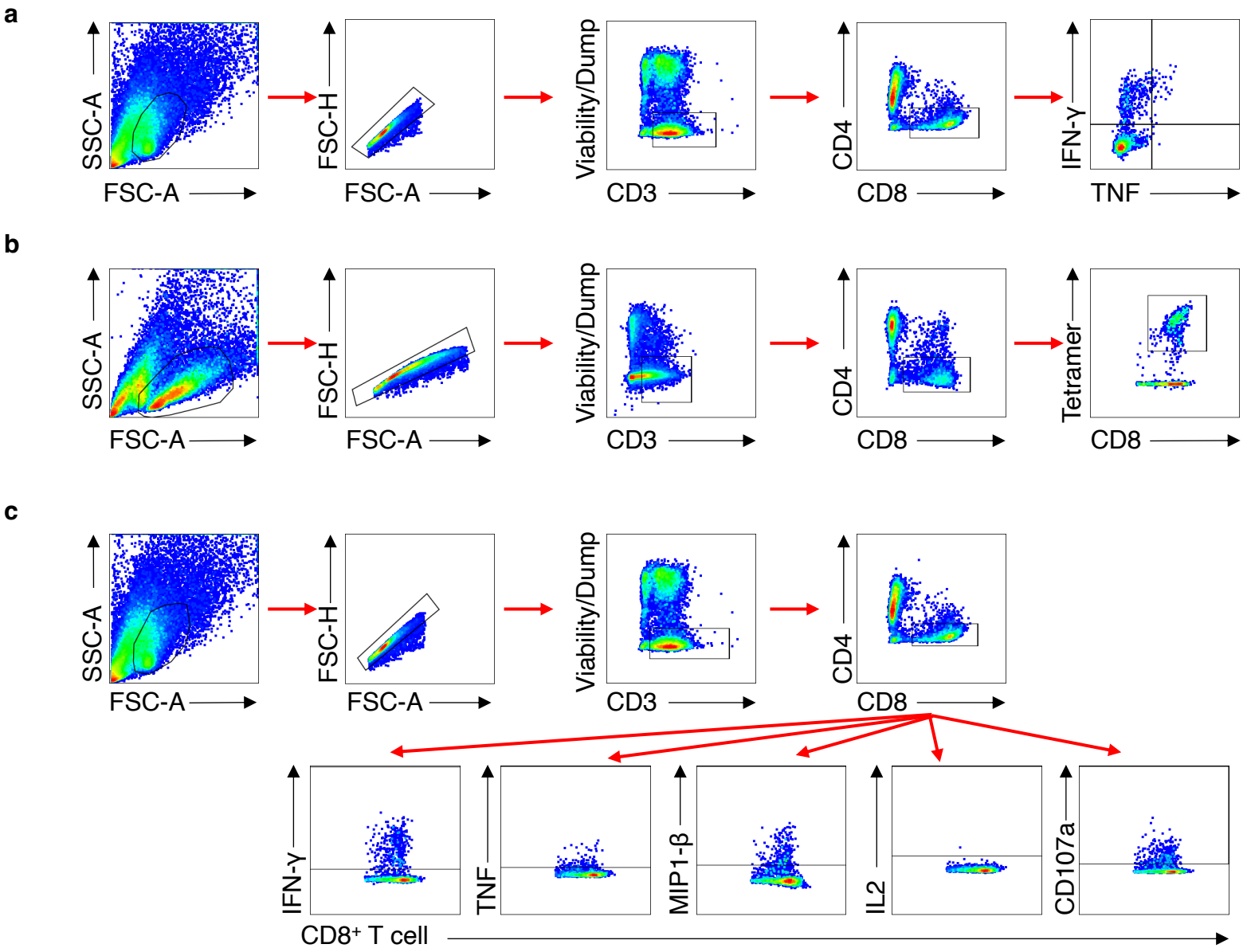
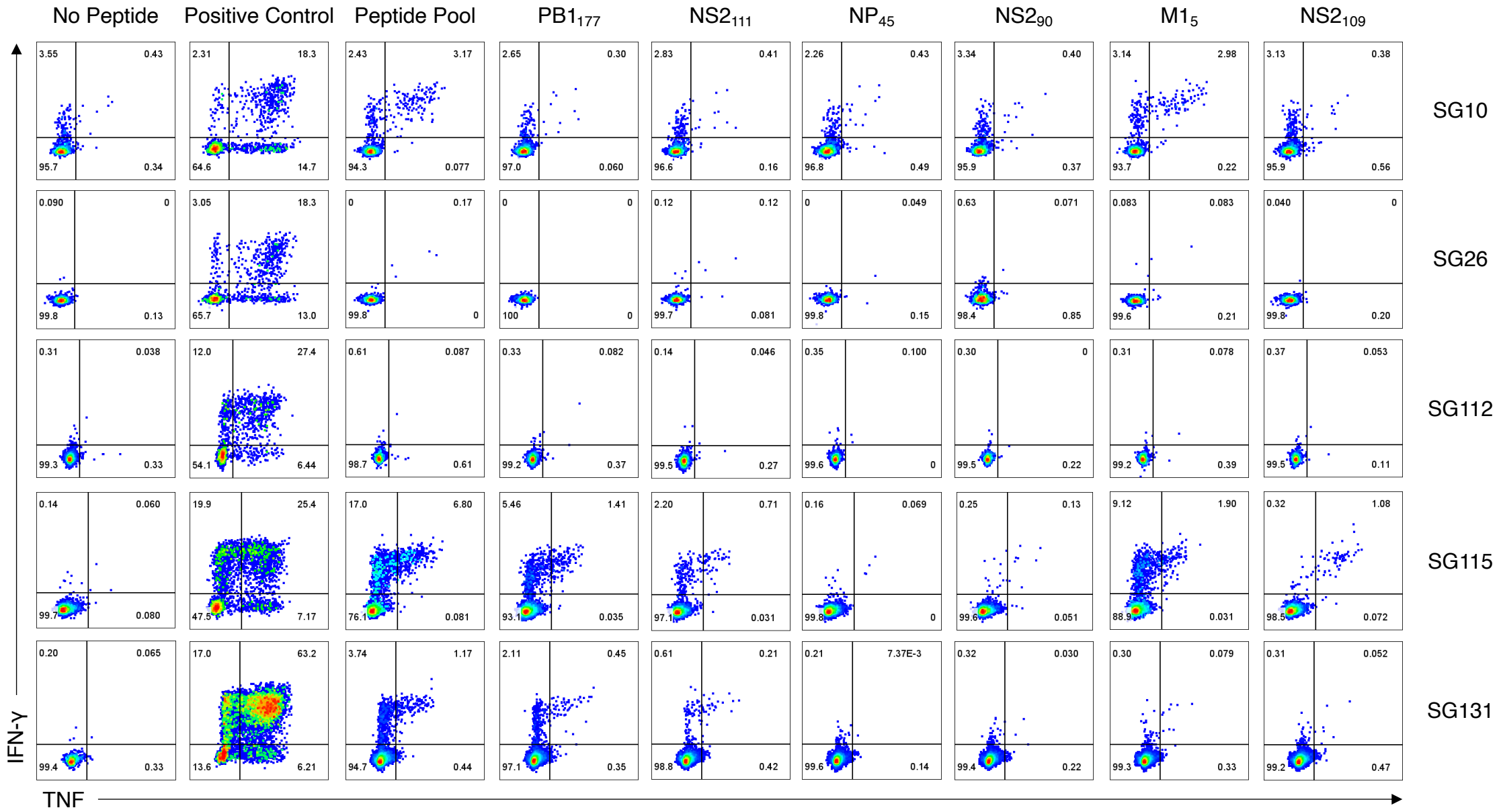
