## UC San Diego UC San Diego Previously Published Works

## Title

IEDB-3D: structural data within the immune epitope database

## Permalink

https://escholarship.org/uc/item/1ft2565z

### Journal

Nucleic Acids Research, 39(suppl\_1)

## ISSN

0305-1048

## **Authors**

Ponomarenko, Julia Papangelopoulos, Nikitas Zajonc, Dirk M <u>et al.</u>

## **Publication Date**

2011

## DOI

10.1093/nar/gkq888

Peer reviewed

# **IEDB-3D:** structural data within the immune epitope database

Julia Ponomarenko<sup>1,2,\*</sup>, Nikitas Papangelopoulos<sup>1</sup>, Dirk M. Zajonc<sup>3</sup>, Bjoern Peters<sup>3</sup>, Alessandro Sette<sup>3</sup> and Philip E. Bourne<sup>1,2</sup>

<sup>1</sup>San Diego Supercomputer Center, <sup>2</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, CA 92093 and <sup>3</sup>La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037, USA

Received August 14, 2010; Revised September 16, 2010; Accepted September 19, 2010

#### ABSTRACT

IEDB-3D is the 3D structural component of the Immune Epitope Database (IEDB) available via the 'Browse by 3D Structure' page at http://www.iedb .org. IEDB-3D catalogs B- and T-cell epitopes and Major Histocompatibility Complex (MHC) ligands for which 3D structures of complexes with antibodies. T-cell receptors or MHC molecules are available in the Protein Data Bank (PDB). Journal articles that are primary citations of PDB structures and that define immune epitopes are curated within IEDB as any other reference along with accompanying functional assays and immunologically relevant information. For each curated structure, IEDB-3D provides calculated data on intermolecular contacts and interface areas and includes an application. EpitopeViewer, to visualize the structures. IEDB-3D is fully embedded within IEDB, thus allowing structural data, both curated and calculated, and all accompanying information to be queried using multiple search interfaces. These include aueries for epitopes recognized in different pathogens, eliciting different functional immune responses, and recognized by different components of the immune system. The query results can be downloaded in Microsoft Excel format, or the entire database, together with structural data both curated and calculated, can be downloaded in either XML or MySQL formats.

#### INTRODUCTION

The Immune Epitope Database (IEDB) (1) catalogs experimentally identified B- and T-cell epitopes and MHC ligands through manual curation of the scientific literature. By the end of 2011, all published epitopes from infectious agents [except HIV, which are maintained in (2)], allergens and autoimmune diseases should be included in the IEDB. In addition, all epitopes characterized by 3D structures of immune receptors complexed with antigens found in the Protein Data Bank (PDB) (3) will be included, independent of the disease association of the antigen. This expansion in scope is made as to understand the general structural principles of epitope recognition a large dataset is desirable, but detailed 3D structural information on epitope complexes is rare.

The 3D structural component of IEDB, called here IEDB-3D, is embedded seamlessly within the general IEDB, thus ensuring that each reference describing 3D structure is curated within IEDB as any other journal article describing immune epitopes. The focus on curation of immunological and epitope relevant information is the major difference between IEDB and IMGT/3Dstructure-DB (4). Similar to IMGT/3Dstructure-DB (4) and other related databases, such as MPID-T (5), Epitome (6), BEID (7) and CED (8), IEDB provides calculated data on intermolecular contacts and interface areas and includes an application to visualize the structure, EpitopeViewer (9), which is a high-quality graphic and rendering tool. Among the aforementioned databases, only CED and IEDB curate epitope residues from the literature. In IEDB, antibody, MHC and T-cell receptor (TCR) residues interacting with the epitope are also curated if they are provided in the reference. Thus, within IEDB, curated and derived data on epitope and antigen-receptor interactions can be seen and compared side-by-side on the IEDB web page and also through launching EpitopeViewer.

#### **IEDB-3D OVERVIEW**

IEDB-3D is fully integrated within IEDB, with its data structure, ontology and query interface indistinguishable from those of the IEDB database. IEDB-3D is the 3D structural component of the Immune Epitope Database (IEDB) available via the 'Browse by 3D Structure' page

\*To whom correspondence should be addressed. Tel: +1 858 822 5415; Fax: +1 858 822 0873; Email: jpon@sdsc.edu

© The Author(s) 2010. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/ by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.



Figure 1. The screen shot of IEDB web page.

at http://www.iedb.org (Figure 1). Similar to IEDB, the types of curated data include experiments describing recognition of epitopes or antigens (peptidic or nonpeptidic) by TCRs (T-cell assays describing 3D structures of antigens/epitopes in complexes with MHC and TCR), immunoglobulins or antibodies (B-cell assays describing structures of antibody-antigen complexes), and MHC molecules (MHC binding assays describing structures of antigens/epitopes in complex with MHC) (Figure 2).

