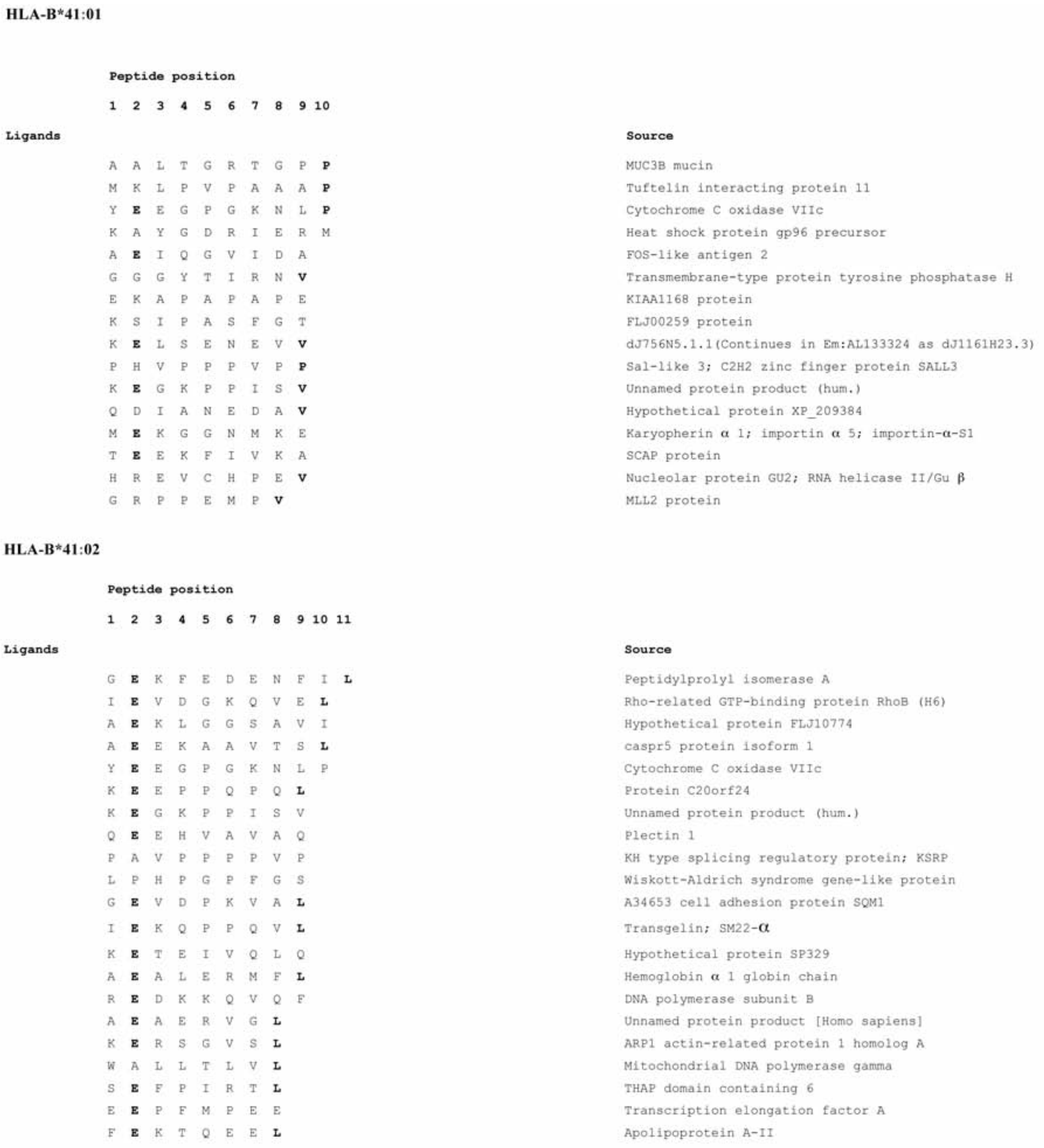
The impact of human leukocyte antigen (HLA) micropolymorphism on ligand specificity within the HLA-B*41 allotypic family

Christina Bade-Döding,[1,2]* Alex Theodossis,[3]* Stephanie Gras,[3] Lars Kjer-Nielsen,[2] Britta Eiz-Vesper,[1] Axel Seltsam,[4] Trevor Huyton,[1] Jamie Rossjohn,[3] James McCluskey,[2#] and Rainer Blasczyk[1#]
*joint first authors; #joint senior authors

[1]Institute for Transfusion Medicine, Hannover Medical School, Hannover, Germany; [2]Department of Microbiology and Immunology, University of Melbourne, Parkville, Australia; [3]Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, Monash University, Clayton, Australia, and [4]German Red Cross Blood Donor Service NSTOB, Institute Springe, Springe, Germany

*Citation: Bade-Döding C, Theodossis A, Gras S, Kjer-Nielsen L, Eiz-Vesper B, Seltsam A, Huyton T, Rossjohn J, McCluskey J, and Blasczyk R. The impact of human leukocyte antigen (HLA) micropolymorphism on ligand specificity within the HLA-B*41 allotypic family. Haematologica 2011;96(1):110-118. doi:10.3324/haematol.2010.030924*

Online Supplementary Table S1. Ligands of the HLA-B*41 variants.



HLA-B*41:01

	Peptide position										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
Ligands											**Source**
	A	A	L	T	G	R	T	G	P	**P**	MUC3B mucin
	M	K	L	P	V	P	A	A	A	**P**	Tuftelin interacting protein 11
	Y	**E**	E	G	P	G	K	N	L	**P**	Cytochrome C oxidase VIIc
	K	A	Y	G	D	R	I	E	R	M	Heat shock protein gp96 precursor
	A	**E**	I	Q	G	V	I	D	A		FOS-like antigen 2
	G	G	G	Y	T	I	R	N	**V**		Transmembrane-type protein tyrosine phosphatase H
	E	K	A	P	A	P	A	P	E		KIAA1168 protein
	K	S	I	P	A	S	F	G	T		FLJ00259 protein
	K	**E**	L	S	E	N	E	V	**V**		dJ756N5.1.1(Continues in Em:AL133324 as dJ1161H23.3)
	P	H	V	P	P	P	V	P	**P**		Sal-like 3; C2H2 zinc finger protein SALL3
	K	**E**	G	K	P	P	I	S	**V**		Unnamed protein product (hum.)
	Q	D	I	A	N	E	D	A	**V**		Hypothetical protein XP_209384
	M	**E**	K	G	G	N	M	K	E		Karyopherin α 1; importin α 5; importin-α-S1
	T	**E**	E	K	F	I	V	K	A		SCAP protein
	H	R	E	V	C	H	P	E	**V**		Nucleolar protein GU2; RNA helicase II/Gu β
	G	R	P	P	E	M	P	**V**			MLL2 protein

