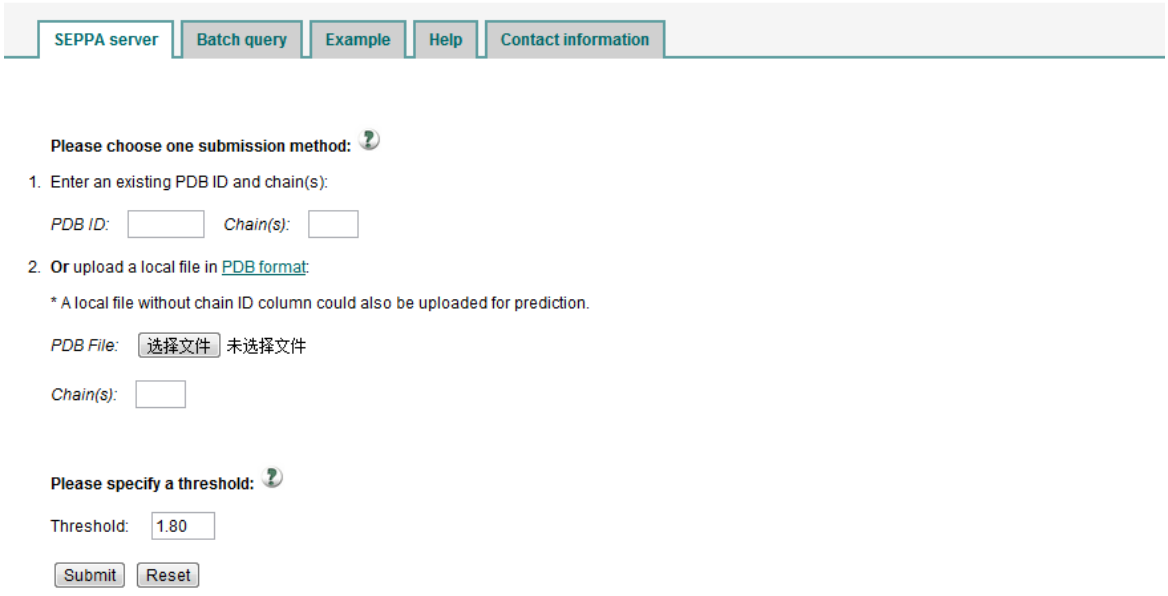
SEPPA: a computational server for spatial epitope prediction of protein antigens

Seppa:

This server is presented as a tool for conformational B-cell epitope prediction. With 3D protein structure as input, each residue in the query protein will be given a score according to its neighborhood residues' information. Higher score corresponds to higher probability the residue to be involved in an epitope.

Homepage:
http://lifecenter.sgst.cn/seppa/index.php



Data format:

The data for query should be an existing PDB ID or a protein structure with a PDB format. This demo uses the data **1NCA** (pdb id) whose information is shown in part 1 below and more detail could be found in PDB.
http://www.rcsb.org/pdb/explore/explore.do?structureId=1NCA

1. Demo Data Information:

1NCA
Refined crystal structure of the influenza virus N9 neuraminidase-NC41 Fab complex.

