



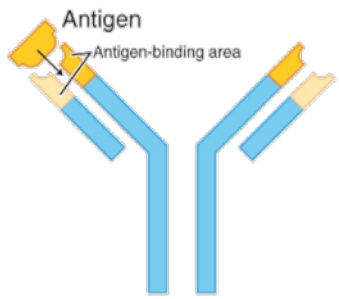
B cell epitope prediction

Swapnil Mahajan
swapnil@lji.org

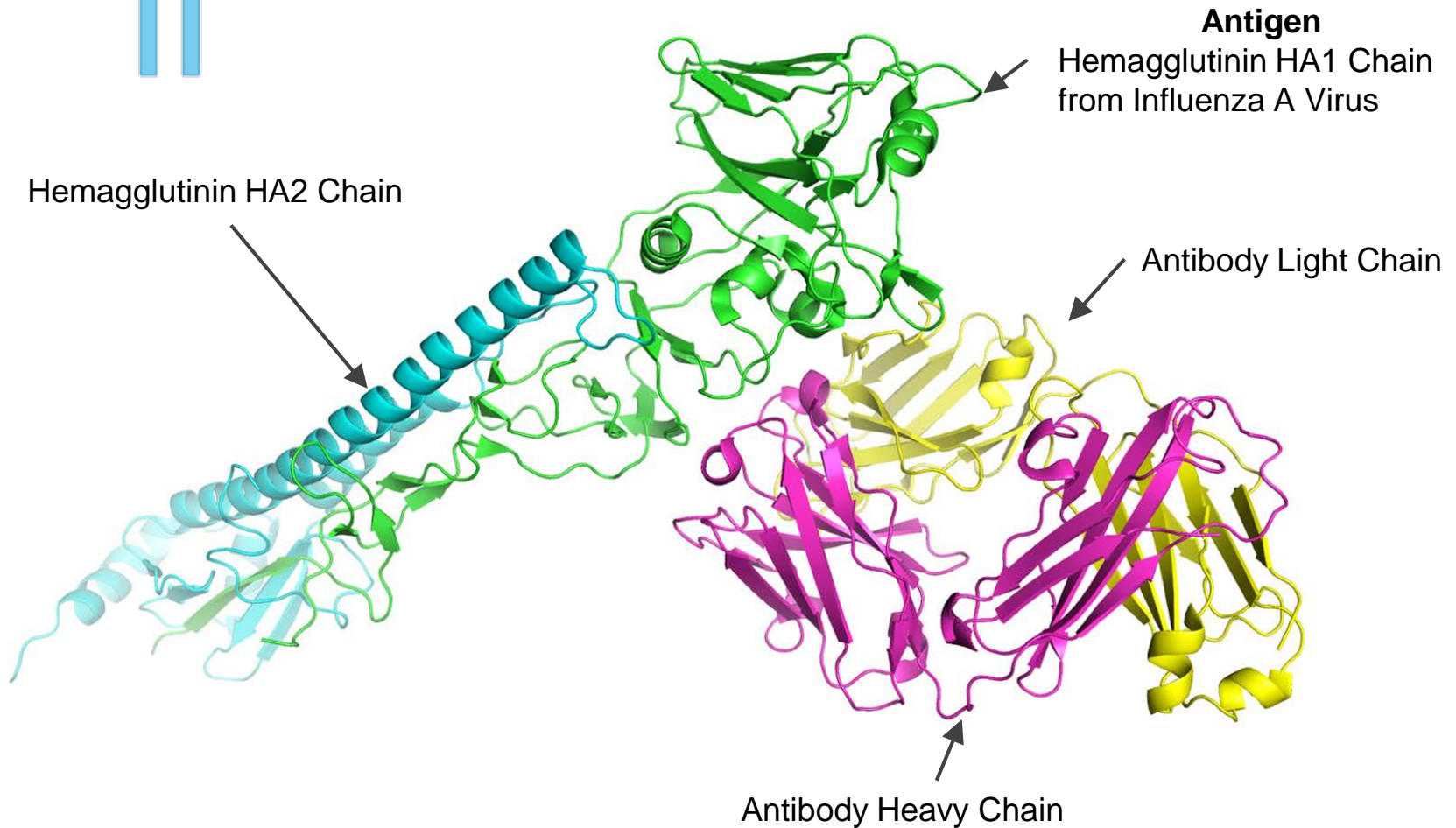
IEDB Workshop, 2018

Outline

- **B cell epitopes:** Discontinuous And linear epitopes
- **Schema** of B cell epitope prediction tools on IEDB
- **Sequence-based epitope prediction methods:**
 - Linear epitope prediction methods
- **3D Structure-based epitope prediction methods:**
 - Discontinuous epitope prediction methods
- **Computational antibody design**
 - Antigen and Antibody structure modelling
 - Antibody-protein docking

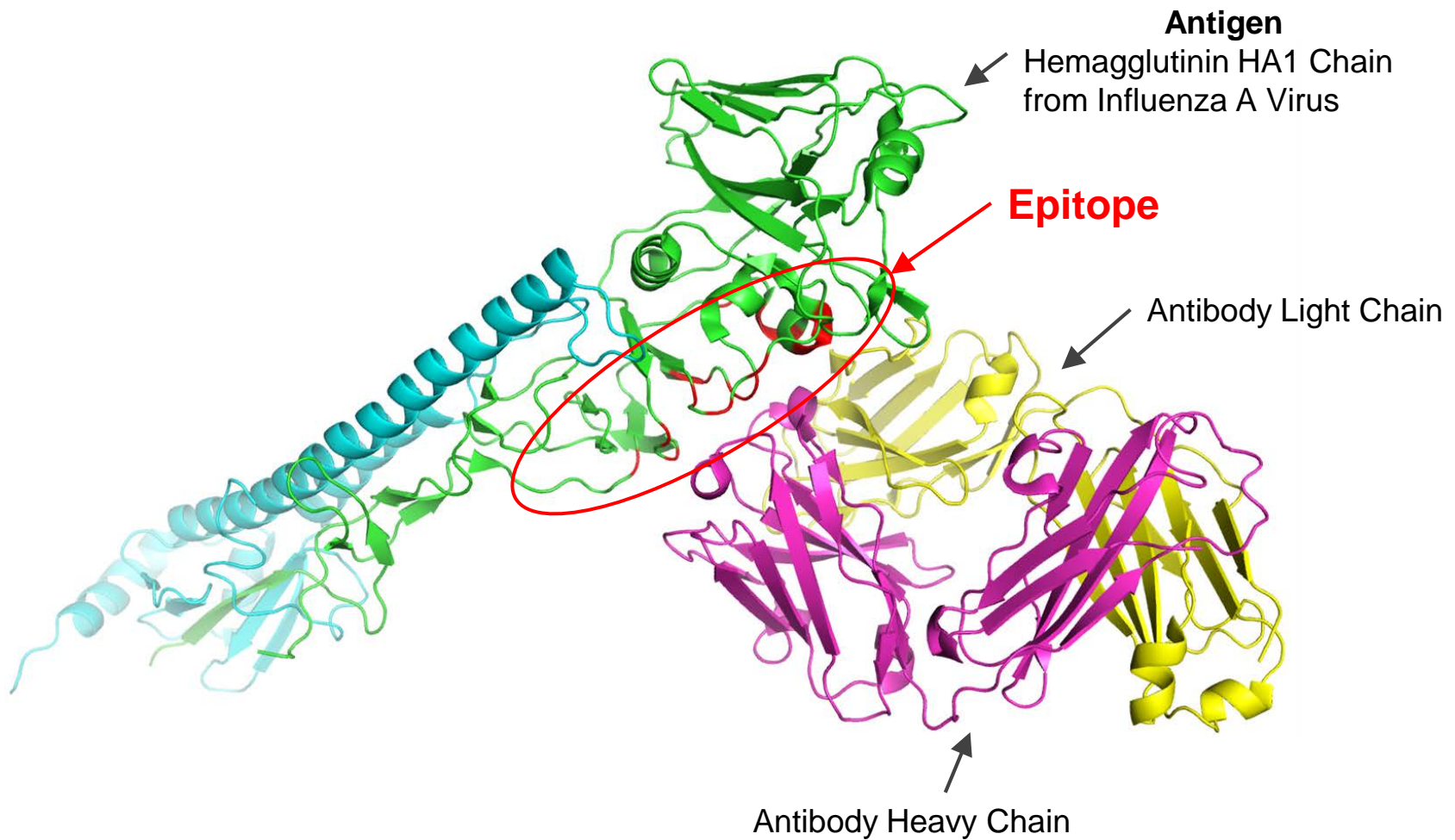


B cell epitopes



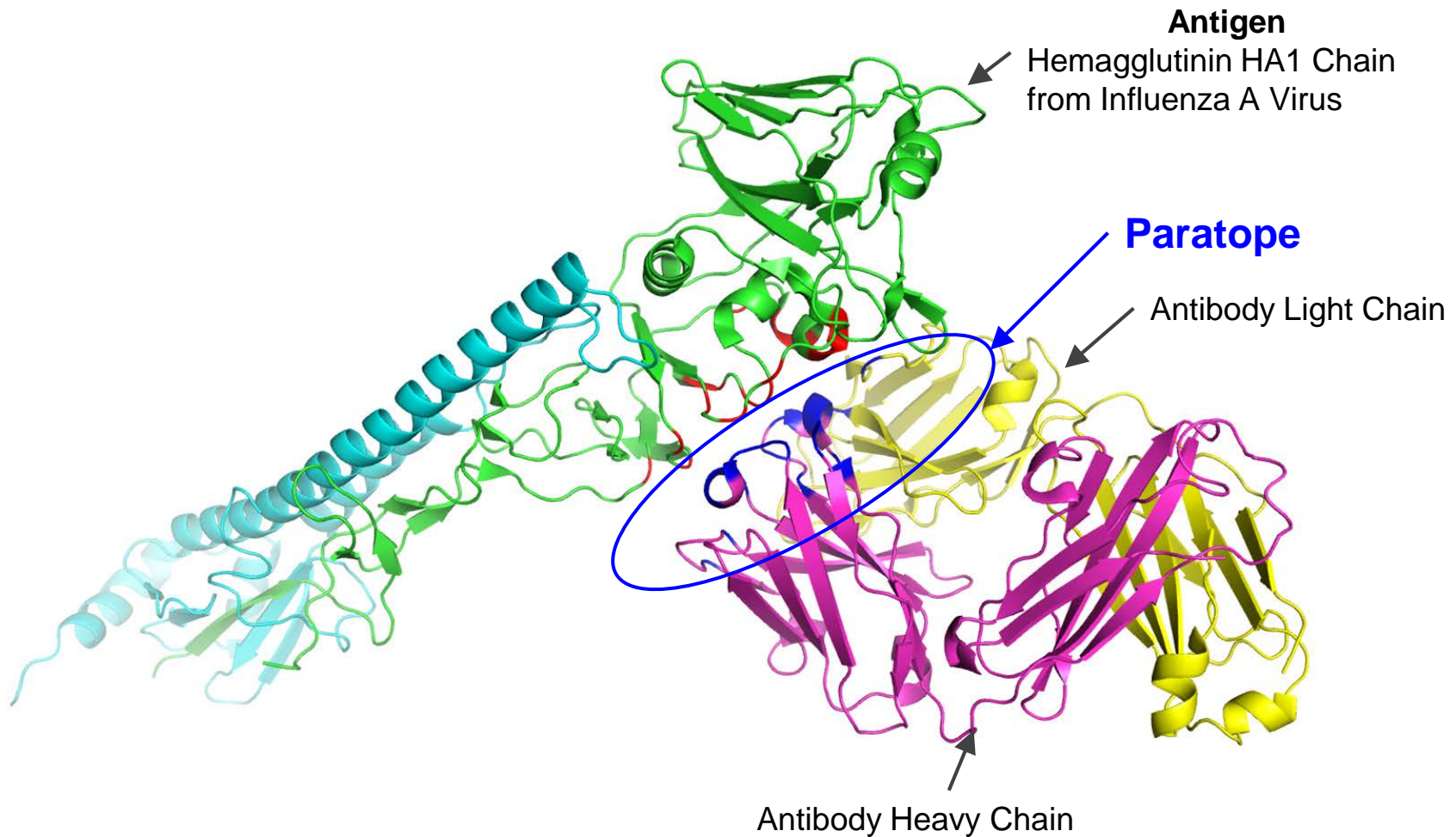
PDB ID: 1EO8

Discontinuous Epitope



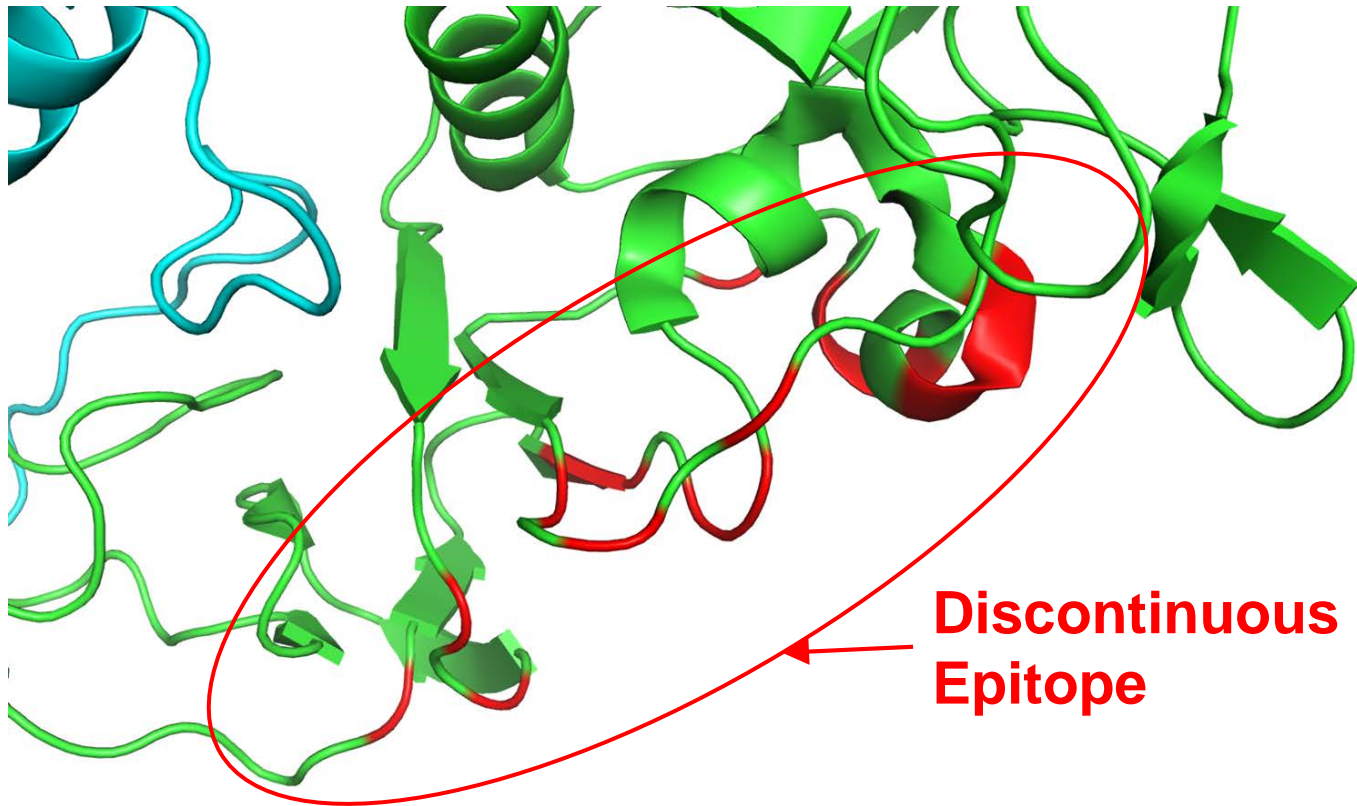
PDB ID: 1EO8

Discontinuous Epitope



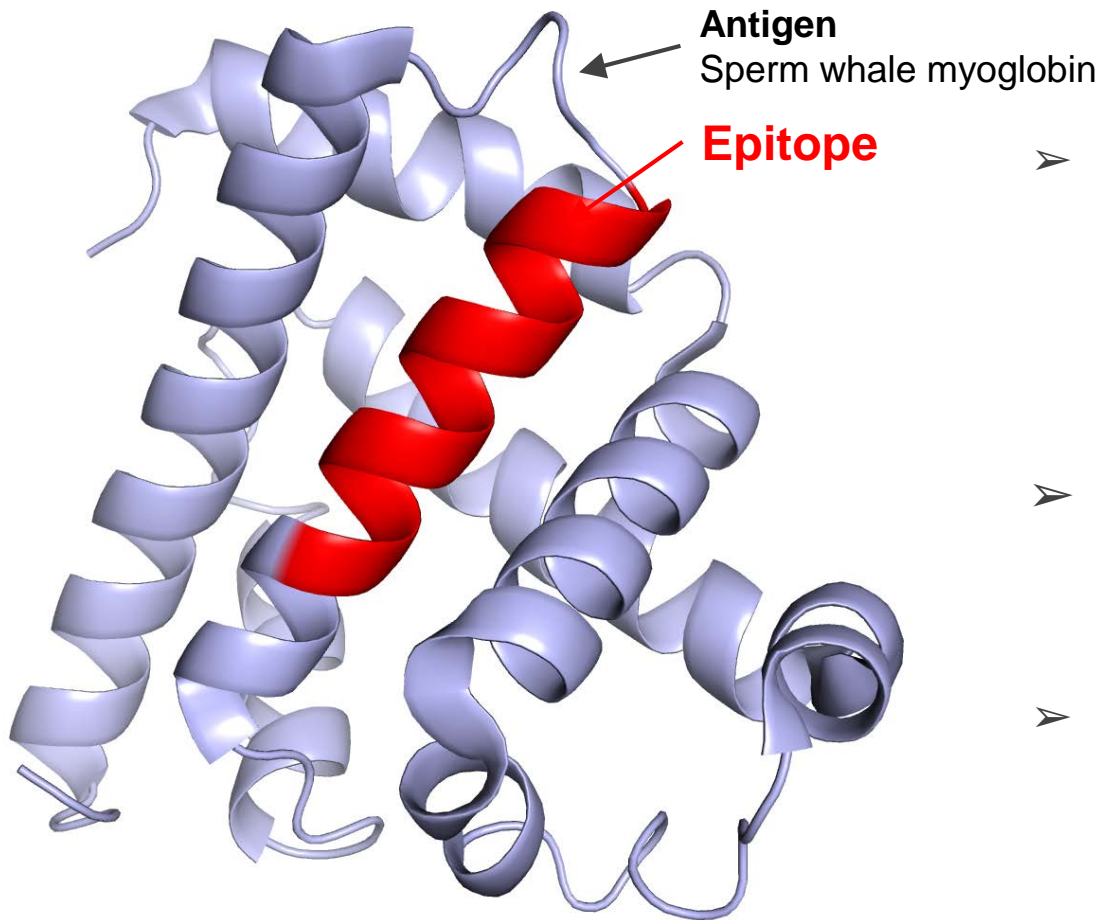
PDB ID: 1EO8

Discontinuous Epitope



PDB ID: 1EO8

B cell epitopes



PDB ID: 1VXG

- Protein antigens usually contain both linear (aka sequential or continuous) epitopes that may still bind to the antibody even when that protein is denatured
- and discontinuous (aka non-sequential or conformational) epitopes
- **More than 90% of the B cell epitopes are estimated to be discontinuous*.**

*Barlow et al, Nature. 1986.

*Van Regenmortel, Methods. 1996.

Schema of B cell epitope prediction tools

Walk through



Schema of B cell epitope prediction tools

The screenshot displays the IEDB website interface. At the top, there is a navigation bar with 'Home', 'Specialized Searches', and 'Analysis Resource'. The 'Analysis Resource' dropdown menu is open, highlighting 'B Cell Epitope Prediction'. Below the navigation, the 'START YOUR SEARCH HERE' section contains several search filters: Epitope, Assay, Antigen, MHC Restriction, Host, and Disease. The 'Assay' filter is checked for 'B Cell Assays'. On the right side, a 'B Cell Epitope Prediction' panel is highlighted with a red box, listing tools like 'Antigen Sequence Properties', 'Discotope', and 'ElliPro'. A '2018 USER WORKSHOP' banner is visible on the left, and a 'Summary Metrics' table is at the bottom left.

Summary Metrics

Peptidic Epitopes	520,928
Non-Peptidic Epitopes	2,685
T Cell Assays	340,887
B Cell Assays	457,157
MHC Ligand Assays	1,059,403
Epitope Source Organisms	3,666
Restricting MHC Alleles	773
References	19,688

Schema of B cell epitope prediction tools

IEDB Analysis Resource

[Overview](#) [T Cell Tools](#) [B Cell Tools](#) [Analysis Tools](#) [Tools-API](#) [Download](#) [Datasets](#) [Contribute Tools](#) [References](#)

B Cell Epitope Prediction Tools

B Cell Epitope Prediction

[Prediction of linear epitopes from protein sequence](#)

A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs.

[Discotope - Prediction of epitopes from protein structure](#)

This method incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential along the length of a protein sequence.

[ElliPro - Epitope prediction based upon structural protrusion](#)

This method predicts epitopes based upon solvent-accessibility and flexibility.

[Methods for modeling and docking of antibody and protein 3D structures](#)

This page provides information on available methods for modeling and docking of antibody and protein 3D structures.

Structure Tools



[LYRA \(Lymphocyte Receptor Automated Modelling\):](#)

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.



: Tools under AR Labs which are experimental and are not quite ready for production yet. They are intended for further research, updates and testing.

Schema of B cell epitope prediction tools

The screenshot shows the 'Antibody Epitope Prediction' web tool interface. At the top, a navigation bar contains links for Home, Help, Example, Reference, Download, and Contact. Below this is the main title 'Antibody Epitope Prediction' and a section titled 'Specify Input'. This section includes a text input field for a Swiss-Prot ID (with an example: P02185) and a larger text area for entering a protein sequence in plain format (up to 50,000 residues). Below the input fields is a section titled 'Choose a method:' which lists seven prediction methods, each with a radio button. The methods are: Bepipred Linear Epitope Prediction, Bepipred Linear Epitope Prediction 2.0, Chou & Fasman Beta-Turn Prediction, Emini Surface Accessibility Prediction, Karplus & Schulz Flexibility Prediction, Kolaskar & Tongaonkar Antigenicity, and Parker Hydrophilicity Prediction. At the bottom right of the form are 'Submit' and 'Reset' buttons. Annotations include a red box around the navigation bar, a bracket on the right side of the 'Specify Input' section labeled 'Input fields', and a red box around the 'Choose a method:' list. An arrow points from the 'Input fields' label to the 'Home' link, and another arrow points from the 'Choose a method:' list to the text 'Do you know these methods?'.

