The nature of peptides presented by an HLA class I low expression allele

Jan Hinrichs, Daniel Föll, Christina Bade-Döding, Trevor Huyton, Rainer Blasczyk, and Britta Eiz-Vesper

Institute for Transfusion Medicine, Hannover Medical School, Hannover, Germany

ABSTRACT

Background
The functional integrity of human leukocyte antigen low expression variants is a prerequisite for considering them as essential in the matching process of hematopoietic stem cell donors and recipients to diminish the risk of serious complications such as graft-*versus*-host disease or graft rejection. The HLA-A*3014L variant has a disulfide bridge missing in the α2 domain which could affect peptide binding and presentation to T cells.

Design and Methods
HLA-A*3014L and HLA-A*3001 were expressed as truncated variants and peptides were eluted and subjected to pool sequencing by Edman degradation as well as to single-peptide sequencing by mass spectrometry. Quantitative analysis of binding peptides presented *in vivo* was performed by a flow cytometric peptide-binding assay using HLA-A*3001 and HLA-A*3014L-expressing B-LCLs.

Results
The truncated HLA-A*3014L protein was secreted in the supernatant and it was possible to elute and sequence peptides. Sequence analysis of these eluted peptides revealed no relevant differences to the peptide motif of HLA-A*3001, indicating that the Cys164Ser substitution does not substantially alter the spectrum of presented peptides. Strong binding of one of the shared *in vivo* identified HLA-A*3001/3014L ligands was confirmed in the peptide-binding assay.

Conclusions
This study is the first to demonstrate that HLA low expression variants are able to present peptides and, thus, can be considered as functionally active. When comparing peptide motifs, it is likely that HLA-A*3014L and HLA-A*3001 represent a permissive mismatch with low allogenicity in hematopoietic stem cell transplantation. These results indicate that surface expression, as well as peptide-binding data of HLA variants with similar disulfide bridge variations (e.g. HLA-A*3211Q) need to be considered as functionally active in an allogeneic hematopoietic stem cell transplantation setting as long as the opposite has not been shown. Otherwise a relevant but not considered HLA mismatch could result in a severe allogeneic T-cell response and graft-*versus*-host disease.

Key words: HLA-A*3014L, peptide motif, HLA low expression alleles, stem cell transplantation.

Introduction

Human leukocyte antigen (HLA) class I molecules present endogenous derived peptides of 8-12 amino acids (aa) in length to CD8$^+$ cytotoxic T lymphocytes (CTL).[1] Most amino acid polymorphisms of different HLA class I molecules are located in the peptide-binding region (PBR) shaped by parts of the $\alpha 1$ and $\alpha 2$ domains and determine the characteristics of presented peptides. Peptide motifs have been reported for the most common HLA-A and B alleles and for some rare variants. Importantly, differences in peptide binding among the alleles of a serological group have also been described.[2,3] Identification and comparison of allele-specific peptide-binding motifs provide important information for donor-recipient matching and prediction of HLA subtype allogenicity in allogeneic hematopoietic stem cell transplantation (HCST).

In order to determine the functionality of HLA low expression (L) alleles, we isolated and sequenced peptides from HLA-A*3014L. Hirv *et al.*[4] originally identified HLA-A*3014L by sequencing-based typing in a patient suffering from chronic myeloid leukemia. The sequence of this low expression allele is identical to that of HLA-A*3001 except for a cysteine (C) to serine (S) substitution at aa position 164, thus impairing disulfide bridge formation in the $\alpha 2$ domain of the mature polypeptide. This alteration of the secondary structure is assumed to decrease the level of protein expression on the cell surface, rendering HLA-A*3014L serologically undetectable.[2,4,5] Although reportedly undetectable by standard microlymphocytotoxicity assays, we analyzed the functional integrity of HLA-A*3014L because weak expression was demonstrated when the corresponding B-lymphoblastoid cell line (B-LCL) was cultured at 30°C or exposed to cytokines at 37°C.[4,6]

Here, we show for the first time that certain HLA low expression alleles present peptides, indicating that they can be functionally active. In order to predict the relevance of such alleles in an allogeneic HSCT setting, it is important to determine the expression levels of HLA variants with similar disulfide rearrangements (e.g. HLA-A*3211Q and B*3565Q). Since HLA mismatches are the main cause of severe graft-*versus*-host-disease (GvHD) and graft rejection, a misinterpretation of these variants as irrelevant could strongly affect transplant-related mortality.[2,7]

Design and Methods

Expression of soluble HLA (sHLA) in a eukaryotic expression system

Human embryonic kidney cell line 293 (HEK293) was maintained in Dulbecco's modified eagle's medium (DMEM; Lonza, Verviers, Belgium) containing 10% fetal calf serum (FCS; Lonza) at 37°C. mRNA of HLA-A*3001 and HLA-A*3014L was amplified by RT-PCR (exons 1-4), ligated into the pcDNA3.1_V5/His eukaryotic expression vector (TOPO TA Expression Kit, Invitrogen, Karlsruhe, Germany), and transfected into HEK293 cells as described previously.[6] Selection of clones resistant to geneticin (G418, Invitrogen) was performed 48h posttransfection using 1000µg/mL G418. After subcloning by limiting dilution, screening for high expression of soluble HLA-A*30 molecules (sHLA) in the supernatants was performed by sandwich ELISA using the anti-HLA-ABC monoclonal antibody (mAb) W6/32 (Serotec, Düsseldorf, Germany) as capture antibody and a horseradish per-oxidase (HRP)-conjugated anti-β_2-microglobulin ($\beta 2m$) antibody (DAKO, Hamburg, Germany) for detection.[6]

HEK293 clones with the highest sHLA expression were used for large-scale production in a CELLine adhere 1000 (Integra Biosciences AG, Chur, Switzerland) two-compartment bioreactor facilitating a high cell density. Supernatants (25 mL/week) were harvested, centrifuged, filtered through a 0.45µm cut-off filter (Vivascience, Göttingen, Germany), and stored at -20°C for further purification. sHLA-A*3014L expression ranged from 600 ng/mL in the first week to more than 1,500 ng/mL during the following 10-week collection period. Compared to the low expression allele, the expression level of HLA-A*3001 was 50 to 100 times higher as determined by ELISA.

