

T Cell processing & immunogenicity predictions

tools.iedb.org

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

2021 IEDB User Workshop

CD8+ T cell epitopes in viral infection



CD8+ T cell epitopes in viral infection



TT 11	ORF 1	<u>MGQIVTMFEALPHIIDEVINIVI</u> IVLIVITGIKAVYN
T cell	ORF 2	MGLKGPD YKGVYQFKSVEFDMSHLNLTMPNACSANN
onitono	ORF 3	MHNFCNLTSAFNKKTFDHTLMSIVSSLHLSIDGNSNY
ephope	ORF 4	<u>MSAQSQCRTFRGRVLDMFRTAFGGKYMRSGWGWTGSD</u>
manning	ORF 5	MHCTYAGPFGMSRILLSQEKTKFFTRRLAGTFTWTLS
mapping	ORF 6	MKCFGNTAVAKCNVNHDAEFCDMLRLIDYNKAALSKF
	ORF 7	MLMRNHLLDLMGVPYCNYSKFWYLEHAKTGETSVPKC ³

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

4

Class I Processing + immunogenicity tools available in the IEDB <u>http://tools.iedb.org/main/tcell/</u>

IEDB Analysis Resource

		, and	y 515 T ((cource	, ,							
	Overview	T Cell Tools	B Cell Tools	Analysis Tools	Tools-API	Usage	Download	Datasets	Contribute Tools	References		
	T Cell E	Epitopes -	Processing	g Prediction								
	These to	ols predict ep	itope candida	ites based upon	the proces	ssing of p	eptides in tl	ne cell.				.
Ν	Proteaso	mal cleavage	e/TAP transpo	ort/MHC class I	combined	predictor						
1	This pept	tool combin ide's intrinsi	es predictors c potential of	f being a T cell	al processi epitope.	ng, IAP t	transport, a	ind MHC b	inding to produc	e an overall	score for each	
	Neural n	etwork based	I prediction of	f proteasomal c	leavage sit	es (NetC	hop) and T	cell epitop	es (NetCTL and	NetCTLpan)		
	Net(epito	Chop is a pre opes along a	dictor of prot protein sequ	teasomal proce Jence. It also e	essing base employs a i	ed upon a neural ne	a neural ne twork arch	twork. Net itecture.	tCTL and NetCTL	pan are pre	dictors of T cell	
	MHC-NP	: Prediction of	of peptides na	aturally process	ed by the N	<u>/HC</u>						
	MHC	-NP employs essed and b	s data obtain inds to a give	ed from MHC e en MHC molecu	elution exp ule. This to	eriments ol was th	in order to	assess th f the 2nd I	e probability that Machine Learnin	at a given pe a Competitio	ptide is naturally	
	<u> </u>											•
	MHC	<u> III-NP</u> :										
	inis	tool utilizes	MHC II ligan	d elution data	to predict	naturaliy	processed	MHC II IIg	gands by scannir	ig the given	peptide sequences	•
		nitonoc	Immunogo	nicity Produ	otion							
	This tool	predicts the r	elative ability	of a peptide/MH	-C complex	c to elicit	an immune	response				
	T cell cla	ss I pMHC in	nmunogenicit	y predictor								1
	This MHC	tool uses an (pMHC) cor	nino acid pro nplex.	perties as well	as their p	osition w	ithin the p	eptide to p	redict the immu	nogenicity o	f a class I peptide	
	beim	munization:										
	The	deimmuniza	tion tool is at	ttempt to ident	tify immun	odomina	nt regions	in a given	therapeutically i	mportant pr	otein, and suggest	
	first	step, the de	immunization	n tool will list a	all the imm	unogenio	regions of	peptides	based on select	ed threshold	These peptides	
	mor	e peptides lis	sted in the re	sults and final	result win	dow will	display the	non-imm	unogenic substit	ution of each	selected peptides.	
	The the f	default three final result w	shold is 8.5 (indow, the to	which is differe	ence in the ake care of	median	of percent that non-i	le rank fro nmunoder	om 26 reference nic substitution i	alleles set fo n the immun	or MHC class II). In ogenic peptides.	
	shou subs	uld not creat stitution on t	e new immur he neighborii	nogenic site in ng peptides.	the neight	oring pe	ptides. The	refore, the	e result window	will also disp	lay the effect of	
	<u> </u>	T cell immun	ogenicity pre	diction:								
	The imm The	server is dev unogenicity combined m	veloped to pr using 7-allele ethod predic	edict the allele e method (<u>Paul</u> ts the final sco	independ l et. al. 20 re that cor	ent CD4 <u>15</u>), imm nbines th	T cell immu nunogenicit	unogenicity y method ons from 7	y at population l and combined m -allele method a	evel. User ca nethod (IEDE nd immunoc	n predict the T cell recommended).	