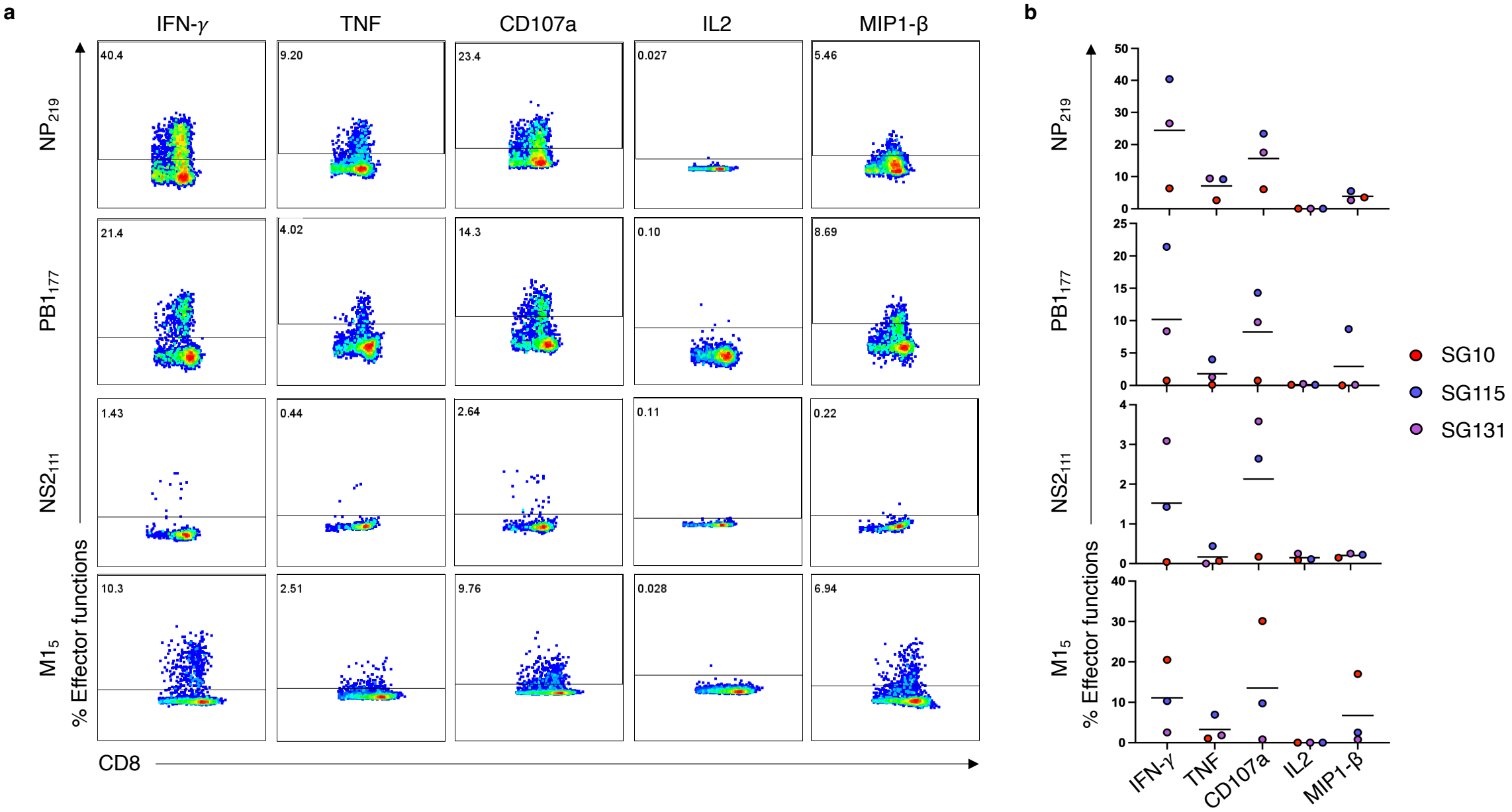


