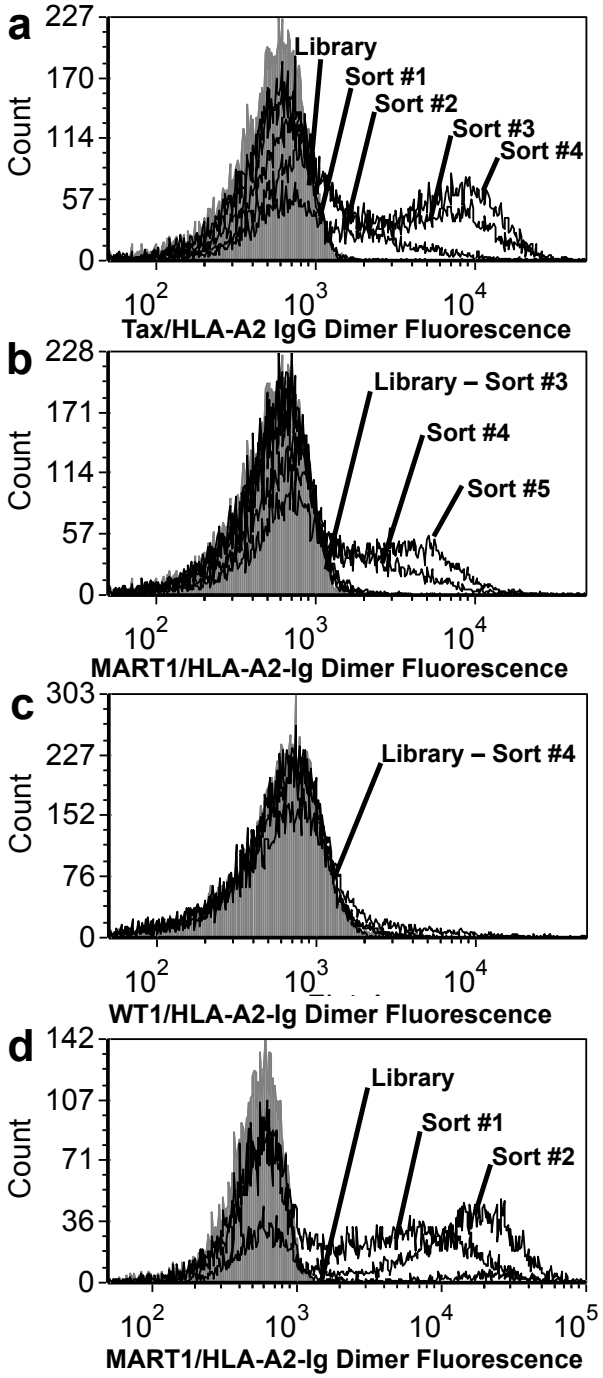
Supplementary Information

Smith S.N., Wang Y., Baylon J.L., Singh N.K., Baker B.M., Tajkhorshid E., & Kranz D.M.

Changing the peptide specificity of a human T cell receptor by directed evolution.

Supplementary Figures



Supplementary Figure 1 | Flow cytometry histograms of the selections of the RD1 library and affinity maturation of RD1-MART1. (**a**) The RD1 library was sorted sequentially with Tax/HLA-A2-Ig dimer. Aliquots of yeast cells after each sort were then incubated with 100 nM Tax /HLA-A2-Ig dimer followed by APC-conjugated goat anti-mouse secondary antibody. (**b**) The RD1 library was sorted sequentially with MART1 (ELAGIGILTV)/HLA-A2-Ig dimer for a total of five sorts. During the 3rd sort, yeast cells were also stained with chicken anti-c-myc antibody, goat anti-chicken IgY alexa 488 secondary antibody and double positives were isolated in order to exclude truncated clones (not shown). Aliquots of yeast cells after each sort were then incubated with 100 nM MART1 (ELAGIGILTV)/HLA-A2-Ig dimer followed by APC-conjugated goat anti-mouse secondary antibody. (**c**) The RD1 library was sorted sequentially with WT1/HLA-A2-Ig dimer. Aliquots of yeast cells after each sort were then incubated with 100 nM WT1/HLA-A2-Ig dimer followed by APC-conjugated goat anti-mouse secondary antibody. (**d**) RD1-MART1 was used as a template for affinity maturation libraries in CDR3 loops, and combined libraries were sorted sequentially with MART1/ (ELAGIGILTV)/HLA-A2-Ig dimer for a total of two sorts. Aliquots of yeast cells after each sort were then incubated with 50 nM MART1 (ELAGIGILTV)/HLA-A2-Ig dimer, followed by APC-conjugated goat anti-mouse secondary antibody. Data is representative of 2 experiments with similar results.

	CDR 1α									CDR 3α										CDR 3β												
	24	25	26	27	28	29	30	31	32	93	98	99	100	101	102	103	104	105	93	94	95	96	97	98	99	100	101	102	103	105	106	107
A6 wild-type	Y	S	D	R	G	S	Q	S	F	T	T	D	S	W	G	K	L	Q	A	S	R	P	G	L	A	G	G	R	P	E	L	Y
A6-X15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	M	S	A	Q	-	-	L	-
RD1 Library	-	-	-	-	-	-	X	-	-	-	X	X	-	-	-	-	-	-	-	-	-	-	-	X	M	S	X	Q	-	-	L	-
RD1-Tax-S1-1	-	-	-	-	-	-	S	-	-	-	P	F	-	-	-	-	-	-	-	-	-	-	-	E	M	S	Q	Q	-	-	L	-
RD1-Tax-S1-2	-	-	-	-	-	-	A	-	-	-	R	S	-	-	-	-	-	-	-	-	-	-	-	P	M	S	W	Q	-	-	L	-
RD1-Tax-S1-3	-	-	-	-	-	-	S	-	-	-	Y	V	-	-	-	-	-	-	-	-	-	-	-	*	M	S	W	Q	-	-	L	-
RD1-Tax-S1-4	-	-	-	-	-	-	E	-	-	-	S	S	-	-	-	-	-	-	-	-	-	-	-	T	M	S	M	Q	-	-	L	-
RD1-Tax-S1-5	-	-	-	-	-	-	I	-	-	-	D	Q	-	-	-	-	-	-	-	-	-	-	-	*	M	S	E	Q	-	-	L	-
RD1-Tax-S1-6	-	-	-	-	-	-	H	-	-	-	I	Q	-	-	-	-	-	-	-	-	-	-	-	L	M	S	G	Q	-	-	L	-
RD1-Tax-S2-1	-	-	-	-	-	-	S	-	-	-	P	H	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S2-2	-	-	-	-	-	-	W	-	-	-	H	S	-	-	-	-	-	-	-	-	-	-	-	W	M	S	A	Q	-	-	L	-
RD1-Tax-S2-3	-	-	-	-	-	-	V	-	-	-	K	S	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S2-4	-	-	-	-	-	-	R	-	-	-	L	M	-	-	-	-	-	-	-	-	-	-	-	Y	M	S	C	Q	-	-	L	-
RD1-Tax-S2-5	-	-	-	-	-	-	M	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S2-6	-	-	-	-	-	-	V	-	-	-	S	S	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S3-1	-	-	-	-	-	-	S	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S3-2	-	-	-	-	-	-	Q	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S3-3	-	-	-	-	-	-	T	-	-	-	S	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S3-4	-	-	-	-	-	-	T	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S3-5	-	-	-	-	-	-	S	-	-	-	G	S	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S3-6	-	-	-	-	-	-	Q	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S4-1	-	-	-	-	-	-	T	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S4-2	-	-	-	-	-	-	T	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S4-3	-	-	-	-	-	-	Q	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S4-4	-	-	-	-	-	-	T	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S4-5	-	-	-	-	-	-	Q	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-S4-6	-	-	-	-	-	-	T	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-1	-	-	-	-	-	-	Q	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-Tax-2	-	-	-	-	-	-	T	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	L	M	S	A	Q	-	-	L	-

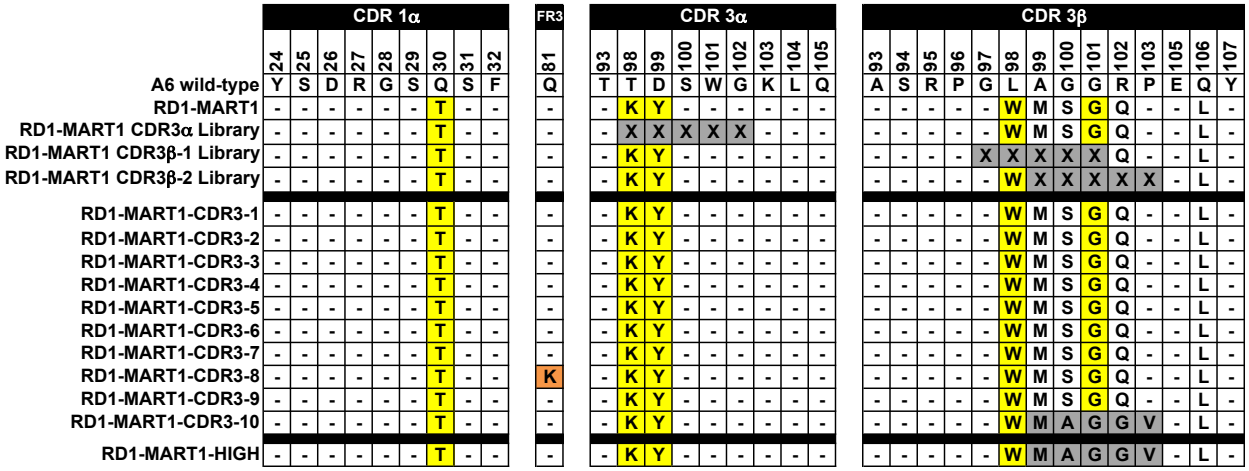
Supplementary Figure 2 | Sequence alignment of clones isolated from the RD1 library following each round of selection with Tax/HLA-A2. The sequences of the A6-X15 template and RD1 library are shown, and degenerate residues in the RD1 library are highlighted in yellow. Clones isolated from each of the rounds of sorting are indicated by S1 to S4. The two Tax-specific variants referenced in Fig. 2, RD1-Tax-1 and RD1-Tax-2, are repeated in the bottom two rows. Residue numbering is consistent with the crystal structures of A6:Tax/HLA-A2 (PDB: 1AO7)[1] and A6-c134:Tax/HLA-A2 (PDB: 4FTV)[2].