HLA-B*41:02

	Peptide position											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
Ligands												**Source**
	G	**E**	K	F	E	D	E	N	F	I	**L**	Peptidylprolyl isomerase A
	I	**E**	V	D	G	K	Q	V	E	**L**		Rho-related GTP-binding protein RhoB (H6)
	A	**E**	K	L	G	G	S	A	V	I		Hypothetical protein FLJ10774
	A	**E**	E	K	A	A	V	T	S	**L**		caspr5 protein isoform 1
	Y	**E**	E	G	P	G	K	N	L	P		Cytochrome C oxidase VIIc
	K	**E**	E	P	P	Q	P	Q	**L**			Protein C20orf24
	K	**E**	G	K	P	P	I	S	V			Unnamed protein product (hum.)
	Q	**E**	E	H	V	A	V	A	Q			Plectin 1
	P	A	V	P	P	P	P	V	P			KH type splicing regulatory protein; KSRP
	L	P	H	P	G	P	F	G	S			Wiskott-Aldrich syndrome gene-like protein
	G	**E**	V	D	P	K	V	A	**L**			A34653 cell adhesion protein SQM1
	I	**E**	K	Q	P	P	Q	V	**L**			Transgelin; SM22-α
	K	**E**	T	E	I	V	Q	L	Q			Hypothetical protein SP329
	A	**E**	A	L	E	R	M	F	**L**			Hemoglobin α 1 globin chain
	R	**E**	D	K	K	Q	V	Q	F			DNA polymerase subunit B
	A	**E**	A	E	R	V	G	**L**				Unnamed protein product [Homo sapiens]
	K	**E**	R	S	G	V	S	**L**				ARP1 actin-related protein 1 homolog A
	W	A	L	L	T	L	V	**L**				Mitochondrial DNA polymerase gamma
	S	**E**	F	P	I	R	T	**L**				THAP domain containing 6
	E	**E**	P	F	M	P	E	E				Transcription elongation factor A
	F	**E**	K	T	Q	E	E	**L**				Apolipoprotein A-II

continued on next page

continued from previous page

HLA-B*41:03

	Peptide position																
Ligands	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Source
	A	**E**	M	Y	G	S	V	T	E	H	P	S	P	S	P	**L**	Heterogeneous nuclear ribonucleoprotein C1/C2
	H	**E**	R	F	P	F	E	I	V	K	M	E	F				KIAA0820 protein
	F	**E**	K	T	Q	E	E	L	T	P	F	F					Apolipoprotein A-II
	F	**E**	P	L	N	K	P	D	S	T	I	Q					Ubiquitin-specific protease 15
	G	**E**	K	F	E	D	E	N	F	I	**L**						Peptidylprolyl isomerase A
	F	I	V	N	C	T	V	N	V	Q	D						Unnamed protein product (hum.)
	I	**E**	V	D	G	K	Q	V	E	**L**							Rho-related GTP-binding protein RhoB (H6)
	G	**E**	E	H	F	K	G	L	V	**L**							Bovine albumin
	E	K	L	F	E	E	K	E	K	P							OG-2 homeodomain protein-like
	R	**E**	L	E	E	T	N	Q	K	**L**							CDR2_HUMAN Cerebellar degeneration-related protein 2
	K	A	Y	G	D	R	I	E	R	M							Heat shock protein gp96 precursor
	K	**E**	D	D	V	R	Q	Y	V	V							Unnamed protein product (hum.)
	K	**E**	N	D	F	D	R	L	V	**L**							Unnamed protein product (hum.)
	K	**E**	M	A	M	L	Q	K	K	I							GRIP coiled-coil protein GCC185 isoform a
	I	N	E	D	R	L	P	H	L	**L**							Signal transducing adaptor molecule 2B
	Y	**E**	D	E	D	S	L	K	T	**L**							DNA topoisomerase II, alpha isozyme
	A	**E**	Y	I	Q	K	N	V	Q	**L**							KIAA1901 protein
	Q	D	I	N	Q	D	N	L	Q	**L**							Cardiac ankyrin repeat protein;
	K	**E**	K	S	Q	L	V	R	P	**L**							Cyclin B3
	N	**E**	Q	G	D	R	V	Y	T	**L**							Nucleolar protein family A, member 3
	G	**E**	D	Y	D	E	R	V	L	P							Cytochrome P450 4F2
	K	**E**	D	E	D	D	K	T	V	**L**							Similar to KIAA1509 protein
	Q	L	Q	S	L	T	L	F	L	Q							Mitochondrial intermediate peptidase
	K	**E**	E	G	P	K	E	M	T	**L**							Adenvlate cvclase 6 isoform a
	M	**E**	D	D	G	Q	P	R	T	**L**							TIA1 cytotoxic granule-associated RNA-binding protein Like 1 isoform 1
