Legends to Supplementary Figures

Figure S1. The characteristics of the peptide anchor residues place Patr-AL in the A02 supertype. The anchor residues found for Patr-AL in this study are shown at the top. Anchor residues reported in the SYFPETHI database for other HLA-A and Patr-A allotypes are grouped by their supertype designation (Sidney et al). U denotes unclassified supertype. Anchor residues are in bold and auxillary anchors in italics. Identities with anchor and auxiliary anchor residues of Patr-AL are given in red and blue, respectively. The residues at positions 9 and 116 of the MHC sequence are shown to the right.

Figure S2. Amino acid sequences of peptides bound to Patr-AL, HLA-A*0201, and HLA-A*0207. Protein of origin and GenBank accession number are shown for each peptide. Anchor residues are boldened.

Figure S3. Alignment of the amino-acid sequences of Patr-AL, HLA-A*0201, and HLA-A*0207. The leader peptide is not included. Domain boundaries are indicated by short vertical lines above the Patr-AL sequence. Patr-AL EAL-2 is the most common allele and was used for the structure determination.

Figure S4. Stereoscopic view of electron density around the peptide ALDKATVLL contoured at 1*σ*. Peptide is colored by *B*-factor on a scale from 20 to 50 Å^2 (blue to red).

Figure S5. Details of the allotypes and peptides compared in Figure 4E. For comparison of peptide conformation, residues 1-180 of Patr-AL were aligned using PyMOL to the following structures (Protein Data Bank ID indicated in brackets): HLA-A*0201-LLFGYPVYV (1HHK), HLA-A*0201 FLWGPRALV (1QEW), HLA-A*0201 ALWGFFPVL (1B0G), HLA-A*0201 GILGFVFTL (1HHI), HLA-A*0201 TLTSCNTSV (1HHG), HLA-B*4401 (1SYV), HLA-B*4403 (1SYS), HLA-A*0101 (1W72), HLA-A*0201 (1HHJ), HLA-B*5301 (1A1O), HLA-C*0401 (1QQD), HLA-B*2705 (1A83), HLA-B*3501 (2CIK), HLA-E*0101 (1MHE), HLA-B*1501 (1XR9), HLA-Cw*0301 (1EFX), HLA-G*0101 (2D31), HLA-B*0801 (1M05), HLA-A*1101 (2HN7), HLA-A*1101 (1X7Q), HLA-A*2402 (2BCK). Root mean-square distances between Ca carbons of peptides in different structures were measured in PyMOL.

Figure S6. Phylogenetic analysis of the gene segment beginning 300bp upstream of the ATG start codon and ending in exon 2 for MHC-A and –AL. Phylogenetic reconstruction was performed using neighbor-joining, maximum-likelihood, and maximum-parsimony methods. Shown is the maximum likelihood tree, with midpoint rooting and black, grey and white circles indicating ≥80, 60 and 50% bootstrap support, respectively, with all three methods. For four nodes where one of the three methods had support of <50% the values for the remaining methods are shown (from top to bottom: neighbor-joining, parsimony, and maximum likelihood).The node support was omitted for the nodes not supported by at least two methods with values ≥50%.

