

# Module 13 Assignment

## 585.751.81 Immunoengineering

1. Read the following paper on models of T cell activation:  
<https://www.nature.com/articles/nri3728> at least up to “Extensions of phenotypic models” on page 623. Describe each of the five possible models for T cell activation listed in the paper (in 2-3 sentences each). Which model best describes T cell activation and why? (50 points)  
The paper describes five models for T cell activation:
  - **Occupancy model:** also known as affinity model, proposes that T cell activation is proportional to the number of T-cell receptors (TCRs) occupied by peptide-MHC (pMHC) complexes. This model states that TCRs become signaling-competent immediately upon pMHC binding. This model predicts that pMHC potency ( $EC_{50}$ ) correlates directly with the dissociation time and the TCR-pMHC dissociation constant ( $K_d$ ) and that the maximum response ( $E_{max}$ ) is independent from the binding parameters which contradicts experimental data.
  - **Kinetic proofreading model** explains how T cells discriminate between ligands based on the dissociation time of the ligand-receptor interaction. It proposes that T cell activation is proportional to the fraction of TCRs that remain bound by pMHC complexes long enough to undergo biochemical modifications, like tyrosine phosphorylation, to reach a signaling-competent state. The delay between pMHC binding and TCR signaling, enables T cells to discern between pMHC complexes based on their dissociation times from the TCR, with the prediction that longer binding times lead to greater T cell activation, which is supported by observed correlations between  $EC_{50}$  and  $K_d$ .
  - **Kinetic proofreading with limited signaling model:** both the kinetic proofreading model and serial triggering models start with the same biochemical assumptions but differ in what triggers T cell activation. In serial triggering models, each TCR can signal only once per pMHC binding event, limiting continuous activation, particularly from pMHC complexes that dissociate slowly. The kinetic proofreading with limited signaling suggests that TCRs can only signal for a limited time after they become active, necessitating multiple pMHCs to bind sequentially for sustained T cell activation. Because of this limited time for signaling, even at high concentrations of pMHCs, each complex can only activate TCRs for a short period. This results in an optimal dissociation time for effective T cell activation, where pMHCs with too long dissociation times fail to sustain activation because they remain bound to non-signaling TCRs.
  - **Kinetic proofreading with sustained signaling model** allows signaling-competent TCRs to sustain signaling even after pMHC unbinding. This sustained signaling capability is supported by experimental evidence suggesting that TCRs, along with their associated signaling complexes, can remain active until they are either dephosphorylated by phosphatases or internalized by the cell. This model modifies the dynamics of T cell activation by allowing pMHC complexes with various dissociation times to produce maximal signaling at high concentrations, without the need for serial binding, thus allowing an optimal dissociation time that varies with pMHC concentration.
  - **Kinetic proofreading with negative feedback model** is an extension of the kinetic proofreading model and introduces a mechanism to regulate the activation of TCRs through negative feedback loops involving phosphorylation. This negative feedback is



mediated by phosphatases like SHP1 and other phosphatases. Such feedback results in T cell activation showing an optimal response function to the pMHC dose, which is modulated by the dissociation time between TCR and pMHC.

2. You are studying HIV and want to identify viral epitopes that may be recognized by CD8+ T cells to kill HIV-infected CD4+ T cells. Use the following database ([http://www.iedb.org/home\\_v3.php](http://www.iedb.org/home_v3.php)) to search for linear peptide epitopes from the organism Human immunodeficiency virus 1 (the more common type of the HIV virus) that bind to HLA-A\*02:01 (the most common HLA-A allele in humans) in human hosts. Once you submit the search, change the linear peptide length to 9 amino acids (as most HLA molecules have a strong preference for binding 9mers) on the left-hand side. Export your results, pick the first 50 epitopes in your search results and input them into the netMHC artificial neural network prediction program using the PEPTIDE format (<https://services.healthtech.dtu.dk/service.php?NetMHC-4.0>). For this problem, look at the binding affinity of your peptides to HLA-A\*02:01 allele in the HLA-A species/loci. (50 points)
  - a. Please list/provide a screenshot of the peptides that you tested in the software and list the peptides that were predicted to be strong binders to the HLA molecule (those labeled “SB”). (15 points)

