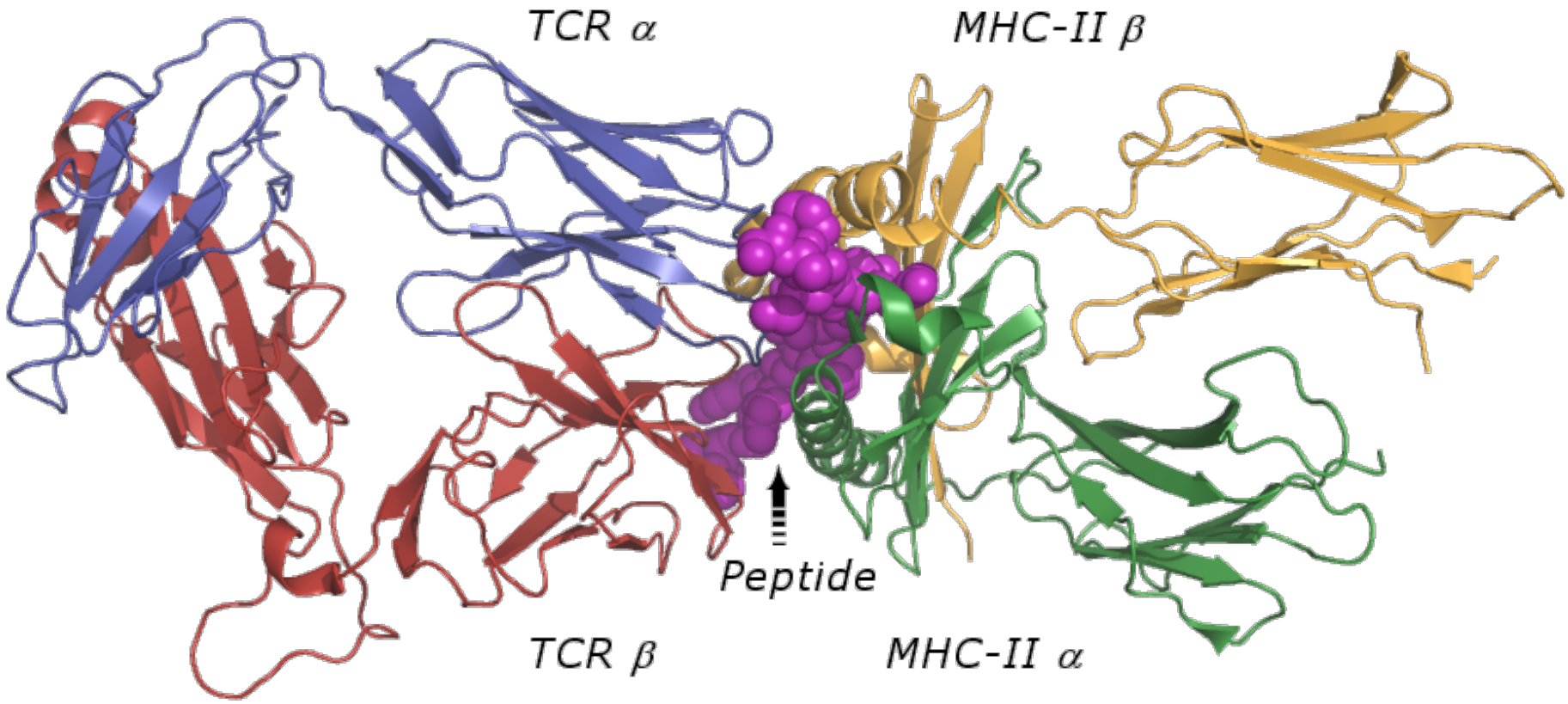
HLA class II peptide binding, T cell recognition and the deimmunization of proteins

T cells activation depends on the formation of a trimolecular complex



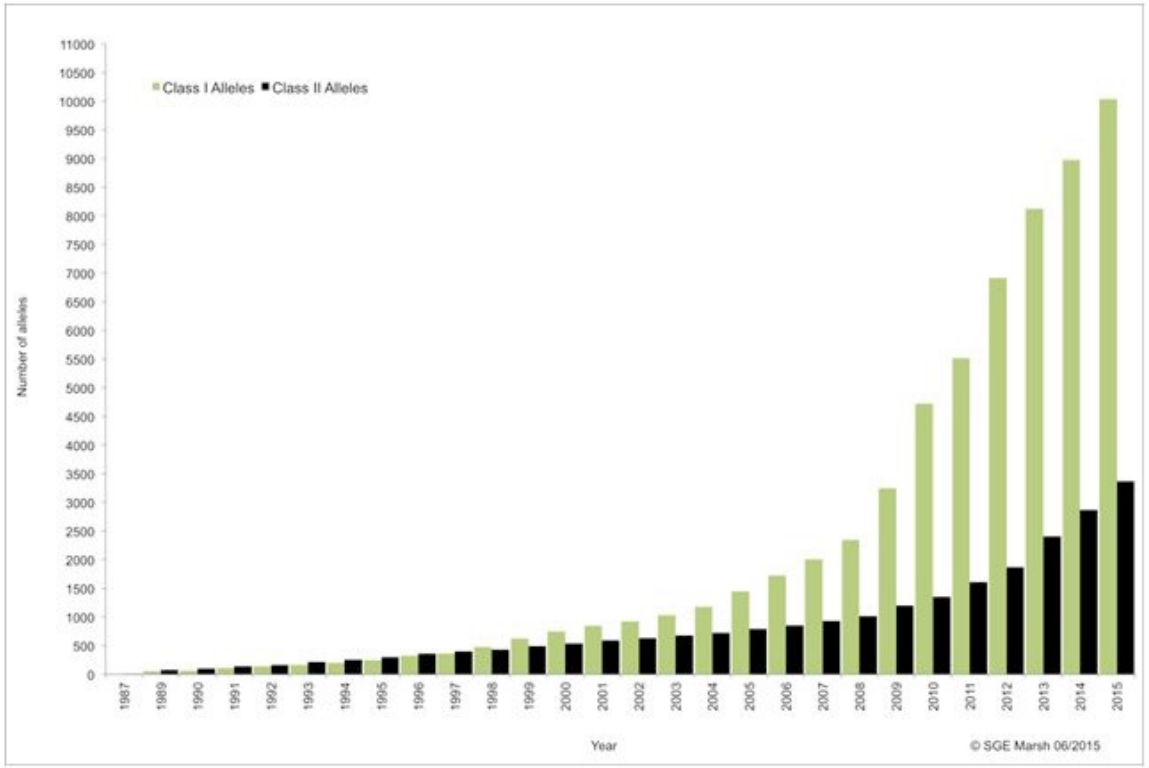
Some key points for strategies to predict HLA class II binding and TCR recognition

- HLA binding is necessary but not sufficient for TCR recognition
- HLA binding predictions predict binding but not necessarily TCR recognition
- HLA binding predictions are allele specific
- However, most applications require predictions at the level of
  - individual subjects>8 alleles
  - responding/treated population>hundreds of alleles
- What is required is a actionable strategies to target not alleles, but individuals and populations

Reducing/eliminating HLA class II binding is a common de-immunization strategy

- Which alleles?
- Which substitutions?

Which HLA alleles should be considered for deimmunization approaches?



IMGT/HLA database

12.672 HLA alleles

A mean of 450 new alleles/ Year

Panel of HLA class II alleles to allow for global coverage

Locus	Molecule	Phenotype frequency
DRB1	DRB1*0101	5.4
	DRB1*0301	13.7
	DRB1*0401	4.6
	DRB1*0405	6.2
	DRB1*0701	13.5
	DRB1*0802	4.9
	DRB1*0901	6.2
	DRB1*1101	11.8
	DRB1*1201	3.9
	DRB1*1302	7.7
	DRB1*1501	12.2
	Combined	71.1
DRB3/4/5	DRB3*0101	26.1
	DRB3*0202	34.3
	DRB4*0101	41.8
	DRB5*0101	16.0
	Combined	87.7

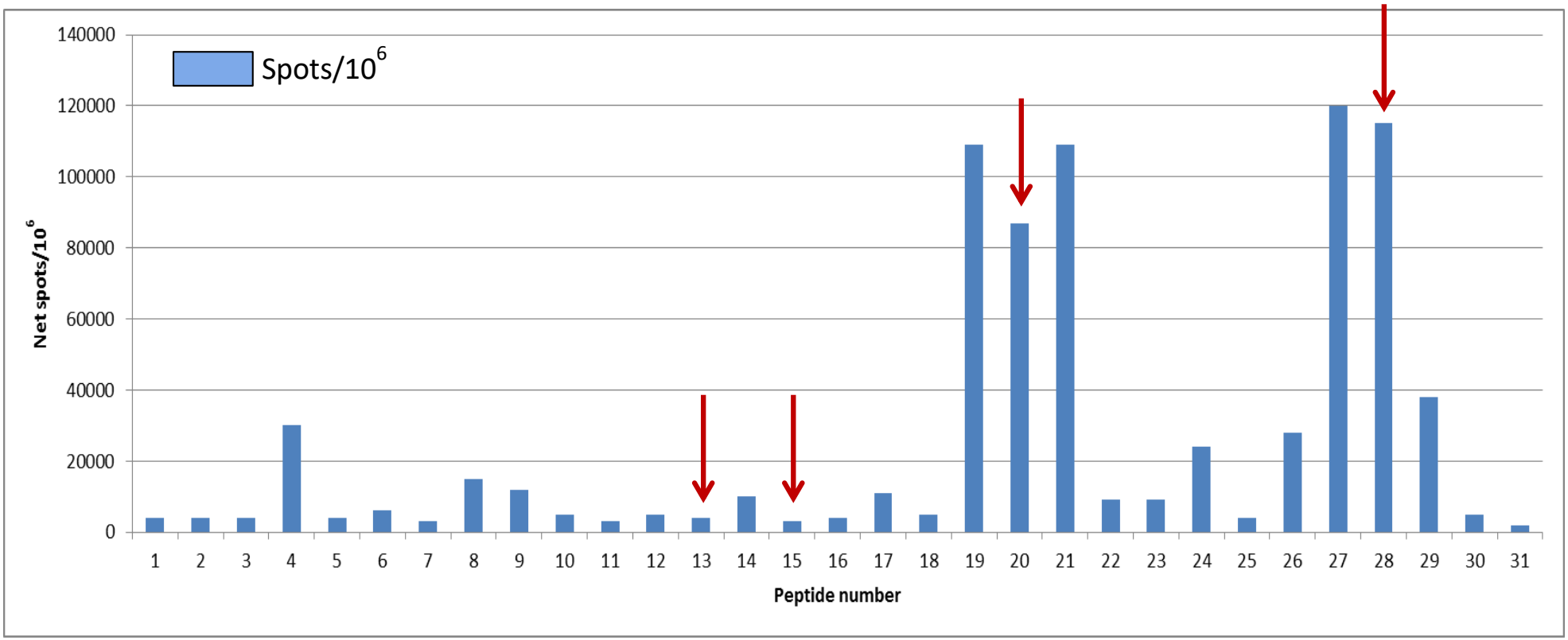
Locus	Molecule	Phenotype frequency
DQA1/DQB1	DQA1*0501/DQB1*0201	11.3
	DQA1*0501/DQB1*0301	35.1
	DQA1*0301/DQB1*0302	19.0
	DQA1*0401/DQB1*0402	12.8
	DQA1*0101/DQB1*0501	14.6
	DQA1*0102/DQB1*0602	14.6
	Combined	81.6
DPA1/DPB1	DPA1*0201/DPB1*0101	16.0
	DPA1*0103/DPB1*0201	17.5
	DPA1*01/DPB1*0401	36.2
	DPA1*0301/DPB1*0402	41.6
	DPA1*0201/DPB1*0501	21.7
	DPB1*1401@	7.4
	Combined	94.5

@ No algorithm available for DPB1*1401

Greenbaum et al., 2011. Immunogenetics

Dominant epitopes correlate with promiscuous binding

Antigenicity of EPO peptides (compiled)



→ Predicted promiscuous binders

Tangri S et al A. Rationally engineered therapeutic proteins with reduced immunogenicity. J Immunol. 2005 Mar 15;174(6):3187-96. PubMed PMID: 15749848.