Supplemental figure S1

		Peptide residue									MHC residue	
supertype	allotype	1	2	3	4	5	6	7	8	9 (C-term)	9	116
A02	Patr AL		**L***QVMI*	*D*						**LV** *IFA*	F	Y
	HLA-A*0201		**LM**				*V*			**VL**	F	Y
	HLA-A*0202		**LV**							**L**	F	Y
	HLA-A*0204		**L**							**L**	F	Y
	HLA-A*0205		*V L IMQ*				*VLA*			**L**	Y	Y
	HLA-A*0206		**V**							**V**	Y	Y
	HLA-A*0207		**L**	*D*						**L**	F	Y
	HLA-A*0214		**VQL**				*LVF*			**L**	Y	Y
	HLA-A*0217		**L**	*P*						**L**	F	Y
	HLA-A*6802		*T V*							*VL*	Y	Y
	HLA-A*6901		**VTA**	*IFLM*			*IFL*			**VL**	Y	Y
A24	HLA-A*24		**Y**			*IV*	*F*			**ILF**	S	Y
	HLA-A*2402		**YF**							**LFI**	S	Y
	Patr-A*0701		**YFMLI**							**VLMAFI**	S	L
	Patr-A*0901		**WYF**							**MLIVA**	S	Y
A01	HLA-A*01		*TS*	**DE**				*L*		**Y**	F	D
	HLA-A*3002		**YFLV**							**Y**	S	H
	HLA-A*3003		**FYIVL**							**Y**	S	H
	HLA-A*3004		*V YT QMI F*							**YML**	S	H
	HLA-A*2601		**VTILF**				*ILV*			**YF**	Y	D
	HLA-A*2602		**VTILF**				*ILV*			**YF**	Y	N
	HLA-A*2603	*E*	**VTILF**				*FILVY*			**YFML**	Y	D
A03	HLA-A*03		**LVM**	*FY*			*IMFVL*	*ILMF*		**KYF**	F	D
	HLA-A*1101		*VI FY*	*MLFYIA*				*LIYVF*		**KR**	Y	D
	HLA-A*3101		*L V YF*	*FLYW*			*LFVI*			**R**	T	D
	HLA-A*3303		*A I L FY V*							**R**	T	D
	HLA-A*6601	*ED*	**TV**							**RK**	Y	D
	HLA-A*6602	*S*	**VT**							**R**	Y	D
	HLA-A*6603	*S*	**VT**							**R**	Y	D
	HLA-A*6801	*DE*	**VT**							**RK**	Y	D
	Patr-A*0101		**FMLI**							**RK**	F	D
	Patr-A*0301		**AST***G IV*							**RK***Y*	Y	D
A01 A03	HLA-A*3001		**YF**							**L**	S	H
A01 A24	HLA-A*2902		**E**	*F*						**Y**	T	D
U	Patr-A*0401	**HKR***ACMSY*				**Y***FMRW*				**R***K*	S	D
U	Patr-A*0602			**ANS***QT*			**P**	**L***IVM*		**Y**	S	D