Purification and detection of soluble HLA

We purified separately 250 mL supernatant containing sHLA-A*3014L or sHLA-A*3001 (adjusted to pH 8.0) using N-hydroxy-succinimide-(NHS)-activated HiTrap columns (GE Healthcare, Uppsala, Sweden) coupled with the anti-HLA-ABC mAb W6/32 (Serotec) in a BioLogic DuoFlow System (Bio-Rad, Hercules, USA) as described previously.[2,8,9] Trimeric sHLA complexes (heavy chain, $\beta 2m$, peptide) were eluted using 0.1 M glycine-HCL buffer (pH 2.7), yielding approximately 0.3 mg of purified sHLA-A*3014L and 10 mg of purified sHLA-A*3001.

Elution, sequencing and analysis of peptides from soluble HLA-A*30 alleles

Peptides from soluble HLA-A*30 complexes were eluted by 0.1% trifluoroacetic acid followed by filtration through a 10 kDa cut-off YM membrane (Vivascience). The flowthrough was concentrated using C18 ZipTips (Millipore, Billerica, Massachusetts, USA). The final preparation was monitored by MALDI-TOF-MS (Proteomics Analyzer 4800, Applied Biosystems) and further analyzed by pool sequencing applying 12 automated cycles of Edman degradation using a Procise 492A protein sequencer (Applied Biosystems, Forster City, CA, USA). Finally, single peptide sequences were identified by LC-ESI-MS/MS technology using an Eksigent nano-LC Ultra 2D HPLC coupled to an Orbitrap ion trap (Thermo Fischer, Waltham, Massachusetts, USA), resulting in a very high mass accuracy (<5 ppm). Mascot database queries[10] were performed using the IPI human database and the respective decoy database.

Flow cytometric verification of peptide binding

In order to verify the binding of the isolated HLA-A*3001/3014L peptides we used a flow cytometric peptide-binding assay. Three Epstein-Barr virus (EBV)-transformed B-lymphoblastoid cell lines (B-LCLs) expressing either HLA-A*3014L,*0201 (Ulm-241539),(4) HLA-A*3001,*0201 (EBRCC-256),[6] or only HLA-A*0201 (EBRCC-2296) were used. B-LCLs were maintained in RPMI-1640 (Lonza) containing 10% FCS (Lonza) supplemented with penicillin/streptomycin (c.c.pro, Oberdorla, Germany).

The peptide-binding assay was performed as previously described with small modifications.[11-13] Briefly, the naturally HLA-presented peptides were eluted from viable B-LCLs using a pH 3.3 acid elution technique. After washing the cells in Hank's balanced salt solution (HBSS Sigma, Deisenhofen, Germany), cells were incubated with 5 ml of pH 3.3 citrate-phosphate buffer (0.131 M citric acid, 0.066 M Na$_2$HPO$_4$) at room temperature for 2 min. We added 20 ml culture medium to neutralize the pH, and the cells were subsequently pelleted and resuspended in culture medium at a concentration of 1×10^6 cells/mL. Acid-treatment resulted in the dissociation of the ternary HLA class I complex by releasing $\beta 2$-microglobulin and bound endogenous peptides into the super-

natant, leaving free class I heavy chains on the cell surface. A subsequent washing step removed the released endogenous peptides and β2-microglobulin. Reconstitution of the HLA class I molecules was carried out by adding synthetic fluorescein isothiocyanate (FITC)-labeled HLA peptides and excess of recombinant β2m. The synthetic FITC-labeled peptide EITALAK(FITC)PSTMK (HLA-A*3001/3014L) and the immunodominant HLA-A*0201-restricted CMVpp65[495-503] peptide (NLVPMK(FITC)VATV) were used (GL Biochem, Shanghai, China). The additional aa lysine (K) in both peptide sequences was required for FITC-labeling of the peptides. Purity was greater than 95%.

Acid-treated and non-acid treated B-LCLs were placed into 24-well tissue culture plates using a pipette at a volume of 250 μL per well (1×10[6] cells). Peptide solutions containing 5 μg/mL or 10 μg/mL of the synthetic peptides (EITALAK(FITC)PSTMK or NLVPMK(FITC)VATV) were added to the respective sample. Simultaneously, 5 μg/mL of β2m (Sigma, Deisenhofen, Germany) was added to the wells. As negative controls, acid-treated and non-acid treated cells were cultured in 250 μL culture medium without peptides. Cells were incubated at 37°C for 3 h in a humidified 5% CO_2 atmosphere. For FACS analysis, cells were stained on ice for 30 min with 10 μL monoclonal phycoerythrin (PE)-labeled anti-HLA-ABC antibody (clone W6/32, Serotec). All samples were analyzed using a flow cytometer (FACScanto, BD Biosciences, Heidelberg, Germany) with live gating on B-lymphocytes during acquisition. At least 50,000 gated events were included in each analysis. Quadrants were set based upon the negative controls. Four independent experiments were performed for each cell line and each peptide, respectively.