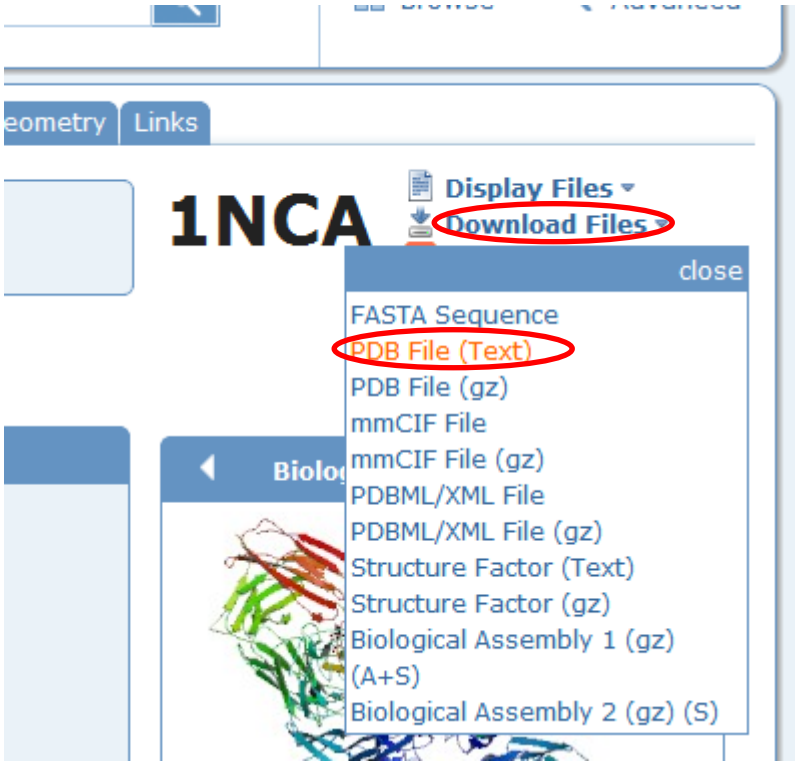
Molecular Description

Molecular Description Hide

Classification: Hydrolase(o Glycosyl)
Structure Weight: 92741.77

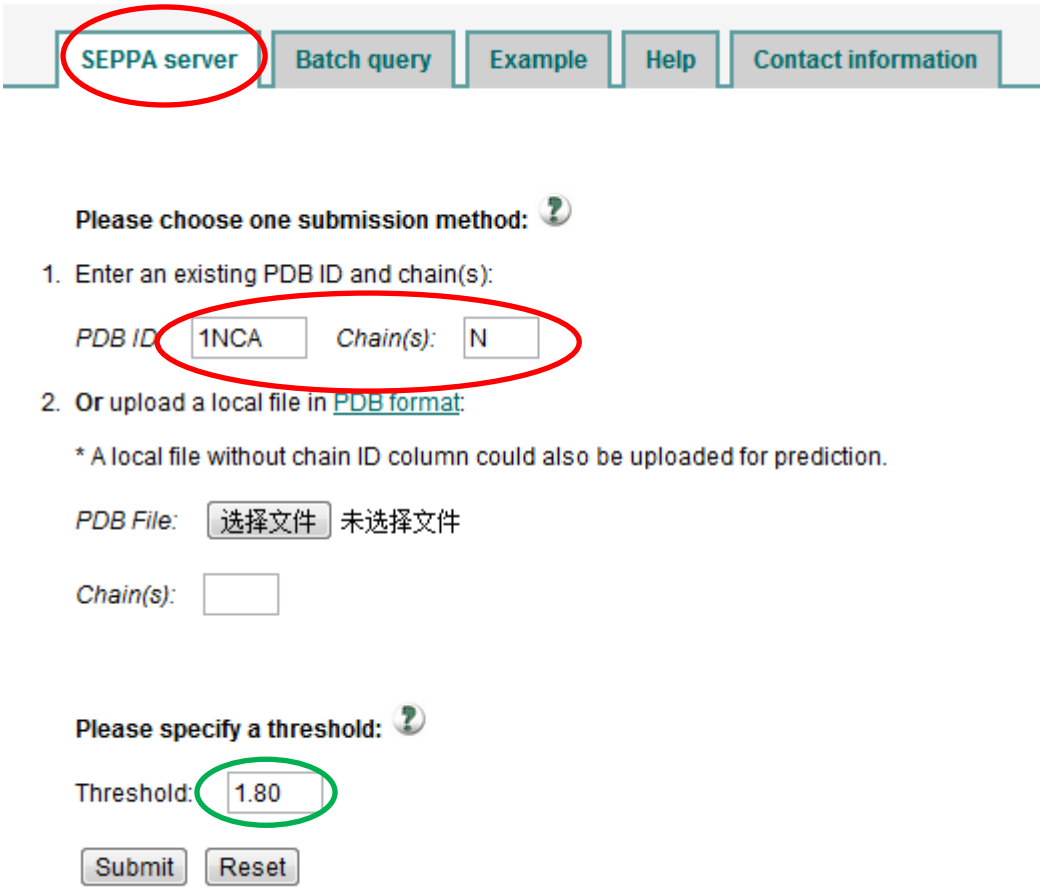
Molecule:	INFLUENZA A SUBTYPE N9 NEURAMINIDASE				
Polymer:	1	**Type:**	protein	**Length:**	389
Chains:	N				
EC#:	3.2.1.18				
Organism	Influenza A virus (A/tern/Australia/G70C/1975(H11N9))				
UniProtKB:	P03472				
Molecule:	IGG2A-KAPPA NC41 FAB (LIGHT CHAIN)				
Polymer:	2	**Type:**	protein	**Length:**	214
Chains:	L				
Organism	Mus musculus				
Molecule:	IGG2A-KAPPA NC41 FAB (HEAVY CHAIN)				
Polymer:	3	**Type:**	protein	**Length:**	221
Chains:	H				
Organism	Mus musculus				
UniProtKB:	P01865				

