La Jolla Institute FOR ALLERGY AND IMMUNOLOGY | Life Without Disease.

Deimmunization, Class II Immunogenicity (CD4EpiScore), Naturally Processed (MHCII-NP) peptide predictions

Sandeep Kumar Dhanda

sdhanda@lji.org

IEDB user workshop 2018

Oct 22-23, 2018

Deimmunization tool

Deimmunization tool: Background

- Native and engineered proteins are widely used to treat a variety of ailments.
- Non-self nature of engineered proteins may cause immunogenicity
- Self protein (e.g. EPO) can also be immunogenic, when administrated in high dose
- Immunogenicity of therapeutic proteins may be associated with safety issues.
- These therapeutic proteins can be de-immunized by removal of T cell epitopes (King et al. 2014)

Deimmunization tool

- To remove the epitope based on binding affinities to a set of reference MHC II alleles.
- The tool is optimized on EPO data (Tangri et. al. 2005).
- The concept of deimmunization was experimentally tested on '**Vatreptacog alpha**': a drug discontinued in clinical trial phase III due to immunogenicity issues.
- Steps
  - Predict the immunogenic regions in given sequence
  - Suggest amino acid substitution.
- Published in Immunology (Dhanda et al. 2018)

Predicting immunogenic regions

http://tools.iedb.org/deimmunization/

Home | Help | Example | Reference | Contact

Deimmunization tool

Step 1/2 (Predicting Immunogenic regions in the given protein sequence/s)

Specify Sequence(s)

Enter epitope sequence(s) in FASTA format

```
>sp|P01588|EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1
APPRLICDSRVLERYLLEAKEAENITTGCAEHC
SLNENITVPDTKVNFYAWKRMEVGQQAVEVWQGLALLSEAVLRGQALLVNSSQPWEPLQ
L
HVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRTITADTFRKLFRVYSNFLRGKL
KLYTGEACRTGDR
```

Or upload epitope sequence(s) from a file — Choose File — No file chosen

Select Median Percentile Rank Threshold

Select maximum median percentile rank threshold: 20

Go to Next step | Reset

Predicting immunogenic regions

APPRLICDSRVLERYLLEAKEAENITTGCAEHCSLNENITVPDTKVNFYAWKRMEVGQQAVEVWQG
LALLSEAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLLRALGAQKEAISPPDAASAAPLRT
ITADTFRKLFRVYSNFLRGKLKLYTGEACRTGDR

Generating overlapping peptides from given protein sequence.

All 15mer, overlapping 10 amino acids.

APPRLICDSRVLERY	1-15	38.89	Non-immunogenic
ICDSRVLERYLLEAK	6-20	60.39	Non-Immunogenic
VLERYLLEAKEAENI	11-25	23.46	Non-immunogenic
LLEAKEAENITTGCA	16-30	21.76	Non-immunogenic
----------	----------	----------	----------
DTFRKLFRVYSNFLR	136-150	6.08	Immunogenic

Allele : A set of 26 reference MHC class II alleles (Greenbaum et al. 2011) covering 95% of populations.

Method : IEDB Recommended, which is a consensus of three methods for predicting MHC class II binding.