	V	**E**	D	D	R	T	L	Q	S	**L**							ELAVL4 protein
	M	V	E	I	S	R	T	Q	P	**L**							Tumor necrosis factor receptor superfamily
	V	**E**	I	E	E	R	G	V	K	**L**							Septin-2 (Protein NEDD5)
	H	**E**	D	I	K	K	G	I	L	**L**							MCM4_HUMAN DNA replication licensing factor MCM4
	L	L	W	S	V	T	V	S	I	F							ALKB_HUMAN alkylated DNA repair protein alkB homolog
	S	**E**	V	N	P	N	T	R	V	M							Hypothetical protein FLJ20323
	K	**E**	E	P	S	N	N	V	K	**L**							Hypothetical protein KIAA0612
	K	**E**	I	E	S	T	K	G	E	**L**							Unnamed protein product (hum.)
	K	**E**	K	N	G	D	V	V	E	**L**							Hypothetical protein (hum. herpesvirus 7)
	G	**E**	V	D	V	E	Q	H	T	**L**							Cyclin B1
	S	**E**	Y	P	I	V	D	G	K	**L**							3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2
	N	S	R	V	Q	D	A	S	K	**L**							Unnamed protein product (hum.)
	A	**E**	A	P	V	Q	E	E	K	**L**							Jumping translocation breakpoint; PAR protein
	A	**E**	N	G	F	L	P	I	H	**L**							Claudin-18A2.1
	G	A	E	P	K	P	K	P	Q	**L**							HS1-binding protein 3, iso. 1
	E	A	E	A	V	M	L	G	Q	P							Calcineurin-binding protein calsarcin-2
	L	S	E	G	A	L	A	I	E	V							bA93B14.1.1 solute carrier fam. 21 org. anion transp.
	A	**E**	Q	Q	G	K	G	V	A	**L**							IF4G_HUM. euk. transl. initiation factor 4γ (eIF-4-γ)
	E	S	A	A	V	G	S	V	V	**L**							Filamin C, gamma (actin-binding protein-280)
	G	T	G	A	S	G	S	F	K	**L**							H1 histone family, member 5
	S	**E**	A	P	P	T	N	Q	A								FLJ00163 protein
	A	**E**	D	P	A	G	L	K	V								Major capsid protein (hum. herpesvirus 4)
	S	N	S	P	S	G	N	N	I								dJ309K20.1.1 (nov. protein sim. to dysferlin, iso. 1)
	G	**E**	G	K	P	T	P	A	**L**								IMMT_HUM. mitoch. membr. protein (Mitofilin) (p87/89)
	G	**E**	A	P	P	G	T	P	P								R28379_3
	G	**E**	V	D	P	K	V	A	**L**								A34653 cell adhesion protein SQM1
	G	**E**	I	I	G	V	K	V	**L**								Similar to KIAA0606 protein
	A	**E**	V	V	V	K	E	G	**L**								Tolloid-like 2
	A	G	L	Q	F	P	V	G	R								AATK protein
	I	**E**	D	G	K	V	V	T	V								JC2186 hippocalcin
	A	**E**	A	S	M	I	S	K	**L**								Hypothetical protein KIAA0052
	G	**E**	E	G	Y	G	R	S	**L**								Anaphase-promoting complex subunit 5
	F	D	S	D	A	T	S	P	R								Axonal transp. synaptic vesicles; kinesin, hc, 1A
	A	**E**	N	E	A	E	T	K	**L**								Glutamyl-prolyl tRNA synthetase; glutamate tRNA lig.
	T	**E**	D	I	P	V	K	T	**L**								Isoleucyl-tRNA synthetase, cytoplasmic
	V	**E**	D	L	A	R	I	S	**L**								Unnamed protein product (hum.)
	V	**E**	G	K	D	L	P	E	H								S22765 heterogeneous ribonuclear particle protein U
	A	**E**	E	L	G	H	R	D	V								Hypothetical protein FLJ40243
	E	D	K	I	P	L	L	V	V								Similar to melanoma chondroitin sulfate proteoglycan
	A	**E**	I	N	N	I	I	K	I								KIAA0853 protein
	E	P	Q	N	A	D	K	I	K								Muscle-specific serine kinase 1
	I	**E**	E	N	T	G	H	T	F								ASC-1 complex subunit P200
	G	**E**	E	L	Q	M	E	P	V								Hypothetical protein XP_291339
	D	T	V	K	V	K	E	S	F								Probable ribosome biogenesis protein
	K	**E**	V	N	F	T	S	E	**L**								Targeting protein for Xklp2
	M	L	L	K	G	H	K	K	**L**								Hypothetical protein (hum.)
	K	**E**	Y	V	G	I	V	R	**L**								Dyskerin
	P	G	Y	D	S	I	I	Y	R								HT028