You can download the data as a PDB File format like below:

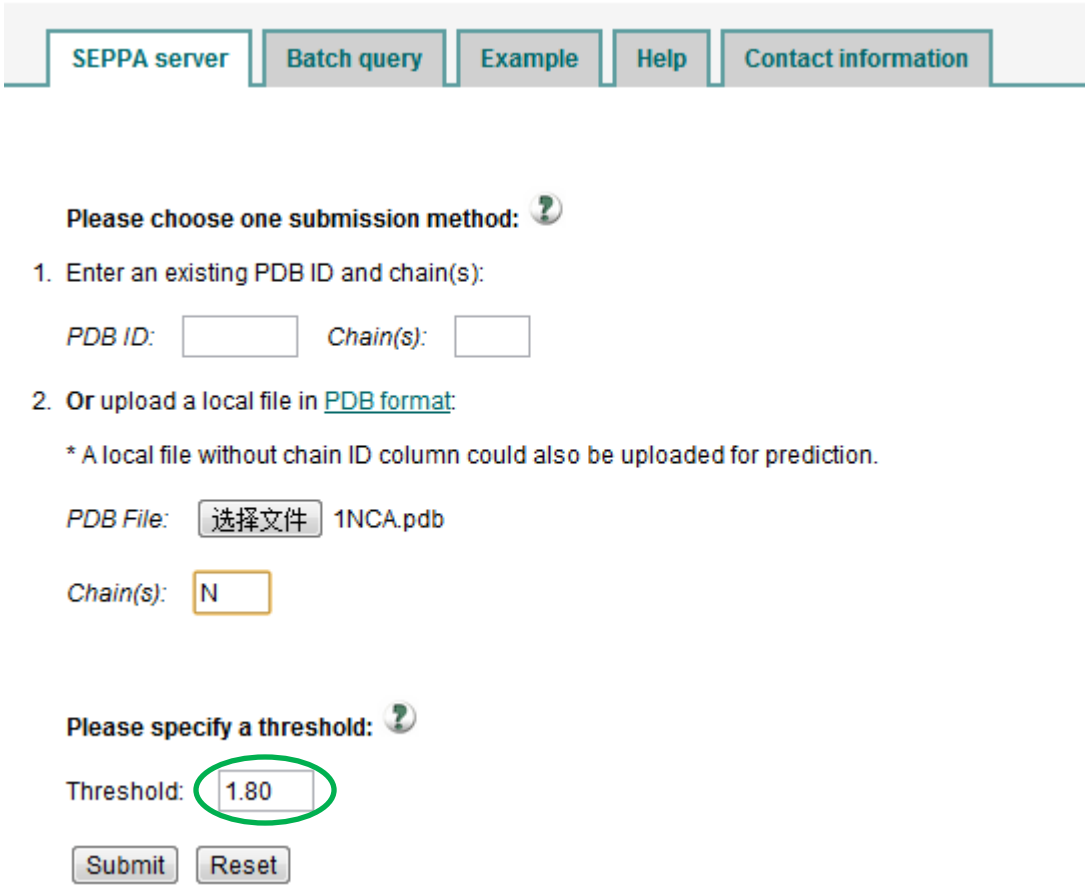


2. Data input

If your query has an existing PDB ID, you could submit the PDB id and the chain(s) directly.



Or you can also submit a protein structure with a PDB format



Note 1: The value of threshold (the green circle)

The default value of THRESHOLD is set at 1.80 to help to specify the epitope residues. Under the default threshold, a sensitivity of 0.568 was received while the specificity was 0.740 on SEPPA training dataset.

As a reference, other thresholds with corresponding sensitivity/specificity/accuracy values are listed below in Table 1.

Users can set different thresholds. Under a lower threshold, more residues will be included as predicted epitope residues. That always results in the increasement of the true positive rate and the false positive rate at the same time.