IEDB ID	Epitope	Antigen	Organism	# References	# Assays
1955	AIIRILQQL	Protein Vpr	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	6	6
2644	ALIRILQQL	Protein Vpr	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	2	2
2835	ALQDSGLEV	Gag-Pol polyprotein	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	1	1
2954	ALVEICTEM	Gag-Pol polyprotein	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	2	2
125703	ALVEMGHIV	Protein Vpu	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	1	1
5091	ATPQDLNTM	Gag polyprotein	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	1	1
5295	AVDLSHFLK	Protein Nef	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	1	1
125823	DIKDTKEAL	Gag polyprotein	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	2	2
125830	DLADQLIHL	Virion infectivity factor	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	1	1
9663	DPKVKQWPL	Gag-Pol polyprotein	Human immunodeficiency virus 1 (human immunodeficiency virus 1 HIV-1)	1	1

IEDB search for linear peptides of length 9mers from Human immunodeficiency virus 1 that bind to HLA-A\*02:01

# NetMHC version 4.0

# Input is in PEPTIDE format

# Rank Threshold for Strong binding peptides 0.500

# Rank Threshold for Weak binding peptides 2.000

pos	HLA	peptide	Core	Offset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0201	AIIRILQQL	AIIRILQQL	0	0	0	0	0	AIIRILQQL	PEPLIST	0.383	793.35	3.50	
0	HLA-A0201	ALIRILQQL	ALIRILQQL	0	0	0	0	0	ALIRILQQL	PEPLIST	0.591	83.87	0.90	<= WB
0	HLA-A0201	ALQDSGLEV	ALQDSGLEV	0	0	0	0	0	ALQDSGLEV	PEPLIST	0.686	29.75	0.40	<= SB
0	HLA-A0201	ALVEICTEM	ALVEICTEM	0	0	0	0	0	ALVEICTEM	PEPLIST	0.637	50.60	0.60	<= WB
0	HLA-A0201	ATPQDLNTM	ATPQDLNTM	0	0	0	0	0	ATPQDLNTM	PEPLIST	0.113	14771.97	21.00	
0	HLA-A0201	AVDSLHFLK	AVDSLHFLK	0	0	0	0	0	AVDSLHFLK	PEPLIST	0.071	23295.88	33.00	
0	HLA-A0201	DPKVKQWPL	DPKVKQWPL	0	0	0	0	0	DPKVKQWPL	PEPLIST	0.034	34473.86	60.00	
0	HLA-A0201	DPNPQEVVL	DPNPQEVVL	0	0	0	0	0	DPNPQEVVL	PEPLIST	0.027	37415.11	75.00	
0	HLA-A0201	ELHPDKWTV	ELHPDKWTV	0	0	0	0	0	ELHPDKWTV	PEPLIST	0.387	758.61	3.50	
0	HLA-A0201	ELRSLYNTV	ELRSLYNTV	0	0	0	0	0	ELRSLYNTV	PEPLIST	0.187	6589.51	12.00	
0	HLA-A0201	EPIVGAETF	EPIVGAETF	0	0	0	0	0	EPIVGAETF	PEPLIST	0.046	30464.55	48.00	
0	HLA-A0201	EVIPMFSAI	EVIPMFSAI	0	0	0	0	0	EVIPMFSAI	PEPLIST	0.163	8567.42	14.00	
0	HLA-A0201	FIMIVGGIV	FIMIVGGIV	0	0	0	0	0	FIMIVGGIV	PEPLIST	0.558	129.53	1.30	<= WB
0	HLA-A0201	FPVVRQVPL	FPVVRQVPL	0	0	0	0	0	FPVVRQVPL	PEPLIST	0.088	19342.74	26.00	
0	HLA-A0201	FPVTPQVPL	FPVTPQVPL	0	0	0	0	0	FPVTPQVPL	PEPLIST	0.110	15206.25	21.00	
0	HLA-A0201	GPCHKARVL	GPCHKARVL	0	0	0	0	0	GPCHKARVL	PEPLIST	0.024	38633.61	80.00	
0	HLA-A0201	GSEELRSLY	GSEELRSLY	0	0	0	0	0	GSEELRSLY	PEPLIST	0.037	33626.54	60.00	
