IEDB
Immune Epitope Database & Tools

B Cell Epitope Prediction

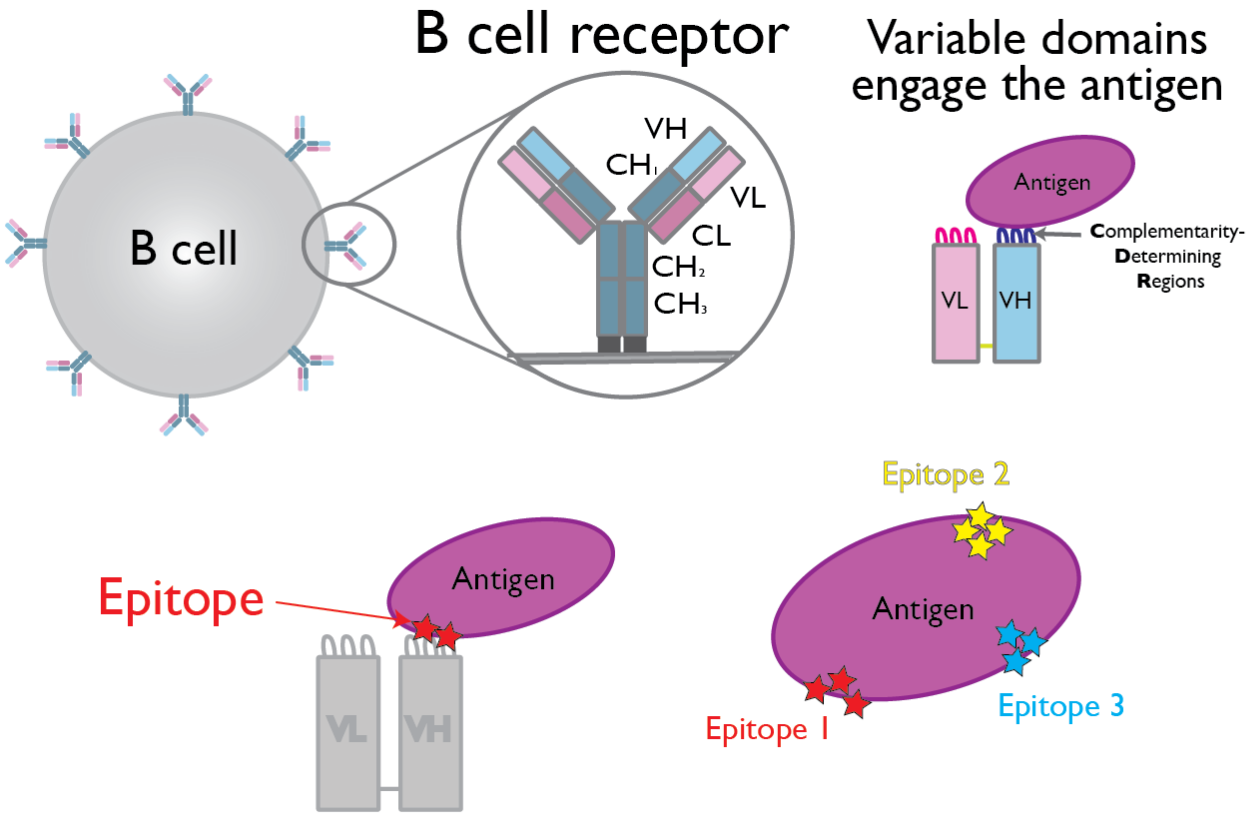
tools.iedb.org

Presented by: Eve Richardson, PhD
Postdoctoral Researcher

Outline of Topics

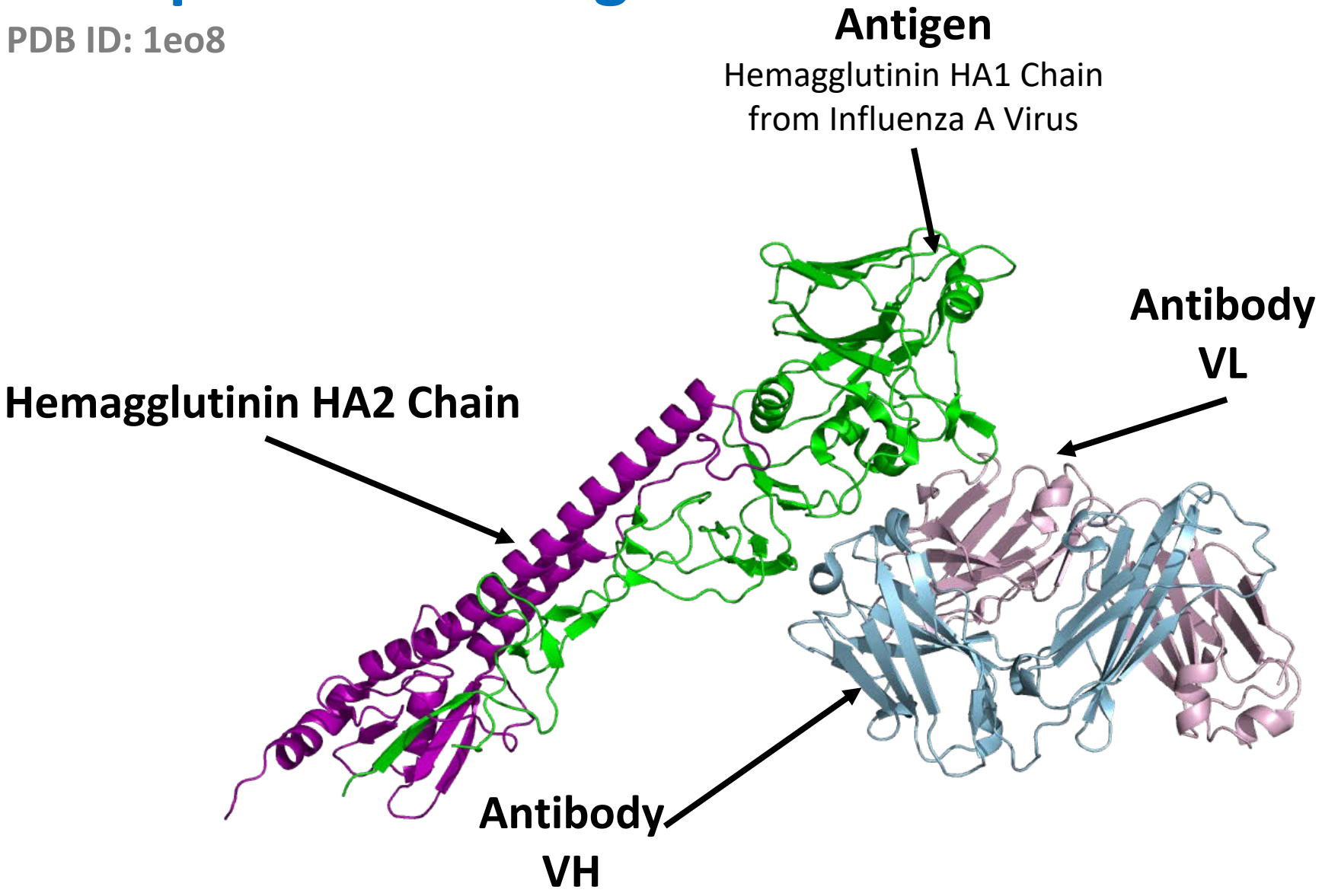
1. B cell epitope biology recap
2. Prediction tools on IEDB
3. Linear sequence-based epitope prediction methods
4. Discontinuous 3D structure-based epitope prediction methods

B Cell Epitopes



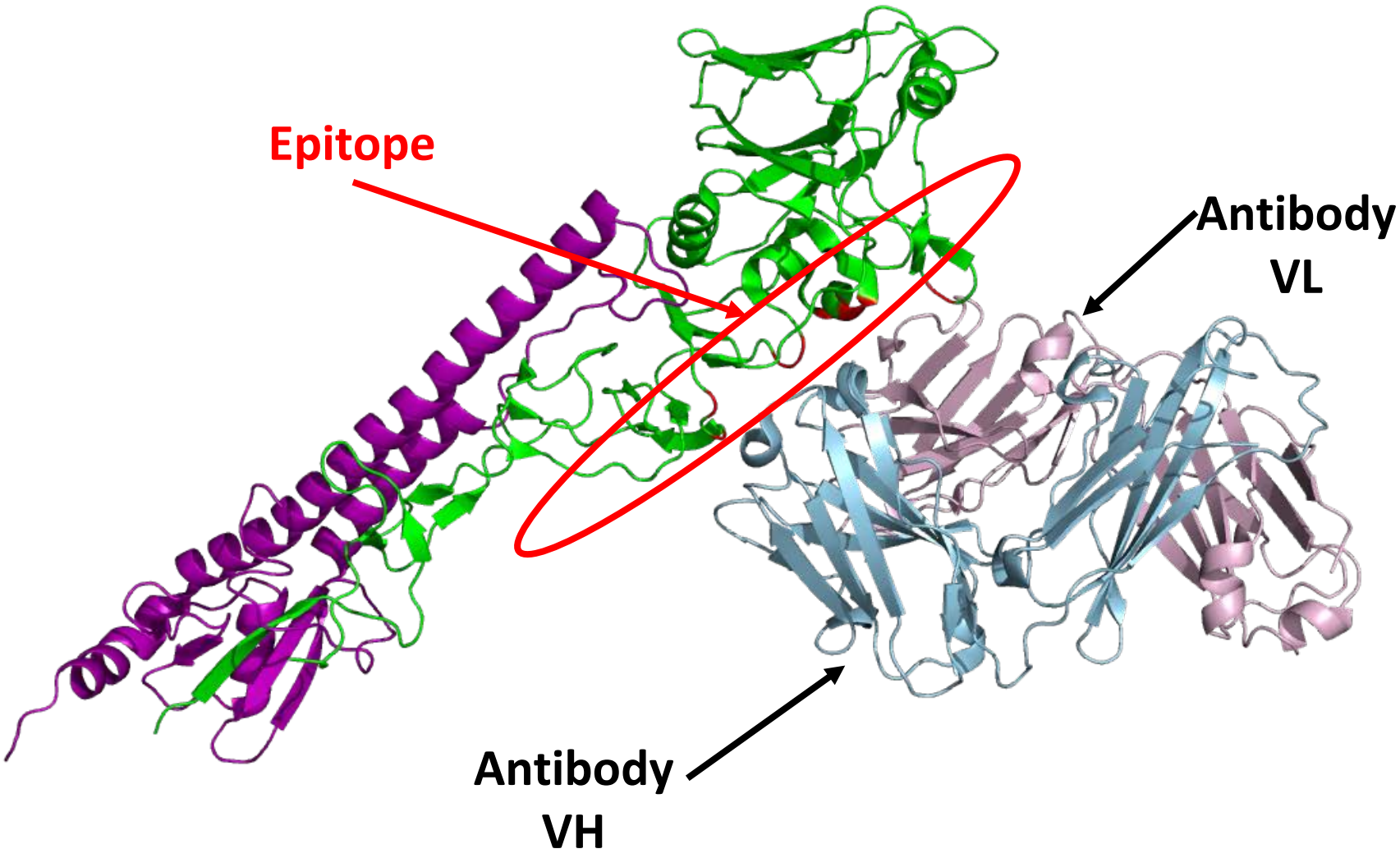
Example: Ab Binding HA1

PDB ID: 1eo8



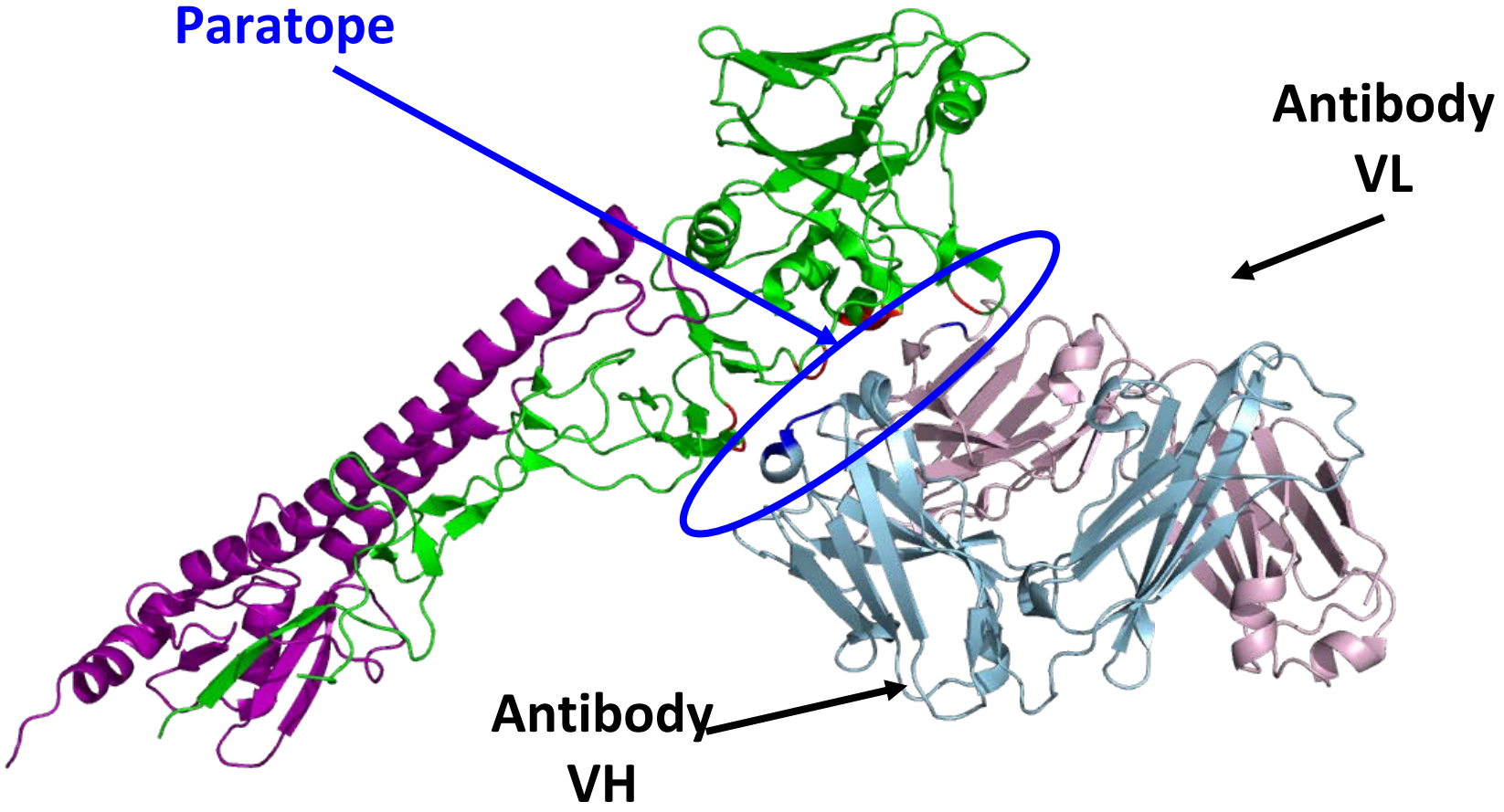
Example: Ab Binding HA1

PDB ID: 1eo8

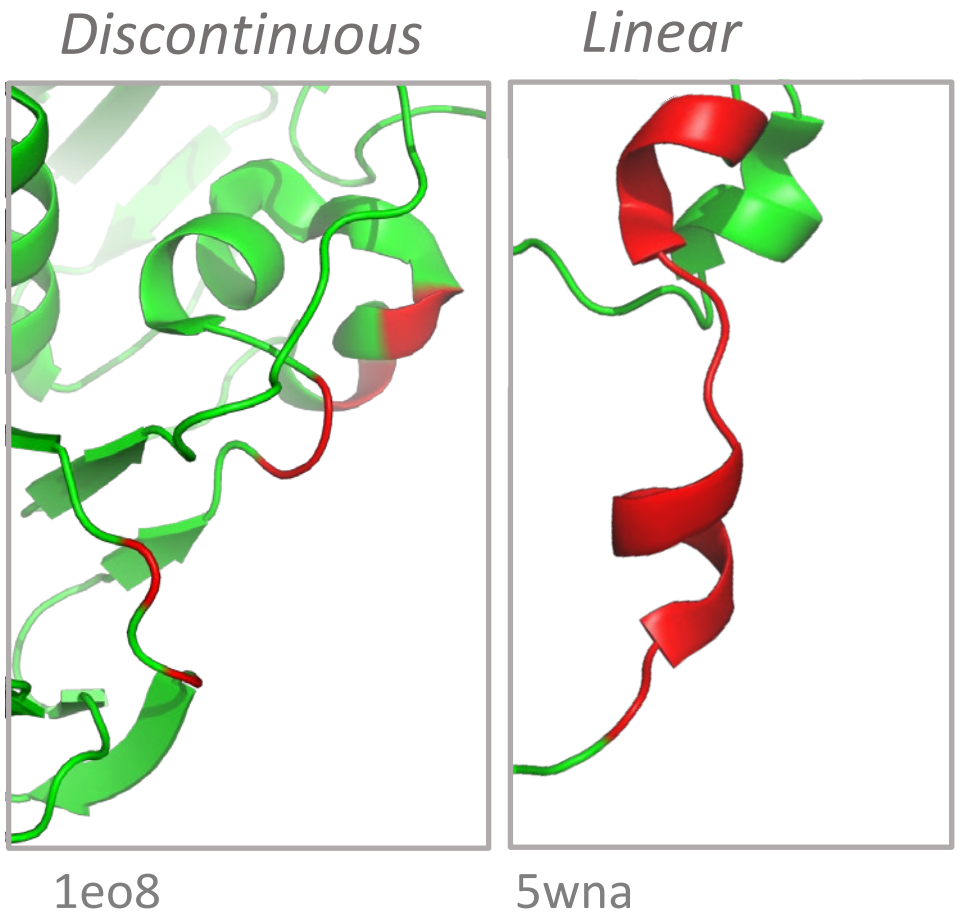


Example: Ab Binding HA1

PDB ID: 1eo8



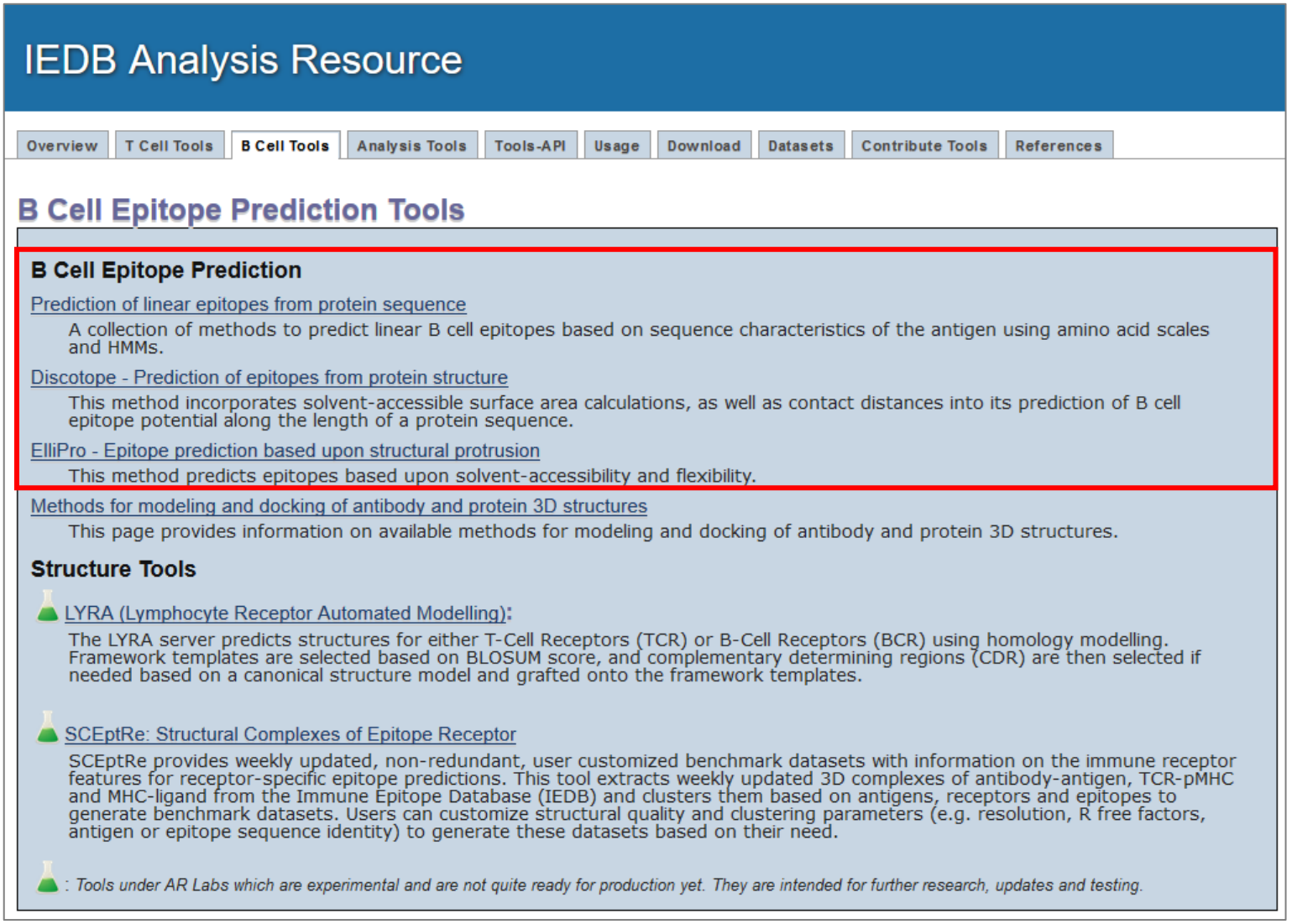
Linear and Discontinuous Epitopes



- Protein antigens usually contain both linear & discontinuous epitopes
  - "Linear" aka sequential or continuous
  - "Discontinuous" aka non-sequential or conformational epitopes
- More than 90% of B cell epitopes are estimated to be discontinuous

Barlow et al, Nature. 1986.
Van Regenmortel, Methods. 1996.