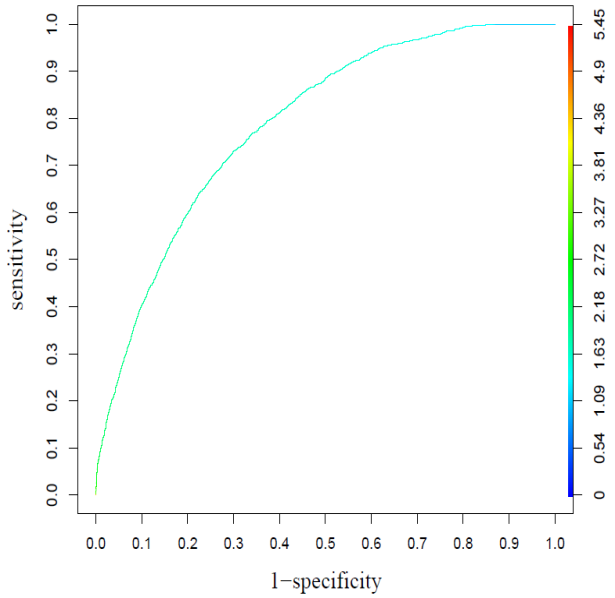


Figure 1. ROC curve of SEPPA. The ROC curve for training dataset.

Table 1. Performance of SEPPA. The sensitivity/specificity/accuracy with different thresholds.

Threshold	Sensitivity	Specificity	Accuracy
1.55	**0.959**	**0.259**	**0.377**
1.60	**0.927**	**0.345**	**0.448**
1.65	**0.859**	**0.452**	**0.531**
1.70	**0.778**	**0.558**	**0.612**
1.75	**0.672**	**0.658**	**0.684**
1.80	**0.568**	**0.740**	**0.739**
1.85	**0.459**	**0.810**	**0.782**
1.90	**0.363**	**0.860**	**0.809**
1.95	**0.278**	**0.900**	**0.829**
2.00	**0.217**	**0.927**	**0.844**

3. The Output Page

Antigenic Prediction for 1NCA.pdb:

Chain: N
Threshold: 1.80
Number of total residues: 389
Number of predicted epitope residues: 39

View 3D structure in Jmol

1-50	IRDFNNLTKG LCTiNSWHIY GKDNAvRIgE DSDVLvTREp YvsCDPDECR
51-100	fyaLSQGTTI RGKHSNGTIH DRSQYRAlIs WPLSSPPTVY NSRVECIGWS
101-150	stsCHDgKTR MSiciSGPNN NaSaViWYNR RPVTEINTWA RNIlRTQEsE
151-200	CVCHNgVCPv VfTdGSATGP AETRiYyfKE gKILKWEPLA GTAKHIEECS
201-250	CYgERAEITc tcRdNWQGsN RpViRIDPVA MTHTSQyICS pVLTdNPRPN
251-300	DPTVGKCNDP YPGNNNNgVK GFSyLDGVNT wlGRTISIAS RSgYEmLKvP
301-350	NaLTDDKSKP TQGQTivLNT DWsGYSgSfm DYWAEGECYR aCfYvelIRG
351-400	RPKEDKVWWT SNsIvsMCSS TEFLGQWDWP DGAKIEYFL