Supplementary Figure 1. Gating Strategy. Gating strategy used for all data in this manuscript. **(a)** Gating strategy for the analysis of specificity of $CD8^+$ T cell lines for Figures 1-2. Cells are gated on lymphocytes, singlets, Live $CD3^{mid\text{-}high}$, $CD8^{mid\text{-}high}$ T cells and IFNγ vs TNF, with the top two quadrants added to report the percentage of $IFN\gamma^+$ of $CD8^+$ T cells. **(b)** Gating strategy for the specificity of $CD8^+$ T cell lines for Figure 2. Cells are gated on lymphocytes, singlets, Live $CD3^{mid\text{-}high}$, $CD8^{mid\text{-}high}$ T cells and $tetramer^+CD8^+$ T cells **(c)** Gating strategy for the analysis of polyfunctionality of $CD8^+$ T cell lines for Figures 3. Cells are gated on lymphocytes, singlets, Live $CD3^{mid\text{-}high}$, $CD8^{mid\text{-}high}$ T cells and IFN-γ, TNF, MIP1-β, IL2 or CD107a which were subsequently analysed with Boolean gating.



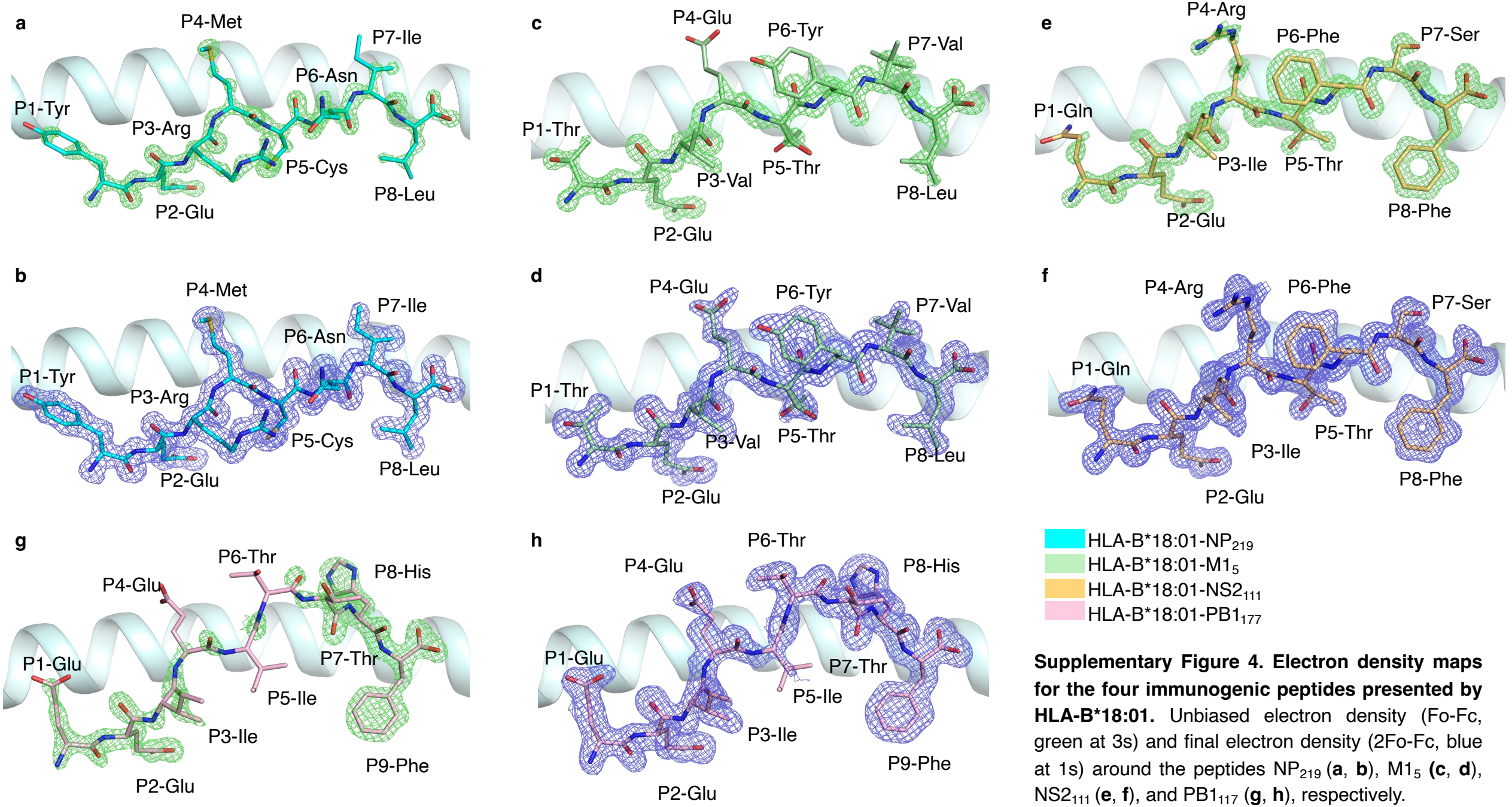
Supplementary Figure 2. Summary of all intracellular cytokine staining from the initial peptide pool screening.