B Cell Epitope Prediction Tools in the IEDB



http://tools.iedb.org/main/bcell/

B Cell Epitope Prediction

When to use epitope prediction methods?

- You have verified thoroughly that no information is available in the IEDB on the antigen of your interest
- You want to know all the candidate antigenic determinants in an antigen of your interest other than epitopes provided in the IEDB

Types of B Cell Epitope Prediction

Sequence-based epitope prediction

- 3D solved structure or structural model of antigen is not required
- Relies on prediction methods for surface accessibility

Structure-based epitope prediction

- 3D solved structure or structural model of antigen is available
- Surface accessibility and secondary structure are known

Sequence-based Epitope Prediction – Amino Acid Scale Methods

Amino acid scales combined over windows of the sequence

- β-Turns (*Chou & Fasman*)
- Hydrophilicity (*Parker*)
- Flexibility (*Karplus & Schulz*)
- Surface Accessibility (*Emini*)
- Antigenic propensity: occurrence of residues in epitopes (*Kolaskar & Tongaonkar*)

Scale:

A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
2.1	1.4	10.0	7.8	-9.2	5.7	2.1	-8.0	5.7	-9.2	-4.2	7.0	2.1	6.0	4.2	6.5	5.2	-3.7	-10.0	-1.9

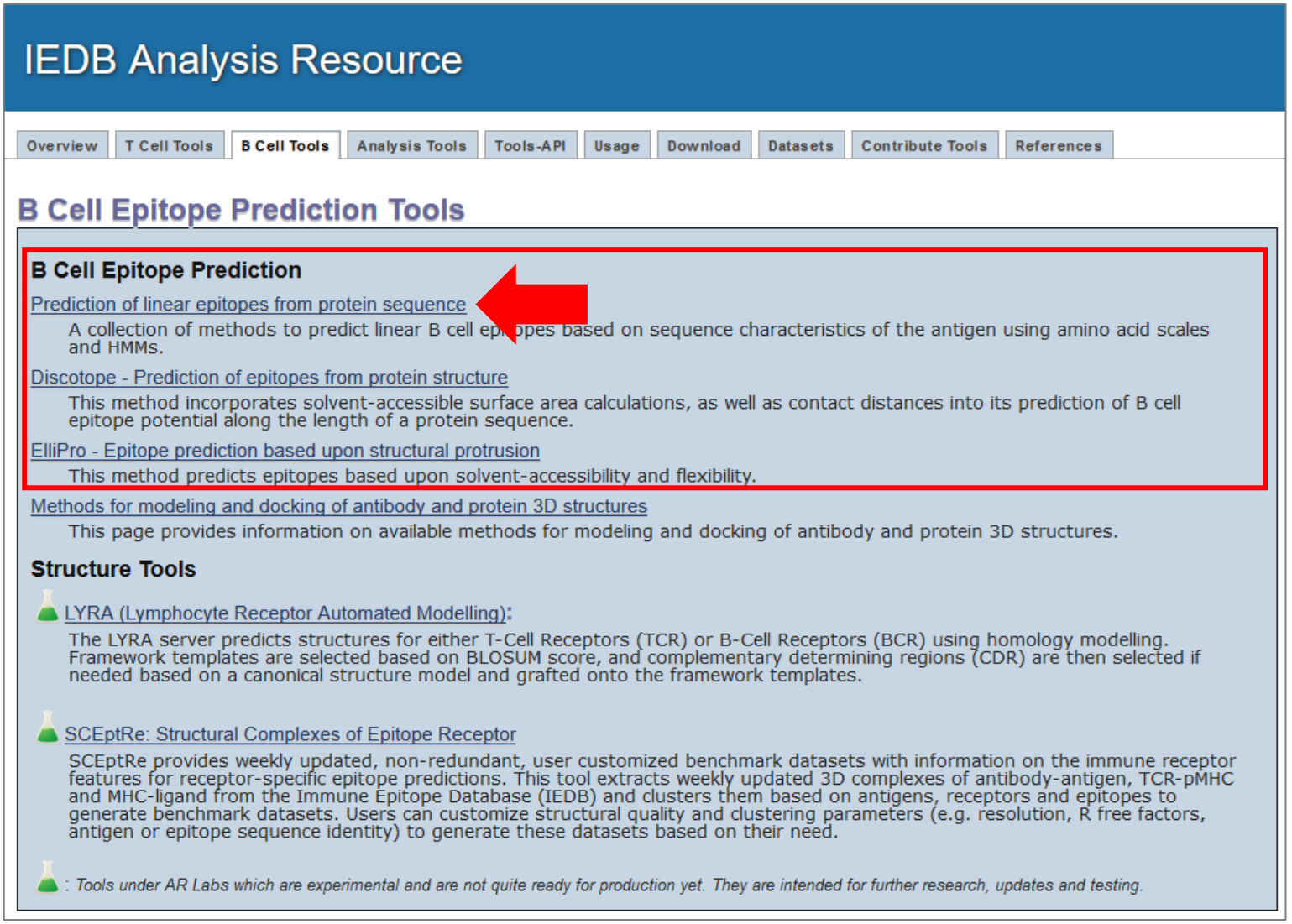
Example of an amino acid scale: *Parker Hydrophilicity Scale, which is averaged over seven residue windows.*

Sequence-based Epitope Prediction – Machine Learning Methods

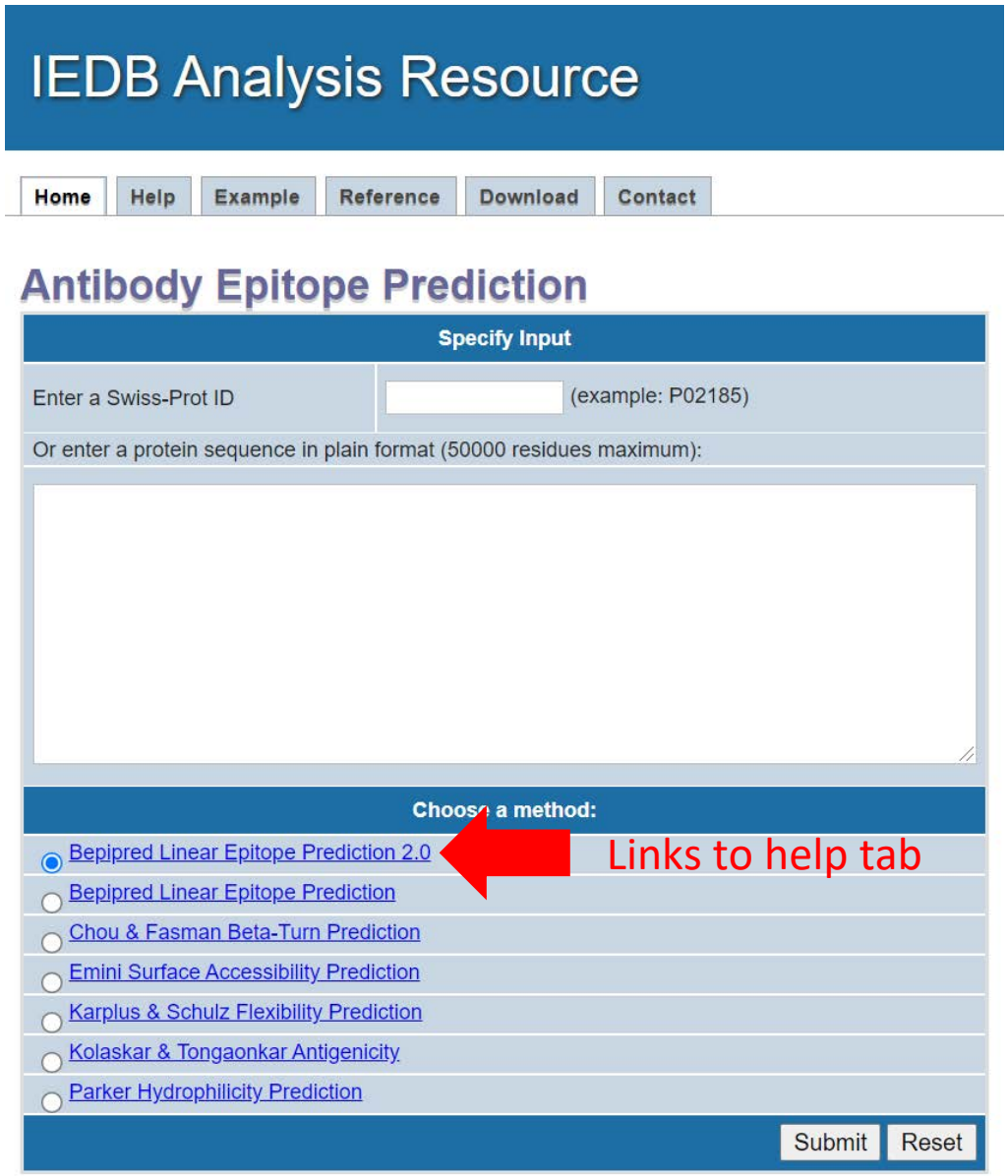
- Positive and negative training datasets are used
- Combination of one or more amino acid scales are used as an input to one of the machine learning algorithms
  - Random Forest (**BepiPred-2.0**)
  - ANN: Artificial Neural Network (ABCpred)
  - SVM: Support Vector Machine (BCpred, FBCpred)
- The model learns to distinguish epitope and non-epitope regions of the sequence and outputs a score for each region

B Cell Prediction Tools on IEDB

http://tools.iedb.org/main/bcell/



Linear B Cell Prediction



http://tools.iedb.org/bcell/

1. Input protein sequence

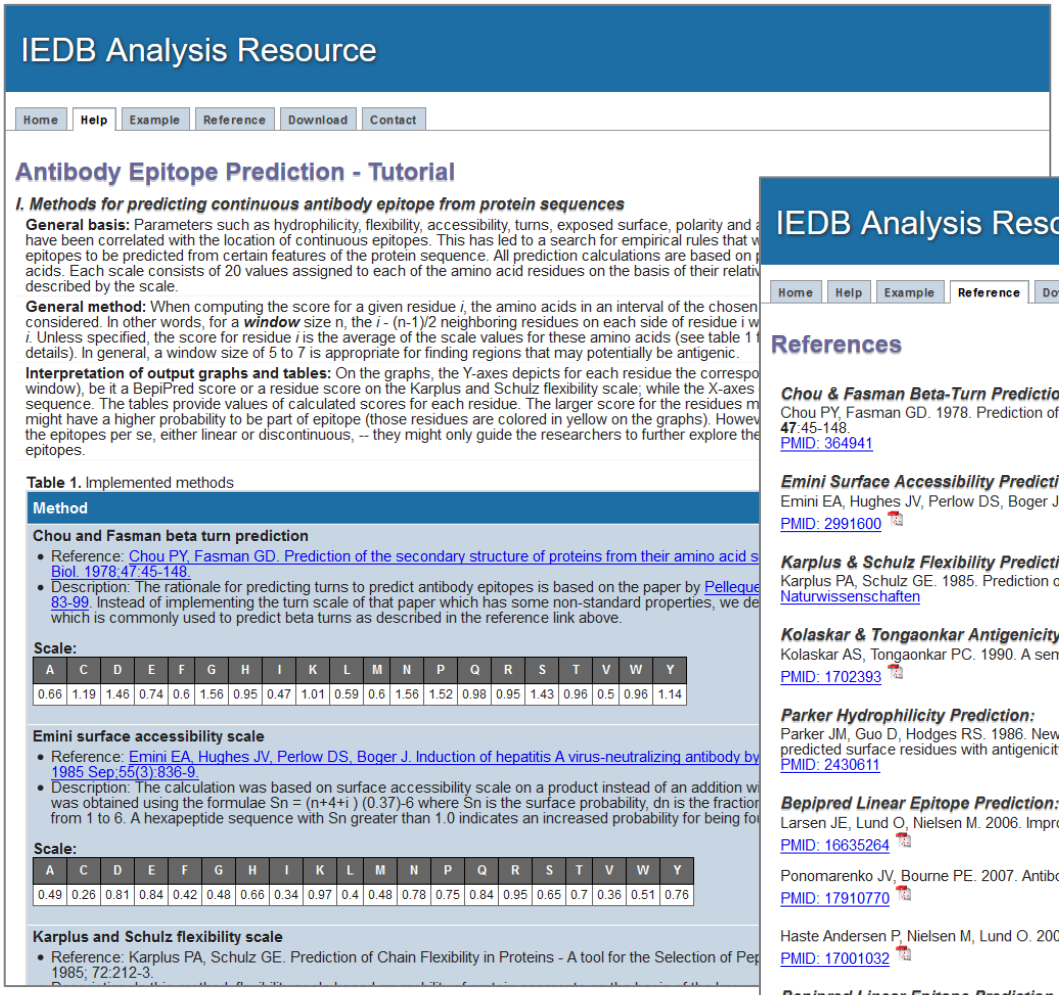
Entry allowed via Swiss-Prot ID or plain format