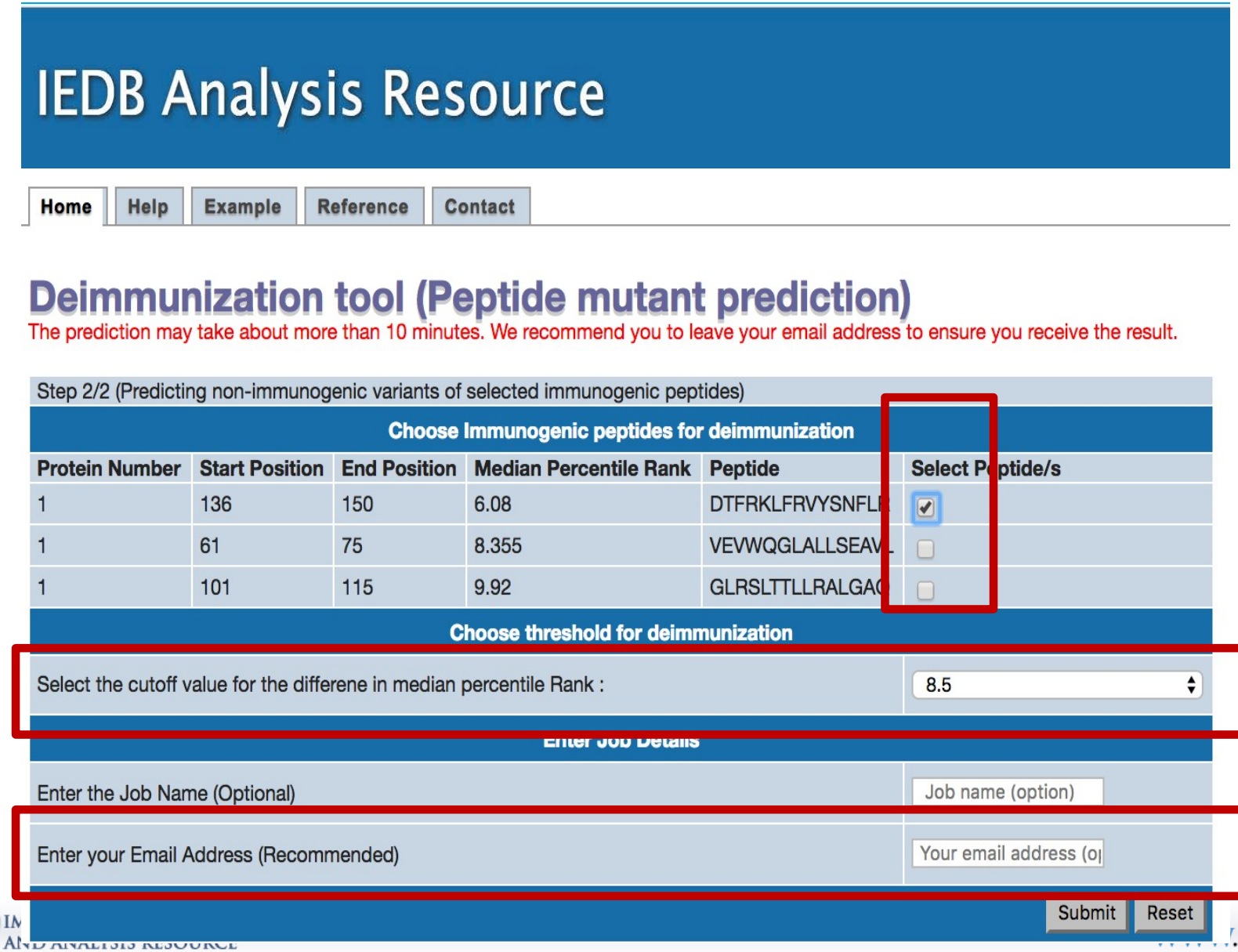
Median Percentile Rank: Take the median value of consensus percentile rank from 26 alleles.

Lower Rank corresponds to MHC binding.

Default threshold of median percentile rank is <=20

User can change the threshold

Deimmunization tool – Intermediate page



Suggest amino acid substitutions

DTFRKLFRVYSNFLR 136-150 6.08 Immunogenic

Generate all possible possible analogs in the selected 15mer (19*15) = 285

Predict immunogenicity with the same approach (as in step 1)

Protein Number	Peptide	Peptide ID	Start Position	End Position	Median Percentile Rank	Median Difference
1	DTFRKLFRVYSNFLR	wild	136	150	6.08	0.0
1	DTFRKLDRVYSNFLR	F142D	136	150	18.88	12.8
1	DTFRKLGRVYSNFLR	F142G	136	150	17.84	11.76
1	DTFRKLERVYSNFLR	F142E	136	150	17.25	11.17
1	DTFRKLFRVYSNFDR	L149D	136	150	17.225	11.145
1	DTFRKLPRVYSNFLR	F142P	136	150	17.11	11.03
1	DTFRKPFRVYSNFLR	L141P	136	150	16.89	10.81
1	DTFRKLFRVYSNFGR	L149G	136	150	16.265	10.185
1	DTFRKLFRCYSNFLR	V144C	136	150	15.405	9.325
1	DTFRKLFRVYSNFNR	L149N	136	150	15.11	9.03
1	DTFRKLHRVYSNFLR	F142H	136	150	14.9	8.82

These substitutions may have affected the neighboring peptides.

In a 15mer (overlapping 10mer), two neighbors can get affected.

RTITADTFRKLDRVY 131-145

DTFRKLDRVYSNFLR 136-150

LDRVYSNFLRGKLKL 141-155

Threshold is based on the difference of median percentile rank

Default is >=8.5 (in this case 6.08 + 8.5 = 14.51)

Any peptide analog with median percentile rank of >=14.51 will be suggested

User can change the threshold

Suggest amino acid substitutions

Protein Number	Peptide	Peptide ID	Start Position	End Position	Median Percentile Rank	Median Difference	C terminal Neighbor 1 (Median)	C terminal Neighbor 2 (Median)	N terminal Neighbor 1 (Median)	N terminal Neighbor 2 (Median)	Deimmunization Score
1	DTFRKLFRVYSNFLR	wild	136	150	6.08	0.0	10.35	24.645	28.525	26.645	NA
1	DTFRKLDRVYSNFLR	F142D	136	150	18.88	**12.8**	18.17	NA	35.715	NA	**3.0**
1	DTFRKLGRVYSNFLR	F142G	136	150	17.84	**11.76**	15.1	NA	30.99	NA	**3.0**
1	DTFRKLERVYSNFLR	F142E	136	150	17.25	**11.17**	15.615	NA	33.405	NA	**3.0**
1	DTFRKLFRVYSNFDR	L149D	136	150	17.225	**11.145**	30.055	45.8	NA	NA	**3.0**
1	DTFRKLPRVYSNFLR	F142P	136	150	17.11	**11.03**	16.61	NA	31.26	NA	**3.0**
1	DTFRKPFRVYSNFLR	L141P	136	150	16.89	**10.81**	14.09	NA	40.215	NA	**3.0**
1	DTFRKLFRVYSNFGR	L149G	136	150	16.265	**10.185**	24.185	43.53	NA	NA	**3.0**
1	DTFRKLFRCYSNFLR	V144C	136	150	15.405	**9.325**	16.455	NA	30.845	NA	**3.0**
1	DTFRKLFRVYSNFNR	L149N	136	150	15.11	**9.03**	15.69	40.485	NA	NA	**3.0**
1	DTFRKLHRVYSNFLR	F142H	136	150	14.9	**8.82**	10.545	NA	29.675	NA	**3.0**