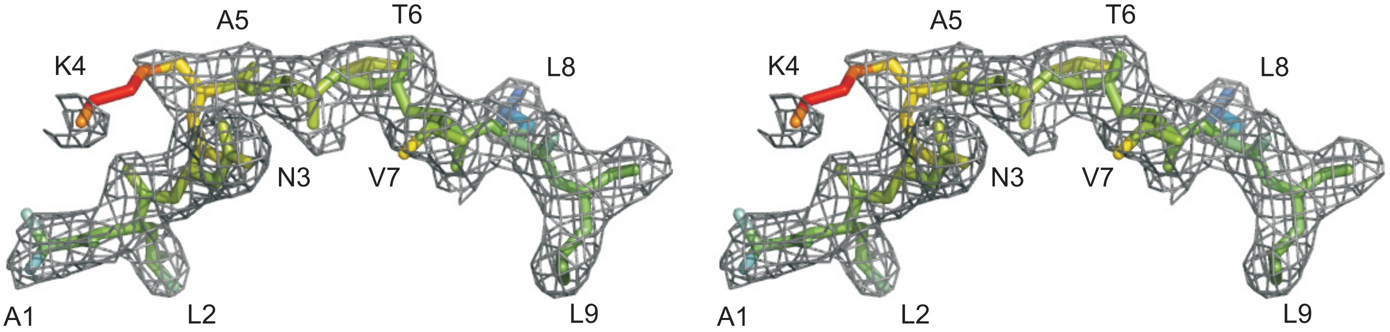
Supplemental figure S2

Accession #	Protein Description	A*0201	A*0207	Patr-AL	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	CT
P61769	β2-microglobulin	■	■	■	Y	**V**	D	D	T	Q	F	V	R	F		**F**
NM_001080856	Hypothetical protein LOC729164	■	■	■	L	**L**	M	S	F	S	V	C	L			**L**
P69905	Hemoglobin-α	■	■	■	S	**L**	D	K	F	L	A	N	V			**V**
P06733	α-enolase	■	■	■	F	**I**	K	D	Y	P	V	V	S	I		**I**
NM_000429	Methionine adenosyltransferase	■	■	■	Y	**L**	D	E	D	T	I	Y	H	L		**L**
P00558	Phosphoglycerate kinase	■	■	■	S	**L**	E	P	V	A	V	E	L			**L**
P14174	Macrophage migration inhibitory factor	■	■	■	F	**L**	S	E	L	T	Q	Q	L			**L**
P05388	Ribosomal protein P0	■	■	■	F	**L**	A	D	P	S	A	F	V	A	A	**A**
P61604	HSP10	■	■	■	V	**L**	D	D	K	D	Y	F	L			**L**
AAC50116	RAK kinase	■	■	■	T	**E**	A	D	K	S	T	V	I			**I**
P04844	Ribophorin II	■	■	■	R	**L**	D	E	L	G	G	V	Y	L		**L**
NM_002343	Transferrin	■	■	■	Y	**L**	G	D	D	Y	V	R	A	M		**M**
EAW82246	Gasdermin domain-containing 1	■	■	■	V	**I**	D	S	D	L	D	V	L	L		**L**
AAS00488	Migration-inducing gene 10 protein	■	■	■	F	**L**	A	I	L	G	G	A	K	V		**V**
NP_009191	Lysophospholipase II	■	■	■	F	**L**	E	K	L	L	P	P	V			**V**
AAS00488	Migration-inducing gene 10 protein	■	■	■	V	**L**	P	G	V	D	A	L	S	N	I	**I**
NP_003519	Histone 2B	■	■	■	F	**V**	N	D	I	F	E	R	I			**I**
NP_631898	Dipeptidylpeptidase 9	■	■	■	F	**L**	D	E	N	V	H	F	F			**F**
P49406	Mitochondrial ribosomal protein L19	■	■	■	R	**L**	D	D	S	L	L	Y	L			**L**
P35602	AP-1 complex subunit μ	■	■		A	**I**	V	D	K	V	P	S	V			**V**
AAD47141	ABC Transporter	■	■		L	**L**	D	E	P	T	N	H	L			**L**
Q9NYB9	Abl interactor 2	■	■		I	**L**	D	D	I	G	H	G	V			**V**
AAR16191	MLAA-42 antigen	■	■		I	**L**	T	D	I	T	K	G	V			**V**
Q969V6	MKL1 Protein	■	■		F	**L**	D	G	H	D	L	Q	L			**L**
NP_002113	MHC Class II DQα	■	■		F	**I**	I	Q	G	L	R	S	V			**V**
P83876	Spliceosomal protein DIM1	■	■		K	**M**	D	E	V	L	Y	S	I			**I**
BAA02430	Dipeptidase precursor	■	■		K	**V**	A	S	L	I	G	V				**V**
BAA34511	KIAA0791 protein	■	■		A	**V**	D	G	I	S	L	H	L			**L**
P13796	L-plastin	■	■		A	**L**	P	E	D	L	V	E	V			**V**
EAW95142	Chloride intracellular channel 4	■	■		F	**L**	D	G	N	E	M	T	L			**L**
