BMC Structural Biology

Research article

Open Access

Antibody-protein interactions: benchmark datasets and prediction tools evaluation

Julia V Ponomarenko*[1,2] and Philip E Bourne[1,2]

Address: [1]San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA and [2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA

Email: Julia V Ponomarenko* - jpon@sdsc.edu; Philip E Bourne - bourne@sdsc.edu

* Corresponding author

Published: 2 October 2007

BMC Structural Biology 2007, **7**:64 doi:10.1186/1472-6807-7-64

Received: 9 April 2007
Accepted: 2 October 2007

This article is available from: http://www.biomedcentral.com/1472-6807/7/64

© 2007 Ponomarenko and Bourne.; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The ability to predict antibody binding sites (aka antigenic determinants or B-cell epitopes) for a given protein is a precursor to new vaccine design and diagnostics. Among the various methods of B-cell epitope identification X-ray crystallography is one of the most reliable methods. Using these experimental data computational methods exist for B-cell epitope prediction. As the number of structures of antibody-protein complexes grows, further interest in prediction methods using 3D structure is anticipated. This work aims to establish a benchmark for 3D structure-based epitope prediction methods.

Results: Two B-cell epitope benchmark datasets inferred from the 3D structures of antibody-protein complexes were defined. The first is a dataset of 62 representative 3D structures of protein antigens with inferred structural epitopes. The second is a dataset of 82 structures of antibody-protein complexes containing different structural epitopes. Using these datasets, eight web-servers developed for antibody and protein binding sites prediction have been evaluated. In no method did performance exceed a 40% precision and 46% recall. The values of the area under the receiver operating characteristic curve for the evaluated methods were about 0.6 for ConSurf, DiscoTope, and PPI-PRED methods and above 0.65 but not exceeding 0.70 for protein-protein docking methods when the best of the top ten models for the bound docking were considered; the remaining methods performed close to random. The benchmark datasets are included as a supplement to this paper.

Conclusion: It may be possible to improve epitope prediction methods through training on datasets which include only immune epitopes and through utilizing more features characterizing epitopes, for example, the evolutionary conservation score. Notwithstanding, overall poor performance may reflect the generality of antigenicity and hence the inability to decipher B-cell epitopes as an intrinsic feature of the protein. It is an open question as to whether ultimately discriminatory features can be found.

Background

A B-cell epitope is defined as a part of a protein antigen recognized by either a particular antibody molecule or a particular B-cell receptor of the immune system [1]. The main objective of B-cell epitope prediction is to facilitate the design of a short peptide or other molecule that can be

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

synthesized and used instead of the antigen, which in the case of a pathogenic virus or bacteria, may be harmful to a researcher or experimental animal [2]. A B-cell epitope may be continuous, that is, a short contiguous stretch of amino acid residues, or discontinuous, comprising atoms from distant residues but close in three-dimensional space and on the surface of the protein.

Synthetic peptides mimicking epitopes, as well as anti-peptide antibodies, have many applications in the diagnosis of various human diseases [3-7]. Also, the attempts have been made to develop peptide-based synthetic prophylactic vaccines for various infections, as well as therapeutic vaccines for chronic infections and noninfectious diseases, including autoimmune diseases, neurological disorders, allergies, and cancers [8-10]. The immunoinformatics software and databases developed to facilitate vaccine design have previously been reviewed [11,12].

During the last 25 years B-cell epitope prediction methods have focused primarily on continuous epitopes. They were mostly sequence-dependent methods based upon various amino acid properties, such as hydrophilicity [13], solvent accessibility [14], secondary structure [15-18], and others. Recently, several methods using machine learning approaches have been introduced that apply hidden Markov models (HMM) [19], artificial neural networks (ANN) [20], support vector machine (SVM) [21], and other techniques [22,23]. Recent assessments of continuous epitope prediction methods demonstrate that "single-scale amino acid propensity profiles cannot be used to predict epitope location reliably" [24] and that "the combination of scales and experimentation with several machine learning algorithms showed little improvement over single scale-based methods" [25].

As crystallographic studies of antibody-protein complexes have shown, most B-cell epitopes are discontinuous. In 1984, the first attempts at epitope prediction based on 3D protein structure was made for a few proteins for which continuous epitopes were known [26-28]. Subsequently, Thornton and colleagues [29] proposed a method to locate potential discontinuous epitopes based on a protrusion of protein regions from the protein's globular surface. However, until the first X-ray structure of an antibody-protein complex was solved in 1986 [30], protein structural data were mostly used for prediction of continuous rather than discontinuous epitopes.