PBMCs of five HLA-B*18:01^{+} donors (n=5) were stimulated with 2 mM per peptide pool condition ($PB1_{177}$, $NS2_{111}$, NP_{45}, $NS2_{90}$, $M1_{5}$, and $NS2_{109}$) for a period of 10^{+} days. Intracellular cytokine staining was completed after 10^{+} days by restimulating cells individually at 10 mM and 2mM/peptide for peptide pool conditions. All cell line conditions incorporated a x500 positive control and a no peptide negative control.

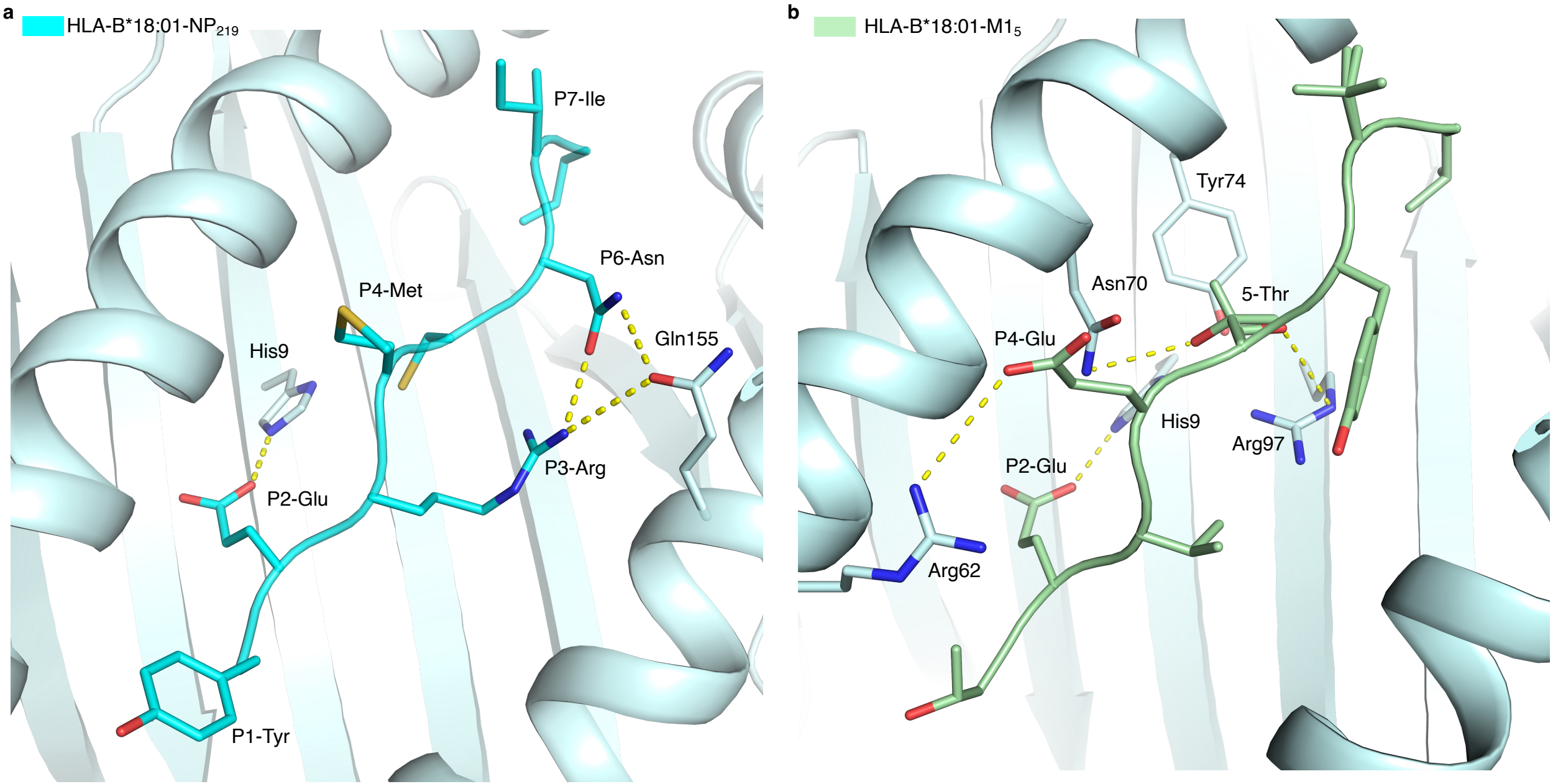


Supplementary Figure 3. Effector function of peptide-specific CD8⁺ T cell lines.