NP_116231	Hypothetical protein FLJ14803	■	■		Y	**L**	F	E	R	I	K	E	L			**L**
P62906	Ribosomal protein L10a	■	■		F	**L**	A	S	E	S	L	I	K	Q	I	**I**
AAD47141	ABC Transporter	■	■		F	**L**	D	E	P	T	N	H	L			**L**
P84098	Ribosomal protein L19	■	■		I	**L**	M	E	H	I	H	K	L			**L**
BAD02827	Cytochrome P450	■	■		M	**K**	H	L	P	G	P	Q				**Q**
Q96QK1	Vacuolar sorting protein VPS35	■	■		V	**M**	A	P	H	E	P	L	V			**V**
P84098	Ribosomal protein L19	■	■		L	**M**	E	H	I	H	K	L				**L**
NP_006806	Coated vesicle membrane protein	■	■		K	**M**	G	L	I	F	E	V				**V**
AAH09524	PSMD14 protein	■	■		F	**V**	D	D	Y	T	V	R	V			**V**
Q9BYN0	Sulfiredoxin-1	■	■		T	**L**	S	D	L	R	V	Y	L			**L**
Q4KMQ2	Transmembrane protein 16F	■	■		V	**L**	D	D	K	L	V	F	V			**V**
O43818	U3 snoRNP protein 3	■	■		A	**L**	L	D	K	L	Y	A	L			**L**
O75460	IRE1 kinase precursor	■		■	T	**L**	D	G	S	L	H	A	V			**V**
CAG29343	Ras-related RAN	■		■	A	**Q**	Y	E	H	D	L	E	V			**V**
AAH29387	Hemoglobin, γ-G	■		■	H	**L**	D	D	L	K	G	T	F	A		**A**
P00558	Phosphoglycerate kinase	■		■	A	**L**	M	D	E	V	V	K	A			**A**
P50579	Methionine aminopeptidase 2	■		■	A	**Q**	F	E	H	T	I	L	L			**L**
P00338	Lactate dehydrogenase variant A	■		■	G	**L**	Y	G	I	K	D	D	V	F	L	**L**
Q13151	Ribosomal protein A0		■	■	A	**V**	D	G	N	T	V	E	L			**L**
NP_006754	B cell translocation gene 2		■	■	K	**M**	D	P	I	I	S	R	V			**V**
P13639	eEF2		■	■	K	**M**	D	R	A	L	L	E	L			**L**
P08708	Ribosomal protein S17		■	■	A	**L**	D	Q	E	I	I	E	V			**V**
EAW79040	Coatomer protein complex subunit β2		■	■	H	**L**	D	R	T	M	Y	L	L			**L**
CAI14173	Chaperonin CCT3		■	■	A	**L**	D	D	M	I	S	T	L			**L**
Q02809	Lysyl hydroxylase 1		■	■	A	**L**	D	E	V	V	L	K	F			**F**
Q99567	Numan nucleoporin 88		■	■	I	**L**	D	P	H	V	V	L	L			**L**
NP_031399	Human suppressor of clear homolog (SOC-2)		■	■	R	**L**	D	S	L	T	T	L	Y	L		**L**
P19237	Troponin I	■			Y	**L**	A	E	R	I	P	T	L			**L**
CAA60786	Proteasome subunit LMP7	■			K	**V**	I	E	N	P	Y	L				**L**
P42696	RNA-binding protein 34	■			V	**L**	F	E	N	T	D	S	V	H	L	**V**
CAG38481	DSIPI protein	■			S	**L**	L	G	G	D	V	V	S	V		**V**
Q92621	Nucleoporin NUP205	■			A	**L**	L	D	R	I	V	S	V			**V**
Q04637	eIF4g	■			V	**L**	M	T	E	D	I	K	L			**L**
P14174	Macrophage migration inhibitory factor	■			F	**L**	S	E	L	T	Q	Q	L	A		**A**