	CDR 1α								
	24	25	26	27	28	29	30	31	32
A6 wild-type	Y	S	D	R	G	S	Q	S	F
A6-X15	-	-	-	-	-	-	-	-	-
RD1 Library	-	-	-	-	-	-	X	-	-
RD1-MART1-S1-1	-	-	-	-	-	-	C	-	-
RD1-MART1-S1-2	-	-	-	-	-	-	S	-	-
RD1-MART1-S1-3	-	-	-	-	-	-	F	-	-
RD1-MART1-S1-4	-	-	-	-	-	-	V	-	-
RD1-MART1-S1-5	-	-	-	-	-	-	M	-	-
RD1-MART1-S2-1	-	-	-	-	-	-	C	-	-
RD1-MART1-S2-2	-	-	-	-	-	-	T	-	-
RD1-MART1-S2-3	-	-	-	-	-	-	V	-	-
RD1-MART1-S2-4	-	-	-	-	-	-	V	-	-
RD1-MART1-S2-5	-	-	-	-	-	-	R	-	-
RD1-MART1-S3-1	-	-	-	-	-	-	G	-	-
RD1-MART1-S3-2	-	-	-	-	-	-	K	-	-
RD1-MART1-S3-3	-	-	-	-	-	-	N	-	-
RD1-MART1-S3-4	-	-	-	-	-	-	R	-	-
RD1-MART1-S3-5	-	-	-	-	-	-	N	-	-
RD1-MART1-S3-6	-	-	-	-	-	-	K	-	-
RD1-MART1-S4-1	-	-	-	-	-	-	T	-	-
RD1-MART1-S4-2	-	-	-	-	-	-	T	-	-
RD1-MART1-S4-3	-	-	-	-	-	-	T	-	-
RD1-MART1-S4-4	-	-	-	-	-	-	T	-	-
RD1-MART1-S4-5	-	-	-	-	-	-	T	-	-
RD1-MART1-S4-6	-	-	-	-	-	-	T	-	-
RD1-MART1-S5-1	-	-	-	-	-	-	T	-	-
RD1-MART1-S5-2	-	-	-	-	-	-	T	-	-
RD1-MART1-S5-3	-	-	-	-	-	-	T	-	-
RD1-MART1-S5-4	-	-	-	-	-	-	T	-	-
RD1-MART1-S5-5	-	-	-	-	-	-	T	-	-
RD1-MART1	-	-	-	-	-	-	T	-	-

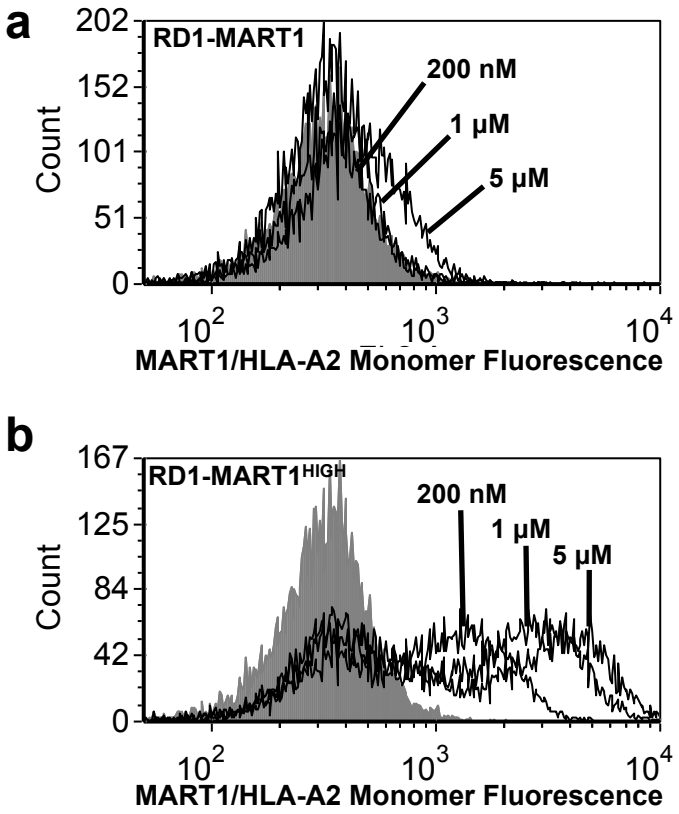
	CDR 3α								
	93	98	99	100	101	102	103	104	105
A6 wild-type	T	T	D	S	W	G	K	L	Q
A6-X15	-	-	-	-	-	-	-	-	-
RD1 Library	-	X	X	-	-	-	-	-	-
RD1-MART1-S1-1	-	Q	S	-	-	-	-	-	-
RD1-MART1-S1-2	-	I	V	-	-	-	-	-	-
RD1-MART1-S1-3	-	*	Y	-	-	-	-	-	-
RD1-MART1-S1-4	-	H	Y	-	-	-	-	-	-
RD1-MART1-S1-5	-	C	L	-	-	-	-	-	-
RD1-MART1-S2-1	-	L	H	-	-	-	-	-	-
RD1-MART1-S2-2	-	H	P	-	-	-	-	-	-
RD1-MART1-S2-3	-	F	T	-	-	-	-	-	-
RD1-MART1-S2-4	-	N	M	-	-	-	-	-	-
RD1-MART1-S2-5	-	R	H	-	-	-	-	-	-
RD1-MART1-S3-1	-	K	M	-	-	-	-	-	-
RD1-MART1-S3-2	-	R	H	-	-	-	-	-	-
RD1-MART1-S3-3	-	R	M	-	-	-	-	-	-
RD1-MART1-S3-4	-	K	H	-	-	-	-	-	-
RD1-MART1-S3-5	-	R	M	-	-	-	-	-	-
RD1-MART1-S3-6	-	K	H	-	-	-	-	-	-
RD1-MART1-S4-1	-	K	Y	-	-	-	-	-	-
RD1-MART1-S4-2	-	K	Y	-	-	-	-	-	-
RD1-MART1-S4-3	-	K	Y	-	-	-	-	-	-
RD1-MART1-S4-4	-	K	Y	-	-	-	-	-	-
RD1-MART1-S4-5	-	K	Y	-	-	-	-	-	-
RD1-MART1-S4-6	-	K	Y	-	-	-	-	-	-
RD1-MART1-S5-1	-	K	Y	-	-	-	-	-	-
RD1-MART1-S5-2	-	K	Y	-	-	-	-	-	-
RD1-MART1-S5-3	-	K	Y	-	-	-	-	-	-
RD1-MART1-S5-4	-	K	Y	-	-	-	-	-	-
RD1-MART1-S5-5	-	K	Y	-	-	-	-	-	-
RD1-MART1	-	K	Y	-	-	-	-	-	-

	CDR 3β													
	93	94	95	96	97	98	99	100	101	102	103	105	106	107
A6 wild-type	A	S	R	P	G	L	A	G	G	R	P	E	Q	Y
A6-X15	-	-	-	-	-	-	M	S	A	Q	-	-	L	-
RD1 Library	-	-	-	-	-	X	M	S	X	Q	-	-	L	-
RD1-MART1-S1-1	-	-	-	-	-	T	M	S	L	Q	-	-	L	-
RD1-MART1-S1-2	-	-	-	-	-	G	M	S	L	Q	-	-	L	-
RD1-MART1-S1-3	-	-	-	-	-	L	M	S	A	Q	-	-	L	-
RD1-MART1-S1-4	-	-	-	-	-	Q	M	S	R	Q	-	-	L	-
RD1-MART1-S1-5	-	-	-	-	-	M	M	S	N	Q	-	-	L	-
RD1-MART1-S2-1	-	-	-	-	-	S	M	S	P	Q	-	-	L	-
RD1-MART1-S2-2	-	-	-	-	-	H	M	S	E	Q	-	-	L	-
RD1-MART1-S2-3	-	-	-	-	-	Q	M	S	P	Q	-	-	L	-
RD1-MART1-S2-4	-	-	-	-	-	L	M	S	L	Q	-	-	L	-
RD1-MART1-S2-5	-	-	-	-	-	L	M	S	E	Q	-	-	L	-
RD1-MART1-S3-1	-	-	-	-	-	T	M	S	S	Q	-	-	L	-
RD1-MART1-S3-2	-	-	-	-	-	H	M	S	S	Q	-	-	L	-
RD1-MART1-S3-3	-	-	-	-	-	Q	M	S	T	Q	-	-	L	-
RD1-MART1-S3-4	-	-	-	-	-	G	M	S	R	Q	-	-	L	-
RD1-MART1-S3-5	-	-	-	-	-	E	M	S	P	Q	-	-	L	-
RD1-MART1-S3-6	-	-	-	-	-	Q	M	S	R	Q	-	-	L	-
RD1-MART1-S4-1	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S4-2	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S4-3	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S4-4	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S4-5	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S4-6	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S5-1	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S5-2	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S5-3	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S5-4	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1-S5-5	-	-	-	-	-	W	M	S	G	Q	-	-	L	-
RD1-MART1	-	-	-	-	-	W	M	S	G	Q	-	-	L	-

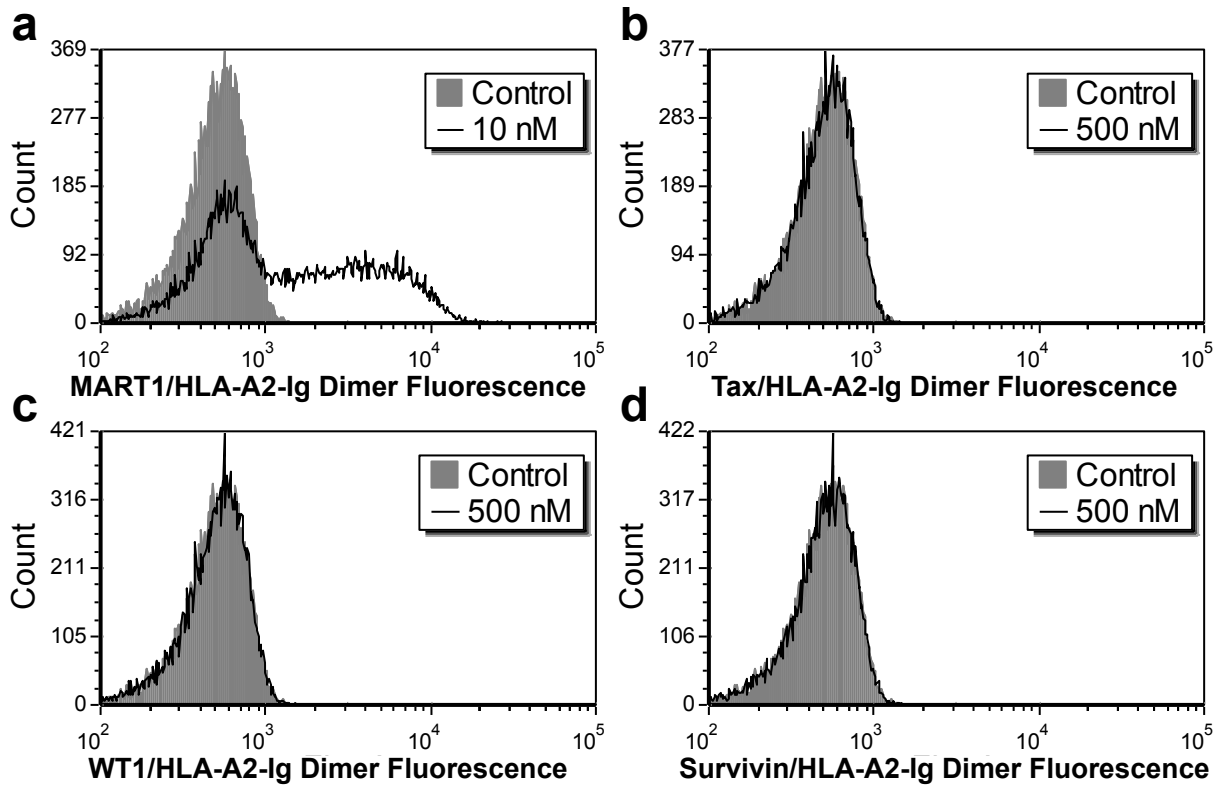
Supplementary Figure 3 | Sequence alignment of clones isolated from the RD1 library following each round of selection with MART1/HLA-A2. The sequences of the A6-X15 template and RD1 library are shown, and residues made degenerate in the RD1 library are highlighted in yellow. Clones isolated from each round of sorting are indicated by S1 to S5. The MART1-specific variant referenced in Fig. 2, RD1-MART1, is repeated in the bottom row. Residue numbering is consistent with the crystal structures of A6:Tax/HLA-A2 (PDB: 1AO7)[1] and A6-c134:Tax/HLA-A2 (PDB: 4FTV)[2].