Class I 'combined predictor'

Home Help Example Reference	e Download Contact
MHC-I Processing Pro	edictions
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)	Browse No file selected.
Choose sequence format	auto detect format
	Choose a Prediction Method
Prediction Method	IEDB recommended V Help on prediction method selections
	Specify what to make binding predictions for
MHC source species	human 🗸
	Allele Length
Select MHC allele(s)	Upload allele file ?
	Proteasomal cleavage prediction
Specify proteasome type	immuno
	TAP transport predictions
Maximum precursor extension	1
Alpha factor	0.2
	Specify Output
Output format	XHTML table 🔽
	Submit Reset

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

Proteasomal cleavage

Proteasomal cleavage prediction											
Specify proteasome type	immuno	~									
	TAP transport predictions	immuno									
Maximum precursor extension	1	constitutive									
Alpha factor	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

Proteasomal cleavage prediction									
Specify proteasome type	immuno								
	TAP transport predictions								
Maximum precursor extension	1								
Alpha factor	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

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 MGQIVTMFEALPHIIDEVINIVIIVLIVITGIKAVYNFATCGIFALISFLLLAGRSCGM YGLKGFDIVKGVYQFKSVEFDMSHLNIIMENACSANNSHHVISMGTSGLELFTFNDSII SHNFCNLTSAFNKKTFDHTLMSIVSSLHLSIRGNSNYKAVSCDFNNGITIQYNLTFSDA QSAQSQCRTFRGRVLDMFRTAFGKYMRSGMGWTGSDGKTTMCSQTSYQLIIQNRTWE NHCTYAGPFGMSRILLSQEKTKFFTRRLAGTFTWTLSDSSGVENPGGYCLTKWMILAAE LKCFGNTAVAKCNVNHDAEFCDMLRLIDYNKAALSKFKEDVESALHLFKTTVNSLISDQ LLMRNHLRDLMGVPYCNYSKFWYLEHAKTGETSVPKCWLVTNGSYLNETHFSDQIEQEA DNMITEMLRKDVIKRQGSTPLALMDLLMFSTSAYLVSIFLHLVKIPTHRHIKGGSCPKP HRLTNKGICSCGAFKVPGVKTVWKRR
Or select file containing sequence(s)	Browse 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 Length H-2-Kb 10 O
	Proteasomal cleavage prediction
Specify proteasome type	immuno
	TAP transport predictions
Maximum precursor extension	1
Alpha factor	0.2
	Specify Output
Output format	XHTML table
	Submit Reset