Possible scenarios

1. Neighboring peptides absent (for terminal residues or small peptides)
2. Immunogenicity is reduced for peptides
3. Immunogenicity remain the same (No effect)
4. Increased immunogenicity

Deimmunization score

Immunogenicity for Neighboring peptide (1)	Immunogenicity for Neighboring peptide (2)	Score
Absent	Absent	1
Absent	Reduced	2
Reduced	Reduced	3
Absent	Neutral	4
Reduced	Neutral	5
Neutral	Neutral	6
Absent	Increased	7
Reduced	Increased	8
Neutral	Increased	9
Increased	Increased	10

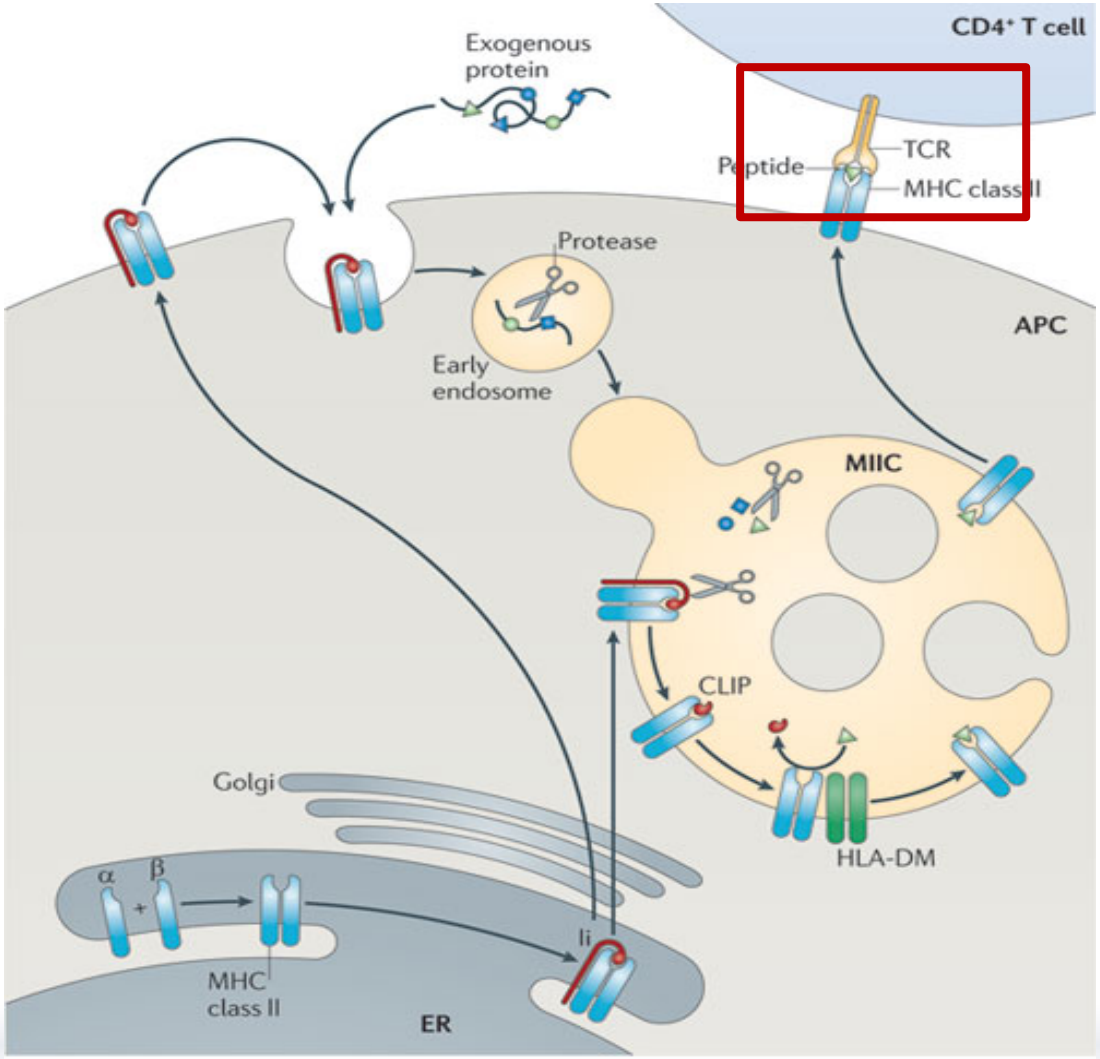
CD4EpiScore tool

Prediction of CD4 T cell epitope

T cell epitopes: Current approaches

- MHC binding prediction
  - Need allele information
  - Several methods available at IEDB
  - Individual based approach
- Promiscuous binders
  - Binding to several alleles
  - 27 Reference alleles (Greenbaum et al. 2011)
  - Population based approach
- Immunodominant epitopes
  - Immunogenic in broader population
  - Based on 7-allele method (Paul et al. 2015)
  - Population based approach