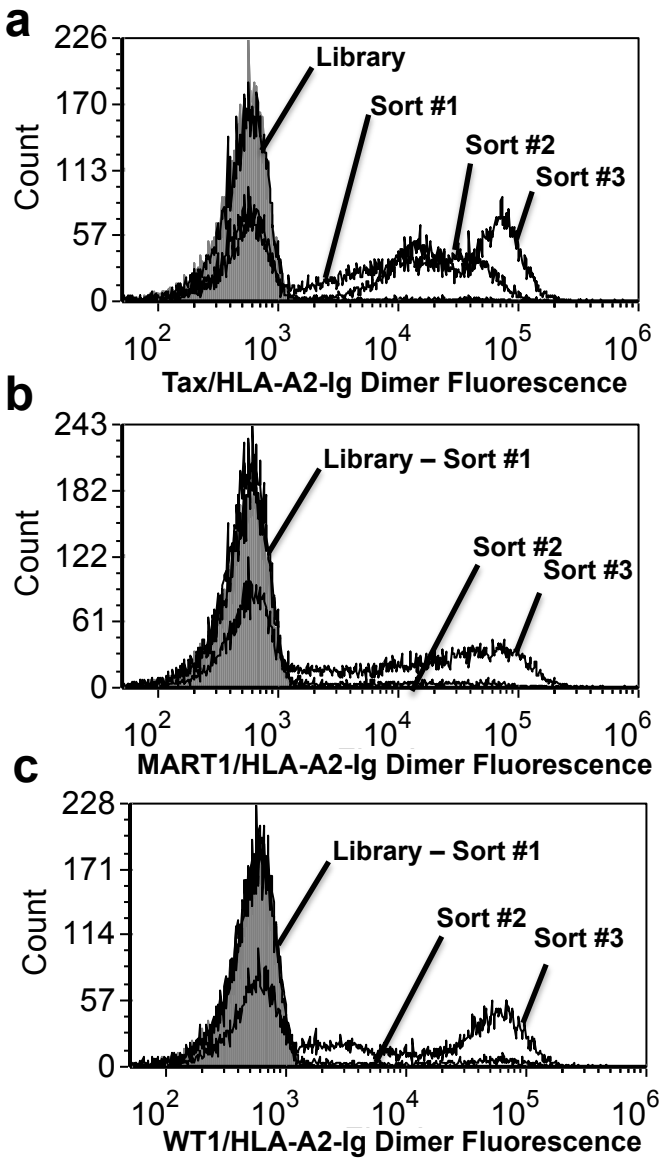
Supplementary Figure 4 | Sequence alignment of clones isolated from the RD1-MART1 CDR3 affinity maturation library following two selections with MART1/HLA-A2. Three five-codon (NNK) libraries, CDR3α, CDR3β–1, and CDR3β-2, shown in gray were generated and combined to create the RD1-MART1 CDR3 affinity-maturation library. Following two selections with MART1/HLA-A2 dimer (200 nM and 1 nM), 10 clones were isolated and sequenced (shown as RD1-MART1-CDR3 #1-10). Eight clones were parental, one clone called RD1-MART1-CDR3-8 contained a Q81K PCR-based framework mutation in the Vα-domain (shown in orange), and one clone, called RD1-MART1^{HIGH}, was derived from the CDR3α library and had three mutations that differed from the parental sequence (S100A, Q102G, and P103V; shown in gray). Residue numbering is consistent with the crystal structures of A6:Tax/HLA-A2 (PDB: 1AO7)[1] and A6-c134:Tax/HLA-A2 (PDB: 4FTV)[2].



Supplementary Figure 5 | MHC monomer staining of RD1-MART1 and RD1-MART1^{HIGH}. (**a**) The yeast-displayed RD1-MART1 clone stained at indicated concentrations of UV-exchanged MART1/HLA-A2 monomers, PE-conjugated streptavidin. (**b**) The yeast-displayed, affinity-matured RD1-MART1^{HIGH} clone stained with indicated concentrations of UV-exchanged MART1/HLA-A2 monomers, PE-conjugated streptavidin.

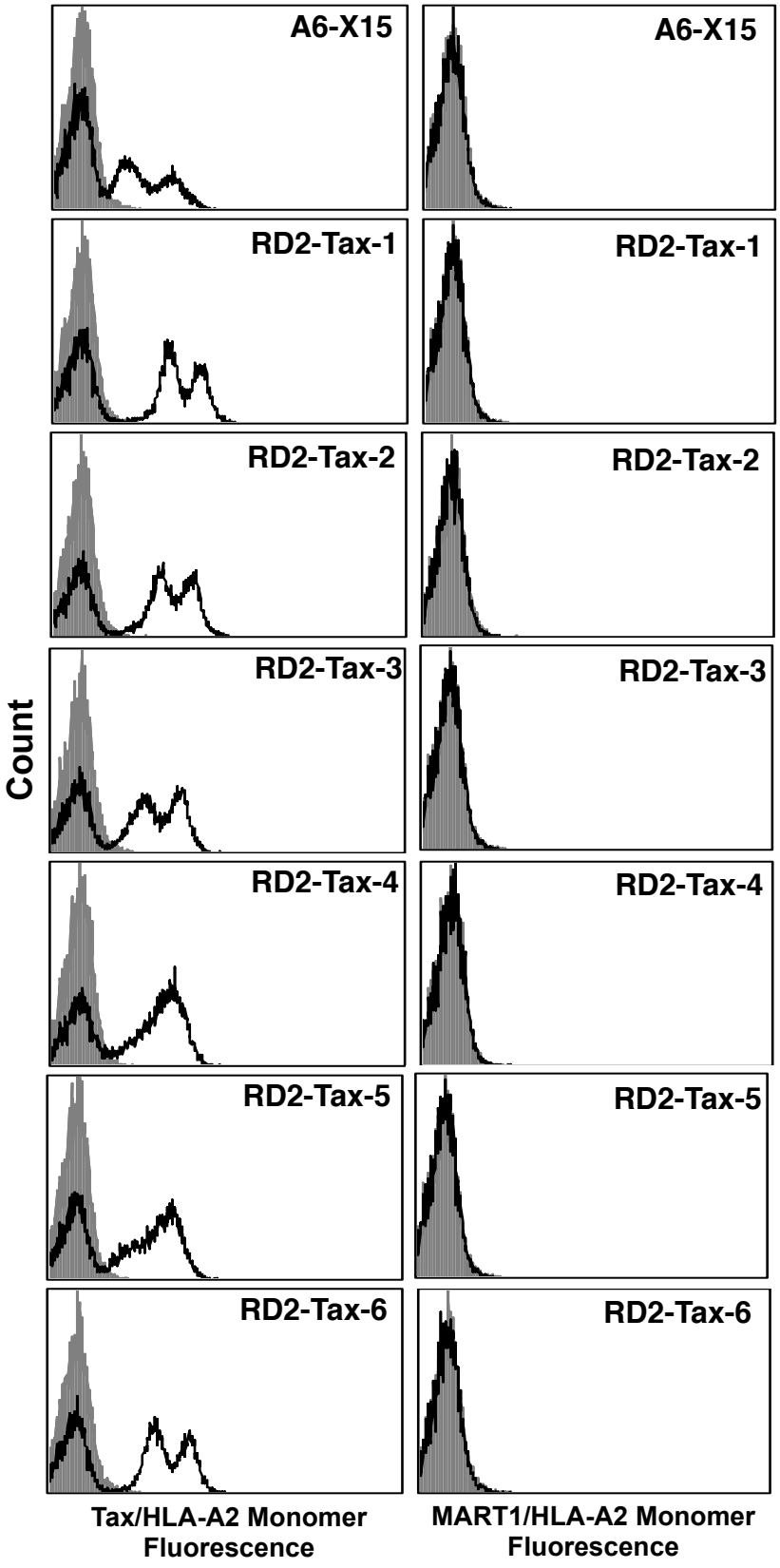


Supplementary Figure 6 | Staining of the RD1-MART1^{HIGH} clone to assess cross reactivity with other HLA-A2 restricted peptides. The yeast-displayed RD1-MART1^{HIGH} clone was stained with 10 nM MART1/HLA-A2-Ig dimer (**a**) or 500 nM Tax/HLA-A2-Ig dimer (**b**), WT1/HLA-A2-Ig dimer (**c**), or Survivin/HLA-A2-Ig dimer (**d**), followed by APC-conjugated goat anti-mouse secondary antibody. Gray indicated histograms of yeast stained with secondary reagent only. Data is representative of 2 experiments with similar results.

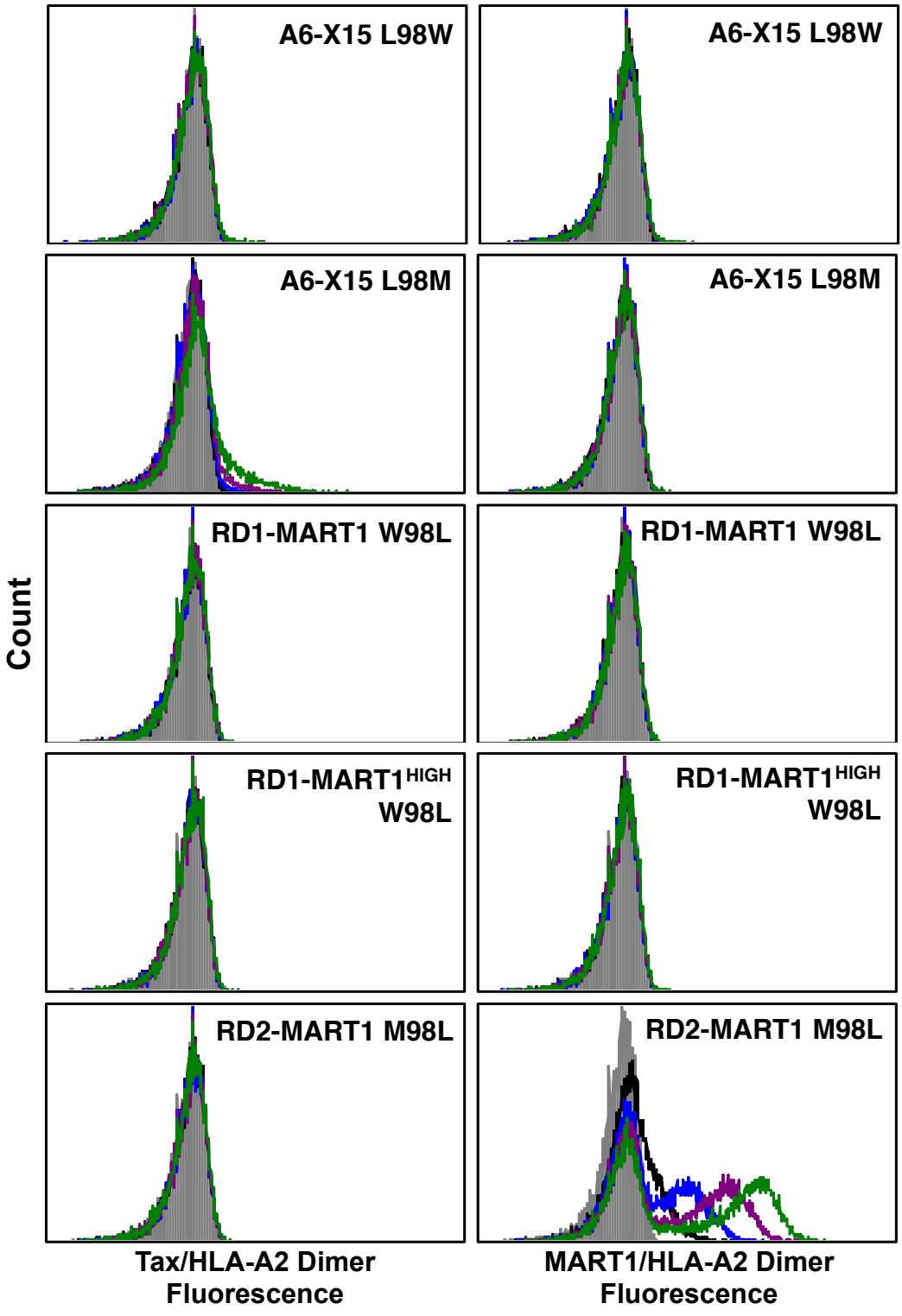


Supplementary Figure 7 | Selection of a second A6 library called RD2. Two sequential magnetic bead selections of the RD2 library were performed following incubation with 1 μM Tax/HLA-A2 or 5 μM MART1 or WT1/HLA-A2UV-exchanged monomers and streptavidin MACS beads (Miltenyi Biotec). A third selection was performed with FACS following incubation with 1 nM Tax/HLA-A2-Ig dimer or 100 nM MART1 or WT1/HLA-A2-Ig dimer, respectively. Aliquots of yeast cells after each selection were incubated with 50 nM selecting/HLA-A2-Ig dimer, APC-conjugated goat anti-mouse secondary antibody. (**a**) Flow cytometry histograms of the RD2

