

BEID: Database for sequence-structure-function information on antigen-antibody interactions

Joo Chuan Tong^{1*}, Chun Meng Song¹, Paul Thiam Joo Tan², Ee Chee Ren^{3,4}, Animesh A. Sinha⁵

¹Institute for Infocomm Research, 1 Fusionopolis Way, #21-01 Connexis, South Tower, Singapore 138632; ²Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA; ³Department of Microbiology, Faculty of Medicine, National University of Singapore, Singapore 117597; ⁴Singapore Immunology Network, 60 Biopolis Street, Genome, #02-01, Singapore 138672; ⁵Center for Investigative Dermatology, Division of Dermatology and Cutaneous Sciences, Michigan State University, East Lansing, MI, USA; Joo Chuan Tong - E-mail: jctong@i2r.a-star.edu.sg; Phone: 65 6408 2156; * Corresponding author

received September 18, 2008, accepted September 21, 2008; published October 13, 2008

Abstract:

The B-cell Epitope Interaction Database (BEID; <http://datam.i2r.a-star.edu.sg/BEID>) is an open-access database describing sequence-structure-function information on immunoglobulin (Ig)-antigen interactions. The current version of the database contains 164 antigens, 126 Ig and 189 Ig-antigen complexes extracted from the Protein Data Bank (PDB). Each entry is manually verified, classified, and analyzed for intermolecular interactions between antigens and the corresponding bound Ig molecules. Ig-antigen interaction information that is stored in BEID includes solvent accessibility, hydrogen bonds, non-hydrogen bonds, gap volume, gap index, interface area and contact residues. The database can be searched with a user-friendly search tool and schematic diagrams for Ig-antigen interactions are available for download in PDF format. The ultimate purpose of BEID is to enhance the understanding of the rules of engagement between antigen and the corresponding bound Ig molecules. It is also a precious data source for developing computational predictors for B-cell epitopes.

Keywords: database; epitope; antigen; antibody; sequence-structure function

Background:

Immunoglobulins (Ig), or antibodies, are proteins generated by B-cells in response to antigenic substances. The site of contact between antigens and Ig molecules are called B-cell epitopes [1]. Approximately 10% of these antigenic determinants are linear, consisting of a single continuous stretch of amino acids along the polypeptide chain [2]. Most B-cell epitopes, though, are thought to be conformational, where distantly separated residues of the polypeptide chain are brought into spatial proximity by protein folding [3]. Mutational analysis of B-cell epitopes showed that antibody binding could be reduced or eliminated by single-site amino acid substitution [4]. On the other hand, solution structures of antigen-antibody complexes showed that antibodies with dissimilar binding site structures may exhibit similar specificities for common epitopes [5] and not all residues within an epitope are functionally important for binding [6]. As such, a detailed understanding of the sequence-structure-function relationship between antigens and their corresponding bound antibodies is essential for effective vaccine design.

Several databases currently exist to facilitate the characterization of antibodies: IMGT/HLA [7], IMGT/3Dstructure-DB [8] and the Immune Epitope Database and Analysis Resource (IEDB) [9] store information on antibody sequences and structures annotated according to the IMGT-ONTOLOGY [10]. AntiJen [11] provides quantitative binding data for both continuous and discontinuous B-cell epitope molecules.

The Conformational Epitope Database (CED) [12] details information of 225 conformational epitopes derived from the literature. Bcipep [13] contains records of 3031 experimentally determined linear B-cell epitopes collected from literature and other public databases. Epitome [14] maintains a list of antibody and antigen residues that are involved in specific interactions, while the Summary of Antibody Crystal Structures (SACS) [15] provides fully automated web-based summaries of the latest antibody crystal structures in the Protein Data Bank (PDB) [16]. Despite these rigorous efforts, most existing resources do not focus on in depth characterization of the antigen-antibody interface.

Here, we describe BEID, a manually curated database containing detailed characterization of 189 antigen-antibody interaction sites. The database contains both linear and conformational epitopes available in the PDB. Each entry describes a specific antigen-antibody interaction in terms of a set of sequence and structural parameters representative of molecular recognition. BEID will provide a valuable resource for investigators working in the areas of vaccine design and allergy research. The sequence and structural parameters also serve as important data sources for the development of B-cell epitope prediction tools.

Methodology:

Construction and content:

BEID is a MySQL relational database hosted on a UNIX server (SunOS 5.10, Apache 2.0.59). The database can be searched by antibody name and PDB code, with options for formulating complex queries and customizing the output. Currently, BEID contains only the structures of experimentally determined antigen-antibody complexes derived from the PDB. Each entry in BEID bears a unique identifier, name and category of the antigen, name and category of the bound antibody, experiment method, resolution of the structure, release year and bibliographic reference. The intermolecular hydrogen bonds, gap volume and gap index are computed using SURFNET [17] and the interface area of the complex is calculated based on the program NACCESS [18]. Schematic diagrams based on the plotting program LIGPLOT [19] are also provided under the fieldname "Interaction Map" to illustrate explicit antigen-antibody interactions, and are particularly useful for analysis of discontinuous intermolecular contact residues.