Class I 'combined predictor' - example

MHC-I Processing Prediction Results

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

Input Sequences

#	Name	Sequence
1	LCMV Armstrong, Protein GP	eq:mgqivtmfealphiideviniviivlivlivlivdikavynfatcgifalisfllagrs CGMYGLKGPDIYKGVYQFKSVEFDMSHLNLTMPNACSANNSHHYISMGTSGLELTF TNDSIISHNFCNLTSAFNKKTFDHTLMSIVSSLHLSIRGNSNYKAVSCDFNNGITI QYNLTFSDAQSAQSQCRTFRGRVLDMFRTAFGGKYMRSGWGWTGSDGKTTWCSQTS YQYLIIQNRTWENHCTYAGPFGMSRILLSQEKTKFFTRRLAGTFTWTLSDSSGVEN PGGYCLTKWMILAAELKCFGNTAVAKCNVNHDAEFCDMLRLIDYNKAALSKFKEDV ESALHLFKTTVNSLISDQLLMRNHLRDLMGVPYCNYSKFWYLEHAKTGETSVPKCW LVTNGSYLNETHFSDQIEQEADNMITEMLRKDYIKRQGSTPLALMDLLMFSTSAYL VSIFLHLVKIPTHRHIKGGSCPKPHRLTNKGICSCGAFKVPGVKTVWKRR
2	LCMV Armstrong, Protein NP	MSLSKEVKSFQWTQALRRELQSFTSDVKAAVIKDATNLLNGLDFSEVSNVQRIMRK EKRDDKDLQRLRSLNQTVHSLVDLKSTSKKNVLKVGRLSAEELMSLAADLEKLKAK IMRSERPQASGVYMGNLTTQQLDQRSQILQIVGMRKPQQGASGVVRVWDVKDSSLL NNQFGTMPSLTMACMAKQSQTPLNDVVQALTDLGLLYTVKYPNLNDLERLKDKHPV LGVITEQQSSINISGYNFSLGAAVKAGAALLDGGNMLESILIKPSNSEDLLKAVLG AKRKLNMFVSDQVGDRNPYENILYKVCLSGEGWPYIACRTSIVGRAWENTTIDLTS EKPAVNSPRPAPGAAGPPQVGLSYSQTMLLKDLMGGIDPNAPTWIDIEGRFNDPVE IAIFQPQNGQFIHFYREPVDQKQFKQDSKYSHGMDLADLFNAQPGLTSSVIGALPQ GMVLSCQGSDDIRKLLDSQNRKDIKLIDVEMTREASREYEDKVWDKYGWLCKMHTG IVRDKKKKEITPHCALMDCIIFESASKARLPDLKTVHNILPHDLIFRGPNVVTL

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	LLYTVKYPNI	1.79	0.50	-2.01	2.28	0.27	103.5
H-2-Kb	1	116	125	10	SIISHNFCNI	1.51	0.46	-1.90	1.97	0.08	78.6
H-2-Kb	2	7	16	10	VKSFQWTQAI	1.42	0.39	-2.01	1.81	-0.21	102.9
H-2-Kb	2	235	244	10	INISGYNFSI	1.53	0.41	-2.18	1.95	-0.23	149.7
H-2-Kb	1	35	44	10	VYNFATCGIE	1.42	1.27	-3.12	2.68	-0.44	1327
H-2-Kb	1	75	84	10	KSVEFDMSHI	1.50	0.47	-2.45	1.96	-0.49	283.6
H-2-Kb	1	369	378	10	YCNYSKFWYI	1.45	0.31	-2.32	1.76	-0.56	206.8
H-2-Kb	1	449	458	10	VSTELHLVKI	1.33	0.33	-2.25	1.66	-0.59	178.2

Class I 'combined predictor' - example

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

- Higher scores = higher efficiency for MHC-I presentation
- MHC binding score = -log10(IC50) (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
- <u>Recommendation</u>: 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 <u>http://tools.iedb.org/main/tcell/</u>