CD4 T cell epitopes



MHC class II binding and immunogenicity

The Majority of Immunogenic Epitopes Generate CD4$^+$ T Cells That Are Dependent on MHC Class II-Bound Peptide-Flanking Residues

Paula Y. Arnold, Nicole L. La Gruta, Tim Miller, Kate M. Vignali, P. Scott Adams, David L. Woodland and Dario A. A. Vignali

J Immunol July 15, 2002, 169 (2) 739-749; DOI: https://doi.org/10.4049/jimmunol.169.2.739

PNAS

A hairpin turn in a class II MHC-bound peptide orients residues outside the binding groove for T cell recognition

Zarixia Zavala-Ruiz[†], Iwona Strug[‡], Bruce D. Walker[§], Philip J. Norris[§¶], and Lawrence J. Stern[‡‖]

European Journal of Immunology

Article | Free Access

Recognition of core and flanking amino acids of MHC class II-bound peptides by the T cell receptor

Derek B. Sant'Angelo, Eve Robinson, Charles A. Janeway, Jr., Lisa K. Denzin

First published: 22 August 2002

CD4EpiScore

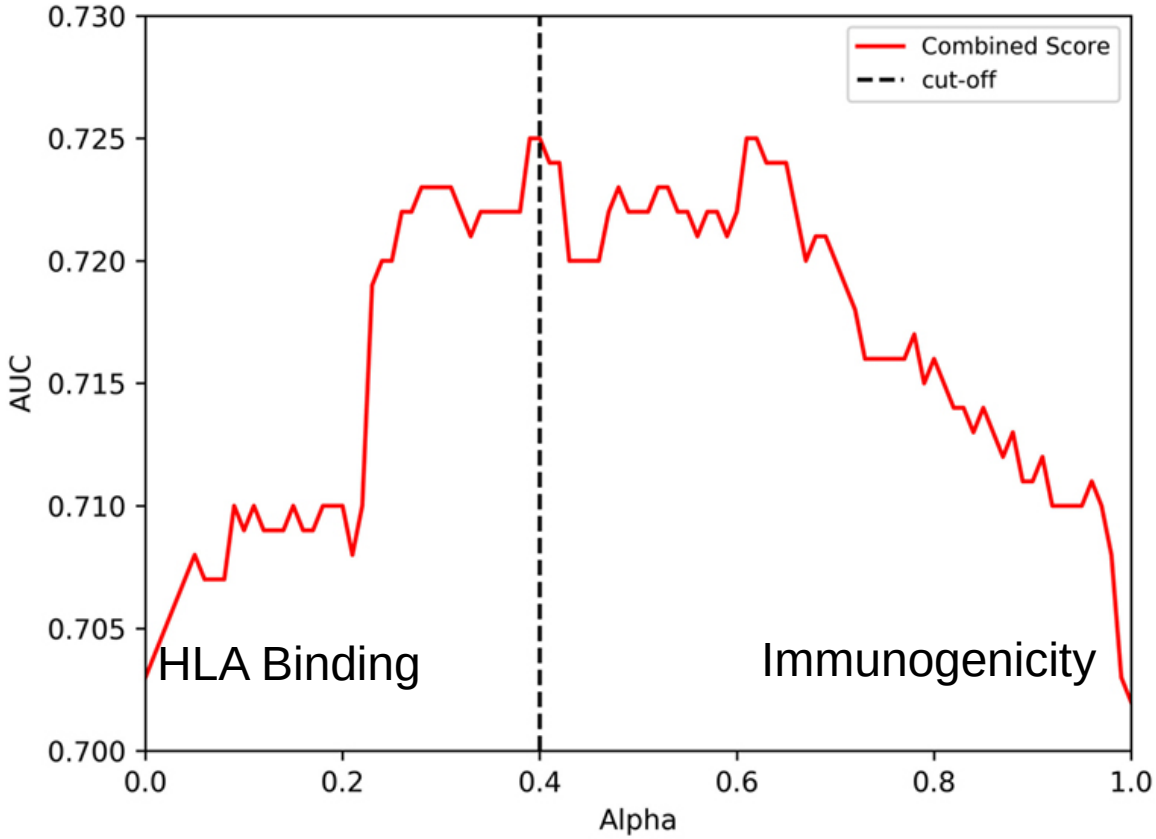
- To predict epitope score for CD4 T cells or class II immunogenicity.
- Based on Neural network model trained on
  - In house dataset for different antigens tested on different population cohorts
  - Tetramer dataset- derived from IEDB
- Validated on 57 independent studies from different groups across the world
- Implemented three approaches
  - 7-allele method (Paul et. al. 2015)
  - Immunogenicity predictions
  - Hybrid approach

Immunogenicity predictions

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).