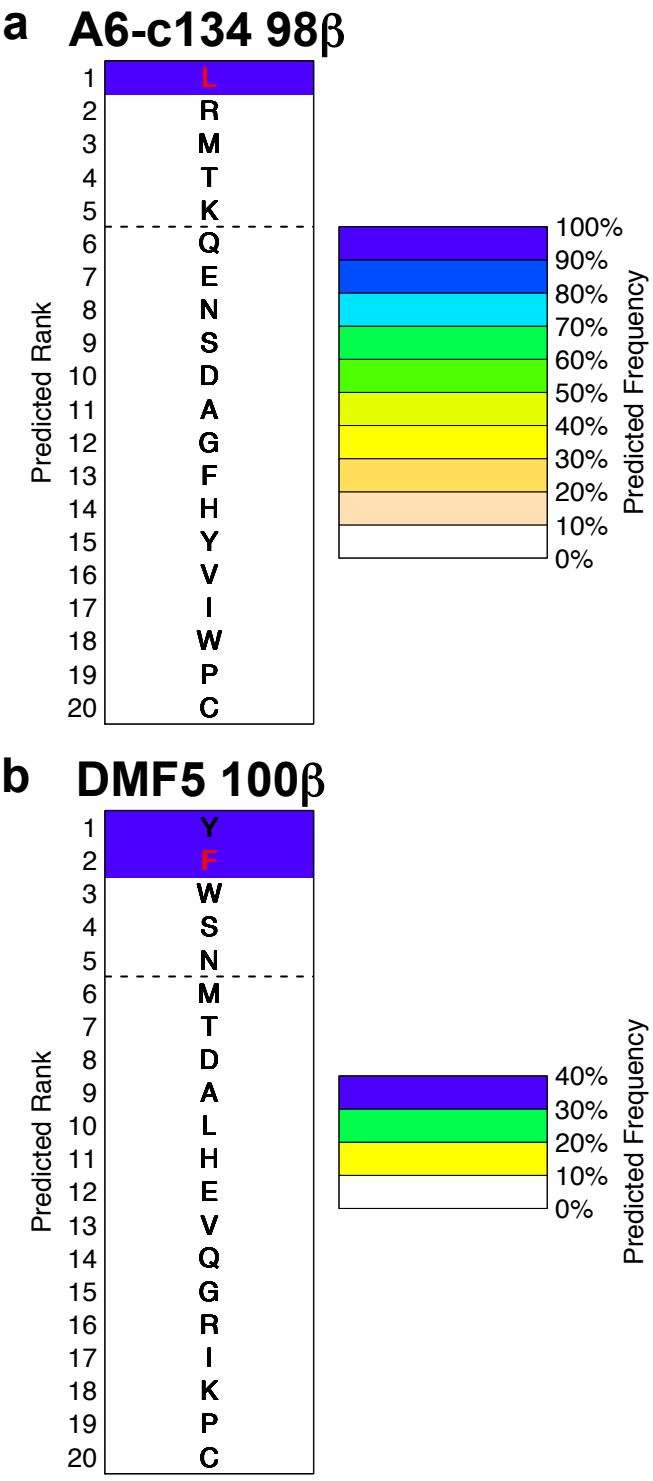
library after sorting with the cognate antigen, Tax (LLFGYPVYV)/HLA-A2. Gray indicates histograms of yeast cells stained with secondary antibody only. (**b**) Flow cytometry histograms of the RD2 library after sorting with the non-cognate antigen, MART1 (ELAGIGILTV)/HLA-A2. (**c**) Flow cytometry histograms of the RD2 library after sorting with the non-cognate antigen, WT1 (RMFPNAPYL)/HLA-A2. Data is representative of 2 experiments with similar results.


A6-X15
A6-X15
RD2-Tax-1
RD2-Tax-1
RD2-Tax-2
RD2-Tax-2
RD2-Tax-3
RD2-Tax-3
RD2-Tax-4
RD2-Tax-4
RD2-Tax-5
RD2-Tax-5
RD2-Tax-6
RD2-Tax-6
Count
Tax/HLA-A2 Monomer Fluorescence
MART1/HLA-A2 Monomer Fluorescence

Supplementary Figure 8 | Staining of RD2 clones isolated with Tax/HLA-A2. The six clones isolated following three sorts of the RD2 library with Tax/HLA-A2 were stained with 2 μM Tax/HLA-A2 monomers (left panels) or MART1/HLA-A2 monomers (right panels) followed by SA-PE. Gray filled histograms indicate yeast cells stained with SA-PE only.

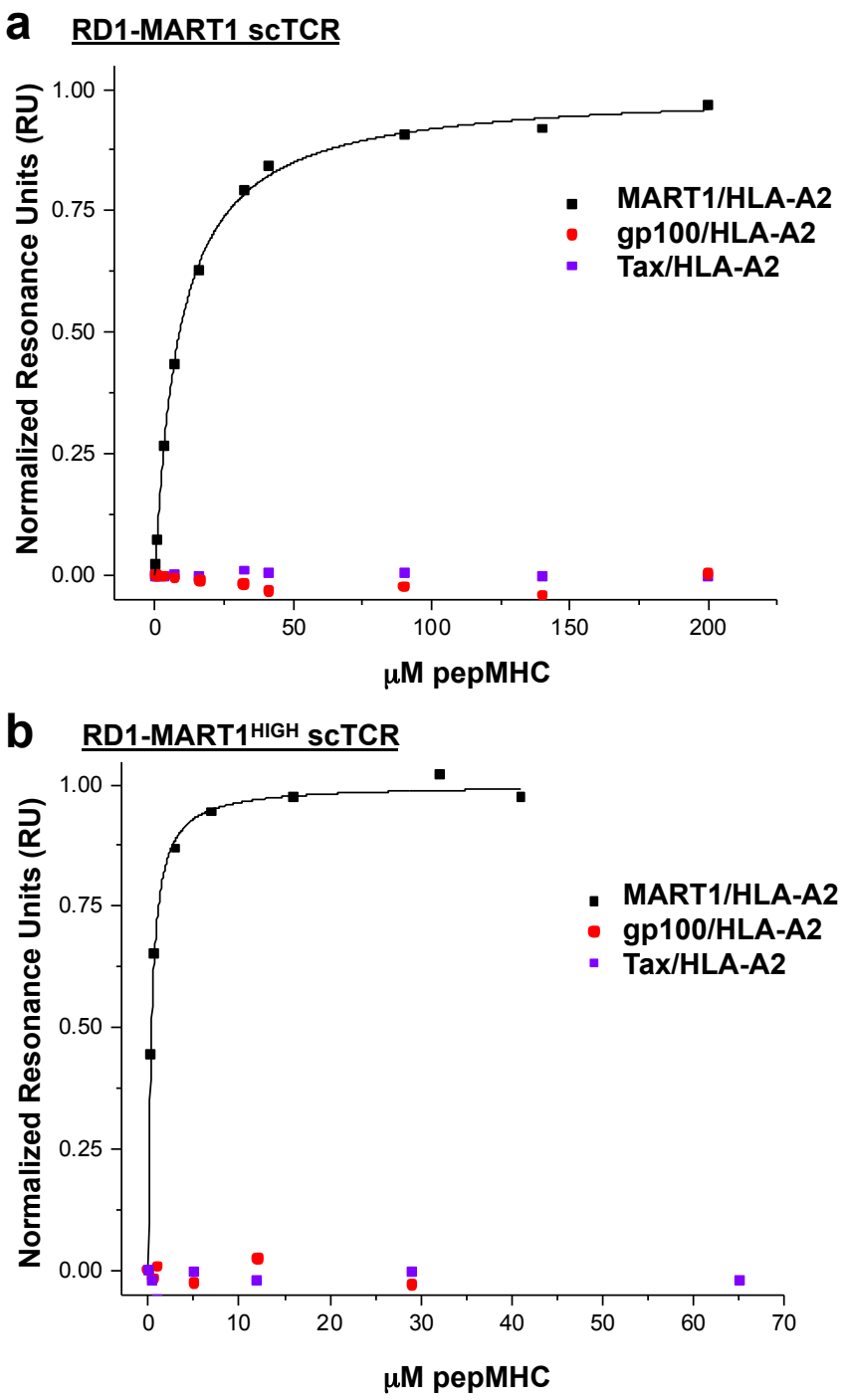

A6-X15 L98W
A6-X15 L98W
A6-X15 L98M
A6-X15 L98M
RD1-MART1 W98L
RD1-MART1 W98L
RD1-MART1HIGH W98L
RD1-MART1HIGH W98L
RD2-MART1 M98L
RD2-MART1 M98L
Count
Tax/HLA-A2 Dimer Fluorescence
MART1/HLA-A2 Dimer Fluorescence

Supplementary Figure 9 | Binding of position 98β mutants with pepHLA-A2 dimers. Flow cytometry histograms of yeast-displayed mutants of A6-X15, RD1-MART1, RD1-MART1^{HIGH}, and RD2-MART1. Mutants were stained with 18.5 nM (black), 55.6 nM (blue), 167 nM (purple), and 500 nM (green) Tax/HLA-A2-Ig (left panels) and MART/HLA-A2-Ig (right panels) dimers, APC-conjugated goat anti-mouse secondary antibody. This experiment was also performed with PE-conjugated streptavidin tetramers with similar results (Fig. 5).



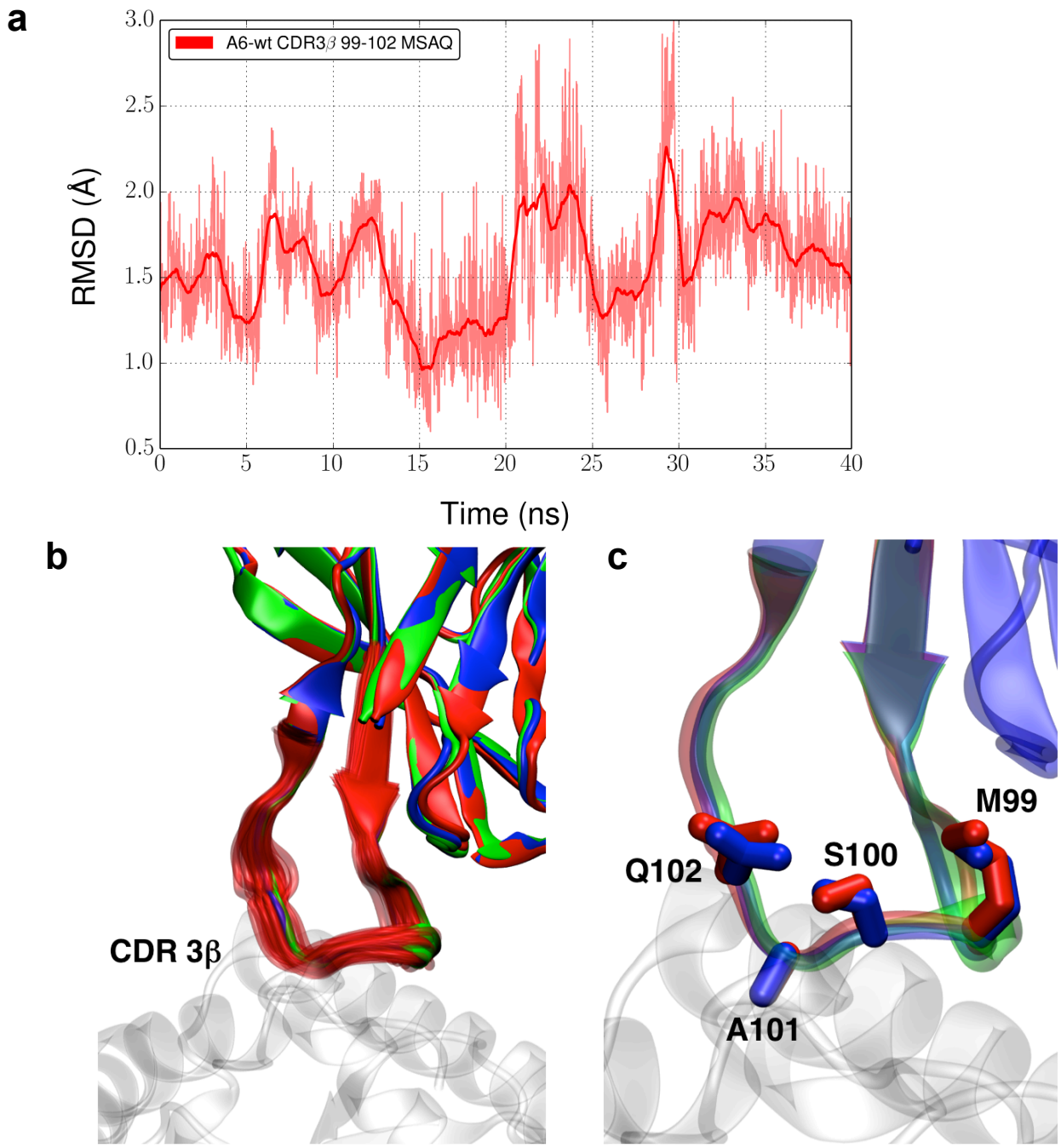
Supplementary Figure 10 | Prediction of TCR reaction specificity using Rosetta sequence tolerance algorithms. Crystal structures of the A6-c134 (4FTV) (**a**) and DMF5 (3QDG) (**b**)