IEDB Analysis Resource

	Ana	y 313 T V	cource	-							
Overview	T Cell Tools	B Cell Tools	Analysis Tools	Tools-API	Usage	Download	Datasets	Contribute Tools	References	1	
T Cell E	pitopes -	Processin	g Prediction								
These too	ols predict ep	bitope candida	ates based upon	the proces	sing of p	peptides in ti	ne cell.				
This pepti	tool combir ide's intrins	es predictors ic potential o	s of proteasoma f being a T cell	al processii epitope.	ng, TAP	transport, a	ind MHC b	inding to produ	ce an overall	score for each	
Neural ne	etwork base	d prediction o	f proteasomal c	leavage sit	es (NetC	hop) and T	cell epitop	es (NetCTL and	NetCTLpan)		
NetC epito	hop is a pre pes along a	edictor of pro a protein sequ	teasomal proce uence. It also e	essing base employs a r	ed upon neural ne	a neural ne etwork arch	twork. Net itecture.	CTL and NetCT	Lpan are pre	dictors of T cell	
MHC-NP:	Prediction	of peptides na	aturally process	ed by the N	<u>/HC</u>						
MHC	-NP employ	s data obtair	ed from MHC e	elution exp	eriments	s in order to	assess th	e probability th	at a given pe	eptide is naturally	
proce		inus to a giv	en Mile molect	ne. mis to	or was ti		i the <u>zhu</u>		ig competition	on in minimunology.	
A MHCI	II-NP:										
This	tool utilizes	MHC II ligar	nd elution data	to predict	naturally	y processed	MHC II lig	jands by scanni	ng the given	peptide sequence	s.
T Cell E	pitopes -	Immunoge	enicity Predic	ction							
his tool r	predicts the	relative ability	of a peptide/MI	HC complex	to elicit	an immune	response.				-
cell clas	<u>ss I pMHC II</u>	mmunogenici	ty predictor					10 1		c 1 - c 1	
I his t MHC	tool uses ai (pMHC) coi	mino acid pro mplex.	operties as well	as their po	osition w	within the p	eptide to p	redict the immi	inogenicity o	of a class I peptide	
	(┛╎
Deimi	munization:										
The c amin first s will b	deimmuniza o-acid subs step, the de oe generate	ation tool is a stitutions that eimmunizatio d from the pi	ttempt to ident t create non-im n tool will list a rotein with 15m	tify immun imunogenio all the imm ner window	odomina c versior unogeni / size an	ant regions ns of the pr c regions of d 10mer ov	in a given oteins. So r peptides rerlap. 2) I	therapeutically we have opted based on select n the second st	important pr a two steps ed threshold ep, the user	rotein, and sugges process; 1) In the I. These peptides can select one or	t
more The c	e peptides li	sted in the re	esults and final	result win	dow will	display the	non-immi	unogenic substi	tution of eac	h selected peptide	S.
the fi shou subst	inal result v ld not creat titution on t	vindow, the t e new immu the neighbori	ools will also ta nogenic site in ng peptides.	ake care of the neighb	the fact oring pe	that non-ineptides. The	mmunoger refore, the	ic substitution result window	will also disp	play the effect of	
<u>CD4 -</u>	T cell immu	nogenicity pre	ediction:								
The s immu The o	server is de unogenicity combined m	veloped to pr using 7-allel nethod predic	redict the allele e method (<u>Pau</u> ts the final sco	independe l et. al. 20: re that cor	ent CD4 <u>15</u>), imn nbines t	T cell immu nunogenicit he predictio	unogenicity y method ons from 7	at population and combined r -allele method a	level. User ca nethod (IEDI and immuno	an predict the T ce B recommended). genicity method.	ell

Neural Network based predictors

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

IEDB Analysis Resource									
Home Help Example Reference	Download Contact								
Proteasomal Cleavag NetChop/NetCTL/NetCTLpan	e Prediction								
	Specify Sequence(s)								
Prediction Method	NetChop	~							
Enter protein sequence(s) in FASTA format		NetChop netCTL netCTLpan	н						
Or select file containing sequence(s)	Browse No file selected.								
	Method Specific Options								
Method	C term 3.0 V								
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 <u>http://tools.iedb.org/main/tcell/</u>

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
1	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 <u>2nd Machine Learning Competition in Immunology</u> .
	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.
	This tool uses amino acid properties as well as their position within the pentide to predict the immunogenicity of a class I pentide
	MHC (pMHC) complex.
	Leimmunization:
	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 spetides, 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 <u>http://tools.iedb.org/mhcnp/</u>



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 V¹, Paul S², Andreatta M³, Marcatili P¹, Peters B², Nielsen M^{4,3}.