Curation of 3D complex data is handled like the remaining data in the IEDB, except that all epitopes are considered in scope, including those from HIV or cancer. IEDB does not handle HIV data, if the paper provides only information on immunological assay involving HIV antigen explicitly, since these data are curated in HIV database (2). However, if 3D structural data are available for HIV epitopes, they are curated, but with a lower priority compared to other 3D structural data. In IEDB-3D, in addition to the immunological context curated for all epitopes in IEDB, information on 3D structure is provided in the section '3D Structure of Complex' (Figure 3).

An overview of IEDB-3D content is given in Table 1. Current statistics on the number of distinct epitope structures curated by assay type and source organism can be accessed from the 'Browse by 3D Structure' page (Figure 1, red box).

#### **CURATION OF 3D STRUCTURAL DATA**

The process of curating epitopes from 3D structures is shown in Figure 2. To be curated, the paper describing



Figure 2. Objects, processes and roles as represented in IEDB-3D.

the structure (primary citation in PDB) should be published; likewise the paper describing the structure cannot be curated if the structure has not been deposited in PDB, unless this article describes the epitope in the

Epitope Information			Enitone			
Epitope	77404	L	Ерноре			
Epitope ID: Chemical Type:	Discontinuous protein residues	_				
Discontinuous Residues:	T85, A86, L87, E89, P91, T92, Y95, N98, Q10	0, S109,	C110, T111, S112, K113,			
Source Name:	A114, V115, I162, K163, G164 Major surface aptigen p30 precursor					
Source Accession:	129348					
Source Organism ID:	5811					
Source Organism: Enitone Reference Details	Toxoplasma gondii					
Epitope Structure Defines:	Epitope containing region/antigenic site					
Epitope Evidence Code:	Exact match to reference information					
Epitope Name:	4F11E12 Fab epitope of SAG1	V2047 P	12050 02052 52061			
Reference Region:	C2062, T2063, S2064, K2065, A2066, V2067,	12114.1	2115, G2116			
B Cell Assay Information						
Host Organism ID:	10090	Im	munization			
Host Organism:	Mus musculus					
1st In Vivo Process	Administration in vive					
1st In Vivo Process Administratio	n Details					
Route:	Intraperitoneal (i.p.)					
1st Immunogen	Courses Organization					
Epitope Relation: Object Type:	Organism					
Organism ID:	5811					
Organism:	Toxoplasma gondii					
Immunogen Details	Not determined					
Immunogen Reference Name:	Toxoplasma gondii					
Immunization Comments						
Immunization Comments:	mAb 4F1E12 was generated by intraperitoneal 76K T. gondii strain, boosted one month later	infection	of mice with cysts from the usly with formalin-treated.			
and comments.	fixed and purified tachyzoites.		,			
B Cell Assay			Assav			
Qualitative Measurement:	Positive X-Ray Crystallography		лэзау			
Assay Type: Assay Type Group:	Characterization of Ab binding					
Assay Type Units:	Angstroms					
Measurement Details	-					
Quantitative measurement:	3.1					
Assayed Antibody			Anthradia			
Assayed Antibody Source Material:	Purified Immunoglobulin		Antibody			
Assayed Antibody	Cale					
Immunoglobulin Domain:	rap					
Assayed Antibody Purification Status:	Monoclonal					
Assayed Antibody Name:	4F11E12					
Assayed Antibody Heavy Chain Type:	IgG1					
Assayed Antibody Light Chain	Kanna					
Antigen		-				
Epitope Relation:	Source Antigen		Antigen			
Chemical Type:	Protein					
Molecule Name: Molecule Accession	Major surface antigen p30 precursor 129348					
Organism ID:	5811					
Organism:	Toxoplasma gondii					
Antigen Details	Exact match to reference information					
Antigen Conformation:	Native					
	Ivauve					
3D Structure of Complex		ctur	a of Complex			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain	IYNT 🖛 3D Strue	ctur	e of Complex			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain:	IYNT A 3D Strue	ctur	e of Complex			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain:	1YNT 📥 3D Strue	ctur	e of Complex			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments:	INNT A 3D Strue B A F There is no drastic conformational change in the observed comparing the unbound homodimer's	ctur he D1 or	D2 domain of SAG1			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments:	INIT         3D Struct           B         B           A         F           Three is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.         Initial change in the structure in the structure.	cture	D2 domain of SAG1 D2: 1KZQ] and Fab-bound			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts	INIT  B A F F There is no drastic conformational change in th observed comparing the unbound homodimeri monomeric structure. E: T2037. A2038. L2039. E2041. P2043. T2044	<b>ctur</b> he D1 or ic form [1	D2 domain of SAG1 D2 domain of SAG1 DBB: 11KZQ] and Fab-bound			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues:	A B A B B B B B B B B B B B B B B B B B	cture he D1 or ic form [1 14, Y204] 12114, J	e of Complex D2 domain of SAG1 PDB: 1KZQ] and Fab-bound 7, N2050, Q2052, S2061, (2115, G2116			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting	A F There is no drastic conformational change in th observed comparing the unbound homodimeri monomeric structure. F: T2037, A2038, L2039, E2041, P2043, T204 C2062, T2033, S2064, K2065, A2066, V2067, A: Y32L1, Y5012, FS312, N9213, T9313, L GS33H1, T50249, S25940 - T56340 - Y6441	cture he D1 or ic form [1 14, Y204', 12114, 1 94:L3; B	e of Complex			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen:	INT  B B A F Three is no drastic conformational change in th observed comparing the unbound homodimeri monomeric structure. F: T2037, A2038, L2039, E2041, P2043, T204 C2062, T2063, S2064, K2065, A2066, V2067, A: V32L1, V533:L2, K533:L2, K533:L2, K543:L3, K5542, K5542, K5542, K5544 S5542, K5542, K5542, K5544 C35542, K5542, K5544 C35642, S55442, S56443, S6043, T601	cture he D1 or ic form [I .12114, ] .94:L3; B H2, S555 H3, W60	e of Complex			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody:	INIT         A           B         B           A         F           There is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.         