Reconstitution of HLA class I molecules with the respective FITC-labeled peptide on the acid-treated B-LCLs was determined by the percentage of PE-HLA-ABC+/FITC-peptide+ cells. Furthermore, relative peptide-binding intensities of the HLA-A*3001/3014L ligand (EITALAK(FITC)PSTMK) compared to the HLA-A*0201 ligand (NLVPMK(FITC)VATV) were calculated for each of the three cell lines as followed: [double positive A*30 binding cells/double positive A*02 binding cells]*100. Background expression of cells, which were cultured without the synthetic peptides, was used as the blank value.

Results

Sequences of peptides eluted from the recombinantely expressed HLA-A*30 alleles

The conformational anti-HLA-ABC mAb W6/32 was used for detection and purification of soluble HLA-A*30 molecules. W6/32 is able to detect the HLA-A*3014L allele, ensuring that the molecules are correctly folded, β2-microglobulin-assembled and peptide-loaded. Therefore, we hypothesized that peptides can be presented by HLA low expression variants.[6] To verify this hypothesis, we eluted peptides from affinity purified low expression alleles and analyzed the samples by Edman pool degradation and single-peptide sequencing (Tables 1 and *Online Supplementary Table S1*).

By Edman pool degradation the described HLA-A*30 peptide motif[15,16] was confirmed for eluted ligands of HLA-A*3001 and HLA-A*3014L, respectively (Table 1). The C-terminal position (PΩ) of the peptides was identified as a primary anchor position. The preferred residues of the HLA-A*30 peptide epitopes at this position are lysine (K), valine (V) and arginine (R). The preference for lysine (K) as the top amino acid at the PΩ position of the bound pep-

tides, as described[15,16] by positional scanning combinatorial peptide libraries (PSCPL) analysis, could be consolidated by our peptide sequence data. Position P3 of the peptides was identified as a primary-secondary anchor showing a high preference for the basic amino acids K and R. Six amino acids are reportedly favored at position P2: phenylalanine (F), serine (S), threonine (T), valine (V), isoleucine (I) and leucine (L).

The obtained peptides ranged from 8 to 14 aa in length, but most were 9 to 10 aa (*Online Supplementary Table S1A and B*). The sequences of 200 HLA-A*3001 ligands and 100 HLA-3014L ligands were identified. Three peptide epitopes (3%) were presented by both HLA-A*3001 and HLA-A*3014L: 1) VLDTPGPPV, a nonameric peptide derived from titin (isoform N2-A, aa position 19783-19791), a protein of human muscle ultrastructure and elasticity; 2) EITALAPSTMK, an 11-mer peptide derived from human muscle protein ACTA1 (actin, alpha 1, skeletal muscle; aa position 301-311); and 3) DNIQGITKPAIR, a 12-mer peptide derived from a histone protein (HIST2H4A; aa position 25-36) (*Online Supplementary Table S1C*). Regarding the three shared epitopes of HLA-A*3001 and A*3014L, the hydrophobic residue valine (V) was located at PΩ and the hydrophobic residue leucine (L) was the preferred anchor amino acid at P2 in the case of the nonameric peptide (VLDTPGPPV). In the 11-mer peptide (EITALAPSTMK), the basic residue K was detected at PΩ and the hydrophobic residue I was a preferred anchor at P2. Exclusively in the sequence of the 12-mer peptide (DNIQGITKPAIR) the basic residue R was found as a main anchor for PΩ. We generated the binding spectra of the three peptides for each of the two alleles to show that the Edman degradation analyses worked rationally. The binding spectra of the 11-mer peptide are shown as a representative example (Figure 1).

Verification of peptide binding to HLA-A*3001 and HLA-A*3014L by a flow cytometric peptide-binding assay

To verify the presentation of naturally isolated peptides from recombinant HLA-A*3001/3014L molecules we used a flow cytometric peptide-binding assay[11-13] on three EBV-transformed B-LCLs expressing either HLA-

Table 1. Peptide motif of HLA-A*3001 and HLA-A*3014L alleles.

	Position										
	1	2	3	4	5	6	7	8	9....12		
Anchor			**K**						**K**		
			R						**V**		
									R		
<u>Preferred</u>	<u>K</u>	<u>F</u>		<u>P</u>		<u>F</u>	<u>I</u>	<u>F</u>			
	<u>R</u>	<u>S</u>				<u>R</u>	<u>M</u>	<u>Y</u>			
		<u>T</u>					<u>P</u>				
		<u>V</u>									
		<u>I</u>									
		<u>L</u>									

*Peptide motif of HLA-A*3001 and HLA-A*3014L alleles obtained by pool sequencing in 12 automated cycles of Edman degradation. Motif definition was performed as described by Falk et al.[23] Bold print: anchors as defined by a 100% increase in pool sequencing signals and by their exclusive occurrence in natural ligands. Underlined: Auxiliary anchors as defined by a >50% increase in pool sequencing signals and by virtue of their occurrence in a significant number of ligands.*