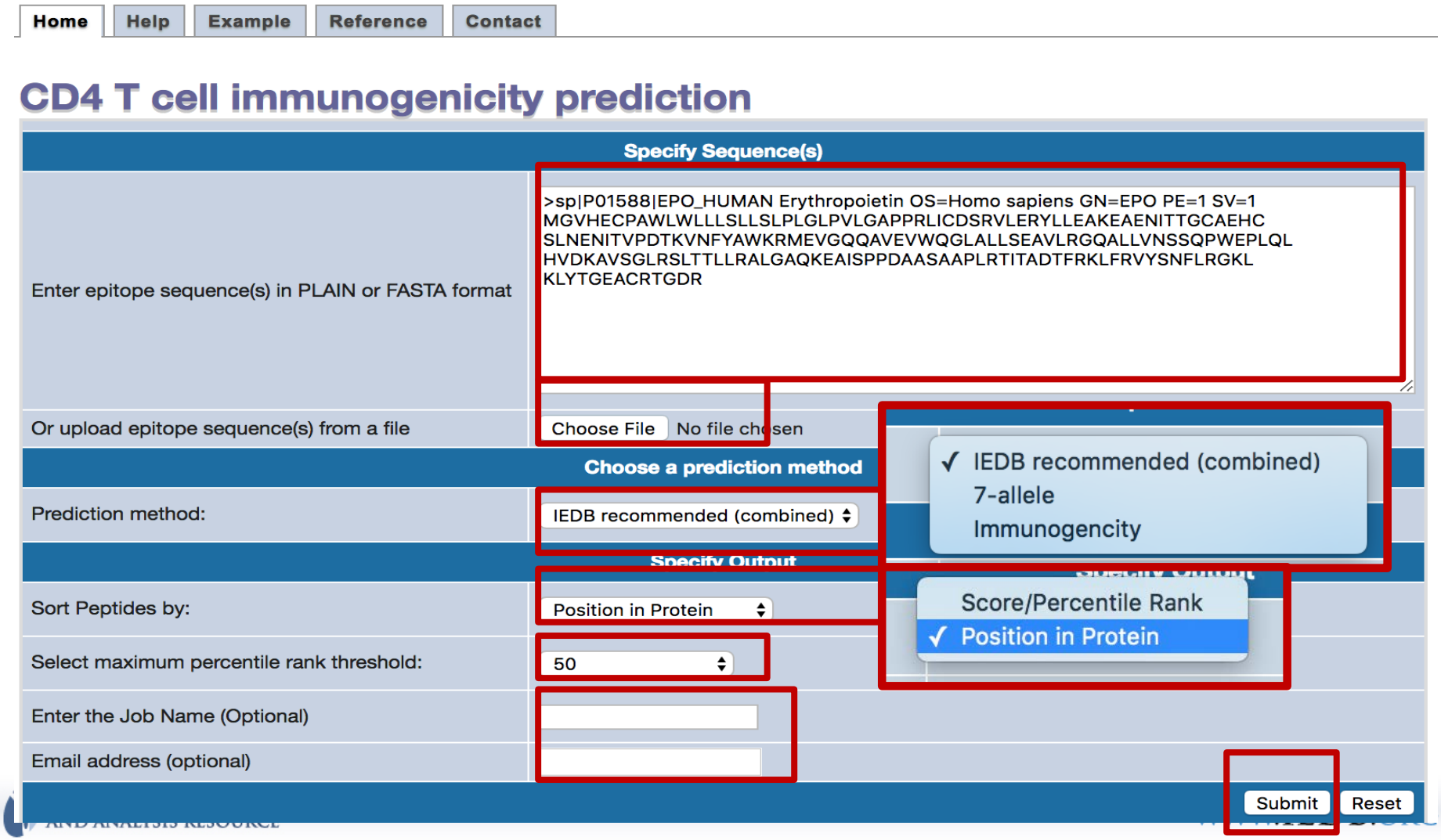
$$\text{Combined score} = \alpha \times \text{Imm score} + (1 - \alpha) \times \text{HLA score.}$$