F: T2037, A2038, L2039, E2041, P2043, T204           F: T2037, A2038, L2039, E2041, P2043, T204         C2062, T2063, S2064, K2065, A2066, V2067, A: V32L1, V5514, S552; H2, S	ctur he D1 or k form [I 14, Y204' . I2114, I .94:L3; B H2, S555 H3, W60	D2 domain of SAG1 D2 domain of SAG1 DB: 1KZQ] and Fab-bound 7, N2050, Q2052, S2061, (2115, G2116 : T530/H1, D531:H1, H2, G556:H2, D557:H2, 2:H3, Y603:H3;			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Calculated Contacts	INIT     3D Strue       B     A       F     There is no drastic conformational change in th observed comparing the unbound homodimeri monomeric structure.       F: T2037, A2038, L2039, E2041, P2043, T204 C2062, T2063, S2064, K2065, A2066, V2067, A: Y32:L1, Y50:H2, S552:H2, T553:H2, Y554:H4, S553:H1, 1550:H2, S552:H2, T553:H2, Y554:H4, S554:H4, S559:H2, S599:H3, S600:H5, T601:H 330       Yiew 3D Structure]     [Yiew Curated Contact	cture he D1 or c form [1 .12114, 1 .94:L3; B H2, S555 H3, W60 ts XML fi	e of Complex D2 domain of SAG1 D8: 1KZQ] and Fab-bound 7, N2050, Q2052, S2061, (2115, G2116 1530:H1, D531:H1, H2, G556:H2, D557:H2, 2:H3, Y603:H3; e]			
3D Structure of Complex Complex PDB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Calculated Contacts	INIT         3D Strue           B         A           F         There is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204           G533:H1, 1550:L2, N53:L2, N92:L3, T631:L330           G533:H1, 2550:H2, S559:H3, S600:H3, T601:           1330           [View 3D Structure]           [View 3D Structure]           F: T2037, A2038, L2039, E2041, P2043, T204	cture he D1 or ic form [I 12114, J 94:L3; B H2, S555 H3, W60 ts XML fi 4, Y2047	e of Complex D2 domain of SAG1 PDB: 1KZQ] and Fab-bound 7, N2050, Q2052, S2061, (2115, G2116 : T530:H1, D531:H1, H2, G556:H2, D557:H2, 2:H3, Y603:H3; e] , N2050, Q2052, C2062, (2062, C2062, C2062,			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Calculated Contacts Epitope Residues:	INIT         3D Strue           B         3D Strue           F         There is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204         C2062, T2063, S2064, K2065, A2066, V2067, T353;H2, T553;H2, T554;H2, T554;H2, T555;H2, T555;H2, T555;H2, T555;H2, T555;H2, T555;H2, T555;H2, T5	cture he D1 or ic form [1 12114, 1 94:L3; B H2, S555 H3, W60 ts XML fi 4, Y2047 12114, F	e of Complex			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Calculated Contacts Epitope Residues: Antibody Residues Interacting with Antigen:	INIT         3D Strue           B         3D Strue           A         F           There is no drastic conformational change in to observed comparing the unbound homodimer monomeric structure.         F: T2037, A2038, L2039, E2041, P2043, T204 (2062, T2063, S2064, K2065, A2066, V2067, A358:H2, S559:H2, S559:H2, S59:H2, S59:	cture he D1 or c form [1 12114, 1 94:13; B H2, S555 H3, W60 ts XML fi 4, Y2047 I2114, k i31, Y53; 601, W6	Complex     C			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Calculated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antigen:	INIT         3D Strue           B         A           F         There is no drastic conformational change in th observed comparing the unbound homodimeri monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204 C2062, T2063, S2064, K2065, A2066, V2067, A: V32L1, Y5012, R5312, U3213, T39132, L G533H1, I550H2, S552H2, T553H2, V3541 A556H2, S559H2, S559H3, S600H3, T601: 1330           J         Image: Comparing the unbound homodimeri monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204 T2063, S2064, K2065, A2066, V2067, T2068, AS50, V2077, A2038, L2039, E2041, P2043, T204 T2063, S2064, K2065, A2066, V2067, T2068, A: S30, V32, Y50, N92, T39, L94; B: T530, D5 S555, D557, A558, S559, Y560, S599, S600, T 97.3	cture he D1 or c form [I 12114, I 94:13; B H2, S555 H3, W60 ts XML fi 4, Y2047 I2114, F i31, Y53; 601, W6	Complex     Complex     D2 domain of SAG1     D3: 1K2Q] and Fab-bound     N2050, Q2052, S2061,     C2115, G2116     T530:H1, D551:H1,     H2, G556:H2, D557:H2,     2:H3, V603:H3;     e]     N2050, Q2052, C2062,     2115, G2116     , G533, S552, T553, Y554,     02, Y603			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Calculated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antigen: Contact Area for Antigen: Contact Area for Antigen: Contact Area for Antigen:	INIT         3D Struct           B         A           F         There is no drastic conformational change in th observed comparing the unbound homodimeri monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204           C2062, T2063, S2064, K2065, A2066, V2067, A: Y32:L1, Y50:L2, R53:L2, N52:L3, T53:L4, Y54:L4, A555:H2, S559:H2, S559:H2, S559:H3, S60:H3, T601: 