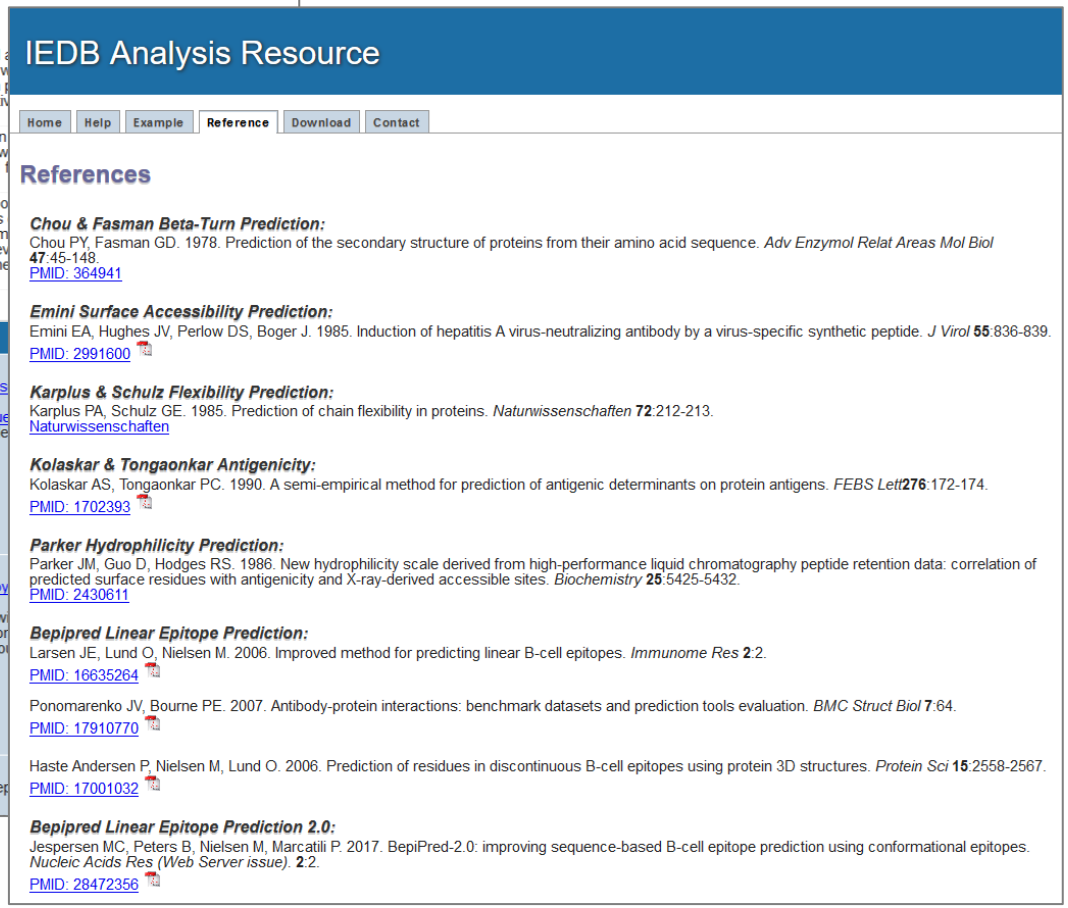
2. Select prediction method

BepiPred is the default & recommended method

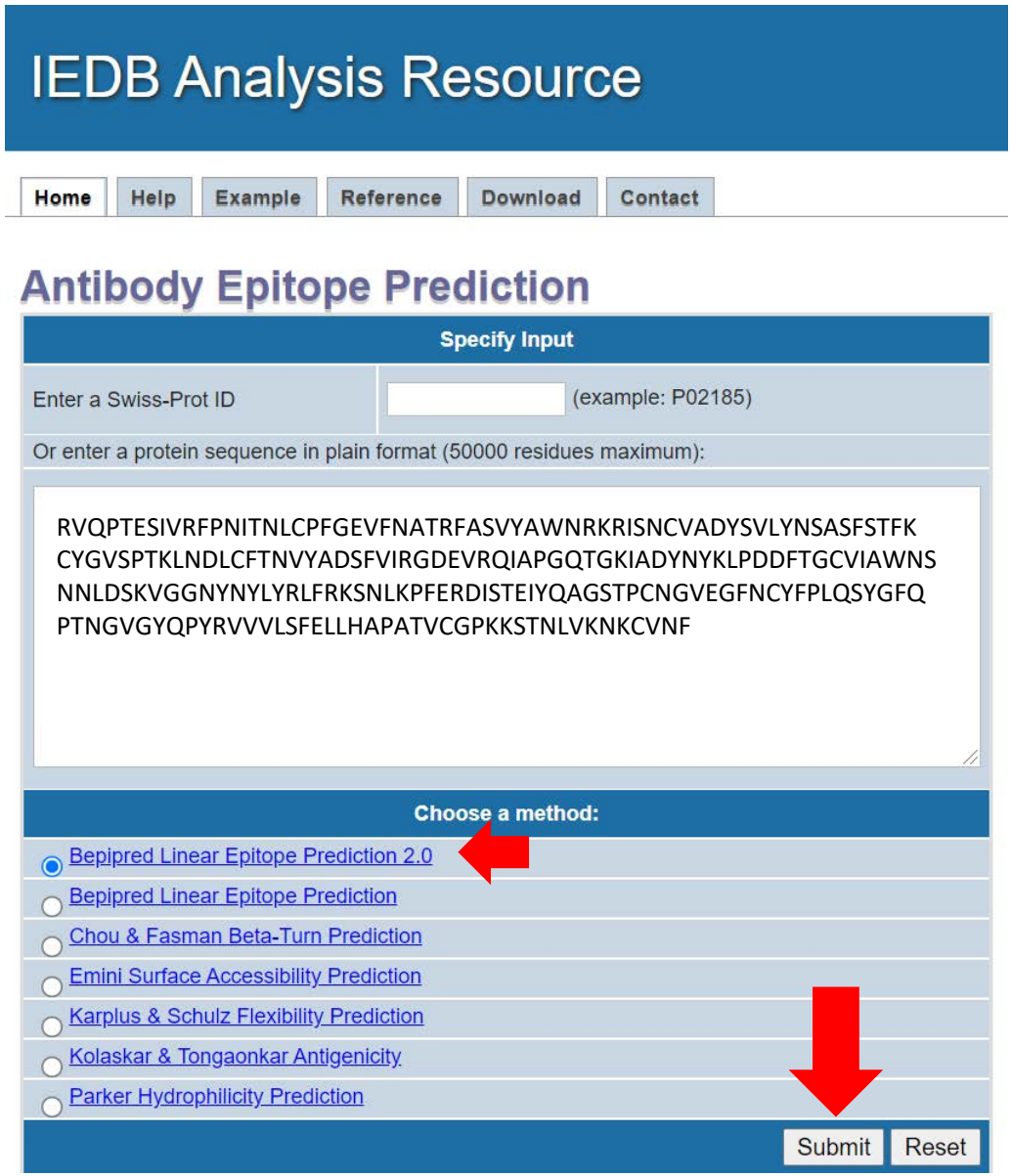
Visit Help & Reference Tabs to learn about a prediction method

http://tools.iedb.org/bcell/





Linear B Cell Prediction – Example



http://tools.iedb.org/bcell/

Example Sequence:

RBD region from SARS-CoV-2 Spike glycoprotein

Swiss-Prot ID: P0DTC2

Linear B Cell Prediction – Example

http://tools.iedb.org/bcell/

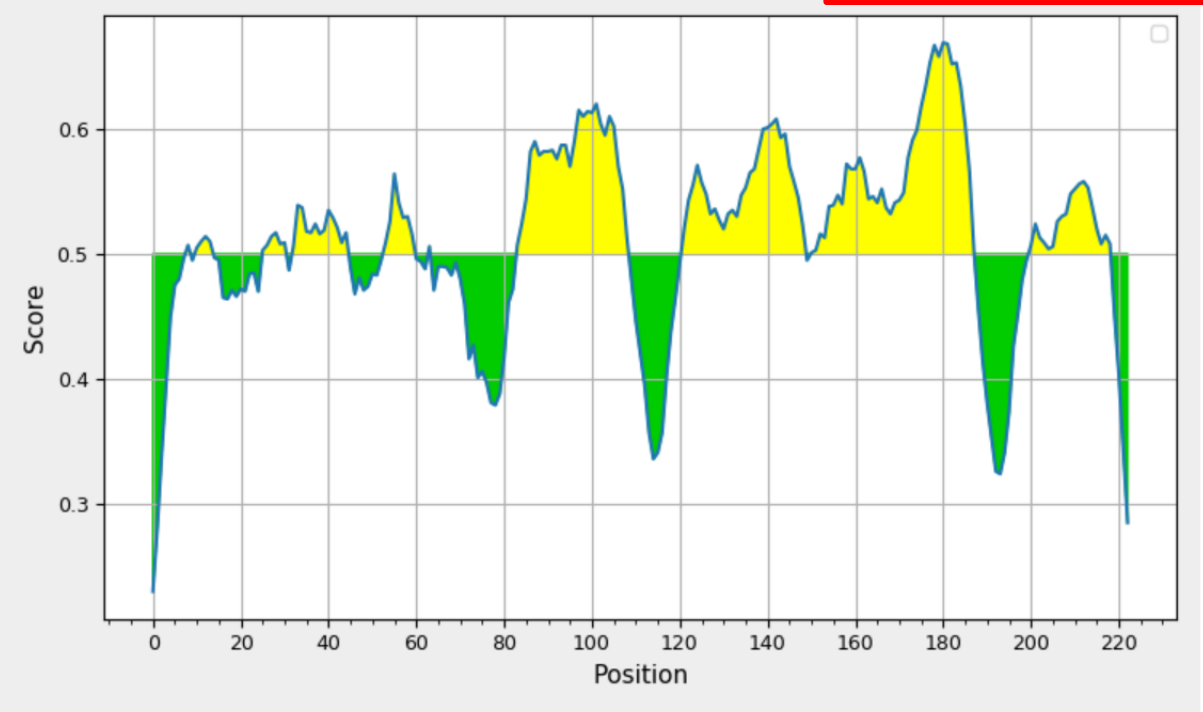
Bepipred Linear Epitope Prediction 2.0 Results

Input Sequences

```
  1 RVQPTESIVR FPNITNLCPF GEVFNATRFA SVYAWNRKRI SNCVADYSVL YNSASFSTFK
 61 CYGVSPTK LNDLCFTNVY ADSFVIRGDE VRQIAPGQTG KIADYNYKLP DDFTGCVIAW
121 NS NNLDSK VGGNYNYLYR LFRKSNLKPF ERDISTEIYQ AGSTPCNGVE GFNCYFPLQS
181 YGFQ PTNG VGYQPYRVVV LSFELLHAPA TVCGPKKSTN LVKNKCVNF
```

Center position: 4 **Threshold:** 0.500 Recalculate

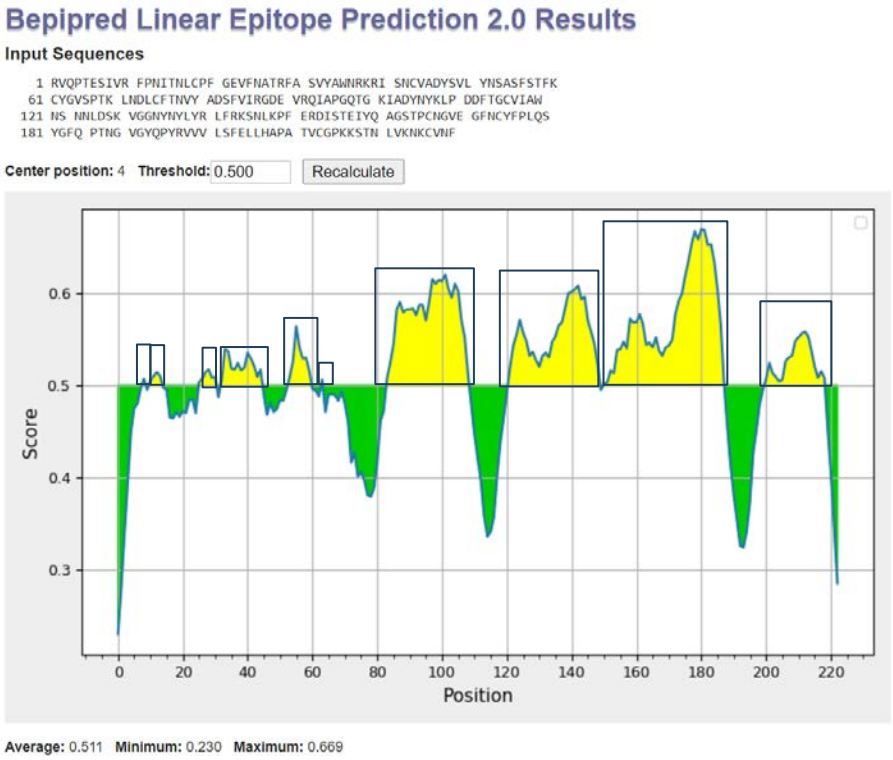
Adjustable threshold



Average: 0.511 **Minimum:** 0.230 **Maximum:** 0.669

Linear B Cell Prediction – Example

http://tools.iedb.org/bcell/



Predicted peptides

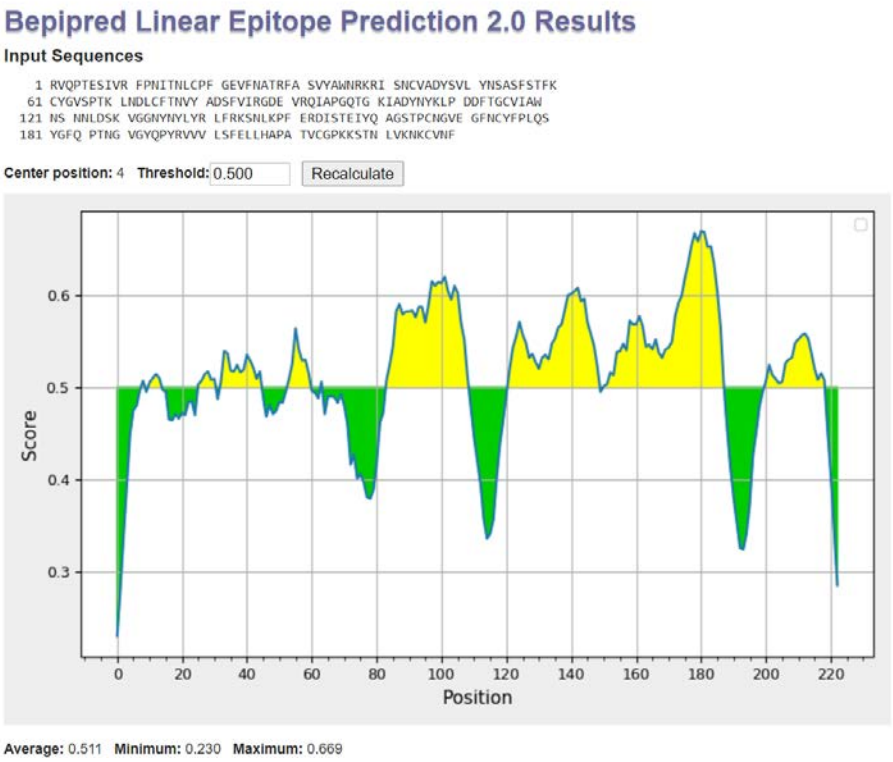
Contiguous sequence segments with scores above the threshold

Predicted peptides:

No.	Start	End	Peptide	Length
1	9	9	V	1
2	11	14	FPNI	4
3	26	31	ATRFAS	6
4	33	45	YAWNRKRISNCVA	13
5	54	60	ASFSTFK	7
6	64	64	V	1
7	84	109	IRGDEVRQIAPGQTGKIADYNYKLPD	26
8	122	149	NLDSKVGGNYNYLYRLFRKSNLKPFERD	28
9	151	188	STEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ	38
10	201	219	HAPATVCGPKKSTNLVKNK	19

Linear B Cell Prediction – Example

http://tools.iedb.org/bcell/



Predicted residue scores

Score for each residue in the input sequence

Predicted residue scores:

Position	Residue	Score	Assignment
0	R	0.230	.
1	V	0.280	.
2	Q	0.339	.
3	P	0.394	.
4	T	0.449	.
5	E	0.475	.
6	S	0.480	.
7	I	0.496	.
8	V	0.507	E
9	R	0.495	.
10	F	0.505	E
11	R	0.510	E

E = Epitope

Download result

Download as a CSV file

Method Comparisons

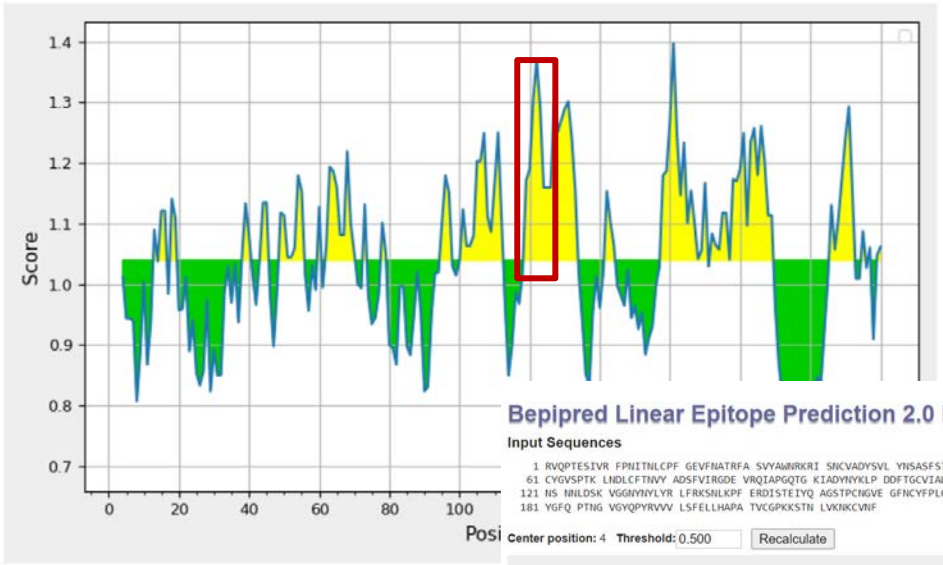
http://tools.iedb.org/bcell/

Chou & Fasman Beta-Turn Prediction Results

Input Sequences

```
  1 RVQPTESIVR FPNITNLCPF GEVFNATRFA SVYAWNRKRI SNCVADYSVL YNSASFSTFK
 61 CYGVSPTK LNDLCFTNVY ADSFVIRGDE VRQIAPGQTG KIADYNYKLP DDFTGCVIAW
121 NS NNLDSK VGGNYNYLYR LFRKSNLKPF ERDISTEIYQ AGSTPCNGVE GFNCYFPLQS
181 YGFQ PTNG VGYQPYRVVV LSFELLHAPA TVCGPKKSTN LVKNKCVNF
```

Center position: 4 Window size: 7 Threshold: 1.039 Recalculate



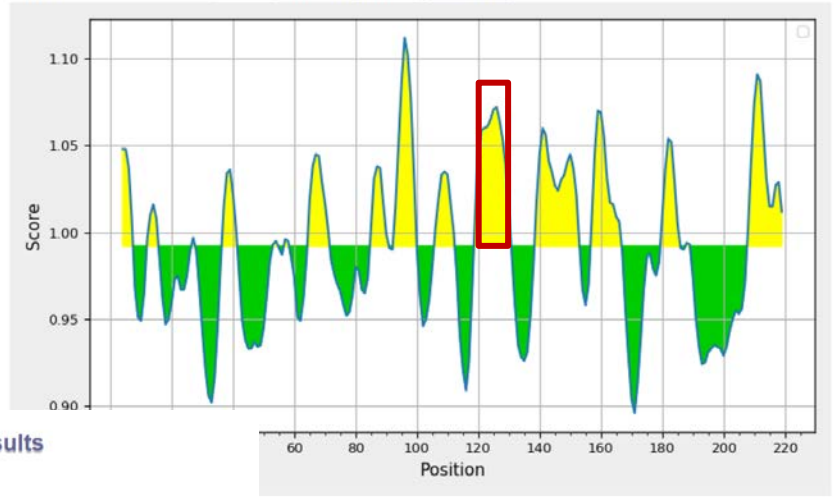
Average: 1.039 **Minimum:** 0.694 **Maximum:** 1.397

Karplus & Schulz Flexibility Prediction Results

Input Sequences

```
  1 RVQPTESIVR FPNITNLCPF GEVFNATRFA SVYAWNRKRI SNCVADYSVL YNSASFSTFK
 61 CYGVSPTK LNDLCFTNVY ADSFVIRGDE VRQIAPGQTG KIADYNYKLP DDFTGCVIAW
121 NS NNLDSK VGGNYNYLYR LFRKSNLKPF ERDISTEIYQ AGSTPCNGVE GFNCYFPLQS
181 YGFQ PTNG VGYQPYRVVV LSFELLHAPA TVCGPKKSTN LVKNKCVNF
```

Center position: 4 Window size: 7 Threshold: 0.992 Recalculate



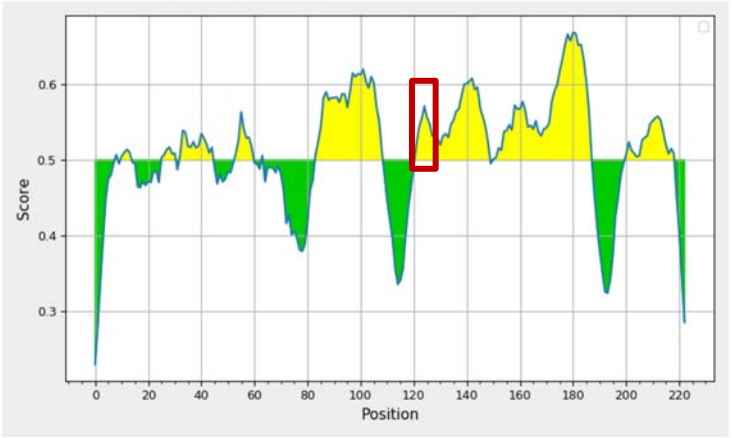
Average: 1.112

Bepipred Linear Epitope Prediction 2.0 Results

Input Sequences

```
  1 RVQPTESIVR FPNITNLCPF GEVFNATRFA SVYAWNRKRI SNCVADYSVL YNSASFSTFK
 61 CYGVSPTK LNDLCFTNVY ADSFVIRGDE VRQIAPGQTG KIADYNYKLP DDFTGCVIAW
121 NS NNLDSK VGGNYNYLYR LFRKSNLKPF ERDISTEIYQ AGSTPCNGVE GFNCYFPLQS
181 YGFQ PTNG VGYQPYRVVV LSFELLHAPA TVCGPKKSTN LVKNKCVNF
```