Dhanda et. al. 2018, Frontiers in Immunology, 9, 1369

CD4EpiScore: Input page

http://tools.iedb.org/CD4EpiScore



CD4EpiScore: Results page

CD4 Immunogenicity prediction results

Number of proteins: 1

Number of 15mer (overlapping 10mer): 37

Threshold : 50.0%

Method : combined

Download result

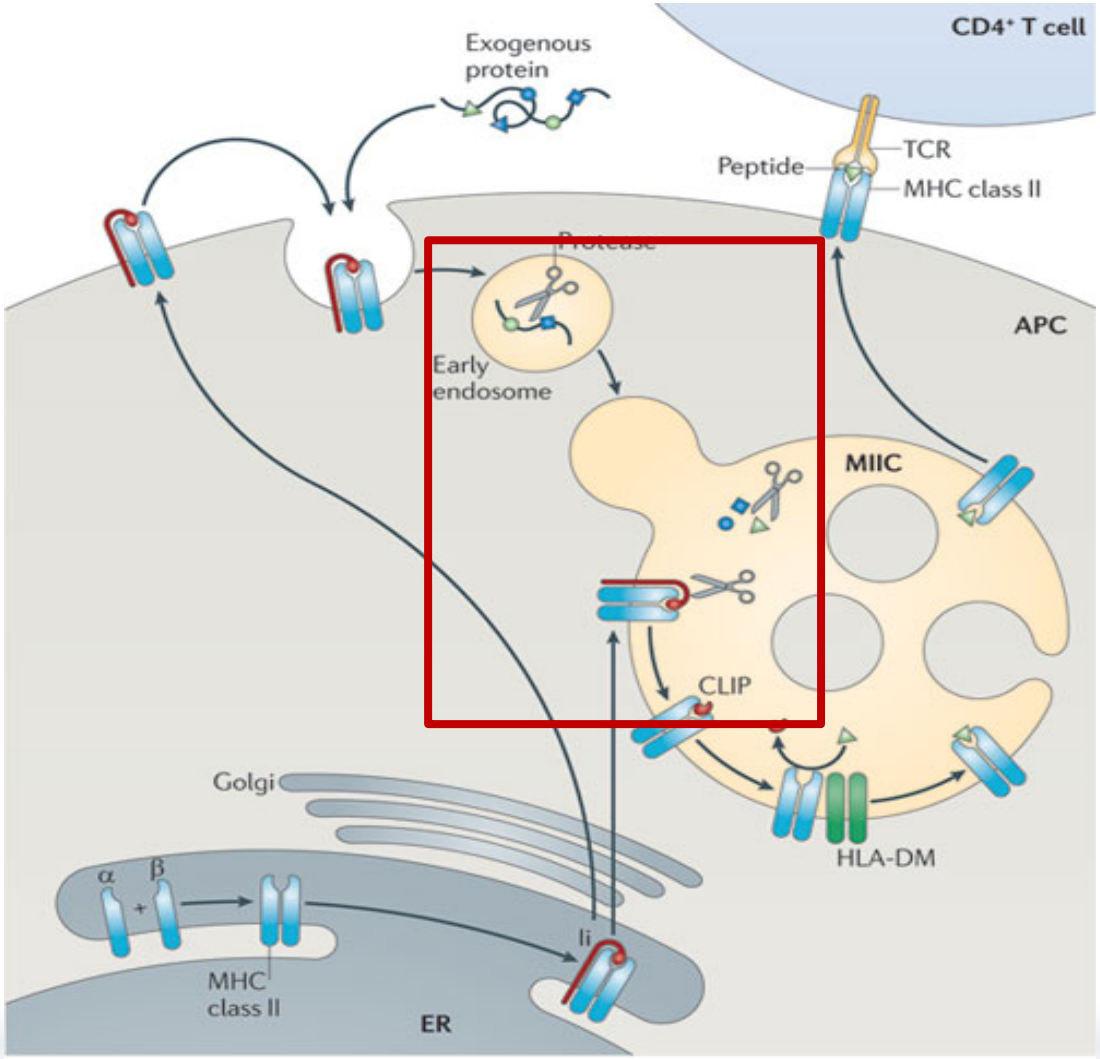
Citations

Protein Number	Protein Description	Peptide	Start	End	Combine 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	CPAWLWLLLSLLSLP	6	20	48.0978	99.7096	LWLLLSLLS	13.69	3.93	13.69	1.25	51.32	63.91	44.08	1.21
1	sp|P01588|EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	WLLLSLLSLPLGLPV	11	25	40.3825	95.0613	LLSLLSLPL	8.93	3.93	4.19	0.65	71.75	38.37	2.98	1.16
1	sp|P01588|EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	LGLPVLGAPPRLICD	21	35	48.0036	99.1292	LPVLGAPPR	13.92	10.76	13.92	15.07	22.97	9.63	49.41	1.87
1	sp|P01588|EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	RLICDSRVLERYLLE	31	45	49.6284	93.921	LICDSRVLE	20.1	0.17	41.0	20.1	33.39	16.1	23.4	19.74
1	sp|P01588|EPO_HUMAN Erythropoietin OS=Homo sapiens GN=EPO PE=1 SV=1	ITVPDTKVNFYAWKR	66	80	49.2266	88.9416	TKVNFYAWK	22.75	22.75	35.9	13.47	22.15	60.35	61.81	9.97