1330           View 3D Structure]         [View Curated Contact           F: T2037, A2038, L2039, E2041, P2043, T204           F: T2037, A2038, L2039, E2041, P2043, T204           F: S555, D557, A558, S559, Y560, S599, S600, T 977.3           942.1           Dimention	cture he D1 or ic form [I 14, Y204; 12114, I 94;L3; B H2, S555 H3, W60 ts XML fi 14, Y2047 12114, K 331, Y532 601, W6	Complex     Complex     D2 domain of SAG1     DB: 1KZQ] and Fab-bound     N2050, 02052, 52061,     C115, 62116     T530:H1,     H42, G556:H2,     D55:H2,     H42, G556:H2,     S51:H1,     Y603:H3; e]     N2050, 02052, C2062,     2115, 62116     , 0533, S552, T553, Y554,     02, Y603			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antibody Chain 2 PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Contact Area for Antibody: Contact Area for Antibody: View 3D Structure: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Antibody Residues Interacting	INIT         3D Struct           B         3D Struct           F         There is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.           F: Z0237, A2038, L2039, E2041, P2043, T204           Structure.           F: T2037, A2038, L2039, E2041, P2043, T204           Stassicity, Structure.           F: T2037, A2038, L2039, E2041, P2043, T204           Stassicity, Structure.           [Wiew 3D Structure]           [View 3D Structure]	cture he D1 or [] (1, 2204) (1, 2214, 1, (1, 2214, 1, (1, 2214, 1, (1, 2214, 1, (1, 2214, 1,	e of Complex D2 domain of SAG1 D2: 1KZQ] and Fab-bound 7, N2050, Q2052, S2061, (2115, G2116 : T530:H1, D551:H1, H2, G556:H2, D557:H2, 2:H3, Y603:H3; e] , N2050, Q2052, C2062, 2115, G2116 , G533, S552, T553, Y554, 02, Y603 file]			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antibody Chain 2 PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Contact Area for Antibody: Contact Area for Antibody: View 3D Structure: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Assay Reference Details	INIT         3D Struct           B         3D Struct           F         There is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.         File           F 2037. A2038. L2039. E2041. P2043. T204. C2062. T2063. S2064. K2065. A2066. V2067. A355142. J953142. J953143. T204. S355142. J553142. J953143. J550145. J953143. J550145. J953143. J550145. J953145. J553145.	cture he D1 or [1] (4, Y204') H2, Y204', Y204', H3, W60 H3, W60H3, W60 H3, W60 H3, W60H3, W60 H3, W60 H3, W60H3, W60 H3, W60H3, W6	Complex     Complex     D2 domain of SAG1     DB: 1KZQ] and Fab-bound     N2050, 02052, 52061,     C2115, G2116     T530:H1,     H42, G556:H2,     D55:H2,     H42, G556:H2,     D55:H2,     C2052, C2062,     C115, G2116     (, G533, H552, T553, Y554,     02, Y603     file]     rymmetric unit; therefore,     served in the Fab-SAG1			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Courated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Contact Area for Antibody: Contact Area for Antibody: View 3D Structure: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Assay Reference Details	INIT         3D Struct           B         3D Struct           F         These is no drastic conformational change in the observed comparing the unbound homodimeri monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204         C2062, T2063, S2064, K2065, A2066, V2067, A: V32L1, V5314, S2012, R5312, BV213, F3312, L           F: T2037, A2038, L2039, E2041, P2043, T204         C3062, T2063, S2064, K2065, A2066, V2067, T2068, S550, H557, A5012, R5532, H553, H553, F553, F553	cture he D1 or c form [1 14, Y204 <sup>1</sup> , 12114, 1, 94:L3; B H3, W60 ts XML fit 12114, 1, 12114, 1,	Complex     C			
3D Structure of Complex Complex POB ID: Antibody Chain 1 PDB Chain: Antibody Chain 2 PDB Chain: Antigen PDB Chain: Comments: Curated Contacts Epitope Residues: Antibody Residues Interacting with Antigen: Contact Area for Antibody: View 3D Structure: Contact Area for Antigen: Contact Area for Antigen: Co	INIT         3D Strue           B         A           F         There is no drastic conformational change in to observed comparing the unbound homodimer monomeric structure.           F: T2037, A2038, L2039, E2041, P2043, T204         C2062, T2063, S2064, K2065, A2066, V2067, A: 9206, V2067, A: 9206, V2067, A: 9206, V2067, A: 9206, V2067, V2068, A558; H2, S559, H2, S559, H2, S559, H3, S600H3, T601:           Total         Statuture         [View Guested Contact           F: T2037, A2038, L2039, E2041, P2043, T204         T2063, S2064, K2065, A2066, V2067, T2068, A: S30, Y32, Y50, N92, T33, L94; B: T530, D5           S555, D557, A556, S559, S500, T907,3         S42.1         [View Gacutated Contact           The structure contains two identical complexes only one (chains A, B, F) was curated. SA61 moromer in the structure is the sourcure is the source is the	Cturn he D1 or from [1 12114, 1 2412, 5555 H3, W60 ts XML fit 31, Y53/3 G601, W6 acts XML fit 31, Y53/3 G601, W6 acts XML fit acts XML	Complex     Complex     D2 domain of SAG1     D2: domain of SAG1     D2: JK2Q] and Fab-bound     V2050, Q2052, S2061,     (2115, G2116     TS30:H1, D531:H1,     H2, G556:H2, D557:H2,     2:H3, V0505, Q2052, C2062,     2115, G2116     (533, S552, T553, Y554,     02, Y603     file]      vymmetric unit; therefore,     served in the Fab-SAG1 of Fab. Epitope is located at     an that is not involved in to for source antigen.			