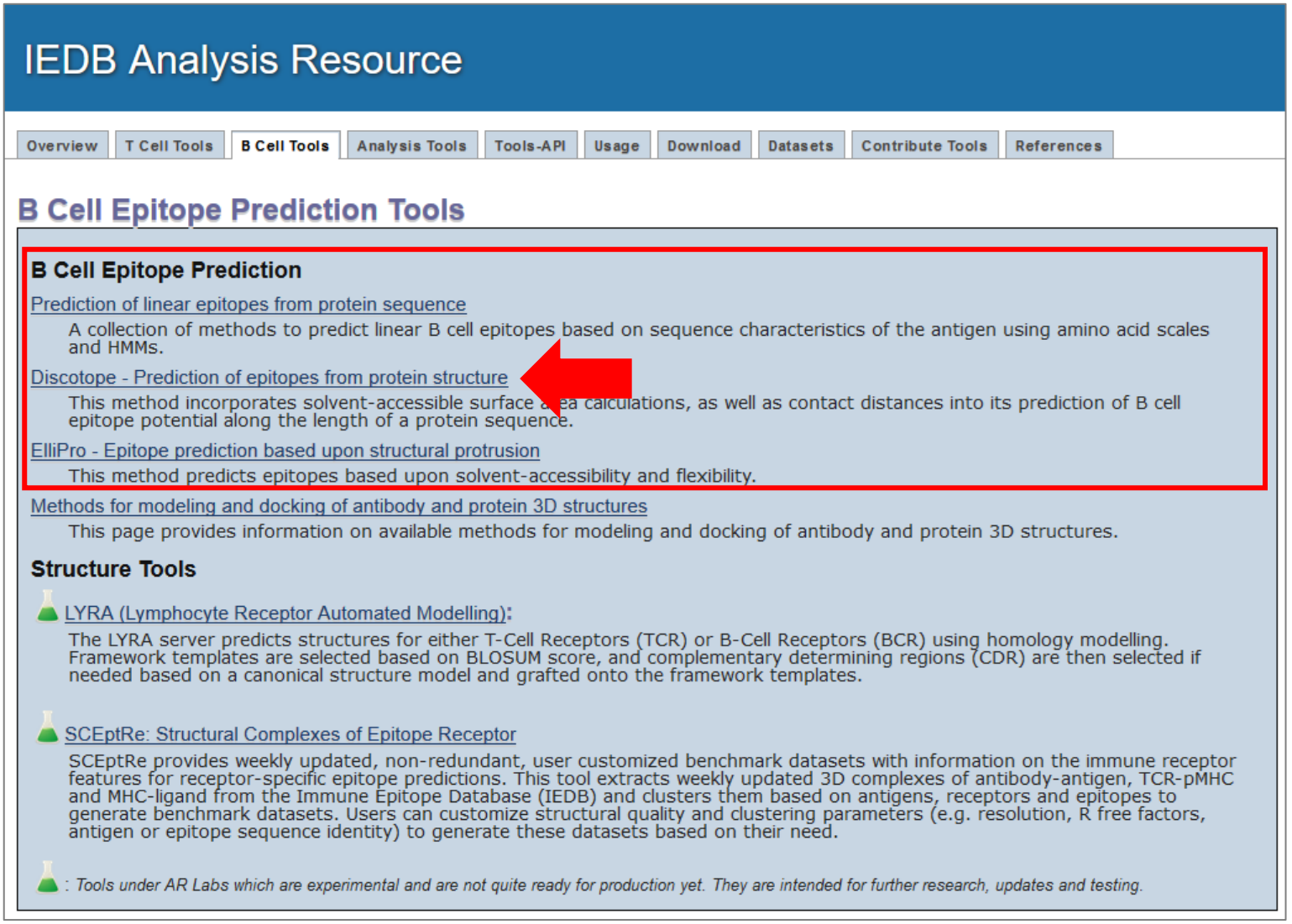
Center position: 4 Threshold: 0.500 Recalculate



Average: 0.511 **Minimum:** 0.230 **Maximum:** 0.669

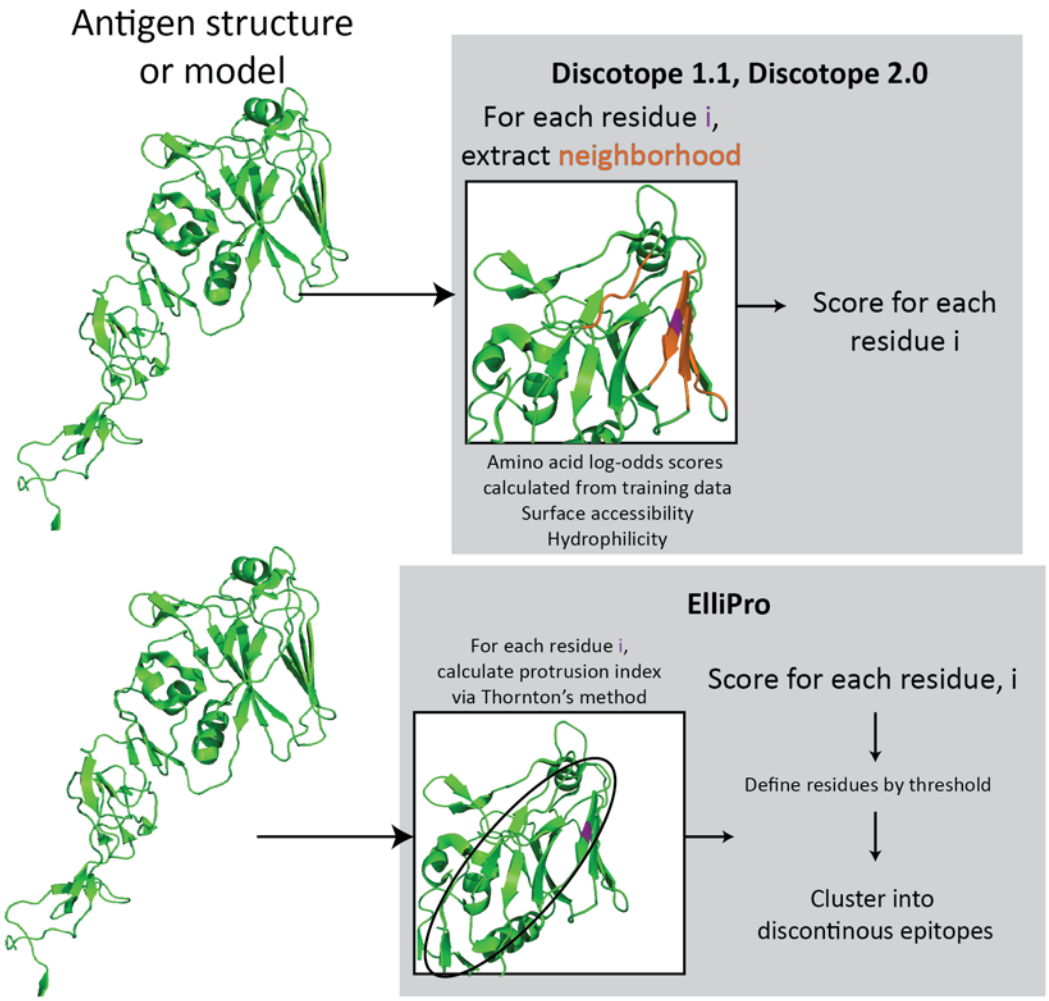
It is better to use **consensus** of different methods rather than relying on a single method

B Cell Prediction Tools on IEDB



http://tools.iedb.org/main/bcell/

3D Structure-based Epitope Prediction



DiscoTope

- 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
  - Parker's hydrophilicity scale
  - Log-odds ratio of amino acid in epitope vs. non-epitope
  - Number of contacts within 10Å (DiscoTope 1) or area of relative solvent accessibility (Discotope 2)

Protein Sci. 2006 Nov;15(11):2558-67. Epub 2006 Sep 25.

Prediction of residues in discontinuous B-cell epitopes using protein 3D structures.

Haste Andersen P[1], Nielsen M, Lund O.

PMID: 17001032 PMCID: PMC2242418 DOI: 10.1110/ps.062405906

DiscoTope 1

DiscoTope 2

PLoS Comput Biol. 2012;8(12):e1002829. doi: 10.1371/journal.pcbi.1002829. Epub 2012 Dec 27.

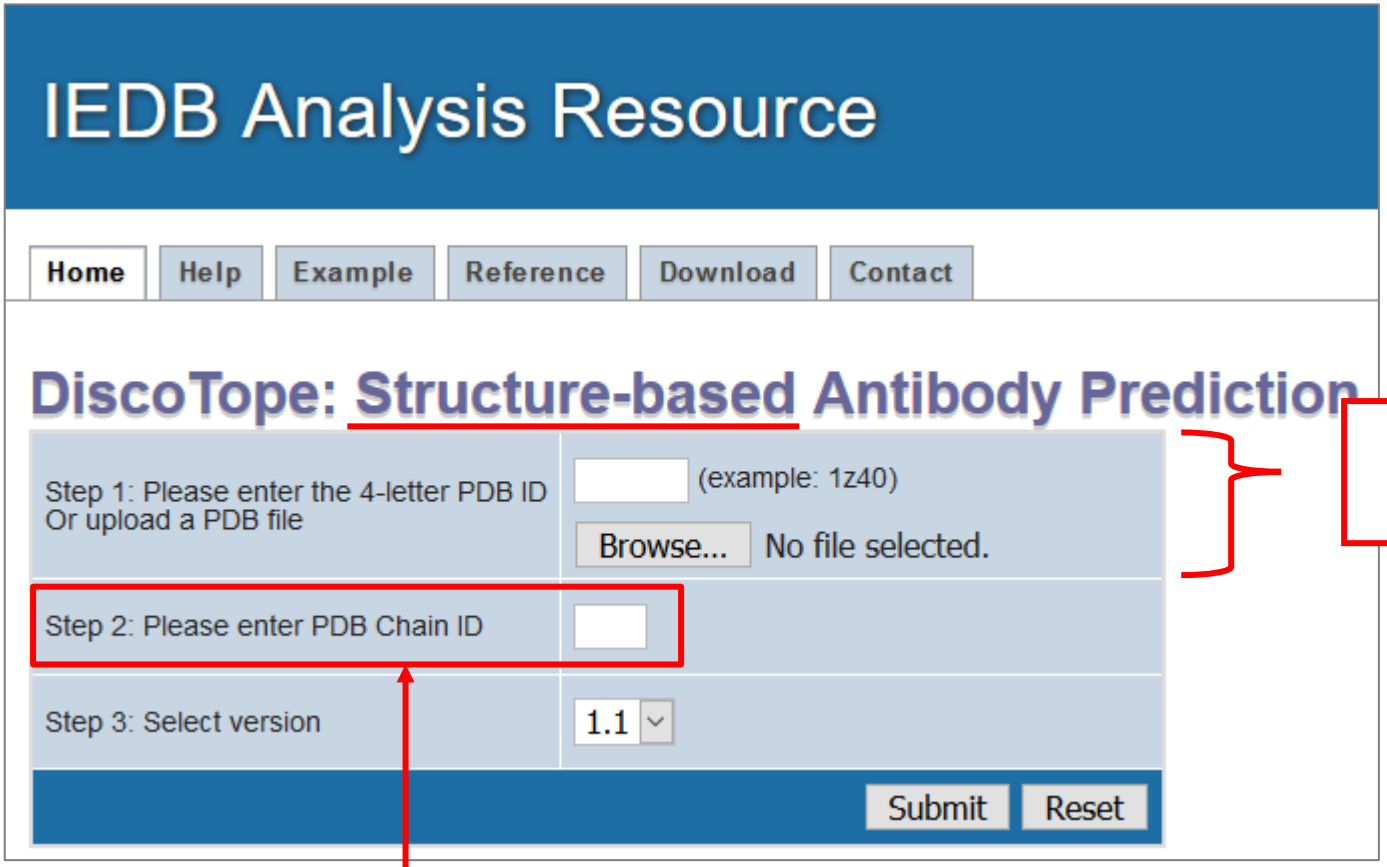
Reliable B cell epitope predictions: impacts of method development and improved benchmarking.

Kringelum JV[1], Lundegaard C, Lund O, Nielsen M.

PMID: 23300419 PMCID: PMC3531324 DOI: 10.1371/journal.pcbi.1002829

DiscoTope

http://tools.iedb.org/discotope/



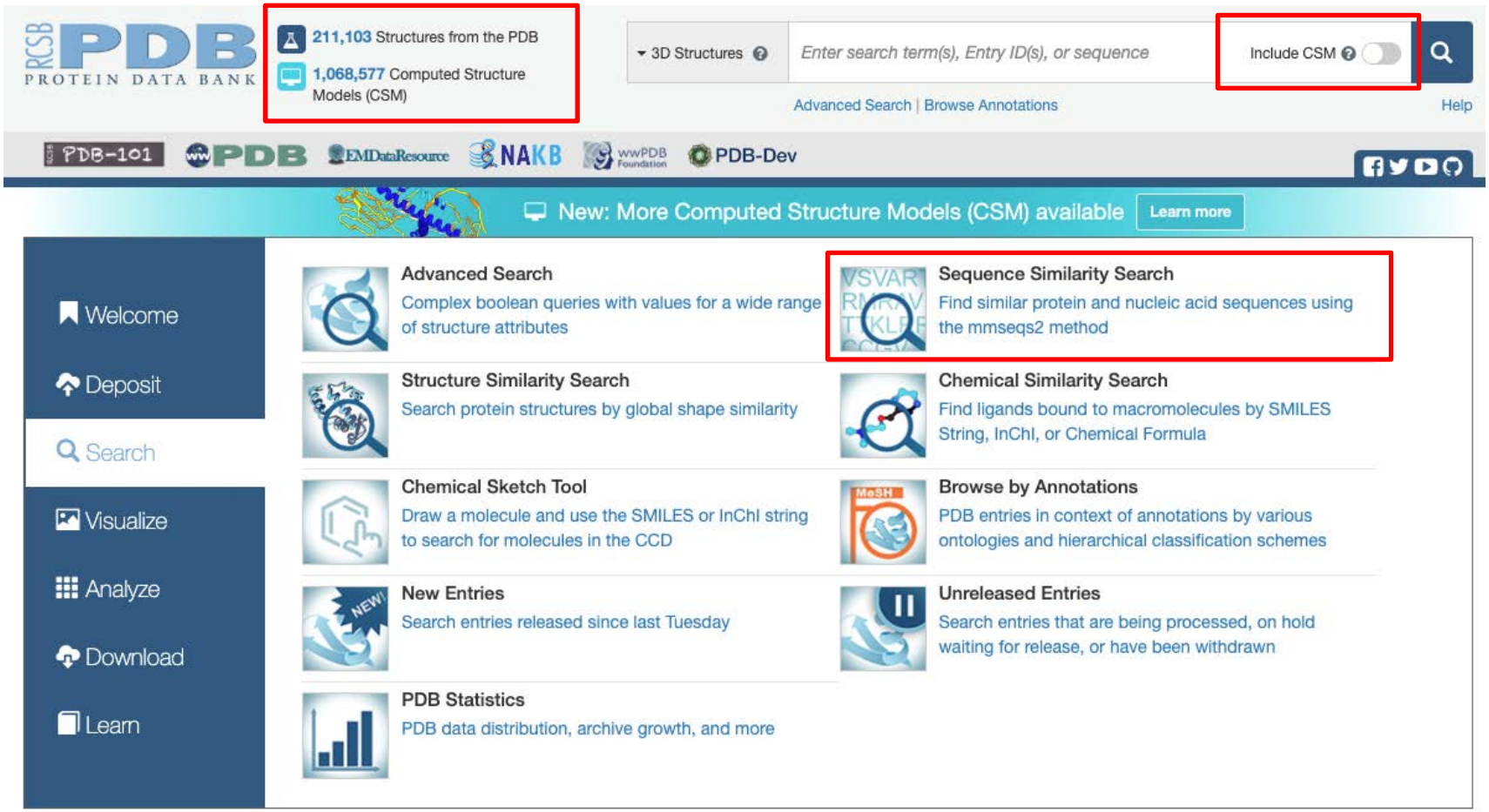
Identify structure

Input PDB id or upload file

Specific to the protein chain of interest

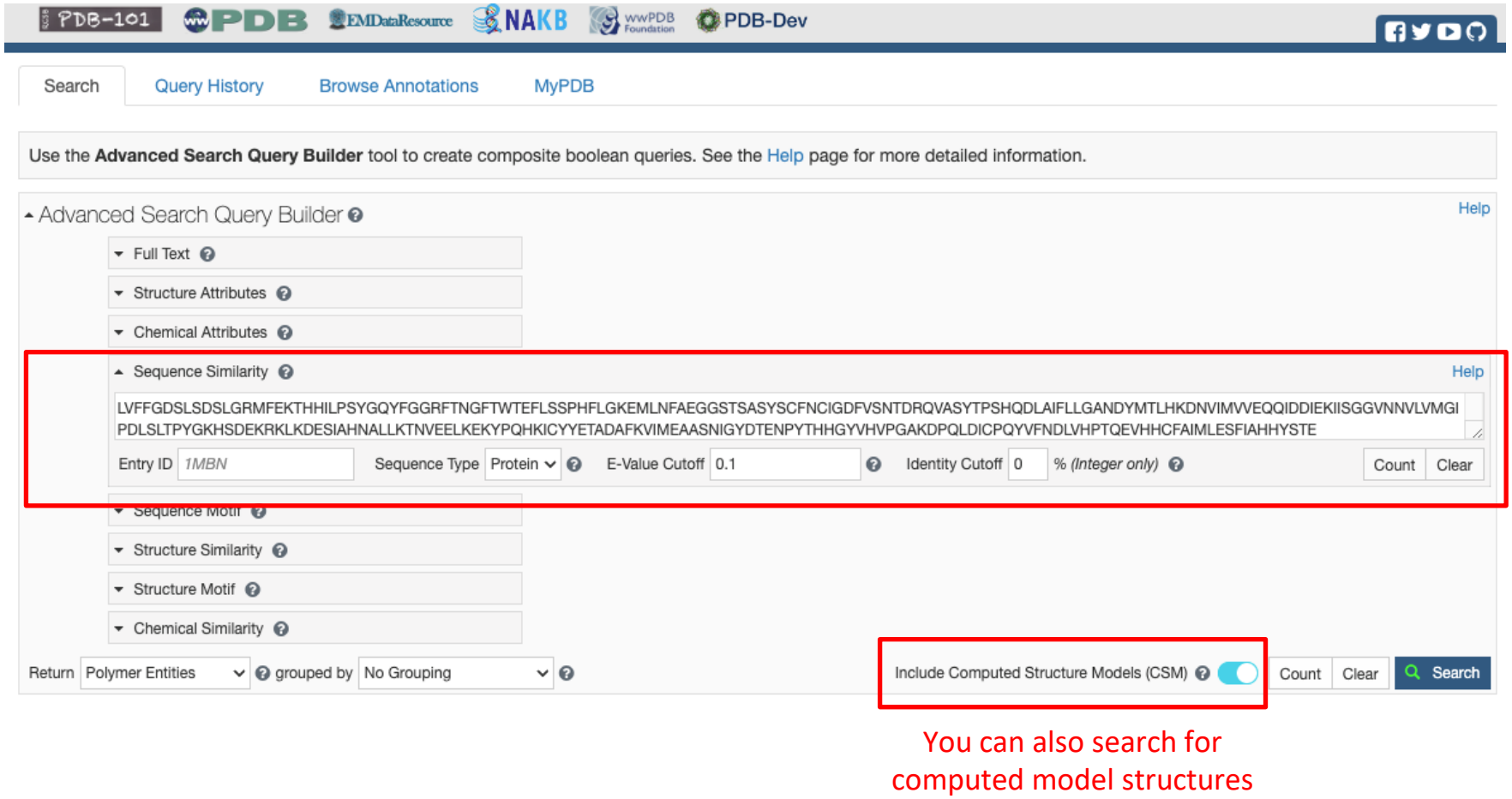
Search in PDB to identify inputs: structures and now, models

