

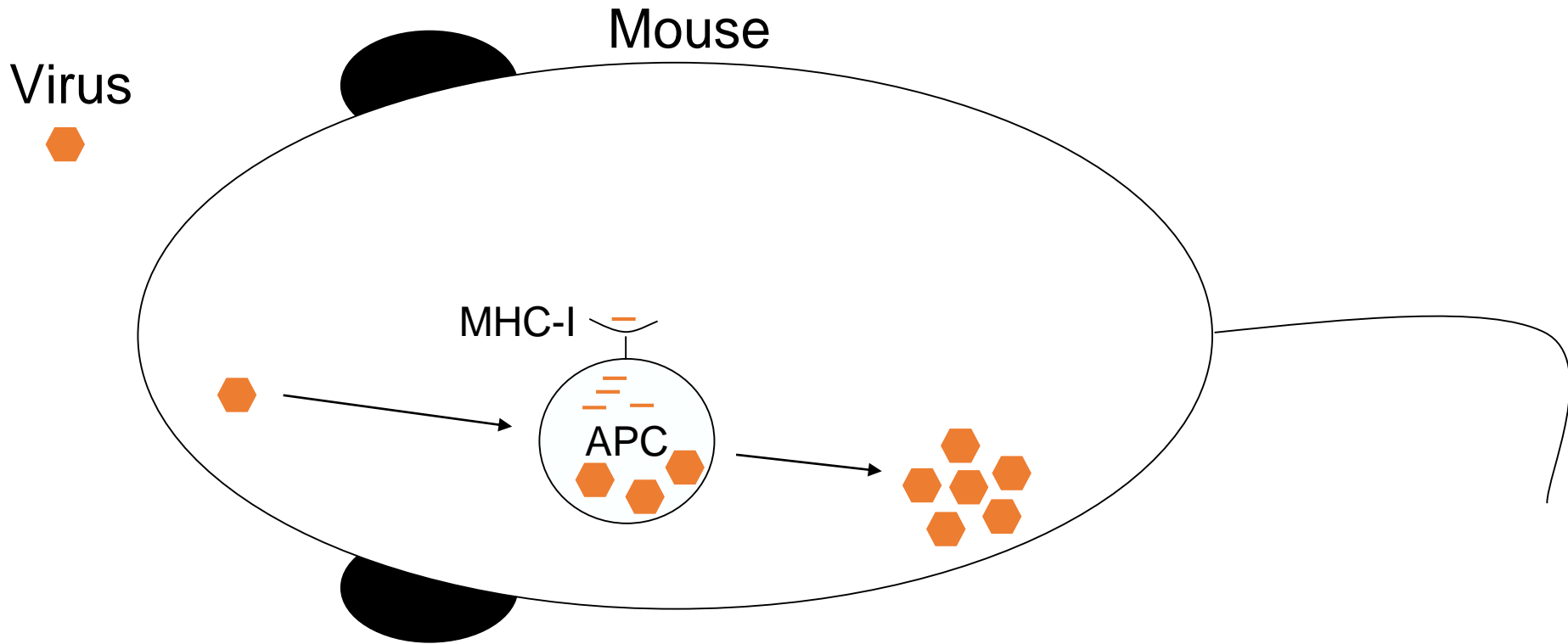


# T Cell processing & immunogenicity predictions

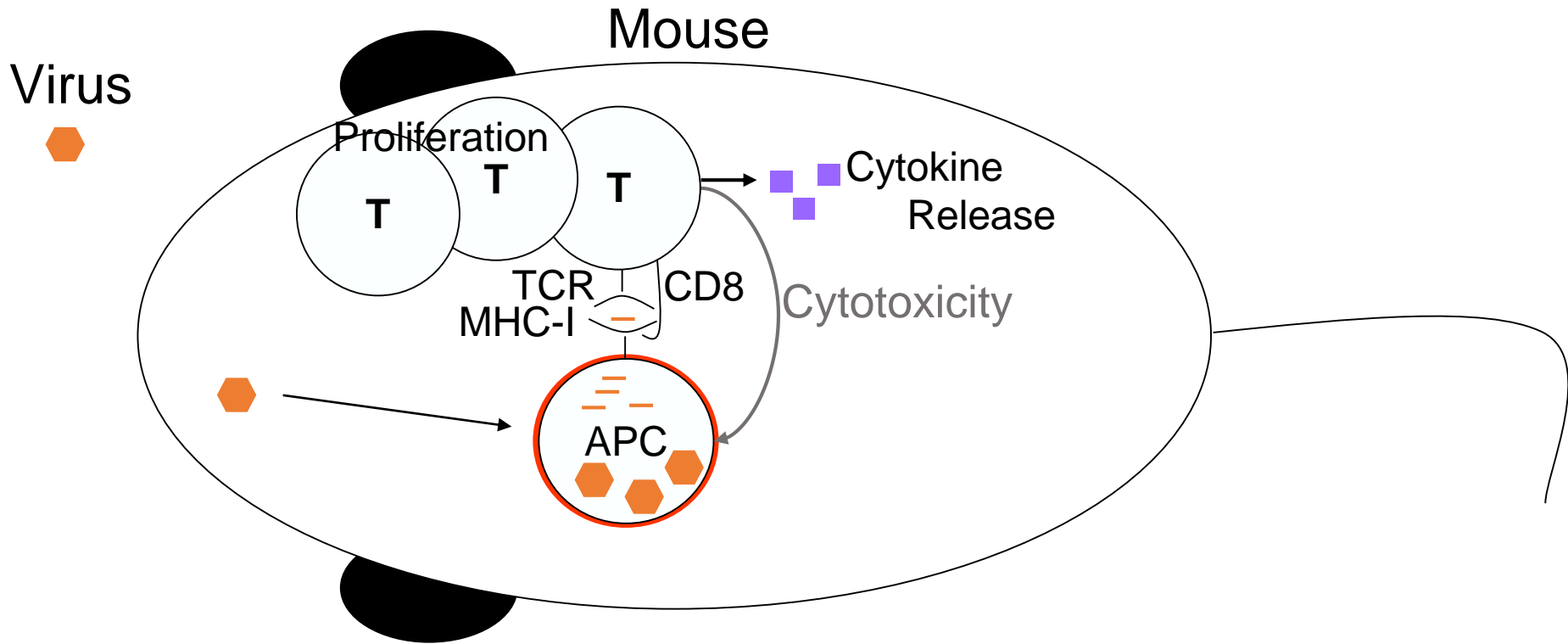
[tools.iedb.org](https://tools.iedb.org)