Is predicting binding to a lot of HLAs necessarily best? A heuristic approach

- Peptide datasets spanning entire proteins associated with measured immune responses in exposed humans

Dataset		No. of Antigens	Total peptides	No. of donors		Reference
Der p/f (House dust mite)		4	156	20		Hinz,2015
Phl p (Timothy grass)		10	425	25		Oseroff, 2010
TB-1		4	71	18		Arlehamn, 2012
TB-2		11	499	32		Arlehamn,2015
Cockroach		6	463	19		Oseroff, 2012
Pertussis		9	785	23		Dillon, 2015
TOTAL		44	2399	137		

Paul S et al. Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. J Immunol Methods. 2015 PubMed PMID: 25862607

Immunogenicity data to train binding predictions strategies

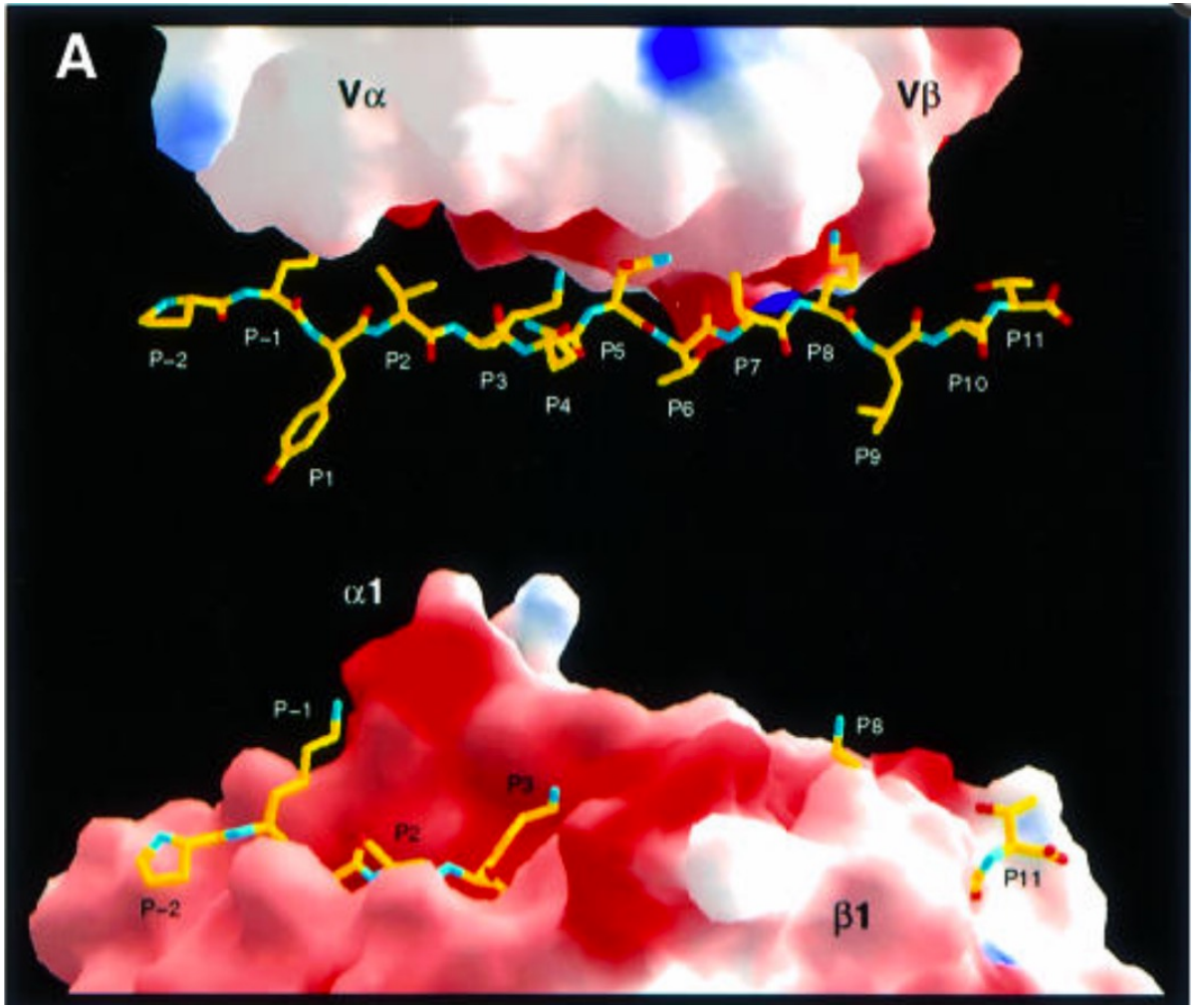
- Optimal results obtained calculating median consensus percentile ranks
- Surprisingly the results were optimal with a set of 7 alleles
- These alleles are representatives of a variety of binding modes and supertypes

Paul S et al. Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes. J Immunol Methods. 2015 PubMed PMID: 25862607

Reducing/eliminating HLA class II binding is a common de-immunization strategy

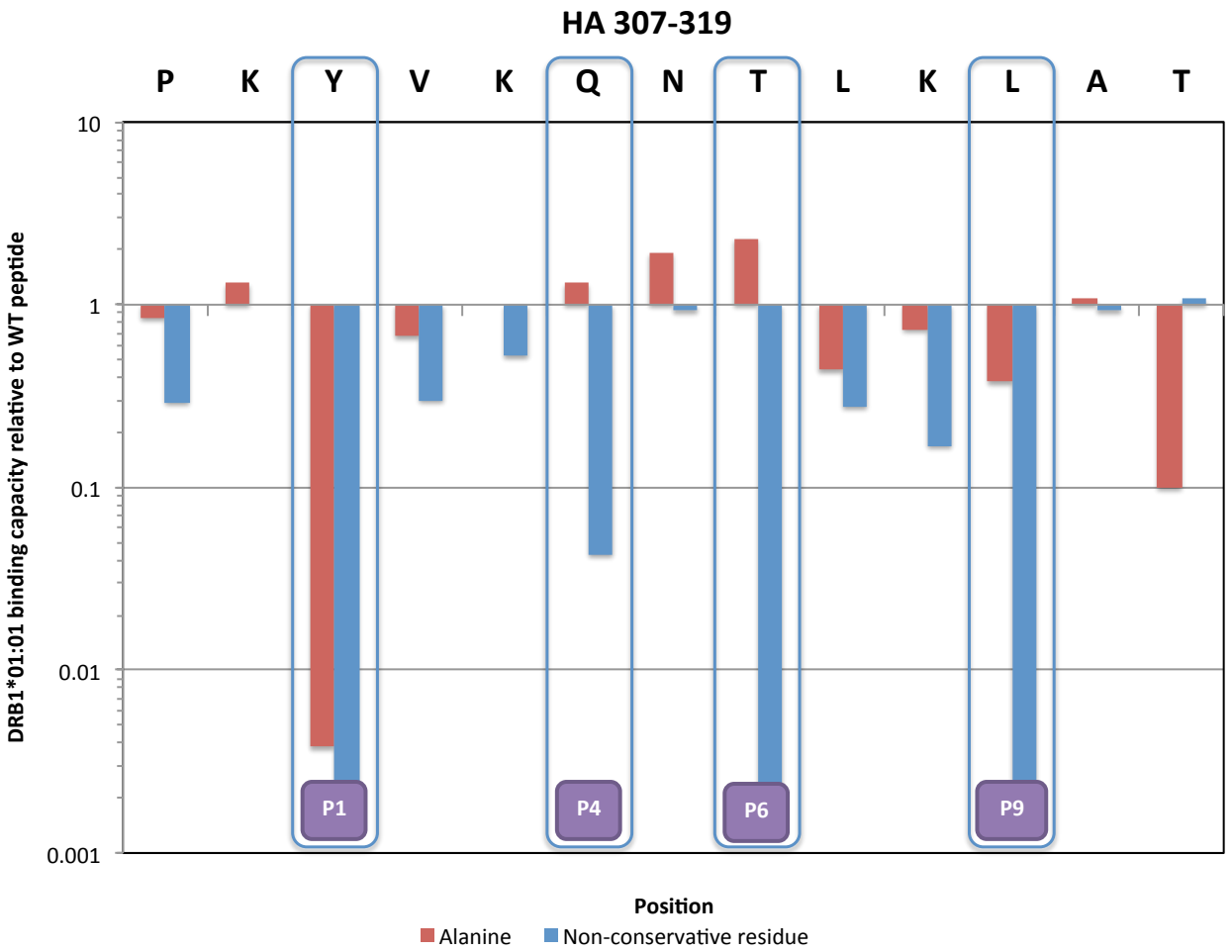
- Often rely on introducing alanine substitutions
- Preferred because the goal is to not to disrupt drug biological potency
- This goal needs to be balanced with the need to disrupt HLA binding
- What do alanine substitutions do? Let's look at the data….