A*3014L,*0201 (Ulm-241539), HLA-A*3001,*0201 (EBRCC-256) or HLA-A*0201 (EBRCC-2296).[4,6] Acid-treatment of the cell lines resulted in the dissociation of the naturally bound peptides and the release of β2-microglobulin from the HLA class I heavy chain. By adding synthetic FITC-labeled HLA peptide ligands and excess of recombinant β2-microglobulin the HLA class I molecules were reconstituted. The *in vivo* isolated HLA-A*30-restricted 11-mer peptide EITALAK(FITC)PSTMK (Table 2) and for control purposes the immunodominant HLA-A*0201-restricted CMVpp65$_{495-503}$ peptide (NLVPMK(FITC)VATV) were used to calculate the relative peptide-binding intensities to the respective HLA-A alleles.
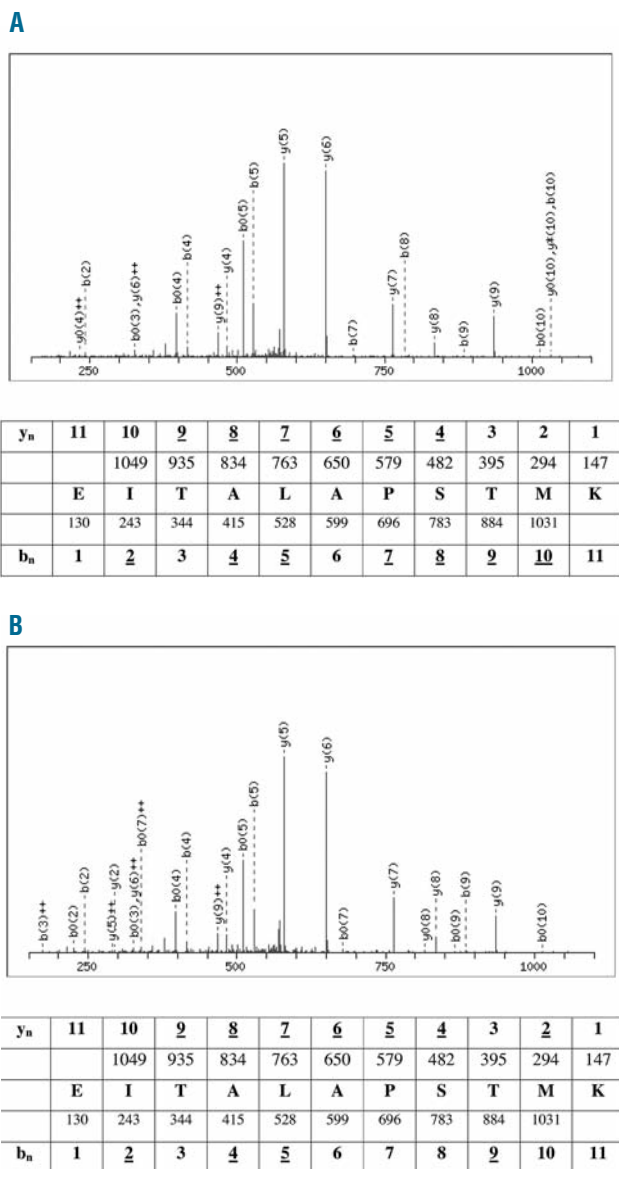
After incubation of the cells with 10μg of the HLA-A*02-restricted peptide about 64.4% HLA-ABC/FITC-peptide+ cells were detected for the non-acid treated HLA-A*0201-homozygous B-LCL EBRCC-2296 (Figure 2). In contrast, after adding the HLA-A*30-restricted ligand a mean of just 9.1% HLA-ABC/FITC-peptide+ cells was detectable. For the acid-treated cells, a reconstitution of 23.5% and 5.9% HLA-ABC/FITC-peptide+ complexes after adding the A*02- and A*30-restricted peptides, respectively, were found. For the HLA-A*0201/A*3001 non-acid treated cell line (EBRCC-256) 12.4% HLA-ABC/A*02-peptide+ and 4.5% HLA-ABC/A*30-peptide+ cells were detected. Acid-treatment and peptide incubation resulted in a reconstitution of 5.0% of HLA-ABC/A*02-peptide+ and 3.7% of HLA-ABC/A*30-peptide+ cells. Similar results were observed for the HLA-A*3014L,*0201-expressing B-LCL (Ulm-241539). On the non-treated cells, 44.1% of the HLA molecules bound the A*02-restricted peptide and 7.2% the A*30 ligand. After acid-treatment, 10.4% of the cells were stained positive for HLA-ABC/A*02-peptide and approximately 5.3% were positive for HLA-ABC/A*30. Calculation of the relative peptide-binding intensities was performed to compare the binding rates of the A*02- and A*30-restricted peptide ligands to the respective HLA molecules expressed on the B-LCLs (Figure 2). The reconstitution of HLA with the HLA-A*30 ligand mounts up to 51% (Ulm-241539) and up to 74% (EBRCC-256), compared to 25% for the HLA-A*0201 homozygous cell line (EBRCC-2296). The binding was higher on those cells expressing the normal HLA-A*3001 allele compared to cells expressing the low expression variant HLA-A*3014L (Figure 2). The results confirmed that this *in vivo* isolated A*30 peptide binds to HLA-A*30 on the cell surface. Peptide binding was found for the A*30 peptide on the HLA-A*3014L-expressing cell line (Ulm-241539), indicating the stability of the HLA-A*3014L cell surface expression.

Similar results, but with a dose-dependent reduction of the HLA complex reconstitution, was observed after decreasing the peptide concentration to 5 μg/mL. Compared to the binding of the HLA-A*0201-restricted peptide, for the HLA-A*30 ligand peptide-binding intensities of 27% for EBRCC-2296 (A*0201), 78% for EBRCC-256 (HLA-A*3001, A*0201) and 68% for Ulm-241539 (HLA-A*3014L, A*0201) were calculated. In relation to the A*02-restricted peptide, the peptide-binding intensity for the A*30-ligand was strongly increased on those cell lines expressing a HLA-A*30 allele.