Home Help Example Reference Download Contact

Antibody Epitope Prediction

Specify Input

Enter a Swiss-Prot ID (example: P02185)

Or enter a protein sequence in plain format (50000 residues maximum):

Choose a method:

- [Bepipred Linear Epitope Prediction](#)
- [Bepipred Linear Epitope Prediction 2.0](#)
- [Chou & Fasman Beta-Turn Prediction](#)
- [Emini Surface Accessibility Prediction](#)
- [Karplus & Schulz Flexibility Prediction](#)
- [Kolaskar & Tongaonkar Antigenicity](#)
- [Parker Hydrophilicity Prediction](#)

Submit Reset

Input fields

Do you know these methods?

Click “Help”, if you want to know more about a prediction method

[Home](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

Antibody Epitope Prediction - Tutorial

1. Methods for predicting continuous antibody epitope from protein sequences

General basis: Parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity and antigenic propensity of polypeptides chains allow the position of continuous epitopes to be predicted from certain features of the protein sequence. All prediction calculations are based on propensity scale on the basis of their relative propensity to possess the property described by the scale.

General method: When computing the score for a given residue i , the amino acids in an interval of the chosen length, centered around residue i , are considered. To compute the score for residue i . Unless specified, the score for residue i is the average of the scale values for these amino acids (see table 1 for specific methods to be antigenic).

Interpretation of output graphs and tables: On the graphs, the Y-axis depicts for each residue the correspondent score (averaged in the specified window around the residue positions in the sequence). The tables provide values of calculated scores for each residue. The larger score for the residues might be interpreted as the most antigenic (however, the presented methods do not predict the epitopes per se, either linear or discontinuous, -- they might only guide the researchers to further

Table 1. Implemented methods

Method

Chou and Fasman beta turn prediction

- Reference: [Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol Biol. 1974;37:1-61.](#)
- Description: The rationale for predicting turns to predict antibody epitopes is based on the paper by [Pellequer et al, Immunology Letters, 36 \(1993\) 83-99.](#) It uses the Chou and Fasman scale which is commonly used to predict beta turns as described in the reference link above.

Scale:

A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
0.66	1.19	1.46	0.74	0.6	1.56	0.95	0.47	1.01	0.59	0.6	1.56	1.52	0.98	0.95	1.43	0.96	0.5	0.96	1.14

Emini surface accessibility scale

- Reference: [Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol. 1985;53:101-105.](#)
- Description: The calculation was based on surface accessibility scale on a product instead of an addition within the window. The accessibility profile was converted to a probability value, and i vary from 1 to 6. A hexapeptide sequence with S_n greater than 1.0 indicates an increased probability for being found on the surface.

Do you want to see how the result page of a prediction method looks?



Antibody Epitope Prediction - Example data

Select one of the example sequences below.

Choose an example sequence

Select Sequence	Sequence Name	Database	Sequence ID
<input checked="" type="radio"/>	Sperm Whale Myoglobin	SwissProt	P02185
<input type="radio"/>	Lysozyme	PDB	5LYM:A
<input type="radio"/>	VP1 [Hepatitis A virus]	NCBI	252562
<input type="radio"/>	Myohemerythin	PDB	2MHR

Submit

Click submit

Citations


[Home](#)[Help](#)[Example](#)[Reference](#)[Download](#)[Contact](#)

References

Chou & Fasman Beta-Turn Prediction:

Chou PY, Fasman GD. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* **47**:45-148.
[PMID: 364941](#)


Emini Surface Accessibility Prediction:

Emini EA, Hughes JV, Perlow DS, Boger J. 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol* **55**:836-839.
[PMID: 2991600](#) 

Karplus & Schulz Flexibility Prediction:

Karplus PA, Schulz GE. 1985. Prediction of chain flexibility in proteins. *Naturwissenschaften* **72**:212-213.
[Naturwissenschaften](#)


Kolaskar & Tongaonkar Antigenicity:

Kolaskar AS, Tongaonkar PC. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* **276**:172-174.
[PMID: 1702393](#) 

Parker Hydrophilicity Prediction:


Parker JM, Guo D, Hodges RS. 1986. 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* **25**:5425-5432.
[PMID: 2430611](#)

Bepipred Linear Epitope Prediction:

Larsen JE, Lund O, Nielsen M. 2006. Improved method for predicting linear B-cell epitopes. *Immunome Res* **2**:2.
[PMID: 16635264](#) 

Ponomarenko JV, Bourne PE. 2007. Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *BMC Struct Biol* **7**:64.

[PMID: 17910770](#) 
[Video presentation on SciVee site](#)

Haste Andersen P, Nielsen M, Lund O. 2006. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci* **15**:2558-2567.
[PMID: 17001032](#) 

Do you need standalone versions?

[Home](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

Antibody Epitope prediction - Download

Antibody Epitope Prediction tool contains collection of python scripts, specific binary for BepiPred and a pickled file containing residue scales for different methods.

- Chou & Fasman Beta-Turn Prediction
- Emini Surface Accessibility Prediction
- Karplus & Schulz Flexibility Prediction
- Kolaskar & Tongaonkar Antigenicity
- Parker Hydrophilicity Prediction
- Bepipred Linear Epitope Prediction

License Agreements

By downloading the standalone tool, you are consenting to be bound by and become a party as the "Licensee" for the use of NetMHC 3.0.

Also you are consenting the terms and conditions of the Non-Profit Open Software License ("Non-Profit OSL") version 3.0

Please read following two agreements before proceeding. If you do not agree to all of the terms of these two agreements, you must not click the download button, not install the product nor use the product, and you do not become a LICENSEE under these agreements.

```
ACADEMIC SOFTWARE LICENSE AGREEMENT FOR END-USERS AT PUBLICLY FUNDED
ACADEMIC, EDUCATION OR RESEARCH INSTITUTIONS FOR THE USE OF NetMHC 3.0
By downloading the Software you are consenting to be bound by and become
a party to this agreement as the "Licensee". If you do not agree to all
of the terms of this agreement, you must not click the Acceptance
button, not install the product nor use the product, and you do not
```

```
This Non-Profit Open Software License ("Non-Profit OSL") version 3.0 (the "Licen
se") applies to any original work of authorship (the "Original Work") whose owne
r (the "Licensor") has placed the following licensing notice adjacent to the cop
yright notice for the Original Work:
```

```
Licensed under the Non-Profit Open Software License version 3.0
```

To download the tools in tar.gz format:

To return to the main page:

IEDB Tools Downloads

Complete Download: IEDB Analysis Resource Virtual Machine Image

For users that would like to run the entire analysis resource locally, a virtual machine image file is available with a paid commercial license. The image is kept in sync with the current version of the IEDB Analysis Resource and is updated on a six month cycle. Please [contact us](#) for details on licensing options.

“Help will always be given at Hogwarts
IEDB to those who ask for it.”



IEDB solution Center
help@iedb.org

Home Help Example Reference Download **Contact**

Contact the developers

Should you require assistance beyond the help provided for each individual tool, please visit the [IEDB Solutions Center](#) or submit a request by clicking on the following link:
help@iedb.org

Outline

- **B cell epitopes:** Discontinuous And linear epitopes
- **Schema** of B cell epitope prediction tools on IEDB
- **Sequence-based epitope prediction methods:**
 - Linear epitope prediction methods
 - Amino acid physicochemical property-based methods
 - Machine learning approaches
- **3D Structure-based epitope prediction methods:**
 - Discontinuous epitope prediction methods
- **Computational antibody design**
 - Antigen and Antibody structure modelling
 - Antibody-protein docking

Epitope prediction

- **When to use epitope prediction methods?**
- When you have verified thoroughly that **no information is available in the IEDB** on the antigen of your interest or
- When you want to **know all the candidate antigenic determinants** in an antigen of your interest other than epitopes provided in the IEDB

Sequence-based epitope prediction

➤ Linear epitope prediction:

- **Amino acid physicochemical properties-based methods:**
 - Features which have been correlated with the location of continuous epitopes
 - β -Turns (**Chou & Fasman**),
 - Surface Accessibility (**Emini**),
 - Flexibility (**Karplus & Schulz**),
 - antigenic propensity: occurrence of residues in epitopes (**Kolaskar & Tongaonkar**),
 - Hydrophilicity (**Parker**).
- Only provide information on protein regions which might be of interest for antibody binding.
- These are not prediction methods in a strict sense.

Linear epitope prediction

➤ Linear epitope prediction methods:

○ Machine learning algorithms

- Positive and negative training datasets are used
- Combination of one or more amino acid scales are used as an input to one of the machine learning algorithms
 - Random Forest (**BepiPred-2.0**)
 - ANN: Artificial Neural Network (ABCpred)
 - SVM: Support Vector Machine (BCpred, FBCpred)
- Prediction accuracy is optimized

B cell epitope prediction tools

IEDB Analysis Resource

[Overview](#) [T Cell Tools](#) [B Cell Tools](#) [Analysis Tools](#) [Tools-API](#) [Download](#) [Datasets](#) [Contribute Tools](#) [References](#)

B Cell Epitope Prediction Tools

B Cell Epitope Prediction

[Prediction of linear epitopes from protein sequence](#)

A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs.

[Discotope - Prediction of epitopes from protein structure](#)

This method incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential along the length of a protein sequence.

[ElliPro - Epitope prediction based upon structural protrusion](#)

This method predicts epitopes based upon solvent-accessibility and flexibility.

[Methods for modeling and docking of antibody and protein 3D structures](#)

This page provides information on available methods for modeling and docking of antibody and protein 3D structures.

Structure Tools



[LYRA \(Lymphocyte Receptor Automated Modelling\):](#)

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.



: Tools under AR Labs which are experimental and are not quite ready for production yet. They are intended for further research, updates and testing.

Linear epitope prediction

Home Help Example Reference Download Contact

Antibody Epitope Prediction

Specify Input

Enter a Swiss-Prot ID (example: P02185)

Or enter a protein sequence in plain format (50000 residues maximum):

```
VLSEGEWQLVLHWAKVEADVAGHGQDILIRLFKSHPETLEKFDRLFHLKTEAEMKASEDLKKHGVTVL  
ALGAILKKKGHHEAELKPLAQSHATKHKIPIKYLEFISEAIIHVLHSRHPGNFGADAGGAMNKALELFRK  
DIAAKYKELGYQG
```

Choose a method:

- [Bepipred Linear Epitope Prediction](#)
- [Bepipred Linear Epitope Prediction 2.0](#)
- [Chou & Fasman Beta-Turn Prediction](#)
- [Emini Surface Accessibility Prediction](#)
- [Karplus & Schulz Flexibility Prediction](#)
- [Kolaskar & Tongaonkar Antigenicity](#)
- [Parker Hydrophilicity Prediction](#)

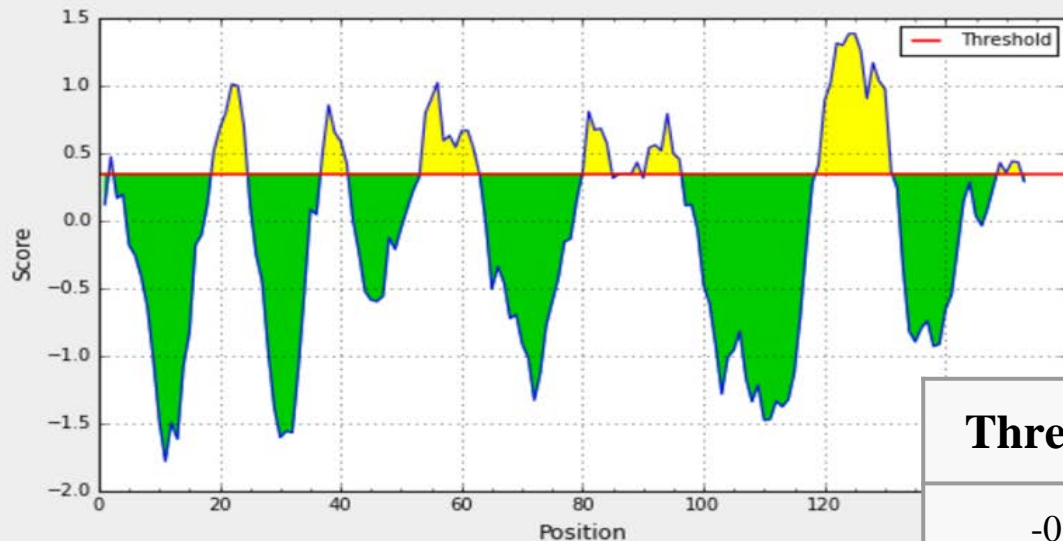
- **BepiPred** is the default and recommended method.
- It is based on a combination of Hidden Markov model (**HMM**) and two amino acid scales,
- **Parker's** hydrophilicity scale and **Levitt's** secondary structure scale.
- Reported AUC was **0.66**. (Larsen et al., 2006)

Bepipred Linear Epitope Prediction Results

Input Sequences

1 VLSEGEWQLV LHWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
 61 LKKHGVTVL T ALGAILKKKG HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
 121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG

Center position: 4 Window size: 7 Threshold: 0.35 Recalculate



Average: -0.105 Minimum: -1.784 Maximum: 1.390

Predicted peptides:

No. ↕	Start ↕	End ↕	Peptide ↕	Length ↕
1	2	2	L	1
2	19	24	ADVAGH	6
3	37	41	PETLE	5
4	54	62	EMKASEDLK	9
5	80	84	GHHEA	5
6	87	87	K	1
7	89	89	L	1

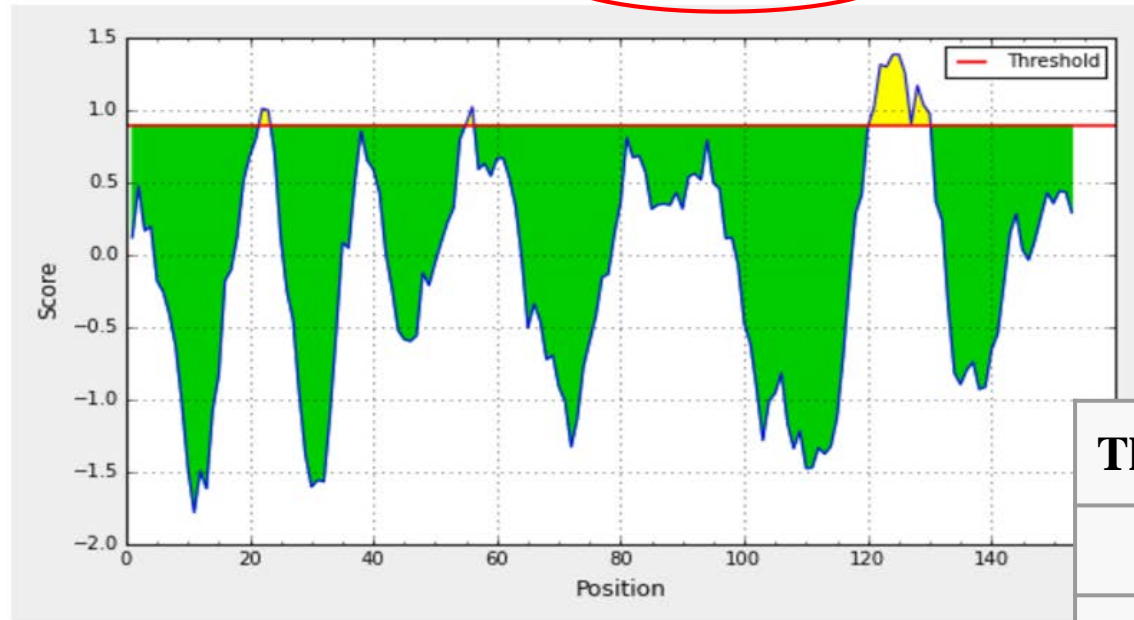
Threshold	Sensitivity	Specificity
-0.20	0.75	0.50
0.20	0.56	0.68
0.35	0.49	0.75
0.90	0.25	0.91
1.30	0.13	0.96

Bepipred Linear Epitope Prediction Results

Input Sequences

1 VLSEGEWQLV LHWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
 61 LKKHGVTVLT ALGAILKKGK HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
 121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG

Center position: 4 Window size: 7 **Threshold: 0.9** Recalculate



Average: -0.105 Minimum: -1.784 Maximum: 1.390

Predicted peptides:

No.	Start	End	Peptide	Length
1	22	23	AG	2
2	55	56	MK	2
3	121	130	GNFGADAGGA	10

Predicted residue scores:

Position	Residue	Score	Assignment
1	V	0.121	.

Threshold	Sensitivity	Specificity
-0.20	0.75	0.50
0.20	0.56	0.68
0.35	0.49	0.75
0.90	0.25	0.91
1.30	0.13	0.96

Predicted peptides:

No. ↕	Start ↕	End ↕	Peptide ↕	Length ↕
1	2	2	L	1
2	19	24	ADVAGH	6
3	37	41	PETLE	5
4	54	62	EMKASEDLK	9
5	80	84	GHHEA	5
6	87	87	K	1
7	89	89	L	1
8	91	96	QSHATK	6
9	119	131	HPGNFGADAGGAM	13
10	149	152	LGYQ	4

Predicted residue scores:

Position ↕	Residue ↕	Score ↕	Assignment ↕
1	V	0.121	.
2	L	0.476	E
3	S	0.168	.
4	E	0.198	.
5	G	-0.180	.
6	E	-0.255	.
7	W	-0.412	.
8	Q	-0.631	.
9	L	-1.022	.
10	V	-1.482	.
11	L	-1.784	.
12	H	-1.496	.

Sequence-based epitope prediction

146	Y	-0.035	.
147	K	0.098	.
148	E	0.265	.
149	L	0.433	E
150	G	0.357	E
151	Y	0.442	E
152	Q	0.436	E
153	G	0.292	.

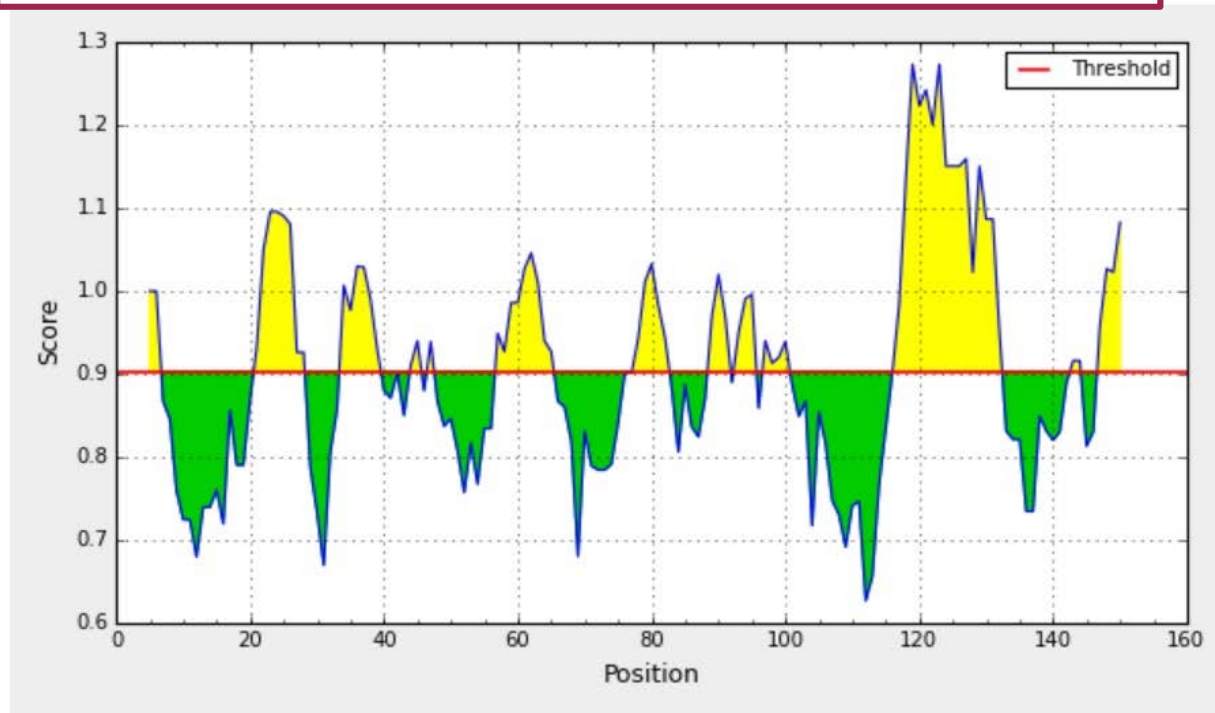
[Download result](#) 

Chou & Fasman Beta-Turn Prediction

Input Sequences

1 VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
61 LKKHGVTVLT ALGAILK KKG HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG

Center position: 4 Window size: 7 Threshold: 0.903 Recalculate



Average: 0.903 Minimum: 0.626 Maximum: 1.274

Predicted residue scores:

Position	Residue	Start	End	Peptide	Score
4	E	1	7	VLSEGEW	0.931
5	G	2	8	LSEGEWQ	1.0
6	E	3	9	SEGEWQL	1.0

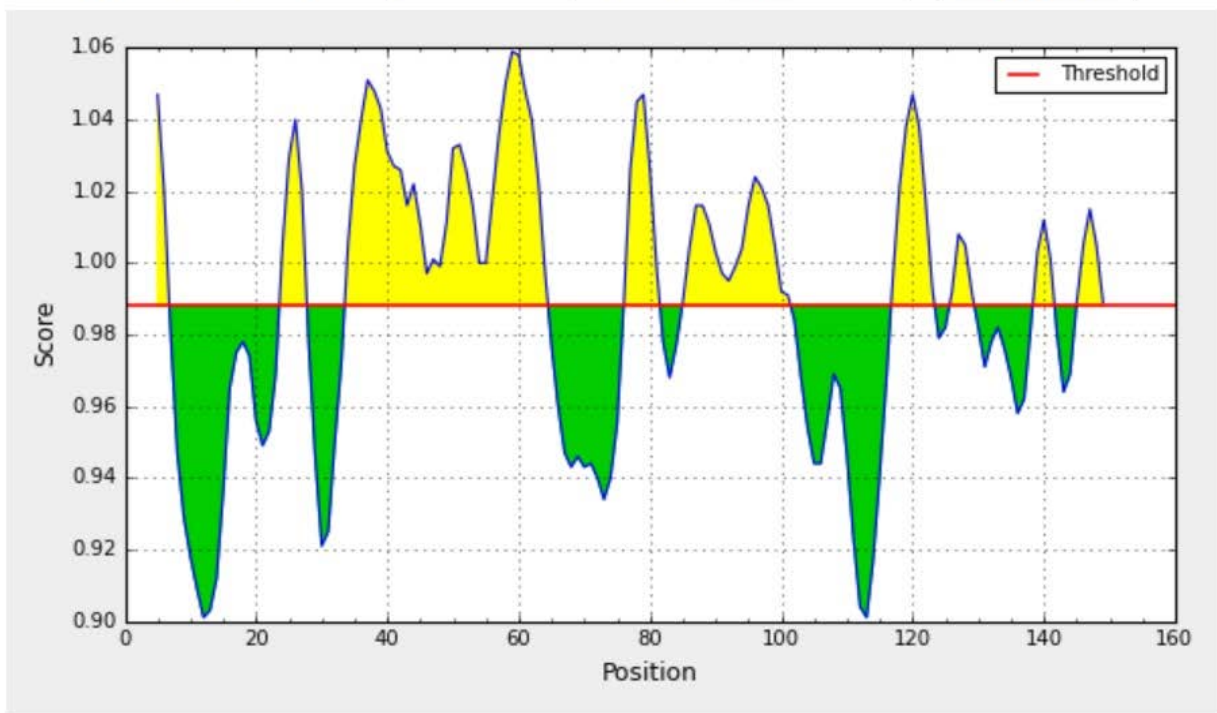
Average score of a protein is chosen as a threshold by default.