continued on next page

continued from previous page

K	**E**	T	E	I	V	Q	L	Q	Hypothetical protein SP329
E	T	V	N	R	I	M	T	**L**	Laminin β 1 related protein
A	**E**	A	L	E	R	M	F	**L**	Hemoglobin α 1 globin chain
K	**E**	I	D	K	N	D	H	**L**	Hypothetical protein FLJ32343
A	**E**	K	E	F	N	E	T	M	Jumonji domain containing 2
K	**E**	K	P	D	I	F	Q	**L**	SACS_HUMAN sacsin
E	**E**	Q	Q	K	Q	Q	L	Q	Hypothetical protein LOC139135
H	**E**	Q	E	E	I	F	K	**L**	Similar to tousled-like kinase 2
E	K	L	E	R	K	E	I	Q	Protein kinase C β (5' partial) splice form 1
F	**E**	D	E	E	I	K	R	**L**	Osteoblast specific factor 2 (fasciclin I-like)
Y	**E**	E	I	F	V	K	N	M	Tissue factor pathway inhibitor
R	**E**	E	Y	P	D	R	I	M	Unnamed protein product (hum.)

HLA-B*41:04

	Peptide position															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	
Ligands																**Source**
	T	W	Q	R	D	G	E	D	Q	T	Q	D	T	E	**L**	Unnamed protein product (hum.)
	A	**E**	Y	G	R	L	G	L	G	E	G	A	E	E	K	Regulator of chromosome condensation
	A	**E**	N	P	G	K	Y	N	I	L	G	T	N	T		Heterogenous nuclear ribonucleoprotein
	A	**E**	V	E	G	K	D	L	P	E	H	A	V	**L**		Heterogeneous ribonuclear particle protein U
	G	**E**	Y	L	P	S	K	G	A	G	N	N	V	**L**		Unnamed protein product (hum.)
	K	**E**	E	P	P	Q	P	Q	L	A	N	G	A	**L**		Protein C20orf24
	A	**E**	Q	A	E	R	Y	D	D	M	A	S	A	M		14-3-3 protein eta (Protein AS1)
	A	**E**	A	G	A	G	S	A	T	E	F	Q	F			40S ribosomal protein S10
	M	**E**	E	D	P	Q	T	S	R	E	I	F	N			Unnamed protein product (hum.)
	L	**E**	I	N	P	D	H	P	I	V	E	T	**L**			Heat shock protein HSP 90-β (HSP 90)
	S	**E**	M	E	V	Q	D	A	E	L	K	A	**L**			Proliferation-associated protein 2G4
	Y	**E**	D	F	K	E	E	G	S	E	N	A	V			Putative GTP-binding protein PTD004
	D	**E**	V	G	G	E	A	L	G	R	L	**L**				Hemoglobin beta subunit
	A	**E**	V	G	D	G	T	T	S	V	T	**L**				T-complex protein 1
	T	**E**	I	D	E	K	E	Y	I	S	**L**					Proteasome activator complex subunit 3
	Y	**E**	E	N	E	E	F	L	R	T	M					UPF0315 protein AD-001
	G	**E**	K	F	E	D	E	N	F	I	**L**					Peptidyl-prolyl cis-trans isomerase A
	I	**E**	V	D	P	D	T	K	E	M	**L**					40S ribosomal protein S17
	K	**E**	F	D	G	K	S	L	V	S	V					Heat shock protein HSP 90-β
	R	**E**	E	L	S	N	V	L	A	A	M					60S ribosomal protein L36
	K	**E**	L	S	E	D	E	I	R	T	**L**					52 kD repressor of protein kinase inhibitor
	H	**E**	E	A	V	S	V	D	R	V	**L**					S-methyl-5-thioadenosine phosphorylase
	E	D	A	L	P	P	I	V	L	R	S					Serine/threonine protein kinase 19
	S	L	I	D	R	S	S	A	P	E	**L**					Cytoplasmic linker protein 2
	A	**E**	E	K	A	A	V	T	S	**L**						caspr5 protein isoform 1
	Y	**E**	D	Y	Y	Y	H	P	P	P						Heterogeneous nuclear ribonucleoprotein R
	I	**E**	**V**	**D**	**G**	**K**	**Q**	**V**	**E**	**L**						Rho-related GTP-binding protein RhoB (H6)
	A	**E**	E	L	P	H	I	H	A	F						BolA-like protein 2
	A	**E**	K	V	E	I	A	T	**L**							Proteasome subunit α type 4
	S	**E**	E	D	F	I	R	S	**L**							FYVE finger-contain. phosphoinositide kinase
	Y	G	**A**	**E**	**A**	**L**	**E**	**R**	**M**							Hemoglobin α subunit
	Y	**E**	E	I	F	V	K	N	M							Tissue factor pathway inhib. (TFPI)
	V	**E**	K	I	D	F	D	S	V							Eukaryotic transl. initiation factor 3 subunit 12
	A	**E**	E	D	F	Y	E	K	**L**							DNA replication licensing factor MCM7
	T	**E**	D	I	P	V	K	T	**L**							Isoleucyl-tRNA synthetase, cytoplasmic
	M	L	L	A	F	Y	E	K	I							Fucose-1-phosphate guanylyltransferase
	F	**E**	E	I	Y	P	P	E	V							Cyclin-A2
	D	T	E	V	K	T	L	K	**L**							Zinc finger protein HRX (ALL-1)
	R	**E**	E	P	D	L	V	L	**L**							Menin
	N	**E**	D	W	E	N	P	Q	**L**							G2/mitotic-specific cyclin-B2
	A	**E**	L	I	V	Q	P	E	**L**							Desmoplakin
	K	**E**	L	G	I	T	A	**L**								40S ribosomal protein S14
	F	**E**	D	E	S	F	A	V								Peptidyl-prolyl cis-trans isomerase G
	G	**E**	A	L	G	R	L	**L**								Hemoglobin beta subunit
	H	W	P	F	G	A	L	**L**								Somatostatin receptor type 1 (SS1R)
	Y	S	D	M	K	R	A	**L**								Splicing factor, arginine/serine-rich 6
	D	M	E	N	E	F	L	**L**								Zinc finger protein 198
	I	R	L	L	E	E	A	**L**								DNA-dependent protein kinase catalytic subunit
	G	**E**	I	E	A	I	E	**L**								Heterogeneous nuclear ribonucleoprotein A/B
	G	**E**	D	V	V	T	L	**L**								Hexokinase-2
	F	**E**	D	E	N	F	I	**L**								Peptidyl-prolyl cis-trans isomerase A
	I	**E**	E	R	G	V	K	**L**								Septin-2 (protein NEDD5)