http://www.rcsb.org/



Search in PDB to Identify Inputs

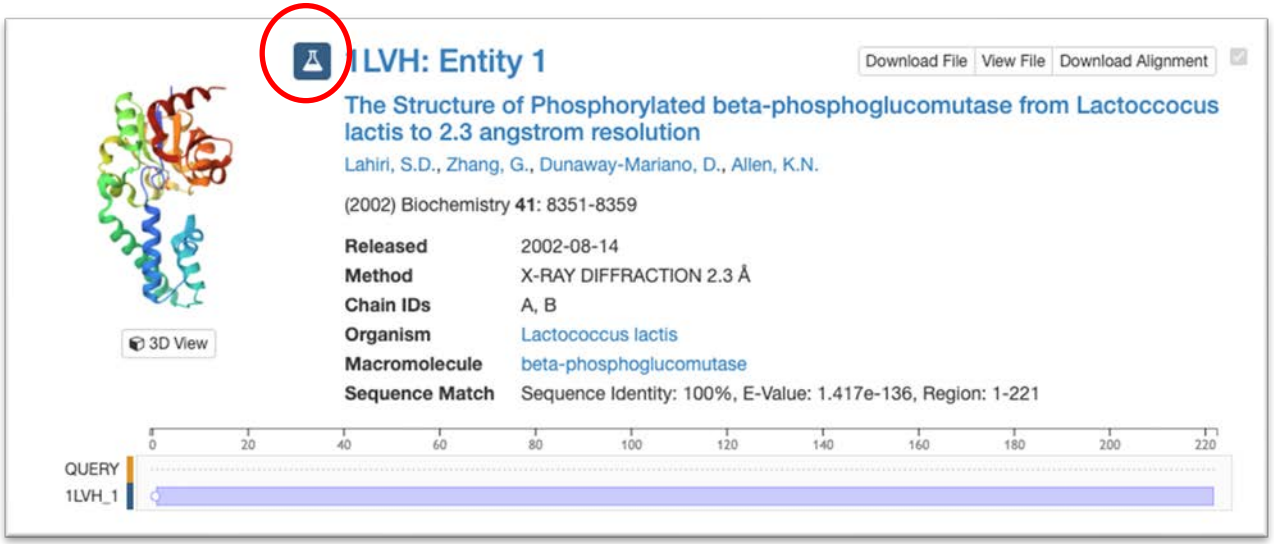
http://www.rcsb.org/



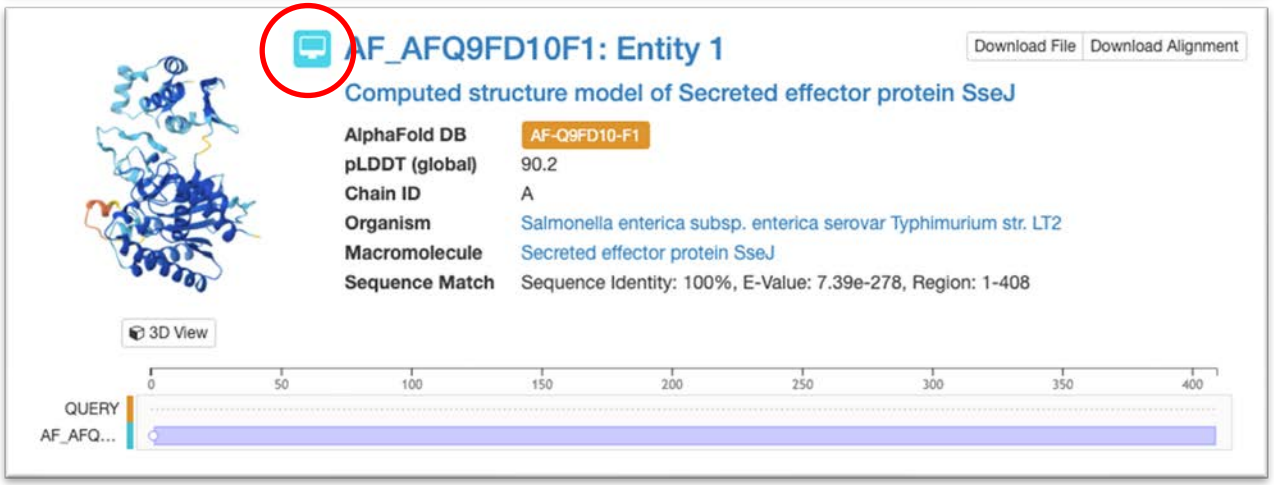
You can also search for computed model structures