Data clustering

Information in the database is classified into five main categories: i) antibody (name, source), ii) isotype (IgA,

IgG, IgE, IgM, IgD), iii) bound ligand (protein, hapten, sugar, steroid, others), iv) computed interaction parameters (intermolecular hydrogen bonds, gap volume, gap index, interface area, contact residues), and v) links to external related databases including IMGT/3Dstructure-DB, AntiJen and Bcipep. For PDB entries containing many molecular assemblies, all antigen-antibody complexes are characterized and stored in individual entries to facilitate analysis.

Definition of molecular descriptors

Detailed description of the interaction parameters have been described elsewhere [20]. A brief outline of the descriptors follows: (1) Number of intermolecular hydrogen bonds **minus** the total number of hydrogen bonds between the antibody and its corresponding bound ligand; (2) Interface area is the change in solvent accessible surface area on complexation from an unbound antibody to a bound antigen-antibody complex state; (3) Gap volume is the volume enclosed by the interacting antibody and its corresponding bound ligand. Gap index (\AA^3) is the ratio of Gap volume between Ig-antigen (\AA^3) by Interface ASA (\AA^2) (per complex).

Figure 1: BEID search page. The user-friendly search page allows queries for Ig-antigen interaction records to be retrieved based on various properties of antigen, antibody, or PDB.

Utility:**User interface**

A user-friendly web query interface allows users to search for specific antigen-antibody interactions. A help page for browsing BEID is provided. Users can query the database by 1) antigen-antibody data or 2) PDB data. An antigen-antibody search is formed by selecting the following fields: antibody source, ligand category, antibody isotype, antibody name and ligand name. Query results can also be customized by selecting only the parameters of interest, which include interaction data, antibody information, ligand information, resolution and bibliographical references. A PDB search allows users to formulate complex queries based on the PDB ID, resolution, release year, and ligand name, and combining them logically using the “and” or “or” radio buttons (Figure 1). Search results are presented in tabular forms that display detailed information of each antigen-antibody interaction site. Selecting the bond number in the summary table opens a new window that details the atomic contacts of the specific interactions. Visual analysis of antigens, antibodies and antigen-antibody complexes can be performed through the use of freely available graphics applications such as CHIME, RASMOL or Jmol. Antigen-antibody interaction maps can also be downloaded as Adobe PDF files. References to PubMed are also provided for easy access to experimental evidence of interaction.

Conclusion:

BEID provides a platform for studying the rules of engagement between antigens and their corresponding bound antibodies, which is a key to understanding the mechanisms of humoral immune responses toward foreign antigens. The molecular descriptors, together with biochemical and functional studies can help direct future research into antibody binding studies, with direct implications in disease diagnosis, treatment and vaccine design. Further investigations are being pursued for incorporating additional molecular descriptors characterizing the antigen-antibody interaction region.

References:

- [01] M. H. Van Regenmortel, *Ann. Biol. Clin.*, 51: 39 (1993) [PMID: 7687834]
- [02] J. E. Larsen *et al.*, *Immunome Res.*, 2: 2 (2006) [PMID: 16635264]
- [03] D. J. Barlow *et al.*, *Nature*, 322: 747 (1986) [PMID: 2427953]
- [04] R. M. Helm *et al.*, *J. Allergy Clin. Immunol.*, 105: 378 (2000) [PMID: 10669862]
- [05] D. T. Nair *et al.*, *J. Immunol.*, 168: 2371 (2002) [PMID: 11859128]
- [06] Selo *et al.*, *Clin. Exp. Allergy*, 29: 1055 (1999) [PMID: 10457108]
- [07] J. Robinson *et al.*, *Nucleic Acids Res.*, 29: 210 (2001) [PMID: 11125094]
- [08] Q. Kaas *et al.*, *Nucleic Acids Res.*, 32: D208 (2004) [PMID: 14681396]
- [09] B. Peters *et al.*, *PLoS Biol.*, 3: e91 (2005) [PMID: 15760272]
- [10] V. Giudicelli and M. Lefranc, *Bioinformatics*, 15: 1047 (1999) [PMID: 10745995]
- [11] C. P. Toseland *et al.*, *Immunome Res.*, 1: 4 (2005) [PMID: 16305757]
- [12] J. Huang and W. Honda, *BMC Immunol.*, 7: 7 (2006) [PMID: 16603068]
- [13] S. Saha *et al.*, *BMC Genomics*, 6: 79 (2005) [PMID: 15921533]
- [14] A. Schlessinger *et al.*, *Nucleic Acids Res.*, 34: D777 (2006) [PMID: 16381978]
- [15] L. C. Allcorn and A. C. Martin, *Bioinformatics*, 18: 175 (2002) [PMID: 11836226]
- [16] H. M. Berman *et al.*, *Nucleic Acids Res.*, 28: 235 (2000) [PMID: 10592235]
- [17] R. A. Laskowski, *J. Mol. Graph.*, 13: 323 (1995) [PMID: 8603061]
- [18] <http://www.bioinf.manchester.ac.uk/naccess/>
- [19] A. C. Wallace *et al.*, *Protein Eng.*, 8: 127 (1995) [PMID: 7630882]
- [20] J. C. Tong *et al.*, *Appl. Bioinformatics*, 5: 111 (2006) [PMID: 16722775]

Edited by P. Kanguane

Citation: Tong *et al.*, *Bioinformatics* 3(2): 58-60 (2008)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.