In cases where the three-dimensional structure of the protein or its homologue is known, a discontinuous epitope can be derived from functional assays by mapping onto the protein structure residues involved in antibody recognition [31]. However, an epitope identified using an immunoassay may be an artefact of measuring cross-reactivity of antibodies due to the presence of denatured or degraded proteins [32,33], or due to conformational changes in the protein caused by residue substitutions that may even lead to protein mis-folding [34]. Therefore, structural methods, particularly X-ray crystallography of antibody-antigen complexes, generally identify B-cell epitopes more reliably than functional assays [35].

B-cell epitopes can be thought of in a structural and functional sense. Structural epitopes (also called antigenic determinants) are defined by a set of residues or atoms in the protein antigen contacting antibody residues or atoms [33,36]. In contrast, a functional epitope consists of antigen residues that contribute significantly to antibody binding [36,37]. Functional epitopes are determined through functional assays (e.g., alanine scanning mutagenesis) or calculated theoretically using known structures of antibody-protein complexes [38,39]. Thus, functional and structural epitopes are not necessary the same. Functional epitopes in proteins are usually smaller than structural epitopes; only three to five residues of the structural epitope contribute significantly to the antibody-antigen binding energy [40]. This work focuses on structural epitopes inferred from known 3D structures of antibody-protein complexes available in the Protein Data Bank (PDB) [41].

Antibody-protein complexes can be categorized as intermediate transient non-obligate protein-protein complexes [40,42]. Non-obligate complexes, implying that individual components can be found on their own *in vivo*, are classified as either permanent or transient depending on their stability under particular physiological and environmental conditions [43]. For example, many enzyme-inhibitor complexes are permanent non-obligate complexes. Transient non-obligate complexes range from weak (e.g., electron transport complexes), to intermediate (e.g., signal transduction complexes), and to strong (e.g., bovine G protein forming a stable trimer upon GDP binding) [44]. Most antibodies demonstrate intermediate affinity for their specific antigens [45]. Based on this classification, general methods for the prediction of intermediate transient non-obligate protein-protein interactions have been applied to the prediction of structural epitopes [40,42]. For example, Jones and Thornton, using their method for predicting protein-protein binding sites [46], successfully predicted B-cell epitopes on the surface of the β-subunit of human chorionic gonadotropin (βhCG) [47].

Since the number of available structures of antibody-protein complexes remains limited, thus far only a few methods, CEP (Conformational Epitope Prediction) [48] and DiscoTope [49], for B-cell epitope prediction using a protein of a given three-dimensional structure have been

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

developed. In the near future, with growth in the number of available structures of antibody-protein complexes, extensive development in this area is expected. Existing and new methods for epitope prediction demand a benchmark which will set the standard for the future comparison of methods. To facilitate the further development of this standard, we have developed B-cell epitope benchmark datasets inferred from existing 3D structures of antibody-protein complexes. Further, using the benchmark datasets, we evaluated CEP, DiscoTope, and six recently developed publicly available web-servers for generalized protein-protein binding site prediction using various approaches: protein-protein docking (ClusPro [50], DOT [51] and PatchDock [52]); structure-based methods applying different principals and trained on different datasets (PPI-PRED [53], PIER [54] and ProMate [55]), and residue conservation (ConSurf [56]).

Results and discussion

Structural epitope definition

Three definitions of an epitope inferred from the X-ray structures of antibody-protein complexes were considered: (1) The epitope consists of protein antigen residues in which any atom of the residue looses more than 1Å^2 of accessible surface area (ASA) upon antibody binding. ASA was calculated using the program NACCESS [57]; (2) The epitope consists of protein antigen residues in which any atom of the epitope residue is separated from any antibody atom by a distance ≤ 4Å; (3) The epitope consists of protein antigen residues in which any atom of the epitope residue is separated from any antibody atom by a distance ≤ 5Å. These three definitions were used for two reasons. First, the methods evaluated in this work use one of these three definitions, second, we wished to study how the epitope definition influenced the results.

Results (not shown) indicated that the structural epitope definition did not influence the outcome. Hence, unless otherwise specified, results are based on the second epitope definition.