Presented by:  
Bjoern Peters, PI  
Austin Crinklaw, Research Technician

# CD8+ T cell epitopes in viral infection



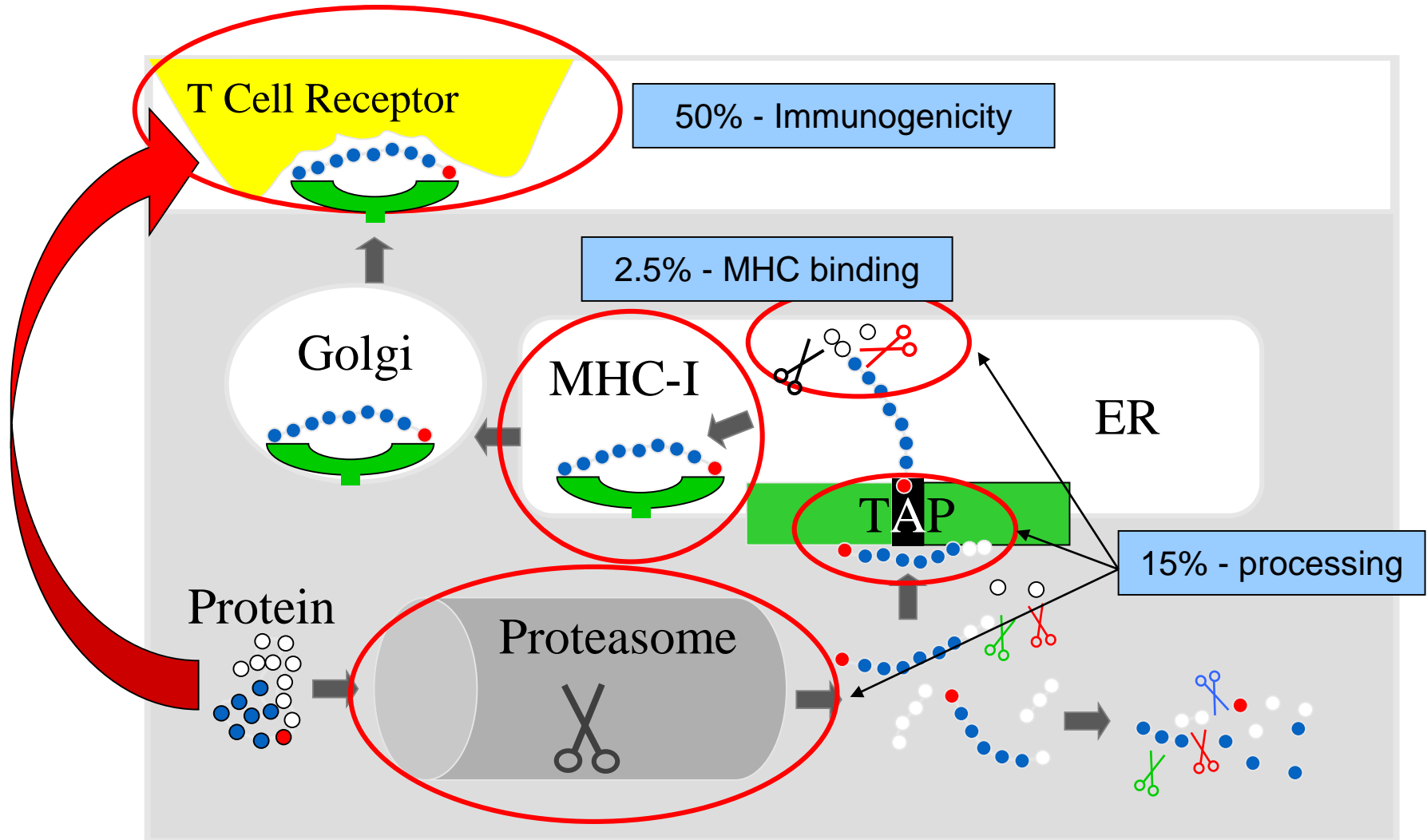
# CD8+ T cell epitopes in viral infection



T cell  
epitope  
mapping

ORF 1	M G Q   V T M F E A L P H I   <b>I D E V I N I V I</b>   I V L   V I T G I K A V Y N ...
ORF 2	M G L K G P D   Y K G V Y Q F K S V E F D M S H L N L T M P N A C S A N N ...
ORF 3	M H N F C N L T S A F N K K T F D H T L M S I V S S L H L S I D G N S N Y ...
ORF 4	M S A Q S Q C R T F R G R V L D M F R T A F G G K Y M R S G W G W T G S D ...
ORF 5	M H C T Y A G P F G M S R I L L S Q E K T K F F T R R L A G T F T W T L S ...
ORF 6	M K C F G N T A V A K C N V N H D A E F C D M L R L I D Y N K A A L S K F ...
ORF 7	M L M R N H L   <b>L D L M G V P Y C N Y</b>   S K F W Y L E H A K T G E T S V P K C ...

# MHC I - Antigen processing and presentation pathway



Peters et al, J Mol Biol 2002, Bioinformatics 2003, J Immunol. 2003; CMLS 2005;

Assarson, J Immunol 2007

# Class I Processing + immunogenicity tools available in the IEDB

<http://tools.iedb.org/main/tcell/>

## IEDB Analysis Resource

Overview **T Cell Tools** B Cell Tools Analysis Tools Tools-API Usage Download Datasets Contribute Tools References

### T Cell Epitopes - Processing Prediction

These tools predict epitope candidates based upon the processing of peptides in the cell.

#### [Proteasomal cleavage/TAP transport/MHC class I combined predictor](#)

This tool combines predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential of being a T cell epitope.

#### [Neural network based prediction of proteasomal cleavage sites \(NetChop\) and T cell epitopes \(NetCTL and NetCTLpan\)](#)

NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL and NetCTLpan are predictors of T cell epitopes along a protein sequence. It also employs a neural network architecture.

#### [MHC-NP: Prediction of peptides naturally processed by the MHC](#)

MHC-NP employs data obtained from MHC elution experiments in order to assess the probability that a given peptide is naturally processed and binds to a given MHC molecule. This tool was the winner of the [2nd Machine Learning Competition in Immunology](#).



#### [MHCII-NP:](#)

This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences.

### T Cell Epitopes - Immunogenicity Prediction

This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.

#### [T cell class I pMHC immunogenicity predictor](#)

This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a class I peptide MHC (pMHC) complex.



#### [Deimmunization:](#)

The deimmunization tool is attempt to identify immunodominant regions in a given therapeutically important protein, and suggest amino-acid substitutions that create non-immunogenic versions of the proteins. So we have opted a two steps process; 1) In the first step, the deimmunization tool will list all the immunogenic regions or peptides based on selected threshold. These peptides will be generated from the protein with 15mer window size and 10mer overlap. 2) In the second step, the user can select one or more peptides listed in the results and final result window will display the non-immunogenic substitution of each selected peptides. The default threshold is 8.5 (which is difference in the median of percentile rank from 26 reference alleles set for MHC class II). In the final result window, the tools will also take care of the fact that non-immunogenic substitution in the immunogenic peptides, should not create new immunogenic site in the neighboring peptides. Therefore, the result window will also display the effect of substitution on the neighboring peptides.



#### [CD4 T cell immunogenicity prediction:](#)

The server is developed to predict the allele independent CD4 T cell immunogenicity at population level. User can predict the T cell immunogenicity using 7-allele method (Paul et. al. 2015), immunogenicity method and combined method (IEDB recommended). The combined method predicts the final score that combines the predictions from 7-allele method and immunogenicity method.

# Class I 'combined predictor'

<http://tools.iedb.org/processing/>

Home Help Example Reference Download Contact

## MHC-I Processing Predictions

Prediction Method Version 2013-02-22 [\[Older versions\]](#)

**Specify Sequence(s)**

Enter protein sequence(s) in FASTA format  
[\(Browse for sequences in NCBI\)](#)

Or select file containing sequence(s)  No file selected.

Choose sequence format

**Choose a Prediction Method**

Prediction Method  [Help on prediction method selections](#)

**Specify what to make binding predictions for**

MHC source species

Show only frequently occurring alleles:  [?](#)  
Select MHC allele(s)    [Upload allele file](#) [?](#)

**Proteasomal cleavage prediction**

Specify proteasome type

**TAP transport predictions**

Maximum precursor extension

Alpha factor

**Specify Output**

Output format

- Combines predictions for:
  - proteasomal cleavage
  - TAP transport
- Trained on specific in vitro datasets

# Proteasomal cleavage

<http://tools.iedb.org/processing/>

Proteasomal cleavage prediction	
Specify proteasome type	<input type="text" value="immuno"/> → <input type="text" value="immuno"/> <input type="text" value="constitutive"/>
TAP transport predictions	
Maximum precursor extension	<input type="text" value="1"/>
Alpha factor	<input type="text" value="0.2"/>

- Proteasomes create the C-terminal end of peptides
- Prediction looks for sequence motive up and downstream of potential cleavage site
- Cleavage sequence motif determined based on in vitro protein digests by proteasomes

# TAP Transport

<http://tools.iedb.org/processing/>

Proteasomal cleavage prediction	
Specify proteasome type	<input type="text" value="immuno"/>
TAP transport predictions	
Maximum precursor extension	<input type="text" value="1"/>
Alpha factor	<input type="text" value="0.2"/>

- TAP transport efficiency of peptides is sequence dependent; motif derived based on in vitro assays
- Overall TAP transport efficiency of a presented MHC ligand can be result of a collection of precursors
- Unless paper specifically read and details about the precursor length distribution are known, **keep parameters unchanged**



# Class I 'combined predictor' - example

<http://tools.iedb.org/processing/>

**MHC-I Processing Predictions**

Prediction Method Version: 2013-02-22 [\[Older versions\]](#)

