [PubMed: 23066154]
42. Sircar A, Kim ET, Gray JJ. RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Research*. 2009; 37:W474–W479. [PubMed: 19458157]
43. Marcatili P, Olimpieri PP, Chailan A, Tramontano A. Antibody structural modeling with prediction of immunoglobulin structure (PIGS). *Nat Protoc*. 2014; 9:2771–2783. [PubMed: 25375991]
44. Zhao L, Li JY. Mining for the antibody-antigen interacting associations that predict the B cell epitopes. *Bmc Structural Biology*. 2010; 10:56.
45. Zhao L, Wong L, Li JY. Antibody-Specified B-Cell Epitope Prediction in Line with the Principle of Context-Awareness. *Ieee-Acm Transactions on Computational Biology and Bioinformatics*. 2011; 8:1483–1494.
- 46*. Sela-Culang I, Benhnia MR, Matho MH, Kaever T, Maybeno M, Schlossman A, Nimrod G, Li S, Xiang Y, Zajonc D, et al. Using a combined computational-experimental approach to predict antibody-specific B cell epitopes. *Structure*. 2014; 22:646–657. This paper shows how to groupings of antibodies with similar specificity can be utilized to improve prediction performance of antibody specific predictions. [PubMed: 24631463]
47. Sela-Culang I, Ashkenazi S, Peters B, Ofra Y. PEASE: Predicting B-Cell Epitopes utilizing Antibody Sequence. *Bioinformatics*. 2014
48. Rubinstein ND, Mayrose I, Martz E, Pupko T. EpiToPIA: a web-server for predicting B-cell epitopes. *Bmc Bioinformatics*. 2009; 10
49. Liang SD, Zheng DD, Standley DM, Yao B, Zacharias M, Zhang C. EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. *Bmc Bioinformatics*. 2010; 11:6. [PubMed: 20047656]
50. Ansari HR, Raghava GP. Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Res*. 2010; 6:6. [PubMed: 20961417]
- 51*. Chuang GY, Acharya P, Schmidt SD, Yang Y, Louder MK, Zhou T, Kwon YD, Pancera M, Bailer RT, Doria-Rose NA, et al. Residue-level prediction of HIV-1 antibody epitopes based on neutralization of diverse viral strains. *J Virol*. 2013; 87:10047–10058. This paper uses neutralization for epitope mapping which could be used complementary to other antibody-specific epitope prediction methods. [PubMed: 23843642]

Highlights

- antibody-independent epitope prediction methods have a poor performance
- methods predicting an epitope for a certain antibody were recently introduced
- the prediction in these methods is done by using antibody sequence and/or structure
- incorporation of antibody cross-blocking data improves prediction performance

Figure 1. known epitopes of the Hemagglutinin antigen. The 3D structure of Hemagglutinin antigen (space-fill representation, PDB ID 1EO8) is shown together with a neutralizing antibody (purple ribbon representation, PDB ID 1KEN). Hemagglutinin epitope residues of the shown neutralizing antibody are colored orange. Other epitope residues (i.e. epitope residues of other antibodies) are colored cyan. The figure was made by superimposing 16 structures of Hemagglutinin co-crystal with an antibody (PDB IDs 1EO8, 1KEN, 1QFU, 2VIR, 2VIS, 2VIT, 3SDY, 3WHE, 3ZTJ, 4FP8, 4FQR, 4FQY, 4GMS, 4KVN, 4NM8 and 4O58) based on the Hemagglutinin structure. Residues were defined as in an epitope if at least one of their non-hydrogen atoms is within a distance of 6 Å from any of the antibody atoms.