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

Class I Processing + immunogenicity tools available in the IEDB <u>http://tools.iedb.org/main/tcell/</u>

IEDB Analysis Resource

ILDD 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.	
<u>Neural network based prediction of proteasomal cleavage sites (NetChop) and T cell epitopes (NetCTL and NetCTLpan)</u> 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 <u>2nd Machine Learning Competition in Immunology</u> .	
This tool utilizes MHC II ligand elution data to predict naturally processed MHC II ligands by scanning the given peptide sequences. TCell Epitopes - Immunogenicity Prediction This tool predicts the relative ability of a peptide/MHC complex to elicit an immune response.	
<u>T cell class I pMHC immunogenicity predictor</u> 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.	
<u>CD4 T cell immunogenicity prediction</u> : 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 (<u>Paul et. al. 2015</u>), 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



Class I immunogenicity prediction -example

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

IEDB Analy	sis Resource	
Home Help Example	Reference Download Contact	
Class I Immun	ogenicity	
	Specify sequence(s) *	
Enter peptide sequence(s) (Browse for sequences in NCBI)	FIAGLIAIV LITGRLQSL RLNEVAKNL KAVYNFATC FQPQNGQFI	
Or select file containing sequence(s)	Browse No file selected.	
	Choose which positions to mask	
Specify which positions to mask	Default (1st, 2nd, and C-terminus amino acids) Custom User Defined (Comma separated numbers) Peptide lengths must be equal when using custom masking.	Mask positions that are MHC anchors
	Submit Reset	
*The tool was only validate	ed for 9-mer peptides. However, predictions can be made for peptides of any length.	

Class I immunogenicity prediction -example

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

IEDB Analysis Resource											
Home He	elp Exa	mple	Reference	Download	Contact						
Class I Immunogenicity Masking: default Masked variables: [1, 2, 'cterm']											
CIASS Masking: defa Masked variat	ault bles: [1, 2, '	'cterm']	genien	.y							
CIASS Masking: defa Masked variat Peptide \$	ault bles: [1, 2, ' Length \$	'cterm'] Score •	genici	.y							
GIASS Masking: defa Masked varial Peptide + FIAGLIAIV	ault bles: [1, 2, ' Length \$ 9	'cterm'] Score • 0.27206	genici	.y							
Glass Masking: defa Masked varial Peptide FIAGLIAIV KAVYNFATC	ault bles: [1, 2, ¹ Length \$ 9 9	"cterm"] Score ▼ 0.27206 0.16928	genici	.y							
Class Masking: def: Masked variat Peptide • FIAGLIAIV KAVYNFATC RLNEVAKNL	ault bles: [1, 2, ' Length 9 9 9 9	'cterm'] Score ▼ 0.27206 0.16928 -0.0101		.y							
Masking: defa Masked varial Peptide • FIAGLIAIV KAVYNFATC RLNEVAKNL LITGRLQSL	ault bles: [1, 2, 5 Length \$ 9 9 9 9 9 9	Score 0.27206 0.16928 -0.0101 -0.10776		.y							

- 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 ~ 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 <u>http://tools.iedb.org/main/tcell/</u>