continued on next page

continued from previous page

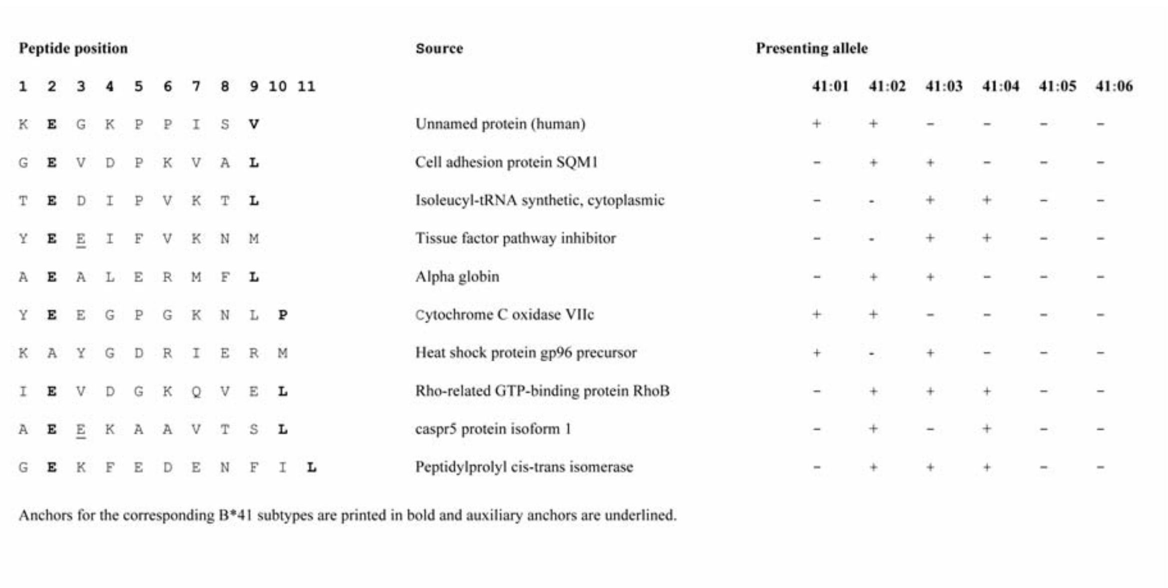
HLA-B*41:05

	Peptide position													
Ligands	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**Source**
	P	**E**	P	L	P	S	W	V	A	R	L	D	**L**	Leucine-rich repeats, immunoglobulin-like domains 3
	F	G	S	G	A	G	A	S	C	V	N	L	**L**	Neuroligin 2
	V	**E**	Y	R	G	Q	A	Q	A	I	E	F	**L**	Cytoplasmic linker 2
	L	**E**	I	N	P	D	H	P	I	V	E	T	**L**	Heat shock protein HSP 90-β (HSP 90)
	Y	F	P	N	M	S	V	G	F	T	A	**V**		TatD DNase domain containing deoxyribonuclease 2
	A	**E**	V	K	K	P	G	A	S	M	K	**V**		Immunoglobulin heavy chain variable region
	A	**E**	V	K	R	P	G	E	S	L	R	I		IgE heavy chain
	L	S	L	S	S	P	Q	V	P	G	S	**P**		KIAA1076 protein
	N	G	D	V	G	A	V	S	E	P	**P**			Unnamed protein product [Homo sapiens]
	V	C	G	K	P	F	H	S	L	S	**P**			Zinc finger protein 136
	A	**E**	V	T	V	S	A	S	G	L	**L**			SQFE253
	D	**E**	A	T	I	R	A	I	I	A	**V**			Unnamed protein product (hum.)
	K	**E**	L	E	L	M	P	E	K	G	**L**			CCCTC-binding factor-like protein
	Y	S	S	T	W	S	G	G	Y	G	**L**			Immunoglobulin alpha heavy chain variable region
	L	L	P	G	G	G	L	L	P	T	**P**			SFRS11 protein
	S	**E**	E	A	R	K	L	M	V	R	**L**			Neuroblastoma-amplified protein
	A	**E**	P	L	E	I	I	L	H	L	**P**			NHP2 non-histone chromosome protein 2-like 1
	A	**E**	A	F	E	A	I	P	R	A	**L**			CCT8 protein
	S	**E**	M	E	V	Q	D	A	E	**L**				Proliferation-associated protein 2G4
	D	N	V	P	M	F	I	S	Q	N				Similar to cadherin-related tumor suppressor precursor
	K	**E**	S	C	E	S	L	S	C	**L**				MYST histone acetyltransferase
	Q	**E**	D	A	E	M	D	A	E	**L**				KIAA1120 protein
	V	**E**	G	G	A	Q	V	Q	Q	**V**				Adaptor-related protein complex 2, α 2 subunit variant
	V	K	S	G	N	P	L	Q	**P**					Chromosome 10 open reading frame 18
	A	**E**	N	P	D	I	F	A	**V**					REC 1L
	P	V	P	A	W	A	R	A	**L**					Cholinergic receptor, α polypeptide 10
	S	P	S	G	W	S	N	V	F					Glycerol 3-phosphate permease
	S	T	C	G	S	H	F	M	**L**					Olfactory receptor, family 56, subfamily A, member 3
	L	V	A	L	S	Y	S	S	**P**					Nucleoporin 133kDa
	F	**E**	D	E	L	H	P	D	**L**					Peptidylprolyl isomerase-like 1
	H	A	G	L	D	A	F	L	**V**					HSPC320
	L	**E**	K	S	L	R	K	S	**L**					Putative glycine-N-acyltransferase
	N	V	V	D	S	S	Q	K	**P**					ALL-1 protein
	R	**E**	G	E	D	K	E	**L**						Titin isoform novex-3
	L	K	K	C	C	K	I	**L**						Plasma membrane Ca(2+)-ATPase isoform 2
	L	**E**	E	G	K	G	G	**P**						Neuronal pentraxin 1
	H	L	S	V	S	T	P	**V**						KIAA0929 protein
	T	**E**	E	A	K	R	Q	**P**						Similar to extensin-like protein
	T	S	F	R	L	A	L	I						Nitrilase family, member 2
	L	**E**	M	M	I	K	E	**L**						SnoN2
	Y	V	A	S	V	L	G	**L**						RAN-binding protein 2-like 1 isoform 2

HLA-B*41:06

	Peptide position														
Ligands	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**Source**
	K	**E**	K	E	T	P	T	P	G	E	D	I	Q	**V**	Hypothetical protein
	A	**E**	V	W	S	T	T	D	A	V	T	Q	**V**		Similar to GRIK2 protein
	I	K	L	Y	K	K	K	T	G	K	D	**V**			Glucose-regulated protein
	D	K	L	L	V	N	Q	T	E	L	F	**V**			Von Hippel-Lindau tumor suppressor-like
	L	**E**	G	R	T	H	T	I	S	L	**P**				KIAA0564 protein isoform a
	P	**E**	P	V	I	L	V	A	C	V	**P**				Phosphodiesterase H1
	A	**E**	M	I	E	K	Y	F	V	S	**P**				KIAA1356 protein
	A	**E**	R	V	G	A	G	A	P	**V**					H2A histone family, member J isoform 2
	C	**E**	S	C	V	D	L	L	F	**V**					Cyclin G1 interacting protein
	G	A	N	A	L	L	F	V	G	**V**					Adenylate cyclase type I
	L	S	L	L	L	F	S	A	**P**						Galanin receptor-like protein
	R	**E**	R	R	D	N	Y	V	**P**						Ribosomal Protein S17
	V	**E**	T	T	F	S	T	E	**P**						Chondroitin sulfate proteoglycan 2
	G	**E**	P	R	F	I	T	**V**							MHC class I antigen
	A	**E**	T	S	A	E	R	**V**							Protocadherin 10 isoform 1 precursor
	Y	**E**	T	D	V	F	V	**V**							GTF2I repeat domain containing 2
	L	**E**	A	L	L	F	R	**V**							KIAA0546 protein
	K	L	K	M	K	S	P	**P**							NADH dehydrogenase subunit 5
	A	**E**	G	F	L	F	V	**V**							ARNTL2 protein
	N	**E**	P	K	F	L	D	G							Cadherin-7 precursor
	Q	G	P	Y	G	M	D	**V**							Immunoglobulin heavy chain
	H	**E**	V	L	M	L	L	**P**							Unnamed protein product (human)