Search in PDB to Identify Inputs

http://www.rcsb.org/

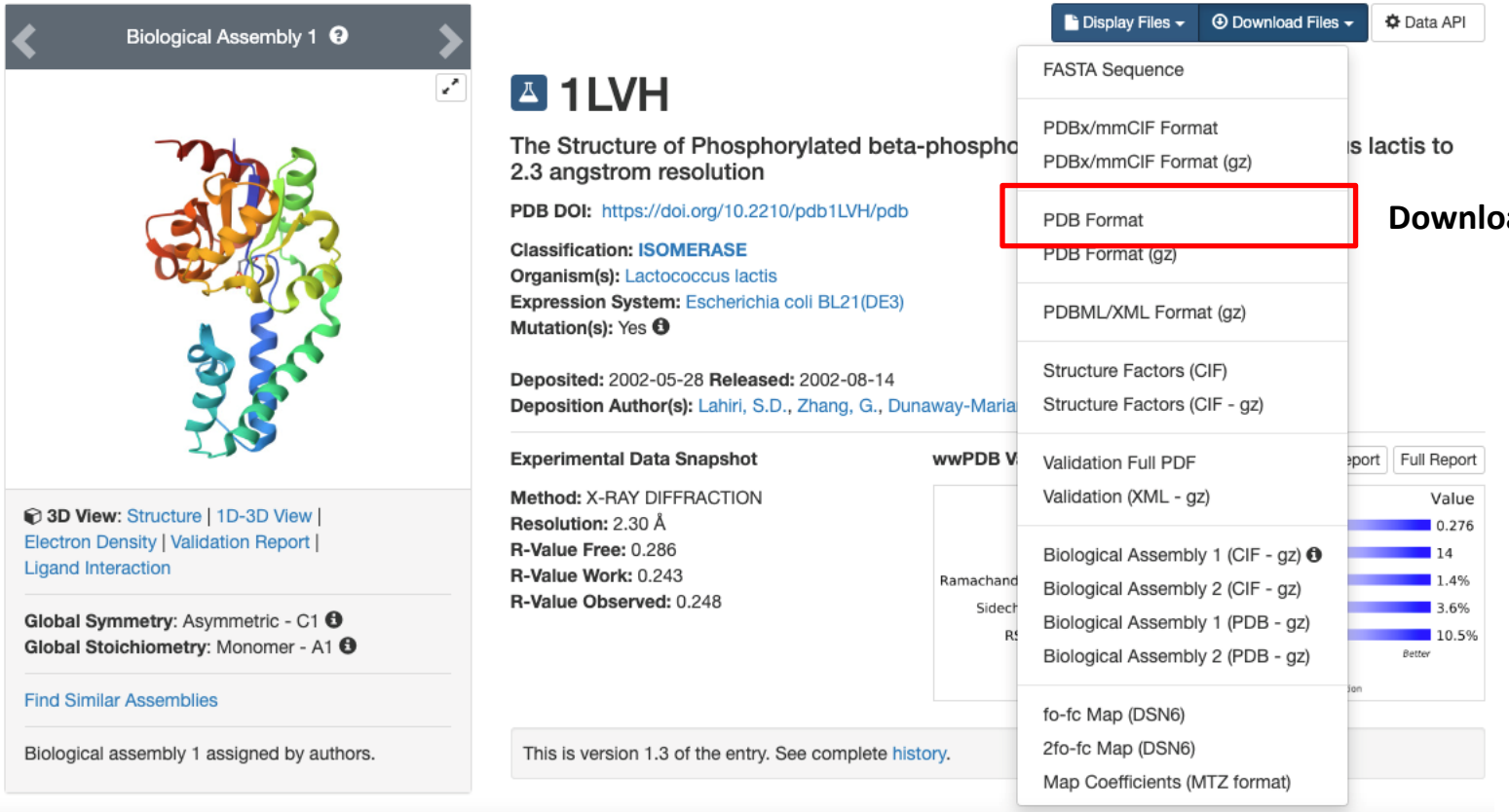


An example of a query with an **experimentally-solved structure**



An example of a query with no solved structure but an **AlphaFold2 model (CSM)**

Search in PDB to Identify Inputs: Experimentally-solved Structure

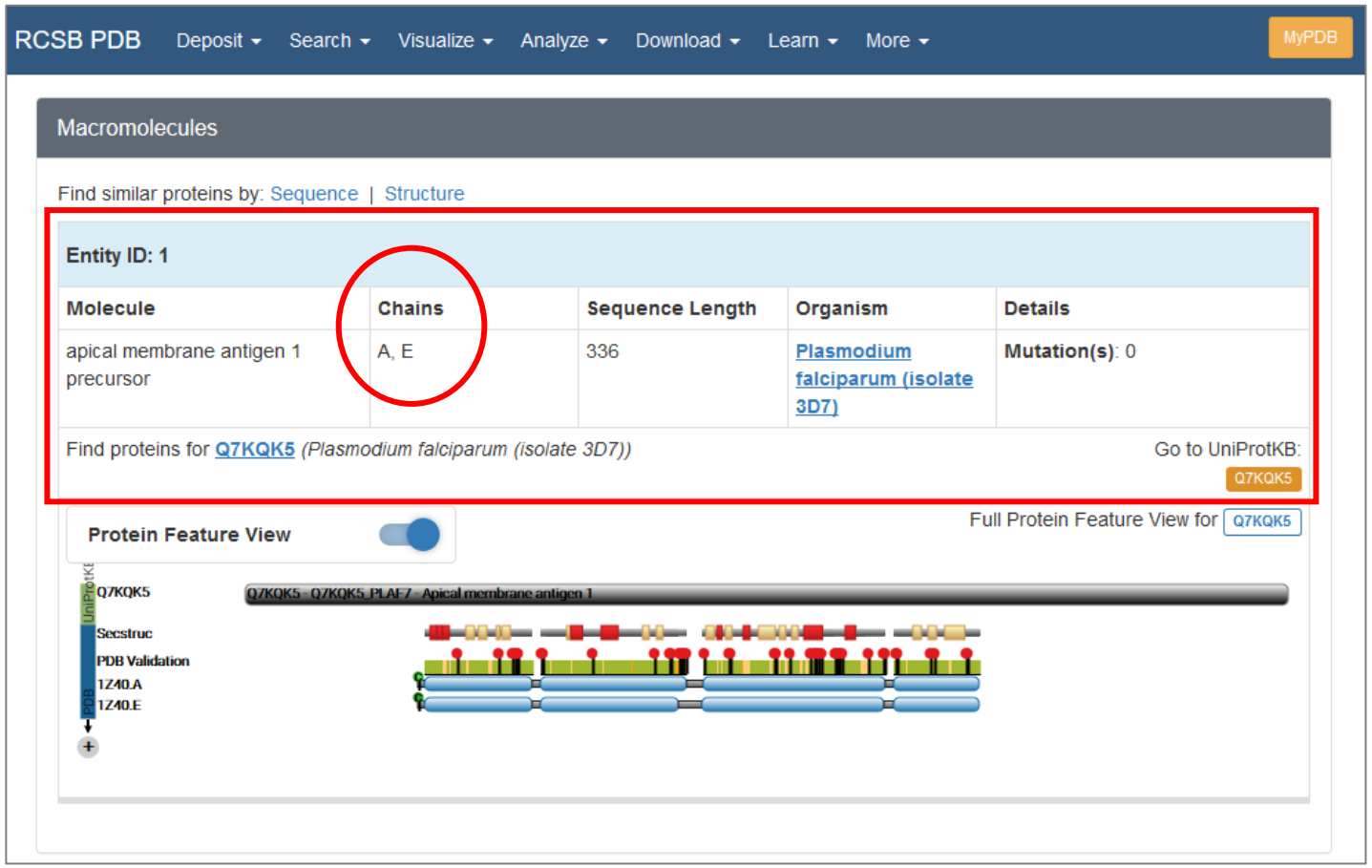


Download PDB file format

http://www.rcsb.org/

Search in PDB to Identify Inputs

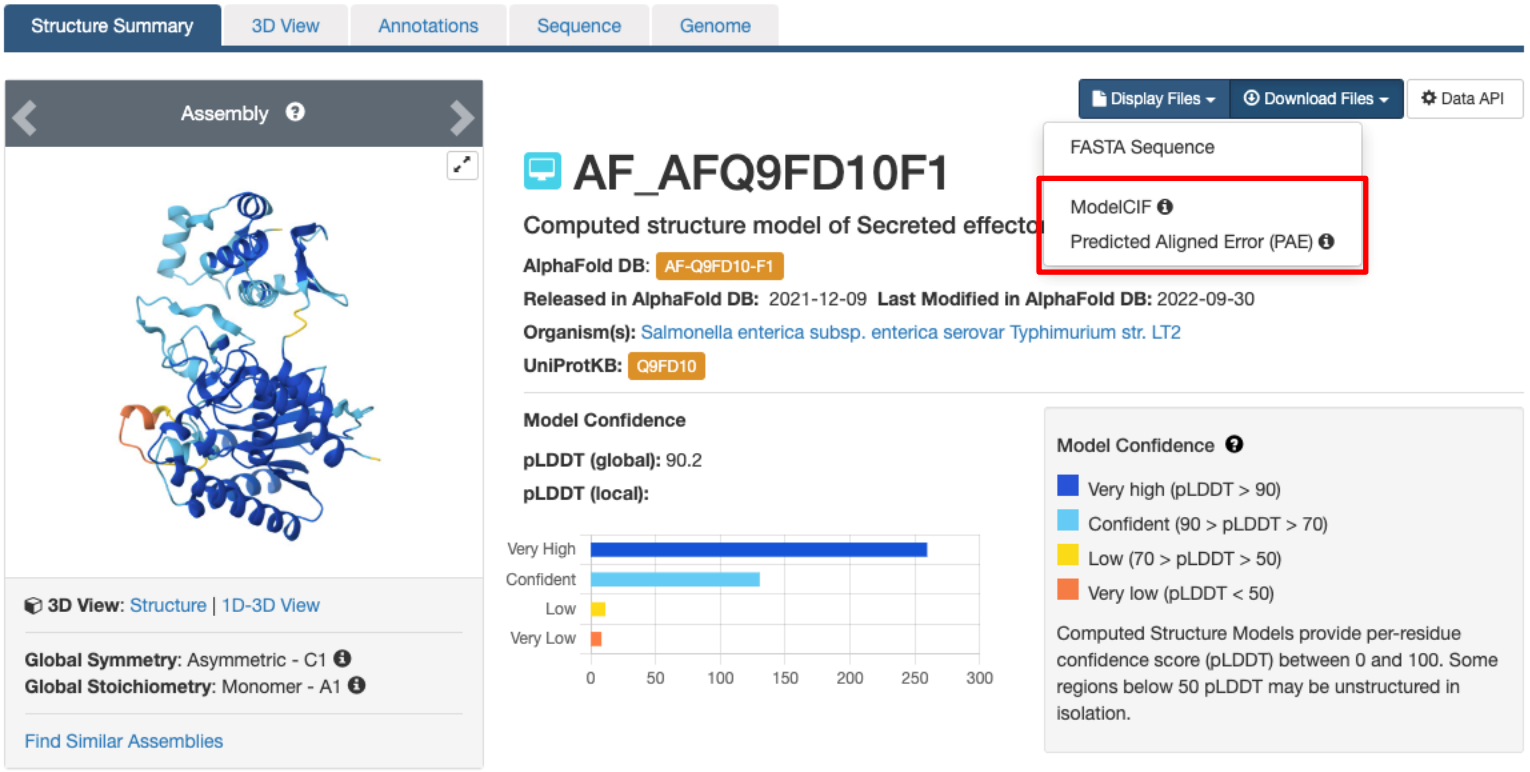
http://www.rcsb.org/



Identify chain

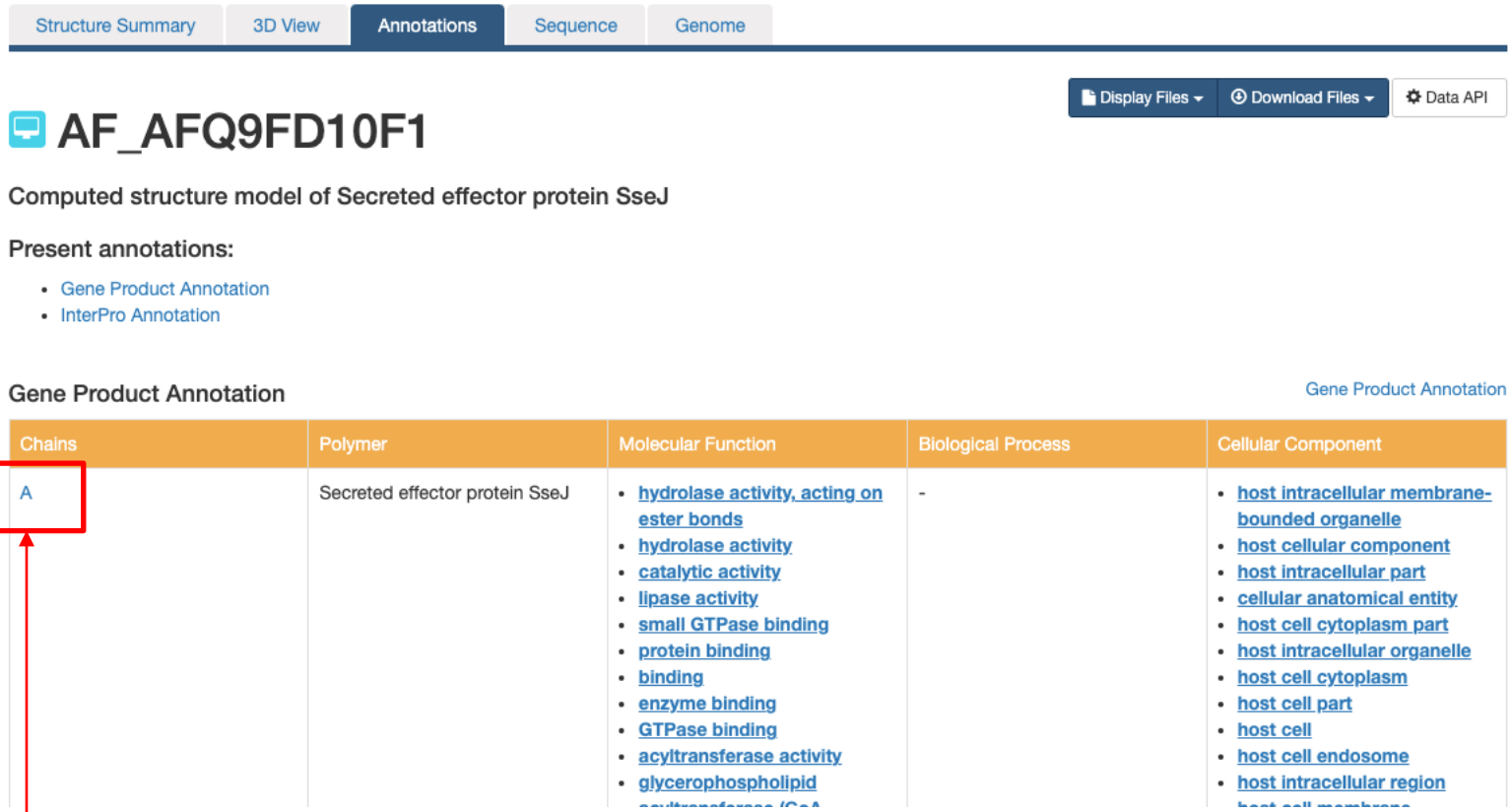
Search in PDB to Identify Inputs: CSM

http://www.rcsb.org/



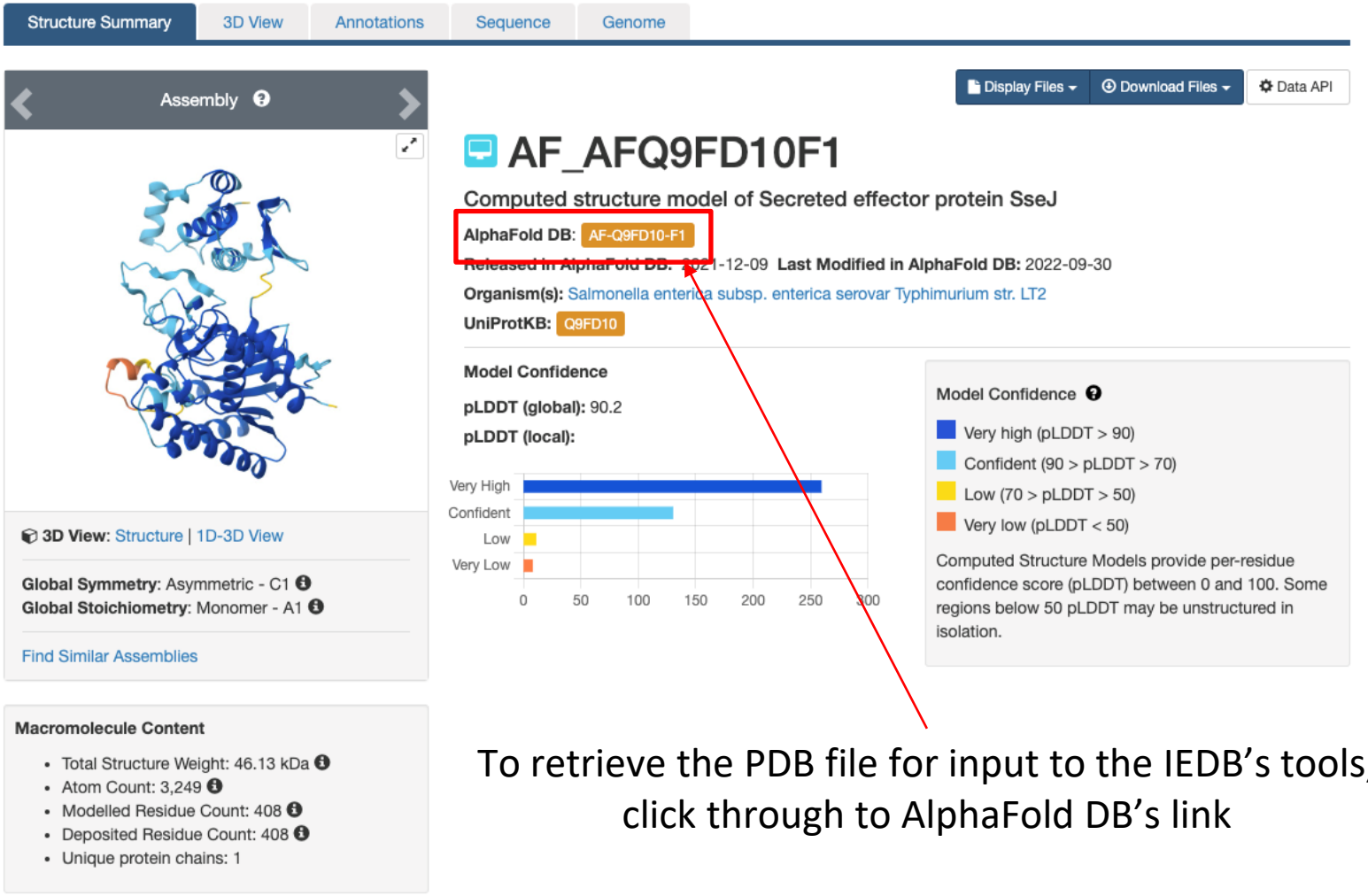
The PDB provides a different file format (so that it is clear that it is a model), ModelCIF.
ModelCIF is not accepted by IEDB currently

Search in PDB to Identify Inputs: CSM



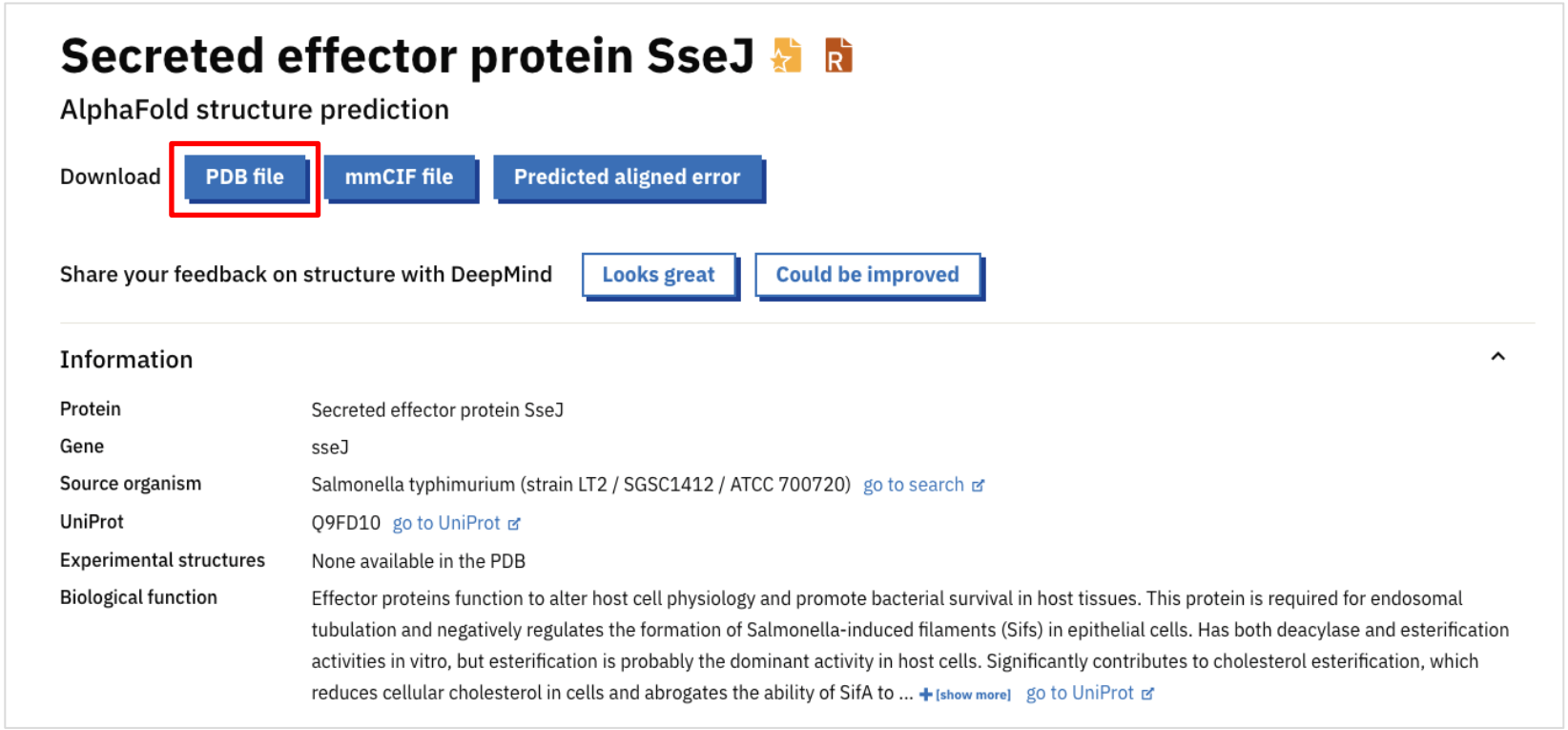
PDB can be used to check the chain (this is **A** in monomeric proteins)

Search in PDB to Identify Inputs: CSM



To retrieve the PDB file for input to the IEDB's tools, click through to AlphaFold DB's link

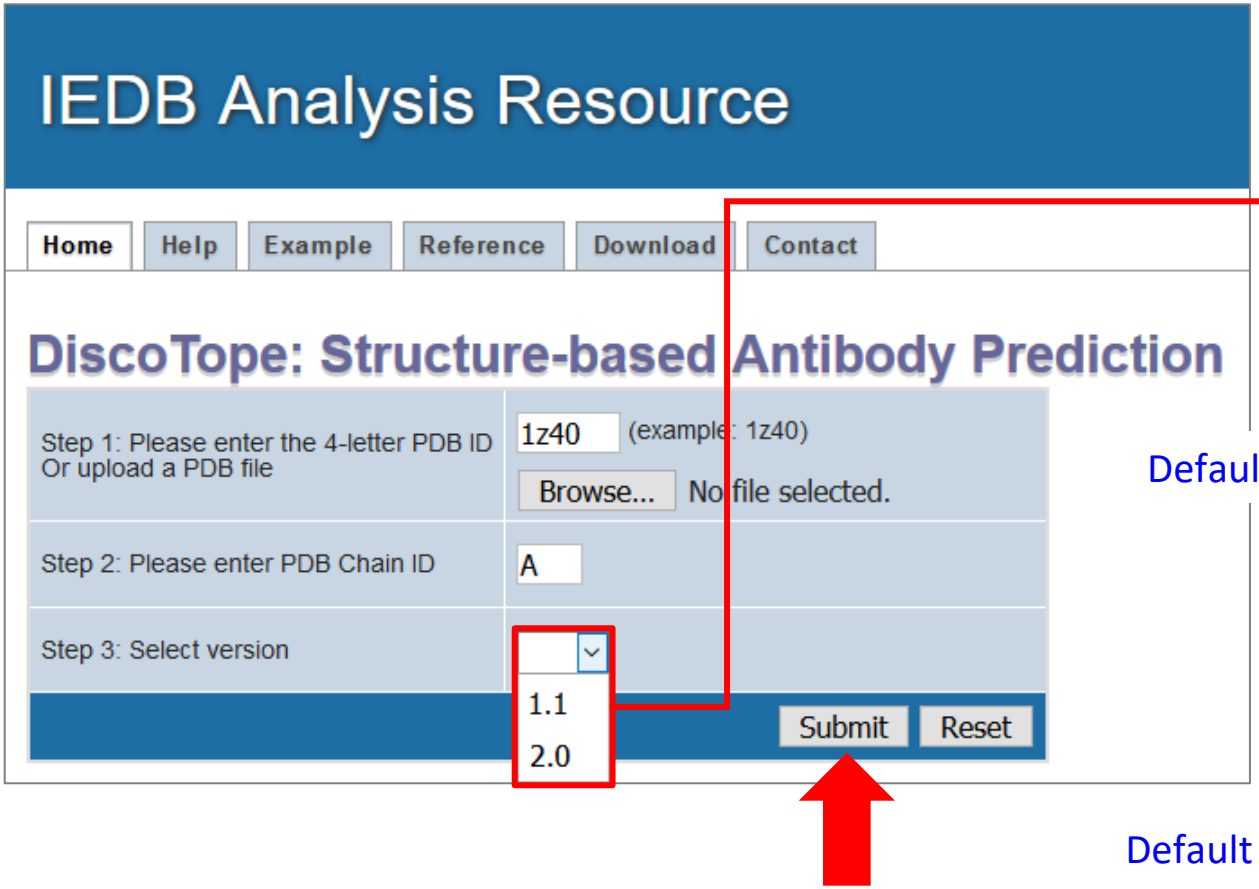
Search in PDB to Identify Inputs: CSM



Click to download the PDB file

DiscoTope – Example

http://tools.iedb.org/discotope/



DiscoTope-1.1

Score	Sensitivity	Specificity
>-3.1	0.16	0.95
>-4.7	0.24	0.90
>-6.0	0.32	0.85
>-6.9	0.40	0.80
>-7.7	0.47	0.75