Figure 3. Example of the B-cell response page. Detailed information is provided for the epitope (blue box) and B-cell response, including data on the immunization (purple box), assay (teal box), antibody (yellow box), antigen used in the X-ray crystallography experiment (green box) and the 3D structure of antigen-antibody complex (red box). The antigen-antibody interactions were curated from (12) and calculated from the coordinates provided in the PDB file [PDB: 1YNT]. The results of clicking on the links 'View 3D Structure' and 'View Contacts XML file' (red arrows) are shown in Figure 4.

Table 1. IEDB-3D content

	Epitope type	Total number of references <sup>a</sup>	Curated references <sup>b</sup>
IEDB category			
Infectious diseases	Peptidic	131	131
	Non-peptidic	11	4
Allergy	Peptidic	8	8
	Non-peptidic	5	3
Autoimmunity	Peptidic	39	30
2	Non-peptidic	11	11
Cancer	Peptidic	40	16
	Non-peptidic	8	5
Transplant	Peptidic	22	18
*	Non-peptidic	2	1
Other	Peptidic	241	130
	Non-peptidic	82	2
HIV	* *	55	4
Total		655	365
Structural category			
B-cell epitopes	Peptidic	311	168
	Non-peptidic	108	20
T-cell epitopes	Peptidic	217	177
* *	Non-peptidic	19	0
Total		655	365

<sup>a</sup>PDB and PubMed search as of January 7, 2010.

<sup>b</sup>Curation status as of January 9, 2010.

context of immunological assays. Upon curation, both a PDB structure and journal article, which is a primary reference of the structure are considered as sources of information for curation.

Similar to other assays, for 3D structures, data on epitope molecular structure, epitope source, immunization that led to recognition of the epitope, assay, antigen, antibody, TCR or MHC are curated according to IEDB general procedure (1). An example of a B-cell response is given in Figure 3. The difference of IEDB-3D data compared to regular IEDB curations is the information captured under the '3D Structure of Complex' heading.

In general, studies often do not provide the exact epitope structure recognized by an antibody/B-cell receptor, especially in the case of discontinuous B-cell epitopes. Therefore, discontinuous epitopes are often curated as partial epitopes, with only a few key residues captured. In cases where the epitope is curated from the structure this can be avoided; if the authors of the paper stated that only residues for part of the epitope were provided, we calculate the epitope residues from the PDB structure as the antigen residues separated from the antibody by 4Å atomic distance (T-cell epitopes are curated as whole peptides or nonpeptidic structures). The mapping between the numbers of the epitope residues provided in the paper and the external database is done manually, using the mapping between the PDB and UniProt numberings provided on the PDB website. For example, the epitope shown in Figure 3 had different residue numbers in the paper ('Reference Region' field at the top in Figure 3) compared to the numbering that matches the source antigen sequence stored in the external database ('Discontinuous Residues' field).

Information on the 3D structure is provided in the section '3D Structure of Complex'. There are three types of tables with different fields for B-cell, T-cell and MHC-binding, respectively. The fields curated for B-cell responses, or structures of antibody-antigen complexes, are shown in Figure 3. The following structural information is curated: PDB ID, antibody and antigen chain IDs, epitope and antibody residues in PDB numbering, as well as CDR loops for the antibody residues and contact surface areas for the antibody and antigen if they are provided in the paper.