Homology-based modeling to compare the binding characteristics of the two HLA-A*30 alleles

We performed homology-based modeling for each of the HLA-A*30 alleles with the shared 9, 11 and 12-mer peptide epitopes to investigate the binding features of the sequenced ligands. The HLA-A*3001 and HLA-A*3014L models are essentially identical with the Cys164>Ser substitution simply adopting an alternate rotamer conformation upon breaking of the disulfide bond breakage. Therefore, only the HLA-A*3014L model is illustrated (Figure 3).

Using CONTACT software [CCP4 (1994)],[22] we analyzed the 9, 11 and 12-mer peptide epitopes presented by both HLA-A*3014L and HLA-A*3001 and the respective models to identify residues in contact with the peptide (cut-off <4 Å). Based on the comprehensive pocket definition by Chelvanayagam,[19] we compared heavy chain

A

y$_n$	11	10	9	8	7	6	5	4	3	2	1
		1049	935	834	763	650	579	482	395	294	147
	E	I	T	A	L	A	P	S	T	M	K
	130	243	344	415	528	599	696	783	884	1031	
b$_n$	1	2	3	4	5	6	7	8	9	10	11

B

| y$_n$	11	10	9	8	7	6	5	4	3	2	1
		1049	935	834	763	650	579	482	395	294	147
	E	I	T	A	L	A	P	S	T	M	K
	130	243	344	415	528	599	696	783	884	1031	
b$_n$	1	2	3	4	5	6	7	8	9	10	11

Figure 1. Spectra of the undecameric peptide EITALAPSTMK eluted from HLA-A*3001 (A) and HLA-A*3014L (B). These spectra show the peptide summary for MS/MS fragmentation of the EITALAPSTMK peptide eluted from HLA-A*3001 (A) and HLA-A*3014L (B) produced using MASCOT software. Predicted masses for b and y fragment ions are listed, and observed ions are labeled in the respective spectra. N- and C-terminal peptide fragments present in all ESI-MS/MS spectra are labeled according to standard nomenclature.[17,18]

residues contacting positions P2, P3 and PΩ of the bound peptides (Table 2) to find the reason for the lack of an auxiliary anchor at P2 of the A*30 derived peptides as described by Lamberth *et al.*[15] When compared with the modules for P2 (Module Explorer, www.histocheck.org),[23,24] no auxiliary anchor at P2 was found for A*1101 or A*3001/14L. According to the Chelvanayagam pocket definition, a module is defined as all amino acids of an HLA molecule forming the respective pocket which binds a certain position of the peptide.[19] A*1101 and A*3001/14L can be distinguished by polymorphism Y9S and R163T relating to pocket B. Residue 9 is a contact residue of P2 of the A*1101_11-mer structure but not of the other five structures. Residue 163 obviously is not in contact with the peptides of any of the six structures. Therefore, these polymorphic residues are unlikely to cause the alteration of P2. The same holds true for pocket D. Within module 3 (all amino acids assigned to bind P3 of the peptide according to the Chelvanayagam definition), the following five variations distinguish A*1101 and A*3001/14L: T9S, R114E, A152W, Q156L, R163T. However, contact analysis showed that these residues are not in contact with P3 and, thus, are of minor importance here (Table 2). Amino acid 156 is part of pocket D of A*1101 but not of A*30 molecules.