IEDB Analysis Resource

Overview T	T Cell Tools	B Cell Tools	Analysis Tools	Tools-API	Usage	Download	Datasets	Contribute Tools	Reference	95	
T Cell Ep	itopes - I	Processing	Prediction								
These tools	s predict epi	tope candida	tes based upon	the process	ing of pe	ptides in the	e cell.				
Proteasoma	al cleavage	/TAP transpo	ort/MHC class I	combined pr	redictor						
This to peptide	ol combine e's intrinsio	es predictors potential of	of proteasoma being a T cell	l processing epitope.	g, TAP tr	ansport, ar	nd MHC b	inding to produ	ice an over	all score for	each
leural netw	work based	prediction of	proteasomal cl	eavage sites	s (NetCh	op) and T c	ell epitop	es (NetCTL and	I NetCTLpa	<u>n)</u>	
NetCho epitope	op is a pre es along a	dictor of prot protein sequ	teasomal proce ience. It also e	ssing based mploys a ne	l upon a eural net	neural net work archit	work. Net tecture.	CTL and NetC	"Lpan are p	redictors of	T cell
MHC-NP: P	Prediction o	f peptides na	turally process	ed by the MH	<u>HC</u>						
MHC-N	IP employs	data obtain	ed from MHC e	lution exper	iments	in order to	assess th	e probability t	nat a given	peptide is r	aturally
This to	ol utilizes	MHC II ligan	d elution data	to predict na	aturally	processed I	MHC II lig	ands by scann	ing the giv	en peptide s	equences.
This to F Cell Ep This tool pre	ool utilizes bitopes - I edicts the re	MHC II ligan mmunoge elative ability	d elution data nicity Predic of a peptide/MF	to predict na :tion IC complex t	aturally to elicit a	processed I n immune re	MHC II lig esponse.	ands by scann	ing the giv	en peptide s	equences.
This to T Cell Ep This tool pre	ool utilizes bitopes - l edicts the re I pMHC im	MHC II ligan mmunoge elative ability munogenicity	d elution data nicity Predic of a peptide/MH y predictor	to predict na :tion IC complex t	aturally to elicit a	processed I n immune re	MHC II lig esponse.	ands by scann	ing the giv	en peptide s	equences.
This to Cell Ep This tool pre- <u>Cell class</u> This to MHC (p	bol utilizes bitopes - I edicts the re I <u>pMHC im</u> pol uses am pMHC) con	MHC II ligan mmunoge elative ability munogenicith nino acid pro nplex.	d elution data nicity Predic of a peptide/MH <u>y predictor</u> perties as well	to predict na t tion IC complex t as their pos	aturally to elicit a sition wit	processed I n immune re thin the pep	MHC II lig esponse. ptide to p	ands by scann redict the imm	ing the giv	en peptide s y of a class i	equences.
This to T Cell Ep This tool pre T cell class This to MHC (p Deimm	itopes - I edicts the re pol uses am pMHC) con unization:	MHC II ligan mmunoge elative ability munogenicity nino acid pro nplex.	d elution data nicity Predic of a peptide/MF <u>y predictor</u> perties as well	to predict na c tion IC complex t as their pos	aturally to elicit a sition wit	processed I n immune re thin the pep	MHC II li <u>c</u> esponse. ptide to p	ands by scann redict the imm	ing the giv	en peptide s y of a class i	equences.
This to T Cell Ep This tool pre Cell class This too MHC (p Deimme The de amino- first stu will be more p The de the fina should substit	bol utilizes itopes - l edicts the re- i pMHC im- pol uses arr pMHC) con- unization: emmunizat -acid subst ep, the dei generated peptides liss- efault thres al result w not creater unized on the cution on the cutio	MHC II ligan mmunoge elative ability munogenicit nino acid pro aplex. tion tool is at itutions that mmunization from the pr ted in the re hold is 8.5 (i indow, the to a new immur an eneighbori	d elution data nicity Predic of a peptide/MH- <u>y predictor</u> perties as well ttempt to ident create non-im n tool will list a otein with 15m sults and final which is differe pols will also ta loggenic site in n peptides.	to predict na tion IC complex t as their pos ify immunogenic Il the immun er window s result windo nce in the n ke care of th the neighbo	aturally to elicit a sition with sition with size and ow will of nedian of he fact t ring pep	n immune re thin the per thin the per of the proi regions or 10mer ove lisplay the chat non-im tides. Ther	MHC II lig esponse. ptide to p teins. So peptides rrlap. 2) I non-immu e rank fro imunoger efore, the	redict the imm therapeutically we have opted based on selec n the second s unogenic subst m 26 reference ic substitution e result window	ing the giv important a two step ted threshc tep, the us itution of e a alleles se in the imm v will also d	of a class i protein, and process; i old. These ach selected thor MHC cl nunogenic pr isplay the e	equences. I peptide d suggest l) In the eptides t one or d peptides. ass II). In eptides, ffect of
This to T Cell Ep This tool pre T cell class This to MHC (p Deimme The de amino- first st will be more p The de the fina should substit	bol utilizes itopes - I edicts the re- i I pMHC im- ool uses am- pMHC) con unization: eimmunizat -acid subst ep, the dei generated generated sefault thres al result we not create ution on the cell immun	MHC II ligan mmunoge elative ability munogenicith nino acid pro pplex. tion tool is at itutions that mmunization from the pr ted in the re hold is 8.5 (' indow, the to e new immur te new immur te new immur	d elution data nicity Predic of a peptide/MH y predictor perties as well ttempt to ident create non-im n tool will list a otein with 15m sults and final which is differed bols will also ta logenic site in the peptides.	to predict na tion IC complex t as their pos ify immunod munogenic II the immun er window s result windo result windo result windo result windo the neighbo	aturally to elicit a sition with versions nogenic size and ow will o nedian o he fact t ring pep	n immune for thin the per thin the per of the prot regions or 10mer ove lisplay the of percentile that non-im otides. Ther	MHC II lig esponse. ptide to p teins. So peptides relap. 2) I non-immu e rank fro imunoger efore, the	redict the imm therapeutically we have opted based on selec n the second s unogenic subst m 26 reference ic substitution e result window	ing the giv unogenicity a two step ted thresht itution of e alleles se in the imm v will also d	en peptide s y of a class i protein, and s process; i old. These p er can select ach selected t for MHC cl nunogenic p isplay the e	equences. I peptide d suggest 1) In the eptides t one or d peptides. ass II). In eptides, ffect of