Interacting residues and contact areas are also calculated. The former are defined based on 4-Å atomic distance, according to the definition from (10), the latter are calculated using the NACCESS program (11). Pairwise atomic contacts are provided in the XML files available via hyperlinks (Figure 3) 'View Curated Contacts XML file' and 'View Calculated Contacts XML file'. The first link is provided only if the contacts were specified in the journal article and curated, the latter link allows the user to view and download the file providing calculated antibody-antigen (antigen-MHC, or antigen-TCR and MHC-TCR) inter-molecular contacts (Figure 4C). The following types of contacts are calculated: hydrogen bonds, salt bridges, van der Waals, hydrophobic and 4-Å interactions (interactions are defined in the EpitopeViewer tutorial at http://spdc .sdsc.edu/iedb/epitopeViewer/EpitopeViewerTutorial v2 .0.htm; this link is also available inside the EpitopeViewer application). The calculation is done during curation, using a php program that takes as an input curated data and PDB file and outputs two XML files with the interacting residues, contact areas and pairwise contacts. These files are also used as input for EpitopeViewer (Figure 4). The curators use the php tool and EpitopeViewer as part of the curation process and to check for errors.

EpitopeViewer allows the user to visualize, render and analyze the structure and save structural and contact views as high-quality pictures for publication (9). Figur 4 shows an example of how the EpitopeViewer can be used to analyze specific inter-molecular contacts. The residue Arg53 of the antibody light chain (Figure 3) was curated as part of the paratope; however, it can be seen that it is located relatively far from the nearest epitope atom (5.4 Å) (Figure 4A). At the same time, another light chain residue, Ser30, which was not curated as part of the paratope, contacts the epitope residues through van der Waals and 4-Å interactions (Figure 4B).

Additional detailed information on the 3D structure is captured in free text form in the 'Assay Comments' field (bottom of Figure 3). This can, for example, provide information on antigen and receptor conformational changes and comparison with other relevant structures if this information is mentioned in the paper. Also, the user can get further details on the PDB website via the PDB ID hyperlink in the 'Complex PDB ID' field.

#### **QUERYING IEDB-3D**

Since IEDB-3D is fully integrated within IEDB, structural data, both curated and calculated, and all accompanying information can be queried using the multiple search capabilities implemented as part of the IEDB web site and described in this section.

On the IEDB home page (Figure 1), epitopes can be searched by epitope sequence, epitope source organism, source antigen and the immune recognition context, including the type of response (B-cell, T-cell or MHCbinding), host organism and MHC allele. If epitopes curated from 3D structures satisfy the search criteria, they will be retrieved together with all other epitopes curated in IEDB.

Using the keyword search option on the IEDB hom page (the box is located in the right top corner of the page, Figure 1, red arrow), data are retrieved by keywords, PDB ID or SwissProt/GenBank ID; the keyword search runs throughout all fields in the database. In addition, the user can explicitly qualify what identifiers to search for when using the option 'Search by Identifier' from the 'Search' dropdown menu on the home page (Figure 1, yellow box). This type of search can be done, using IEDB internal identifiers, PDB ID, PubMed ID and CHEBI ID.

The 'Search' dropdown menu on the home page allows advanced search of any specified field in the database, including the fields related to 3D structure. For example, on the 'B Cell Search' webpage, the search fields related to 3D structure are made visible by clicking the '+' sign next to '3D Structure of Complex' within the 'B Cell Assay' subsection. Figure 5A shows how to search on the 'B Cell Search' page for the structures of antigens in complex with antibodies obtained via *in vivo* administration/immunization with *Toxoplasma gondii* and containing tryptophan (W) in the paratope (residues in the antibody interacting with the antigen). The result of this query is a single epitope that was curated from the structure with the PDB ID 1YNT (Figure 3).

Alternatively, epitopes for which 3D structural information is available can also be searched for by specifying a particular type of assay. For example, the epitope shown in Figure 3 can be found by specifying the assay type 'X-ray crystallography' using the 'Assay Finder' on the 'B Cell Search' form. Additionally, the resolution for the 3D structure can be specified.

In addition to the keyword, simple and advanced searches, epitopes curated from 3D structures can be accessed via the 'Browse by 3D Structure' page provided in the 'Browse' dropdown menu on the IEDB home page (Figure 1, red box). By expanding the 'B Cell', 'T Cell' and 'MHC Binding' trees the epitopes can be browsed by the organism that is the source of the antibody, T cell, and MHC molecule, respectively. This page also provides an up to date overview on the number of distinct epitope structures curated in IEDB-3D.

When clicking on an epitope ID (Figure 5C), all assays in which that epitope is characterized are returned.