Construction of the benchmark datasets

Two benchmark datasets were derived from the 3D structures of antibody-protein complexes available from the PDB [41]:

• Dataset #1 – Representative 3D structures of protein antigens with structural epitopes inferred from 3D structures of antibody-protein complexes. This dataset is intended for the study of the antigenic properties of proteins as well as for development and evaluation of the methods based on protein structure alone, or protein-protein unbound docking methods, that is, if the structure of the antibody is known or can be modeled. Here this dataset was used for the evaluation of scale-based methods (DiscoTope, PIER, ProMate and ConSurf). The dataset contains 62 antigens, 52 of which are one-chain antigen proteins.

• Dataset #2 – Representative 3D structures of antibody-protein complexes presenting different epitopes. This dataset is useful for the study of the properties of individual epitopes as well as for development and evaluation of protein-protein bound docking methods. Since the current work attempts to compare the methods of different types, including protein-protein docking methods, this dataset was used to compare the performance of all methods to each other. The dataset contains 70 structures of proteins in complexes with two-chain antibodies and 12 structures of proteins in complexes with one-chain antibodies.

The flowchart describing the construction of the benchmark datasets is shown in Figure 1. Steps from 1 to 4 relate to dataset #1; steps 1–6 relate to dataset #2.

Step 1 – crystal structures of protein antigens of length ≥30 amino acids at a resolution ≤ 4Å in complex with antibody fragments containing variable regions (Fab, VHH, Fv, or scFv fragments) were collected from the Protein Data Bank (PDB) [41]. Structures in which the antibody binds antigen but involves no CDR residues have been excluded from the analysis; there were four such structures [PDB: 1MHH, 1HEZ, 1DEE, 1IGC]. If a structure contained several complexes in one asymmetric unit and there was no structural difference observed between these complexes, only one complex was selected. In this way 166 structures containing 187 antibody-protein complexes were selected: 24 complexes were formed by one-chain antibody fragments and 163 complexes by two-chain antibody fragments.

Step 2 – all antigen protein chains were structurally aligned to one another using the CE algorithm [58]. Two protein chains were considered similar if all the following conditions applied: (i) rmsd ≤3Å, (ii) z-score ≥4.0, (iii) number of residue-residue matches relative to the length of the longest chain ≥80%, (iv) sequence identity in the structural alignment (not considering gaps) ≥80%. The z-score takes into account overall structural similarity and number of gapped positions. Two protein molecules were considered similar if each chain in one protein had a similar chain in another protein. Figure 2 demonstrates how the last two parameters, number of matches and sequence identity in the structural alignment, are defined.

The structural alignment rather than sequence alignment was used because protein structure is more conserved than sequence, and there can be expected regions in proteins with low sequence similarity that cannot be aligned by

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

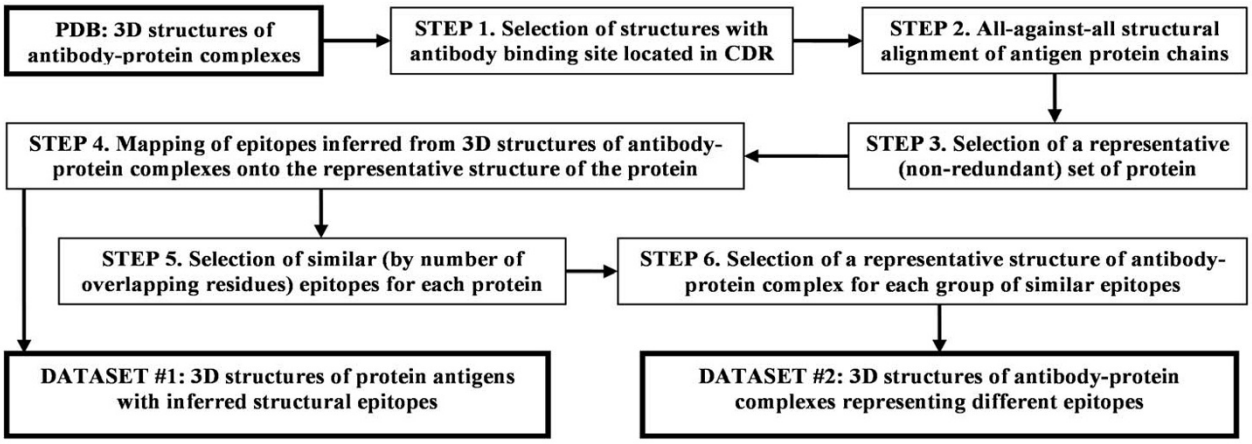

Figure 1
Flowchart for building benchmark datasets.

sequence alone. The structural alignment also avoids considering two proteins as similar if they have similar sequences but different structures (possible over short regions). The threshold values were chosen empirically based on previous experience working with the CE algorithm. As a result, the chosen threshold values separated human and bird lysozymes (61% sequence identity) and neuraminidases of different influenza virus strains, H3N2 and H11N9 (47% sequence identity).