Karplus & Schulz Flexibility Prediction

Input Sequences

1 VLSEGEWQLV LHWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
 61 LKKHGVTULT ALGAILKKGK HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
 121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG

Center position: 4 Window size: Threshold:



Average: 0.988 Minimum: 0.901 Maximum: 1.059

Predicted residue scores:

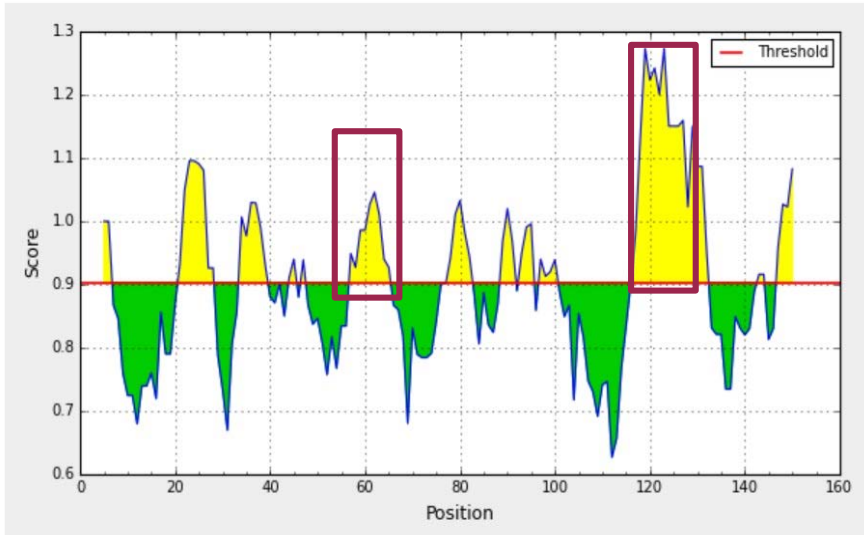
Position	Residue	Start	End	Peptide	Score
4	E	1	7	VLSEGEW	1.053
5	G	2	8	LSEGEWQ	1.047
6	E	3	9	SEGEWQL	1.02

Chou & Fasman Beta-Turn Prediction

Input Sequences

```
1 VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
61 LKKHGVTVLT ALGAILKKGK HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG
```

Center position: 4 Window size: Threshold:

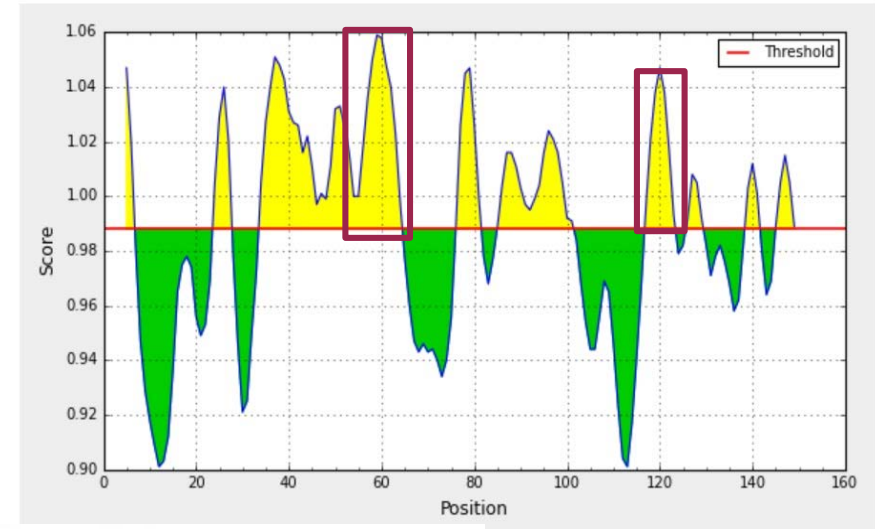


Karplus & Schulz Flexibility Prediction

Input Sequences

```
1 VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
61 LKKHGVTVLT ALGAILKKGK HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG
```

Center position: 4 Window size: Threshold:

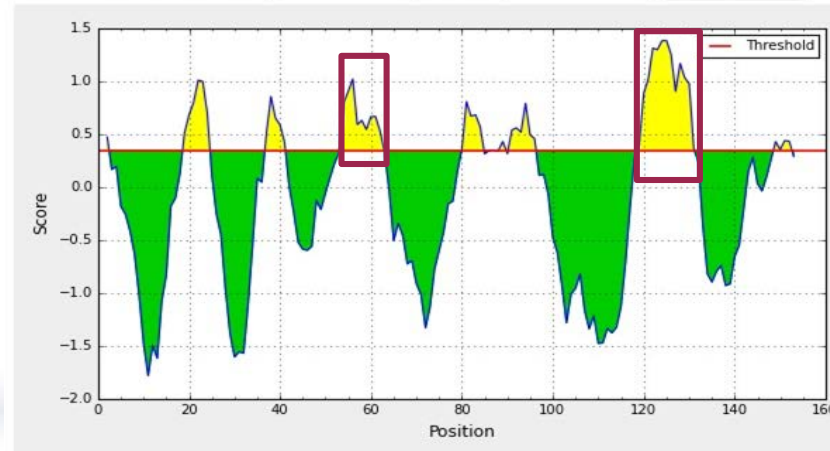


Bepipred Linear Epitope Prediction

Input Sequences

```
1 VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK TEAEMKASED
61 LKKHGVTVLT ALGAILKKGK HHEAELKPLA QSHATKHKIP IKYLEFISEA IIHVLHSRHP
121 GNFGADAGGA MNKALELFRK DIAAKYKELG YQG
```

Center position: 4 Window size: Threshold:



- It is better to use **consensus** of different methods rather than relying on a single method.

Sequence-based epitope **prediction**

- Evaluation of amino acid scales (Greenbaum et al., 2007):
No method gave AUC above 0.60.
- Poor performance might be explained by the benchmark datasets containing long 15-25 aa peptides which along with epitope residues contain non-epitope residues.

15-mer AVVLYHNSACCPKWA

Outline

- **B cell epitopes:** Discontinuous And linear epitopes
- **Schema** of B cell epitope prediction tools on IEDB
- **Sequence-based epitope prediction methods:**
 - Linear epitope prediction methods
 - Amino acid physicochemical property-based methods
 - Machine learning approaches
- **3D Structure-based epitope prediction methods:**
 - Discontinuous epitope prediction methods
- **Computational antibody design**
 - Antigen and Antibody structure modelling
 - Antibody-protein docking

3D Structures of Ab-Ag complexes

Methods for 3D structure determination:

- **X-ray crystallography** (This method provides the most accurate identification of epitopes)
- **Nuclear magnetic resonance (NMR)**
- **Electron microscopy (EM)**

- **Where to get 3D Ab-Ag complexes??**

IEDB 3D export (1791 3D BCR assays)

- **Where to get 3D coordinates of proteins?**

Biomolecular 3D structural data is deposited into **PDB (Protein Data Bank)**



An Information Portal to 133920 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligands

Go

Advanced Search | Browse by Annotations



Take the RCSB PDB User Survey



A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

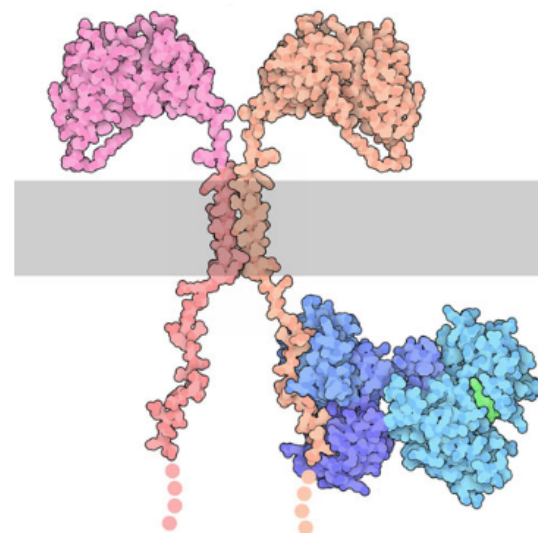
As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

2017 RCSB PDB User Survey



October Molecule of the Month



Chimeric Antigen Receptors

Search in PDB

RCSB PDB Deposit **Search** Visualize Analyze Download Learn More MyPDB Login

RCSB PDB PROTEIN DATA BANK An Information Portal to 121821 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligands **Go**

Advanced Search | Browse by Annotations

EMDataBank NUCLEIC ACID DATABASE StructuralBiology Knowledgebase Worldwide Protein Data Bank Foundation

Facebook Twitter YouTube Apple Android

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

Search Options

- Drill Down by Categories
- Advanced Search
- Sequences
- Ligands
- Drugs & Drug Targets
- Unreleased & New Entries
- Browse by Annotation
- PDB Statistics

Explore the PDB Archive

- Organism
- Taxonomy
- X-ray Resolution
- Polymer Type
- SCOP Classification
- Protein Stoichiometry
- UniProt Molecule Name
- Experimental Method
- Release Date
- Enzyme Classification
- Protein Symmetry
- Membrane Proteins

Select a Organism category below:



- Homo sapiens (32644)
- Escherichia coli (8201)
- Mus musculus (5200)
- Saccharomyces cerevisiae (3434)
- Bos taurus (2548)
- Rattus norvegicus (2517)
- synthetic construct (2031)
- Other (64686)

Search in PDB

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

Search Options

- Drill Down by Categories
- Advanced Search
- Sequences**
- Ligands
- Drugs & Drug Targets
- Unreleased & New Entries
- Browse by Annotation
- PDB Statistics

Search by Sequences

Choose Option A or B to search for protein and nucleic acid sequences. [Read Tutorial](#) [Advanced Sequence Searching](#)

NOTE Parameters: BLAST method, E-value cutoff: 10.0, Mask Low Complexity: On.

Option A: Use PDB Sequence

Select Associated Chain

...

or Option B: Paste Sequence

Run Sequence Search

Search in PDB

Advanced Search Interface

Sequence (BLAST/FASTA/PSI-BLAST) ?

Sequence search (BLAST or FASTA)

Structure Id

Chain Id

Sequence

Search Tool

Mask Low Complexity

E-Value Cutoff

Sequence Identity Cutoff (% , Integer Only)

Retrieve only representatives at sequence identity ?

Match of the above conditions.

Downloading a PDB file

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

RCSB PDB PROTEIN DATA BANK An Information Portal to 121821 Biological Macromolecular Structures


Search by PDB ID, author, macromolecule, sequence, or ligands Go

Advanced Search | Browse by Annotations | Search History (1) | Previous Results (1)

PDB-101 WORLDWIDE PDB PROTEIN DATA BANK EMDatabank NUCLEIC ACID DATABASE Structural Biology Knowledgebase Worldwide Protein Data Bank Foundation

- Structure Summary
- 3D View
- Annotations
- Sequence
- Sequence Similarity
- Structure Similarity
- Experiment
- Literature

Biological Assembly 1 ?



View in 3D: JSmol or PV (in Browser)

Standalone Viewers

Simple Viewer Protein Workshop Ligand Explorer

1Z40

AMA1 from *Plasmodium falciparum*

DOI: 10.2210/pdb1z40/pdb

Classification: **UNKNOWN FUNCTION**

Deposited: 2005-03-14 Released: 2005-08-16

Deposition author(s): [Bai, T.](#), [Becker, M.](#), [Gupta, A.](#), [Strike, P.](#), [Murphy, V.J.](#), [Anderson](#)

Organism: [Plasmodium falciparum](#)

Expression System: Escherichia coli BL21(DE3)

Structural Biology Knowledgebase: 1Z40 (2 models >14 annotations) [SBKB.org](#)

Experimental Data Snapshot

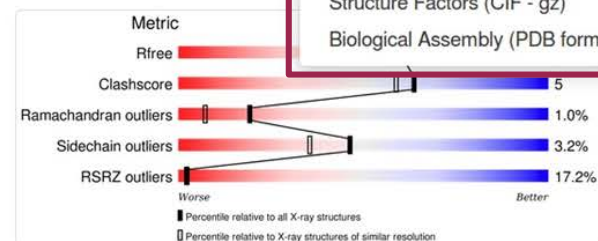
Method: X-RAY DIFFRACTION

Resolution: 1.9 Å

R-Value Free: 0.236

R-Value Work: 0.192

wwPDB Validation



Display Files Download Files

- FASTA Sequence
- PDB Format
- PDB Format (gz)
- PDBx/mmCIF Format
- PDBx/mmCIF Format (gz)
- PDBML/XML Format (gz)
- Structure Factors (CIF)
- Structure Factors (CIF - gz)
- Biological Assembly (PDB format - gz) (A)

Protein Chains

Macromolecules

Classification: [UNKNOWN FUNCTION](#)

Sequence Display for 1Z40

Total Structure Weight: 76870.31

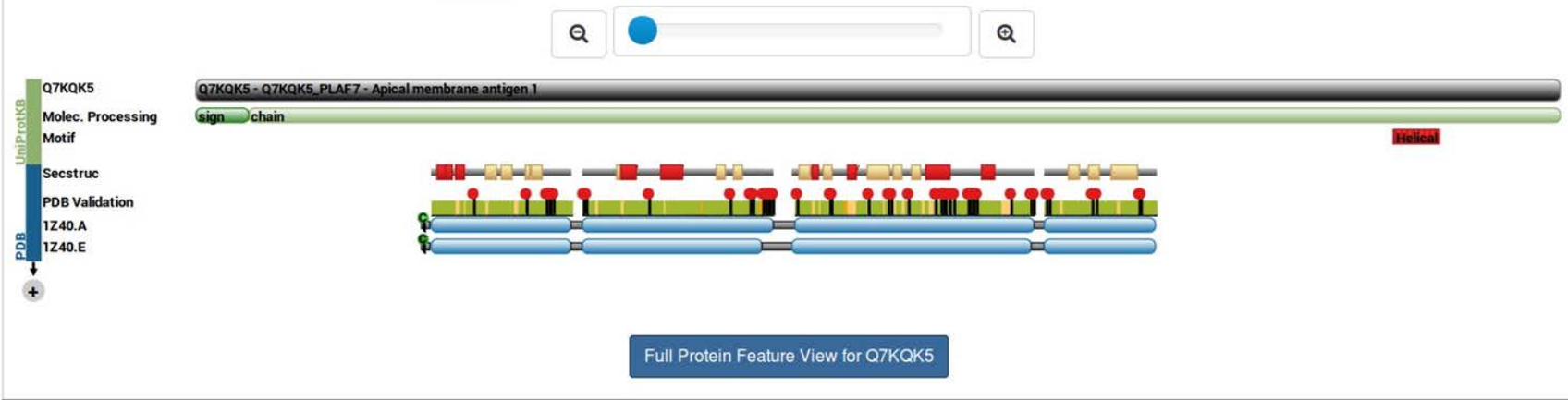
Macromolecule Entities

Toggle Protein Feature View

Molecule	Chains	Length	Organism	Details
apical membrane antigen 1 precursor	A, E	336	Plasmodium falciparum	Fragment: domain I & domain II Gene Name(s): PF11_0344

Protein Feature View - UniProtKB AC: [Q7KQK5](#)

Find similar proteins by: [Sequence](#) | [Structure](#)



What if the 3D structure of a protein of your interest is not available in PDB?

- **Homology or comparative modeling methods, servers and databases**

3D Structure-based epitope prediction

➤ Discontinuous epitope prediction

- Structure-based epitope prediction using:
 - Geometrical properties combined with amino acid scales (**DiscoTope**, **ElliPro**, CEP)
 - Geometrical properties and amino acid scales used as input to machine learning approaches (EPSVR)
 - Protein-protein docking algorithms
- Sequence-based epitope prediction using
 - machine learning approaches (CBTope)

3D Structure-based epitope prediction

➤ DiscoTope:

- DiscoTope 1 – Andersen *et al*, 2006, Protein Science
DiscoTope 2 – Kringelum *et al*, 2012, PLoS Comp. Biol.
- Trained on 75 X-ray structures of antibody-protein complexes (DiscoTope 2 took into account multiple epitopes in an antigen).
- Assigns each residue a score value calculated as a linear combination of normalized values from
 - Parker's hydrophilicity scale,
 - amino acid occurrence,
 - the number of contacts within 10Å, and
 - the area of relative solvent accessibility.
- AUC 0.71 for DiscoTope 1 and 0.73 for DiscoTope 2

3D Structure-based epitope prediction

IEDB Analysis Resource

[Overview](#) [T Cell Tools](#) [B Cell Tools](#) [Analysis Tools](#) [Tools-API](#) [Download](#) [Datasets](#) [Contribute Tools](#) [References](#)

B Cell Epitope Prediction Tools

B Cell Epitope Prediction

[Prediction of linear epitopes from protein sequence](#)

A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs.

[Discotope - Prediction of epitopes from protein structure](#)

This method incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential along the length of a protein sequence.

[ElliPro - Epitope prediction based upon structural protrusion](#)

This method predicts epitopes based upon solvent-accessibility and flexibility.

[Methods for modeling and docking of antibody and protein 3D structures](#)

This page provides information on available methods for modeling and docking of antibody and protein 3D structures.

Structure Tools



[LYRA \(Lymphocyte Receptor Automated Modelling\):](#)

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.



: Tools under AR Labs which are experimental and are not quite ready for production yet. They are intended for further research, updates and testing.

DiscoTope

[Home](#)[Help](#)[Example](#)[Reference](#)[Download](#)[Contact](#)

DiscoTope: Structure-based Antibody Prediction

Step 1: Please enter the 4-letter PDB id Or upload a PDB file	<input type="text" value=""/> (example: 1z40)
	<input type="button" value="Browse..."/> No file selected.
Step 2: Please enter PDB chain id:	<input type="text" value=""/> (example: A)
Step 3: Select version 	1.1 <input type="button" value="v"/>
<input type="button" value="Submit"/>	

Please cite this [reference](#) when using DiscoTope.

DiscoTope

[Home](#)[Help](#)[Example](#)[Reference](#)[Download](#)[Contact](#)

DiscoTope: Structure-based Antibody Prediction

Step 1: Please enter the 4-letter PDB ID
Or upload a PDB file

(example: 1z40)

No file chosen

Step 2: Please enter PDB Chain ID

Step 3: Select version

DiscoTope Chart View

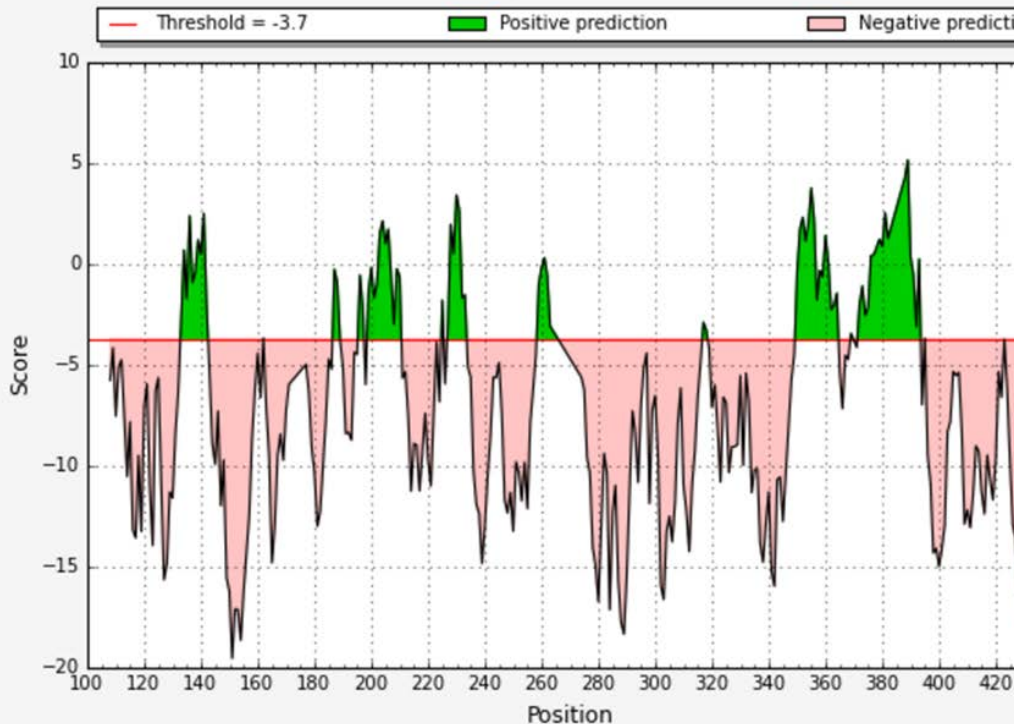
Home Help Example Reference Download Contact

DiscoTope: Structure based antibody prediction.