Figure 4. Visualization of 3D structures and contacts. EpitopeViewer screen-shots for the curated (A) and calculated (B) data for the epitope shown in Figure 3. Epitope residues are in blue. (A) The residue Arg53 of the antibody light chain, which was curated as part of the paratope, is shown in an all-atom presentation and colored in yellow; the red line shows the distance between this residue and the nearest epitope atom. Another light chain residue Ser30, which was not curated as part of the paratope, is shown in magenta. This residue is highlighted in yellow in (B) and its calculated contacts with the epitope are shown as broken lines (cyan, van der Waals interactions; white, 4Å interactions). (C) The fragment of the XML file providing calculated contacts between the antibody and antigen [PDB: 1YNT].

For example, the query shown in Figure 5A returns one epitope (Figure 5B and C), for which two B-cell response assays were curated (Figure 5D), one of which describes the 3D structure (Figure 3) and another, the immunological assay (not shown).

(all fields are present except pairwise contacts) or compact form (Figure 5C and D).

The query results obtained using either simple or advanced search options can be exported as a Microsoft Excel formatted file and downloaded in either full

#### DATA ACCESS

The entire database, together with structural data both curated and calculated (including pairwise contacts), can be downloaded in XML and MySQL formats.

			B Cell Sea	arch							
⊞ Reference											
⊞-Epitope	A										
- Immunization											
Host Organism				Organism f	Finder						
<b>⊞</b> Host Details						Search Parameters:					
→ Ist In Vivo Process <sup>○</sup> may or	may not 💿	MUST 🔍 MUS	ST NOT be presen	t in search i	results.	<ul> <li>1st In Vivo Process</li> </ul>	is present				
🖃 <b>1st Immunogen</b> $^{\circ}$ may or ma	y not 💿 ML	ізт 🗢 мизт і	NOT be present in	search resu	ılts.	<ul> <li>1st Immunogen is</li> </ul>	present				
Epitope Relation Source	e Organism		~			In Vivo1 Immunog     In Vivo1 Immunog	en Epitope Relation <b>equ</b> en Object Subtype <b>equ</b>	als 'Source Org als 'Organism'	janism		
Organism Toxop	asma gondii;		Organism Finder			In Vivo1 Immunog	en Object Organism Na	me contains 'T	oxoplasma gondii'		
🕑 Immunogen Details							i Antibody Antigen Resi	dues contains	vv		
						Revise Search					
<b>⊡ 2nd Immunogen</b> <sup>®</sup> may or m	ay not 🔍 M	UST 🔍 MUST	NOT be present in	n search res	ults.	Epitopes	Positive* Ne	gative**	All		
⊞ Immunization Comments						Peptidic	Peptidic 1 0 1				
	nay not 🔍 M	IUST O MUST	T NOT be present	in search re	sults.	Non-Peptidic	0	0	0		
E-B Cell Assay											
Qualitative Measurement	â					Assays	Positive	Negative			
Assay Type	* <u>. M</u>			Assay Finder B Cell Response			ponse 1				
Measurement Details						MHC Ligand E	lution 0	0	0		
						МНС Ві	nding 0	0	0		
= 3D Structure of Complex											
Complex PDB ID											
Antibody Chain 1 PDB Chain				INE EP	ITOP	DATABASE					
Antibody Chain 2 PDB Chain			AND	ANALY	SIS RE	SOURCE					
Viewer Flag O d	nv 🧿 Yes	O No	Home			Browco	aarab	Toolo	С		
	1, 100	110	ноше			Blowse 3		10015	subt		
C							Epit	ope			
Comments											
			1 item(s) found	I, displayin	g 1 to 1	(Click the column headers to adju	st the sorting)				
E-Curated Contacts			Export all result	s: 🗷 (full	)						
E-Calculated Contacts	Г		Epitope ID	Structu	ire	Sou	rce Antigen		Source Organism		
			77401	T85, A8	6,L87,E	89, P91, T92, Y95, N98,	r a rface antigen n 20 pr		Tovoplasma gondi		
Epitop	e Residues		77491	Q100, S	109	Inajo	i surface anuger pso pre				
			1 item(s) found	l, displayin	g 1 to 1						
		/	Export all result	s: 🗷 (full	)						
Antibody Residues Interacting w	th Antigen										
Anabody Residies Including w	ar Anagen										
IMMUN	E EPITOR	PE DATABA	SE								
ANDAN	ALYSIS R	ESOURCE						Keyword Sea	rch Search		
Home		Browse	s	earch		Tools Su	pport Mo	re IEDB			
					Epit	tope Information					
Search Clear					2,510						
Distinct Epitope											
			Epitope ID	77491	1.07 500	DO1 TO2 YOE NOR 0100 5100 61	110				
		Discor	ntinuous Residues	T111, S1	, LO7, EO9 12, K113,	A114, V115, I162, K163, G164					
			Source Antigen	Major sur	face antige ma gondii	en p30 precursor		ן ט ן			
			oodi ee organism	Textopido	na gonai						
			7								
Source (1) Refer	ence (1) B	Cell Assay (2)									
2 item(s) found, dis	laying 1 to 2	(Click the colur	mn headers to adjus	t the sorting	g)						
Export all results: 2	(compact	ful)			1		1				
B Cell ID 🕇 Ref	erence Epite	ope		Host	Immuni	zation	Assay Antigen	Antigen Epitope Relation	Assay Description		
Mar	T85.	A86.187.E89.	P91. T92. Y95.						Ab Competition of		
1245875 Grai	e; N98,	Q100, S109		Mus	Administr	ation in vivo with Toxoplasma gondii	Toxoplasma gondii	Source Organism	Ab Binding Detection of Ab/Ag		
J Mc   200	5 Majo	plasma gondii	ripso precursor	musculus	(Source (	organism)	roxopiasma gondii		binding Positive		
	тог	486 1 87 590	P91 T92 VOE				Major surface antigen p30		X-Ray		
1246572 Mar View 3D Grai	e; N98,	Q100, S109	, F 91, 192, 195,	Mus	Administr	ation in vivo with Toxoplasma gondii	precursor Mator surface antigen p30	Source Antigen	Crystallography Characterization of		
Structure 200	Biol Major Toxe	r surface antige plasma qondii	n p30 precursor	musculus	(Source (	Organism)	precursor	oouroo Anugan	Ab binding		
							i oxopiasma gondii		Positive		
2 item(s) found, dis Export all requires the	laying 1 to 2	f.ID									
Export ai results:	a (compace)	rcal)									