**Specify Sequence(s)**

Enter protein sequence(s) in FASTA format ([Browse for sequences in NCBI](#))

```
>LCMV Armstrong, Protein GP
MGQIVTMFEALPHIIDEVINIVIIVLIVITGIRKAVYNFATCGIFALISFLLLAGRSCGM
YGLKGPDIYKGVYQFKSVEFDMSHLNLTPNACSAANSHHYISMGTSGLELTFNTDSII
SHNFCNLTSAPNKKTFDHTLMSIVSSLHLSIRGNSNYKAVSCDFNNGITIQYNLTFSDA
QSAQSQCRTRFRGRVLDMPRTAFGGKYMRSWGWTGSDGKTTWCSQTSYQYLIIQNRWE
NHCTYAGPFGMSRIILLSQEKTKFFTRRLAGFTWTLSDDSGVENPGGYCLTKWMLAAE
LKCFGNTAVAKCNVNHDAEFCDMLRLIDYNKAALSFKFEDVESALHLFKTTVNSLISDQ
LLMRNHLRDLMGVPPYCNYSKFWYLEHARTGETSVPKCWLVTNGSYLNETHPSDQIEQEA
DNMITEMLRKDYIKRQGSTFLALMDLDMFSTSAYLVSIFLHLVKIPTHRIKGGSCPKP
HRLTRKIGICSGAFKVPVGVKTVWKRK
```

Or select file containing sequence(s)  No file selected.

Choose sequence format: auto detect format

**Choose a Prediction Method**

Prediction Method: IEDB recommended [Help on prediction method selections](#)

**Specify what to make binding predictions for**

MHC source species: mouse

Select MHC allele(s):  
Allele: H-2-Kb Length: 10  ?

**Proteasomal cleavage prediction**

Specify proteasome type: immuno

**TAP transport predictions**

Maximum precursor extension: 1

Alpha factor: 0.2

**Specify Output**

Output format: XHTML table



# Class I 'combined predictor' - example

## MHC-I Processing Prediction Results

<http://tools.iedb.org/processing/>

Input Sequences

#	Name	Sequence
1	LCMV Armstrong, Protein GP	MGQIVTMFEALPHIIDEVINIVIIVLIVITGIKAVYNFATCGIFALISFLLLAGRS CGMYGLKGPDIYKGVYQFKSVEFDMSHLNLTMPNACSANNHHYISMGTSGLELTF TNDSEIISHNFCNLTSAFNKKTFDHTLMSIVSSLHLSIRGNSNYKAVSCDFNNGITI QYNLTFSDAQSAQSQCRTFRGRVLDMFRTAFGGKYMRSGWGTGSDGKTWCSQTS YQYLIIQNRTWENHCTYAGPFGMSRILLSQEKTKFTRRLAGTFTWTLSDSSGVEN PGGYCLTKWMI LAAELKCFGNTAVAKCNVNHDAEFCDMLRLIDYNKAALSKFKEDV ESALHLFKTTVNSLISDQLLMRNHLRDLMGVPYCNYSKFWYLEHAKTGETSVPKCW LVTNGSYLNETHFSDOIEQEADNMI TEMLRKDYIKRQGSTPLALMDLLMFST SAYL VSI FLHLVKIPTHRRHIGGSCPKPHRLTNKGICSCGAFKVPGVKTVVKRR
2	LCMV Armstrong, Protein NP	MSLSKEVKSFWQTQALRRELQSFTSDVKA AVIKDATNLLNGLDFSEVSNVQIRMRK EKRDDKDLQRLRSLNQTVHSLVDLKSTSKKNVLKVGRLSAEELMSLAADLEKLRKAK IMRSERPQASGVYMGNLTTQQLDQRSQILQIVGMRKPQQGASGVVRVWDVKDSLL NNQFGTMPSLTMACMAKQSTPLNDVVQALTDLGLLYTVKYPNLNDRERLKDHPV LGVITEQQSSINISGYNFSLGA AVKAGAALLDGGNMLESILIKPSNSEDLLKAVLG AKRKLNMFVSDQVGDNRNYPENILYKVCLSGEGWPYIACRTSIVGRAWENTIDLTS EKPAVNSPRPAPGAAGPPQVGLSYSQTMLLKDLKDLGGIDPNAPTWIDIEGRFNDPVE IAIFQPQNGQFIHFYREPV DQKQFKQDSKYSHGMDLADLFNAQPGLTSSVIGALPQ GMVLS CQGSDDIRKLLDSQNRKDIKLIDVEMTREASREYEDKVDWVKYGLWCKMHTG IVRDKKKEITPHCALMDCII FESASKARLPDLKTVHNILPHDLIFRGNVVTL

Prediction method: recommended | High Score = high efficiency

[Download result](#)

### Citations

Allele	#	Start	End	Peptide Length	Peptide	Proteasome Score	TAP Score	MHC Score	Processing Score	Total Score	MHC IC50[nM]
H-2-Kb	2	203	212	10	LLYTVKYPNL	1.79	0.50	-2.01	2.28	0.27	103.5
H-2-Kb	1	116	125	10	SIISHNFCNL	1.51	0.46	-1.90	1.97	0.08	78.6
H-2-Kb	2	7	16	10	VKSFWTQAL	1.42	0.39	-2.01	1.81	-0.21	102.9
H-2-Kb	2	235	244	10	INISGYNFSL	1.53	0.41	-2.18	1.95	-0.23	149.7
H-2-Kb	1	35	44	10	VYNFATCGIF	1.42	1.27	-3.12	2.68	-0.44	1327
H-2-Kb	1	75	84	10	KSVEFDMSHL	1.50	0.47	-2.45	1.96	-0.49	283.6
H-2-Kb	1	369	378	10	YCNYSKFWYL	1.45	0.31	-2.32	1.76	-0.56	206.8
H-2-Kb	1	449	458	10	VSI FLHLVKI	1.33	0.33	-2.25	1.66	-0.59	178.2

# Class I 'combined predictor' - example

<http://tools.iedb.org/processing/>

Allele	#	Start	End	Peptide Length	Peptide	Proteasome Score	TAP Score	MHC Score	Processing Score	Total Score	MHC IC50[nM]
H-2-Kb	2	203	212	10	LLYTVKYPNL	1.79	0.50	-2.01	2.28	0.27	103.5
H-2-Kb	1	116	125	10	SIISHNFCNL	1.51	0.46	-1.90	1.97	0.08	78.6
H-2-Kb	2	7	16	10	VKSFQWTQAL	1.42	0.39	-2.01	1.81	-0.21	102.9
H-2-Kb	2	235	244	10	INISGYNFSL	1.53	0.41	-2.18	1.95	-0.23	149.7
H-2-Kb	1	35	44	10	VYNFATCGIF	1.42	1.27	-3.12	2.68	-0.44	1327
H-2-Kb	1	75	84	10	KSVEFDMSHL	1.50	0.47	-2.45	1.96	-0.49	283.6
H-2-Kb	1	369	378	10	YCNYSKFWYL	1.45	0.31	-2.32	1.76	-0.56	206.8
H-2-Kb	1	449	458	10	VSIFLHLVKI	1.33	0.33	-2.25	1.66	-0.59	178.2

- Higher scores = higher efficiency for MHC-I presentation
- MHC binding score =  $-\log_{10}(\text{IC}_{50})$  (sign change)
- Combined scores are additive
  - Processing = proteasome + TAP
  - Total = proteasome + TAP + MHC
- Different variance in scores reflects different selectivity

# Caveats / performance of processing predictions

- Processing predictions beat MHC binding predictions when predicting **eluted peptides**
- No clear evidence that processing predictions are better at predicting **epitopes**
- Eluted peptides may over represent ‘best possible’ ligands, and the difference in processing may not be relevant in practice
- Co-evolution of MHC molecules to bind peptides with motifs that are generated by proteasome and TAP means that most high affinity MHC binders are also efficiently processed
- Recommendation: Use MHC binding predictions alone by default
  - If resources require limiting the number of peptides considered, use total score of processing predictions as additional filter

# Class I Processing + immunogenicity tools available in the IEDB

<http://tools.iedb.org/main/tcell/>

## IEDB Analysis Resource

Overview **T Cell Tools** B Cell Tools Analysis Tools Tools-API Usage Download Datasets Contribute Tools References

### T Cell Epitopes - Processing Prediction

These tools predict epitope candidates based upon the processing of peptides in the cell.