HA 307-319 and DRB1*01:01
structure identifies P1, P4, P6 and P9 as HLA contact residues



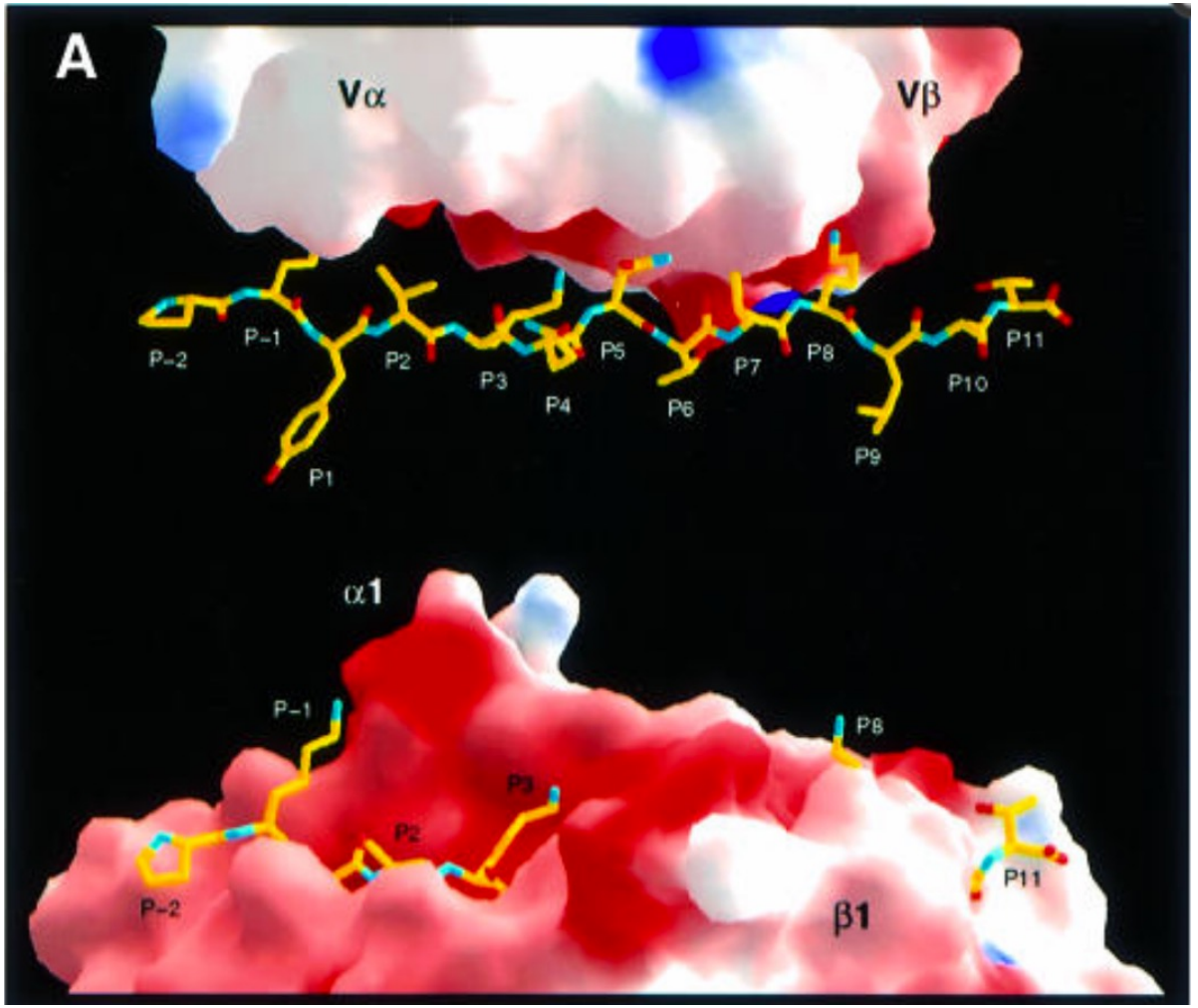
Hennecke et al., The EMBO Journal 19:5611 (2000)

Non-conservative substitutions, not alanine, identify HA 307-319 and DRB1*01:01 HLA contacts and disrupt HLA binding



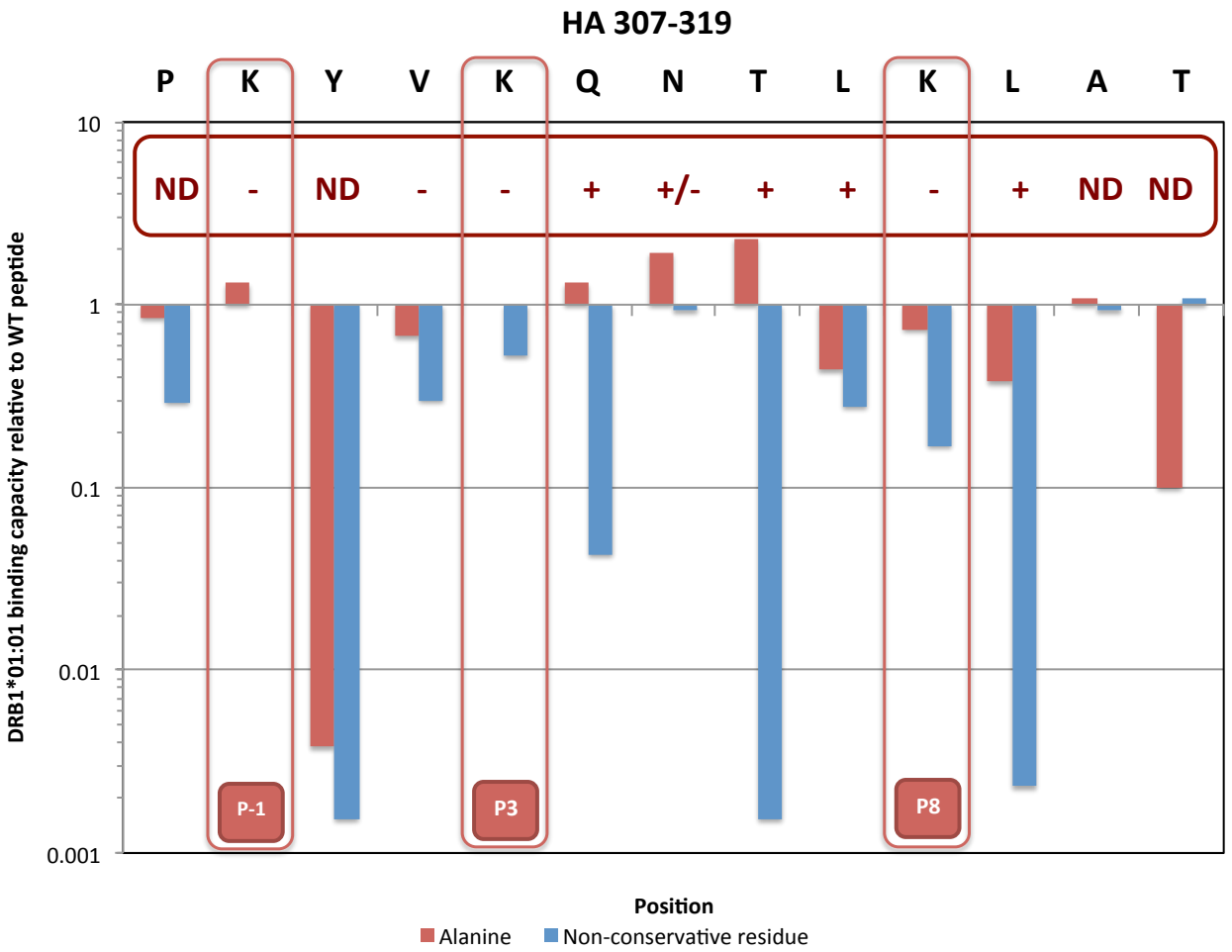
Alexander et al., J Immunol 150:1 (1993)

HA 307-319 and DRB1*01:01
structure identifies P-1, P3 and P8 as main TCR contact residues



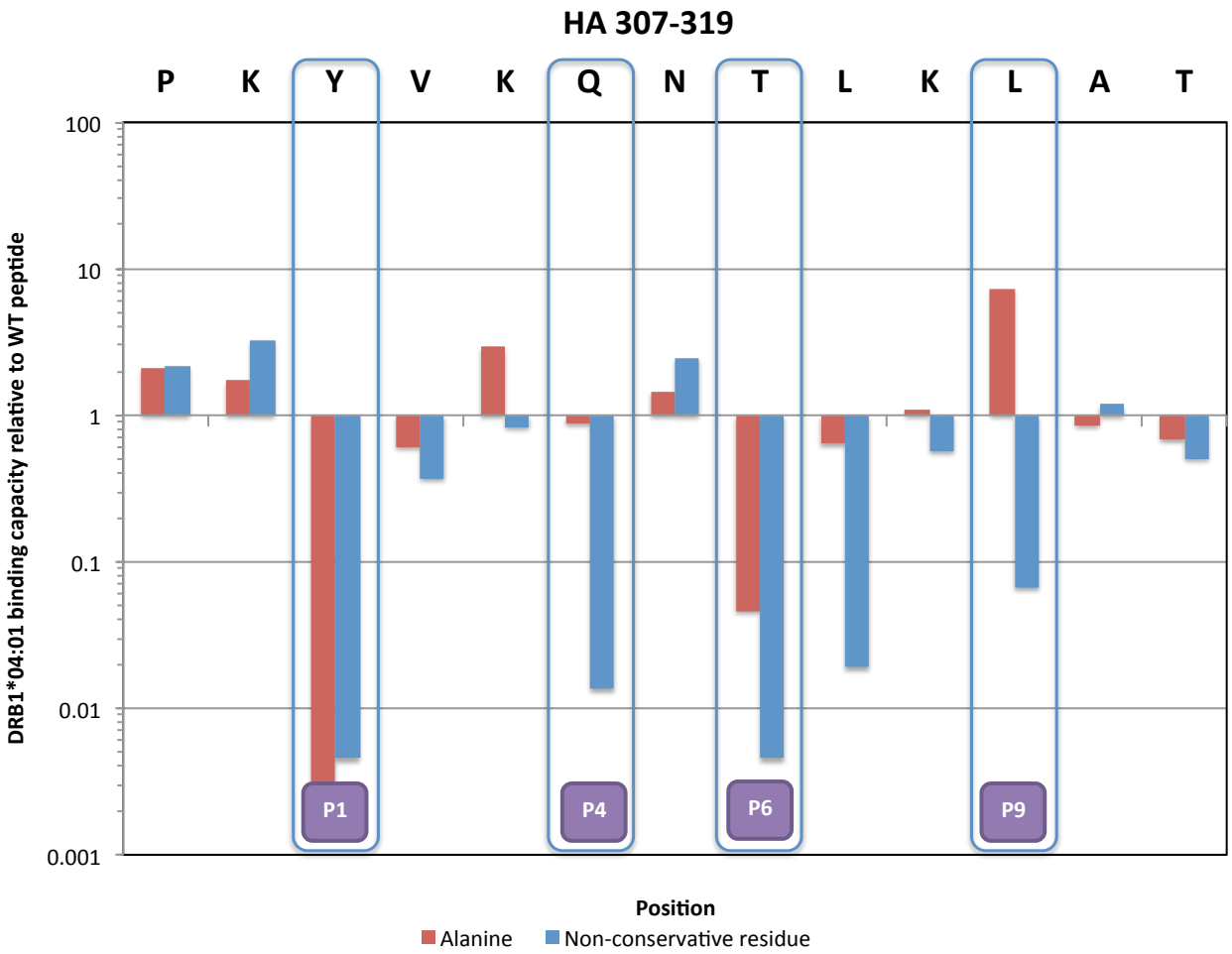
Hennecke et al., The EMBO Journal 19:5611 (2000)