Predicted result format: EPITOPE RESIDUE | NON-EPITOPE RESIDUE | core residue

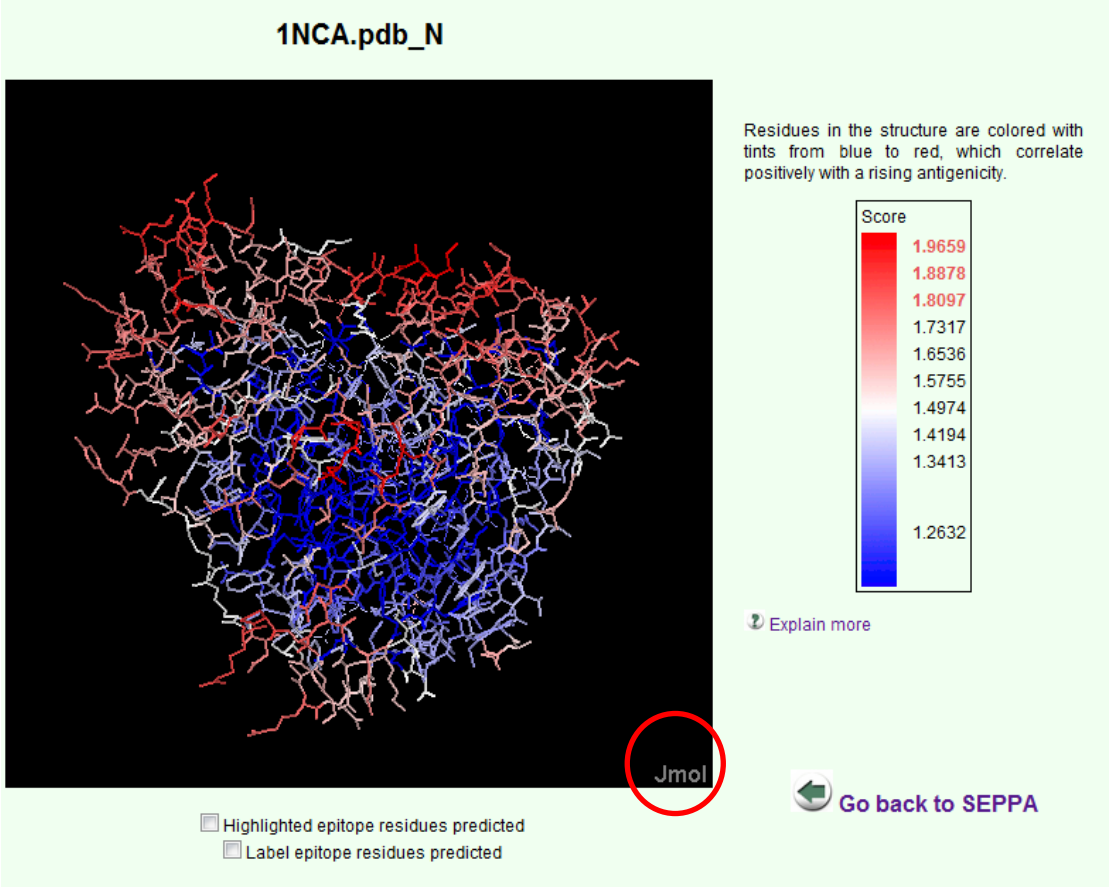
Download the score file

Explain the result

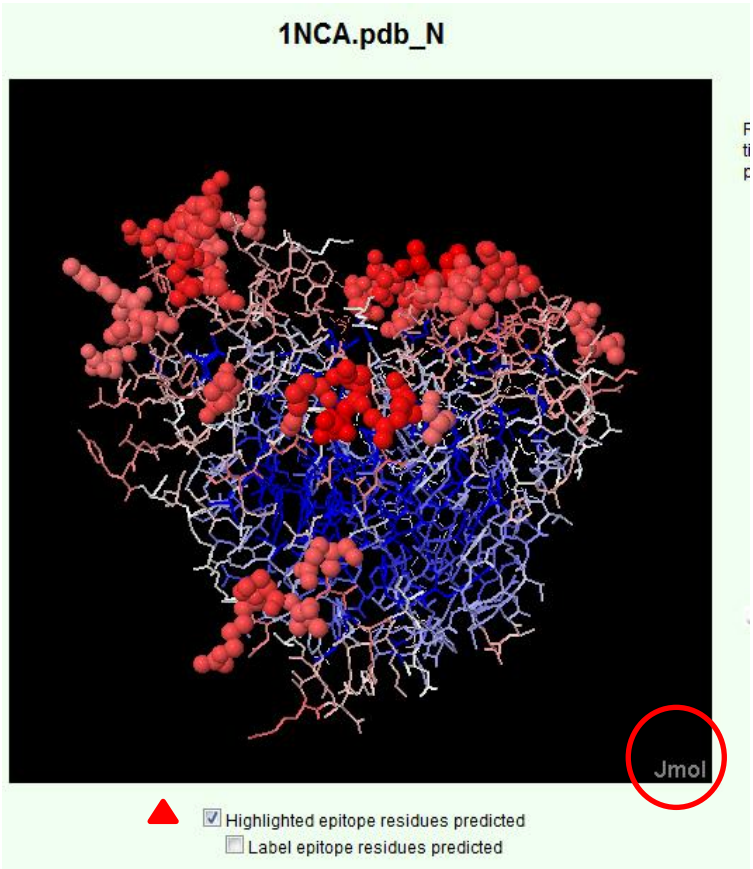
(1) Summary for the prediction result.