DiscoTope 2.0 prediction for structure: 1z40 & Chain ID: A

Threshold:

DiscoTope Prediction



DiscoTope 2

Score	Sensitivity	Specificity
1.9	0.17	0.95
0.5	0.23	0.90
-1.0	0.30	0.85
-2.5	0.39	0.80
-3.7	0.47	0.75

DiscoTope Table View

DiscoTope - Result

DiscoTope 2.0 prediction for structure: 1z40 & Chain ID: A

The positive predictions are displayed in green.

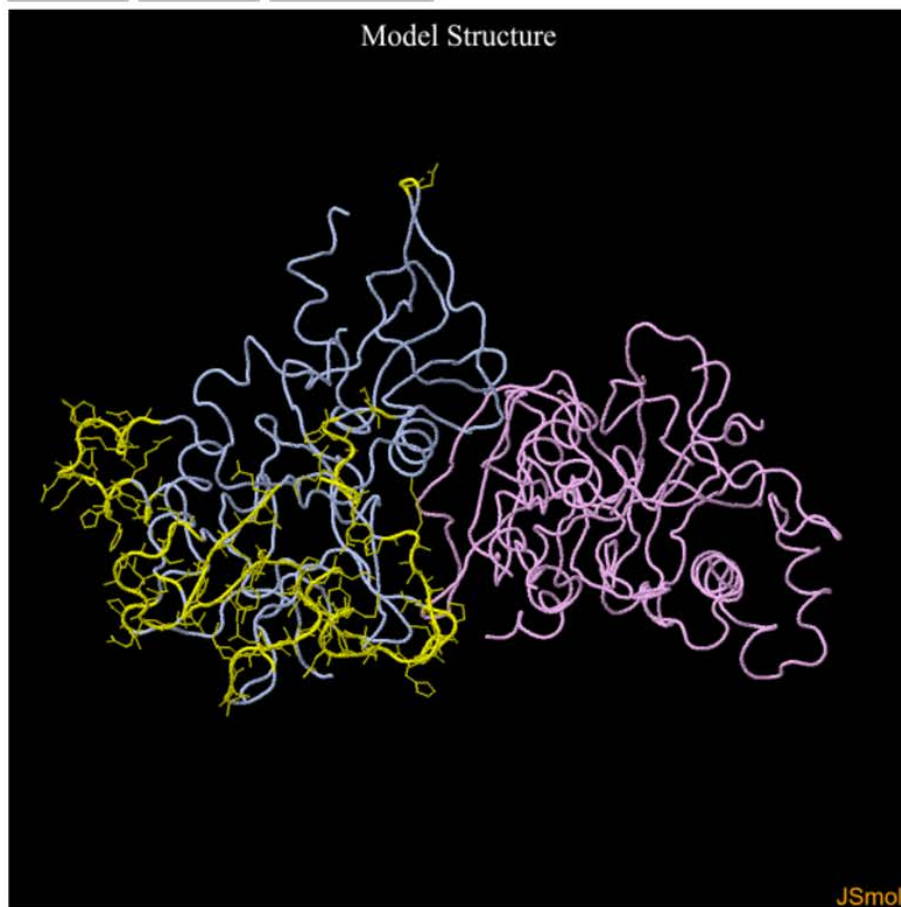
Chart View 3D View Save Prediction

Chain ID ↕	Residue ID ↕	Residue Name ↕	Contact Number ↕	Propensity Score ↕	DiscoTope Score ↕
A	108	ASN	17	-4.287	-5.749
A	109	PRO	7	-3.77	-4.141
A	110	TRP	23	-5.522	-7.532
A	111	THR	4	-5.226	-5.085
A	112	GLU	2	-5.1	-4.744
A	113	TYR	20	-5.97	-7.584
A	114	MET	20	-9.295	-10.526
A	115	ALA	10	-7.532	-7.816
A	116	LYS	23	-11.888	-13.166
A	117	TYR	33	-11.038	-13.564
A	118	ASP	12	-9.15	-9.478
A	119	ILE	35	-10.419	-13.246
A	120	GLU	10	-6.657	-7.042
A	121	GLU	13	-5.019	-5.937
A	122	VAL	30	-8.676	-11.129
A	123	HIS	43	-10.161	-13.938
A	124	GLY	17	-4.876	-6.27
A	125	SER	25	-3.112	-5.629
A	126	GLY	33	-8.05	-10.92
A	127	ILE	39	-12.601	-15.637
A	128	ARG	27	-13.272	-14.85
A	129	VAL	25	-9.506	-11.288
A	130	ASP	39	-8.027	-11.589
A	131	LEU	29	-5.732	-8.408
A	132	GLY	30	-3.241	-6.318
A	133	GLU	20	-0.565	-2.8
A	134	ASP	19	3.255	0.695
A	135	ALA	28	1.707	-1.709

DiscoTope 3D View

JSmol-Rendered PDB Structure

Chart View Table View Save Prediction

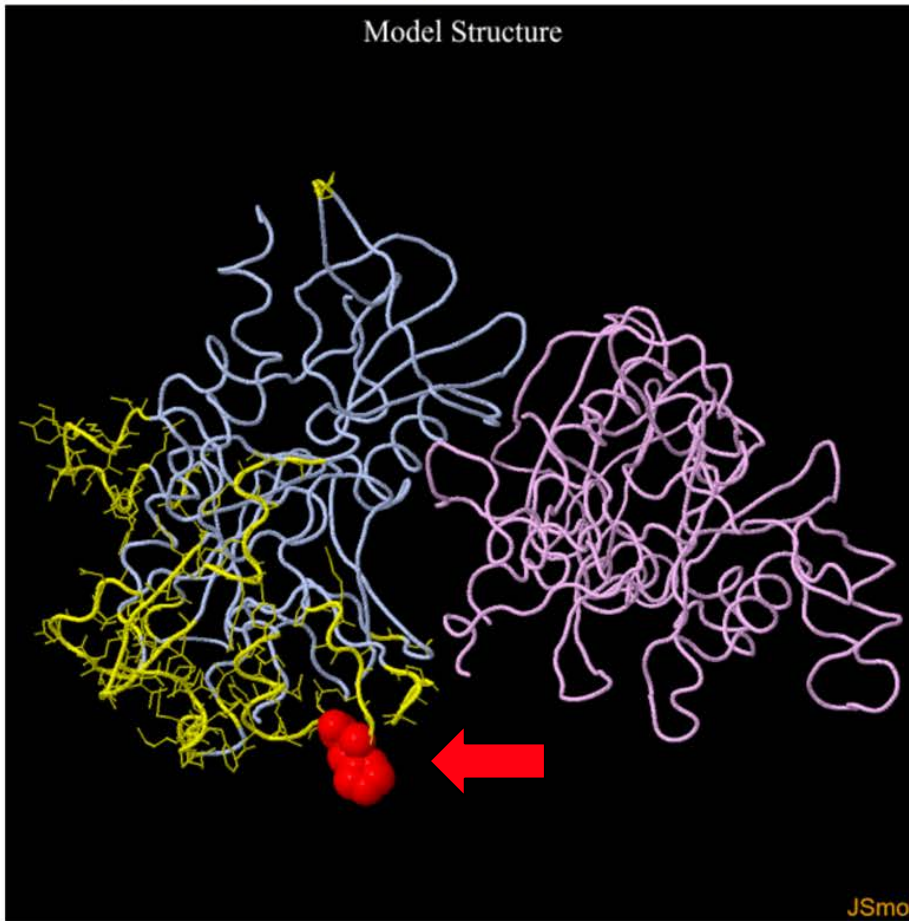


Chain ID	Residue ID	Residue Name	Contact Number	Propensity Score	DiscoTope Score	View
A	133	GLU	20	-0.565	-2.8	CPK
A	134	ASP	19	3.255	0.695	CPK
A	135	ALA	28	1.707	-1.709	CPK
A	136	GLU	9	3.871	2.391	CPK
A	137	VAL	26	2.342	-0.917	CPK
A	138	ALA	14	1.33	-0.433	CPK
A	139	GLY	6	2.141	1.204	CPK
A	140	THR	23	3.529	0.478	CPK
A	141	GLN	12	4.393	2.508	CPK
A	142	TYR	31	1.669	-2.088	CPK
A	162	ASN	1	-4.008	-3.662	CPK
A	187	GLU	0	-0.296	-0.262	CPK
A	188	PRO	10	0.275	-0.907	CPK
A	196	ASP	2	-0.359	-0.548	CPK
A	197	GLU	16	-0.325	-2.128	CPK
A	199	ARG	10	-0.233	-1.356	CPK
A	200	HIS	9	0.977	-0.171	CPK
A	201	PHE	23	1.105	-1.667	CPK
A	202	TYR	27	2.394	-0.986	CPK
A	203	LYS	4	2.272	1.551	CPK
A	204	ASP	1	2.544	2.136	CPK

DiscoTope 3D View

JSmol-Rendered PDB Structure

Chart View Table View Save Prediction



Chain ID	Residue ID	Residue Name	Contact Number	Propensity Score	DiscoTope Score	View
A	133	GLU	20	-0.565	-2.8	CPK
A	134	ASP	19	3.255	0.695	CPK
A	135	ALA	28	1.707	-1.709	CPK
A	136	GLU	9	3.871	2.391	CPK
A	137	VAL	26	2.342	-0.917	CPK
A	138	ALA	14	1.33	-0.433	CPK
A	139	GLY	6	2.141	1.204	CPK
A	140	THR	23	3.529	0.478	CPK
A	141	GLN	12	4.393	2.508	CPK
A	142	TYR	31	1.669	-2.088	CPK
A	162	ASN	1	-4.008	-3.662	CPK
A	187	GLU	0	-0.296	-0.262	CPK
A	188	PRO	10	0.275	-0.907	CPK
A	196	ASP	2	-0.359	-0.548	CPK
A	197	GLU	16	-0.325	-2.128	CPK
A	199	ARG	10	-0.233	-1.356	CPK
A	200	HIS	9	0.977	-0.171	CPK
A	201	PHE	23	1.105	-1.667	CPK
A	202	TYR	27	2.394	-0.986	CPK
A	203	LYS	4	2.272	1.551	CPK
A	204	ASP	1	2.544	2.136	CPK

DiscoTope 3D View

Chart View Table View Save Prediction

Model Structure

The screenshot displays the DiscoTope 3D View interface. At the top, there are three tabs: 'Chart View', 'Table View', and 'Save Prediction'. Below the tabs is a black rectangular area labeled 'Model Structure'. Inside this area, a protein structure is shown in a 3D ribbon representation. The structure is divided into three color-coded regions: a yellow region on the left, a light blue region in the middle, and a pink region on the right. A context menu is open over the yellow region. The menu has a main list on the left and a sub-menu on the right. The main list includes: 'model 1/1', 'Configurations', 'Select (599)', 'View', 'Style', 'Color', 'Surfaces', 'Symmetry', 'Scenes', 'Zoom', 'Spin', 'Vibration', 'Spectra', 'Animation', 'Measurements', 'Set picking', 'Console', 'JavaScript Console', 'Show', 'File', 'Computation', 'Language', and 'About...'. The sub-menu, which is currently open, contains: 'Display Selected Only', 'Selection Halos', 'All', 'None', 'Invert Selection', 'Element', 'Symmetry', 'Protein', 'Nucleic', 'Hetero', 'Carbohydrate', and 'None of the above'. The 'All' option is highlighted in blue. In the bottom right corner of the black area, the text 'JSmol' is visible.

model 1/1
Configurations
Select (599)
View
Style
Color
Surfaces
Symmetry
Scenes
Zoom
Spin
Vibration
Spectra
Animation
Measurements
Set picking
Console
JavaScript Console
Show
File
Computation
Language
About...

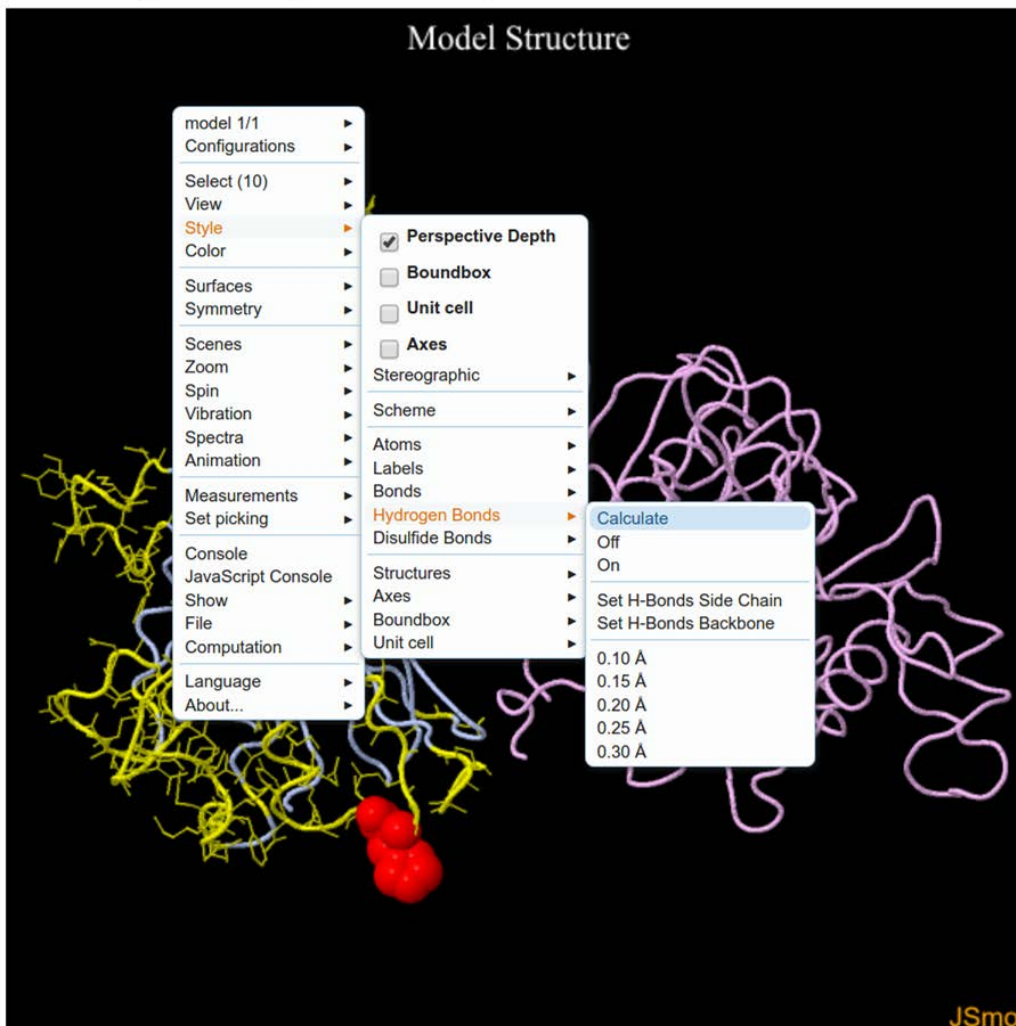
Display Selected Only
Selection Halos
All
None
Invert Selection
Element
Symmetry
Protein
Nucleic
Hetero
Carbohydrate
None of the above

JSmol

DiscoTope 3D View

JSmol-Rendered PDB Structure

Chart View Table View Save Prediction



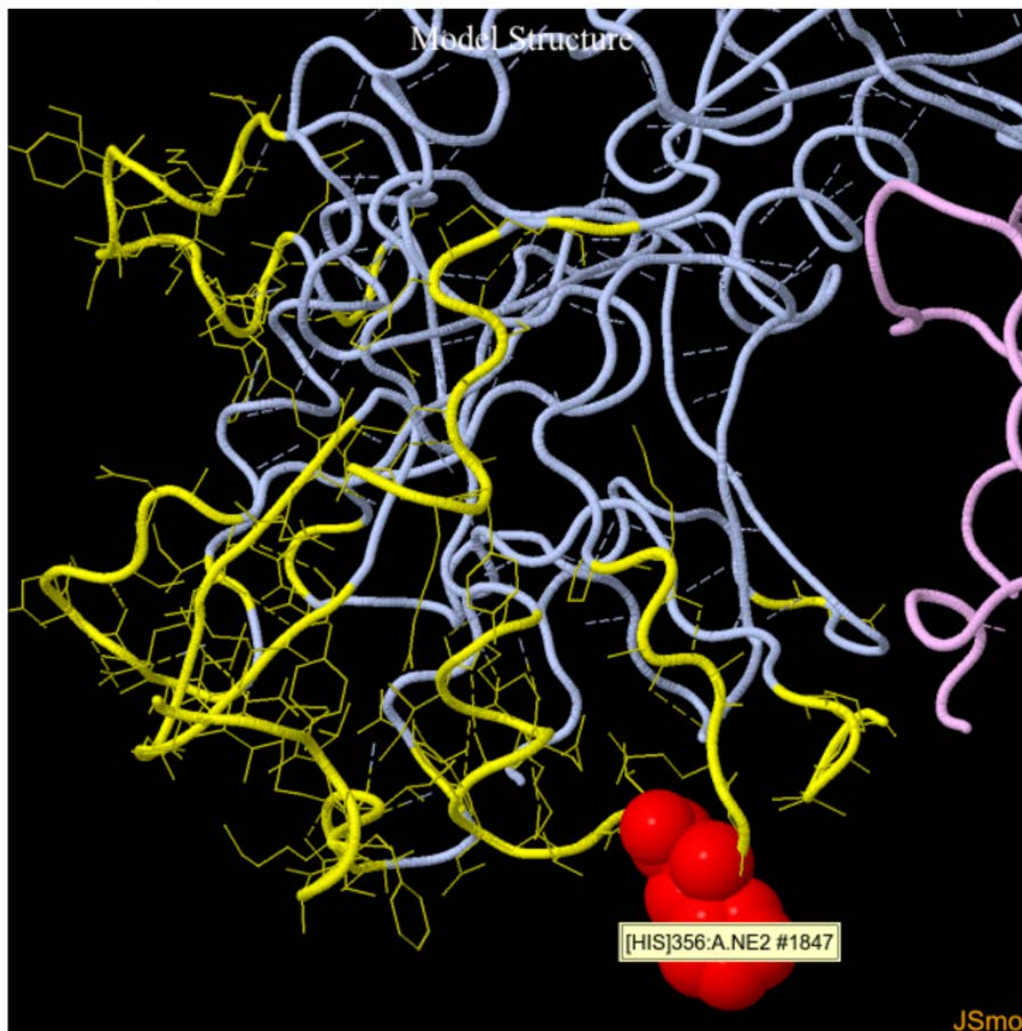
DiscoTope 3D View

JSmol-Rendered PDB Structure

Chart View

Table View

Save Prediction



3D Structure-based epitope prediction

IEDB Analysis Resource

[Overview](#) [T Cell Tools](#) [B Cell Tools](#) [Analysis Tools](#) [Tools-API](#) [Download](#) [Datasets](#) [Contribute Tools](#) [References](#)

B Cell Epitope Prediction Tools

B Cell Epitope Prediction

[Prediction of linear epitopes from protein sequence](#)

A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs.

[Discotope - Prediction of epitopes from protein structure](#)

This method incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential along the length of a protein sequence.

[ElliPro - Epitope prediction based upon structural protrusion](#)

This method predicts epitopes based upon solvent-accessibility and flexibility.

[Methods for modeling and docking of antibody and protein 3D structures](#)

This page provides information on available methods for modeling and docking of antibody and protein 3D structures.

Structure Tools



[LYRA \(Lymphocyte Receptor Automated Modelling\):](#)

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.



: Tools under AR Labs which are experimental and are not quite ready for production yet. They are intended for further research, updates and testing.

ElliPro

IEDB Analysis Resource

[Home](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

ElliPro: Antibody Epitope Prediction

Specify Sequence(s)

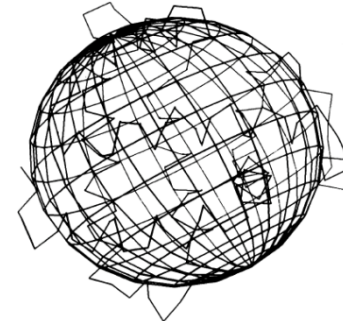
Enter PDB ID(s) or upload PDB file No file chosen

Select Epitope Prediction Parameters

Minimum score: (Default is 0.5)

Maximum distance (Angstrom): (Default is 6)

- Implements three algorithms:
 - approximation of the protein shape as an ellipsoid
 - calculation of the residue protrusion index (PI)
 - clustering of neighboring residues based on PI values



Ponomarenko *et al*, BMC Bioinformatics 2008.
Thornton *et al*, EMBO J. 1986.