DRB1*01:01 antigenicity of alanine containing HA 307-319 analogs



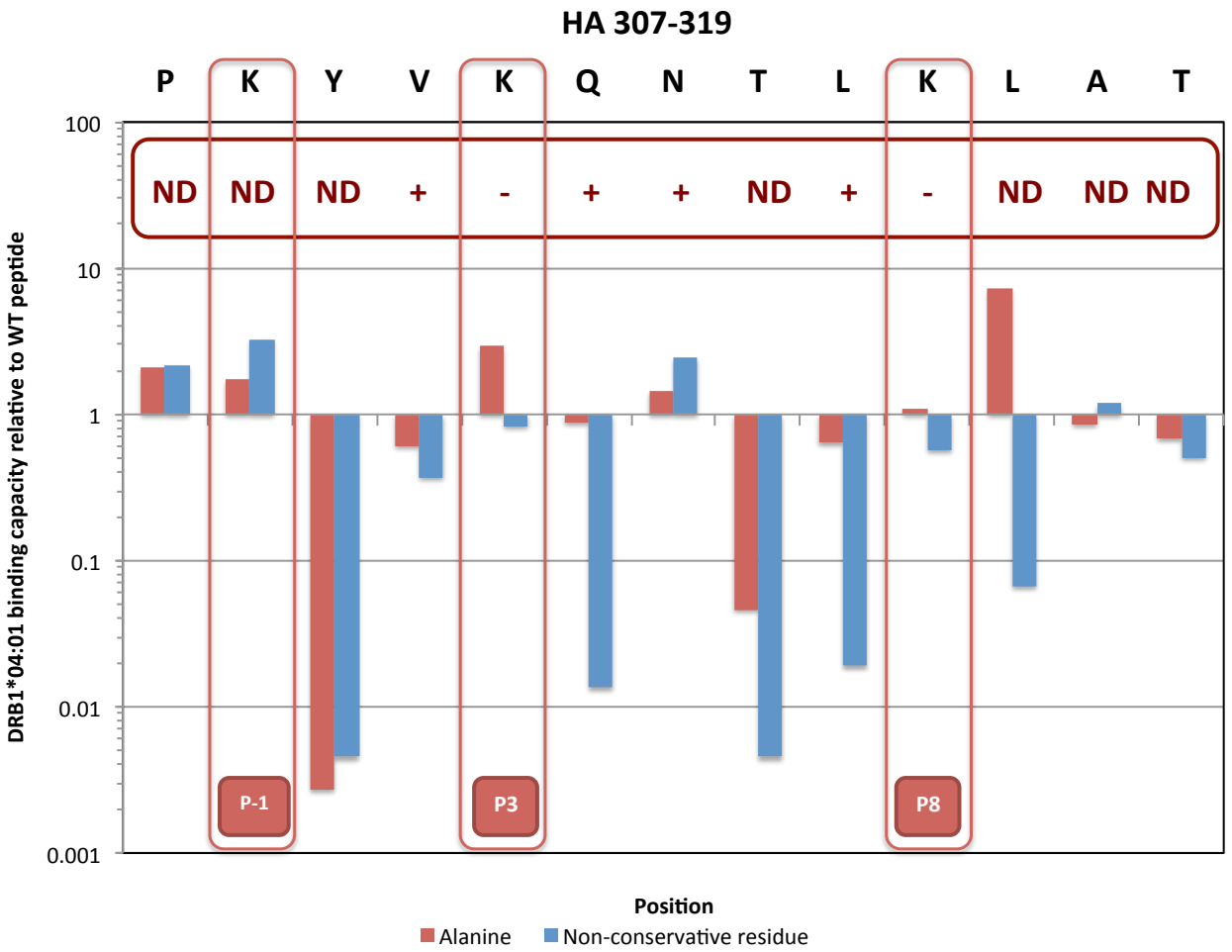
Alexander et al., J Immunol 150:1 (1993)

Non-conservative substitutions, not alanine, identify HA 307-319 and DRB1*04:01 HLA contacts and disrupt HLA binding



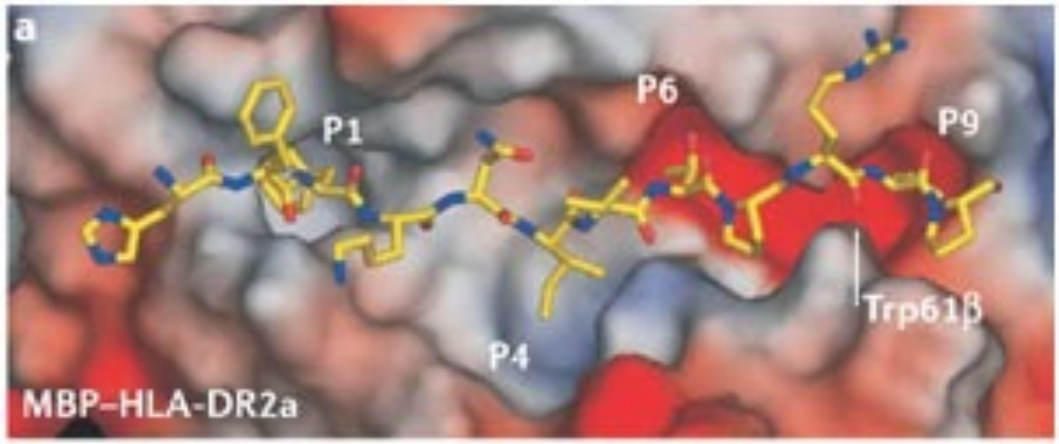
Alexander et al., J Immunol 150:1 (1993)

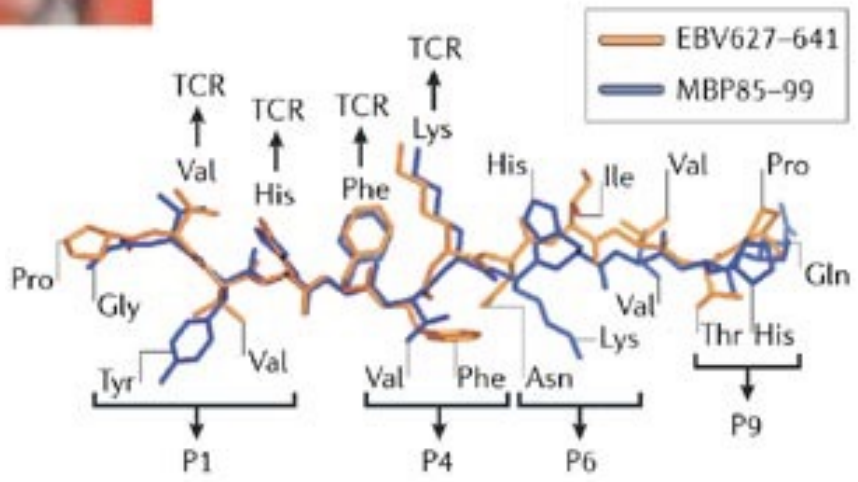
DRB1*04:01 antigenicity of alanine containing HA 307-319 analogs



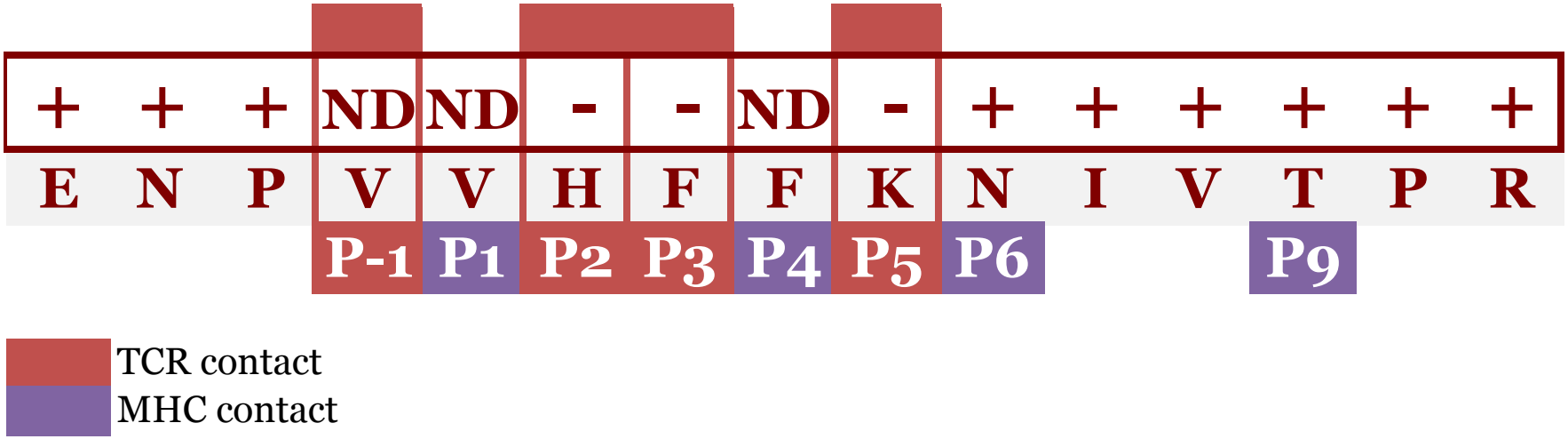
Alexander et al., J Immunol 150:1 (1993)