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/



MHCII-NP -example

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

IEDB Analysis Resource - Labs

Home Help

Example Reference D

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	SPIP15848IARSB_HUMAN ARYLSULFATASE B OS=HOMO S	510	524	15	VPVYFPAQDPRCDPK	SVP	PKA	1.75814	0.00		
2	SPIP15848IARSB_HUMAN ARYLSULFATASE B OS=HOMO S	2	16	15	GPRGAASLPRGPGPR	MGP	PRR	1.73735	0.02		
3	SPIP15848IARSB_HUMAN ARYLSULFATASE B OS=HOMO S	247	261	15	VPEEYLKPYDFIQDK	QVP	DKN	1.48840	0.04		
4	SPIP15848IARSB_HUMAN ARYLSULFATASE B OS=HOMO S	384	398	15	SPSPRIELLHNIDPN	GSP	PNF	1.40420	0.05		
5	SPIP15848 ARSB_HUMAN ARYLSULFATASE B OS=HOMO S	12	26	15	GPGPRRLLLPVVLPL	RGP	PLL	1.33714	0.07		
Cor	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).

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^{1*}©, Bruno Alvarez¹⁺, Sinu Paul², Alessandro Sette², Bjoern Peters², Massimo Andreatta¹, Søren Buus³ and Morten Nielsen^{1,4*}

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



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 S¹, Karosiene E¹, Dhanda SK¹, Jurtz V², Edwards L¹, Nielsen M^{2,3}, Sette A^{1,4}, Peters B^{1,4}.