were used as input for Rosetta sequence tolerance algorithms[3,4]. Structural models of the predicted 10 lowest energy level configurations were made and mutated to the other possible 19 amino acids. The average predicted frequencies of each amino at the indicated TCRβ positions in the ensemble are presented as a ranked list with range specified by color. The top 5 residues are indicated by a dotted line. According to Smith and Kortemme[3,4], 42-82% of amino acids from experimental phage display data occur above the dashed line in experimental validation datasets.



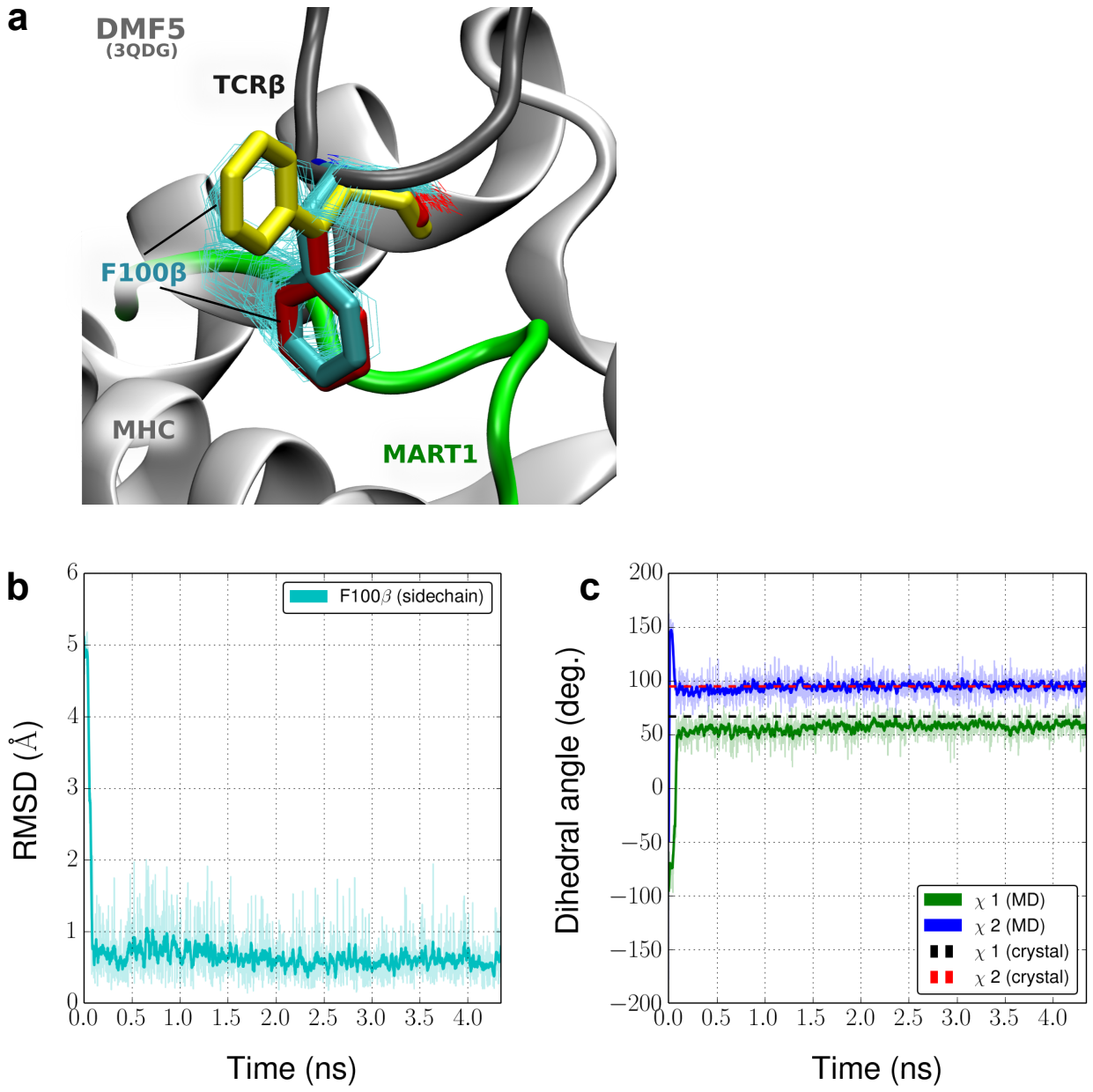
Supplementary Figure 11 | Equilibrium Binding of RD1-MART1 and RD1-MART1^{HIGH}. SPR binding of immobilized RD1-MART1 (**a**) and RD1-MART1^{HIGH} (**b**) scTCRs with MART1/HLA-A2 (selecting, non-cognate; black), Tax/HLA-A2 (non-selecting, cognate; purple), and gp100 (non-selecting, non-cognate; red). Data at each concentration are representative of two injections.

Averages and standard deviations of all SPR experiments (including those in the reverse orientation) are shown in Table 1.



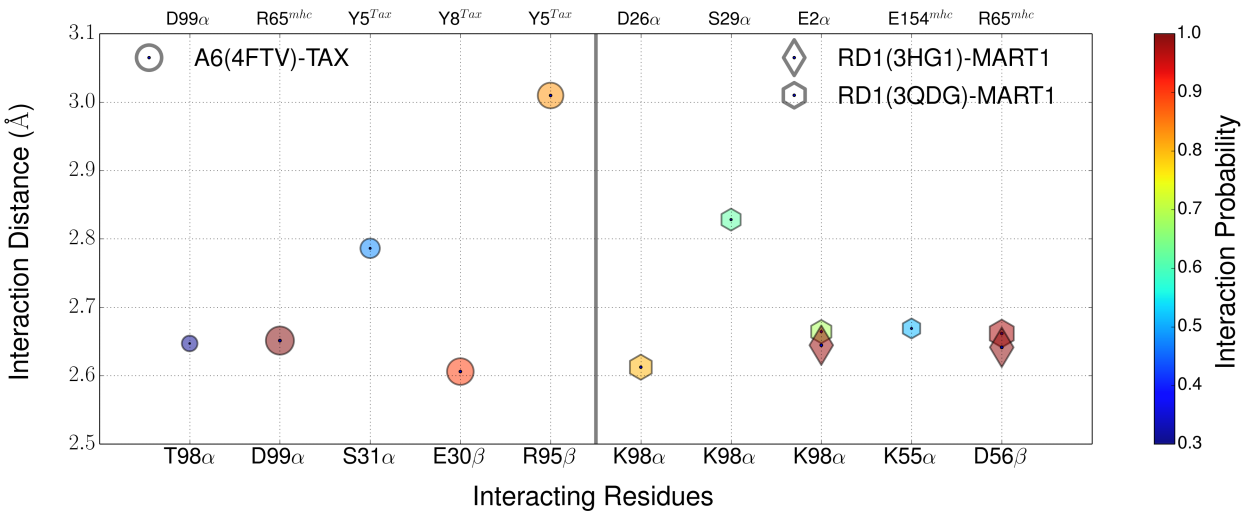
Supplemental Figure 12 | MD simulation of modeled A6-c134 CDR3β loop. (**a**) Root mean square deviation (RMSD) of the CDR3β modeled positions 99-102 (based on the A6 wild-type crystal structure, PDB: 1AO7) over 40 ns. The A6-c134 crystal structure (4FTV) was employed as the reference. An RMSD of <1.5 Å suggests a good side chain overlap with respect to the crystal structure at the indicated positions. Raw data is shown in lighter color, and

the block-averaged data is shown by the darker line. (**b**) Overlay of the backbone configurations of the A6-wt ‘MSAQ’ trajectory are shown in red, with the starting A6 wild type configuration (green) and A6-c134 crystal structure (blue). (**c**) Side chain orientations of an A6-wt ‘MSAQ’ trajectory configuration (RMSD of 0.6 Å) compared to their orientations in the A6-c134 crystal structure (4FTV).

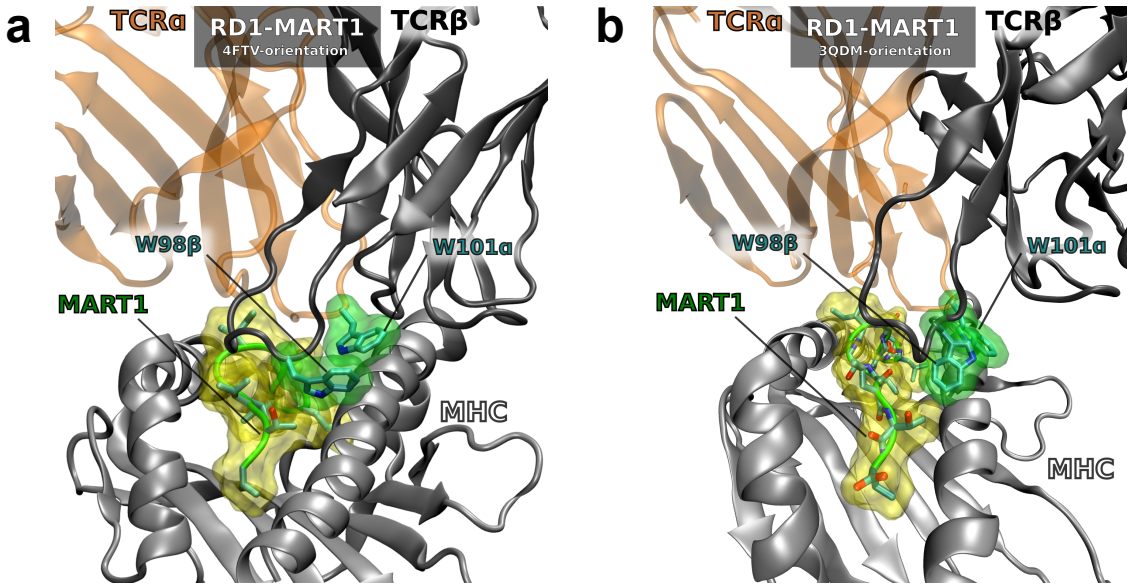


Supplemental Figure 13 | MD of DMF5 TCR Residue F100β. (**a**) Snapshots of F100β from the simulations compared with the original crystal structure conformation. The conformation of F100β from the crystal is shown in red; starting conformation shown in yellow; final conformation shown in cyan and intermediates shown in cyan with stick representation. (**b**) Root mean square deviation (RMSD) of the heavy atoms in F100β side-chain from the crystal structure orientation.