MHCII-NP

Prediction of class II naturally processed peptides

CD4 T cell epitopes



MHCII-NP

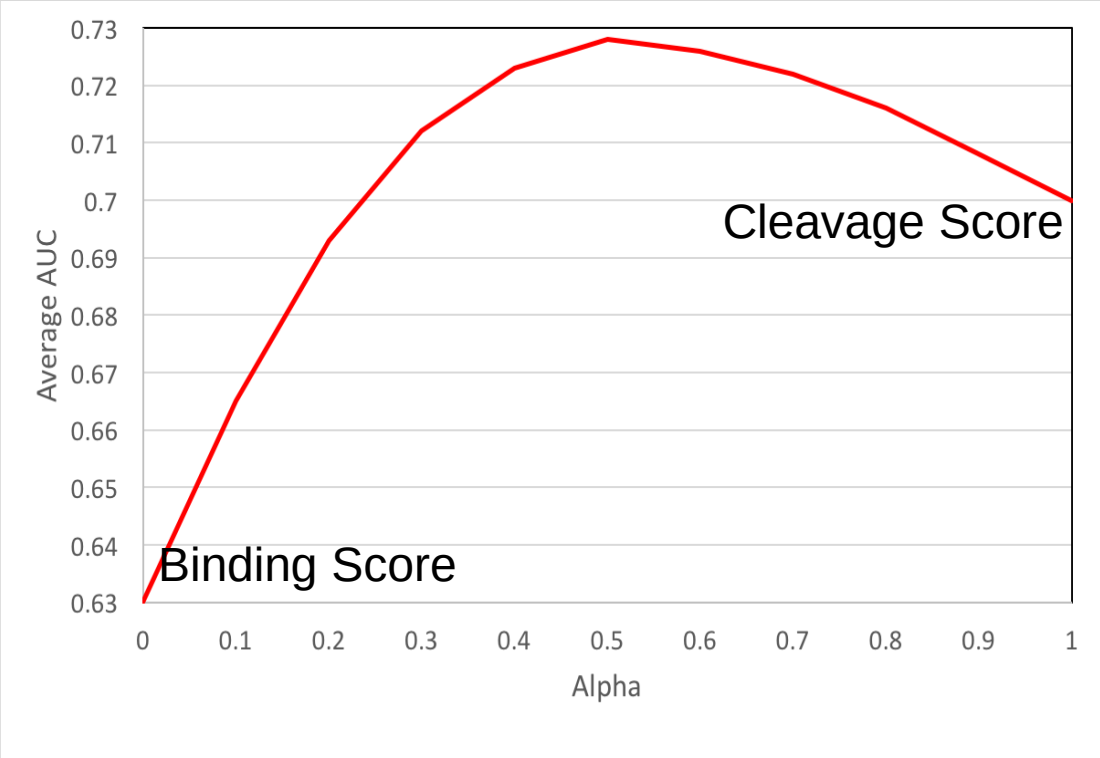
- Predicting the naturally processed peptides for MHC class II
- Based on
  - cleavage motif analysis at C and N terminal of peptides
  - Ligand elution data derived from IEDB
  - Ligand predictions is improved by combining the binding and cleavage motifs.
- T cell epitope prediction is not improved

Ligand predictions

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

Cleavage Score: Derived from the cleavage motif analysis in ligand elution data

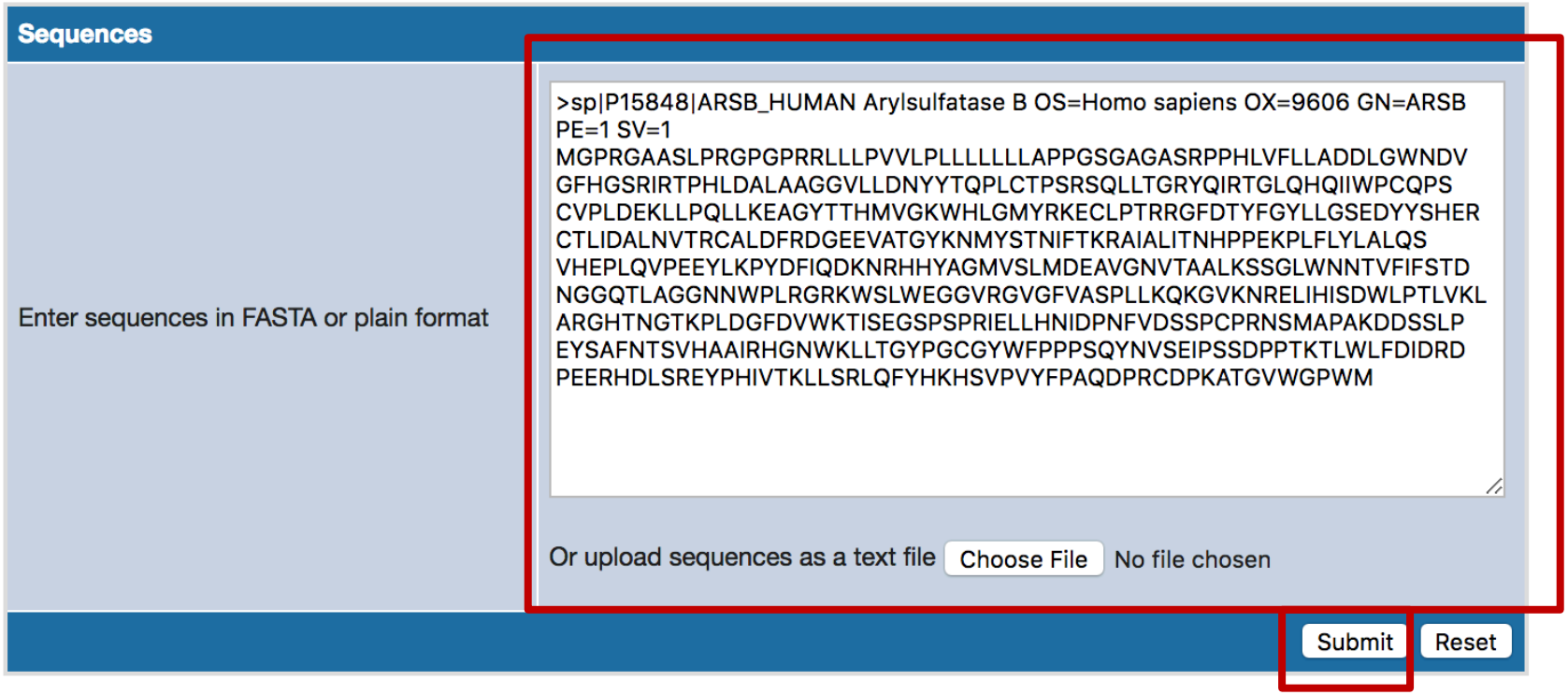
Binding Score: Derived from HLA binding affinity using 7-allele method (Paul et. al. 2015).