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

Class II Processing + immunogenicity tools available in the IEDB <u>http://tools.iedb.org/main/tcell/</u>

IEDB Analysis Resource

	<i>i</i> nan	y 515 T (5000100								
					,					1	
Overview	T Cell Tools	B Cell Tools	Analysis Tools	Tools-API	Usage	Download	Datasets	Contribute Tools	References		
	nitones -	Processing									
These to	ols predict ep	pitope candida	tes based upon	the proces	ssing of p	eptides in t	ne cell.				
Proteaso	mal cleavag	e/TAP transpo	ort/MHC class I	combined	predictor						
This pept	tool combin ide's intrins	ies predictors ic potential of	of proteasoma being a T cell	al processi epitope.	ng, TAP	transport, a	and MHC b	inding to produc	e an overall	score for each	
Neural ne	etwork based	d prediction of	proteasomal cl	eavage sit	es (NetC	hop) and T	cell epitop	es (NetCTL and I	NetCTLpan)		
NetC epito	Chop is a pre opes along a	edictor of prot protein sequ	easomal proce ence. It also e	essing base mploys a r	ed upon neural ne	a neural ne etwork arch	twork. Net itecture.	tCTL and NetCTL	pan are pre	dictors of T cell	
/HC-NP	: Prediction	of peptides na	turally processe	ed by the N	<u>/HC</u>						
MHC	C-NP employ ressed and b	s data obtain inds to a give	ed from MHC e en MHC molecu	lution exp le. This to	eriments ol was th	s in order to he winner o	o assess th of the 2nd	ie probability tha Machine Learnin	at a given pe a Competitio	eptide is naturally	
This Cell E	tool utilizes	MHC II ligan	d elution data t nicity Predic	to predict	naturally	/ processed	MHC II lig	gands by scannir	ng the given	peptide sequences	J
cell cla	ss I pMHC ir	nmunogenicit	v predictor	ro compios	to oncit		rooponoo.				
This MHC	tool uses ar C (pMHC) cor	nino acid pro mplex.	perties as well	as their p	osition w	ithin the p	eptide to p	redict the immu	nogenicity o	f a class I peptide	
Deim	munization:										
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.											
	T cell immu	nogenicity pre	diction:								
The	server is de unogenicity	veloped to pr using 7-allel	edict the allele method (<u>Paul</u>	independe et. al. 20	ent CD4 <u>15</u>), imn	T cell immi nunogenicit	unogenicity y method	y at population le and combined m -allele method a	evel. User ca nethod (IEDE	an predict the T cell 3 recommended).	

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 I cell Imm	lunogenicity prediction									
	Specify Sequence(s)									
Enter epitope sequence(s) in PLAIN or FASTA format										
Or upload epitope sequence(s) from a file	Browse No file selected.									
	Choose a prediction method									
Prediction method:	IEDB recommended (combined)									
Sort Peptides by:	Position in Protein									
Select maximum percentile rank threshold:	50 V Position in Protein									
Enter the Job Name (Optional)										
Email address (optional)										
	Submit Reset									

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	spIP01588IEPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	WLLLSLLSLPLGLPV	11	25	42.16452	95.0613	LLSLLSLPL	6.9	25.0	3.2	3.6	73.0	33.0	6.9	6.5
1	splP01588 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	splP01588IEPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	VSGLRSLTTLLRALG	126	140	44.95964	86.8991	LTTLLRALG	17.0	12.0	17.0	9.3	70.0	20.0	20.0	1.3
1	splP01588lEPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	SLTTLLRALGAQKEA	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	PLRTITADTFRKLFR	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	TADTFRKLFRVYSNF	161	175	46.66984	44.6746	FRKLFRVYS	48.0	63.0	58.0	23.0	48.0	33.0	53.0	24.0
1	splP01588 EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	RKLFRVYSNFLRGKL	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).

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



Predicting HLA CD4 Immunogenicity in Human Populations.

<u>Dhanda SK¹, Karosiene E¹, Edwards L¹, Grifoni A¹, Paul S¹, Andreatta M², Weiskopf D¹, Sidney J¹, Nielsen M^{2,3}, Peters B^{1,4}, Sette A^{1,4}.</u>

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