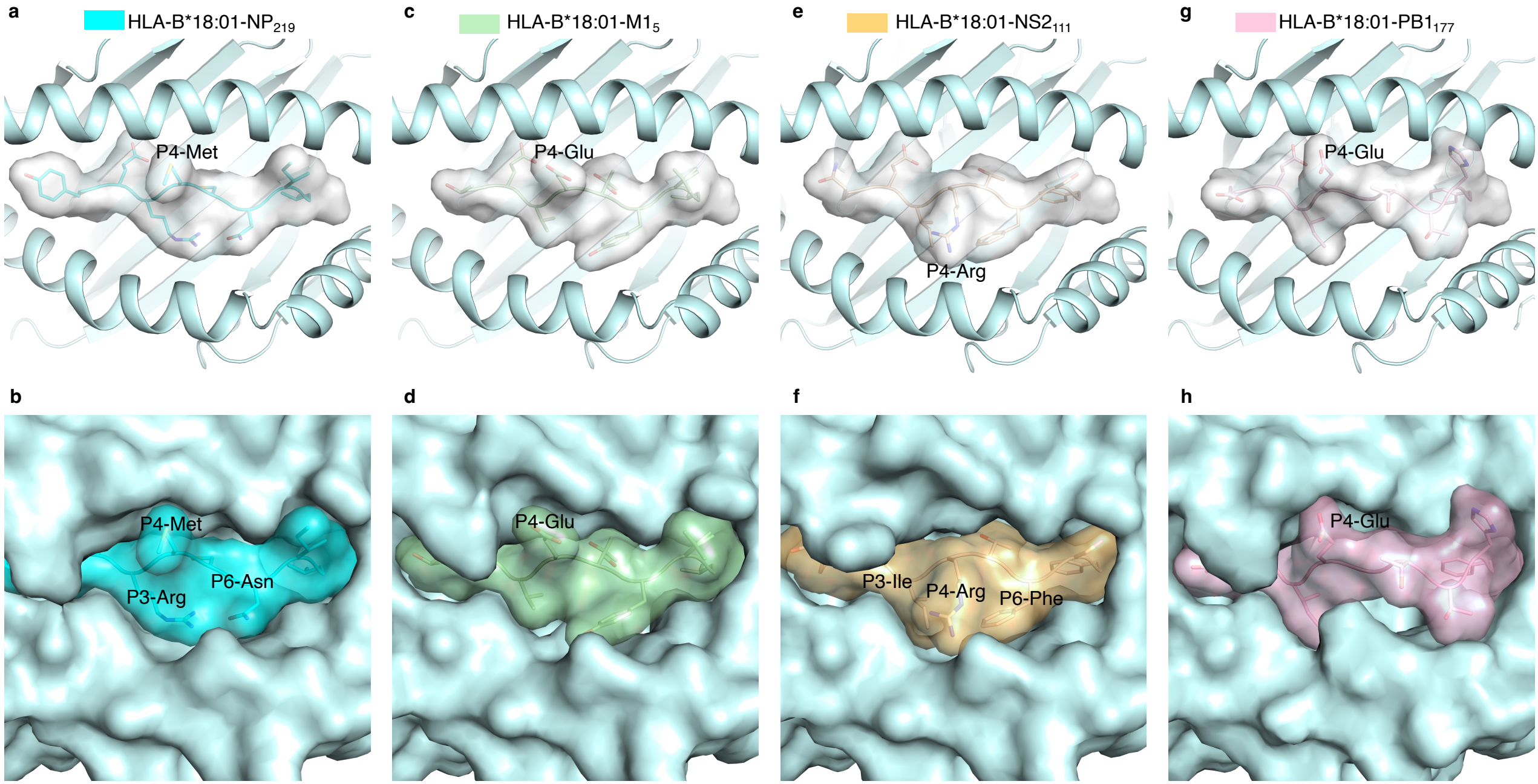
PBMCs from three HLA-B*18:01⁺ donors (n=3) were stimulated with 10 mM of the NP_{219}, $PB1_{177}$, $NS2_{111}$ or $M1_5$ peptides individually and CD8⁺ T cell function was measured in an ICS assay **(a)** Representative FACS plots and **(b)** Summary of all cytokines produced by CD8⁺ T cell lines reported in Figures 2C and 3. The coloured dots represents a particular donor's effector response (red, SG10; blue, SG115; and purple, SG131).



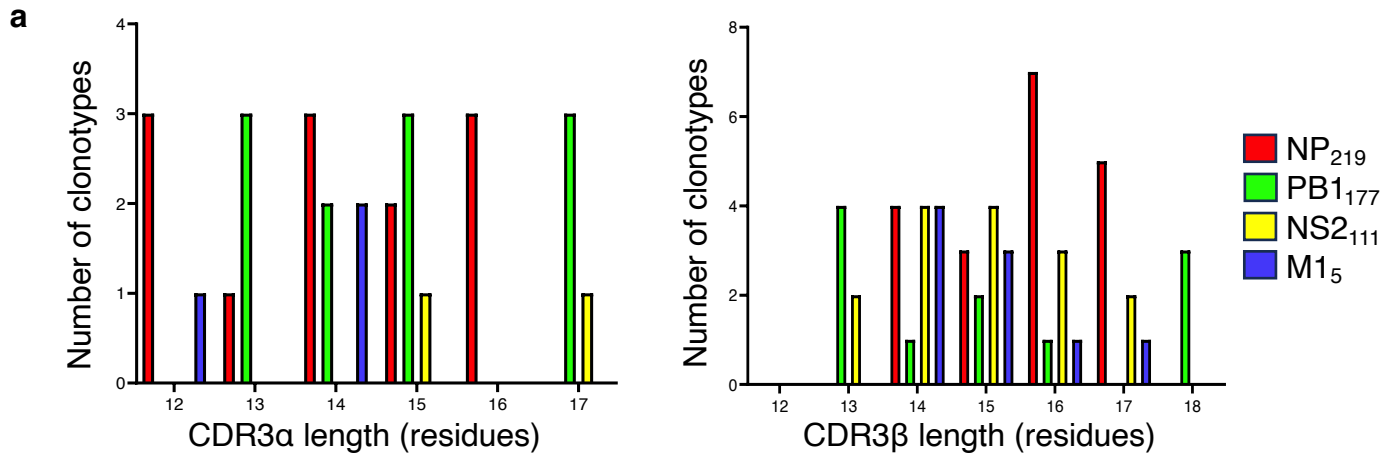
Supplementary Figure 4. Electron density maps for the four immunogenic peptides presented by HLA-B*18:01. Unbiased electron density (Fo-Fc, green at 3s) and final electron density (2Fo-Fc, blue at 1s) around the peptides NP_{219} (**a**, **b**), $M1_5$ (**c**, **d**), $NS2_{111}$ (**e**, **f**), and $PB1_{117}$ (**g**, **h**), respectively.