EliPro Input

IEDB Analysis Resource

[Home](#)[Help](#)[Example](#)[Reference](#)[Download](#)[Contact](#)

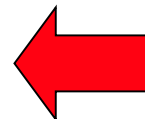
EliPro: Antibody Epitope Prediction

Select PDB chain(s) for calculation

Input pdb id: **5LYM**

Number of chain(s) found: **2**

Select	No.	Chain	Chain type	Number of residues
<input type="checkbox"/>	1	A	Amino acid	129
<input type="checkbox"/>	2	B	Amino acid	129

[Submit](#)[Reset](#)

EliPro Output

[Home](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

Input Sequences: 5LYM

Chain: A
1 KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS TDYGILQINS
61 RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIYS DNGMNAWVA WRNRCKGTDV
121 QAWIRGCRL

Predicted Linear Epitope(s):

No.	Chain	Start	End	Peptide	Number of residues	Score	3D structure
1	A	45	50	RNTDGS	6	0.78	View
2	A	112	129	RNRCKGTDVQAWIRGCRL	18	0.771	View
3	A	100	103	SDGN	4	0.76	View
4	A	64	81	CNDGRTPGSRNLCNIPCS	18	0.666	View
5	A	1	7	KVFGRCE	7	0.597	View
6	A	13	23	KRHGLDNYRGY	11	0.574	View
7	A	85	88	SSDI	4	0.504	View

Click to view 3D structure
in JSMol

Predicted Discontinuous Epitope(s):

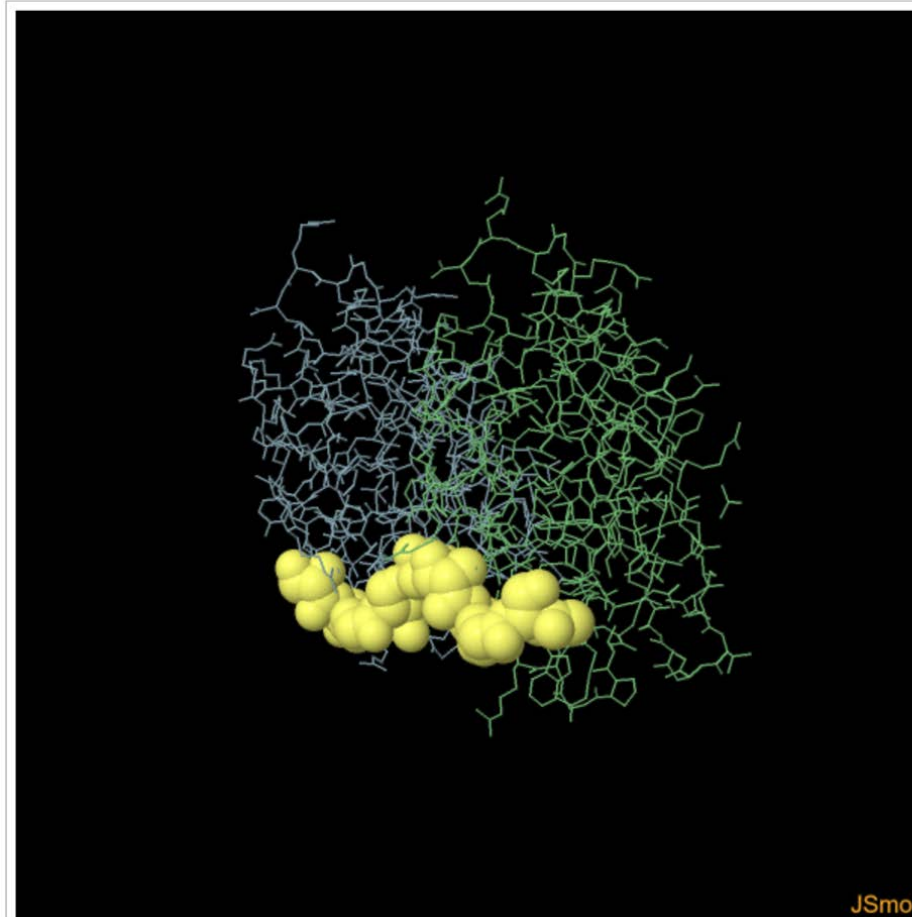
No.	Residues	Number of residues	Score	3D structure
1	A:S100, A:D101, A:G102, A:N103, A:N106	5	0.727	View
2	A:K1, A:V2, A:F3, A:G4, A:R5, A:C6, A:E7, A:F38, A:N39, A:T40, A:Q41, A:A42, A:S85, A:S86, A:D87, A:I88, A:R112, A:N113, A:R114, A:C115, A:K116, A:G117, A:T118, A:D119, A:Q121, A:A122, A:I124, A:R125, A:G126, A:C127, A:R128, A:L129	32	0.657	View
3	A:R45, A:N46, A:T47, A:D48, A:G49, A:S50, A:N59, A:S60, A:R61, A:W62, A:W63, A:C64, A:N65, A:D66, A:G67, A:R68, A:T69, A:P70, A:G71, A:S72, A:R73, A:N74, A:I75, A:C76, A:N77, A:I78, A:P79, A:S81	28	0.648	View

ElliPro Output

ElliPro: Epitope 3D Structures for 5LYM

No.	Residues	Number of residues	Score
1	A:S100, A:D101, A:G102, A:N103, A:N106	5	0.727

JSmol-Rendered PDB Structure



ELIPro Output

Predicted Linear Epitope(s):

No.	Chain	Start	End	Peptide	Number of residues	Score	3D structure
1	A	45	50	RNTDGS	6	0.78	View
2	A	112	129	RNRCKGTDVQAWIRGCRL	18	0.771	View
3	A	100	103	SDGN	4	0.76	View
4	A	64	81	CNDGRTPGSRNLCNIPCS	18	0.666	View
5	A	1	7	KVFGRC	7	0.597	View
6	A	13	23	KRHGLDNRYGY	11	0.574	View
7	A	85	88	SSDI	4	0.504	View

Predicted Discontinuous Epitope(s):

No.	Residues	Number of residues	Score	3D structure
1	A:S100, A:D101, A:G102, A:N103, A:N106	5	0.727	View
2	A:K1, A:V2, A:F3, A:G4, A:R5, A:C6, A:E7, A:F38, A:N39, A:T40, A:Q41, A:A42, A:S85, A:S86, A:D87, A:I88, A:R112, A:N113, A:R114, A:C115, A:K116, A:G117, A:T118, A:D119, A:Q121, A:A122, A:I124, A:R125, A:G126, A:C127, A:R128, A:L129	32	0.657	View
3	A:R45, A:N46, A:T47, A:D48, A:G49, A:S50, A:N59, A:S60, A:R61, A:W62, A:W63, A:C64, A:N65, A:D66, A:G67, A:R68, A:T69, A:P70, A:G71, A:S72, A:R73, A:N74, A:L75, A:C76, A:N77, A:I78, A:P79, A:S81	28	0.648	View
4	A:A10, A:K13, A:R14, A:G16, A:L17, A:D18, A:N19, A:Y20, A:R21, A:G22, A:Y23, A:S24	12	0.564	View

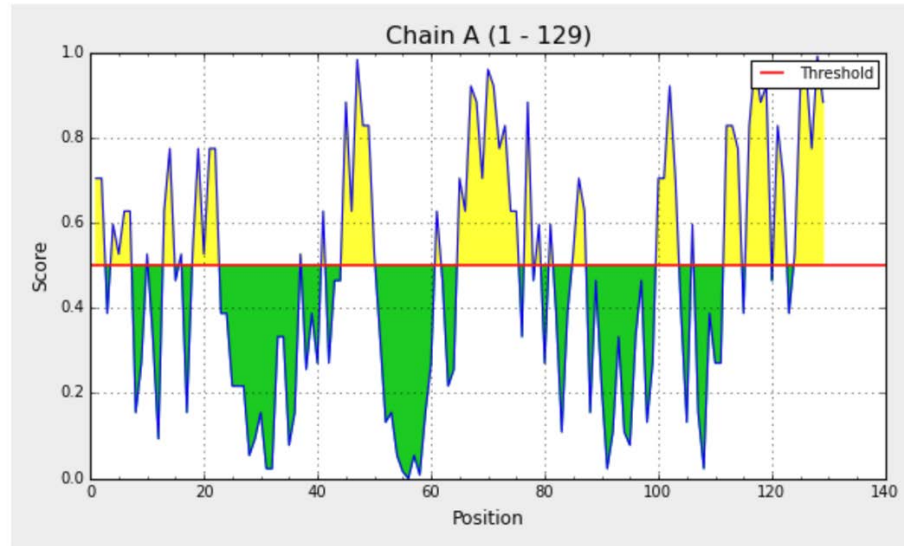
[Click here to view residue scores](#)

© 2005-2017 | [IEDB Home](#)

Supported by a contract from the [National Institute of Allergy and Infectious Diseases](#), a component of the National Institutes of Health in the Department of Health and Human Services.

EliPro Output

EliPro: 2D Score Chart(s) for 5LYM



Data table

No.	Chain	Residue number	Residue name	Score
1	A	1	LYS	0.705
2	A	2	VAL	0.705
3	A	3	PHE	0.388
4	A	4	GLY	0.597
5	A	5	ARG	0.527
6	A	6	CYS	0.628
7	A	7	GLU	0.628
8	A	8	LEU	0.155
9	A	9	ALA	0.271

Exercise

- Use BepiPred and DiscoTope to predict B cell epitopes of dengue 2 virus envelope glycoprotein.
- Download crystal structure and sequence of dengue 2 virus envelope glycoprotein from PDB (PDB ID: [4UTC](#)).

Exercise

Bepipred Linear Epitope Prediction Results

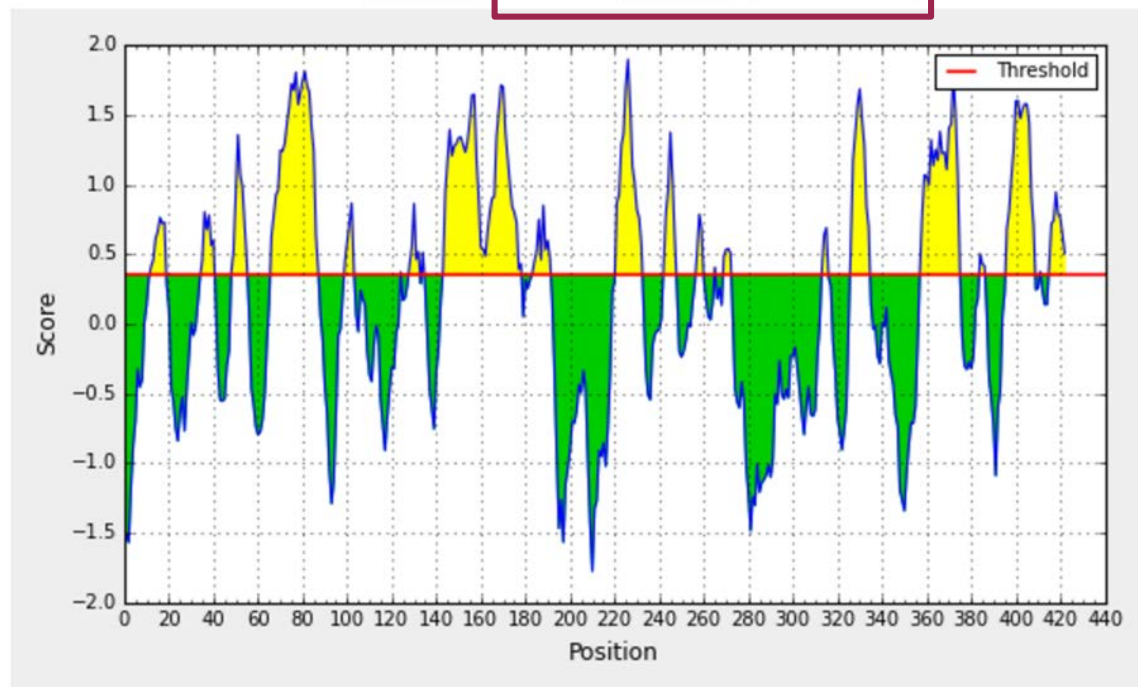
Input Sequences

```
1 MRCIGISNRD FVEGVSGGSW VDIVLEHGSC VTTMAKNKPT LDFELIKTEA KQPATLRKYC
61 IEAKLTNTTT ESRCPQTQEP SLNEEQDK RFICKHSMVD RGWNGCGLF GKGGIVTCAK
121 FTCKKNMEGK IVQPENLEYT IVITPHSGEE HAVGNDTGKH GK EIKITP QSSTTEAELT
181 GYGTVTMECS PRTGLDFNEM VLLQMEDKAW LVHRQWFLDL PLPWLPGADT QGSNWIQKET
241 LVTF KNPH AKKQDVVVLG SQEGAMHTAL TGATEIQMSS GNLLFTGHLK CRLRMDKLQL
301 KGMSYSMCTG KFKIVKEIAE TQHGTI VI RVQYEGDGSP CKIPFEITDL EKRHVLGRLI
361 TVNPIVTEKD SPVNIEAEPF FGDSYIIVGV EPGQLKLNWL RPLESRGP FEGKPIPNNL
421 LGLDSTRTGH HH
```

Center position: 4 Window size: 7

Threshold: 0.35

Recalculate



Average: 0.188 Minimum: -1.776 Maximum: 1.900

Too many epitope candidates?

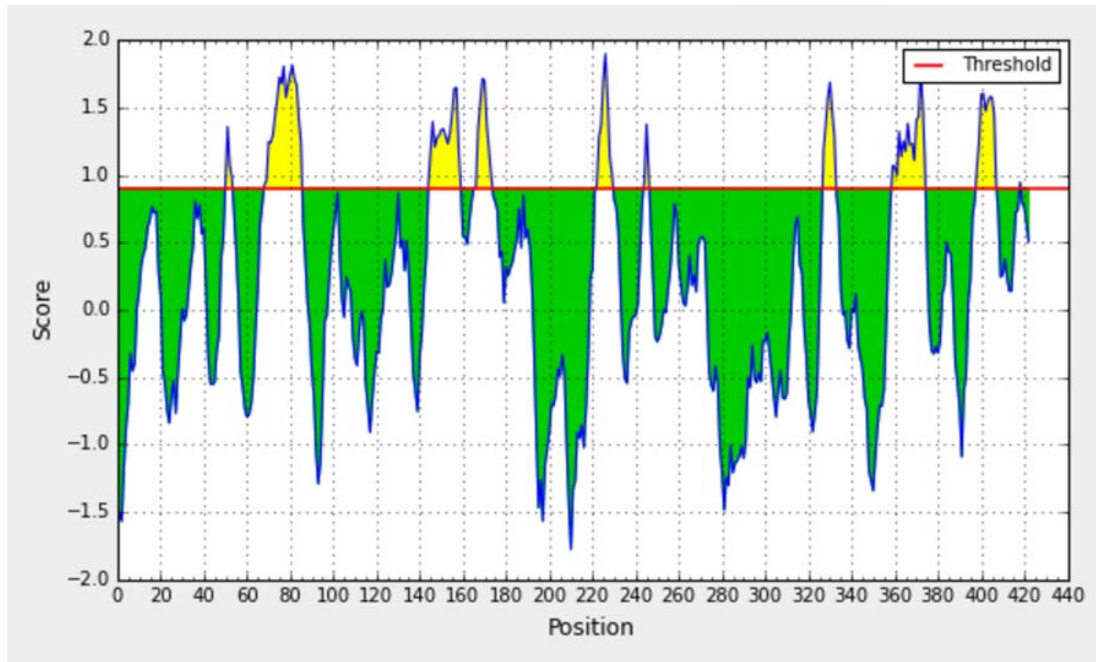
Exercise

Bepipred Linear Epitope Prediction Results

Input Sequences

```
1 MRCIGISNRD FVEGVSGGSW VDIVLEHGSC VTTMAKNKPT LDFELIKTEA KQPATLRKYC
61 IEAKLTNTTT ESRCPQTQEP SLNEEQDK RFICKHSMVD RGWNGCGGLF GKGGIVTCAK
121 FTCKKNMEGK IVQPENLEYT IVITPHSGEE HAVGNDTGKH GK EIKITP QSSTTEAELT
181 GYGTVTMECS PRTGLDFNEM VLLQMEDKAW LVHRQWFLDL PLPWLPGADT QGSNWIQET
241 LVTF KNPH AKKQDVVVLG SQEGAMHTAL TGATEIQMSS GNLLFTGHLK CRLRMDKLQL
301 KGMSYSMCTG KFKIVKEIAE TQHGTI VI RVQYEGDGP CKIPFEITDL EKRHVLGRLI
361 TVNPIVTEKD SPVNIEAEPF FGDSYIIVGV EPGQLKLNWL RPLESRGP FEGKPIPNNL
421 LGLDSTRTGH HH
```

Center position: 4 Window size: Threshold:



Average: 0.188 Minimum: -1.776 Maximum: 1.900

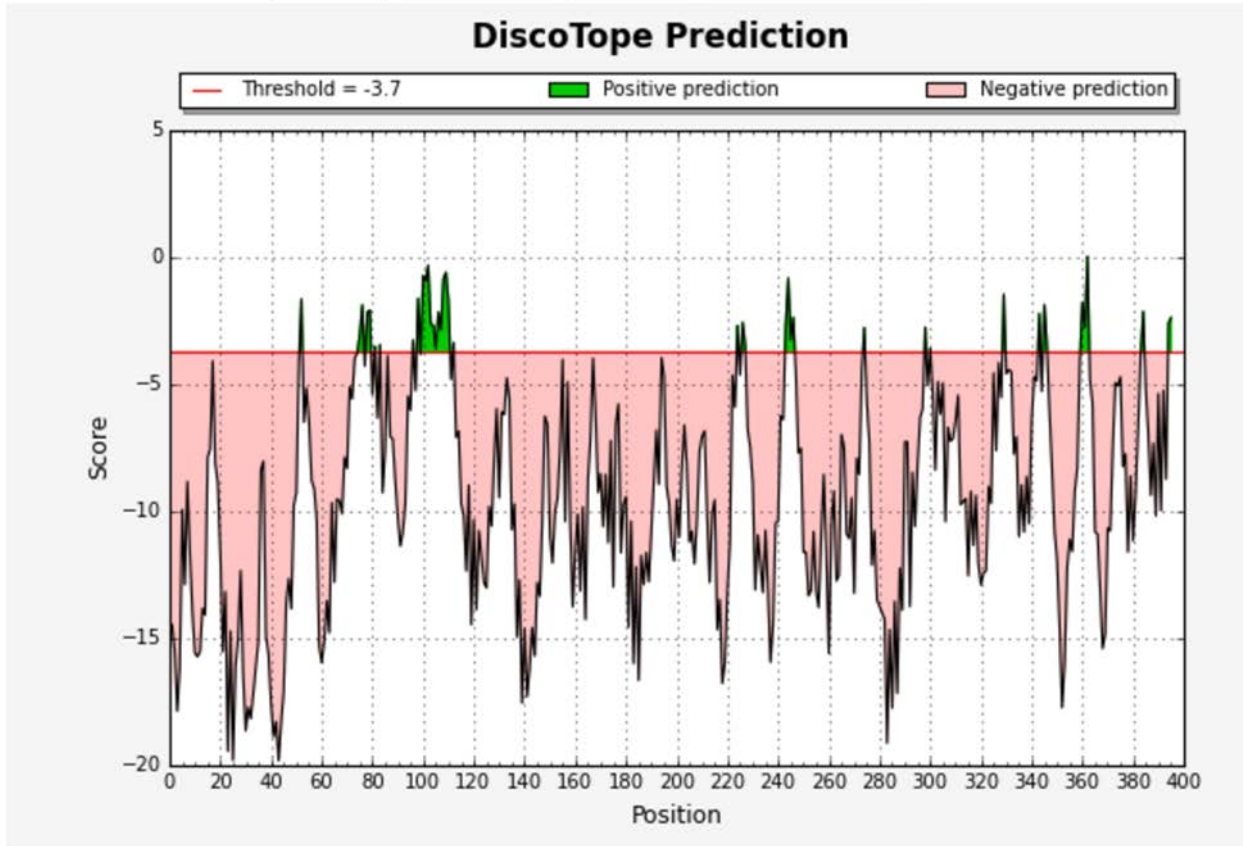
Score threshold of 0.9 corresponds to 90% specificity

Exercise

DiscoTope: Structure based antibody prediction.

DiscoTope 2.0 prediction for structure: 4utc & Chain ID: A

Threshold:



Did you notice the length difference between **BepiPred** (length **422**) and **DiscoTope** (length **391**) outputs?