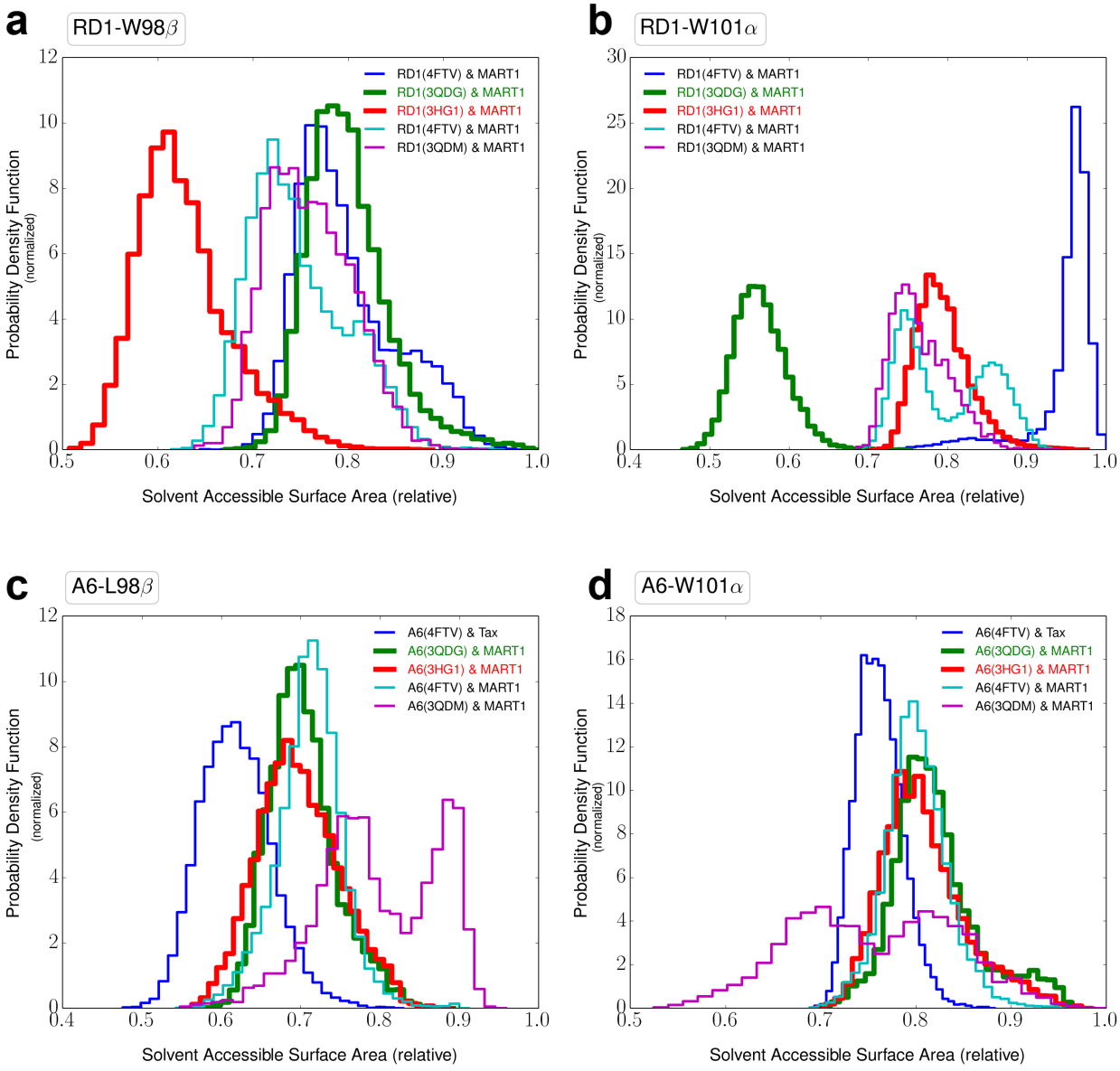
(**c**) χ1 and χ2 dihedral angles of F100β. For both (**b**) and (**c**) raw data is shown in lighter color, and the block-averaged data is shown by darker lines.



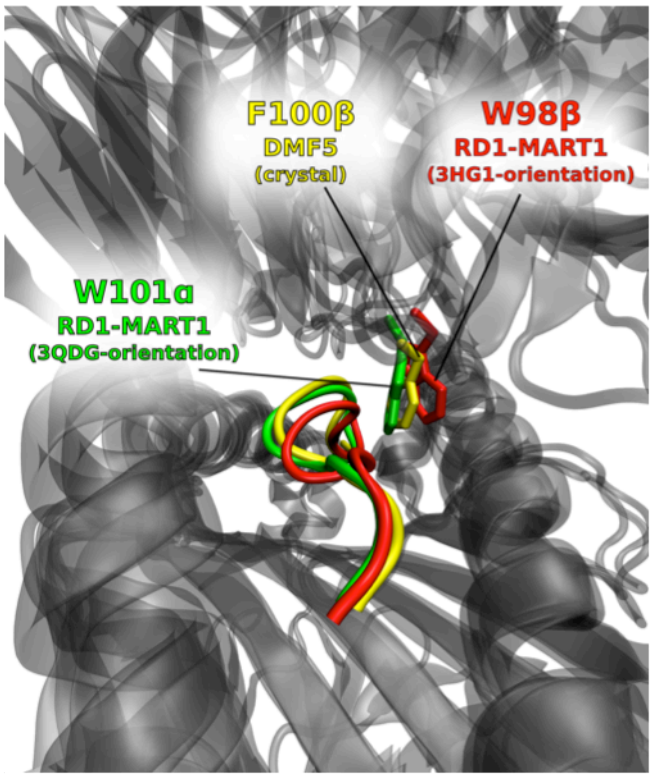
Supplementary Figure 14 | Molecular dynamics predicted interactions between TCR and peptide or MHC. The colors indicate the probability of strong interaction, which is measured by the fraction of time residues interact with a distance less than 3.5 angstroms during the simulation. The size of each marker is also proportional to the interaction probability. The optimal interaction distance during interactions is indicated on the y-axis.



Supplementary Figure 15 | Molecular dynamics of interactions of alternative RD1-MART1 orientations. No significant hydrophobic interactions exist between W98β and W101α and MART1 peptide when RD1-MART1 is in 4FTV/A6-c134-orientation (**a**) or 3QDM/DMF4-orientation (**b**).

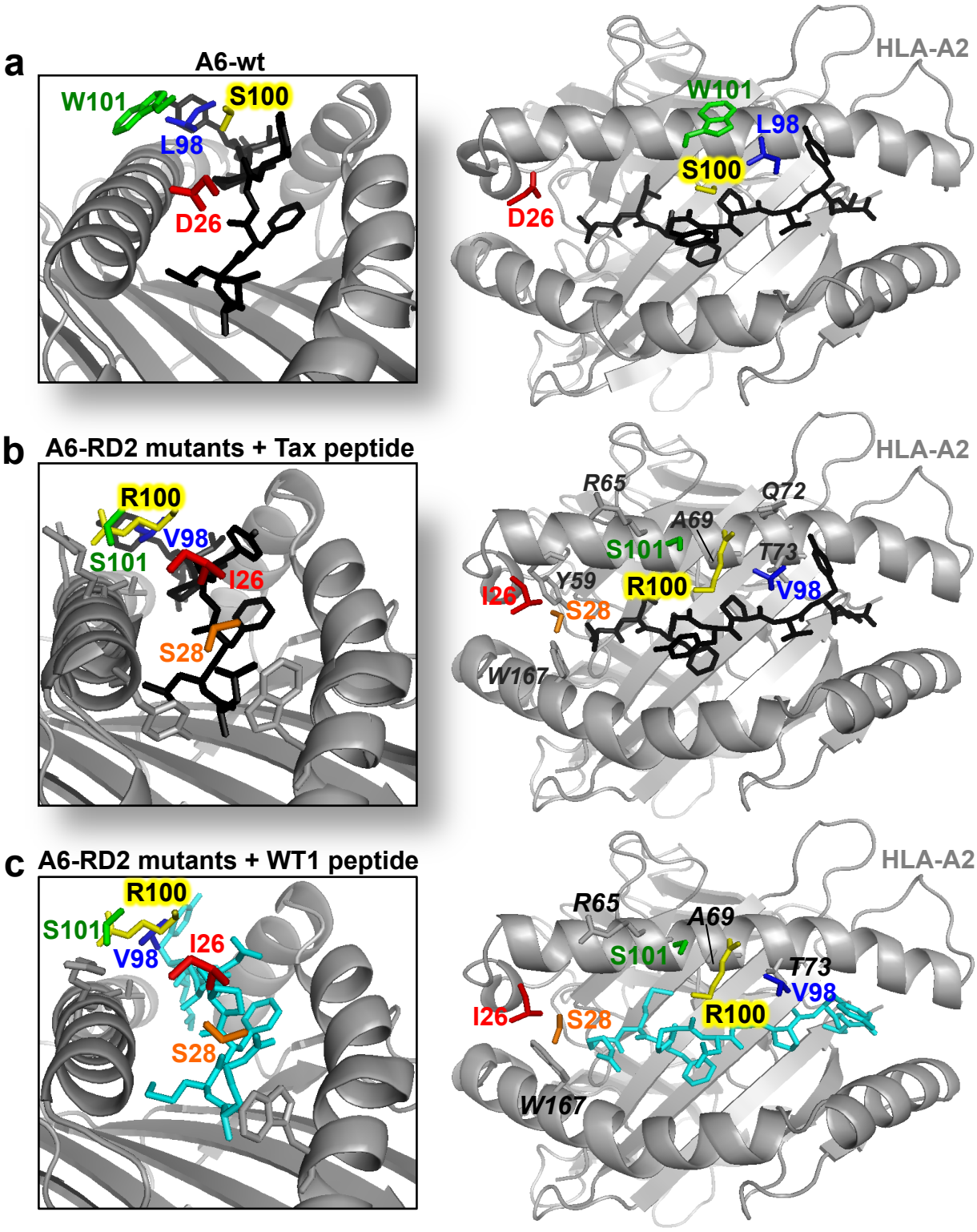


Supplementary Figure 16 | Molecular dynamics of solvent accessible surface area in modeled complexes. Hydrophobic interactions between W98β (**a**,**c**) or W101α (**b**,**d**) and the peptide-MHC complex, indicated by solvent accessible surface area (relative). A6-c134 (**c**,**d**) and RD1-MART1 TCR (**a**,**b**) are abbreviated as A6 and RD1, respectively.



Supplementary Figure 17 | Overlay of RD1-MART1 models and the DMF TCR.

Superposition of the DMF5 TCR with RD1-MART1 in the 3QDG/DMF5-orientation and RD1-MART1 in the 3HG1/Mel5-orientation, using the MHC as an alignment reference. The key aromatic residue at position 98β (for the 3HG1/Mel5-orientation) and 101α (for the 3QDG/DMF5-orientation) are highlighted in red and green, respectively. The aromatic F100β of the aligned DMF5 TCR crystal structure is in yellow. The MART1 peptide in each structure or model is shown with matched color.


a
A6-wt
W101
S100
L98
D26
HLA-A2
W101
L98
S100
D26
b
A6-RD2 mutants + Tax peptide
R100
V98
S101
I26
S28
HLA-A2
R65
Q72
A69
S101
T73
Y59
R100
V98
I26
S28
W167
c
A6-RD2 mutants + WT1 peptide
R100
S101
I26
V98
S28
HLA-A2
R65
A69
S101
T73
V98
I26
S28
R100
W167

Supplementary Figure 18 | Modeling of RD2-WT1 that yielded cross-reactivity with non-selecting ligands. Rosetta Backrub flexible backbone modeling algorithms were used to model the RD2-WT1 mutations (D26αI, red; G28αS, orange; S100αR, yellow; W101αS, green; and L98βV, blue) into the wild type A6 crystal structure (PDB: 1AO7)[1]. HLA-A2 residues within 3.2Å of RD2-WT1 residues are labeled in italics and shown as sticks (gray) (**a**) The wild type A6 crystal structure (PDB: 1AO7)[1] showing RD2-WT1 positions prior to mutation. (**b**) The A6 crystal structure (PDB: 1AO7)[1] with mutated RD2-WT1 residues. Tax peptide is shown in black. (**c**) The A6 crystal structure (PDB: 1AO7)[1] with mutated RD2-WT1 residues and mutated WT1 peptide in the HLA-A2 binding groove (cyan).