#### [Proteasomal cleavage/TAP transport/MHC class I combined predictor](#)

This tool combines predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential of being a T cell epitope.

#### [Neural network based prediction of proteasomal cleavage sites \(NetChop\) and T cell epitopes \(NetCTL and NetCTLpan\)](#)

NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL and NetCTLpan are predictors of T cell epitopes along a protein sequence. It also employs a neural network architecture.

#### [MHC-NP: Prediction of peptides naturally processed by the MHC](#)

MHC-NP employs data obtained from MHC elution experiments in order to assess the probability that a given peptide is naturally processed and binds to a given MHC molecule. This tool was the winner of the [2nd Machine Learning Competition in Immunology](#).

#### [MHCII-NP:](#)

This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences.

### T Cell Epitopes - Immunogenicity Prediction

This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.

#### [T cell class I pMHC immunogenicity predictor](#)

This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a class I peptide MHC (pMHC) complex.

#### [Deimmunization:](#)

The deimmunization tool is attempt to identify immunodominant regions in a given therapeutically important protein, and suggest amino-acid substitutions that create non-immunogenic versions of the proteins. So we have opted a two steps process; 1) In the first step, the deimmunization tool will list all the immunogenic regions or peptides based on selected threshold. These peptides will be generated from the protein with 15mer window size and 10mer overlap. 2) In the second step, the user can select one or more peptides listed in the results and final result window will display the non-immunogenic substitution of each selected peptides. The default threshold is 8.5 (which is difference in the median of percentile rank from 26 reference alleles set for MHC class II). In the final result window, the tools will also take care of the fact that non-immunogenic substitution in the immunogenic peptides, should not create new immunogenic site in the neighboring peptides. Therefore, the result window will also display the effect of substitution on the neighboring peptides.

#### [CD4 T cell immunogenicity prediction:](#)

The server is developed to predict the allele independent CD4 T cell immunogenicity at population level. User can predict the T cell immunogenicity using 7-allele method (Paul et. al. 2015), immunogenicity method and combined method (IEDB recommended). The combined method predicts the final score that combines the predictions from 7-allele method and immunogenicity method.



# Neural Network based predictors

<http://tools.iedb.org/netchop/>

IEDB Analysis Resource

Home Help Example Reference Download Contact

## Proteasomal Cleavage Prediction

NetChop/NetCTL/NetCTLpan

Specify Sequence(s)

Prediction Method: NetChop

Enter protein sequence(s) in FASTA format

Or select file containing sequence(s): Browse... No file selected.

Method Specific Options

Method: C term 3.0

Threshold: 0.5

Submit Reset

- **NetChop:** proteasomal cleavage
- **NetCTL:** combines NetChop, TAP transport, NetMHC
- **NetCTLpan:** combines NetChop, TAP transport, NetMHCpan

Key difference is the use of NetChop

# Class I Processing + immunogenicity tools available in the IEDB

<http://tools.iedb.org/main/tcell/>

## IEDB Analysis Resource

Overview **T Cell Tools** B Cell Tools Analysis Tools Tools-API Usage Download Datasets Contribute Tools References

### T Cell Epitopes - Processing Prediction

These tools predict epitope candidates based upon the processing of peptides in the cell.

#### [Proteasomal cleavage/TAP transport/MHC class I combined predictor](#)

This tool combines predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential of being a T cell epitope.

#### [Neural network based prediction of proteasomal cleavage sites \(NetChop\) and T cell epitopes \(NetCTL and NetCTLpan\)](#)

NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL and NetCTLpan are predictors of T cell epitopes along a protein sequence. It also employs a neural network architecture.

#### [MHC-NP: Prediction of peptides naturally processed by the MHC](#)

MHC-NP employs data obtained from MHC elution experiments in order to assess the probability that a given peptide is naturally processed and binds to a given MHC molecule. This tool was the winner of the [2nd Machine Learning Competition in Immunology](#).

#### [MHCII-NP:](#)

This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences.

### T Cell Epitopes - Immunogenicity Prediction

This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.

#### [T cell class I pMHC immunogenicity predictor](#)

This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a class I peptide MHC (pMHC) complex.

#### [Deimmunization:](#)

The deimmunization tool is attempt to identify immunodominant regions in a given therapeutically important protein, and suggest amino-acid substitutions that create non-immunogenic versions of the proteins. So we have opted a two steps process; 1) In the first step, the deimmunization tool will list all the immunogenic regions or peptides based on selected threshold. These peptides will be generated from the protein with 15mer window size and 10mer overlap. 2) In the second step, the user can select one or more peptides listed in the results and final result window will display the non-immunogenic substitution of each selected peptides. The default threshold is 8.5 (which is difference in the median of percentile rank from 26 reference alleles set for MHC class II). In the final result window, the tools will also take care of the fact that non-immunogenic substitution in the immunogenic peptides, should not create new immunogenic site in the neighboring peptides. Therefore, the result window will also display the effect of substitution on the neighboring peptides.

#### [CD4 T cell immunogenicity prediction:](#)

The server is developed to predict the allele independent CD4 T cell immunogenicity at population level. User can predict the T cell immunogenicity using 7-allele method ([Paul et. al. 2015](#)), immunogenicity method and combined method (IEDB recommended). The combined method predicts the final score that combines the predictions from 7-allele method and immunogenicity method.





# MHC-NP: Prediction of peptides naturally processed by the MHC

<http://tools.iedb.org/mhcnp/>

IEDB Analysis Resource

Home Help Example Reference Download Contact

## MHC-NP: Prediction of Peptides Naturally Processed by the MHC

Developed by: Sébastien Giguère, Alexandre Drouin, Alexandre Lacoste, Mario Marchand, Jacques Corbeil and François Lavolette

Specify Sequence(s)

Enter protein sequence(s) in FASTA format  
[\(Browse for sequences in NCB\)](#)

Or select file containing sequence(s)  No file selected.

Choose a Prediction Method

Prediction Method  [Help on prediction method selections](#)

MHC source species

Select MHC allele(s) Allele  Length  [Upload allele file](#) ?

Specify Output

Sort peptides by

Output format

Pan-predictions trained on both binding + eluted ligand data now available!

*J Immunol.* 2017 Nov 1;199(9):3360-3368. doi: 10.4049/jimmunol.1700893. Epub 2017 Oct 4.

**NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data.**

Jurtz V<sup>1</sup>, Paul S<sup>2</sup>, Andreatta M<sup>3</sup>, Marcatili P<sup>1</sup>, Peters B<sup>2</sup>, Nielsen M<sup>4,3</sup>.

PMID: 28978689 PMCID: PMC5679736 DOI: 10.4049/jimmunol.1700893



# Class I Processing + immunogenicity tools available in the IEDB

<http://tools.iedb.org/main/tcell/>

## IEDB Analysis Resource

Overview **T Cell Tools** B Cell Tools Analysis Tools Tools-API Usage Download Datasets Contribute Tools References

### T Cell Epitopes - Processing Prediction

These tools predict epitope candidates based upon the processing of peptides in the cell.

#### Proteasomal cleavage/TAP transport/MHC class I combined predictor

This tool combines predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential of being a T cell epitope.

#### Neural network based prediction of proteasomal cleavage sites (NetChop) and T cell epitopes (NetCTL and NetCTLpan)

NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL and NetCTLpan are predictors of T cell epitopes along a protein sequence. It also employs a neural network architecture.

#### MHC-NP: Prediction of peptides naturally processed by the MHC

MHC-NP employs data obtained from MHC elution experiments in order to assess the probability that a given peptide is naturally processed and binds to a given MHC molecule. This tool was the winner of the [2nd Machine Learning Competition in Immunology](#).



#### MHCII-NP:

This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences.

### T Cell Epitopes - Immunogenicity Prediction

This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.

#### T cell class I pMHC immunogenicity predictor

This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a class I peptide MHC (pMHC) complex.