Exercise

Protein alignment	Nucleotide alignment	Web services	Help & Documentation
#	#	#-----	#-----
4UTC_A_seqres	1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA		50
4UTC_A_atomse	1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA		50
4UTC_A_seqres	51 KQPATLRKYCIEAKLTNTTTSRCPTQGEP SLNNEEQDKRFICKHSMVDRG		100
4UTC_A_atomse	51 KQPATLRKYCIEAKLTNTTTSRCPTQGEP SLNNEEQDKRFICKHSMVDRG		100
4UTC_A_seqres	101 WGNCGCLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA		150
4UTC_A_atomse	101 WGNCGCLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA		150
4UTC_A_seqres	151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPTGLDFNEMVLLQ		200
4UTC_A_atomse	151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPT - -DFNEMVLLQ		198
4UTC_A_seqres	201 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV		250
4UTC_A_atomse	199 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV		248
4UTC_A_seqres	251 VVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCLRMDKLQKLGMSYS		300
4UTC_A_atomse	249 VVLGSQEGAMHTALTGATEIQMSSGNLLF - -HLKCLRMDKLQKLGMSYS		296
4UTC_A_seqres	301 MCTGKFKIVKEIAETQHGTVIRVQYEGDGS PCKIPFEITDLEKRHLVGR		350
4UTC_A_atomse	297 MCTGKFKIVKEIAETQHGTVIRVQYEGDGS PCKIPFEITDLEKRHLVGR		346
4UTC_A_seqres	351 LITVNP I VTEKDSPVNI EAEPFGDSYIIVGVEPGQLKLNWLRPLESRGP		400
4UTC_A_atomse	347 LITVNP I VTEKDSPVNI EAEPFGDSYIIVGVEPGQLKLNWLRPL - - - - -		391
4UTC_A_seqres	401 FEGKPIP NLLGLDSTRTGHHH	422	
4UTC_A_atomse	392 - - - - -	391	
#	#	#-----	#-----

- There might be missing residues in PDB coordinate file compared to the sequence file provided by PDB.
- These missing residues are not resolved properly in the structure.
 - e.g. flexible loops

Exercise

```

4UTC_A_BepiPred 1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA 50
|
4UTC_A_DiscoTope 1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA 50
|
4UTC_A_BepiPred 51 KQPATLRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFICKHSMVDRG 100
|
4UTC_A_DiscoTope 51 KQPATLRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFICKHSMVDRG 100
|
4UTC_A_BepiPred 101 WNGGCGLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA 150
|
4UTC_A_DiscoTope 101 WNGGCGLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA 150
|
4UTC_A_BepiPred 151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPTGLDFNEMVLLQ 200
|
4UTC_A_DiscoTope 151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPT--DFNEMVLLQ 198
|
4UTC_A_BepiPred 201 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV 250
|
4UTC_A_DiscoTope 199 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV 248
|
4UTC_A_BepiPred 251 VVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCLRMDKLQKGM SYS 300
|
4UTC_A_DiscoTope 249 VVLGSQEGAMHTALTGATEIQMSSGNLLF--HLKCLRMDKLQKGM SYS 296
|
4UTC_A_BepiPred 301 MCTGKFKIVKEIAETQHGTVIRVQYEGDGSPPCKIPFEITDLEKRHV LGR 350
|
4UTC_A_DiscoTope 297 MCTGKFKIVKEIAETQHGTVIRVQYEGDGSPPCKIPFEITDLEKRHV LGR 346
|
4UTC_A_BepiPred 351 LIITVNPIVTEKISFVNIEAEPFPGDSYIIVGVEPGQLKNWLRPLESRGP 400
|
4UTC_A_DiscoTope 347 LIITVNPIVTEKISFVNIEAEPFPGDSYIIVGVEPGQLKNWLRPL----- 391
|
4UTC_A_BepiPred 401 FEGKPIPNPLLGLDSTRIGHHH 422

```

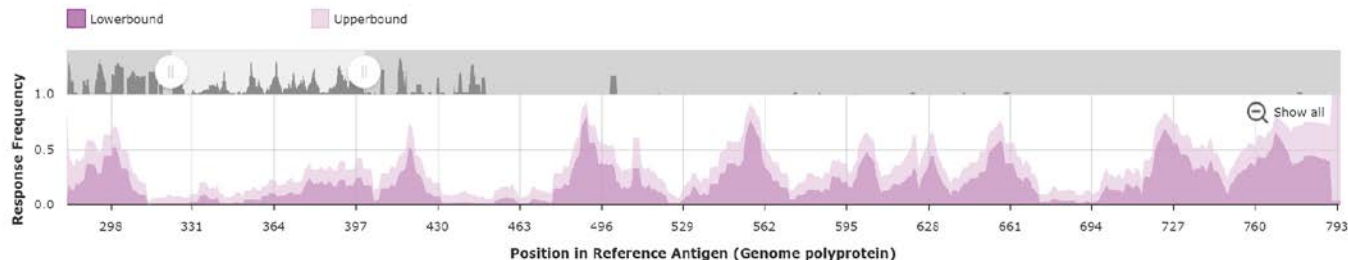


Predicted



Correctly predicted

Epitope residues from 3D B cell assay 3319631 (PDB ID: 2R69) were mapped on Dengue envelope glycoprotein.



```

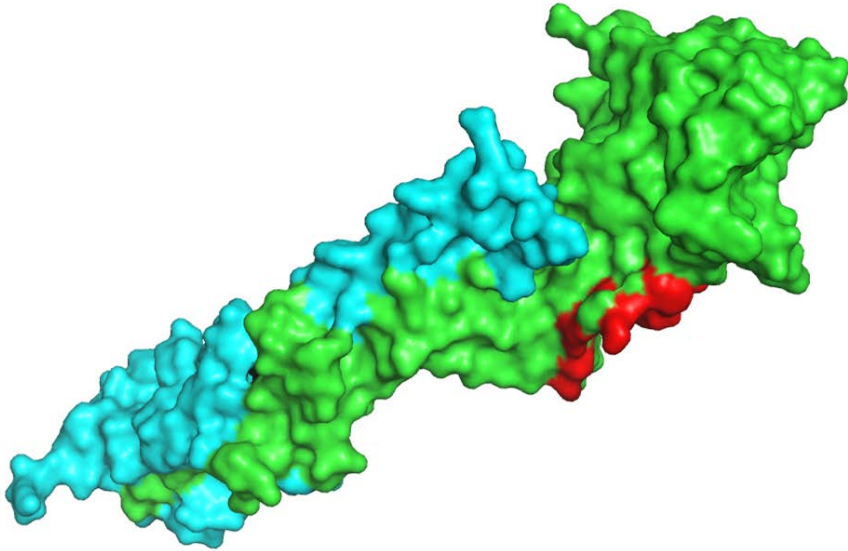
4UTC_A_BepiPred 1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA 50
4UTC_A_DiscoTope 1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA 50
4UTC_A_BepiPred 51 KQPAILRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFICKHSMVDRG 100
4UTC_A_DiscoTope 51 KQPATILRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFICKHSMVDRG 100
4UTC_A_BepiPred 101 WGNCGLFGKGGIVICAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA 150
4UTC_A_DiscoTope 101 WGNCGLFGKGGIVICAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA 150
4UTC_A_BepiPred 151 VGNDIGKHGKEIKITPQSSITTEAELTGYGVTIMECSPTGLDFNEMVLLQ 200
4UTC_A_DiscoTope 151 VGNDIGKHGKEIKITPQSSITTEAELTGYGVTIMECSPT--DFNEMVLLQ 198
4UTC_A_BepiPred 201 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLIVTFKNPHAKKQDV 250
4UTC_A_DiscoTope 199 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLIVTFKNPHAKKQDV 248
4UTC_A_BepiPred 251 VVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCLRMDKLQKGMSSYS 300
4UTC_A_DiscoTope 249 VVLGSQEGAMHTALTGATEIQMSSGNLLF--HLKCLRMDKLQKGMSSYS 296
4UTC_A_BepiPred 301 MCTGKFKIVKEIAETQHGTVIRVQYEGDGSFCKIPFEITDLEKRHVLGR 350
4UTC_A_DiscoTope 297 MCTGKFKIVKEIAETQHGTVIRVQYEGDGSFCKIPFEITDLEKRHVLGR 346
4UTC_A_BepiPred 351 LITVNPVITKDSFVNIEAEPFPGDSYIIVGVEPGQLKLNWLRPLESRGP 400
4UTC_A_DiscoTope 347 LITVNPVITKDSFVNIEAEPFPGDSYIIVGVEPGQLKLNWLRPL----- 391
4UTC_A_BepiPred 401 FEGKPIPNPLGLDSTRIGHHH 422
4UTC_A_DiscoTope 392 ----- 391

```

Predicted
 Correctly predicted

Epitope residues from the IEDB in Dengue envelope glycoprotein.

3D Structure-based epitope prediction



<http://www.ofranlab.org/PEASE>

- Theoretically, the whole exposed surface of an antigen can be targeted by different antibodies.
- Therefore, antibody sequence based B cell epitope prediction method called PEASE was developed.
 - Users must provide antigen structure and antibody sequence.

HOME LAB MEMBERS SERVICES PUBLICATIONS CONTACT

PEASE: Predicting Epitopes using Antibody Sequence

This automated tool predicts the epitope for a given antigen structure and an antibody sequence.

Upload the structure of the antigen in a PDB format, or choose an existing PDB file:

Antigen Structure: No file chosen

Antigen Structure ID:

If the antigen structure is a computational model, please upload in addition the antigen sequence, in order to identify residues with no coordinates:
Note: The antigen sequence should be in a fasta format, and the title line should be the chain ID (e.g. ">C" for antigen chain C).

Antigen sequence: No file chosen

Sela-Culang *et al*, Bioinformatics 2015.

Benchmark on 42 X-ray structures of Ab-protein complexes (Ponomarenko & Bourne, 2008)

Average AUC values

- **0.73 ElliPro** (Ponomarenko et al., 2008)
- **0.65 Epitopia** (Rubinstein et al., 2008)
- **0.63 PEPITO** (Sweredoski & Baldi, 2008)
- **0.60 DiscoTope 1** (Andersen et al, 2006)

- **0.59 DOT** (1st model, bound **Ab-protein docking**)
- **0.58 PatchDock** (1st model, bound **Ab-protein docking**)

Benchmark on 52 X-ray structures of Ab-protein complexes (Kringelum et al., 2012, *PLoS Comp. Biol.*)

Average AUC values

(* means p-value < 0.05 in comparison with DiscoTope 2)

- **0.73 DiscoTope 2** (Kringelum et al., 2012)
- **0.73 PEPITO** (Sweredoski & Baldi, 2008)
- **0.73 Epitopia** (Rubinstein et al., 2008)
- **0.72 SEPPA** (Sun et al., 2009)
- **0.71 DiscoTope 1** (Andersen et al., 2006)
- **0.69* ElliPro** (Ponomarenko et al., 2008)
- **0.65* EPCES** (Liang et al., 2009)
- **0.59* EPSVR** (Liang et al., 2010)

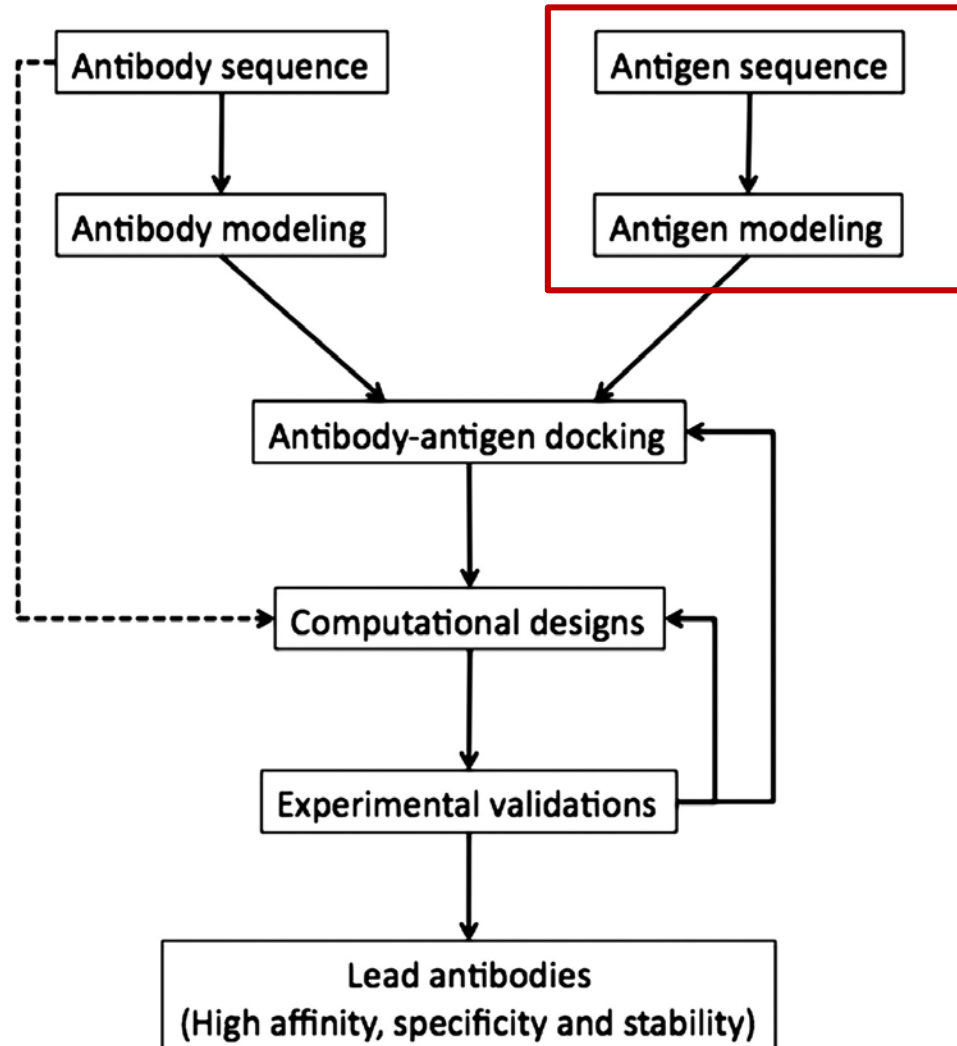
3D Structure-based epitope prediction

- The reason for the relatively poor performance is in the quality of the benchmark datasets.
 - structural information on the entire “biological unit“ is often not available
 - existence of well characterized epitopes from very few antigens

Outline

- **B cell epitopes:** Discontinuous And linear epitopes
- **Schema** of B cell epitope prediction tools on IEDB
- **Sequence-based epitope prediction methods:**
 - Linear epitope prediction methods
 - Amino acid physicochemical property-based methods
 - Machine learning approaches
- **3D Structure-based epitope prediction methods:**
 - Discontinuous epitope prediction methods
- **Computational antibody design**
 - Antigen and Antibody structure modelling
 - Antibody-protein docking

Computational antibody design

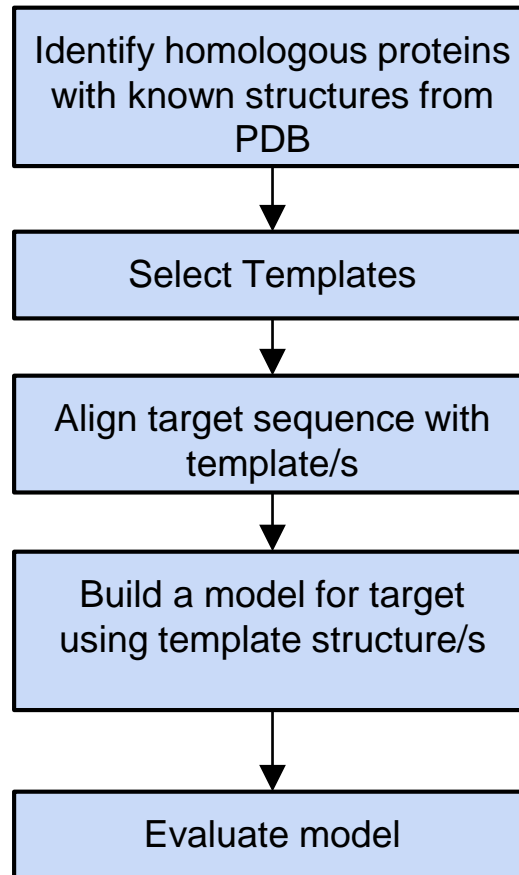


Kuroda D *et al*, PEDS 2012.

What if the 3D structure of a protein of your interest is not available in PDB?

- Protein structures are more conserved than protein sequences
- Homologous proteins have similar structures
- **Homology or comparative modeling:**
method,
databases (**PMP**, ModBase) and
web-servers (**I-TASSER**, ROBETTA, HHPred).

Homology modeling algorithm



Protein Modeling Portal

(www.proteinmodelportal.org)

PSI | The Protein Model Portal

YLDVGFDTTRVAVIQFMHLEEK
SDFSNDVFPFADRSKQVRAIP
SVVVKRGGAVPIGIGRADTTIS



Home Interactive Modeling Quality Estimation Protein Modeling 101 More ▾

Please enter your query

Welcome to the

Protein Model Portal (PMP)

PMP gives access to various models computed by comparative modeling methods provided by different partner sites, and provides access to various interactive services for model building, and quality assessment.

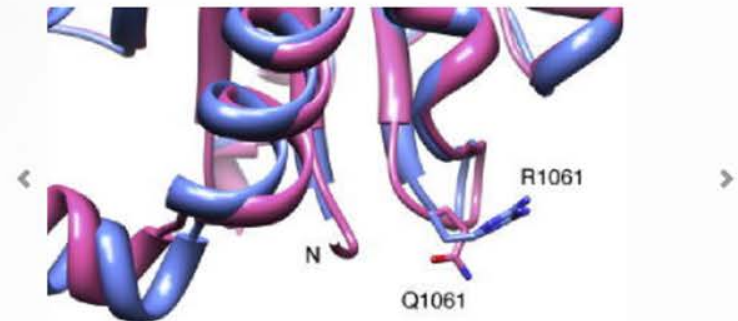
```
EDWGPCTEHGEHRIRTPRTPARVTGGVFLVDKNPHNTTESRLVDFSQFSRGNTRVSW  
PKFAVPNLQSLTNLLSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARLSS  
NSRIINNQHRTMQNLHNSCSRNLVSLMLLYKTYGRKHLHLYSHPIILGFRKIPMGVGLSP  
FLLAQFTSAICSVVRRAFPCHLAFSYMDDVVLGAKSVQHLESLYAAVTNFLSLGIHLNP  
HKTKRWGYSLNFMGYVIGGWGTLPEHIVQKIKMCFRKLVPNRPIDWKVCQRIVGLLG
```

Search



Examples: [UniProt AC] [UniProt ID] [RefSeq] [PDBID] [Sequence] [Free Text]

Modeling Highlights (Show all)



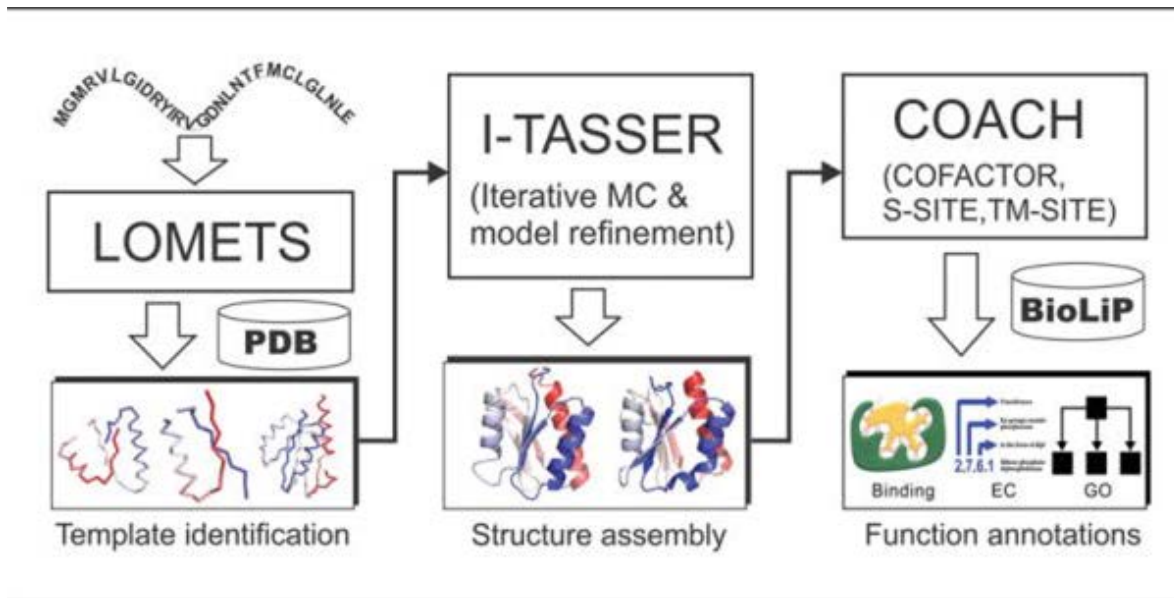
A Structure of a Collagen VI VWA Domain Displays N and C Termini at Opposite Sides of the Protein

Becker AA. *Structure* (2014) 22(2), 199–208

Haas *et al*, Database 2013

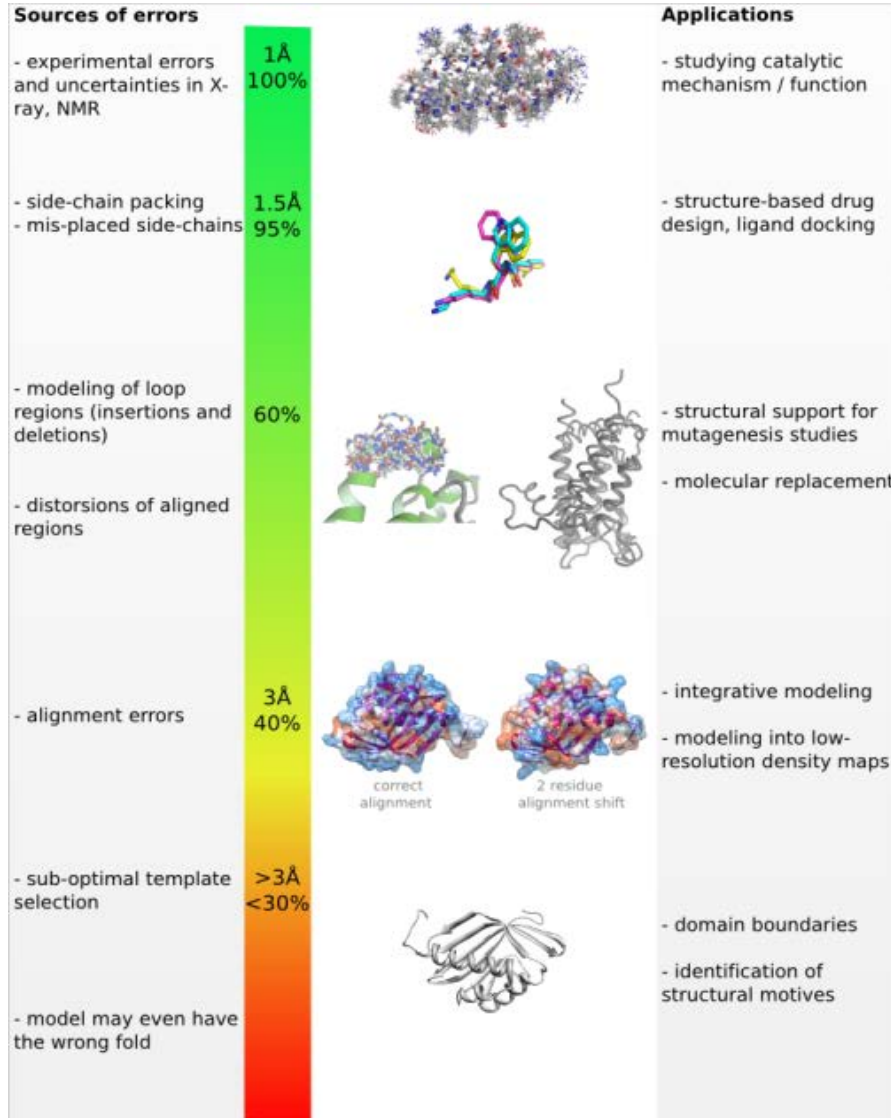
74

I-TASSER



- One of the best server for protein structure prediction in recent community-wide CASP7 to CASP12 experiments.
- Ranked as the best for function prediction in CASP9.
- Relatively simple user interface and parameters to understand.