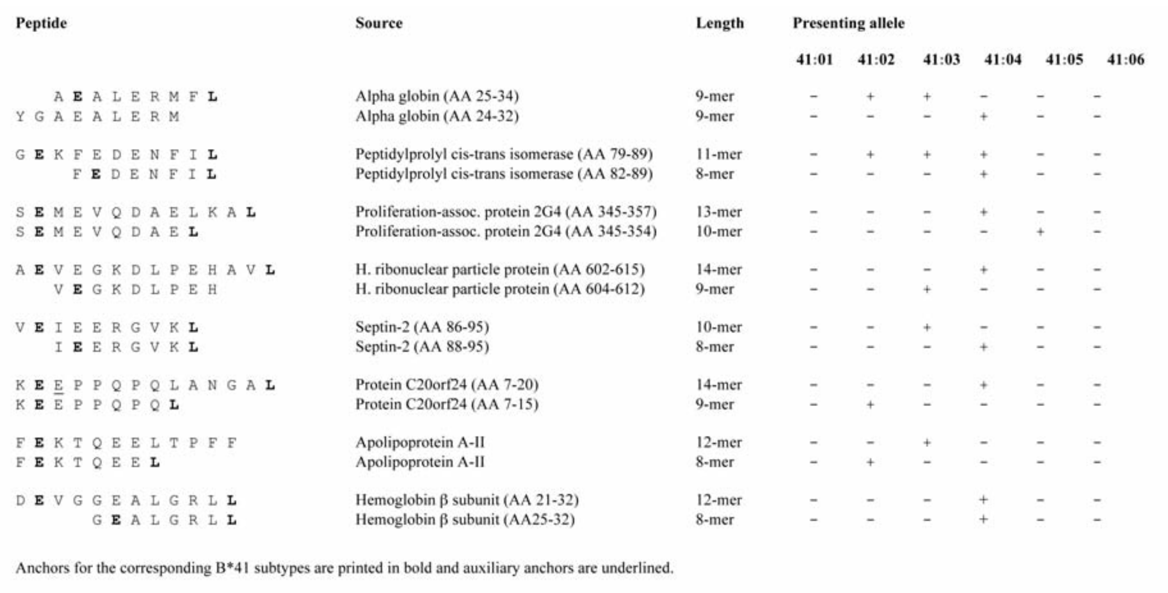
Online Supplementary Table S2. List of promiscuous peptides.



Peptide position											Source	Presenting allele					
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**		**41:01**	**41:02**	**41:03**	**41:04**	**41:05**	**41:06**
K	**E**	G	K	P	P	I	S	**V**			Unnamed protein (human)	+	+	–	–	–	–
G	**E**	V	D	P	K	V	A	**L**			Cell adhesion protein SQM1	–	+	+	–	–	–
T	**E**	D	I	P	V	K	T	**L**			Isoleucyl-tRNA synthetic, cytoplasmic	–	-	+	+	–	–
Y	**E**	E	I	F	V	K	N	M			Tissue factor pathway inhibitor	–	-	+	+	–	–
A	**E**	A	L	E	R	M	F	**L**			Alpha globin	–	+	+	–	–	–
Y	**E**	E	G	P	G	K	N	L	**P**		Cytochrome C oxidase VIIc	+	+	–	–	–	–
K	A	Y	G	D	R	I	E	R	M		Heat shock protein gp96 precursor	+	-	+	–	–	–
I	**E**	V	D	G	K	Q	V	E	**L**		Rho-related GTP-binding protein RhoB	–	+	+	+	–	–
A	**E**	E	K	A	A	V	T	S	**L**		caspr5 protein isoform 1	–	+	–	+	–	–
G	**E**	K	F	E	D	E	N	F	I	**L**	Peptidylprolyl cis-trans isomerase	–	+	+	+	–	–

Anchors for the corresponding B*41 subtypes are printed in bold and auxiliary anchors are underlined.

Online Supplementary Table S3. Differentially selected peptides by B*41 subtypes.



Peptide	Source	Length	Presenting allele					
			41:01	**41:02**	**41:03**	**41:04**	**41:05**	**41:06**
A **E** A L E R M F **L**	Alpha globin (AA 25-34)	9-mer	–	+	+	–	–	–
Y G A E A L E R M	Alpha globin (AA 24-32)	9-mer	–	–	–	+	–	–
G **E** K F E D E N F I **L**	Peptidylprolyl cis-trans isomerase (AA 79-89)	11-mer	–	+	+	+	–	–
F **E** D E N F I **L**	Peptidylprolyl cis-trans isomerase (AA 82-89)	8-mer	–	–	–	+	–	–
S **E** M E V Q D A E L K A **L**	Proliferation-assoc. protein 2G4 (AA 345-357)	13-mer	–	–	–	+	–	–
S **E** M E V Q D A E **L**	Proliferation-assoc. protein 2G4 (AA 345-354)	10-mer	–	–	–	–	+	–
A **E** V E G K D L P E H A V **L**	H. ribonuclear particle protein (AA 602-615)	14-mer	–	–	–	+	–	–
V **E** G K D L P E H	H. ribonuclear particle protein (AA 604-612)	9-mer	–	–	+	–	–	–
V **E** I E E R G V K **L**	Septin-2 (AA 86-95)	10-mer	–	–	+	–	–	–
I **E** E R G V K **L**	Septin-2 (AA 88-95)	8-mer	–	–	–	+	–	–
K **E** E P P Q P Q L A N G A **L**	Protein C20orf24 (AA 7-20)	14-mer	–	–	–	+	–	–
K **E** E P P Q P Q **L**	Protein C20orf24 (AA 7-15)	9-mer	–	+	–	–	–	–
F **E** K T Q E E L T P F F	Apolipoprotein A-II	12-mer	–	–	+	–	–	–
F **E** K T Q E E **L**	Apolipoprotein A-II	8-mer	–	+	–	–	–	–
D **E** V G G E A L G R L **L**	Hemoglobin β subunit (AA 21-32)	12-mer	–	–	–	+	–	–
G **E** A L G R L **L**	Hemoglobin β subunit (AA25-32)	8-mer	–	–	–	+	–	–

Anchors for the corresponding B*41 subtypes are printed in bold and auxiliary anchors are underlined.

Online Supplementary Table S4. Peptide contacts in the B*41:03/16-mer and B*41:04/11-mer structures.