#### Deimmunization:

The deimmunization tool is attempt to identify immunodominant regions in a given therapeutically important protein, and suggest amino-acid substitutions that create non-immunogenic versions of the proteins. So we have opted a two steps process; 1) In the first step, the deimmunization tool will list all the immunogenic regions or peptides based on selected threshold. These peptides will be generated from the protein with 15mer window size and 10mer overlap. 2) In the second step, the user can select one or more peptides listed in the results and final result window will display the non-immunogenic substitution of each selected peptides. The default threshold is 8.5 (which is difference in the median of percentile rank from 26 reference alleles set for MHC class II). In the final result window, the tools will also take care of the fact that non-immunogenic substitution in the immunogenic peptides, should not create new immunogenic site in the neighboring peptides. Therefore, the result window will also display the effect of substitution on the neighboring peptides.



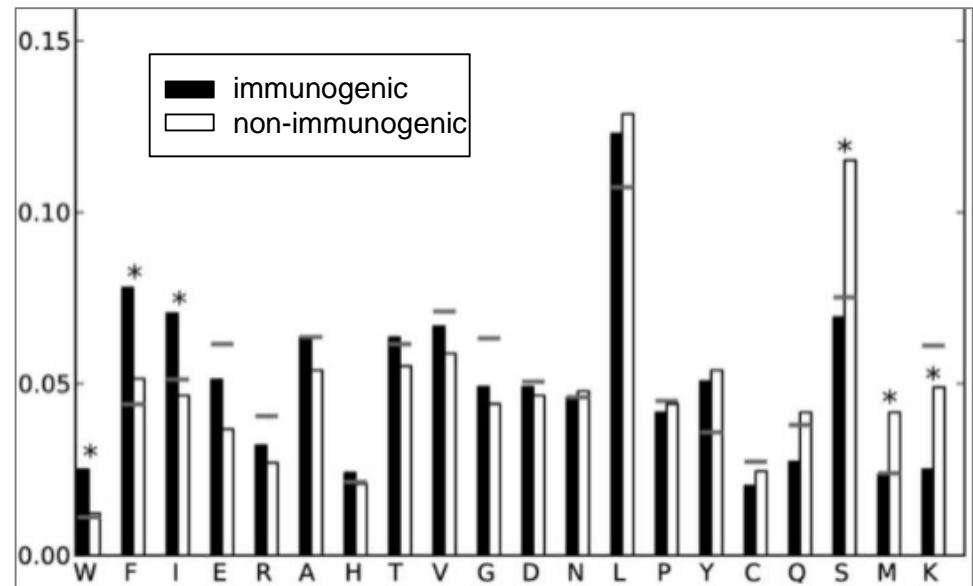
#### CD4 T cell immunogenicity prediction:

The server is developed to predict the allele independent CD4 T cell immunogenicity at population level. User can predict the T cell immunogenicity using 7-allele method (Paul et. al. 2015), immunogenicity method and combined method (IEDB recommended). The combined method predicts the final score that combines the predictions from 7-allele method and immunogenicity method.



# Class I immunogenicity prediction

- Approach: Assemble two datasets of peptides with similar MHC binding affinity, that are (i) recognized or (ii) not recognized by T cells
- Enrichment of W,F,I and depletion of S,M,K in immunogenic peptides
- Use enrichments to calculate propensity scores



[PLoS Comput Biol. 2013 Oct;9\(10\):e1003266. doi: 10.1371/journal.pcbi.1003266. Epub 2013 Oct 24.](https://doi.org/10.1371/journal.pcbi.1003266)

## Properties of MHC class I presented peptides that enhance immunogenicity.

[Calis JJ<sup>1</sup>](#), [Maybeno M](#), [Greenbaum JA](#), [Weiskopf D](#), [De Silva AD](#), [Sette A](#), [Keşmir C](#), [Peters B](#).

PMID: 24204222 PMCID: [PMC3808449](https://pubmed.ncbi.nlm.nih.gov/PMC3808449/) DOI: [10.1371/journal.pcbi.1003266](https://doi.org/10.1371/journal.pcbi.1003266)

# Class I immunogenicity prediction -example

<http://tools.iedb.org/immunogenicity/>

## IEDB Analysis Resource

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

### Class I Immunogenicity

Specify sequence(s) \*

Enter peptide sequence(s)  
[\(Browse for sequences in NCBI\)](#)

```
FIAGLIAIV
LITGRLQSL
RLNEVAKNL
KAVYNFATC
FQPQNGQFI
```

Or select file containing sequence(s)

No file selected.

Choose which positions to mask

Specify which positions to mask

Default (1st, 2nd, and C-terminus amino acids)  
 Custom  (Comma separated numbers)  
Peptide lengths must be equal when using custom masking.

\*The tool was only validated for 9-mer peptides. However, predictions can be made for peptides of any length.

**Mask positions that are MHC anchors**



# Class I immunogenicity prediction -example

<http://tools.iedb.org/immunogenicity/>


## IEDB Analysis Resource

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

### Class I Immunogenicity

Masking: **default**  
Masked variables: [1, 2, 'cterm']

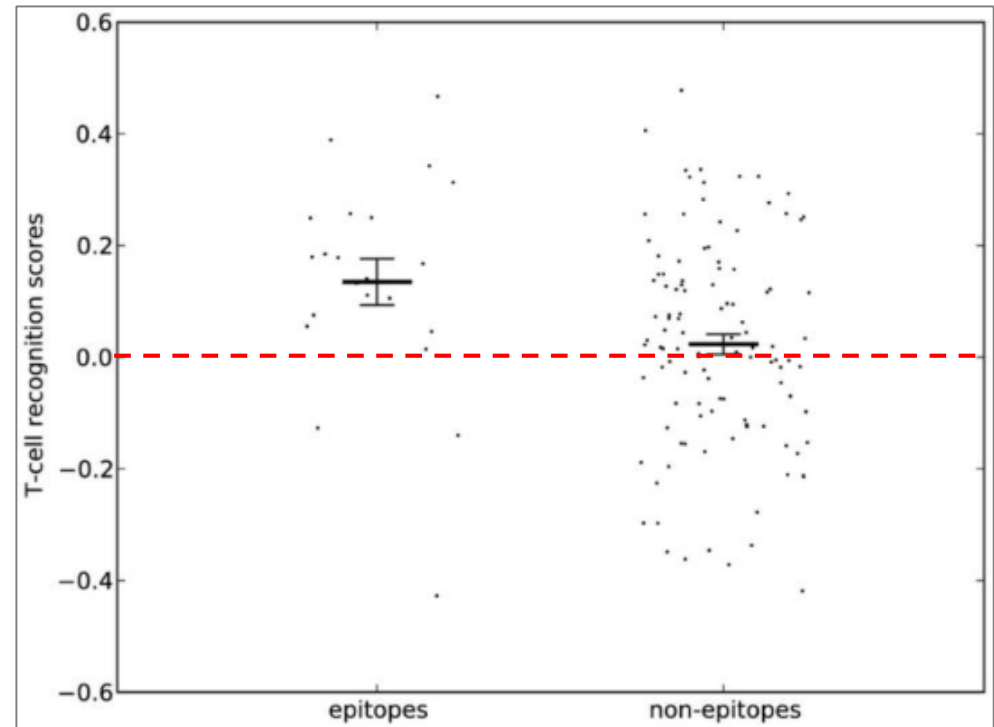
Peptide	Length	Score
FIAGLIAIV	9	0.27206
KAVYNFATC	9	0.16928
RLNEVAKNL	9	-0.0101
LITGRLQSL	9	-0.10776
FQPQNGQFI	9	-0.12392

[Download result](#) 

- Scores are sums of propensity scores at all unmasked positions
- High scores = peptide is more likely to be immunogenic