Chain: N
Threshold: 1.80
Number of total residues: 389
Number of predicted epitope residues: 39

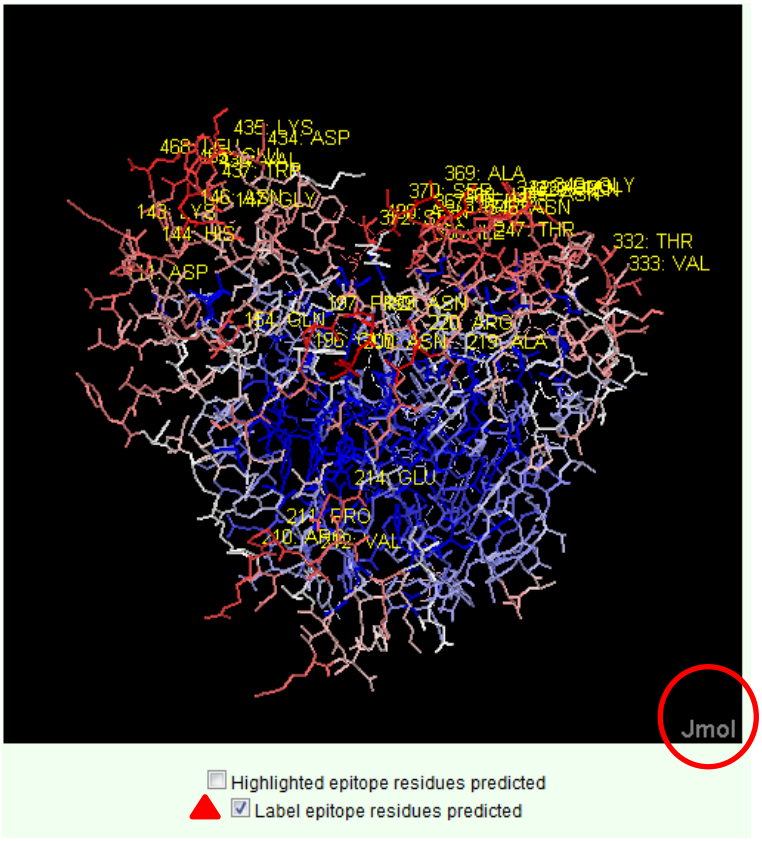
(2) The link to visualization of the prediction result (red circle).
You can adjust the perspective of the structure with your mouse.