BAA95563	TRP channel-related 7	■			R	**V**	A	D	V	I	A	Q	V			**V**
P11216	Glycogen phosphorylase	■			N	**L**	A	E	N	I	S	R	V			**V**
CAA88733	Helicase	■			Y	**L**	D	Q	T	L	P	R	A			**A**
CAI13665	GTP-binding protein 4	■			T	**L**	T	E	E	G	V	I	K	V		**V**
P14923	Junction plakoglobin	■			N	**L**	M	E	Q	P	I	K	V			**V**
NP_057146	Hypothetical protein LOC51647	■			T	**L**	I	G	L	S	I	K	V			**V**
P20591	Interferon-induced protein Mx	■			T	**L**	I	D	L	P	G	I	T	R	V	**V**
P04035	HMG-CoA Reductase		■		F	**L**	D	K	E	L	T	G	L			**L**
AAG44629	Dendritic cell protein DC23		■		I	**I**	D	G	V	K	V	Q	V			**V**
NP_008835	DNA-PK		■		L	**L**	D	E	N	N	V	S	S	Y	L	**L**
AAB58363	Lysyl hydroxylase isoform 2		■		A	**V**	**D**	E	V	V	L	K	F			**F**
EAW97938	Serine/threonine phosphatase 1, γ subunit		■		N	**I**	**D**	S	I	I	Q	R	L			**L**
P49720	Proteasome subunit beta type 3		■		H	**L**	**F**	E	T	I	S	Q	A			**A**
AAH32618	Ring finger protein 20		■		H	**L**	**D**	E	A	R	T	L	L			**L**
AAG00606	RP42 protein		■		A	**L**	**D**	P	A	S	I	S	V			**V**
NP_002255	Karyopherin alpha		■		S	**V**	**D**	D	I	V	K	G	I			**I**
AAH56687	Cancer antigen SDCCAG1		■		A	**V**	**D**	E	F	Y	S	K	I			**I**
Q92538	Golgi brefeldin resistance factor 1		■		F	**V**	**D**	P	N	G	K	I	S	L		**L**
P04406	Glyceraldehyde-3-phoshphate dehydrogenase		■		G	**L**	**M**	T	T	V	H	A	I			**I**
NP_006707	Activator of S-phase kinase		■		V	**V**	**D**	D	I	V	S	K	L			**L**
CAI95003	KIAA1432 protein		■		N	**L**	**D**	N	A	F	A	H	V			**V**
P04406	Glyceraldehyde-3-phoshphate dehydrogenase		■		G	**L**	**M**	T	T	V	H	A	I	T	A	**A**
NP_037471	Asparagine-linked glycosylation 6		■		K	**L**	**P**	D	[	I	L	]	F	S	V	**V**
AAG44629	PHD finger-containing 11		■		L	**L**	**P**	T	G	V	F	Q	V			**V**
AAP74806	Proliferation inducing gene 5		■		L	**L**	**D**	G	V	V	E	K	L			**L**
Q5SW96	LDLR adaptor protein 1		■		G	**L**	**D**	E	A	F	S	R	L			**L**
AAB37682	GCN5L1 protein		■		L	**V**	**D**	H	L	N	V	G	V			**V**
Q9BV68	RING finger protein 126		■		G	**L**	**D**	A	I	I	T	Q	L			**L**
Q92786	Homeobox prospero-like PROX1			■	A	**L**	**D**	K	H	I	L	G	F			**F**
Q9P2M7	Cingulin			■	A	**Q**	**D**	P	T	M	L	Q	F			**F**
Q9NRL3	Zinedin			■	A	**V**	**D**	P	N	G	A	F	L	M		**M**