# Class I immunogenicity prediction caveats / performance

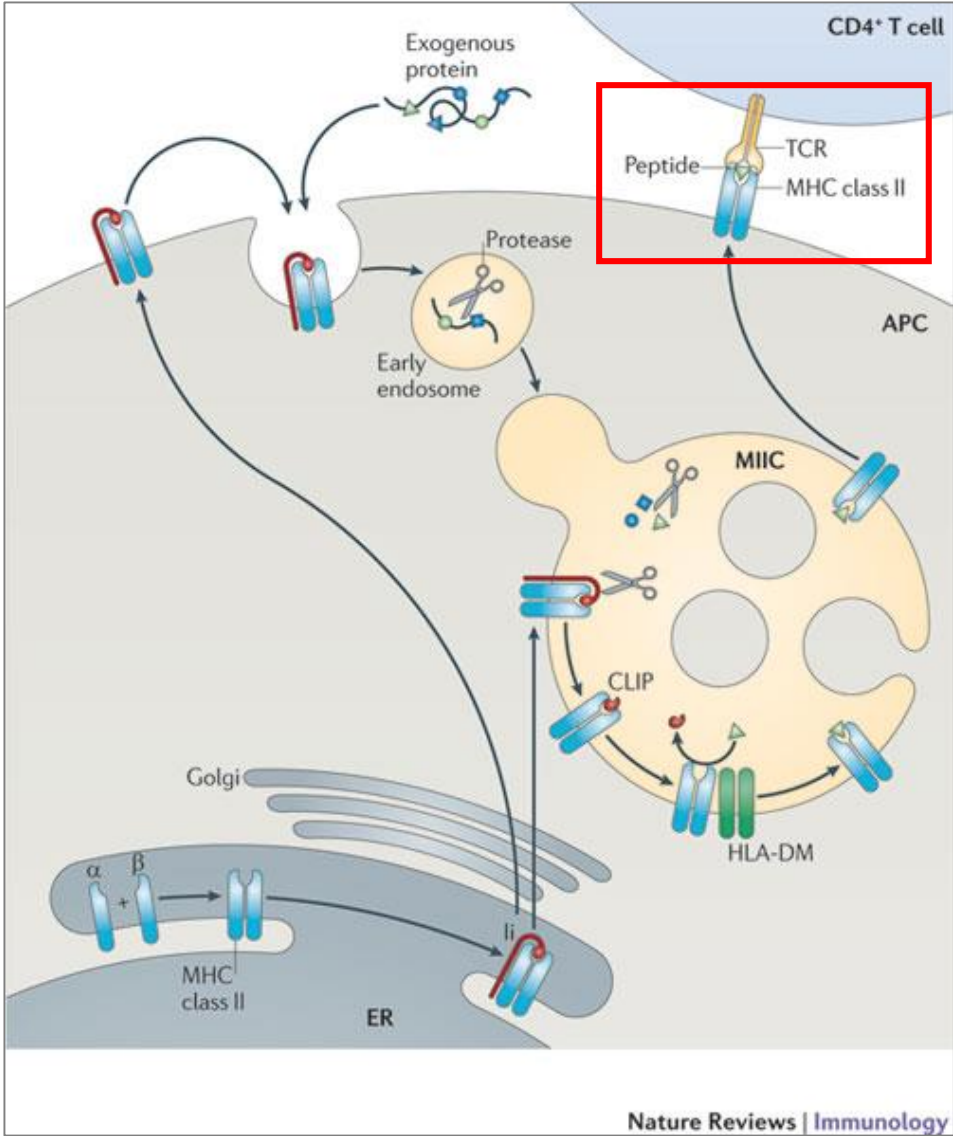
- Experimentally, many MHC binding peptides can be immunogenic (~50%)
- Cross validation gave AUC values  $\sim 0.65$ . Test on independent blind set gave AUC = 0.69
- Recommendation: Use as filter (cutoff 0) if high specificity is desired. Suggested cutoff is 0



# Class I Summary

- Processing predictions are better at identifying naturally processed ligands, but have not been shown to be superior in identifying epitopes compared to MHC binding predictions
- Specific processing and immunogenicity predictions are good additional filters if the only goal is to select high likelihood T cell epitopes
- NetMHCPan 4.0 EL scores, which are trained on both MHC binding and ligand elution data are a straightforward replacement of MHC binding predictions, and show some enhanced performance
  - ❓ Use these, and consider pairing with immunogenicity scores, when predicting epitope candidates

# CD4 T cell epitopes (MHC class II)



# Class II Processing + immunogenicity tools available in the IEDB

<http://tools.iedb.org/main/tcell/>

## IEDB Analysis Resource

Overview **T Cell Tools** B Cell Tools Analysis Tools Tools-API Usage Download Datasets Contribute Tools References

### T Cell Epitopes - Processing Prediction

These tools predict epitope candidates based upon the processing of peptides in the cell.

#### [Proteasomal cleavage/TAP transport/MHC class I combined predictor](#)

This tool combines predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential of being a T cell epitope.

#### [Neural network based prediction of proteasomal cleavage sites \(NetChop\) and T cell epitopes \(NetCTL and NetCTLpan\)](#)

NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL and NetCTLpan are predictors of T cell epitopes along a protein sequence. It also employs a neural network architecture.

#### [MHC-NP: Prediction of peptides naturally processed by the MHC](#)

MHC-NP employs data obtained from MHC elution experiments in order to assess the probability that a given peptide is naturally processed and binds to a given MHC molecule. This tool was the winner of the 2nd Machine Learning Competition in Immunology.

#### [MHCII-NP:](#)

This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences.

### T Cell Epitopes - Immunogenicity Prediction

This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.

#### [T cell class I pMHC immunogenicity predictor](#)

This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a class I peptide MHC (pMHC) complex.

#### [Deimmunization:](#)

The deimmunization tool is attempt to identify immunodominant regions in a given therapeutically important protein, and suggest amino-acid substitutions that create non-immunogenic versions of the proteins. So we have opted a two steps process; 1) In the first step, the deimmunization tool will list all the immunogenic regions or peptides based on selected threshold. These peptides will be generated from the protein with 15mer window size and 10mer overlap. 2) In the second step, the user can select one or more peptides listed in the results and final result window will display the non-immunogenic substitution of each selected peptides. The default threshold is 8.5 (which is difference in the median of percentile rank from 26 reference alleles set for MHC class II). In the final result window, the tools will also take care of the fact that non-immunogenic substitution in the immunogenic peptides, should not create new immunogenic site in the neighboring peptides. Therefore, the result window will also display the effect of substitution on the neighboring peptides.

#### [CD4 T cell immunogenicity prediction:](#)

The server is developed to predict the allele independent CD4 T cell immunogenicity at population level. User can predict the T cell immunogenicity using 7-allele method (Paul et. al. 2015), immunogenicity method and combined method (IEDB recommended). The combined method predicts the final score that combines the predictions from 7-allele method and immunogenicity method.





# MHCII-NP

- Predicting the naturally processed peptides for MHC class II
- Based on
  - Cleavage motif analysis at C and N terminal of peptides
  - Ligand elution data derived from IEDB
- Ligand predictions is improved markedly when combining the binding and cleavage motifs
- T cell epitope prediction is not significantly improved

# MHCII-NP - example

<http://tools.iedb.org/mhciinp/>

## IEDB Analysis Resource - Labs



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

### MHC II NP - Prediction of naturally processed MHC II ligands

#### Sequences

Enter sequences in FASTA or plain format

```
>sp|P15848|ARSB_HUMAN Arylsulfatase B OS=Homo sapiens OX=9606
GN=ARSB PE=1 SV=1
MGPRGAASLPRGPGPRRLLLPVVLPLLLLLLAPPGSGAGASRPPHLVFL LADDLGWNDV
GFHGSRI RTPHLDALAAGGVLLDNYTQPLCTPSRSQLLTGRYQIRTGLQH QI IWPCQPS
CVPLDEKLLPQLLKEAGYTHMVGKWHLGMYRKECLPTRRGFDTYFGYLLGSE DYYSHER
CTLIDALNVTRCALDFRDGEEVATGYKNMYSTNIFTKRAIALITNHPPEKPLFLYLALQS
VHEPLQVP E EYLKPYDFIQDKNRHHYAGMVSLMDEAVGNVTAALKSSGLWNNTV FIFSTD
NNGGQTLAGGNNWPLRGRKWSLWEGGVRGVGFVASPLLKQKGVKNRELIHISDWLPTLVKL
ARGHTNGTKPLDGFVWVTISEGSPSPRIELLNIDPNFVDSSPCPRNSMAPAKDDSSLP
EYSAFNTSVHAAIRHGNWKL LTGYPGCGYWFPPPSQYNVSEIPSSDPPTKTLWLF DIDRD
PEERHDL SREYPHIVTKLLSRLQFYHKHSVPVYFPAQDPRCDPKATGVWGPWM
```

Or upload sequences as a text file  No file selected.