TCR and HLA contact residues for MBP 85-99 and DRB1*15:01



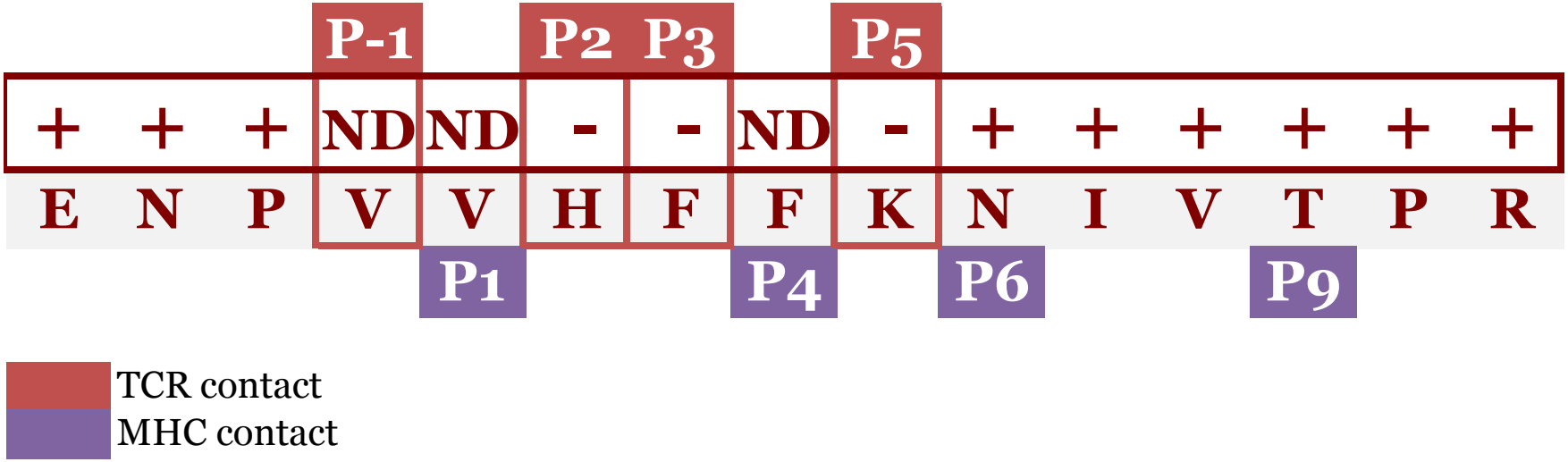


Alanine substitutions identify TCR contact residues for MBP 85-99 and DRB1*15:01



Wucherpfennig et al., J Exp Med 179:279 (1994)

Alanine substitutions identify TCR contact residues for MBP 85-99 and DRB1*15:01

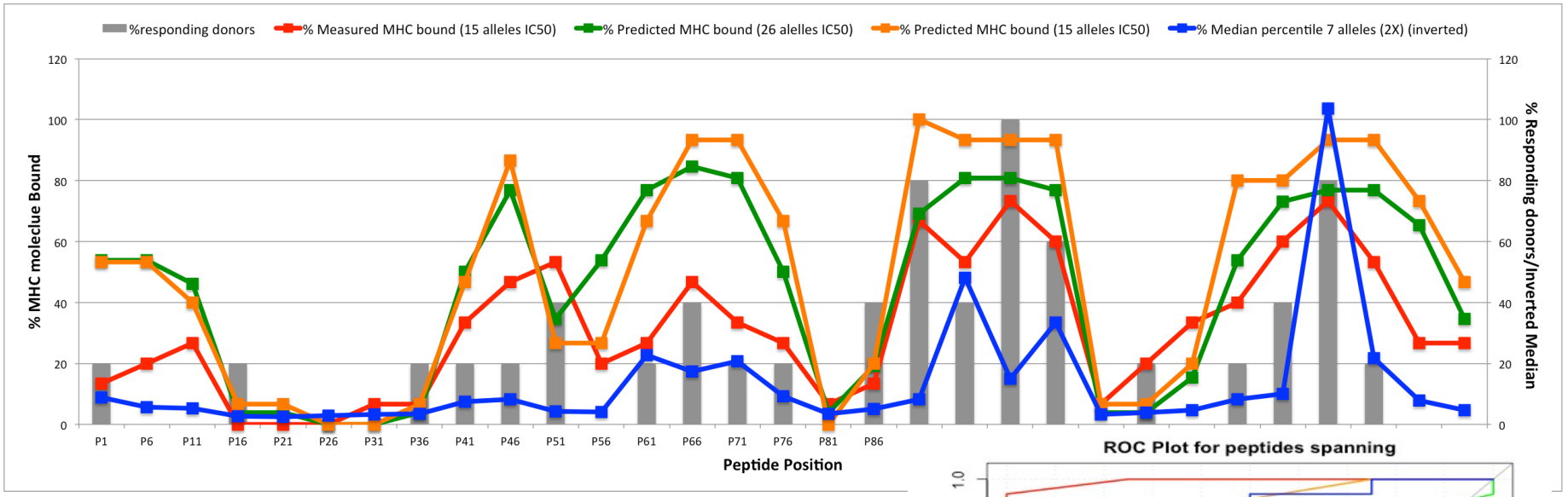


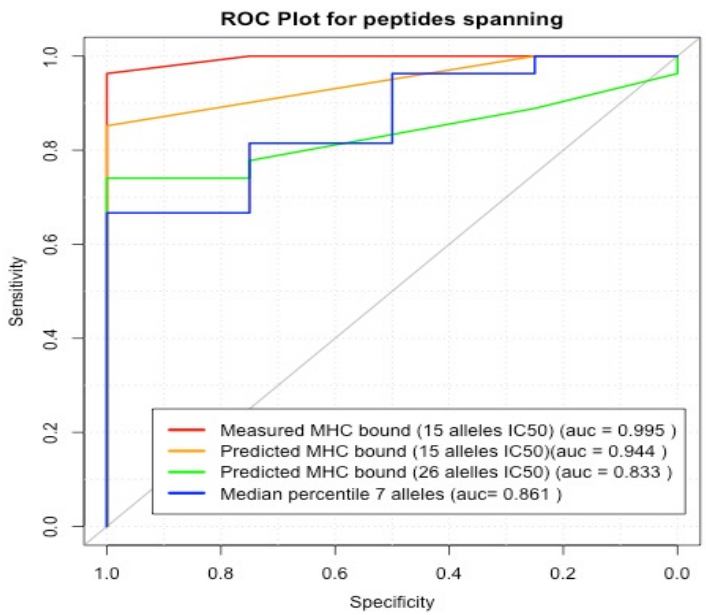
Wucherpfennig et al., J Exp Med 179:279 (1994)

Alanine substitutions

- Not a good way to map HLA contact residues (especially if high peptide concentrations are used in the assay)
- Not a good way to reduce HLA binding
- Likely to affect TCR recognition; testing the analog drug will yield lower activity in individuals exposed to the original drug
- Since the analog drug still binds HLA, the potential for immunogenicity is retained

A computational application to select candidate analogs for de-immunization





- **Used published data that mapped the two main immunogenic regions of EPO**
- **The 7-allele method predicts those two regions**

Dhanda SK et al. A.Development of a strategy and computational application to select candidate protein analogues with reduced HLA binding and immunogenicity. Immunology. 2017Aug 22. PubMed PMID: 28833085.

The effect of substitutions at various positions in the two main epitopes was also known

Table 1: Antigenic regions from Epo and their mutant analogs

	Peptide Sequence	% Responding donors	Immunogenicity
Region 1			
Wild	GLRSLTTLLRALGAQ	100	+
Analog 1	GARSLTTLLRALGAQ	75	+
Analog 2	GERSLTTLLRALGAQ	100	+
Analog 3	GGRSLTTLLRALGAQ	75	+
Analog 4	GPRSLTTLLRALGAQ	0	-
Analog 5	GSRSLTTLLRALGAQ	100	+
Analog 6	GLDSLTTLLRALGAQ	50	+
Analog 7	GLGSLTTLLRALGAQ	75	+
Analog 8	GLRRLTTLLRALGAQ	50	+
Analog 9	GLRSDTTLLRALGAQ	50	+
Analog 10	GLRSLTDLLRALGAQ	25	-
Analog 11	GGRSLTDLLRALGAQ	0	-
Analog 12	GSRSLTDLLRALGAQ	25	-
Region 2			
Wild peptide	DTFRKLFRVYSNFLR	100	+
Analog 1	DTFRKDFRVYSNFLR	25	-
Analog 2	DTFRKLFDVYSNFLR	75	+
Analog 3	DTFRKLFGVYSNFLR	100	+
Analog 4	DTFRKLFRDYSNFLR	50	+
Analog 5	DTFRKLFRGYSNFLR	50	+
Analog 6	DTFRKLFRRYSNFLR	75	+
Analog 7	DTFRKLFRVYDNFLR	0	-
Analog 8	DTFRKLFRVYSDFLR	75	-
Analog 9	DTFRKDFRVYDNFLR	0	-

Tangri S et al A. Rationally engineered therapeutic proteins with reduced immunogenicity. J Immunol. 2005 Mar 15;174(6):3187-96. PubMed PMID: 15749848.

The 7-allele D method predicts the effect of substitutions

Parameter	Region 1	Region2	Both region together
TP	3	1	4
FP	3	0	3
TN	5	6	11
FN	1	2	3
Sensitivity	0.75	0.33	0.57
Specificity	0.63	1.00	0.79
Accuracy	0.67	0.78	0.71
MCC	0.35	0.50	0.36

Dhanda SK et al. A.Development of a strategy and computational application to select candidate protein analogues with reduced HLA binding and immunogenicity. Immunology. 2017Aug 22. PubMed PMID: 28833085.

The issue of intoducing neo-epitope in flanking peptides

Deimmunization strategies must avoid unintended creation of new epitopes in the neighboring sequences

Amino acid substitutions of a 15mer peptide can affect up to two neighboring peptides in the protein sequence