Supplementary Figure 5. Hydrogen bonding network between the HLA-B*18:01 binding pocket (light cyan) and the peptide residue side chains of NP_{219} (cyan) and $M1_5$ (green). (**a**) HLA-B*18:01 (pale cyan) presenting the NP_{219} peptide (cyan), with hydrogen bond as yellow dashed lines. (**b**) HLA-B*18:01 (pale cyan) presenting the $M1_5$ peptide (green), with hydrogen bond as yellow dashed lines.

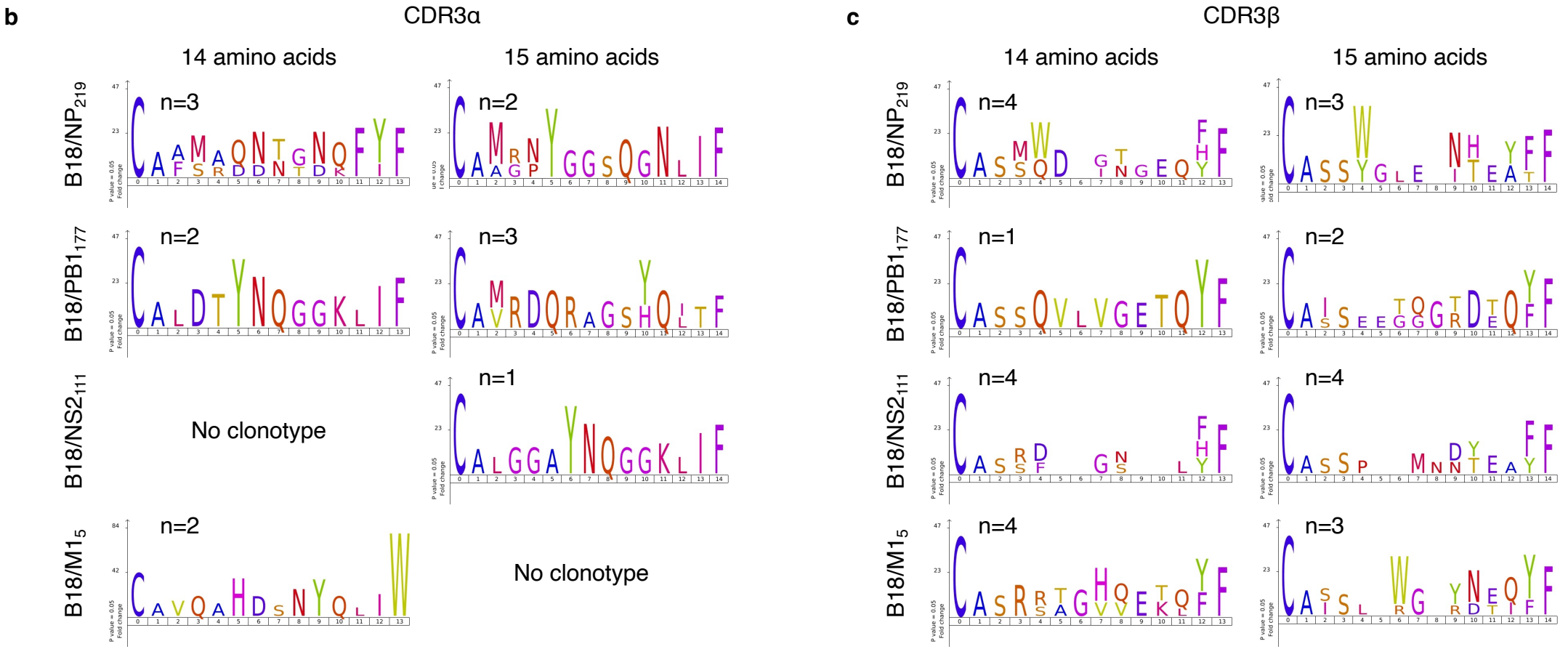


Supplementary Figure 6. Top view of pHLA-B*18:01 structures represented as surface. Top-down view of peptide (**a, b**) NP_{219} (cyan), (**c, d**) $M1_5$ (green), (**e, f**) $NS2_{111}$ (yellow), and (**g, h**) $PB1_{117}$ (pink). The HLA-B*18:01 binding cleft (light cyan) represented as pale cyan cartoon on the top panels and as surface on the bottom panels.



Supplementary Figure 7. Distinct clonotype T cell receptor length and motif identification. $PB1_{177}$, $NS2_{111}$, $M1_5$ and NP_{219} peptide specific T cell lines from three HLA-B*18:01$^+$ donors were tetramer stained with the cognate peptide, single-cell sorted and subject to single-cell multiplex PCR. **(a)** The bar graph illustrates the length of the distinct TCR clonotype **(b-c)** Depicts the **(b)** CDR3a and (C) CRD3b motifs for the most common lengths of 14 and 15 amino acids long. Blank amino acid residues display no amino-acid preferences. Motif viewer of CDR3 was obtained via iceLogo (Colaert *et al*, *Nat Methods* 2009; **6**: 786-787).