0	HLA-A0201	HLEGGKIVL	HLEGGKIVL	0	0	0	0	0	HLEGGKIVL	PEPLIST	0.490	248.06	1.90	<= WB
0	HLA-A0201	HPDIVIYQY	HPDIVIYQY	0	0	0	0	0	HPDIVIYQY	PEPLIST	0.068	23882.63	34.00	
0	HLA-A0201	IISLWDQSL	IISLWDQSL	0	0	0	0	0	IISLWDQSL	PEPLIST	0.347	1169.08	4.50	
0	HLA-A0201	ILKEPVHGV	ILKEPVHGV	0	0	0	0	0	ILKEPVHGV	PEPLIST	0.574	99.95	1.10	<= WB
0	HLA-A0201	ILLEPVHGV	ILLEPVHGV	0	0	0	0	0	ILLEPVHGV	PEPLIST	0.852	4.95	0.03	<= SB
0	HLA-A0201	IPLTEEAEL	IPLTEEAEL	0	0	0	0	0	IPLTEEAEL	PEPLIST	0.054	27961.78	42.00	
0	HLA-A0201	IPRRIRQGL	IPRRIRQGL	0	0	0	0	0	IPRRIRQGL	PEPLIST	0.035	34355.45	60.00	
0	HLA-A0201	IVGAETFYV	IVGAETFYV	0	0	0	0	0	IVGAETFYV	PEPLIST	0.704	24.61	0.40	<= SB
0	HLA-A0201	KAACWAGI	KAACWAGI	0	0	0	0	0	KAACWAGI	PEPLIST	0.370	911.12	4.00	
0	HLA-A0201	KLTPCLCTL	KLTPCLCTL	0	0	0	0	0	KLTPCLCTL	PEPLIST	0.634	52.70	0.70	<= WB
0	HLA-A0201	KLVGKLNWA	KLVGKLNWA	0	0	0	0	0	KLVGKLNWA	PEPLIST	0.563	112.62	1.10	<= WB
0	HLA-A0201	LLNATDIAV	LLNATDIAV	0	0	0	0	0	LLNATDIAV	PEPLIST	0.677	32.88	0.50	<= SB
0	HLA-A0201	LLQLTVWGI	LLQLTVWGI	0	0	0	0	0	LLQLTVWGI	PEPLIST	0.605	71.44	0.80	<= WB
0	HLA-A0201	LLQYWSQEL	LLQYWSQEL	0	0	0	0	0	LLQYWSQEL	PEPLIST	0.648	45.03	0.60	<= WB
0	HLA-A0201	LLWKGEHAV	LLWKGEHAV	0	0	0	0	0	LLWKGEHAV	PEPLIST	0.605	71.44	0.80	<= WB
0	HLA-A0201	LTFGWCFKL	LTFGWCFKL	0	0	0	0	0	LTFGWCFKL	PEPLIST	0.621	60.45	0.70	<= WB
0	HLA-A0201	LVGPTPVNI	LVGPTPVNI	0	0	0	0	0	LVGPTPVNI	PEPLIST	0.225	4389.37	9.00	
0	HLA-A0201	MTNPPPIPV	MTNPPPIPV	0	0	0	0	0	MTNPPPIPV	PEPLIST	0.485	262.89	1.90	<= WB
0	HLA-A0201	NANPDCKTI	NANPDCKTI	0	0	0	0	0	NANPDCKTI	PEPLIST	0.059	26303.06	38.00	
0	HLA-A0201	NPDIVYQY	NPDIVYQY	0	0	0	0	0	NPDIVYQY	PEPLIST	0.066	24360.48	34.00	
0	HLA-A0201	NSSKVSQNY	NSSKVSQNY	0	0	0	0	0	NSSKVSQNY	PEPLIST	0.036	33741.71	60.00	
0	HLA-A0201	PPIPVGDIY	PPIPVGDIY	0	0	0	0	0	PPIPVGDIY	PEPLIST	0.033	35127.12	65.00	
0	HLA-A0201	RATEAQOHL	RATEAQOHL	0	0	0	0	0	RATEAQOHL	PEPLIST	0.176	7430.63	13.00	
0	HLA-A0201	RAMASDFNL	RAMASDFNL	0	0	0	0	0	RAMASDFNL	PEPLIST	0.444	408.75	2.50	
0	HLA-A0201	RIKQIINMW	RIKQIINMW	0	0	0	0	0	RIKQIINMW	PEPLIST	0.073	22633.75	31.00	
0	HLA-A0201	RILQQLFI	RILQQLFI	0	0	0	0	0	RILQQLFI	PEPLIST	0.543	140.54	1.30	<= WB
0	HLA-A0201	RLVNGSLAL	RLVNGSLAL	0	0	0	0	0	RLVNGSLAL	PEPLIST	0.609	68.79	0.80	<= WB
0	HLA-A0201	RMYSPIISL	RMYSPIISL	0	0	0	0	0	RMYSPIISL	PEPLIST	0.567	108.49	1.10	<= WB
0	HLA-A0201	RPIVSTQLL	RPIVSTQLL	0	0	0	0	0	RPIVSTQLL	PEPLIST	0.081	20814.80	29.00	
0	HLA-A0201	RPMTYKAAL	RPMTYKAAL	0	0	0	0	0	RPMTYKAAL	PEPLIST	0.071	23270.69	32.00	
0	HLA-A0201	RVIEVLQRA	RVIEVLQRA	0	0	0	0	0	RVIEVLQRA	PEPLIST	0.401	650.38	3.50	
0	HLA-A0201	SLYNTIAVL	SLYNTIAVL	0	0	0	0	0	SLYNTIAVL	PEPLIST	0.637	50.52	0.60	<= WB
0	HLA-A0201	SLYNTVATL	SLYNTVATL	0	0	0	0	0	SLYNTVATL	PEPLIST	0.632	53.76	0.70	<= WB