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

Application to Vatreptacog alpha

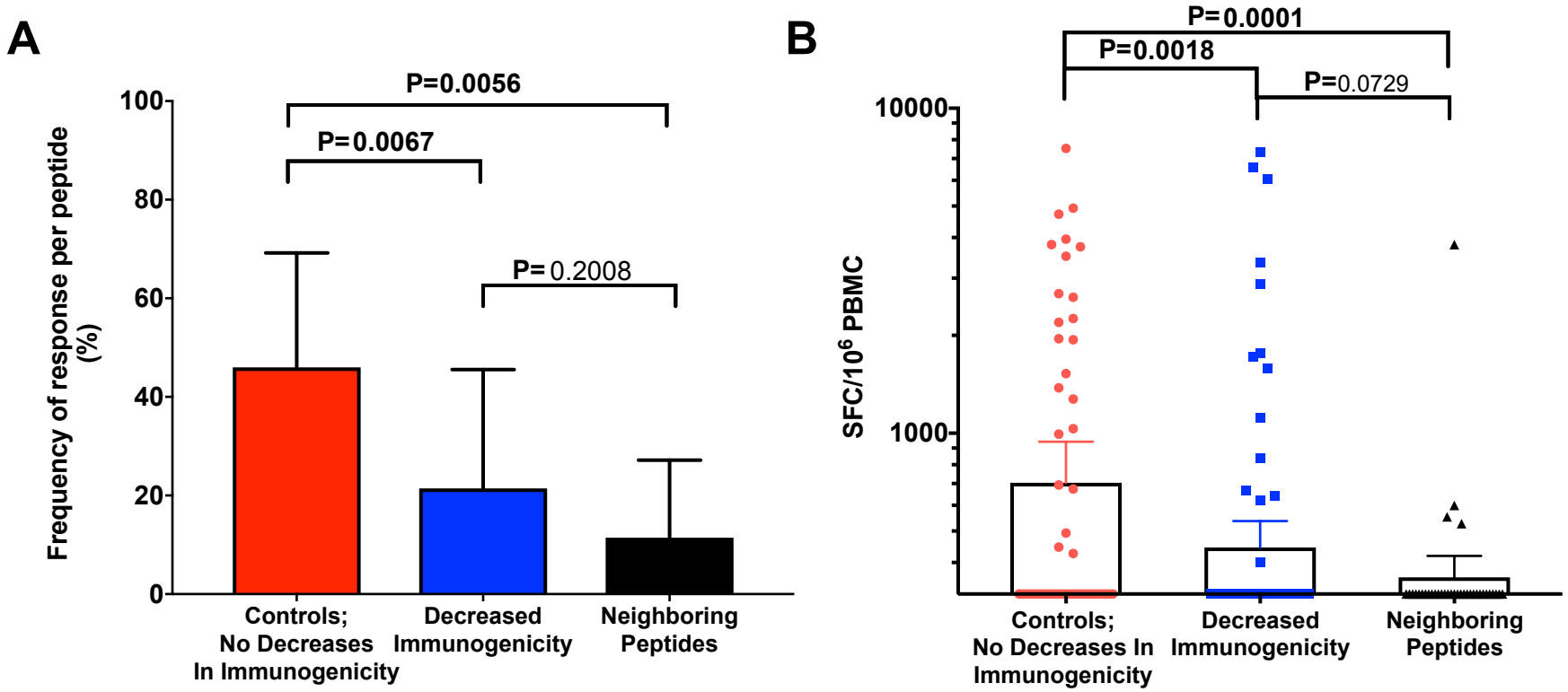
start	end	Median	Mutant pep_seq
166	180	5.25	WQVLLLVNGAQLCGG
181	195	10.34	TLINTIWVVSAAHCF
196	210	8.71	DKIKNWRNLIAVLGE
201	215	14.64	WRNLIAVLGEHDLSE
221	235	18.36	QSRRVAQVIIPSTYV
241	255	6.03	HDIALLRLHQPVVLT
246	260	19.88	LRLHQPVVLTDHVVP
271	285	14.9	RTLAFVRFSLVSGWG
276	290	11.88	VRFSLVSGWGQLLDR
291	305	6.65	GATALVLQVLNVPRL
296	310	15.01	VLQVLNVPRLMTQDC
386	400	13.44	WLQKLMRSEPRPGVL

- A factor VII analog discontinued for immunogenicity problems
- The 7-allele method predicted nine peptides corresponding to six regions
  - Two contiguous peptides were predicted in the 196-215, 241-260 and 291-230 regions
- The region involved in immunogenicity was ranked third out of six regions predicted in the mutant or analog

20 substitutions predicted to have detrimental effect for both peptides

L297A	GATALVAQV	20.31	13.66		24.5	nan		21.77	nan	3
	VAQVLNVPR	24.5	9.49	nan		nan		20.31	21.77	3
L297C	GATALVCQV	18.79	12.14		22.75	nan		26.91	nan	3
	VCQVLNVPR	22.75	7.74	nan		nan		18.79	26.91	3
L297D	GATALVDQV	32.63	25.98		27.1	nan		26.72	nan	3
	VDQVLNVPR	27.1	12.09	nan		nan		32.63	26.72	3
L297E	GATALVEQV	31.58	24.93		26.47	nan		25.78	nan	3
	VEQVLNVPR	26.47	11.46	nan		nan		31.58	25.78	3
L297G	GATALVGQV	29.49	22.84		25.22	nan		22.46	nan	3
	VGQVLNVPR	25.22	10.21	nan		nan		29.49	22.46	3
L297H	GATALVHQV	14.78	8.13		22.33	nan		22.53	nan	3
	VHQVLNVPR	22.33	7.32	nan		nan		14.78	22.53	3
L297P	GATALVPQV	34.15	27.5		28.2	nan		26.45	nan	3
	VPQVLNVPR	28.2	13.19	nan		nan		34.15	26.45	3
L297Q	GATALVQQV	15.71	9.06		24.07	nan		22.78	nan	3
	VQQVLNVPR	24.07	9.06	nan		nan		15.71	22.78	3
L297T	GATALVTQV	19.77	13.12		25.85	nan		22.78	nan	3
	VTQVLNVPR	25.85	10.84	nan		nan		19.77	22.78	3
L300C	GATALVLQV	14.06	7.41		28	nan		22.05	nan	3
	VLQVCNVPR	28	12.99	nan		nan		14.06	22.05	3
L300E	GATALVLQV	14.53	7.88		26.42	nan		21.96	nan	3
	VLQVENVPR	26.42	11.41	nan		nan		14.53	21.96	3
L300G	GATALVLQV	14.84	8.19		31.65	nan		22.78	nan	3
	VLQVGNVPR	31.65	16.64	nan		nan		14.84	22.78	3
L305C	GATALVLQV	19.31	12.66		26.1		71.49	nan	nan	3
	VLQVLNVPR	26.1	11.09		71.49	nan		19.31	nan	3
L305D	GATALVLQV	21.56	14.91		28.07		80.51	nan	nan	3
	VLQVLNVPR	28.07	13.06		80.51	nan		21.56	nan	3
L305E	GATALVLQV	20.41	13.76		26.1		77.68	nan	nan	3
	VLQVLNVPR	26.1	11.09		77.68	nan		20.41	nan	3
L305Q	GATALVLQV	14.81	8.16		22.33		68.57	nan	nan	3
	VLQVLNVPR	22.33	7.32		68.57	nan		14.81	nan	3
Q298D	GATALVLDV	16.56	9.91		22.52	nan		22.52	nan	3
	VLDVLNVPR	22.52	7.51	nan		nan		16.56	22.52	3
R304C	GATALVLQV	14.76	8.11		30.53		47.38	nan	nan	3
	VLQVLNVPC	30.53	15.52		47.38	nan		14.76	nan	3
R304D	GATALVLQV	15.23	8.58		29.25		62.06	nan	nan	3
	VLQVLNVPD	29.25	14.24		62.06	nan		15.23	nan	3
R304E	GATALVLQV	16.34	9.69		29.49		57.87	nan	nan	3
	VLQVLNVPE	29.49	14.48		57.87	nan		16.34	nan	3

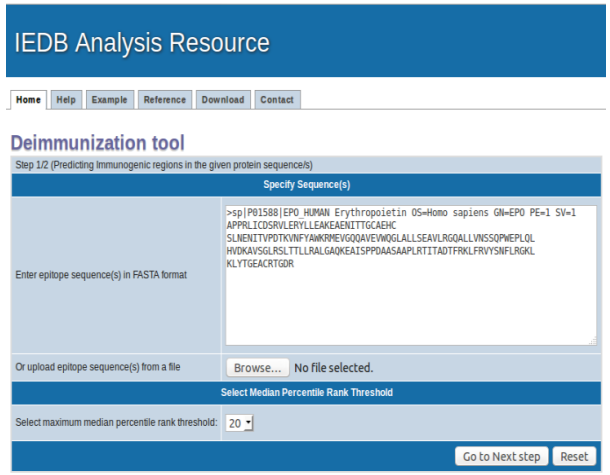
Validation at the level of in vitro immunogenicity



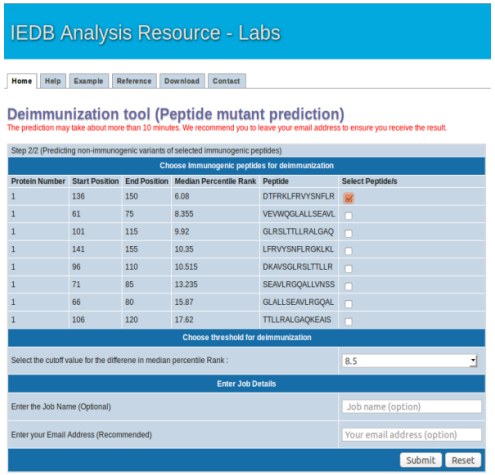
Dhanda SK et al. A.Development of a strategy and computational application to select candidate protein analogues with reduced HLA binding and immunogenicity. Immunology. 2017Aug 22. PubMed PMID: 28833085.