Step 3 – 35 proteins were orphans represented by only one 3D structure. Of the remaining 27 proteins represented by more than one 3D structure, the structure with the best resolution was selected as the representative structure. The final representative dataset contained 62 antigens [see Additional file 1], 52 of which were one-chain antigen proteins.

```
(A)  AVCQ---YWC
(B)  A-CYARTY-C
```

Figure 2
Hypothetical example of the structural alignment of proteins (A) (sequence AVCQYWC) and (B) (sequence ACYARTYC). Number of residue-residue matches = 5, number of residue-residue matches relative to the length the longest chain = 63% (5/8), sequence identity = 80% (4/5).

Step 4 – for each protein, epitopes inferred from the 3D structures of antibody-protein complexes were mapped onto the representative structure of the protein. First, epitope residues were calculated for each complex structure using one of the aforementioned epitope definitions. Second, epitope residues defined for the represented structures were mapped onto the representative structure based on the structure alignments. For example, the hemagglutinin HA1 chain of influenza A virus was represented by six 3D structures of the protein in complexes with Fab fragments of antibodies HC45 [PDB:1QFU], BH151 [PDB:1EO8], HC63 [PDB:1KEN], and HC19 [PDB:2VIR, 2VIS, 2VIT]. Figure 3 illustrates a representative structure [PDB:1EO8] of hemagglutinin HA1 upon which epitopes are mapped having been inferred from six complex structures. In this way, epitopes inferred from 187 structures of antibody-protein complexes were mapped onto the 62 representative protein structures. The resulting dataset is denoted dataset #1. Data on mapped epitope residues are available upon request.

Step 5 – to study the properties of individual epitopes and their prediction a dataset of representative epitopes, dataset #2 derived from 3D structures of antibody-protein complexes defining different epitopes was constructed. An important question to consider is how to define individual epitopes yet avoid bias by over-presentation of particular epitopes? For example (Fig. 3), while HC45 (blue) and BH151 (magenta) epitopes overlap, neither HC63 (green) nor HC19 (red) epitopes overlap, they are separated on the protein surface. Nevertheless, HC45 and BH151 epitopes share residues (orange in Fig. 3), as do HC63 and HC19 epitopes (yellow in Fig. 3). Are HC45