# MHCII-NP -example

<http://tools.iedb.org/mhciinp/>

## IEDB Analysis Resource - Labs



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

### MHC II NP results

Top 5 peptides per protein:

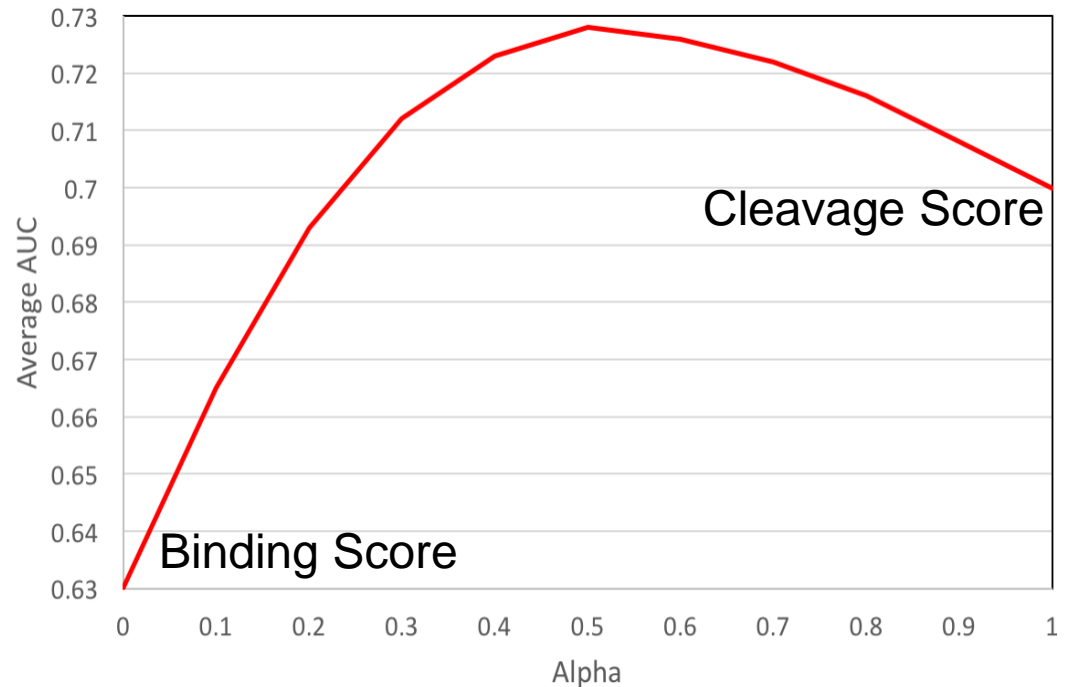
	Seq name	Peptide start	Peptide end	Peptide length	Peptide	N motif	C motif	Cleavage probability score	Cleavage probability percentile rank
1	SP P15848 ARSB_HUMAN ARYLSULFATASE_B OS=HOMO S...	510	524	15	VPVYFPAQDPRCDPK	SVP	PKA	1.75814	0.00
2	SP P15848 ARSB_HUMAN ARYLSULFATASE_B OS=HOMO S...	2	16	15	GPRGAASLPRGPGPR	MGP	PRR	1.73735	0.02
3	SP P15848 ARSB_HUMAN ARYLSULFATASE_B OS=HOMO S...	247	261	15	VPEEYLKPYDFIQDK	QVP	DKN	1.48840	0.04
4	SP P15848 ARSB_HUMAN ARYLSULFATASE_B OS=HOMO S...	384	398	15	SPSPRIELLHNIDPN	GSP	PNF	1.40420	0.05
5	SP P15848 ARSB_HUMAN ARYLSULFATASE_B OS=HOMO S...	12	26	15	GPGPRLLLLPWVLP	RGP	PLL	1.33714	0.07

[Complete results](#)

# MHCII-NP scores

- **Cleavage Score:** Derived from the cleavage motif analysis in ligand elution data
- **Binding Score:** Derived from HLA binding affinity using 7-allele method (Paul et. al. 2015).

$$\text{Combined score} = \alpha \times \text{cleavage probability score} + (1 - \alpha) \times \text{binding score}$$



Barra et al. *Genome Medicine* (2018) 10:84  
<https://doi.org/10.1186/s13073-018-0594-6>

Genome Medicine

RESEARCH

Open Access



Footprints of antigen processing boost MHC class II natural ligand predictions

Carolina Barra<sup>1\*</sup>, Bruno Alvarez<sup>1\*</sup>, Sinu Paul<sup>2</sup>, Alessandro Sette<sup>2</sup>, Bjoern Peters<sup>2</sup>, Massimo Andreatta<sup>1</sup>, Soren Buus<sup>3</sup> and Morten Nielsen<sup>1,4\*</sup>

*Front Immunol.* 2018 Aug 6;9:1795. doi: 10.3389/fimmu.2018.01795. eCollection 2018.

**Determination of a Predictive Cleavage Motif for Eluted Major Histocompatibility Complex Class II Ligands.**

Paul S<sup>1</sup>, Karosiene E<sup>1</sup>, Dhanda SK<sup>1</sup>, Jurtz V<sup>2</sup>, Edwards L<sup>1</sup>, Nielsen M<sup>2,3</sup>, Sette A<sup>1,4</sup>, Peters B<sup>1,4</sup>.

PMID: 30127785 PMCID: PMC6087742 DOI: 10.3389/fimmu.2018.01795

# Class II Processing + immunogenicity tools available in the IEDB

<http://tools.iedb.org/main/tcell/>

**IEDB Analysis Resource**

Overview | **T Cell Tools** | B Cell Tools | Analysis Tools | Tools-API | Usage | Download | Datasets | Contribute Tools | References

### T Cell Epitopes - Processing Prediction

These tools predict epitope candidates based upon the processing of peptides in the cell.

[Proteasomal cleavage/TAP transport/MHC class I combined predictor](#)


This tool combines predictors of proteasomal processing, TAP transport, and MHC binding to produce an overall score for each peptide's intrinsic potential of being a T cell epitope.

[Neural network based prediction of proteasomal cleavage sites \(NetChop\) and T cell epitopes \(NetCTL and NetCTLpan\)](#)

NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL and NetCTLpan are predictors of T cell epitopes along a protein sequence. It also employs a neural network architecture.

[MHC-NP: Prediction of peptides naturally processed by the MHC](#)

MHC-NP employs data obtained from MHC elution experiments in order to assess the probability that a given peptide is naturally processed and binds to a given MHC molecule. This tool was the winner of the 2nd Machine Learning Competition in Immunology.

 **MHCII-NP:**


This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences.

### T Cell Epitopes - Immunogenicity Prediction


This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.

[T cell class I pMHC immunogenicity predictor](#)

This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a class I peptide MHC (pMHC) complex.

 **Deimmunization:**

The deimmunization tool is attempt to identify immunodominant regions in a given therapeutically important protein, and suggest amino-acid substitutions that create non-immunogenic versions of the proteins. So we have opted a two steps process; 1) In the first step, the deimmunization tool will list all the immunogenic regions or peptides based on selected threshold. These peptides will be generated from the protein with 15mer window size and 10mer overlap. 2) In the second step, the user can select one or more peptides listed in the results and final result window will display the non-immunogenic substitution of each selected peptides. The default threshold is 8.5 (which is difference in the median of percentile rank from 26 reference alleles set for MHC class II). In the final result window, the tools will also take care of the fact that non-immunogenic substitution in the immunogenic peptides, should not create new immunogenic site in the neighboring peptides. Therefore, the result window will also display the effect of substitution on the neighboring peptides.