The deimmunization tool is freely available in the IEDB

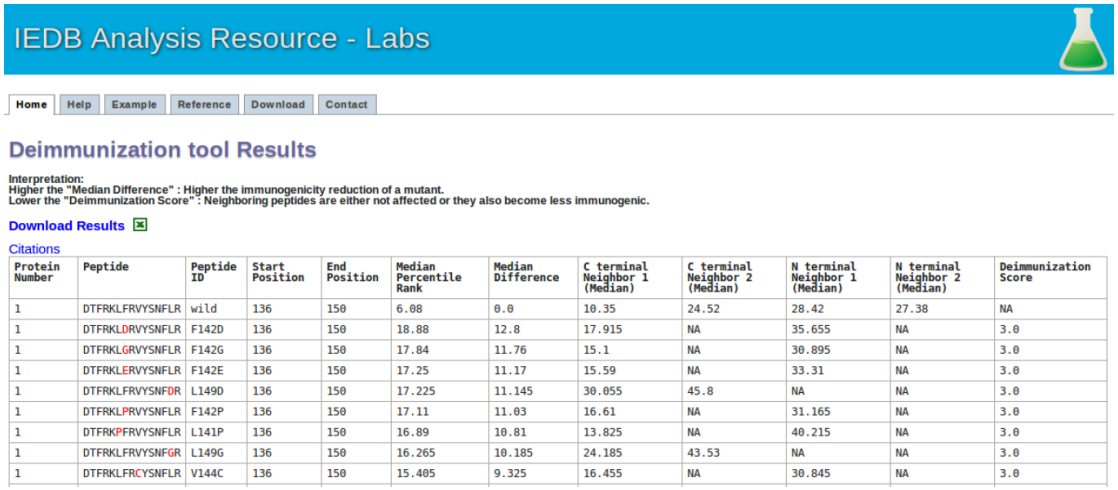
A



B

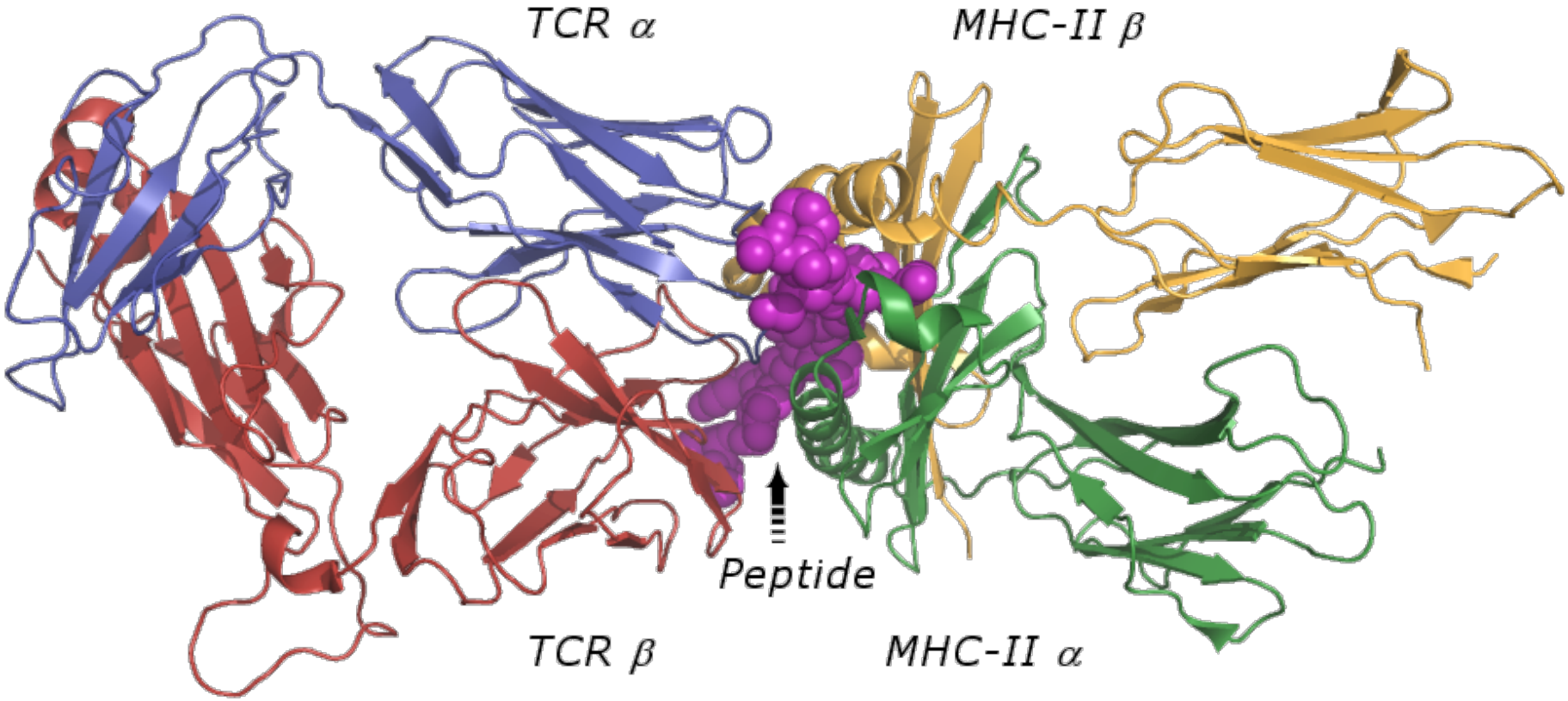


C



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	24.52	28.42	27.38	NA
1	DTFRKLDRVYSNFLR	F142D	136	150	18.88	12.8	17.915	NA	35.655	NA	3.0
1	DTFRKLGRVYSNFLR	F142G	136	150	17.84	11.76	15.1	NA	30.895	NA	3.0
1	DTFRKLERVYSNFLR	F142E	136	150	17.25	11.17	15.59	NA	33.31	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	31.165	NA	3.0
1	DTFRKPFRVYSNFLR	L141P	136	150	16.89	10.81	13.825	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

T cells activation depends on the formation of a trimolecular complex



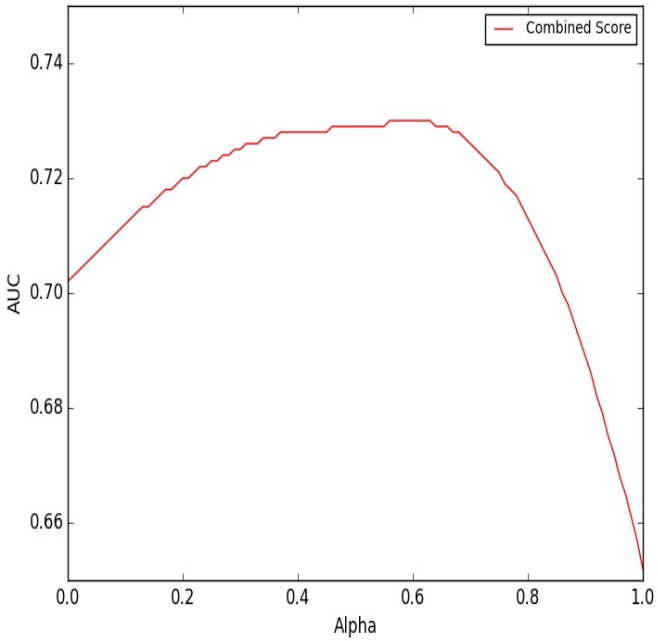
So , what about TCR recognition? Can we predict that???

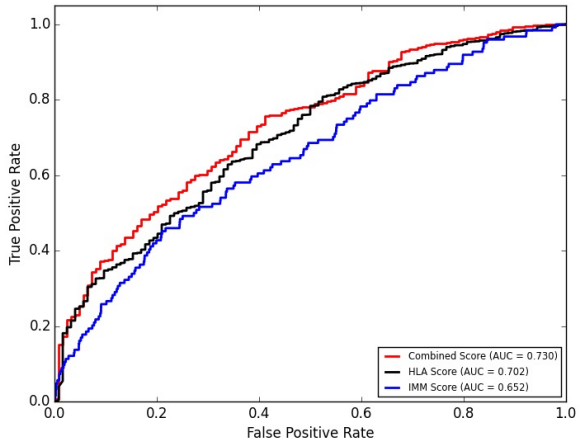
An HLA-binding agnostic approach to trained NN

Table 1a. Initial training data sets

S. No.	Antigen(s)	#donors	Peptide selection method	Reference	#epitopes	#control peptides
1	Mycobacterium tuberculosis	18	Overlapping	(Arlehamn et al., 2012)	65	53
2		28	Predicted	(Lindestam Arlehamn et al., 2013)		1043
3		61	Overlapping	Lindestam Arlehamn et al., 2016)		362
4		61	Confirmed epitopes	Lindestam Arlehamn et al., 2016)		137
5	Timothy Grass	25	Overlapping	(Oseroff et al., 2010)	60	360
6		35	Predicted	(Schulten et al., 2013)		360
7		21	Overlapping	(Westernberg et al., 2016)		6
8		37	Overlapping	(Hinz et al., 2015)		0
9	House Dust Mite	20	Overlapping	(Hinz et al., 2015)	52	6
10	Cockroach	19	Overlapping	(Dillon et al., 2015)	71	521
	Overall				**248**	**2836**

Tested performance on a IEDB-derived tetramer dataset





The Immunogenicity score works

Combination with the 7-allele method further improves performance

Generated a new NN model

Table 1b. Epitope Validation Sets from the Sette Laboratory

S. No.	Antigen (s)	#donors	Peptide Selection method	Reference	#Epitopes	# Control peptides
1	Dengue Polyproteins	150	Predicted	(Weiskopf et al., 2015)	325	140
2	Erythropoietin	5	Overlapping	(Tangri et al., 2005)	9	11
3	CRJ1 and CRJ2	54	Overlapping	(Oseroff et al., 2016)	30	18
4	Mouse allergens	22	Predicted	(Scheulten et al, submitted)	82	885
5	Novel house dust mite antigens	20	Predicted	(Oseroff et al., 2017)	105	186
6	Pertussis Vaccine Antigens	53	Overlapping	(Bancroft et al., 2016)	100	202
7	Ragweed allergens	25	Overlapping	(Pham et al., 2016)	15	183
8	Tetanus	20		(Antunes et al., 2017)	28	98
9	ZIKV polyprotein	18	Overlapping	(Grifoni et al., 2017).	48	529
10	Yellow fever virus polyprotein	42	Overlapping	(Weishkopf et al, unpublished)	42	639
	Overall				**784**	**2891**

- Initial data set
- Tetramer data set
- Additional data sets generated in the meantime

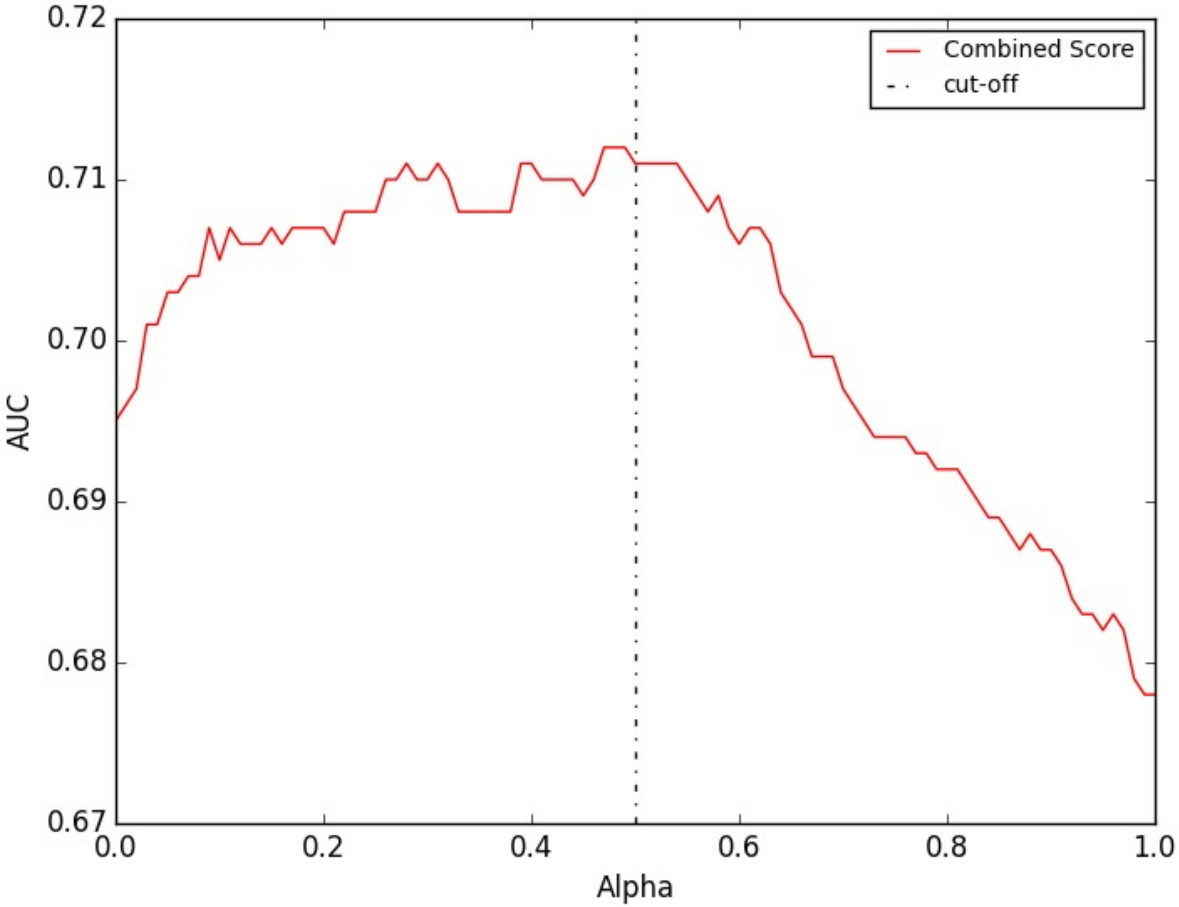
The new NN model was tested on an independent dataset