Discussion

The fact that the monoclonal anti-HLA class I antibody W6/32 used in this study is a conformational antibody implies that only correctly folded, β2-microglobulin-assembled and peptide-loaded MHC complexes can be detected. This suggests that parts of soluble HLA-A*3014L molecules are assembled correctly and secreted by transfectants. Presumably, more soluble HLA-A*3014L is produced and secreted into the supernatant that might be not correctly folded because of the lack of the disulfide bridge in the α2 domain. Therefore, these molecules are not detected by the conformational anti-HLA-ABC mAb. This assumption arose after comparing mRNA and associated protein levels of HLA-A*3014L and HLA-A*3001 alleles.[6] The fact that the mRNA levels measured by RT-PCR are the same for both alleles leads to the conclusion that no transcriptional differences occur. In contrast, the protein levels varied greatly (50 to 100-fold difference). These results are supported by findings of Warburton *et al.*[5] who determined a 95% decrease of cell surface HLA-A*0201 expression after performing side-directed mutagenesis of aa positions 101 and 164 in the HLA molecule. The HLA-A*3014L protein accumulates inside the cells, which indicates that HLA-A*3014L translation is not affected. Consequently, the lack of protein secretion in the supernatant is best explained by posttranslational instability of the HLA-A*3014L molecules because of the missing disulfide bridge.[6] Based on these findings, it is also likely that the intracellular enriched HLA-A*3014L protein is a major substrate for proteasomal cleavage and provides a flood of peptide fragments presented to cytotoxic T lymphocytes. As a result of this indirect allo-recognition pathway, it is possible that GvHD or graft rejection might be promoted in the event of a severe mismatch.[25] Consequently, considering HLA-A*3014L as null allele is, in case of a mismatch with any other HLA-A allele, potentially more dangerous in terms of GvHD and graft rejection according to the
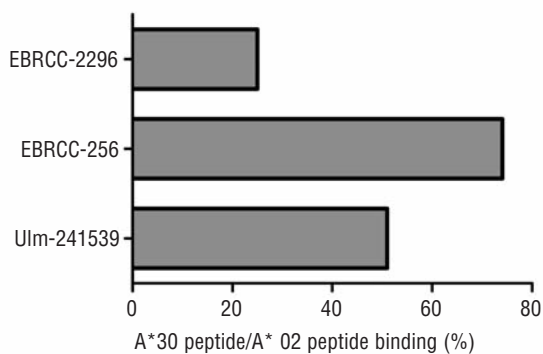
Figure 2. Relative A*30/A*02 peptide-binding intensities for different HLA-expressing B-LCLs. The diagram shows the peptide-binding intensities of the investigated cell lines EBRCC-2296 (HLA-A*0201 homozygous), EBRCC-256 (HLA-A*0201/A*3001), Ulm-241539 (HLA-A*0201/A*3014L) achieved after acid-treatment and reconstitution using FITC-labeled A*3001/3014L-restricted (EITALAK(FITC)PSTMK) and A*0201-restricted (NLVPMK(FITC)VATV) peptides. Relative peptide-binding intensities of the HLA-A*30 ligand compared to the HLA-A*0201 ligand were calculated for each of the three cell lines as followed: [double positive A*30 binding cells/double positive A*0201 binding cells]*100. Background expression of cells, which were cultured without the synthetic peptides, was used as the blank value.
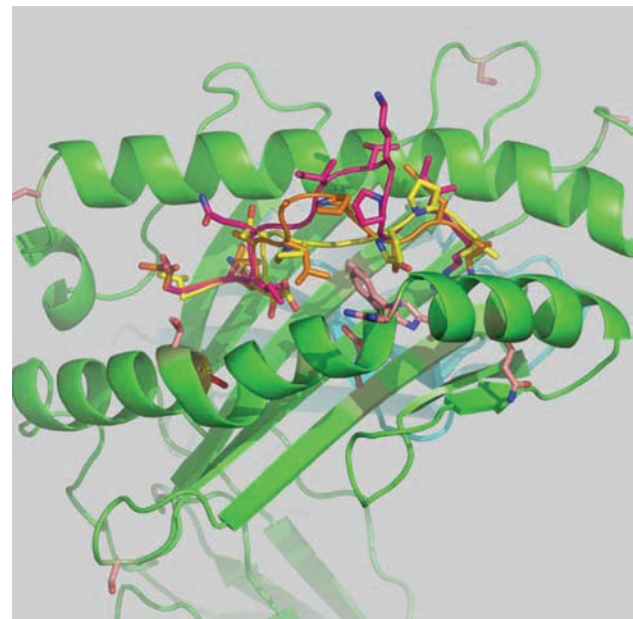
Figure 3. Homology-based model of HLA-A*3014L with the three shared peptide ligands. Modeling of the HLA-A*3001 and HLA-A*3014L structures was carried out using the SCWRL homology-based modeling server[20] using the crystal structure of HLA-A1101 (1Q94) as a template. Peptide templates for 9-mer (1Q94), 11-mer (2BVO) and 12-mer (3BW9) were superposed and merged with the HLA-A*3014L model. Peptide mutagenesis was then performed using DeepView[21] using the rotamer library to find the best side chain orientations with minimum steric clashes. Each model was then subjected to energy minimization as implemented in DeepView. The graphics program PyMOL (*http://www.pymol.org*) was used to generate the structure figures.

direct and indirect allo-recognition pathway than a mismatch with its most related allele HLA-A*3001. Indeed, mistyping HLA-A*3014L as an N allele has led to a severe GvHD in one patient transplanted with hematopoietic stem cells from an HLA-A*0201 homozygous donor.[4]

The peptide-binding region of HLA class I molecules is divided into six specificity pockets (A-F), as demonstrated by X-ray crystallography.[26,27] These pockets bind certain aa side chains from antigenic peptides. Previous functional and Edman degradation analyses of presented peptides have shown that position 2 and the C-terminal position (PΩ) of the presented peptides interacting with pockets B and F are the main anchor positions for binding to most known HLA molecules.[14,28] HLA-A*3001 is reported to be an exception to this classical paradigm.[16] Here the peptide is anchored by P3 in interaction with pocket D instead of P2 in pocket B. Additionally, P3 was identified as a main anchor position for HLA-A*01[29-31] and HLA-B*08.[32-35]

Lamberth *et al.* later described P3 of HLA-A*3001-bound peptides as a primary-secondary anchor to rule out the higher importance of the main (primary) anchor at the C-terminal peptide-binding position (PΩ).[15] These peptide-binding data analyses were achieved using positional scanning combinatorial peptide libraries (PSCPL)[15,16] and a peptide-binding assay[15] showing mutual concordance of the peptide-binding motifs.

A database search (*www.syfpeithi.de*) revealed that neither the HLA-A*3001- nor HLA-A*3014L-isolated peptides have been published previously as ligands for any other HLA class I molecule.[36] However, the search revealed that QKEITALAPSTMKI, a 14-mer peptide from aa positions 299-312 of actin and naturally eluted from the HLA class II molecule HLA-DRB1*1001, has a sequence similar to that of the newly identified 11-mer peptide sequence but is extended by two aa at the N-terminus (glutamine, Q; K) and one aa at the C-terminus (I).[37]

Table 2. Interacting Pocket B, D and F residues for structures and models.