Q13588	GRB2-related adapter protein			■	A	**Q**	**D**	P	S	Q	L	S	F			**F**
P23771	Transcription factor GATA-3			■	A	**Q**	**Y**	P	L	P	E	E	V			**V**
NP_055271	Programmed cell death 4			■	A	**L**	**D**	K	A	T	V	L	L			**L**
AAH48261	Kinase CDC42BPB			■	A	**L**	**Q**	K	E	I	L	M	L			**L**
CAG33546	Trimethyllysine hydroxylase epsilon			■	R	**L**	**D**	E	T	T	L	F	F			**F**
AAH14143	Histone acetyltransferase MYST4			■	V	**Q**	**D**	G	S	V	L	K	V			**V**
AAH42999	ASXL2 protein			■	R	**Q**	**V**	G	P	D	G	L	M			**M**
NP_009057	Valosin containing protein			■	A	**V**	**E**	F	K	V	V	E	T	D		**D**
Q99622	Homeobox C10 protein			■	A	**Q**	**D**	P	E	I	A	S	L			**L**
NP_055531	Centaurin b1			■	K	**L**	**D**	S	H	A	E	L	L			**L**
P61604	HSP10			■	G	**D**	**I**	L	G	K	Y	V	D			**D**
P41091	eIF2G			■	R	**Q**	**D**	L	T	T	L	D	V			**V**
CAA50726	Topoisomerase IIb			■	G	**L**	**D**	K	D	E	Y	T	F			**F**
EAX04806	Hypothetical protein FLJ20035			■	S	**V**	**G**	I	G	P	A	R	F			**F**
Q8TBE7	Transmembrane protein TMEM22			■	A	**L**	**D**	K	F	H	P	A	L			**L**
P20073	Annexin A7			■	G	**Q**	**Y**	P	Y	P	S	G	F			**F**
EAW69005	CD23			■	G	**Q**	**Y**	S	E	I	E	E	L			**L**
P46695	Radiation-inducible gene product PRG1			■	G	**Q**	**H**	G	L	E	E	D	F	M		**M**
P52630	Transcription factor STAT2			■	S	**Q**	**D**	P	E	S	L	L	L			**L**
Q13614	Myotubularin related protein 2			■	R	**Q**	**F**	P	T	A	F	E	F			**F**
Q9P2D0	Inhibitor of Burton tyrosine kinase			■	A	**L**	**N**	R	E	I	L	F	V			**V**
Q15274	Quinolinate phosphoribosyl transferase			■	A	**L**	**D**	F	S	L	K	L				**L**
O95336	6-phosphogluconolactonase			■	S	**L**	**F**	P	D	H	P	L	L			**L**
O60287	Nuclear preribosomal-associated protein 1			■	V	**L**	**D**	Q	K	I	L	L	L			**L**
EAW53437	POGZ protein			■	R	**Q**	**F**	S	T	P	F	Q	L			**L**
NP_036465	C-myc binding protein			■	G	**M**	**D**	V	A	A	S	E	F	F		**F**
P15498	Protooncogene Vav			■	K	**M**	**D**	R	Y	A	F	L	L			**L**
P30043	Flavin reductase			■	G	**Q**	**D**	A	V	I	V	L	L			**L**
NP_11558	Nucleotide-binding oligomerization domain 27			■	R	**L**	**D**	L	S	H	L	L	L			**L**
AAS75801	Phosphoinositide-3 kinase, gamma polypeptide			■	R	**Q**	**D**	M	L	I	L	Q	I			**I**
AAH03099	DPH1 protein			■	A	**Q**	**D**	F	R	V	L	Y	V			**V**