Default

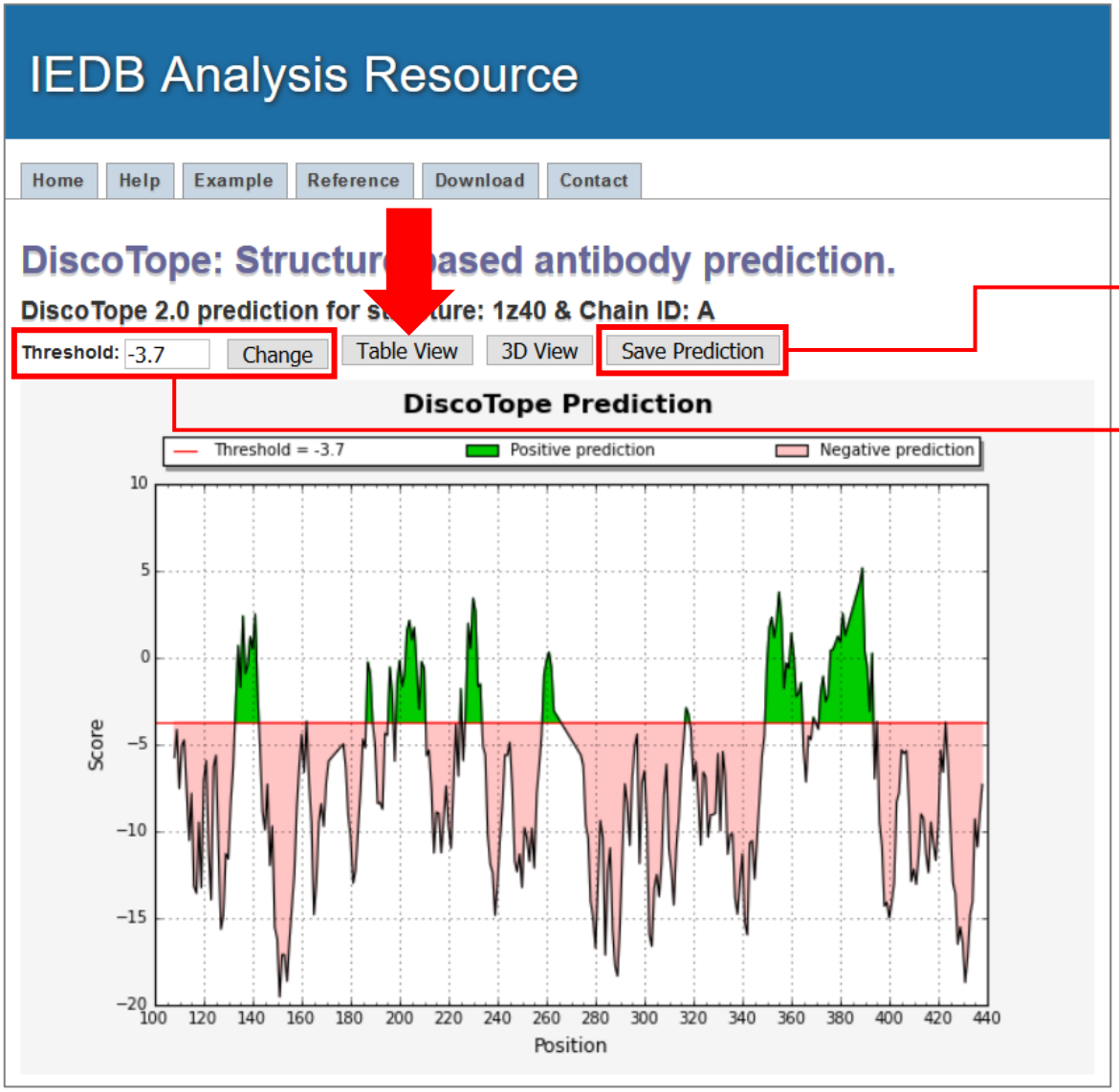
DiscoTope-2.0

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

Default

DiscoTope – Example

http://tools.iedb.org/discotope/

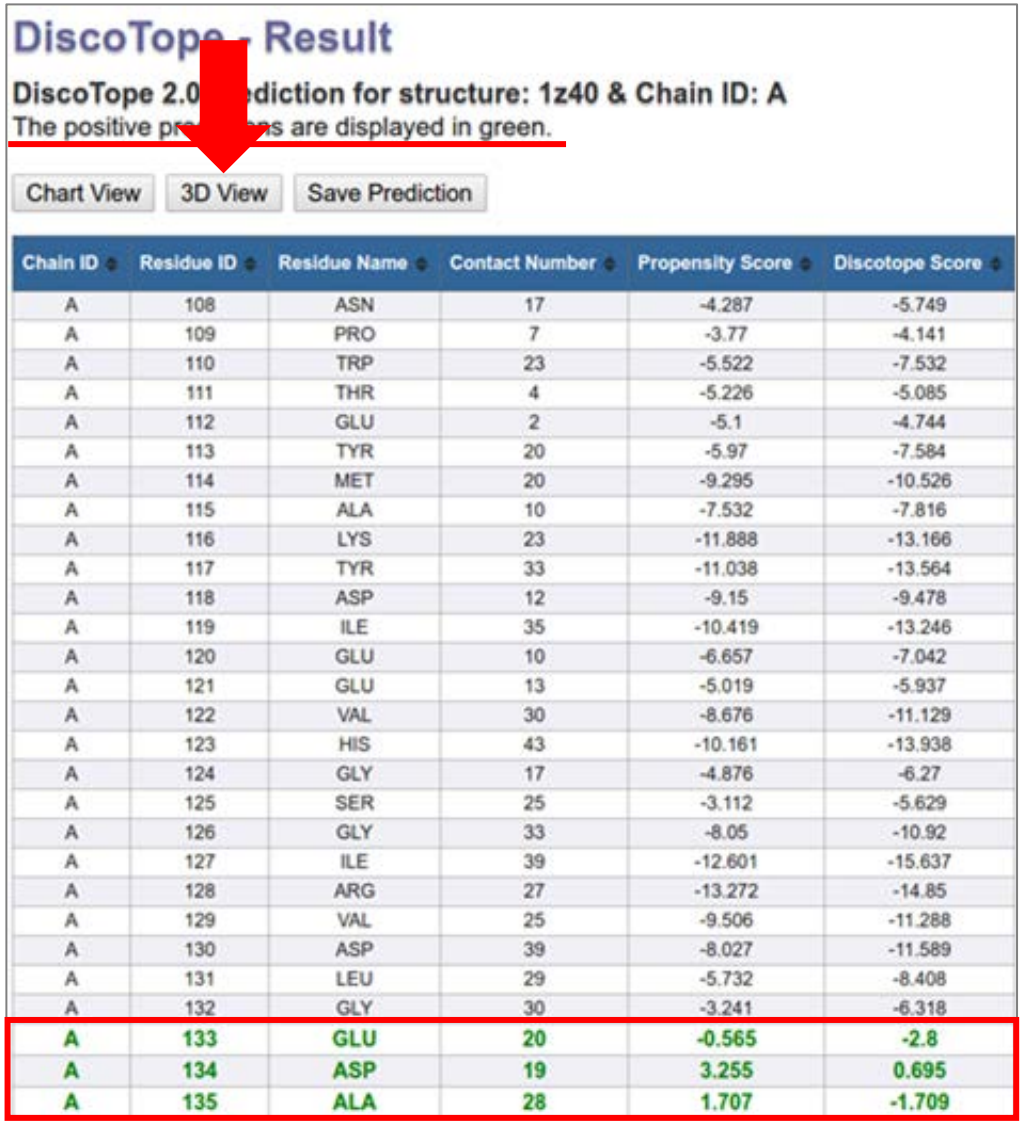


Downloads as csv

Adjustable threshold

DiscoTope - Example

http://tools.iedb.org/discotope/



DiscoTope - Result

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

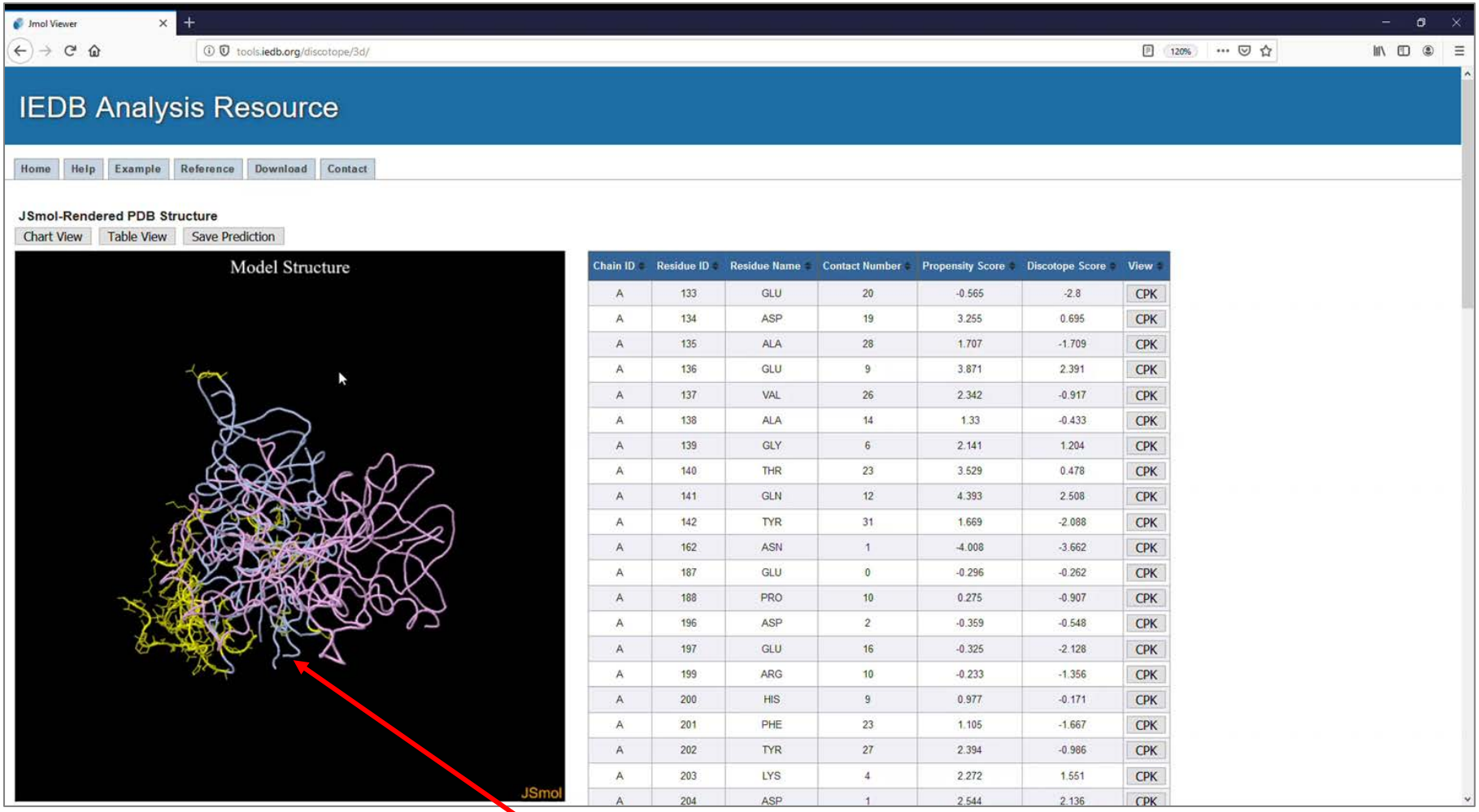
The positive pr ns are displayed in green.

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**

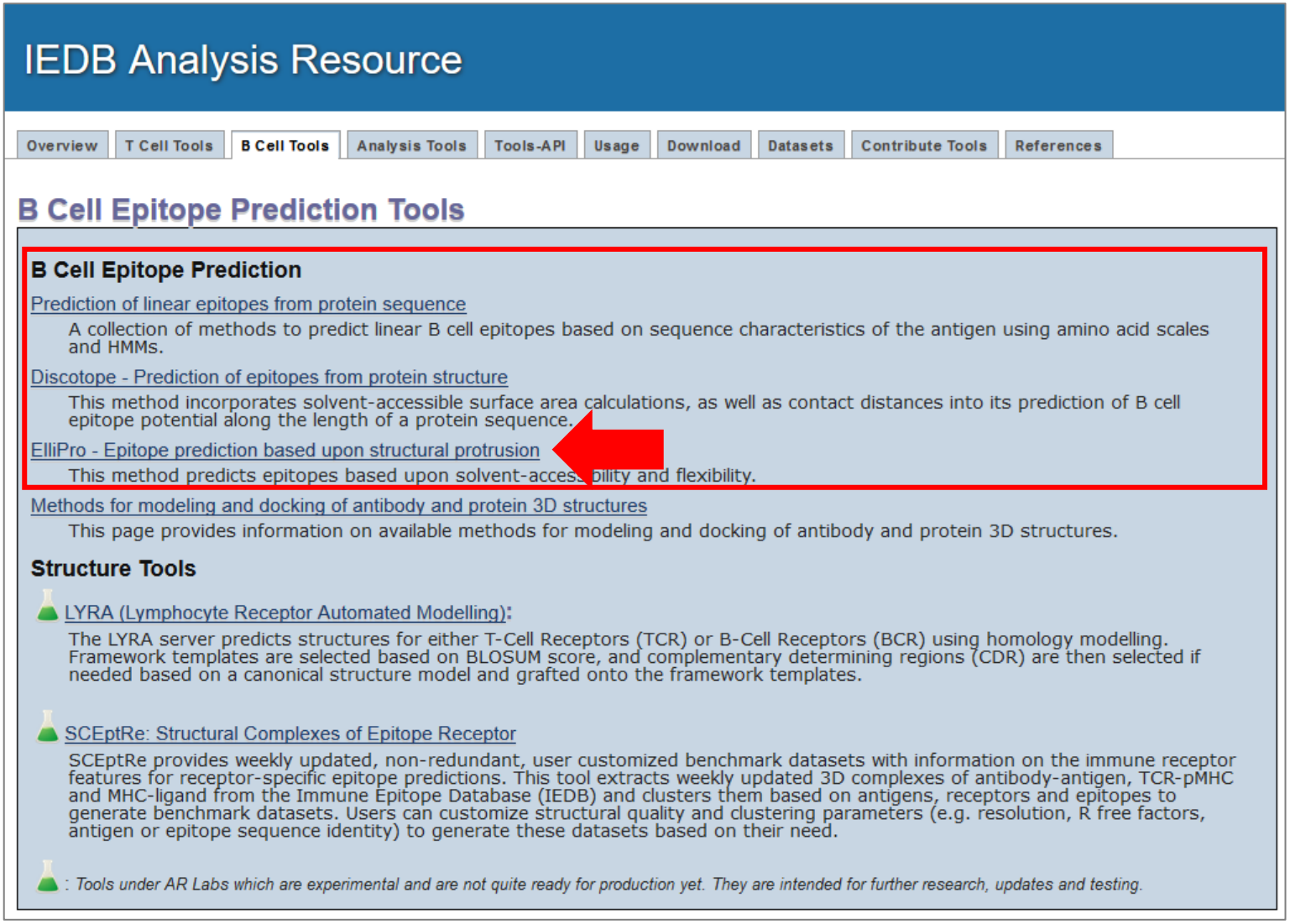
DiscoTope – Example

http://tools.iedb.org/discotope/



Yellow = Predicted epitope

B Cell Prediction Tools on IEDB



http://tools.iedb.org/main/bcell/

ElliPro

- Predicts linear and discontinuous antibody epitopes based on the geometrical properties of protein structure
- Implements three algorithms:
  - Approximation of the protein shape as an ellipsoid (Thornton's method)
  - Calculation of the residue protrusion index (PI)
  - Clustering of neighboring residues based on PI values

EMBO J. 1986 Feb;5(2):409-13.

Location of 'continuous' antigenic determinants in the protruding regions of proteins.

Thornton JM, Edwards MS, Taylor WR, Barlow DJ.

PMID: 2423325 PMCID: PMC1166746

BMC Bioinformatics. 2008 Dec 2;9:514. doi: 10.1186/1471-2105-9-514.

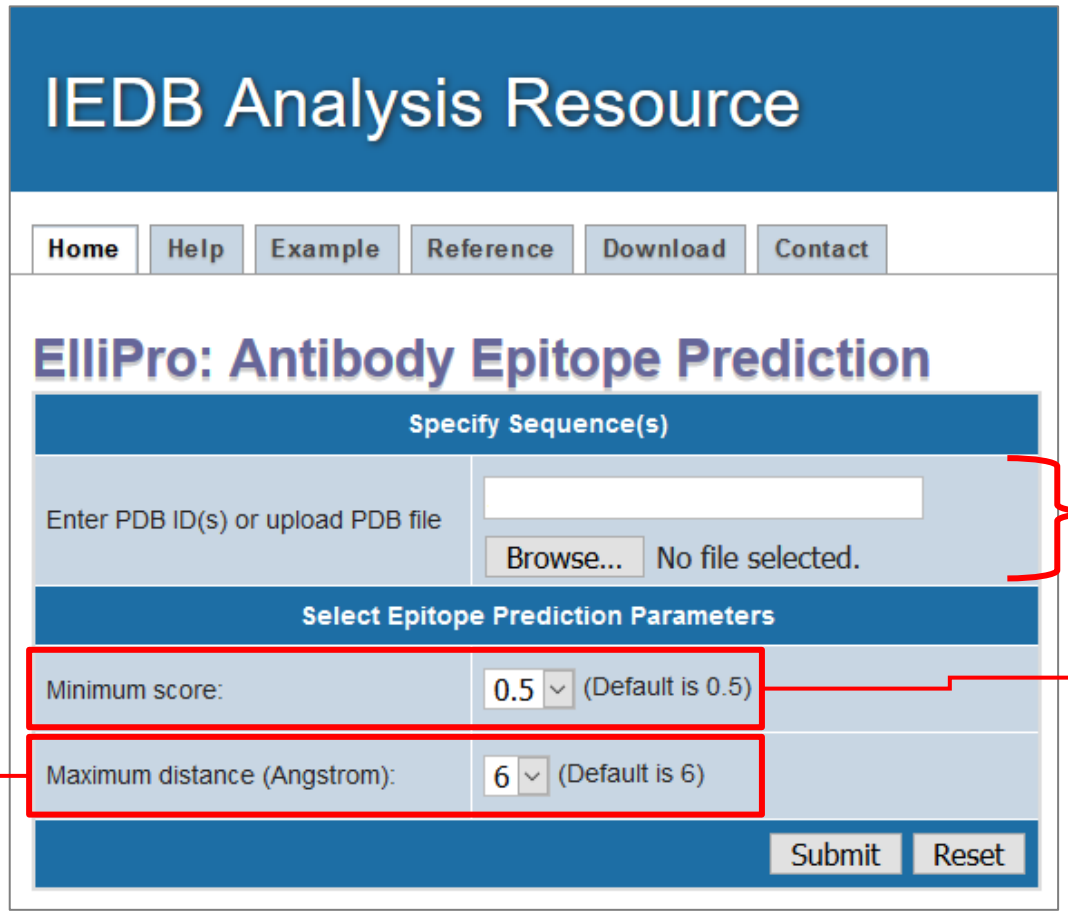
ElliPro: a new structure-based tool for the prediction of antibody epitopes.

Ponomarenko J[1], Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B.

PMID: 19055730 PMCID: PMC2607291 DOI: 10.1186/1471-2105-9-514

ElliPro

http://tools.iedb.org/ellipro/



Identify structure

Input PDB id or upload file

Select min. PI value

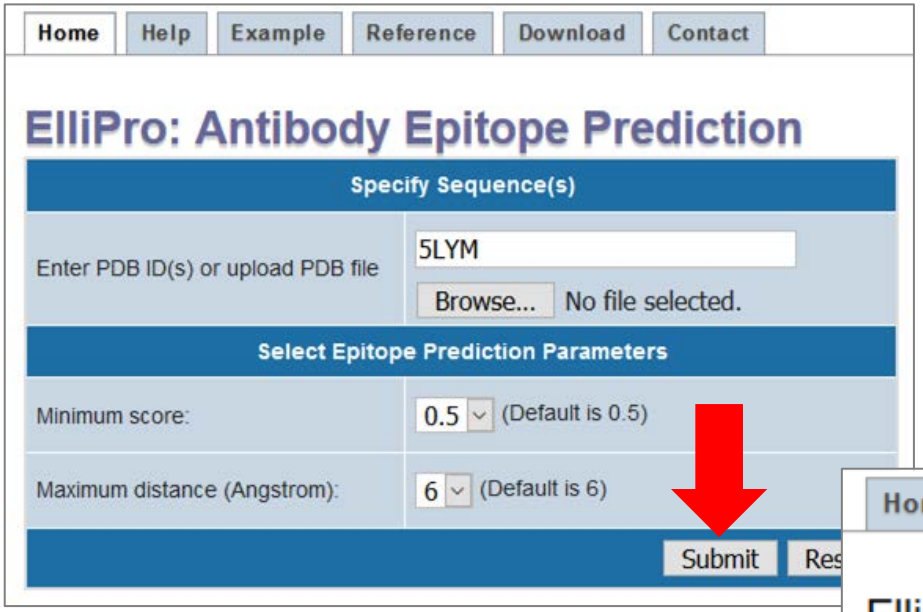
- Averaged over epitope residues
- Higher scores predict fewer epitopes

Specify max distance for predicting (grouping) discontinuous epitopes

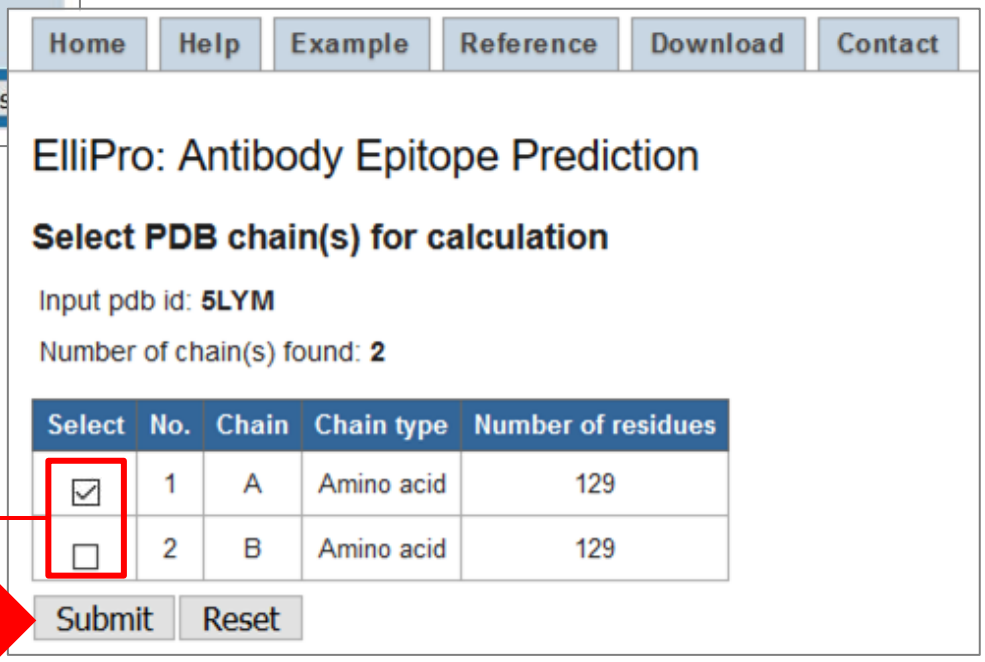
Longer distances predict discontinuous epitopes spanning larger regions