Structure/Model	P2 peptide residue	Interacting Pocket B residues (<4 Å away)																	
		7	9	24	25	26	34	35	36	45	62	63	66	67	70	99	159	163	167
1Q94 9-mer	I	Y	-	-	-	-	-	-	-	-	-	E	-	-	-	Y	Y	-	-
3001 9-mer model	L	Y	-	-	-	-	-	-	-	M	-	E	N	V	-	Y	Y	-	-
2BVO 11-mer	A	Y	Y	-	-	-	-	-	-	-	-	E	N	M	-	Y	Y	-	-
3001 11-mer model	I	Y	-	-	-	-	-	-	-	M	-	E	N	-	-	Y	Y	-	-
3BW9 12-mer	P	Y	-	-	-	-	-	-	-	-	-	N	I	F	-	Y	Y	-	-
3001 12-mer model	N	Y	-	-	-	-	-	-	-	M	-	E	N	-	-	Y	Y	-	-

Structure/Model	P3 peptide residue	Interacting Pocket D residues (<4 Å away)											
		7	9	62	66	97	99	114	152	155	156	159	163
1Q94 9-mer	F	-	-	-	-	-	Y	-	-	Q	Q	Y	-
3001 9-mer model	D	-	-	-	-	-	Y	-	-	-	-	Y	-
2BVO 11-mer	F	-	-	-	N	-	Y	-	-	Q	L	Y	-
3001 11-mer model	T	-	-	-	-	-	Y	-	-	-	-	Y	-
3BW9 12-mer	S	-	-	-	-	-	Y	-	-	-	R	Y	-
3001 12-mer model	I	-	-	-	-	-	Y	-	-	Q	-	Y	-

Structure/Model	PΩ peptide residue	Interacting Pocket F residues (<4 Å away)																	
		70	73	74	76	77	80	81	84	95	96	97	114	116	123	124	142	143	146	147
1Q94 9-mer	K	-	-	-	-	D	T	-	Y	-	-	-	-	D	Y	-	-	T	-	-
3001 9-mer model	V	-	-	-	-	D	T	L	Y	-	-	-	-	D	Y	-	-	T	K	-
2BVO 11-mer	F	-	-	-	-	N	I	-	Y	I	-	-	-	Y	Y	-	-	T	K	-
3001 11-mer model	K	-	-	-	-	D	I	L	Y	-	-	-	-	H	Y	-	-	T	-	W
3BW9 12-mer	Y	-	-	Y	-	S	N	L	Y	-	-	R	-	S	Y	-	-	T	K	W
3001 12-mer model	R	-	-	-	-	D	T	-	Y	I	Q	I	-	H	Y	-	-	T	-	W

The table is shows the sub-sites of residues within the antigen-binding groove which bind side chains from antigenic peptides (pockets) using the pocket definition of Chelvanayagam.[19] For comparison, the table shows the A*11 structure used for the homology-based model of HLA-A*3014L and the 3001. Eighteen residues are designed as part of pocket B and thus contact P2 of bound peptides. In the A*11 structure and model, we confirmed that eight residues contact the P2 residue. In pocket D, formed by all amino acids contacting position P3, five residues contact the peptide at P3. Nineteen amino acids within the peptide-binding region are part of the F pocket and thus contact the C-terminus (PΩ) of the respective peptide. The structure as well as the model for A*3001 confirmed that 13 residues contact the C-terminal position of the designated peptide. Dashes indicate that a residue is not in contact with the peptide residue and thus not included in the respective pocket.

Peptide processing and presentation between class I and II molecules is fundamentally different. Presumably, class II peptides are longer and have completely different preferences for certain amino acids (aa) at the anchor positions. The three shared epitopes of HLA-A*3001 and A*3014L are consistent with the reported and confirmed A*3001 peptide motif, showing at the main anchor position PΩ predominantly the basic amino acids K, R or the hydrophobic residue V, respectively.[15,16] Because the shared epitopes feature sequences matching the peptide-binding motif for the analyzed HLA-A*30 alleles, it may be assumed that HLA-A*3001/HLA-A*3014L mismatches at aa position 164 are permissive and thus not allogeneic in HSCT.

By homology-based modeling, only marginal differences between the two HLA-A*30 alleles were found. Although the models look identical and the alleles appear to bind identical peptides, the Cys164Ser polymorphism could potentially generate additional flexibility within the peptide-binding groove, thereby influencing binding kinetics, particularly of peptides of lower affinity. Such an effect could have serious implications for transplantation. We analyzed the 9, 11 and 12-mer peptide epitopes presented by both HLA-A*3014L and HLA-A*3001 and the respective models to identify residues in contact with the peptide (cut-off <4 Å) using CONTACT software [CCP4 (1994)].[22] These polymorphic residues are unlikely to cause the alteration of P2. It can be assumed that P2 is not anchored deeply in pocket B because only 8 residues are bound at P2 in the PBR. This might explain the promiscuity of A*1101 and A*3001/14L-derived peptides regarding P2. Because the amino acid signature is in contact with P2 and P3, promiscuity might be affected by the lack of residues contacting the peptide within the heavy chain.

Prediction algorithms generally do not consider extraordinarily long peptides because it is known that the proteasome complex generates fragments with a Gaussian-like length distribution, preferring peptides 8 to 11 residues in length.[38] However, recent peptide elution studies have shown that HLA class I molecules are capable of presenting ligands of unusual length. The finding of a low expression allele presenting a long peptide is important and underlines that longer epitopes are possible. Whether or not these longer epitopes have to be considered as potential antigenic epitopes that can be recognized by T cells and should, therefore, be considered in prediction algorithms remains to be clarified.