.



Supplementary Table 1. Data Collection and Refinement Statistics

Data Collection Statistics	**HLA-B*18:01-NS2$_{111}$**	**HLA-B*18:01-PB1$_{177}$**	**HLA-B*18:01-NP$_{219}$**	**HLA-B*18:01-M1$_5$**
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Cell Dimensions (a,b,c) (Å)	50.76, 81.27, 109.89	50.75, 81.63, 110.56	50.85 81.48 111.30	50.65, 81.14, 110.1
Resolution (Å)	46.08 – 1.15 (1.17 – 1.15)	46.12 – 1.60 (1.63 – 1.60)	46.25 – 1.40 (1.42 – 1.40)	46.02 – 1.45 (1.47 – 1.45)
Total number of observations	1182798 (56379)	280508 (13853)	716128 (34694)	738504 (35754)
Number of unique observations	159264 (7596)	59129 (3078)	91186 (4394)	81219 (3941)
Multiplicity	7.4 (7.4)	4.7 (4.5)	7.9 (7.9)	9.1 (9.1)
Data completeness (%)	98.7 (96.5)	97 (98.5)	99.6 (98.2)	100 (100)
I/s_I	13.3 (2.5)	10.3 (1.6)	15.6 (3.3)	14.9 (3.5)
Mn(I) half-set correlation $CC_{(1/2)}$	99.9 (80.0)	99.7 (63.0)	99.9 (63.9)	99.8 (70.7)
R_{pim}[a] (%)	2.5 (29.1)	3.8 (38.9)	2.9 (27.5)	3.4 (26.1)
Refinement Statistics				
R_{factor}[b] (%)	15.24	20.16	15.37	18.60
R_{free}[b] (%)	16.72	23.43	18.86	20.23
Rms deviations from ideality				
Bond lengths (Å)	0.004	0.007	0.009	0.001
Bond angles (°)	0.79	0.92	1.03	1.25
Ramachandran plot (%)				
Allowed region	98.42	98.64	98.68	98.68
Disallowed region	0.00	0.00	0.26	0.26
PDB code	8ROP	8RNH	8ROO	8RNG

[a] $R_{p.i.m} = S_{hkl} [1/(N-1)]^{1/2} S_i | I_{hkl,i} - < I_{hkl} > | / S_{hkl} < I_{hkl} >$.

[b] $R_{factor} = S_{hkl} | | F_o | - | F_c | | / S_{hkl} | F_o |$ for all data except ≈ 5% which were used for R_{free} calculation.

Supplementary Table 2. B18/NP$_{219}$ specific TCR sequences

								Number of sequences		
TRAV	**TRAJ**	**CDR3α**	**Length**	**TRBV**	**TRBJ**	**CDR3β**	**Length**	**SG10**	**SG115**	**SG131**
14/DV4*01	4*01	CSLREEFSGGFNKLIF	16					1		
22*01	36*01	CAPQSPWANNLFF	13					1		
				19*01/02	2-1*01	CASSIDLTSGVYNEQFF	17	1		
				19*01/02	2-5*01	CASSIDYPSGVQETQYF	17	2		
				19*01/02	2-7*01	CASSTDLATGAYEQYF	16	1		
				19*01/02	2-7*01	CASSIDLASGTTYEQYF	17	1		
				2*01	2-5*01	CASSEGSLGGGLETQYF	17	1		
				2*01/02/03	2-5*01	CASMGRDINQPQHF	14	1		
				4-1*01	1-1*01	CASSQDPGGTEAFF	14	1		
38-1*01	34*01	CAFMRDDNTDKFIF	14	4-3*01	2-7*01	CASSQDFAGGSYEQYF	16		1	
29/DV5*01	42*01	CAAGNYGGSQGNLIF	15	7-2*04	1-2*01	CASSLEAVSIHGYTF	15		1	
20*01/02	42*01	CAVGREYGGSQGNLIF	16	4-1*01	2-2*01	CASSQDKTTGELFF	14		1	
				6-5*01	1-1*01	CASSWGLEVNTEAFF	15		1	
				6-5*01	1-1*01	CASSYGLEANTEAFF	15		1	
14/DV4*03	8*01	CVMREAMNTGFQKLVF	16						1	
				4-3*01	1-2*01	CASSQDRGTGAYGYTF	16		1	
14/DV4*03	42*01	CAMRPYGGSQGNLIF	15	4-3*01	2-7*01	CASSQDRASGEYEQYF	16			1
29/DV5*01	49*01	CAASAQNTGNQFYF	14	7-2*01/02/04 or 7-8*01	1-1*01	CASSPSMRSGLAEAFF	16			9
29/DV5*01	49*01	CAASAQNTGNQFYF	14							6
22*01	20*01	CAVVSNDYKLSF	12							2
13-2*01/02	37*02	CADSSNTGKLIF	12							1
	37*02	CAEASNTGKLIF	12							1
				19*03	2-7*01	CASSWDRGTGEQYF	14			1
				4-3*01	2-1*01	CASSQDLASGTYNEQFF	17			3

	7-2*01/02/04 or 7-8*01	1*01	CASSPSMRSGLAEAFF	16			1
	19*01/02	2-7*01	CASSIDLAGGSYEQYF	16			1
			Total number of resolved sequences	**10**	**7**		**26**

The αβTCR repertoire of $CD8^+$ T cells from SG10, SG115 and SG131 specific for the B18/NP_{219} tetramer. Length refers to the number of amino acids in the CDR3.