**Figure 5.** Example of the IEDB-3D query. (A) The advanced B-cell search page showing the query for 3D structures of antigens in complex with antibodies obtained via *in vivo* administration/immunization with *Toxoplasma gondii* and containing tryptophan (W) in the paratope calculated as the antibody residues interacting with the antigen at <4Å atomic distance. The result of clicking on the button 'Search' (red box) is shown in B. (B) The result of the search shown in A. The result of clicking on the number of epitopes (red box) is shown in C. (C) The list of epitopes as the result of the search shown in A. The result of clicking on the epitope ID (red box) is shown in (D). (D) The list of assays for the epitope found using the query shown in A. The button 'View 3D Structure' (red arrow) launches the EpitopeViewer for curated data (Figure 4A). The result of clicking on the assay ID (red box) is shown in Figure 3.

#### FUNDING

Funding for open access charge: National Institute of Health/National Institute of Allergy and Infectious Diseases (contract number HHSN2662004000 0 6C).

Conflict of interest statement. None declared.

#### REFERENCES

- Vita, R., Zarebski, L., Greenbaum, J.A., Emami, H., Hoof, I., Salimi, N., Damle, R., Sette, A. and Peters, B. (2010) The immune epitope database 2.0. *Nucleic Acids Res.*, 38, D854–D862.
- Korber, B.T.M., Brander, C., Haynes, B.F., Koup, R., Moore, J.P., Walker, B.D. and Watkins, D.I. (2007) HIV Molecular Immunology. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico, LA-UR 07-4752.
- Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, 28, 235–242.
- Ehrenmann, F., Kaas, Q. and Lefranc, M.P. (2010) IMGT/ 3Dstructure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF. *Nucleic Acids Res.*, **38**, D301–D307.
- 5. Tong,J.C., Kong,L., Tan,T.W. and Ranganathan,S. (2006) MPID-T: database for sequence-structure-function information on

T-cell receptor/peptide/MHC interactions. *Appl. Bioinformatics*, **5**, 111–114.

- Schlessinger, A., Ofran, Y., Yachdav, G. and Rost, B. (2006) Epitome: database of structure-inferred antigenic epitopes. *Nucleic Acids Res.*, 34, D777–D780.
- Tong, J.C., Song, C.M., Tan, P.T., Ren, E.C. and Sinha, A.A. (2008) BEID: Database for sequence-structure-function information on antigen-antibody interactions. *Bioinformation*, 3, 58–60.
- 8. Huang, J. and Honda, W. (2006) CED: a conformational epitope database. *BMC Immunol.*, **7**, 7.
- 9. Beaver, J.E., Bourne, P.E. and Ponomarenko, J.V. (2007) EpitopeViewer: a Java application for the visualization and analysis of immune epitopes in the Immune Epitope Database and Analysis Resource (IEDB). *Immunome Res.*, **3**, 3.
- Greenbaum, J.A., Andersen, P.H., Blythe, M., Bui, H.H., Cachau, R.E., Crowe, J., Davies, M., Kolaskar, A.S., Lund, O., Morrison, S. *et al.* (2007) Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *J. Mol. Recognit.*, 20, 75–82.
- 11. Hubbard,S.J. and Thornton,J.M. (1993) NACCESS. *Department* of *Biochemistry and Molecular Biology*, University College, London.
- 12. Graille, M., Stura, E.A., Bossus, M., Muller, B.H., Letourneur, O., Battail-Poirot, N., Sibai, G., Gauthier, M., Rolland, D., Le Du, M.H. *et al.* (2005) Crystal structure of the complex between the monomeric form of Toxoplasma gondii surface antigen 1 (SAG1) and a monoclonal antibody that mimics the human immune response. J. Mol. Biol., 354, 447–458.