Note 2: **Tints from blue to red represent a rising antigenicity**

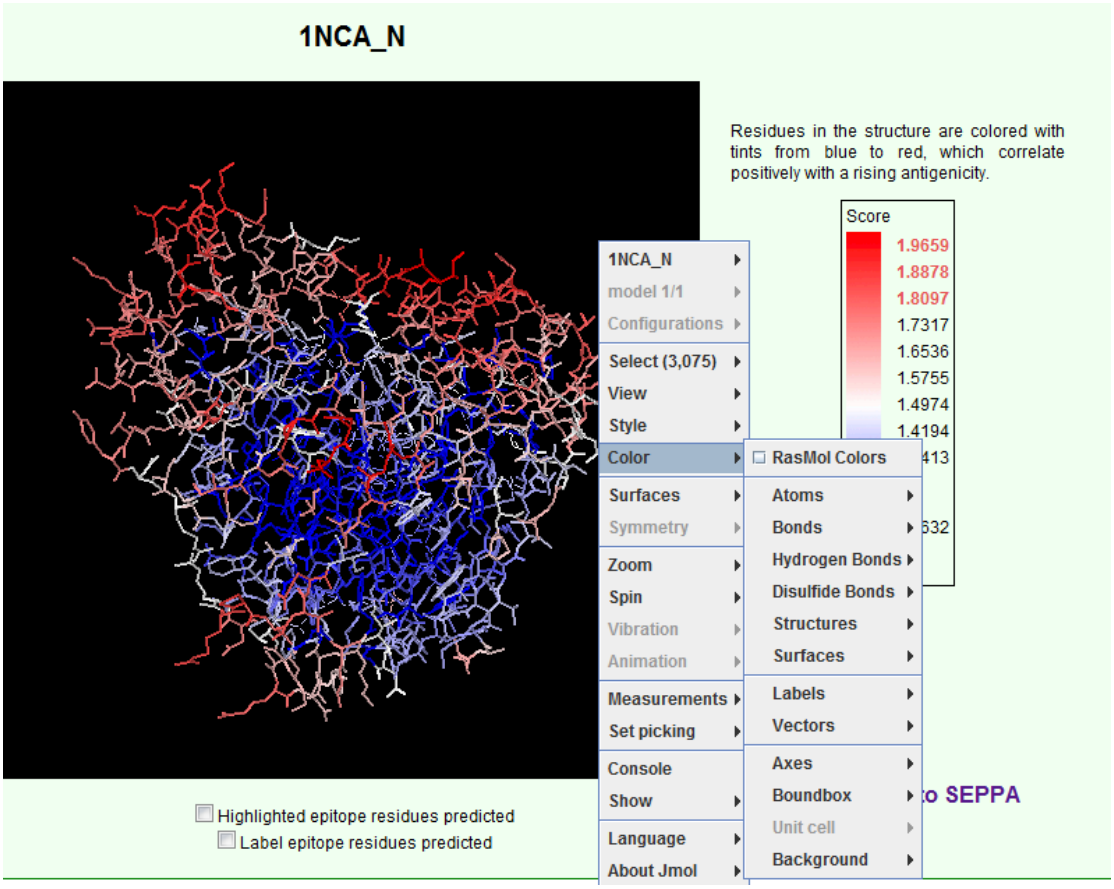


Selecting the "Highlighted epitope residues predicted" checkbox



Selecting the "Label epitope residues predicted" checkbox

You can also click the word "Jmol" in the lower right corner to access the Jmol menu which contains more choses such as zooming in and out, viewing with different perspective, changing colors and style.



(3) A glance of the prediction result

1-50	IRDFNNLTKG LCTiNSWHIY GKDNAvRIgE DSDVLvTREp YvsCDPDECR
51-100	fyaLSQGTTI RGKHSNGTIH DRSQYRAlIs WPLSSPPTVY NSRVECIGWS
101-150	stsCHDgKTR MSiciSGPNN NaSaViWYNR RPVTEINTWA RNIlRTQEsE
151-200	CVCHNgVCPv VfTdGSATGP AETRiYyfKE gKILKWEPLA GTAKHIEECS
201-250	CYgERAEITc tcRdNWQGsN RpViRIDPVA MTHTSQyICS pVLTdNPRPN
251-300	DPTVGKCNDP YPGNNNNgVK GFSyLDGVNT wlGRTISIAS RSgYEmLKvP
301-350	NaLTDDKSKP TQGQTivLNT DWsGYSgSfm DYWAEGECYR aCfYvelIRG
351-400	RPKEDKVWWT SNsIvsMCSS TEFLGQWDWP DGAKIEYFL

Note 3:

Residues are listed sequentially.

The predicted epitope residues are highlighted in yellow.

The core residues are shown in lowercase

(4) You can download the complete score file through the link circled blue in the figure for output page. The head contains the summary of the prediction followed by the score for each residue (like below).

```
SEPPA(Spatial Epitope Prediction of Protein Antigens)
[Designed for B-cell Conformational Epitope Prediction]

Thu Oct 11 01:53:13 2012


----SEPPA Prediction Result-----

Antigenic Prediction for 1NCA.pdb:

Chain: N
Threshold: 1.80
Number of total residues: 389
Number of predicted epitope residues: 39

chainID resSeq  resName score

N    81  ILE 1.69
N    82  ARG 1.79
N    83  ASP 1.66
N    84  PHE 1.65
N    85  ASN 1.50
N    86  ASN 1.53
N    87  LEU 1.43
N    88  THR 1.54
N    89  LYS 1.50
N    90  GLY 1.47
N    91  LEU 1.39
```

4. Multiple PDB ID entries can be submitted in batch query. Each entry should include PDB ID and chain ID(s), which are separated with space(s) in one line. After you set the parameter and submit your queries (like below), you can see a page which combines the prediction result for each query like above.

SEPPA Server | Batch query | Example | Help | Contact information

Batch query with structures of existing PDB IDs:

- Enter PDB IDs and chains:

```
1NCA N
1A14 N
1NDM C
```

Please specify a threshold:

Threshold: 1.80

Submit Reset