Supplementary Methods

Yeast Display Constructs and Libraries

The RD1 and RD2 library constructs consisted of the variable TCR fragments attached by the linker region GSADDAKKDAAKKDGKS[5-8] using the A6-X15 scTCR[5] as a template. N-terminal HA and C-terminal c-myc epitope tags were added to monitor for expression.

The RD1 library was synthesized by Genscript (Piscataway, NJ, USA), where regions indicated by "X" were made degenerate by NNS codons and "*" indicates a stop codon: NAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMAWYRQDPGMGLRLIHYSVGVGITDQGDVP DGYKVSRSTTEDFPLRLLSAAPSQTSVYFCASRPGXMSXQPELYFGPGTRLTVTEDLINGSAD DAKKDAAKKDGKSQKEVEQNSGPLSVPEGAIASLNCTYSDRGSXSFFWYRQYSGKSPELIMSI YSNGDKEDGRFTAQLNKASQYVSLLIRDSQPSDSATYLCAVTXXSWGKLQFGAGTQVVVTPD EQKLISEEDL**. The gene was codon optimized for both yeast and *E. coli* with 5' sequence TCT GCT AGC and 3' sequence CTC GAG ATC TGA.

For homologous recombination in yeast, pCT302 overhangs were added to the synthesized RD1 library using forward primer 5'-CAG GCT AGT GGT GGT GGT GGT TCT GGT GGT GGT GGT TCT GGT GGT GGT GGT TCT GCT AGC AAT GCT GGT GTA ACA CAA ACG CCA A-3' and reverse primer 5'-GGA ACA AAG TCG ATT TTG TTA CAT CTA CAC TGT TGT TAA CAG ATC TCG AGT CAT TAT AAA TCT TCT TCA GAG ATC-3'. Yeast libraries were generated by homologous recombination in EBY100 yeast by electroporation of PCR products along with NheI and XhoI digested yeast display plasmid, pCT302[9-12]. The resultant library size for RD1 was 6 X 10^6.

The RD2 Library was synthesized by DNA2.0 (Menlo Park, CA, USA), where positions indicated by "X" were made degenerate by NNK codons, the positions labeled "1234" were binary allowing for A6 wild type CDR3β loop AGGR or A6-X15 CDR3β loop MSAQ, the position indicated by "#" was binary allowing for either wild type residue Q or mutated T, and positions indicated by "*" were stop codons:

NAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMAWYRQDPGMGLRLIHYSVGVGITDQGDVP DGYKVSRSTTEDFPLRLLSAAPSQTSVYFCASRPGX1234PELYFGPGTRLTVTEDLINGSADD AKKDAAKKDGKSKEVEQNSGPLSVPEGAIASLNCTYSXRXS#SFFWYRQYSGKSPELIMSIYSN GDKEDGRFTAQLNKASQYVSLLIRDSQPSDSATYLCAVTTDXXGKLQFGAGTQVVVTPDIEQK LISEEDL**. The gene was codon optimized for yeast and the following flanking DNA sequences were added which contained overlap with the T7 and Splice4L cloning primers: N-terminal DNA sequence: 5' – GGC AGC CCC ATA AAC ACA CAG TAT GTT TTT AAG GAC AAT AGC TCG ACG ATT GAA GGT AGA TAC CCA TAC GAC GTT CCA GAC TAC GCT CTG CAG GCT AGT GGT GGT GGT GGT TCT GGT GGT GGT GGT TCT GGT GGT GGT GGT TCT GCT AGC – 3', and C-terminal DNA sequence: 5' – CTC GAG ATC TGT TAA CAA CAG TGT AGA TGT AAC AAA ATC GAC TTT GTT CCC ACT GTA CTT TTA GCT CGT ACA AAA TAC AAT ATA CTT TTC ATT TCT CCG TAA ACA ACA TGT TTT CCC ATG TAA TAT CCT TTT CTA TTT TTC GTT CCG TTA CCA ACT TTA CAC ATA CTT TAT ATA GCT ATT CAC TTC TAT ACA CTA AAA AAC TAA GAC AAT TTT AAT TTT GCT GCC TGC CAT ATT TCA ATT TGT TAT AAA TTC CTA TAA TTT ATC CTA TTA GTA GCT AAA AAA AGA TGA ATG TGA ATC GAA TCC TAA GAG AAT TGA GCT CCA ATT CGC CCT ATA GTG AGT CGT ATT A. The delivered PCR product was amplified via PCR using the Splice4L and T7 primers, and yeast libraries were generated by homologous recombination in EBY100 yeast as described above[9-12]. The resultant library size for RD2 was 2.4 X 10^8.

For the affinity maturation of the RD1-MART1 TCR, pre-SOE PCR products were generated for each of the four libraries utilizing the following primer pairs. β1: 5'- GGC AGC CCC ATA AAC ACA CAG TAT -3' (Splice 4L) and 5'- CGG ACG GGA AGC GCA GAA ATA CAC TGA GGT TTG AGA AGG TGC AGC GCT TAA CAG ACG CAG CGG -3', and 5'- ACC TCA GTG TAT TTC TGC GCT TCC CGT CCG <u>NNK NNK NNK NNK NNK</u> CAG CCT GAA CTG TAC TTT GGT CCA GGC ACT AGA C -3' and 5'- TAA TAC GAC TCA CTA TAG GG -3' (T7); β2: Splice 4L and 5'- CGG ACG GGA AGC GCA GAA ATA CAC TGA GGT TTG AGA AGG

TGC AGC GCT TAA CAG ACG CAG CGG -3', and 5'- ACC TCA GTG TAT TTC TGC GCT TCC CGT CCG GGT TGG NNK NNK NNK NNK NNK GAA CTG TAC TTT GGT CCA GGC ACT AGA CTG ACC G -3' and T7; α: Splice 4L and 5'- CGT AAC CGC GCA CAA GTA TGT GGC CGA ATC GGA AGG CTG GGA GTC ACG AAT CAG CAA ACT AAC ATA CTG GC -3', and 5'- TCC GAT TCG GCC ACA TAC TTG TGC GCG GTT ACG NNK NNK NNK NNK NNK AAA CTG CAA TTT GGT GCG GGC ACC CAG GTT GTG G -3' and T7. SOE PCR was performed with each corresponding Pre-SOE along with both T7 and Splice 4L for each library.

Yeast libraries were generated by homologous recombination in EBY100 yeast as described above[9-12]. Library diversity was confirmed at all 5 degenerate positions following sequencing of 6 clones from each library. The resultant library sizes were β1: 2.1 X 10^7, β2: 1.7 X 10^7, and α: 1.1 X 10^7. Libraries were pooled in equal cell numbers in ratios reflecting relative diversity, and expanded in SD-CAA media.

Rosetta Sequence Tolerance

Rosetta sequence tolerance algorithms[3,4] were used to predict the specificity of certain residues in the TCR:pepMHC interactions. The crystal structures for A6-c134 (PDB: 4FTV) and DMF4 (PDB: 3QDG) were uploaded to the Rosetta server and used to generate an ensemble of 20 structures. Interacting partners taken into account of the sequence predictions included the TCRβ chain (chain E), peptide (chain C), and HLA-A2 heavy chains (chain A), using self-interaction energies of 0.4, partner interacting energies of 1.0, and a Boltzmann Factor (kT) of 0.228, according to the recommended published values. The resultant frequencies are presented as a ranked list for the TCRβ positions indicated.

Molecular Dynamics Simulations

In order to determine if MD was a valid method to analyze possible mechanisms of specificity of the RD1 variants, two sets of validation simulations were performed. First, we determined whether MD simulations would correctly predict the backbone and side chain orientations of the CDR3β loop of high-affinity TCR A6-c134 derived from A6. For this, the A6 wild-type crystal structure (PDB: 1AO7)[1] was mutated to the CDR3β loop of A6-c134 (i.e. positions 99-102 mutated from AGGR to MSAQ). Following mutagenesis, the structure of the A6 wild-type with 'MSAQ' was subjected to a 40 ns simulation. The resultant conformations were overlaid with the experimentally characterized A6-c134 high-affinity crystal structure (PDB: 4FTV)[2] and root mean square deviations (RMSDs) were calculated (Supplementary Fig. 12a). Among the configurations sampled during the simulation, there was significant CDR3β loop-overlap between simulated orientations and the A6-c134 crystal structure (RMSD < 1.5 Å for 44.3% of the trajectory) (Supplementary Fig. 12).

In a second validation experiment, we determined if MD simulations could accurately predict the insertion of the side chain of residue F100β in the DMF5 TCR (PDB: 3QDG)[13]. In the crystal structure, this residue inserts into a hydrophobic pocket formed between the HLA-A2 α1 helix and the MART1 peptide. In this MD simulation, the conformation of F100β was initially manipulated to move this residue out of the pocket (180° rotation of χ1 dihedral angle and a 90° rotation for χ2 dihedral angle prior the start of the simulation). Residue F100β was selected due to its predicted key interaction with MART1/HLA-A2 and its similar proposed mechanism of binding for the CDR3 tryptophans (W98β - 3HG1/Mel5-orientation or W100α - 3QDG/DMF5-orientation) of RD1-MART1 (Supplementary Fig. 13 and see below). Within 0.1 ns into the simulation, the F100β re-inserts into the binding pocket of MART1/HLA-A2 and remains in this position for the remainder of the simulation (4.3 ns). This rapid reorganization suggests that the F100β side-chain insertion is a critical feature of the interaction between DMF5 and

MART1/HLA-A2. The consistency between the results of these two test simulations and the prior structural studies, as well as the success of earlier MD simulations in replicating structural and dynamic properties of TCRs and their complexes[14,15] support the validity of MD in exploring the determinants of specificity.

To examine possible mechanisms of binding and specificity involved in the TCRs selected by yeast display, four different TCR:peptide/MHC complexes were modeled and simulated: A6-c134:Tax/HLA-A2, A6-c134:MART1/HLA-A2, RD1-MART1:Tax/HLA-A2, and RD1-MART1:MART1/HLA-A2. In all the simulation systems that included Tax, the starting conformation was based on the crystal structure of A6-c134:Tax/HLA-A2[2] (PDB: 4FTV; denoted as the 4FTV/A6-c134 orientation hereafter). The RD1-MART1 TCR structure was generated by introducing point mutations at respective sites in the A6-c134 TCR. The initial structures for Tax and MART1 peptides were taken from A6-c134:Tax/HLA-A2 (PDB: 4FTV)[2] and DMF5:MART1/HLA-A2 (PDB: 3QDG)[13] structures, respectively. An additional TCR-peptide-MHC complex for A6-c134 was constructed starting from the A6 wild type crystal structure (PDB entry 1AO7), by *in silico* mutation of positions 99-102 of the CDR3β loop (AGGR) to the A6-c134 sequence (MSAQ). In this system, the orientation of the TCR is similar to the 4FTV orientation, therefore this system is also considered to start from the 4FTV/A6-c134 orientation.

In order to reduce the size of the simulation system and thereby allowing for better sampling and statistics of the dynamics and interactions at the interface, remote regions of MHC and TCR components were truncated. Thus, the MHC part was truncated to residue 1-182, and TCR α-and β-chain were truncated to residues 1-119 and 30-122, respectively. To preserve the binding orientation and scaffolding of the structures, positional restraints were applied at the truncation sites; in MHC, Cα atoms of residues 30, 32, 96, 122, and 182 were harmonically restrained (force constant k = 1.0 kcal/mol/Å^2) to their initial positions in all the simulations. Additionally, backbone carbonyl carbons of residue 119 in CDRα and residue 122 in CDRβ were harmonically restrained (force constant k = 1.0 kcal/mol/Å^2) in the following simulations:

A6-c134:Tax/HLA-A2, A6-c134:MART1/HLA-A2 (4FTV/A6-c134 orientation), RD1-MART1:Tax/HLA-A2 and RD1-MART1:MART1/HLA-A2 (4FTV/A6-c134 orientation), in order to preserve the binding orientation of the TCR subunit.