Homology modeling

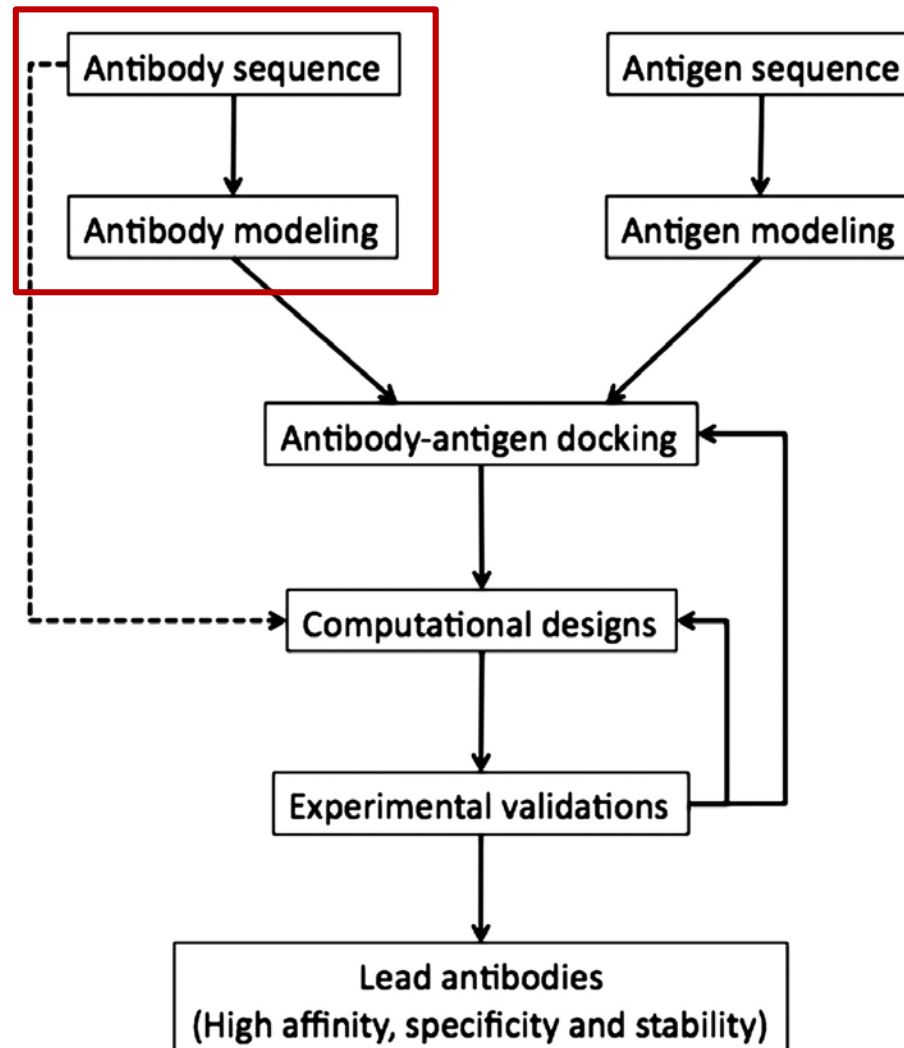


- Models can have errors if target to template sequence identity < 30%
- If the antigen model quality is good then it can be used for B cell epitope prediction using Discotope and ElliPro.

Rost 1999; Martí-Renom *et al.* 2000

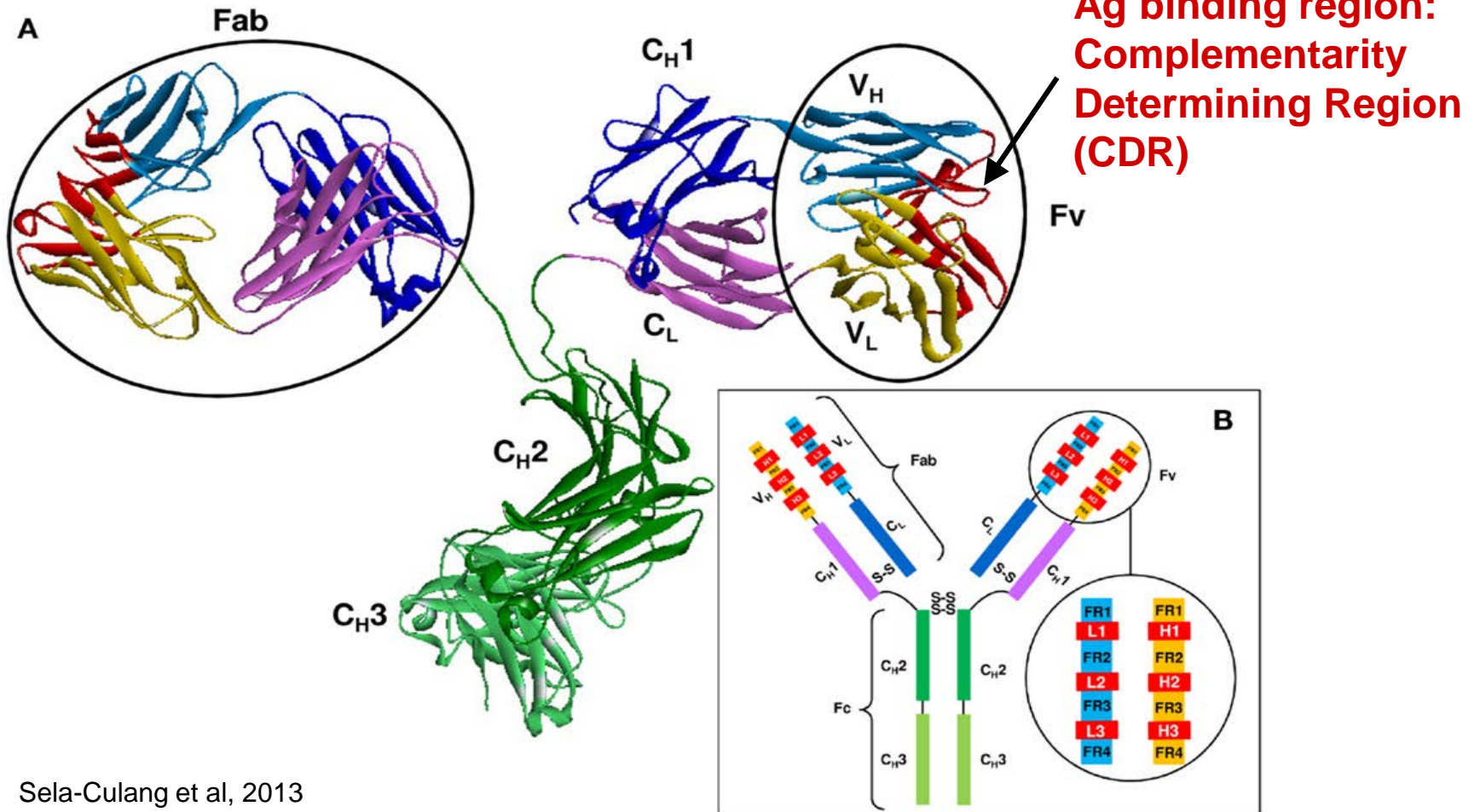
Image from PMP website

Computational antibody design



Kuroda D *et al*, PEDS 2012.

Antibody structure



Sela-Culang et al, 2013

3D Structure-based epitope prediction

IEDB Analysis Resource

[Overview](#) [T Cell Tools](#) [B Cell Tools](#) [Analysis Tools](#) [Tools-API](#) [Download](#) [Datasets](#) [Contribute Tools](#) [References](#)

B Cell Epitope Prediction Tools

B Cell Epitope Prediction

[Prediction of linear epitopes from protein sequence](#)

A collection of methods to predict linear B cell epitopes based on sequence characteristics of the antigen using amino acid scales and HMMs.

[Discotope - Prediction of epitopes from protein structure](#)

This method incorporates solvent-accessible surface area calculations, as well as contact distances into its prediction of B cell epitope potential along the length of a protein sequence.

[ElliPro - Epitope prediction based upon structural protrusion](#)

This method predicts epitopes based upon solvent-accessibility and flexibility.

[Methods for modeling and docking of antibody and protein 3D structures](#)

This page provides information on available methods for modeling and docking of antibody and protein 3D structures.

Structure Tools



[LYRA \(Lymphocyte Receptor Automated Modelling\):](#)

The LYRA server predicts structures for either T-Cell Receptors (TCR) or B-Cell Receptors (BCR) using homology modelling. Framework templates are selected based on BLOSUM score, and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates.



: Tools under AR Labs which are experimental and are not quite ready for production yet. They are intended for further research, updates and testing.

LYRA: Lymphocyte Receptor Automated Modeling

IEDB Analysis Resource

[Home](#) [Result](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

Lymphocyte Receptor Automated Modelling (LYRA)

Specify Chains

First chain sequence:

Second chain sequence:

Or select file containing chains: [?](#) No file chosen

[Advanced options »](#)

Side Chain Modeling Method: [?](#) ▾

Blacklisted PDBs (optional): [?](#)

- B- and T-cell receptor structure modeling
- **Canonical structures (CS):** The hypervariable CDR loops can only assume a limited number of conformations (Chothia & Lesk, JMB 1987).
- Canonical structures can usually be identified by specific sequence features

Klausen MS *et al*, NAR 2015.

LYRA: Lymphocyte Receptor Automated Modeling

IEDB Analysis Resource

[Home](#) [Result](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

Lymphocyte Receptor Automated Modelling (LYRA)

Specify Chains

First chain sequence:

Second chain sequence:

Or select file containing chains: [?](#) No file chosen

[Advanced options »](#)

Side Chain Modeling Method: [?](#) ▾

Blacklisted PDBs (optional): [?](#)

- Templates are identified using BLOSUM62 scores
- Loop modeling: If the canonical structure of target and template loops do not match then the highest scoring identical canonical structure loops from other structures are selected.

Klausen MS *et al*, NAR 2015.

LYRA

IEDB Analysis Resource

[Home](#) [Result](#) [Help](#) [Example](#) [Reference](#) [Download](#) [Contact](#)

Lymphocyte Receptor Automated Modelling (LYRA)

Specify Chains	
First chain sequence:	<pre>DIQMTQSPASLSASVIGATVITCRSENIIDSYLAWYQQRQKSPQLLVYAATNLADGVPSRFSGSGSGTQY SLKINSLQSEDEVARYYCQHYSTTPWTFGGGTQLEIKRADAAPTVISIFPPSSEQLTSGGASVVCFLNRFYPK DINVKWKIDGSEKQNGVLNSWTDQSKDSTYSMSSTLTLTKDEYERHNSYTCETHKTSTSPIVKSFNENE C</pre>
Second chain sequence:	<pre>EVQLQQSGPELVKPGASVKISCKASGYSFTGYMNVKQSPKSLKLEWIGEMSPSTGRITTYNQNFKAKATLT VDQSSSTAYMQLKSLTSEDSAVYYCARSVPLTLLIEDWYFDVWGTTTIVTSSAKTTPPSVYPLAPGAAQ TNSMVTLGCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSDLYTLSSSVTPVPSSTWVPEWTEVCNVAHPAS STKVDKIKVPR</pre>
Or select file containing chains: ? <input type="button" value="Choose File"/> No file chosen	
Advanced options »	
<input type="button" value="Submit"/> <input type="button" value="Reset"/>	

LYRA

Lyra results

Contents [\[hide\]](#)

- [1 Input sequences](#)
- [2 Summary of modelled BCR](#)
- [3 Alignment](#)
 - [3.1 Heavy chain alignment](#)
 - [3.2 Kappa Light chain alignment](#)
- [4 Structure 3D View](#)
- [5 Download PDB file](#)

Input Sequences

Name	Sequence
First chain sequence:	DVVMQTPLSISVTLGQPASISCKSSQSVDLTFAIWVFORPGQSPRKLIFLISKRDSGVPDRFTGS ASGDTFLKISRVEEDVGVYYCWQGFHPHTVGGGKLEIA
Second chain sequence:	GVQLVESGGGVVQPGRSIRLSCAASGFTFSTYAMHWVRQAPGRGLEWVAIISYDGSKKYYADS VKGRFTISRNNKDTLFAQMNSVRAEDTAVYYCARASIAAARVLDYWGRTMVTSS

Summary of modelled BCR

	Heavy chain			Kappa Light chain		
	Template	Template CS	Predicted CS [?]	Template	Template CS	Predicted CS [?]
Framework	1NL0			1NLD		
Loop 1	1NL0	1	1	1QLR	6	6
Loop 2	1NL0	4	4	1NLD	1	1
Loop 3	3I9G	22	22	1NLD	1	1
Packing	1NLD					

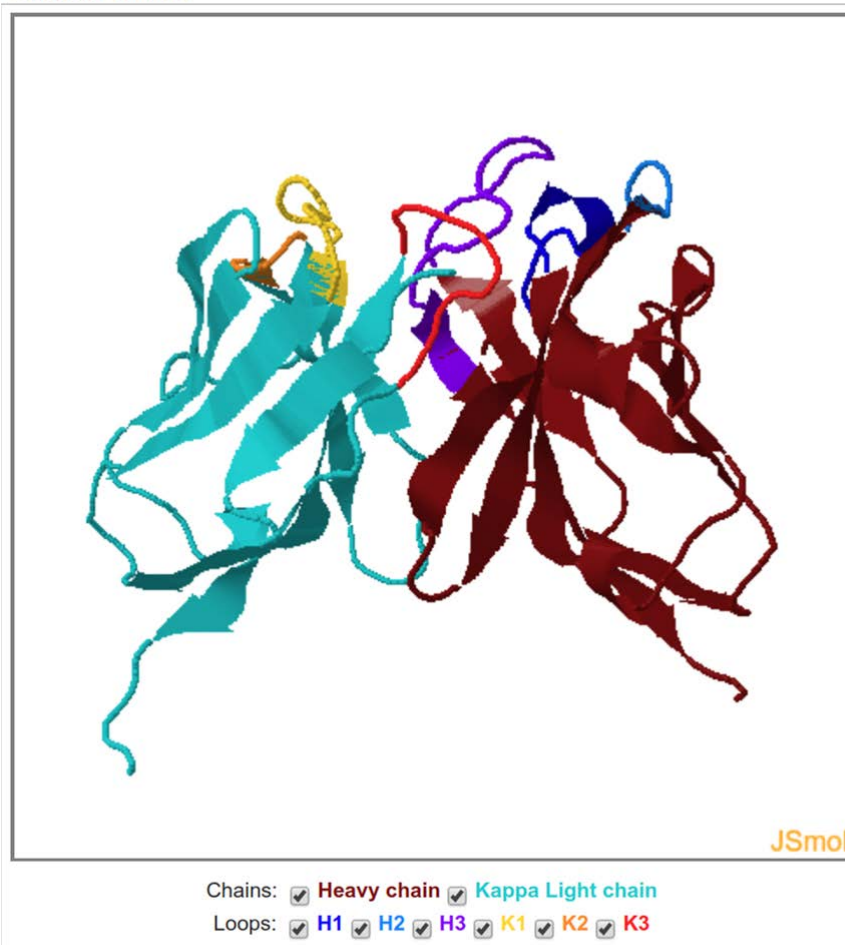
Alignment:

Heavy chain alignment:

	1	10	20	30	33	43	55	65	75	92			
Template:	-VQLVESGGGVVQPGRSIRLSCAASGFTFST-----YAMHWVRQAPGKGLEWVAIISY-----DGSKKYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARAGFYGSTIWF-----			ABCDEF	G		ABCDEF	GHI		ABC		ABCDEFGHIJKLMN	
Input:	-VQLVESGGGVVQPGRSIRLSCAASGFTFST-----YAMHWVRQAPGRGLEWVAIISY-----DGSKKYYADSVKGRFTISRNNKDTLFAQMNSVRAEDTAVYYCARASIAAARVLD-----												
				□	□		□	□		□		□	
				□ H1 □			□ H2 □				□ H3 □		

LYRA

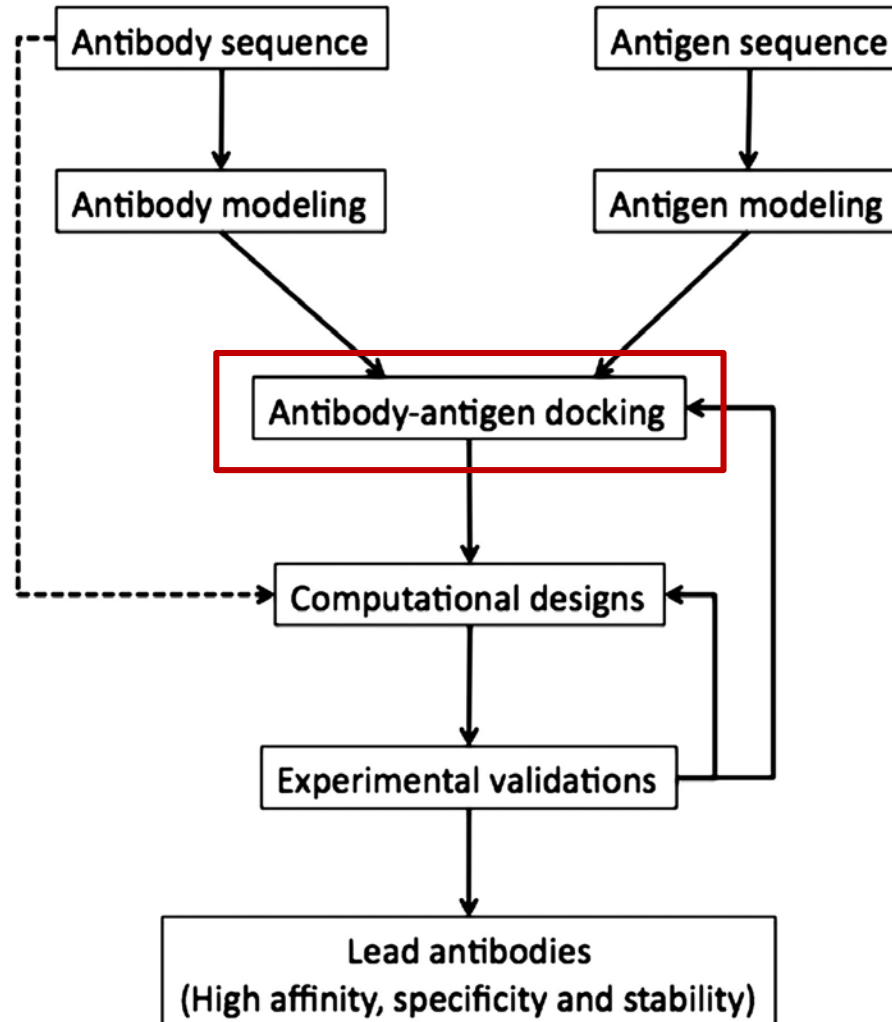
Structure 3D View



Download PDB model file:

[Download](#)

Computational antibody design

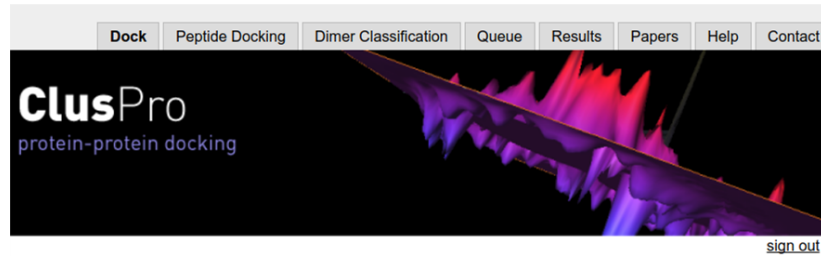


Kuroda D *et al*, PEDS 2012.

ClusPro

<https://cluspro.bu.edu>

- Consistently successful in CAPRI experiments (Critical Assessment of Protein Interactions)



Dock

Note: all jobs by non logged in users will be publicly accessible. Please create an account if data is embargoed and needs to remain confidential

Job Name:

Server:

Accepted PDB Input:
20 standard amino acids and RNA (as receptor only), ref: RNA Select Heparin Mode to use Heparin as Ligand.

Receptor	Ligand
PDB ID: <input type="text"/>	PDB ID: <input type="text"/>
Upload PDB	Upload PDB
Chains: <input type="text"/>	Chains: <input type="text"/>

Whitespace separate desired chains. Leave chains blank to use all chains.

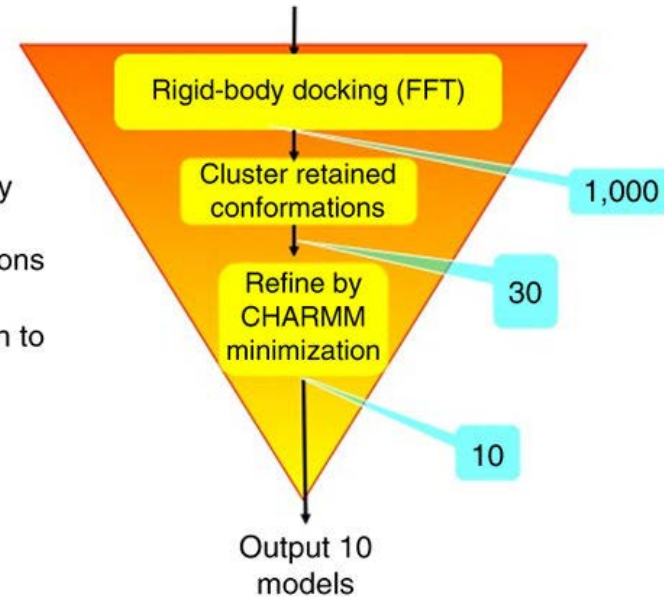
► **Advanced Options**

I agree to use ClusPro only for noncommercial purposes.

FFT-based global sampling on a grid using PIPER

Clustering to find highly populated clusters of low-energy conformations

CHARMM minimization to remove steric clashes



Kozakov et al, Nature Protocols 2017; Brenke et al, Bioinformatics 2012.

ClusPro

[sign out](#)

Dock

Note: all jobs by non logged in users will be publicly accessible. Please create an account if data is embargoed and needs to remain confidential

Job Name:

Server:

Accepted PDB Input:
20 standard amino acids and RNA (as receptor only), ref: [RNA](#) Select Heparin Mode to use Heparin as Ligand.

Receptor	Ligand
<input type="button" value="Choose File"/> mod...db	PDB ID: <input type="text" value="1n26"/>
Use PDB ID	Upload PDB
Chains: <input type="text"/>	Chains: <input type="text" value="A"/>

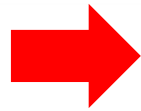
Whitespace separate desired chains. Leave chains blank to use all chains.

▶ Advanced Options



I agree to use ClusPro only for noncommercial purposes.

ClusPro



▼ **Advanced Options**

▶ **Attraction and Repulsion**

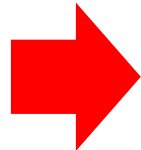
▶ **Structure Modification**

▶ **Multimer Docking (Beta Version)**

▼ **Antibody Mode**

Please set antibody as receptor and antigen as ligand.

It is additionally recommended to provide a receptor mask of the non-CDR regions.



- Use Antibody Mode
- Automatically Mask non-CDR regions

▶ **Others Mode**

▶ **Heparin Ligand**

▶ **Saxs Profile**

I agree to use ClusPro only for noncommercial purposes.