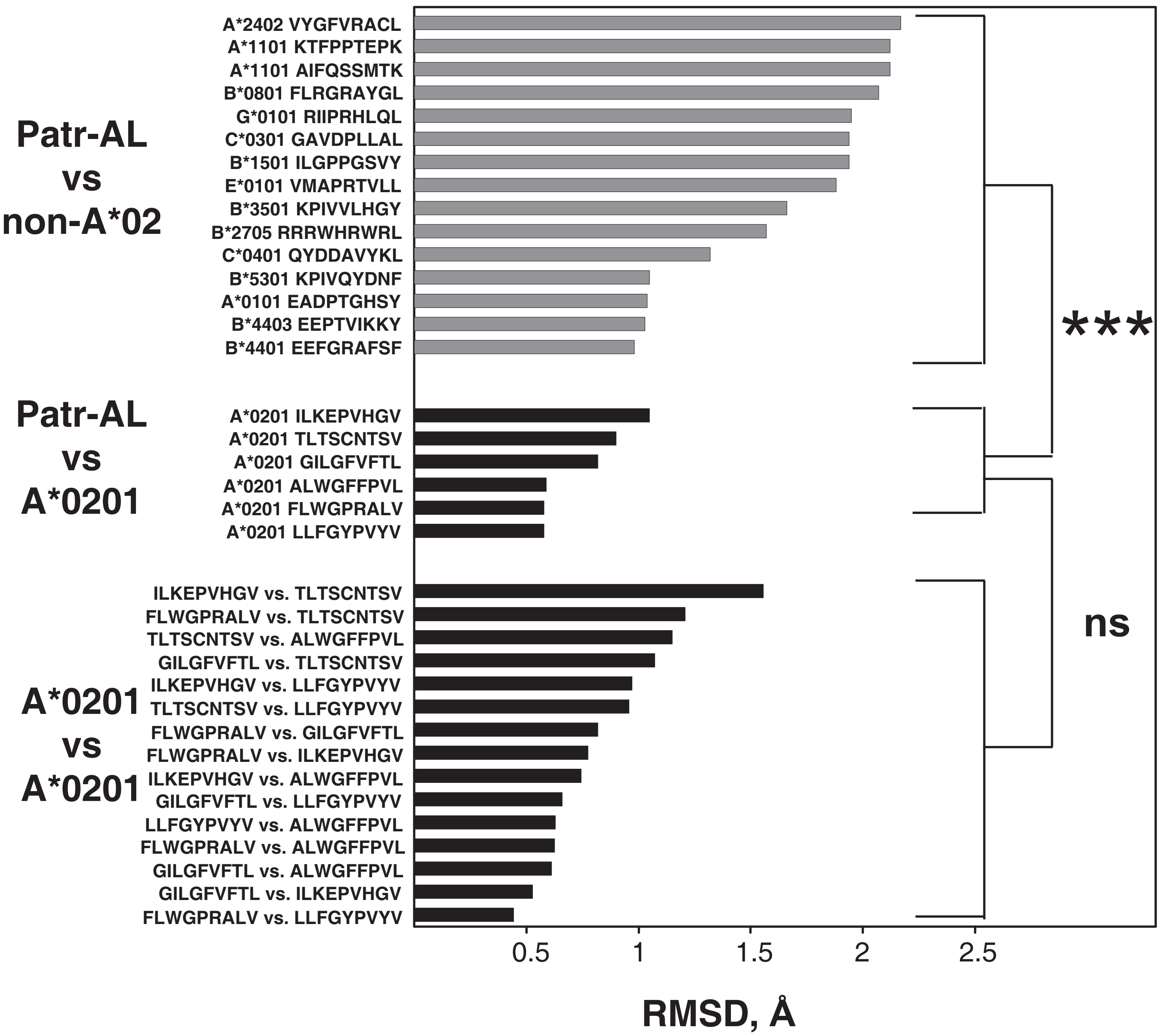
Supplemental figure S3

```
                 1                                                                                                  90
                 | a1                                                                                                |
Patr-AL EAL-2    GSHSMRYFFT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASKRMEPRAP WMEQEEPEYW DQETQISKVY AQNDRVNLGT LRGYYNQSEA
Patr-AL EAL-1    .......... .......... .......... .......... .......... .......... .......... .......... ..........
Patr-AL FRAL-1   .......... .......... .......... .......... .......... .I........ .......... .......... ..........
HLA-A*0201       .......... .......... .......... .......... ..Q........ I...G.... .G..RKV.AH S.TH..D... ..........
HLA-A*0207       .......... .......... .......... .......... ..Q....... .I...G.... .G..RKV.AH S.TH..D... ..........

                 91                                                                                                180
                 | a2                                                                                                |
Patr-AL EAL-2    GSHTIQLMYG CDVGPDGRFL RGYRQYAYDG KDYIALNEDL RSWTAADMAA RITKRKWEAA RRAEQLRVYL KGRCVDWLRR YVENGKEMLQ
Patr-AL EAL-1    D......... .......... .......... .......... .......... .......... .......... .......... ..........
Patr-AL FRAL-1   .......... .......... .......... .......... .......... .......... .......... .......... ..........
HLA-A*0201       ....V.R... ....S.W... ...H...... ......K... .......... QT..H..... HV.....A.. E.T..E.... .L.....T..
HLA-A*0207       ....V.R.C. ....S.W... ...H...... ......K... .......... QT..H..... HV.....A.. E.T..E.... .L.....T..

                 181                                                                                               270
                 | a3                                                                                                |
Patr-AL EAL-2    RTDAPKTRM THHPVSDRETTL RCWALGFYP AEITLTWQRD GEDQTQDTELV ETRPAGDGT FQKWAAVVVPS GKEQRYTCL VQHEGLPEPL
Patr-AL EAL-1    ......... ............ ......... .......... ........... ......... ........... ......... ..........
Patr-AL FRAL-1   ......... ............ ......... .......... ........... ......... ........... ......... ..........
HLA-A*0201       .......H. ...A...H.A.. .....S... .......... ........... ......... ........... .Q......H .......K..
HLA-A*0207       .......H. ...A...H.A.. .....S... .......... ........... ......... ........... .Q......H .......K..

                 281                                                                     341
                 | Stem         | Transmembrane           | Cytoplasmic                    |
Patr-AL EAL-2    TLRWEPSSQP TIPIVGIIAG LVLFVAVITG AVVAAVMWRR KSSDRKGGSY FQAASNDSSQ GSEVSLTACKV
Patr-AL EAL-1    .......... .......... .......... .......... .......... .......... ...........
Patr-AL FRAL-1   .......... .......... .......... .......... .......... .......... ...........
HLA-A*0201       .......... .......... ....G..... .......... .......... S....S..A. ..D........
HLA-A*0207       .......... .......... ....G..... .......... .......... S....S..A. ..D........
```

Supplemental figure S4



Supplemental figure S5



Supplemental figure S6