All the simulations started from the 4FTV/A6-c134 orientation. Each simulation system was minimized for 1000 steps, followed by a 500-ps MD simulation under constant-volume, constant-temperature (NVT) conditions with all the backbone atoms restrained to their initial position (k = 10.0 kcal/mole/$Å^2$). This was followed by a 4-ns MD simulation under constant pressure, constant temperature (NPT) conditions, during which the restraints were gradually decreased and eventually eliminated completely.

After the initial equilibration phase, A6-c134:Tax/HLA-A2 and RD1-MART1:Tax/HLA-A2 were simulated each for another 100 ns of production runs, while, A6-wt MSAQ was simulated for 40 ns of production run. For MART1 simulations, in addition to the 4FTV/A6-c134 orientation described above, three additional initial models were constructed and simulated. Bound MART1/HLA-A2 is present in several crystal structures in complex with other MART1/HLA-A2-specific TCRs, in which the docking orientation of the TCR on the peptide/MHC varies. In order to examine all experimentally observed orientations of MART1/HLA-A2 binding to MART1/HLA-A2-specific TCRs, in addition to the orientation based on the A6-c134:Tax/HLA-A2 structure (4FTV/A6-c134 orientation), initial models representing orientations observed in PDB entries 3QDG[13], 3QDM[13], and 3HG1[16], were also constructed using the protocol described below. We refer to these different orientations as 3QDG/DMF5, 3QDM/DMF4, and 3HG1/Mel5 orientations, respectively.

In order to generate the initial models for the three new orientations, in each case we started from the 4FTV/A6-c134 orientation described above. The model was then slowly morphed into the target orientation using biased simulations with system-specific collective variables (orientation quaternions)[17]. In the first step, a biasing potential with a force constant of k = 50,000 - 100,000 kcal/mol/$radian^2$ was applied to the TCR subunit for 40 ns. This was then

followed by a 60-ns relaxation simulation. In order to accelerate the relaxation to the new orientation, non-bonded interactions were scaled intermittently. The scaling factor was decreased from 1.0 to 0.8 and then increased back to 1.0, with a step size of 0.05. At each step, the structures were equilibrated for about 5 ns, followed by another 5 ns equilibration using a scaling factor 1.0. Once a relaxed structure was achieved, the systems were simulated for production runs: 100ns (for A6-c134:MART1/HLA-A2 in the 4FTV/A6-c134 and 3QDG/DMF5 orientations), and 170ns (for RD1-MART1:MART1/HLA-A2 in the 4FTV/A6-c134, 3QDG/DMF5, 3QDM/DMF4, and 3HG1/Mel5 orientations). The A6-c134:MART1/HLA-A2 complex in the 3HG1/Mel5 and 3QDM/DMF4 orientations was simulated for 50 ns only, since no significant binding interactions were evident for these complexes.

System preparation was done using VMD[18], and all the simulations were performed using NAMD[19]. CHARMM27 force field was used for protein and ions[20]. The simulations employed rigid bonds for bonds involving hydrogens (using SETTLE[21] and RATTLE algorithms[22]), with a 2 fs time step and periodic boundary conditions. For non-bonded interactions, a cutoff of 12 Å was used along with a switching function starting at 10 Å. Electrostatic interactions were computed using the particle mesh Ewald method[23]. Temperature was kept constant at 310 K using the Langevin method[24], with a damping coefficient of 1/ps. Pressure was maintained constant at 1.01325 bar with the isotropic Nosé-Hoover Langevin piston method[25,26], with a barostat oscillation period of 200 fs and damping time scale 50 fs.

MD SASA Analysis

The solvent accessible surface area (SASA) was calculated for the heavy atoms in 1 W101α, W98β (for RD1-MART), or L99β (for A6-134), using the backbone of Tax or MART1 and HLA-A2 as the environment. The results were normalized by the SASA of the same residue type in vacuum. The first 1 ns of the simulations for A6 systems and the first 3 ns of the simulations of the RD1-MART1 systems were excluded from this analysis.

The interaction distances were calculated using the minimum distance between the terminal heavy atoms (e.g., nitrogens in Arg and oxygens in a Glu) in each trajectory frame. Then all the resulting data were histogramed with bin widths ranging 0.1-0.2 Å. The peak location was used as the maximum likelihood interaction distance and shown in Supplementary Fig. 14. The first 3 ns of the trajectories were excluded from distance analysis.

MD Simulation and Analysis of DMF5 Position F100β

The simulation of DMF5:MART1/HLA-A2 complex was prepared using crystal structure with PDB identifier: 3QDG. TCR and MHC were truncated in order to focus on relevant regions: TCRα residues 1—108, TCRβ residues 4—116 and MHC residues 1—182. In order to assess whether the insertion of the side-chain of F100β reoccurs, a rotation of 180° for χ1 dihedral angle and another rotation of 90° for χ2 dihedral angle were made for residue F100β. The following restraints were applied: The backbone of TCR and all heavy atoms in MART1/HLA-A2 were restrained with 1 kcal/mol/ $Å^2$ force constant. The resulting simulation system was first energetically minimized for 1000 steps and then run for ~4.3 ns under constant temperature, volume (NVT) condition.

For analysis the root mean square deviation (RMSD) of the heavy atoms in the sidechain of F100β, and the χ1 and χ2 dihedral angles were measured for comparison with the crystal structure. During the simulation, the aromatic ring of F100β flipped by 180°. For better comparison with the crystal structure, the CD1/CE1 and CD2/CE2 atom names were switched in the crystal structure data file during the data analysis.

Backrub Modeling of RD2-WT1

Rosetta Backrub flexible backbone modeling algorithms (https://kortemmelab.ucsf.edu/backrub/)[27,28] were used to model the RD2-WT1 mutations (D26Iα, G28Sα, S100Rα, W101Sα, and L98Vβ) into the wild type A6 (PDB: 1AO7)[1] crystal

structure along with the wild type Tax, MART1 10mer (ELAGIGILTV), and WT1 (RMFPNAPYL) peptides. Mutated residues were given a 10Å radius of effect for the flexible backbone modeling. For modeling purposes, the residue in position “0” of the MART1 10-mer peptide was omitted from the prediction. PyMOL software was used to visualize overlays of the lowest two energy conformation of each model (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC)[29], and the measurement tool was used to determine which HLA-A2 residues were within 3Å of mutated RD2-WT1 residues.

Supplementary References

1 Garboczi, D. N. *et al.* Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134-141 (1996).

2 Cole, D. K. *et al.* Increased Peptide Contacts Govern High Affinity Binding of a Modified TCR Whilst Maintaining a Native pMHC Docking Mode. *Front Immunol* **4**, 168, doi:10.3389/fimmu.2013.00168 (2013).

3 Smith, C. A. & Kortemme, T. in *PloS one* Vol. 6 e20451 (2011).

4 Smith, C. A. & Kortemme, T. in *J Mol Biol* Vol. 402 460-474 (2010).

5 Aggen, D. H. *et al.* Identification and engineering of human variable regions that allow expression of stable single-chain T cell receptors *Protein Eng. Des. Sel.* **24**, 361-372, doi:10.1093/protein/gzq113 (2011).

6 Kieke, M. C. *et al.* Selection of functional T cell receptor mutants from a yeast surface-display library. *Proc Natl Acad Sci USA* **96**, 5651-5656 (1999).

7 Hoo, W. *et al.* Characterization of a single-chain T-cell receptor expressed in Escherichia coli. *Proc. Natl Acad. Sci. USA* **89**, 4759-4763 (1992).

8 Weber, K. S., Donermeyer, D. L., Allen, P. M. & Kranz, D. M. Class II-restricted T cell receptor engineered in vitro for higher affinity retains peptide specificity and function. *Proc. Natl Acad. Sci. USA* **102**, 19033-19038 (2005).

9 Benatuil, L., Perez, J. M., Belk, J. & Hsieh, C.-M. An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng Des Sel* **23**, 155-159 (2010).

10 Colby, D. W. *et al.* Engineering antibody affinity by yeast surface display. *Meth Enzymol* **388**, 348-358 (2004).

11 Starwalt, S. E., Masteller, E. L., Bluestone, J. A. & Kranz, D. M. Directed evolution of a single-chain class II MHC product by yeast display. *Protein Eng* **16**, 147-156 (2003).

12 Swers, J. S., Kellogg, B. A. & Wittrup, K. D. Shuffled antibody libraries created by in vivo homologous recombination and yeast surface display. *Nucleic Acids Res* **32**, e36, doi:10.1093/nar/gnh030 (2004).

13 Borbulevych, O. Y., Santhanagopolan, S. M., Hossain, M. & Baker, B. M. TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. *J. Immunol.* **187**, 2453-2463 (2011).

14 Scott, D. R., Borbulevych, O. Y., Piepenbrink, K. H., Corcelli, S. A. & Baker, B. M. Disparate degrees of hypervariable loop flexibility control T-cell receptor cross-reactivity, specificity, and binding mechanism. *J. Mol. Biol.* **414**, 385-400 (2011).

15 Scott, D. R., Vardeman, C. F., 2nd, Corcelli, S. A. & Baker, B. M. Limitations of time-resolved fluorescence suggested by molecular simulations: assessing the dynamics of T cell receptor binding loops. *Biophysical journal* **103**, 2532-2540, doi:10.1016/j.bpj.2012.10.037 (2012).

16 Cole, D. K. *et al.* Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J. Biol. Chem.* **284**, 27281-27289 (2009).

17 Fiorin, G., Klein, M. L. & Henin, J. Using collective variables to drive molecular dynamics simulations. *Mol Phys* **111**, 3345-3362, doi:Doi 10.1080/00268976.2013.813594 (2013).

18 Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J Mol Graph Model* **14**, 33-38, doi:Doi 10.1016/0263-7855(96)00018-5 (1996).

19 Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802, doi:Doi 10.1002/Jcc.20289 (2005).

20 Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* **25**, 1400-1415, doi:Doi 10.1002/Jcc.20065 (2004).

21 Miyamoto, S. & Kollman, P. A. Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J Comput Chem* **13**, 952-962, doi:Doi 10.1002/Jcc.540130805 (1992).
22 Andersen, H. C. Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations. *J Comput Phys* **52**, 24-34, doi:Doi 10.1016/0021-9991(83)90014-1 (1983).
23 Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **98**, 10089-10092, doi:Doi 10.1063/1.464397 (1993).
24 Adelman, S. A. & Doll, J. D. Generalized Langevin Equation Approach for Atom-Solid-Surface Scattering - General Formulation for Classical Scattering Off Harmonic Solids. *J Chem Phys* **64**, 2375-2388, doi:Doi 10.1063/1.432526 (1976).
25 Feller, S. E., Zhang, Y. H., Pastor, R. W. & Brooks, B. R. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. *J Chem Phys* **103**, 4613-4621, doi:Doi 10.1063/1.470648 (1995).
26 Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant-Pressure Molecular-Dynamics Algorithms. *J Chem Phys* **101**, 4177-4189, doi:Doi 10.1063/1.467468 (1994).
27 Lauck, F., Smith, C. A., Friedland, G. F., Humphris, E. L. & Kortemme, T. RosettaBackrub--a web server for flexible backbone protein structure modeling and design. *Nucleic Acids Res* **38**, W569-575 (2010).
28 Smith, C. A. & Kortemme, T. Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. *J Mol Biol* **380**, 742-756 (2008).
29 Schrodinger, LLC. *The PyMOL Molecular Graphics System, Version 1.3r1* (2010).