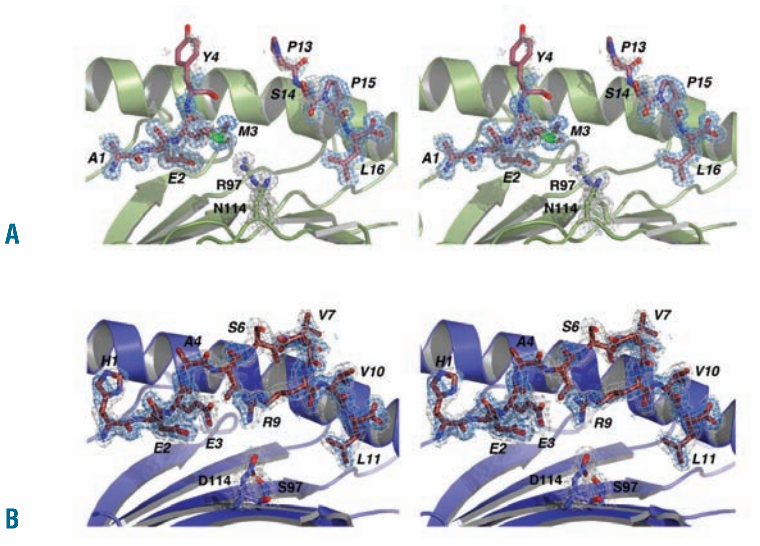
HLA-B*41:03/AEMYGSVTEHPSPSPL			HLA-B*41:04/HEEAVSVDRVL		
Peptide residue	**Interaction type**	**Contact[a,b]**	**Peptide residue**	**Interaction**	**Contact**
Ala$_1$	VDW[c]	Y7, R62, E63, Y159, W167, Y171	His$_1$	VDW	M5, Y7, E63 Y159, W167, Y171
Ala$_1^{N,O}$	HB[d]	Y7$^{O\eta}$, Y159$^{O\eta}$, Y171$^{O\eta}$	His$_1^{N,O}$	HB	Y7$^{O\eta}$, Y159$^{O\eta}$, Y171$^{O\eta}$
			His$_1^{N\delta1}$	HB	*E2*^{O}
Glu$_2$	VDW	Y7, H9, T24, K45, R62, E63, I66, S67, Y99, Y159	Glu$_2$	VDW	Y7, H9, T24, K45, E63, I66, S67, Y99, Y159
Glu$_2^{N,O}$	HB	R62$^{N\eta1}$, $^{N\eta2}$, E63$^{O\varepsilon2}$	Glu$_2^{N,O}$	HB	E63$^{O\varepsilon2}$, H1$^{N\delta1}$
Glu$_2^{O\varepsilon1,O\varepsilon2}$	HB, SB[e]	H9$^{N\eta2}$, K45$^{N\zeta}$, Y99$^{O\eta}$	Glu$_2^{O\varepsilon1,O\varepsilon2}$	HB, SB[e]	H9$^{N\varepsilon2}$, K45$^{N\zeta}$, Y99$^{O\eta}$
Glu$_2^{O\varepsilon1,O\varepsilon2}$	Wm[f]	T24$^{O\gamma1}$, S67$^{O\gamma}$, N70$^{N\delta2}$, *M3*^{O}	Glu$_2^{O\varepsilon1,O\varepsilon2}$	Wm	T24$^{O\gamma1}$, S67$^{O\gamma}$, N70$^{N\delta2}$, H9$^{N\varepsilon2}$, *Glu*^{3O}
Met$_3$	VDW	I66, R97, Y99, Q155, D156, Y159	Glu$_3$	VDW	I66, Y99, Q155, D156, Y159, *V5*
Met$_3^{N,O}$	HB, Wm	Y99$^{O\eta}$, N70$^{N\delta2}$, *Glu2*$^{O\varepsilon2}$	Glu$_3^{N,O}$	HB, Wm	Y99$^{O\eta}$, N70$^{N\delta2}$, *Glu2*$^{O\varepsilon2}$, *V5*^{N}
Met$_3^{S\delta}$	HB, Wm	R97$^{N\eta2}$, D156$^{O\delta1}$	Glu$_3^{O\varepsilon1,O\varepsilon2}$	HB, SB, Wm	D156$^{O\delta1}$, *R9*$^{N\eta1,N\eta2}$
Tyr$_a$	VDW	R62			
			Val$_5$	VDW	N70$^{N\delta2}$, *E3*, *R9*
			Val$_5^{N,O}$	Wm	I66^{O}, N70$^{O\delta1}$
			Ser$_6$	VDW	*D8*, *R9*
			Ser$_6^{O}$	HB	*R*$_9^{N}$
			Ser$_6^{O\gamma}$	HB, Wm	*D8*$^{O\delta1}$
			Val$_7$	VDW	T73, E76, *R9*, *V10*
			Val$_7^{O}$	HB	*V*$_{10}^{N}$
Pro$_{13}$	VDW	T73, *P15*	Asp$_8$	VDW	A150, *S6*, *V10*
			Asp$_8^{O\delta1,O\delta2}$	HB	Q155$^{N\varepsilon2}$, *S6*$^{O\gamma}$, *R9*^{N}
Ser$_{14}$	VDW	W147, V152	Arg$_9$	VDW	W147, V152, Q155, *V5*, *S6*, *V7*
Ser$_{14}^{O}$	Wm	Y116$^{O\eta}$	Arg$_9^{N,O}$	HB, Wm	Y116$^{O\eta}$, *S6*^{O}
			Arg$_9^{N\eta1,N\eta2}$	HB, SB, Wm	Y116$^{O\eta}$, D114$^{O\delta1\ O\delta2}$, Q155$^{O\varepsilon1}$, D156$^{O\delta1\ O\delta2}$, *E3*$^{O\varepsilon1,O\varepsilon2}$
Pro$_{15}$	VDW	T73, E76, S77, W147, *P13*	Val$_{10}$	VDW	T73, E76, S77, N80, W147, *V7*, *D8*
Pro$_{15}^{O}$	HB	W147$^{N\varepsilon1}$	Val$_{10}^{N,O}$	HB	W147$^{N\varepsilon1}$, *V7*^{O}
Leu$_{16}$	VDW	S77, N80 Y84, L95, Y116, Y123, T143, K146, W147	Leu$_{11}$	VDW	S77, N80, L81, Y84, L95, Y116, Y123, T143, W147
Leu$_{16}^{N,O,OXT}$	HB, SB	S77$^{O\gamma}$, N80$^{N\delta2}$, Y84$^{O\eta}$, T143$^{O\gamma1}$, K146$^{N\zeta}$	Leu$_{11}^{N,O,OXT}$	HB, SB	S77$^{O\gamma}$, N80$^{N\delta2}$, Y84$^{O\eta}$, T143$^{O\gamma1}$, K146$^{N\zeta}$