Table 1c: Validation dataset derived from literature

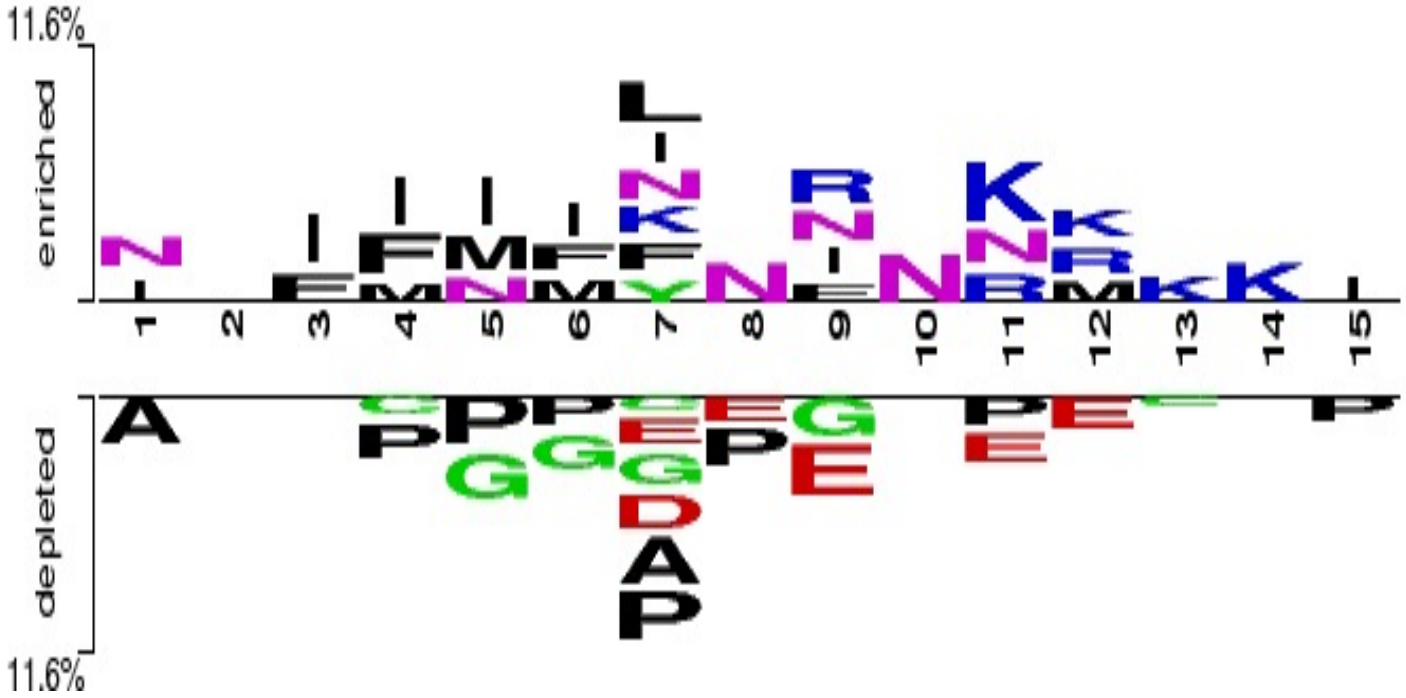
S.No.	Antigen (s)	#Donors	Reference
1	Acetylcholine receptor subunit alpha (UniProt:P02708), Homo sapiens	22	(Manfredi et al., 1992)
2	Circumsporozoite protein, Plasmodinium vivax and Plasmodinium falciparum	22	(Herrera et al., 1992)
3	Conserved hypothetical lipoprotein (UniProt:Q5NGE4), Francisella tularensis	10	(Sjostedt et al., 1990)
4	Other Plasmodium falciparum (malaria parasite P. falciparum) protein	12	(Rzepczyk et al., 1990)
5	Circumsporozoite (CS) protein, Plasmodinium falciparum	64	(Zevering et al., 1990)
6	Circumsporozoite (CS) protein, Plasmodinium falciparum	35	(Good et al., 1988)
7	Api m 1, Apis mellifera	40	(Carballido et al., 1993)
8	Myelin basic protein (UniProt:P02686), Homo sapiens	12	(Salvetti et al., 1993)
9	Circumsporozoite protein, putative, Plasmodinium vivax	52	(Bilsborough et al., 1993)
10	Acetylcholine receptor subunits gamma and delta, Homo sapiens	22	(Manfredi et al., 1993)
11	Acetylcholine receptor subunit alpha (UniProt:P02708), Homo sapiens	22	(Moiola et al., 1993)
12	Glutamate decarboxylase 2, Homo sapiens	44	(Atkinson et al., 1994)
13	(Chaye et al., 1993) Structural polyprotein, Rubella virus	10	(Chaye et al., 1993)
14	Envelope glycoprotein D, Human herpesvirus 1	24	(Damhof et al., 1993)
15	Thyroglobulin and Thyrotropin receptor, Homo sapiens	15	(Kellermann et al., 1995)
16	Fusion glycoprotein F0, Measles morbillivirus	13	(Muller et al., 1996)
17	Poa p 5, Poa pratensis (Kentucky bluegrass)	13	(Zhang et al., 1996)
18	Myelin basic protein (UniProt:P02686), Homo sapiens	20	(Pender et al., 1996)
19	Structural polyprotein, Rubella virus	14	(Marttila et al., 1996)
20	Acetylcholine receptor subunits delta and alpha, Homo sapiens	58	(Wang et al., 1997)
21	Hev b 1, Hevea brasiliensis (rubbertree)	19	(Raulf-Heimsoth et al., 199
22	Api m 1, Apis mellifera	10	(Kammerer et al., 1997)
23	Thrombospondin-related anonymous protein, TRAP, Plasmodinium falciparum	50	(Flanagan et al., 1999)
24	Nucleoprotein, Measles morbillivirus	19	(Marttila et al., 1999)
25	Genome polyprotein, Hepatitis C virus	22	(Lamonaca et al., 1999)
26	Subtilisin-like protease 6, Trichophyton rubrum	38	(Woodfolk et al., 2000)
27	Blood groups Rh(D) and Rh(CE) polypeptides, Homo sapiens	22	(Stott et al., 2000)
28	Myelin proteolipid protein and Myelin basic protein, Homo sapiens	16	(Tejada-Simon et al., 2001
29	Genome polyprotein, Enterovirus B; Glutamate decarboxylase 2, Homo sapiens	22	(Marttila et al., 2001)
30	Gal d 1, Gallus gallus	14	(Holen et al., 2001)
31	Genome polyprotein, Hepatitis C virus	10	(Wertheimer et al., 2003)
32	Hev b 6, Hevea brasiliensis (rubbertree)	16	(de Silva et al., 2004)
33	Bos d 9, Bos taurus (bovine)	10	(Elsayed et al., 2004)
34	Cha o 1, Chamaecyparis obtusa	19	(Sone et al., 2005)
35	Genome polyprotein, Hepatitis C virus	22	(Schulze zur Wiesch et al., 20
36	Genome polyprotein, Hepatitis C virus	41	(Sarobe et al., 2006)
37	Bos d 9, Bos taurus (bovine)	29	(Ruiter et al., 2006)
38	Cytochrome P450 2D6, Homo sapiens	80	(Ma et al., 2006)
39	Capsid protein VP1, Human parvovirus	19	(Kasprowicz et al., 2006)
40	Integrin beta-3, Homo sapiens	31	(Sukati et al., 2007)
41	Genome polyprotein, Hepatitis C virus	44	(Schulze Zur Wiesch et al., 20
42	Equ c 1, Equus caballus	10	(Immonen et al., 2007)
43	Merozoite surface protein 1, Plasmodium falciparum	48	(Malhotra et al., 2008)
44	Cry j 1, Cryptomeria japonica	12	(Masuyama et al., 2009)
45	Cha o 2, Chamaecyparis obtusa	19	(Sone et al., 2009)
46	Capsid protein VP1, Adeno-associated dependoparvovirus A and Adeno-associated virus	16	(Madsen et al., 2009)
47	Non-specific lipid-transfer protein, Prunus persica (peach)	15	(Pastorello et al., 2010)
48	Aquaporin-4, Homo sapiens	32	(Matsuya et al., 2011)
49	UniProt:B8ZU53, Mycobacterium leprae	152	(Chaduvula et al., 2012)
50	Pas n 1 allergen, Paspalum notatum	18	(Etto et al., 2012)
51	Pen a 1 allergen, Farfantepenaeus aztecus	16	(Ravkov et al., 2013)
52	Genome polyprotein, Tick-borne encephalitis virus	47	(Schwaiger et al., 2014)
53	Other Canis lupus (wolf or dog) protein	25	(Ronka et al., 2015)
54	Can f 5, Canis lupus	24	(Kailaanmaki et al., 2016)
55	Botulinum neurotoxin type A, Clostridium botulinum	25	(Oshima et al., 2016a)
56	Genome polyprotein, Rhinovirus A and C	20	(Gaido et al., 2016)
57	Botulinum neurotoxin type A, Clostridium botulinum	14	(Oshima et al., 2016b)

- 57 different studies mined from the IEDB
- Tested sets of overlapping peptides spanning entire proteins
- Manually recurated
- Overall 530 positives and 1759 negatives

The Immunogenicity score works and combination with the 7-allele method further improves performance



A Class II immunogenicity motif



Conclusions.I

- Alanine substitution not a feasible approach to modulate HLA binding
- We have developed a computational tool to select analogs for deimmunization
- Experimental validation in progress

Acknowledgments

- Bjoern Peters
- Sandeep Kumar
- John Sidney
- Sinu Paul

Conclusions. II

- Immunogenicity prediction of at the population level is the most frequent application (HLA types unavailable)
- Improvement by combined HLA binding + immunogenicity
- The increase is relatively small
  - HLA binding is the dominant in shaping the T cell repertoire
  - Possible coordinate evolution
- Small aminoacids avoided in the middle of epitopes, while longer side chains are overrepresented
- Validation over different methodologies, ethnicities, and infectious diseases, allergy and autoimmunity
- Tool will be available on the IEDB website

Acknowledgments

- Bjoern Peters
- Sandeep Kumar Dhanda
- Edita Karosiene
- Lindy Edwards
- Alba Grifoni
- Morten Nielsen