Dock

ClusPro

View Model Scores

Download all Models for all Coefficients

Antibody Mode

Display Models:

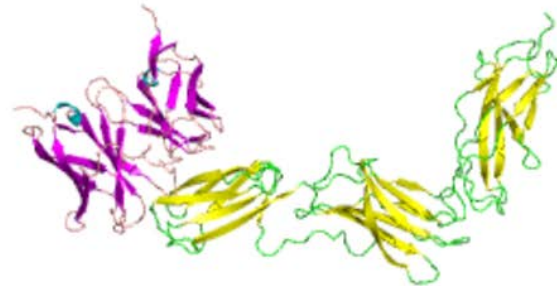
Download Displayed Models

If you use these models in a paper, please cite our papers

00

0

1



ClusPro

[View Models](#)

Balanced | [Electrostatic-favored](#) | [Hydrophobic-favored](#) | [VdW+Elec](#)

[Download Model Scores for this Coefficient](#)

Coefficient Weights

See *Kozakov et. al.* in [Papers](#) for a description of these terms

$$E = 0.40E_{rep} + -0.40E_{att} + 600E_{elec} + 1.00E_{DARS}$$

Cluster Scores

We strongly encourage you to read the [FAQ related to these scores](#) before using them.

Cluster	Members	Representative	Weighted Score
0	224	Center	-847.9
		Lowest Energy	-847.9
1	96	Center	-676.8
		Lowest Energy	-820.9
2	91	Center	-690.0
		Lowest Energy	-851.4
3	59	Center	-657.2
		Lowest Energy	-841.4
4	55	Center	-675.3
		Lowest Energy	-786.9
5	51	Center	-841.2
		Lowest Energy	-841.2
6	50	Center	-647.8
		Lowest Energy	-798.3

- Number of cluster members are used to get the best docked models
- Weighted scores should not be used to get the best models.

SnugDock

http://rosie.rosettacommons.org/snug_dock

Welcome to ROSIE
Rosetta Online Server that Includes Everyone

Welcome Queue About ChangeLog Documentation Support Login Create an account

Rosetta SnugDock Protocol

[\[Submit SnugDock task\]](#) [\[SnugDock Queue\]](#) [\[SnugDock Server Documentation\]](#)

Please cite the following article when referring to results from our ROSIE server:

1. A. Silcar & J. J. Gray, "SnugDock: Paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models," *PLoS Comput. Biol.* 6(1): e1000644 (2010). [Online](#)
2. Lyskov S., Gray J.J. "The RosettaDock server for local protein-protein docking" *Nucleic Acids Research* 36 (Web Server Issue), W233-W238 (2008). [Online](#)
3. Lyskov S, Chou FC, Conchúir SÓ, Der BS, Drew K, Kuroda D, Xu J, Weitzner BD, Renfrew PD, Sripakdeevong P, Borgo B, Havranek JJ, Kuhlman B, Kortemme T, Bonneau R, Gray JJ, Das R., "Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE)". *PLoS One*. 2013 May 22;8(5):e63906. doi: 10.1371/journal.pone.0063906. Print 2013. [Link](#)

Modeling tools developed by [GrayLab](#). The Rosie implementation was developed by Sergey Lyskov.

We welcome scientific and technical comments on our server. For support please contact us at [Rosetta Forums](#) with any comments, questions or concerns.

ROSIE is web front-end for [Rosetta software suite](#). Developed by Sergey Lyskov, [GrayLab at JHU](#). Copyright © 2013 Rosetta Commons Member Institutions.

- Uses RosettaDock
- Simultaneously optimizes
 - antibody-antigen position,
 - CDR loops conformation and
 - heavy and light chain relative position
- Drawback: Requires one or more days to complete job
- Assessment: Weitzner *et al*, Nature Protocols 2017.

SnugDock

Welcome to ROSIE

Rosetta Online Server that Includes Everyone

Welcome Queue About ChangeLog Documentation Support Login Create an account

Submit a new SnugDock job

Job short description (visible in queue):

Input PDB No file chosen

Docking partners list of chains separated by underscore:

SnugDock **protocol** to use:

Job Description (for your own records):

Submit

You have not logged in! If you already have a ROSIE account, please [login](#) to submit your job. Alternatively, you also can submit a job as Guest:

Keep my job-data public (Note that Public Jobs have higher priority and longer life time!)

If you decide to keep your job private and submit it as Guest, your ROSIE Job will be allocated an obscure URL (**anybody with a link will be able to access the results**). For better security please consider to [Create a ROSIE Account](#) and use it to submit private jobs.

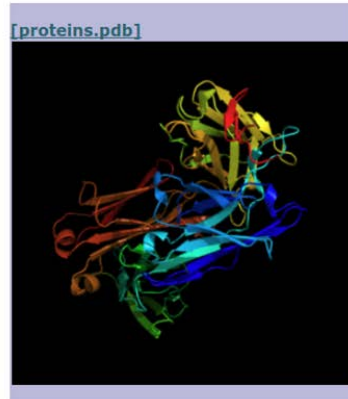
[optional] Notify me when my job completes by sending a mail to:

ROSIE is web front-end for [Rosetta software suite](#). Developed by Sergey Lyskov, [GrayLab at JHU](#). Copyright © 2013 Rosetta Commons Member Institutions.

SnugDock

SnugDock Job 2b1aclean 「№35368」 Details

Inputs

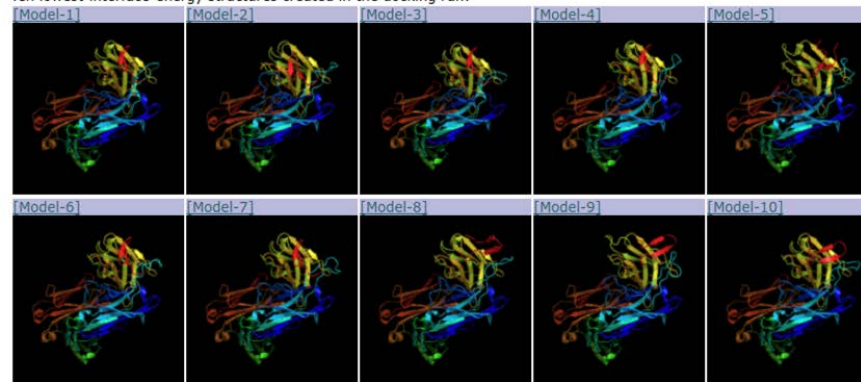


Status

Job ID	35368
Job Name	2b1aclean
Visibility	PUBLIC
Protocol	SnugDock
CPU hours used	75.2
user	jessye
Status	Finished
Description	
docking_partners	LH_P
protocol	thorough_snug_dock
Submitted time	2017-09-22 07:29
Start time	2017-09-22 08:22
End time	2017-09-22 18:21
Daemon	GrayLab.Jazz

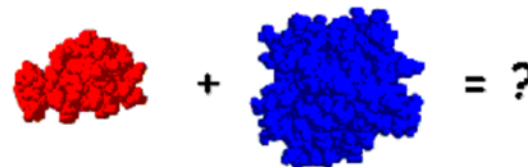
Results

Ten lowest-interface-energy structures created in the docking run:



<https://bioinfo3d.cs.tau.ac.il/PatchDock/>

PATCHDOCK



Molecular Docking Algorithm Based on Shape Complementarity Principles

[\[About PatchDock\]](#) [\[Web Server\]](#) [\[Download\]](#) [\[Help\]](#) [\[FAQ\]](#) [\[References\]](#)

Type PDB codes of receptor and ligand molecules or upload files in PDB format

Receptor Molecule:

(PDB:chainId e.g. 2kai:AB) **or** upload file:

Choose File

model_pigs_...41_dir.pdb



Ligand Molecule:

(PDB:chainId e.g. 2kai:I) **or** upload file:

Choose File

no file selected

e-mail address:

(the results are sent to this address)

Clustering RMSD:

Complex Type:

Be sure to give receptor and ligand in the corresponding order!



Submit Form

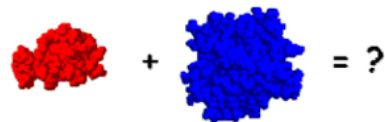
Clear

Advanced Options:

[\[Show\]](#)[\[Hide\]](#)

Schneidman-Duhovny *et al*, NAR 2005.

PATCHDOCK



Molecular Docking Algorithm Based on Shape Complementarity Principles

[\[About PatchDock\]](#) [\[Web Server\]](#) [\[Download\]](#) [\[Help\]](#) [\[FAQ\]](#) [\[References\]](#)

Receptor [model_pigs_de4241_dir.pdb](#) Ligand [1n26A](#) Complex Type [AA](#) Clustering RMSD [4.0](#) User e-mail [jpon@sdsc.edu](#) Receptor Site [CDR](#) Ligand Site [-](#) Distance Constraints [-](#)

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	13410	1900.90	202.79	-1.94 -1.08 2.05 24.73 107.88 63.00	result.1.pdb
2	12948	1731.60	-240.76	1.29 0.87 2.10 -10.48 70.22 38.19	result.2.pdb
3	12770	1655.80	-192.09	0.33 1.24 -2.98 -15.46 65.89 49.76	result.3.pdb
4	11584	1251.70	-26.96	-2.35 0.82 2.55 41.50 17.24 126.21	result.4.pdb
5	11570	1609.00	28.24	2.21 0.04 -1.09 76.26 77.02 40.59	result.5.pdb
6	11348	1407.10	-190.78	-1.57 0.12 -3.13 19.21 88.61 101.82	result.6.pdb
7	11300	1878.80	-225.56	2.48 -1.14 -0.04 49.28 97.02 55.91	result.7.pdb
8	11154	1813.50	-150.37	1.27 0.53 2.01 -10.16 68.70 31.29	result.8.pdb
9	10972	1479.60	47.91	1.17 0.53 2.07 -9.12 72.92 27.93	result.9.pdb
10	10872	1631.70	-316.73	-2.72 0.73 -1.53 27.15 81.70 102.14	result.10.pdb
11	10852	1512.40	-37.31	-1.25 -0.95 2.79 11.50 79.14 115.20	result.11.pdb
12	10848	1336.90	-95.71	-1.41 -0.56 2.68 14.57 79.48 123.92	result.12.pdb
13	10780	1503.70	-119.41	1.15 1.05 2.53 -14.96 65.76 41.33	result.13.pdb
14	10698	1717.60	13.31	3.09 1.53 1.37 -29.84 23.86 101.92	result.14.pdb
15	10696	1584.70	-180.44	0.64 1.01 2.79 -13.04 74.70 40.65	result.15.pdb
16	10680	1377.50	-140.37	-1.02 0.61 -1.48 0.49 60.66 55.36	result.16.pdb
17	10652	1590.40	-200.87	-1.32 1.12 -0.49 -24.09 20.10 109.61	result.17.pdb
18	10646	1502.70	-79.45	0.29 -0.81 -0.18 13.65 40.36 3.49	result.18.pdb
19	10600	1702.10	-319.87	-0.13 1.09 -2.08 -14.18 54.69 49.89	result.19.pdb
20	10484	1641.80	-65.84	-0.19 -0.33 2.38 46.01 63.28 81.06	result.20.pdb

[show next 20 >>>](#)

[NEW: Jmol view](#)



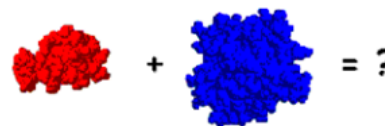
DOWNLOAD best solutions as a ZIP file: (solutions number, from 2 to 100) (this takes few seconds, please wait patiently)

DOWNLOAD [solutions table](#) [transformations file](#)

REFINE best solutions with FireDock: (solutions number, from 1 to 1000)

This option is recommended for protein-protein docking only!

PATCHDOCK



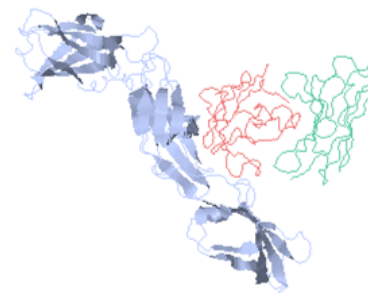
Molecular Docking Algorithm Based on Shape Complementarity Principles

[\[About PatchDock\]](#) [\[Web Server\]](#) [\[Download\]](#) [\[Help\]](#) [\[FAQ\]](#) [\[References\]](#)

Receptor	Ligand	Complex Type	Clustering RMSD	User e-mail	Receptor Site	Ligand Site	Distance Constraints
model_pigs_de4241_dir.pdb	1n26A	AA	4.0	jpon@sdsc.edu	CDR	-	-

Solution No	Score	Area	ACE	PDB file of the complex	show/hide
1	13410	1900.90	202.79	result.1.pdb	<input checked="" type="checkbox"/>
2	12948	1731.60	-240.76	result.2.pdb	<input type="checkbox"/>
3	12770	1655.80	-192.09	result.3.pdb	<input type="checkbox"/>
4	11584	1251.70	-26.96	result.4.pdb	<input type="checkbox"/>
5	11570	1609.00	28.24	result.5.pdb	<input type="checkbox"/>
6	11348	1407.10	-190.78	result.6.pdb	<input type="checkbox"/>
7	11300	1878.80	-225.56	result.7.pdb	<input type="checkbox"/>
8	11154	1813.50	-150.37	result.8.pdb	<input type="checkbox"/>
9	10972	1479.60	47.91	result.9.pdb	<input type="checkbox"/>
10	10872	1631.70	-316.73	result.10.pdb	<input type="checkbox"/>
11	10852	1512.40	-37.31	result.11.pdb	<input type="checkbox"/>
12	10848	1336.90	-95.71	result.12.pdb	<input type="checkbox"/>
13	10780	1503.70	-119.41	result.13.pdb	<input type="checkbox"/>
14	10698	1717.60	13.31	result.14.pdb	<input type="checkbox"/>
15	10696	1584.70	-180.44	result.15.pdb	<input type="checkbox"/>
16	10680	1377.50	-140.37	result.16.pdb	<input type="checkbox"/>
17	10652	1590.40	-200.87	result.17.pdb	<input type="checkbox"/>
18	10646	1502.70	-79.45	result.18.pdb	<input type="checkbox"/>
19	10600	1702.10	-319.87	result.19.pdb	<input type="checkbox"/>
20	10484	1641.80	-65.84	result.20.pdb	<input type="checkbox"/>

show all/hide all



[show next 20 >>>](#)

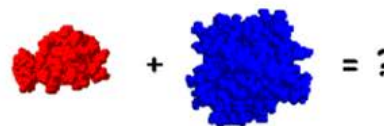
Jmol

DOWNLOAD best solutions as a ZIP file: (solutions number, from 2 to 100) (this takes few seconds, please wait patiently)

DOWNLOAD [solutions table](#) [transformations file](#)

REFINE best solutions with FireDock: (solutions number, from 1 to 1000)

PATCHDOCK



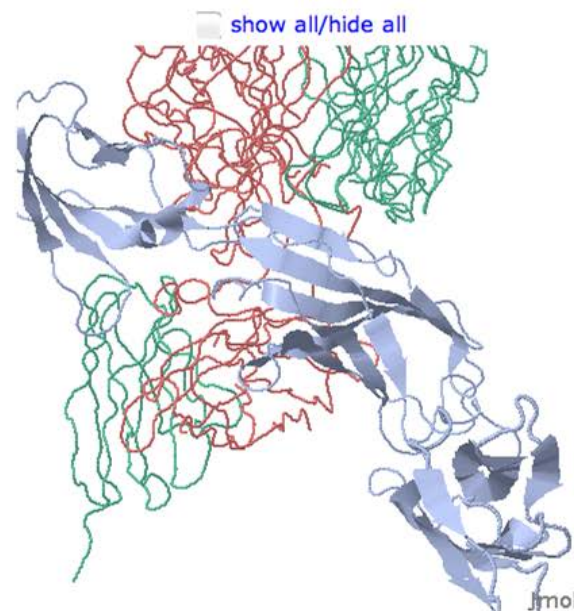
Molecular Docking Algorithm Based on Shape Complementarity Principles

[\[About PatchDock\]](#) [\[Web Server\]](#) [\[Download\]](#) [\[Help\]](#) [\[FAQ\]](#) [\[References\]](#)

Receptor [model_pigs_de4241_dir.pdb](#) Ligand [1n26A](#) Complex Type [AA](#) Clustering RMSD [4.0](#) User e-mail [jpon@sdsc.edu](#) Receptor Site [CDR](#) Ligand Site [-](#) Distance Constraints [-](#)

Solution No	Score	Area	ACE	PDB file of the complex	show/hide
1	13410	1900.90	202.79	result.1.pdb	<input checked="" type="checkbox"/>
2	12948	1731.60	-240.76	result.2.pdb	<input checked="" type="checkbox"/>
3	12770	1655.80	-192.09	result.3.pdb	<input checked="" type="checkbox"/>
4	11584	1251.70	-26.96	result.4.pdb	<input type="checkbox"/>
5	11570	1609.00	28.24	result.5.pdb	<input type="checkbox"/>
6	11348	1407.10	-190.78	result.6.pdb	<input type="checkbox"/>
7	11300	1878.80	-225.56	result.7.pdb	<input type="checkbox"/>
8	11154	1813.50	-150.37	result.8.pdb	<input type="checkbox"/>
9	10972	1479.60	47.91	result.9.pdb	<input type="checkbox"/>
10	10872	1631.70	-316.73	result.10.pdb	<input type="checkbox"/>
11	10852	1512.40	-37.31	result.11.pdb	<input type="checkbox"/>
12	10848	1336.90	-95.71	result.12.pdb	<input type="checkbox"/>
13	10780	1503.70	-119.41	result.13.pdb	<input type="checkbox"/>
14	10698	1717.60	13.31	result.14.pdb	<input type="checkbox"/>
15	10696	1584.70	-180.44	result.15.pdb	<input type="checkbox"/>
16	10680	1377.50	-140.37	result.16.pdb	<input type="checkbox"/>
17	10652	1590.40	-200.87	result.17.pdb	<input type="checkbox"/>
18	10646	1502.70	-79.45	result.18.pdb	<input type="checkbox"/>
19	10600	1702.10	-319.87	result.19.pdb	<input type="checkbox"/>
20	10484	1641.80	-65.84	result.20.pdb	<input type="checkbox"/>

[show next 20 >>](#)

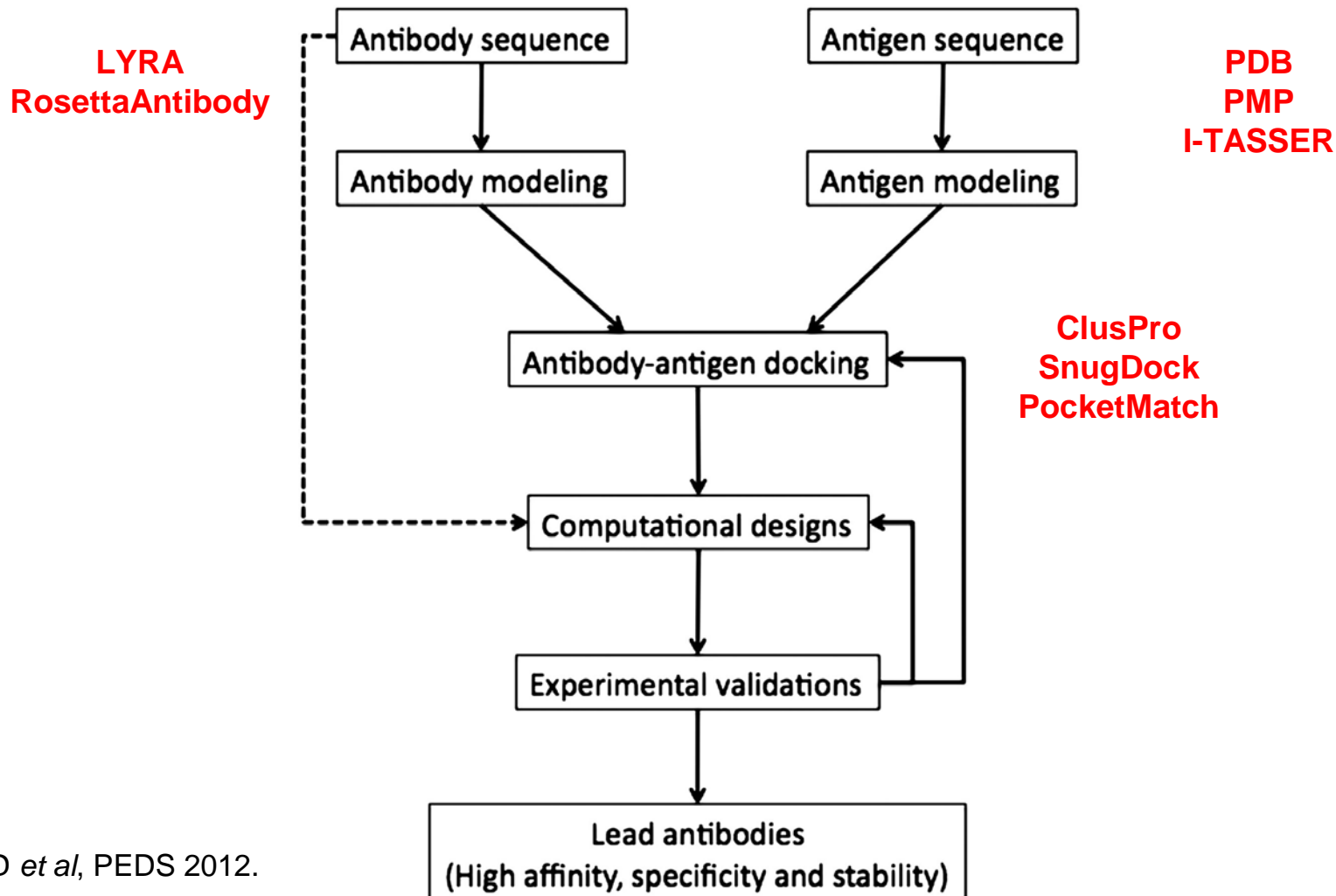


DOWNLOAD best solutions as a ZIP file: (solutions number, from 2 to 100) (this takes few seconds, please wait patiently)

DOWNLOAD [solutions table](#) [transformations file](#)

REFINE best solutions with FireDock: (solutions number, from 1 to 1000)

Computational antibody design



Kuroda D *et al*, PEDS 2012.



IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE

Thank you!

Questions?