ElliPro – Example

http://tools.iedb.org/ellipro/



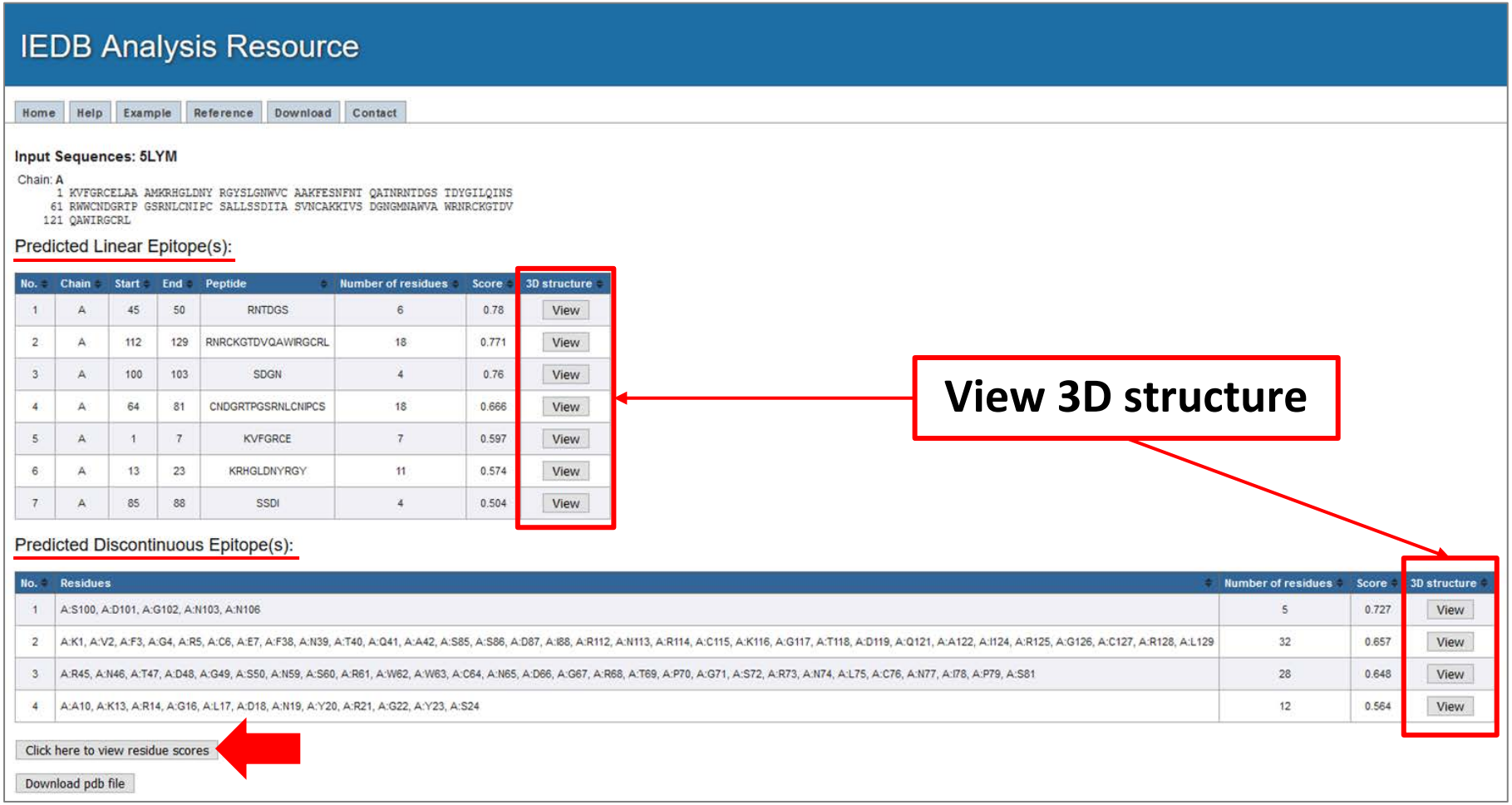
NB: for multimeric complexes, calculations are made for each monomer independently → give each monomer the same chain ID



Select chain(s) of interest

ElliPro – Example

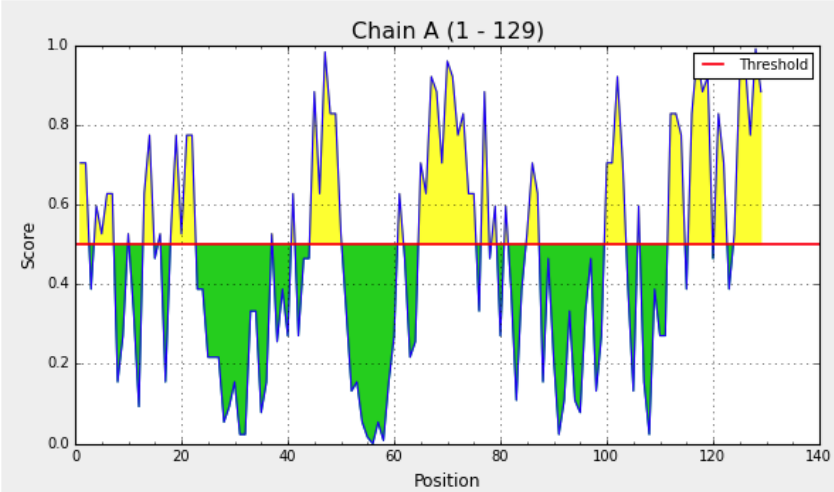
http://tools.iedb.org/ellipro/



ElliPro – Example

http://tools.iedb.org/ellipro/

ElliPro: 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
10	A	10	ALA	0.527
11	A	11	ALA	0.333
12	A	12	MET	0.093
13	A	13	LYS	0.628
14	A	14	ARG	0.775
15	A	15	HIS	0.465
16	A	16	GLY	0.527
17	A	17	LEU	0.155

Summary – B Cell Epitope Predictions

- The IEDB hosts numerous B cell epitope prediction tools
- Our recommended sequence-based epitope prediction tool is the ML model Bepipred-2
- When a structure or good model for the antigen is available, we recommend one of our two structure-based methods, DiscoTope and ElliPro
- The PDB can be used to retrieve structures to input, as well as computational models when structures haven't been solved yet

IEDB
Immune Epitope Database & Tools

THANK YOU!

Practice 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**)

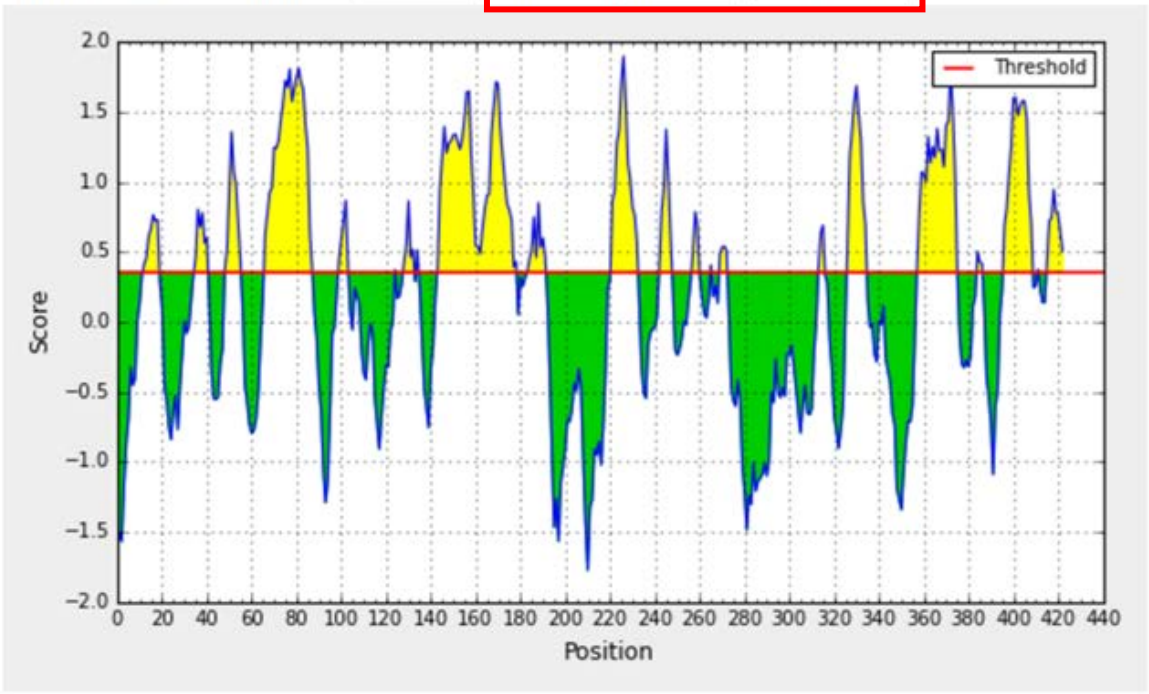
Practice Exercise

Bepipred Linear Epitope Prediction Results

Input Sequences

```
  1 MRCIGISNRD FVEGVSGGSW VDIVLEHGSC VTTMAKNKPT LDFELIKTEA KQPATLRKYC
 61 IEAKLTNTTT ESRCPTQGEP SLNEEQDK RFICKHSMVD RGWGNGCGLF 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 SPVNIEAEPP FGDSYIIVGV EPGQLKLNWL RPLESRGP FEGKPIPNPL
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?

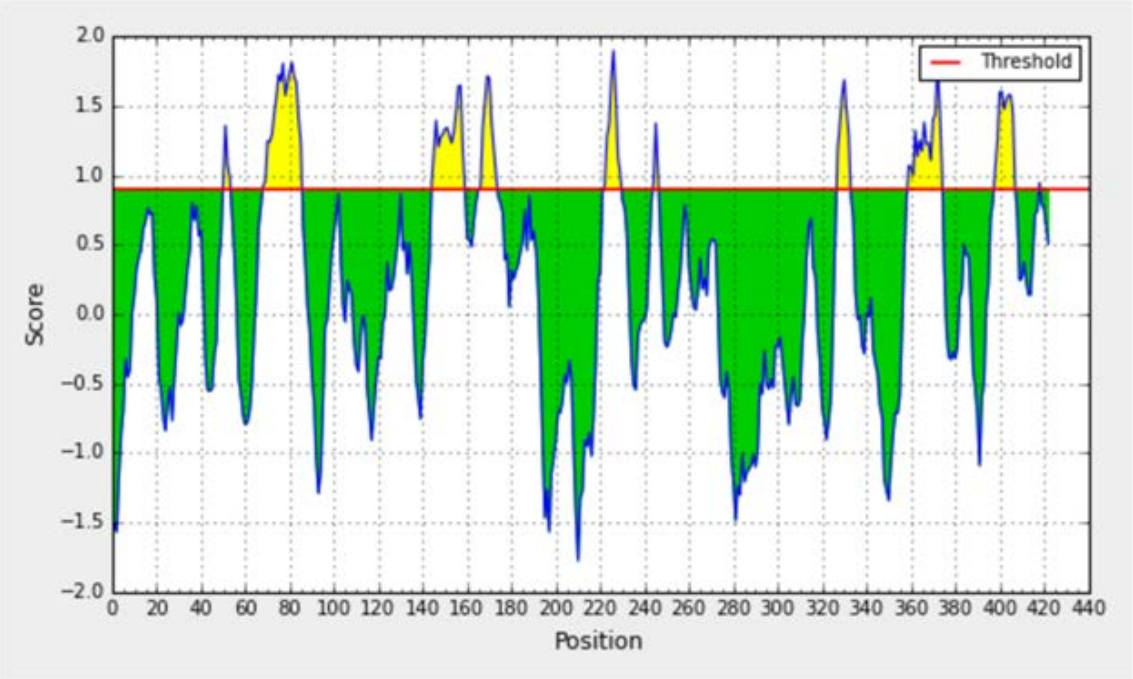
Practice Exercise

Bepipred Linear Epitope Prediction Results

Input Sequences

```
  1 MRCIGISNRD FVEGVSGGSW VDIVLEHGSC VTTMAKNKPT LDFELIKTEA KQPATLRKYC
 61 IEAKLTNTTT ESRCPTQGEP SLNEEQDK RFICKHSMVD RGWGNGCGLF 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 SPVNIEAEPP FGDSYIIVGV EPGQLKLNWL RPLESRGP FEGKPIPNPL
421 LGLDSTRTGH HH
```

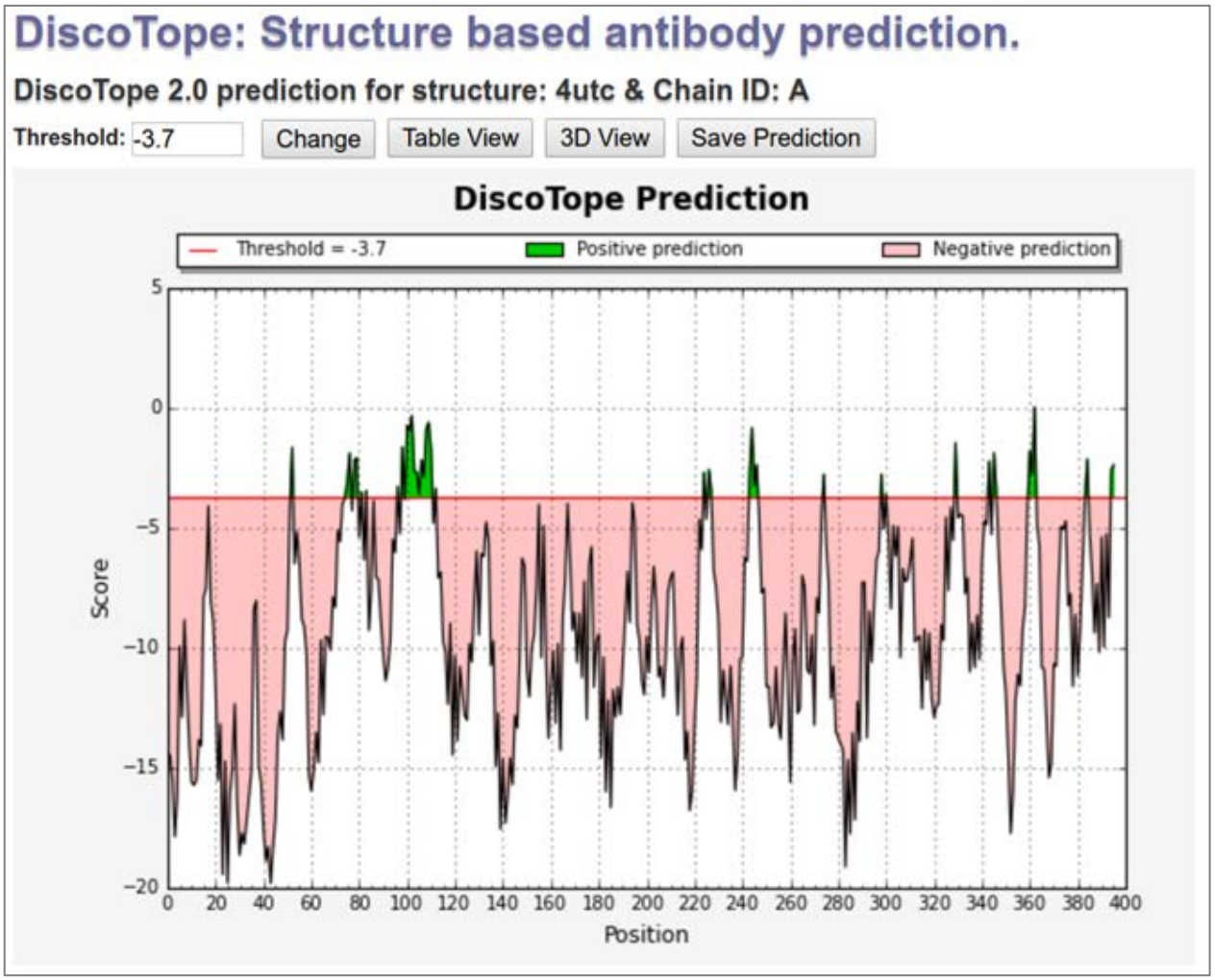
Center position: 4 **Window size:** 7 **Threshold:** 0.9 Recalculate



Average: 0.188 **Minimum:** -1.776 **Maximum:** 1.900

Score **threshold of 0.9** corresponds to **90% specificity**

Practice Exercise



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

Practice Exercise

- 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

Practice Exercise

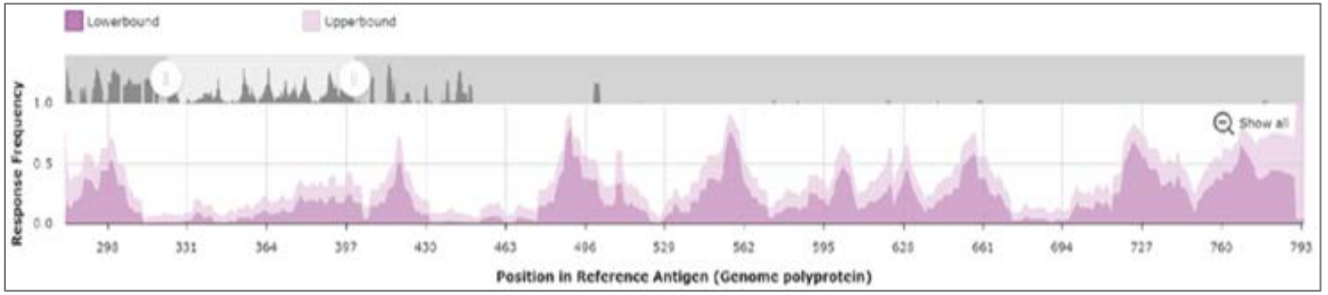
```
4UTC_A_BepiPred     1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA   50
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope    1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA   50

4UTC_A_BepiPred    51 KQPATLRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFICKHSMVDRG  100
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope   51 KQPATLRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFICKHSMVDRG  100

4UTC_A_BepiPred   101 WGNGCGLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA  150
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  101 WGNGCGLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA  150

4UTC_A_BepiPred   151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPRTGLDFNEMVLLQ  200
                      |||||||||||||||||||||||||||||||||||  |||||||||
4UTC_A_DiscoTope  151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPRT--DFNEMVLLQ  198

4UTC_A_BepiPred   201 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV  250
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  199 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV  248

4UTC_A_BepiPred   251 VVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCRLRMDKLQLKGMSYS  300
                      ||||||||||||||||||||||||||||  ||||||||||||||||||
4UTC_A_DiscoTope  249 VVLGSQEGAMHTALTGATEIQMSSGNLLF--HLKCRLRMDKLQLKGMSYS  296

4UTC_A_BepiPred   301 MCTGKFKIVKEIAETQHGTIVIRVQYEGDGSPCKIPFEITDLEKRHVLGR  350
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  297 MCTGKFKIVKEIAETQHGTIVIRVQYEGDGSPCKIPFEITDLEKRHVLGR  346

4UTC_A_BepiPred   351 LITVNPIVTEKDSPVNIEAEPPFGDSYIIVGVEPGQLKLNWLRPLESRGP  400
                      ||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  347 LITVNPIVTEKDSPVNIEAEPPFGDSYIIVGVEPGQLKLNWLRPL-----  391

4UTC_A_BepiPred   401 FEGKPIPNPLLGLDSTRTGHHH    422
```

Predicted

Correctly predicted

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

Practice 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 WGNGCGLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA 150
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  101 WGNGCGLFGKGGIVTCAKFTCKKNMEGKIVQPENLEYTIVITPHSGEEHA 150

4UTC_A_BepiPred   151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPRTGLDFNEMVLLQ 200
                      |||||||||||||||||||||||||||||||||||  |||||||||
4UTC_A_DiscoTope  151 VGNDTGKHGKEIKITPQSSTTEAELTGYGTVTMECSPRT--DFNEMVLLQ 198

4UTC_A_BepiPred   201 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV 250
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  199 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV 248

4UTC_A_BepiPred   251 VVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCRLRMDKLQLKGMSYS 300
                      |||||||||||||||||||||||||||  ||||||||||||||||||
4UTC_A_DiscoTope  249 VVLGSQEGAMHTALTGATEIQMSSGNLLF--HLKCRLRMDKLQLKGMSYS 296

4UTC_A_BepiPred   301 MCTGKFKIVKEIAETQHGTIVIRVQYEGDGSPCKIPFEITDLEKRHVLGR 350
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  297 MCTGKFKIVKEIAETQHGTIVIRVQYEGDGSPCKIPFEITDLEKRHVLGR 346

4UTC_A_BepiPred   351 LITVNPIVTEKDSPVNIEAEPPFGDSYIIVGVEPGQLKLNWLRPLESRGP 400
                      |||||||||||||||||||||||||||||||||||||||||||
4UTC_A_DiscoTope  347 LITVNPIVTEKDSPVNIEAEPPFGDSYIIVGVEPGQLKLNWLRPL----- 391

4UTC_A_BepiPred   401 FEGKPIPNPLLGLDSTRTGHHH    422

4UTC_A_DiscoTope  392 ----------------------    391
```

Epitope residues **from the IEDB** in Dengue envelope glycoprotein

Predicted

Correctly predicted