## netMHC 50 first binding affinity results

- b. Are there any features in common between the peptides listed as “strong binders”? If so, what are those features and why are they conserved between the peptides? If you do not see any features in common or only have 1-2 strong binding peptides, answer the question more generally: what common features would you expect to see between peptides that bind strongly to a given HLA/MHC allele? (20 points)

When examining the epitopes predicted by netMHC to have a strong binding affinity to HLA-A02:01, we observe that they possess low %Rank scores. This observation is in concordance with the definition provided for the %Rank: it is a comparative metric that ranks the predicted affinity of the epitope against a reference set comprising 400,000 random natural peptides. In this context, epitopes characterized by a %Rank score of less than 0.5 are classified as strong binders, whereas those with a %Rank score of less than 2 are considered weak binders.

When including the complete list of epitopes (and not only the first 50) obtained from our prior IEDB query in netMHC and look into the strong binders returned by the neural network, we can identify a pattern of specific amino acids recurring at identical positions across the 9-mer sequences. Below, using the color-coded visualizations of duplicate residues within each position, we can see that a high frequency of leucine (L) at positions 2 and 9, followed by

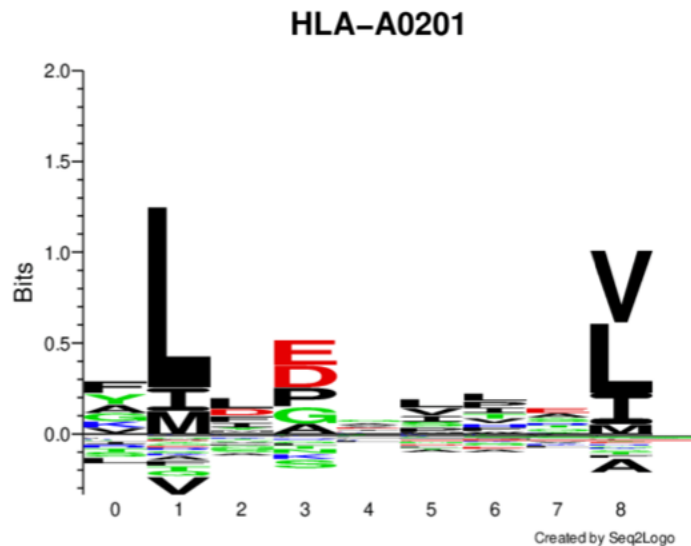


valine (V) at position 9. Research by Vadim Karnaukov et al., support these findings (Fig. 1). Leucine and valine are hydrophobic amino acids, which may suggest the importance of hydrophobic interactions in the binding process. This also, indicates that these positions are key features for strong binding to HLA-A\*02:01, possibly because they are the residues that fit into the binding groove of the HLA molecule. These features are conserved because they contribute to the stability and specificity of the peptide-HLA complex, the peptide needs to bind strongly to the HLA molecule and present its key residues to the T cell receptor in an optimal orientation to trigger an effective immune response.