 **CD4 T cell immunogenicity prediction:**

The server is developed to predict the allele independent CD4 T cell immunogenicity at population level. User can predict the T cell immunogenicity using 7-allele method (Paul et. al. 2015), immunogenicity method and combined method (IEDB recommended). The combined method predicts the final score that combines the predictions from 7-allele method and immunogenicity method.

# MHC-II restricted immunogenicity prediction


- Extracted datasets of proteins from the IEDB for which overlapping peptides were tested for immunogenicity
- Utilized these datasets to train a Neural Network to learn 'motifs' associated with immunogenicity independent of specific MHC alleles expressed
- Resulting score can be combined with '7 allele method' quantifying MHC binding across alleles to predict overall immunogenicity

# Class II immunogenicity prediction

- Based on Neural network model trained on
  - In house dataset for different antigens tested on different population cohorts
  - Tetramer dataset- derived from IEDB
- Validated on 57 independent studies from different groups across the world
- Implemented three approaches
  - 7-allele method (*Paul et. al. 2015*)
  - Immunogenicity predictions
  - Hybrid approach

# Class II immunogenicity prediction -example

<http://tools.iedb.org/CD4episcore/>

**IEDB Analysis Resource - Labs** 

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

## CD4 T cell immunogenicity prediction


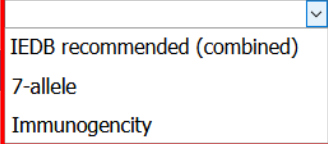
**Specify Sequence(s)**

Enter epitope sequence(s) in PLAIN or FASTA format


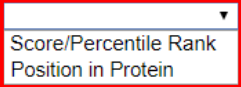
```
>sp|P01588|EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1
MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAENITGCAEHC
SLNENITVPDTKVNFYAWKRMEVGGQAVEVWQGLLALLSEAVLRGQALLVNSSQPWEPLQL
HVDKAVSGLRSLTTLRLALGAQKEAISPPDAASAAPLRTITADTFRKLFVYSNFLRGKL
KLYTGACRTGDR
```


Or upload epitope sequence(s) from a file  No file selected.

**Choose a prediction method**

Prediction method:   

**Specify Output**

Sort Peptides by:   

Select maximum percentile rank threshold:  

Enter the Job Name (Optional)

Email address (optional)



# Class II immunogenicity prediction - example

<http://tools.iedb.org/CD4episcore/>

IEDB Analysis Resource - Labs



Home Help Example Reference Contact

## CD4 Immunogenicity prediction results

Number of proteins: 1

Number of 15mer (overlapping 10mer): 37

Threshold : 50.0%

Method : combined

[Download result](#)

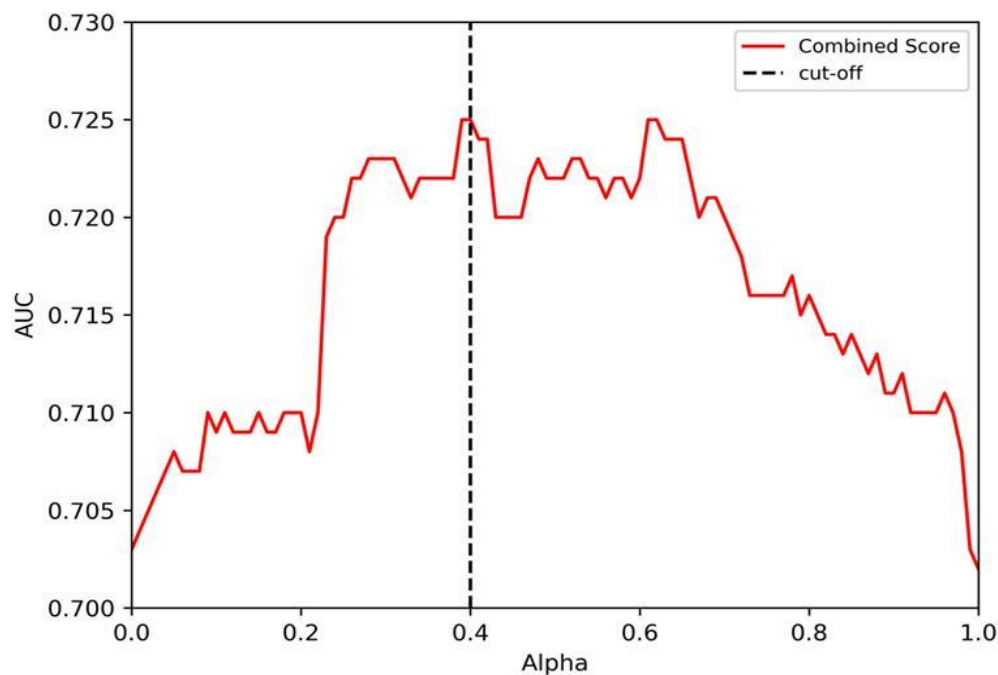
Citations

Protein Number	Protein Description	Peptide	Start	End	Combined Score	Immunogenicity Score	Peptide core	Median Percentile Rank (7-allele)	HLA-DRB1:03:01	HLA-DRB1:07:01	HLA-DRB1:15:01	HLA-DRB3:01:01	HLA-DRB3:02:02	HLA-DRB4:01:01	HLA-DRB5:01:01
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	WLLSLLSLPLGLPV	11	25	42.16452	95.0613	LLSLLSLPL	6.9	25.0	3.2	3.6	73.0	33.0	6.9	6.5
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	TKVNFYAWKRMEVGQ	71	85	47.39488	67.4872	TKVNFYAWK	34.0	52.0	22.0	15.0	71.0	30.0	65.0	34.0
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	EPLQLHVDKAVSGLR	116	130	32.55636	43.8909	LHVDKAVSG	25.0	5.4	59.0	40.0	22.0	7.0	38.0	25.0
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	VSGLRSLTLLRALG	126	140	44.95964	86.8991	LTLLRALG	17.0	12.0	17.0	9.3	70.0	20.0	20.0	1.3
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	SLTLLRALGAQKEA	131	145	42.78744	69.4686	LLRALGAQK	25.0	47.0	46.0	21.0	89.0	25.0	14.0	1.6
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	PLRTITADTRFKLFR	156	170	46.18064	85.4516	LRTITADTF	20.0	6.1	47.0	40.0	8.0	20.0	20.0	20.0
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	TADTRFKLFRVYSNF	161	175	46.66984	44.6746	FRKLFVYS	48.0	63.0	58.0	23.0	48.0	33.0	53.0	24.0
1	sp P01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	RKLFVYSNFLRGKL	166	180	13.8966	22.4415	FRVYSNFLR	8.2	53.0	8.2	0.12	27.0	4.8	26.0	1.5

# Class II immunogenicity prediction scores

- Immunogenicity Score: Derived from the neural network model trained on Immunogenicity data
- HLA Score: Derived from HLA binding affinity using 7-allele method (Paul et. al. 2015).

$$\text{Combined score} = \alpha \times \text{Imm score} + (1 - \alpha) \times \text{HLA score.}$$



[Front Immunol.](#) 2018 Jun 14;9:1369. doi: 10.3389/fimmu.2018.01369. eCollection 2018.

## Predicting HLA CD4 Immunogenicity in Human Populations.

[Dhanda SK](#)<sup>1</sup>, [Karosiene E](#)<sup>1</sup>, [Edwards L](#)<sup>1</sup>, [Grifoni A](#)<sup>1</sup>, [Paul S](#)<sup>1</sup>, [Andreatta M](#)<sup>2</sup>, [Weiskopf D](#)<sup>1</sup>, [Sidney J](#)<sup>1</sup>, [Nielsen M](#)<sup>2,3</sup>, [Peters B](#)<sup>1,4</sup>, [Sette A](#)<sup>1,4</sup>.

PMID: 29963059 PMCID: [PMC6010533](#) DOI: [10.3389/fimmu.2018.01369](#)

# Class II Summary

- Similar to MHC class I, enhancement of epitope prediction efficacy is minor compared to using MHC binding predictions alone
- Prediction of naturally eluted ligands is greatly improved with processing predictions
- As of now, recommendation is to stick to allele specific MHC binding predictions (NetMHCPanII), or the 7-allele method for broad populations

→ Both class I and II epitope predictions are constantly being re-evaluated, and these recommendations are subject to change