Paul et. al. 2018, Frontiers in Immunology, 9, 1795

MHCII-NP

MHC II NP - Prediction of naturally processed MHC II ligands



MHCII-NP Results

Seq name	Peptide start	Peptide end	Peptide length	Peptide N	motif	C motif	Cleavage probability score			Cleavage probability percentile rank	
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	2	16	15	GPRGAASLPRGPGPR	MGP	PRR	1.73735	0.02			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	247	261	15	VPEEYLKPYDFIQDK	QVP	DKN	1.4884	0.04			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	384	398	15	SPSPRIELLHNIDPN	GSP	PNF	1.4042	0.05			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	12	26	15	GPGPRRLLLPVVLPL	RGP	PLL	1.33714	0.07			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	480	493	14	DPEERHDLSREYPH	RDP	PHI	1.2645	0.09			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	2	17	16	GPRGAASLPRGPGPRR		MGP	RRL	1.10023	0.11		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	33	46	14	APPGSGAGASRPPH	LAP	PHL	1.06464	0.12			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	253	268	16	KPYDFIQDKNRHHYAG		LKP	AGM	1.03988	0.14		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	247	260	14	VPEEYLKPYDFIQD	QVP	QDK	0.95759	0.16			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	466	479	14	DPPTKTLWLFDIDR	SDP	DRD	0.95486	0.18			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	462	476	15	IPSSDPPTKTLWLFD	EIP	FDI	0.86742	0.19			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	69	84	16	TPHLDALAAGGVLLDN		RTP	DNY	0.86541	0.21		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	375	388	14	DVWKTISEGSPSPR	FDV	PRI	0.86177	0.23			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	466	481	16	DPPTKTLWLFDIDRDP		SDP	DPE	0.83011	0.25		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	274	288	15	DEAVGNVTAALKSSG	MDE	SGL	0.81861	0.27			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	92	106	15	TPSRSQLLTGRYQIR	CTP	IRT	0.7967	0.3			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	33	45	13	APPGSGAGASRPP	LAP	PPH	0.7844	0.32			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	384	397	14	SPSPRIELLHNIDP	GSP	DPN	0.76819	0.34			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	510	523	14	VPVYFPAQDPRCDP	SVP	DPK	0.76819	0.34			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	230	244	15	KPLFLYLALQSVHEP	EKP	EPL	0.76271	0.37			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	334	349	16	SPLLKQKGVKNRELIH		ASP	IHI	0.75419	0.39		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	334	346	13	SPLLKQKGVKNRE	ASP	REL	0.74385	0.41			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	518	532	15	DPRCDPKATGVWGPW	QDP	PWM	0.73955	0.42			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	415	430	16	DDSSLPEYSAFNTSVH		KDD	VHA	0.72819	0.44		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	247	262	16	VPEEYLKPYDFIQDKN		QVP	KNR	0.72168	0.46		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	411	426	16	APAKDDSSLPEYSAFN		MAP	FNT	0.70948	0.48		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	122	135	14	VPLDEKLLPQLLKE	CVP	KEA	0.69567	0.5			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	53	68	16	DDLGWNDVGFHGSRIR		ADD	IRT	0.67205	0.51		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	416	430	15	DSSLPEYSAFNTSVH	DDS	VHA	0.65617	0.53			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	395	407	13	IDPNFVDSSPCPR	NID	PRN	0.6459	0.55			
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	20	35	16	LPVVLPLLLLLLLAPP		LLP	PPG	0.64372	0.57		
SP|P15848|ARSB_HUMAN ARYLSULFATASE B OS=HOMO SAPIENS OX=9606 GN=ARSB PE=1 SV=1	411	424	14	APAKDDSSLPEYSA	MAP	SAF	0.63924	0.58			

Thank you for attention !!