[a]Atomic contacts determined using the CCP4i implementation of CONTACT[1,2] and a cutoff of 4 Å. [b]Selected contacts between peptide residues are marked in italics. [c]VDW - Van der Waals' interactions, defined as non-hydrogen bond contact distances of 4 Å or less. [d]HB - Hydrogen bond interactions, defined as contact distances of 3.5 Å or less, between suitable atoms at appropriate angles,[3] as determined in Coot. [e]SB - Ionic interactions, defined as contact distances of 4.0 Å or less between suitable residue atoms. [f]Wm - Water mediated hydrogen bond (see "d").

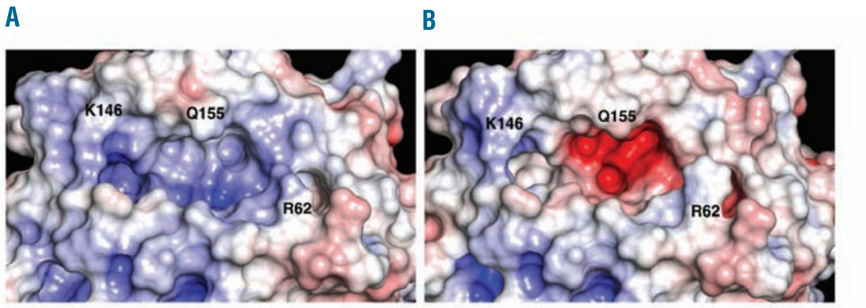
Online Supplementary Table S5. Theoretically calculated ionization states for selected titratable groups in the B*41:03 and B*41:04 structures.

Protein Model	**Predicted Residue p$K_{1/2}$[a]**						
	His9 (6.6)	**Arg97 (12.0)**	**Asp114 (4.0)**	**Asp156 (4.0)**	***pGlu2* (4.4)**	***pGlu3* (4.4)**	***pArg9* (12.0)**
B*41:03	1.8	21.0	-	1.2	-	-	-
B*41:03/16mer[b]	4.1	26.3d	-	0.82	-10.2[d]	-	-
B*41:03 R97S[c]	5.2	-	-	2.3	-	-	-
B*41:04	6.4	-	2.6	7.2	-	-	-
B*41:04/11mer	9.1	-	3.9	11.5	-9.5	5.2	27.9
B*41:04/11mer *pR9A*[c]	11.1	-	4.1	15.9	-8.3	8.1	-

*[a]pK calculations were carried out using the H++ server[4] and a range of starting models based on the crystal structures of the B*41:03/16mer and B*41:04/11mer complexes. The calculated values for selected heavy chain and peptide residues are presented. Standard solution values for the pK of each residue type are also shown in brackets. [b]The full length 16mer peptide conformation derived by molecular modeling was used in the calculations. [c]Residue coordinates for R97 and pR9 in the case of the B*41:03 and B*41:04 structures, respectively, were mutated in COOT. [d]Extreme low/high p$K_{1/2}$ for a given group are representative of very high energetic barriers that prevent changes to that group's protonation state within the range of experimental accessible pH values.*



Online Supplementary Figure S1. Stereographic representations of the antigen-binding cleft of B*41:03/AEMYGSVTEHPSPSPL and B*41:04/HEEAVSVDRVL Stereographic representations of the antigen-binding cleft in (A) the B*41:03/16mer and (B) the B*41:04/11mer complexes. In each case the MHC is presented in cartoon format and the α-2 helix of the heavy chain has been removed for clarity. The peptide residues are drawn as sticks, as are the side chains of the polymorphic positions 97 and 114. Overlayed is the unbiased *Fo-Fc* density in the region of the peptide, contoured at 3.0 σ (marine mesh), as well as the final *2Fo-Fc* density for the peptide and heavy chain residues 97 and 114, contoured at 1.0 σ (gray mesh).



Online Supplementary Figure S2. Polymorphism alters the size and charge of the antigen-binding cleft in B*41:03 and B*41:04 Surface representation of (A) the B*41:03/16mer and (B) the B*41:04/11mer structures, showing the antigen-binding cleft from above with peptide and solvent atoms removed. In each case the ± 6 kT/e electrostatic potential from the solvent accessible surface of the MHC has been rendered on the molecular surface. The N- and C-terminal ends of the cleft are marked by Arg_{62} and Lys_{146}, respectively, while the D pocket aligns with Gln155. The polymorphism at positions 97 and 114 gives rise to an antigen-binding cleft in B*41:04 that is 140 Å^3 larger and more electronegative than B*41:03 in the region of pocket D. The molecular volume of each antigen-binding cleft (corresponding to the Connolly surface) was calculated using the *CASTp* server and a 1.4 Å probe radius.[5] Electrostatics calculations were carried out using only the heavy chain and β-2-microglobulin coordinates of each structure. Briefly coordinate preparation (placing and optimizastion of hydrogen atoms assuming standard protonation at pH 7.0, followed by assignment of atomic charge and radii parameters using the PARSE forcefield) was carried out using the *PDB2PQR* server (v1.5;).[6] Electrostatic calculations were subsequently carried out using the *APBS* plugin in *PyMOL* (v1.1.0[7] and v1.2.x,[8] respectively) with a 0.15 M concentration for the +1 and -1 ion species.

References

1. Bailey S. The CCP4 suite: programs for protein crystallography. Acta Cryst. 1994;50(Pt 5):760-3.
2. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983;22(12):2577-637.
3. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. Prog Biophys Mol Biol. 1984;44(2):97-179.
4. Gordon JC, Myers JB, Folta T, Shoja V, Heath LS, Onufriev A. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. Nucleic Acids Res. 2005;33(Web Server issue):W368-71.
5. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 2006;34 (Web Server issue):W116-8.
6. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. 2004;32(Web Server issue):W665-7.
7. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA. 2001;98(18):10037-41.
8. DeLano WL. The PyMOL Molecular Graphics System, (http://www.pymol.org). 2002 [cited; Available from: http://www.pymol.org