Color-coded duplicates within each column

Peptide	1	2	3	4	5	6	7	8	9
ALQDSGLEV	A	L	Q	D	S	G	L	E	V
ILLEPVHGV	I	L	L	E	P	V	H	G	V
IVGAETFYV	I	V	G	A	E	T	F	Y	V
LLNATDIAV	L	L	N	A	T	D	I	A	V
TLTSCNTSV	T	L	T	S	C	N	T	S	V
VLAEAMSQV	V	L	A	E	A	M	S	Q	V
WLWYIKIFI	W	L	W	Y	I	K	I	F	I
YTAFTIPSI	Y	T	A	F	T	I	P	S	I
ALVEMGHHV	A	L	V	E	M	G	H	H	V
GLADQLIHI	G	L	A	D	Q	L	I	H	I
GLADQLIHL	G	L	A	D	Q	L	I	H	L
GLADQLIHM	G	L	A	D	Q	L	I	H	M
NLADQLIHL	N	L	A	D	Q	L	I	H	L
SLADQLIHL	S	L	A	D	Q	L	I	H	L
SLVKHHMYV	S	L	V	K	H	H	M	Y	V
VLYCVHQRV	V	L	Y	C	V	H	Q	R	V
SLFNTVATL	S	L	F	N	T	V	A	T	L
SLFNAVATL	S	L	F	N	A	V	A	T	L
SLFNTIATL	S	L	F	N	T	I	A	T	L
SLFNTVATV	S	L	F	N	T	V	A	T	V
SLFNTVVTL	S	L	F	N	T	V	V	T	L
SLFNAVAVL	S	L	F	N	A	V	A	V	L
SLFNAVVTL	S	L	F	N	A	V	V	T	L
SLFNTIAVL	S	L	F	N	T	I	A	V	L
SLYNAIATL	S	L	Y	N	A	I	A	T	L
SLYNAVATL	S	L	Y	N	A	V	A	T	L
SLYNAVVTL	S	L	Y	N	A	V	V	T	L
SLYNSVATL	S	L	Y	N	S	V	A	T	L
SLYNTIATL	S	L	Y	N	T	I	A	T	L
SLYNTVVTL	S	L	Y	N	T	V	V	T	L
YTAFTIPSV	Y	T	A	F	T	I	I	S	V
IIIGALVGV	I	I	I	G	A	L	V	G	V





**Figure 1 - Sequence Logo – Ref: [1]**

- c. Describe (in no more than a few sentences) how netMHC could be utilized in an immunoengineering context. (15 points)

Epitope identification is costly and time-consuming as it requires experimental screening of large arrays of potential epitope candidates. Without designing such assays, NetMHC could be utilized to predict epitope binding and understand T cell activation. Mapping the binding sites (or epitopes) of antibodies and their target antigens is critical for understanding their motion and dynamics in response to binding and provides additional insight to advance therapeutic candidates, maximize efficacy, and reduce adverse immune reactions.

[1] V. Karnaukhov *et al.*, “HLA binding of self-peptides is biased towards proteins with specific molecular functions,” *bioRxiv*, p. 2021.02.16.431395, 2021, doi: 10.1101/2021.02.16.431395