Supplementary Table 3. B18/PB1$_{177}$ specific TCR sequences

								Number of sequences	
TRAV	**TRAJ**	**CDR3α**	**Length**	**TRBV**	**TRBJ**	**CDR3β**	**Length**	**SG10**	**SG115**	**SG131**
1-2*01	28*01	CAVRDQYSGAGSYQLTF	17					1		
				10-3*04	2-1*01	CAISEEGQGRDEQFF	15	1		
				10-2*02	1-2*01	CASSLGTGGGYTF	13	2		
				3-2*03	2-5*01	CASSQVLVGETQYF	14	2		
				4-3*02/03	2-3*01	CASSSSGGGAGDTQYF	16	1		
				20-1*02/03	2-3*01	CSASPERGTSVGTDTQYF	18	1		
				19*01/02	1-1*01	CASSIVERAEAFF	13	1		
1-2*01	28*01	CAVRDQRAGSYQLTF	15						3	
1-2*01/03	28*01	CAVRDQRAGSYQLTF	15						12	
1-2*01/03	28*01	CAVRDQLTGAGSYQLTF	17						3	
1-2*01/03	28*01	CAMRDQRGGSHQIPL	15						1	
3*01	13*01	CAVRDSFGSGGYQKVTF	17						1	
				5-4*01	2-3*01	CASSLLTGGTDTQYF	15		2	
				19*01/03	2-3*01	RAGPGQRLWAPVF	13		1	
6*01/07	23*01	CALDTYNQGGKLIF	14	27*01	2-2*01	CASSWVPPGQGRTGELFF	18			8
6*01/07	23*01	CALDTYNQGGKLIF	14							11
				27*01	2-2*01	CASSWVPPGQGRTGELFF	18			1
							Total number of sequences	**9**	**23**	**20**

The αβTCR repertoire of CD8$^+$ T cells from SG10, SG115 and SG131 specific for the B18/PB1$_{177}$ tetramer. Length refers to the number of amino acids in the CDR3.

Supplementary Table 4. B18/NS2$_{111}$ specific TCR sequences

							Number of sequences		
TRAV	**TRAJ**	**CDR3**	**Length**	**TRBV**	**TRBJ**	**CDR3β**	**Length**	**SG10**	**SG115**	**SG131**
6*01/07	23*01	CALGGAYNQGGKLIF	15					1		
				6-4*01/02	2-2*01	CASSDSDGSGELFF	14	1		
				20-1*01	2-1*01	CSARDRQASSYNEQFF	16	2		
				27*01	1-4*01	CASSFLQGEKLFF	13	1		
				27*01	1-2*01	CASSLRPDPPYNEQFF	16	1		
				27*01	1-1*01	CASSPPGREDTEAFF	15	1		
				18*01	1-1*01	CASSPSEGLNTEAFF	15	1		
				20-1*01	1-6*02	CSARDGQGSSPLHF	14	1		
				20-1*01/04/05	2-1*01	CSANDRTSGSNYNEQFF	17	1		
				15*02	1-5*01	CATSRDRGWSNQPQHF	16	1		
				4-1*01/02	1-6*01/02	CASSQAGSLENSPLPLL	17	1		
				2*02	2-7*01	CASSTGRYYEQYF	13	1		
8-6*02	32*02	CAVTEGPYGGATNKLIF	17	4-2*01	2-3*01	CASSKGLAVDYTQYF	15		1	
				12-3/4*01	2-3*01	CASRLVSGTDTQYF	14		7	
				28*01	1-4*01	CASSFLATNEKLFF	14		1	
				20-1*04	2-2*01	CSASLERMNTGELFF	15			1
						Total number of sequences resolved		**13**	**9**	**1**

The αβTCR repertoire of CD8$^+$ T cells from SG10, SG115 and SG131 specific for the B18/NS2$_{111}$ tetramer. Length refers to the number of amino acids in the CDR3.

Supplementary Table 5. B18/M1$_5$ specific TCR sequences

							Number of sequences		
TRAV	**TRAJ**	**CDR3**	**Length**	**TRBV**	**TRBJ**	**CDR3β**	**Length**	**SG10**	**SG115**	**SG131**
20*01/02	33*01	CAVQAHDSNYQLIW	14	4-1*01/02	1-4*01	CASRRTGVVEKLFF	14	3		
20*01/02	33*01	CAVQAHDSNYQLIW	14					5		
				4-1*01/02	1-4*01	CASRRTGVVEKLFF	14	1		
				10-3*04	2-7*01	CAISAPRGERDEQYF	15	5		
				4-3*01	2-3*01	CASSQGPQVRGADTQYF	17		5	
				27*01	1-3*01	CASSLSWGPGNTIYF	15		4	
				18*01	1-2*01	CASSPTPGDPHYGYTF	16		2	
14/DV4*03	47*01	CAMREDGNKLVF	12	27*01	2-5*01	CASRSAGHQETQYF	14			1
				27*01	2-5*01	CASRSAGHQETQYF	14			2
				19*01/02	2-1*01	CASSLLAGVYNEQFF	15			1
						Total number of resolved sequences		**14**	**11**	**4**

The αβTCR repertoire of CD8$^+$ T cells from SG10, SG115 and SG131 specific for the B18/M1$_5$ tetramer. Length refers to the number of amino acids in the CDR3.