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

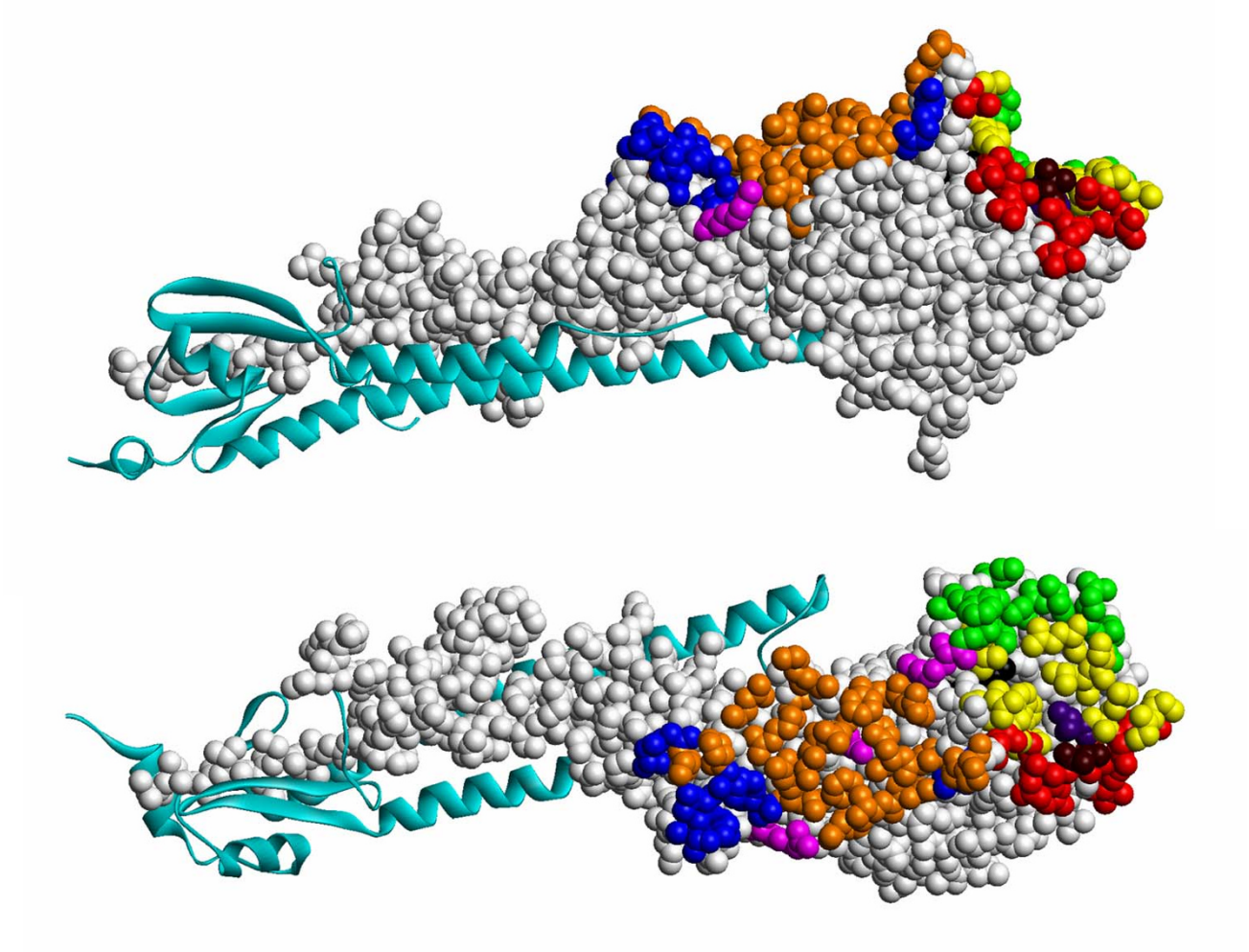

Figure 3
Two orthogonal views of a representative structure, influenza A virus hemagglutinin HA1 chain [PDB:1EO8]. Chain A is shown in light gray upon which are mapped epitope residues inferred from six protein structures in complexes with antibody fragments: HC45 Fab [PDB:1QFU] (blue), BH151 Fab [PDB:1EO8] (magenta), HC63 Fab [PDB:1KEN] (green), HC19 Fab [PDB:2VIR, 2VIS, 2VIT] (red). The hemagglutinin HA2 chain is shown in cyan. Residues common to HC45 and BH151 epitopes are shown in orange; residues common to HC63 and HC19 epitopes are shown in yellow; residue Tyr98 which is a part of HC19 epitope inferred from structure 2VIR but not from 2VIS and 2VIT structures is shown in black; The HC19 epitope residue Thr131 which is mutated to Ile in the 2VIS structure is shown in dark red. The HC19 epitope residue Thr155 which is mutated to Ile in 2VIT structure is shown in violet.

and BH151 epitopes similar or different? This question is answered by considering the degree of overlap.

Two epitopes are deemed similar if, in addition to the aforementioned criteria for epitope definition, they belong to similar protein chains and have >75% residues in common for both epitopes. A cut-off value of 75% for epitope similarity was chosen empirically. Thus, the HC45 and BH151 epitopes on influenza A virus hemagglutinin HA1 (Fig. 3) share 14 residues, that make up 74% and 93% of the size of HC45 and BH151 epitopes, respectively. A cut-off on epitope overlap of less than 75% would define HC45 and BH151 epitopes as similar even though they are known to be different. HC45 and BH151 are antibodies from different germ-lines with variable domains sharing only 56% sequence similarity, their H3 CDR regions adopt distinct conformations and these antibodies are tolerant to different mutations in hemagglutinin [59]. Another example, X5 and 17B epitopes of gp120 share 75% of their residues yet X5 and 17B antibodies are from different genes [60]. A cut-off value for epitope similarity equal to or less than 75% would erroneously define X5 and 17B epitopes as similar. Conversely, a cut-off value of 80% would make epitopes inferred from different

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.