Using the flow cytometric peptide-binding assay[11-13] we were able to confirm the binding of the 11-mer peptide (EITALAPSTMK) to HLA-A*3001 as well as to the HLA-A*3014L. We used B-LCLs naturally expressing the HLA molecules (A*02 and/or A*3001/14L) and showed that stable complexes of HLA-A*3014L and EITALAPSTMK are assembled after acid-treatment and reconstitution. The HLA-A*0201-restricted CMVpp65 (495-503) peptide (NLVPMVATV) has a SYFPEITHI score (www.syfpeithi.de) of 30 indicating a very strong binding. Although all used B-LCLs are HLA-A*0201 positive, the peptide-binding data reaches highest levels for the HLA-A*0201 homozygous B-LCLs compared to the two HLA-A heterozygous cell lines. For the HLA-A*30-restricted 11-mer ligand (EITALAPSTMK) no SYFPEITHI prediction score was available. But for 9-mer and 10-mer peptides derived from this sequence a SYFPEITHI score of less than 15 was reported. Therefore binding of this ligand to HLA-A*0201 molecules is rather unlikely and was also found to be 4-fold weaker compared to the strong A*0201 ligand for the HLA-A*0201-homozygous cell line.

One reason for the difference in the binding intensities for the HLA-A*30 ligand to the HLA-A*3001 versus A*3014L molecules is most probably due to the reduced cell surface expression and assumed instability of the HLA-A*3014L molecule. On the other hand, the reduced expression of HLA-A*3014L might lead to an increase of HLA-A*0201 cell surface expression altering the peptide-binding ratio for the benefit of the A*0201 peptide ligand.

We were able to show for the first time that an HLA low expression allele (HLA-A*3014L) presents peptides with features identical to those of its most related allele HLA-A*3001. Our results indicate that a mismatch at aa position 164 might be permissive. Therefore, mismatching of these alleles will presumably be of low allogenicity in allogeneic HSCT. The fact that a low expression variant is not only functional and able to present peptides, but also shares epitopes with its related variant leads to the conclusion that low expression variants need to be considered in the donor selection as permissive or non-permissive mismatches, respectively. Increasing knowledge of the expression levels of HLA expression variants, such as L and Q alleles, will help to improve HLA allogenicity prediction algorithms like HistoCheck[2] by confirming that these variants are fully functional. Taking all relevant factors into account, the results of our study make it possible to predict the immunogenicity of alternatively expressed alleles in a transplant setting.

In summary, from a clinical view, HLA-A*3014L should be considered to be a fully functional HLA allele and is best mismatched with HLA-A*3001. It is important to identify expression levels of HLA variants with questionable expression (Q alleles); alleles with a similar variation as described for HLA-A*3014L should be considered as relevant variants in matching procedures. This currently applies for HLA-A*3211Q, B*3565Q, B*3938Q, and Cw*0322Q. Due to a substitution of the cysteine residue at aa position 101 or 164, all of those alleles show a disruption of their disulfide bridge in the α2 domain. Based on our results, we hypothesize that these alleles have a low but clinically relevant protein surface expression. We are currently conducting a further study using the established sHLA secretion assay to investigate the expression levels of HLA-A*3211Q, B*3565Q and B*1308Q (Y159C).[6] Because of the additional cysteine at position 159 in B*1308Q the disulfide bond between aa position 101 and 164 is most probably affected altering the conformation and thereby the expression of the HLA-B molecule.

Authorship and Disclosures

References

1. Natarajan K, Li H, Mariuzza RA, Margulies DH. MHC class I molecules, structure and function. Rev Immunogenet. 1999;1(1):32-46.

2. Bade-Doeding C, Elsner HA, Eiz-Vesper B, Seltsam A, Holtkamp U, Blaszczyk R. A single amino-acid polymorphism in pocket A of HLA-A*6602 alters the auxiliary anchors compared with HLA-A*6601 ligands. Immunogenetics. 2004;56(2):83-8.

3. Prilliman KR, Jackson KW, Lindsey M, Wang J, Crawford D, Hildebrand WH. HLA-B15 peptide ligands are preferentially anchored at their C termini. J Immunol. 1999;162(12):7277-84.

4. Hirv K, Pannicke U, Mytilineos J, Schwarz K. Disulfide bridge disruption in the alpha2 domain of the HLA class I molecule leads to low expression of the corresponding antigen. Hum Immunol. 2006;67(8):589-96.

5. Warburton RJ, Matsui M, Rowland-Jones SL, Gammon MC, Katzenstein GE, Wei T, et al. Mutation of the alpha 2 domain disulfide bridge of the class I molecule HLA-A*0201. Effect on maturation and peptide presentation. Hum Immunol. 1994;39(4):261-71.

6. Hinrichs J, Figueiredo C, Hirv K, Mytilineos J, Blaszczyk R, Horn PA, et al. Discrimination of HLA null and low expression alleles by cytokine-induced secretion of recombinant soluble HLA. Mol Immunol. 2009;46(7):1451-7.

7. Oudshoorn M, Horn PA, Tilanus M, Yu N. Typing of potential and selected donors for transplant: methodology and resolution. Tissue Antigens. 2007;69 (Suppl 1):10-2.

8. Brodsky FM, Parham P, Barnstable CJ, Crumpton MJ, Bodmer WF. Monoclonal antibodies for analysis of the HLA system. Immunol Rev. 1979;47:3-61.

9. Barnstable CJ, Bodmer WF, Brown G, Galfre G, Milstein C, Williams AF, et al. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. Cell. 1978;14(1):9-20.

10. Hirosawa M, Hoshida M, Ishikawa M, Toya T. MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. Comput Appl Biosci. 1993;9(2):161-7.

11. Maeurer MJ, Martin D, Elder E, Storkus WJ, Lotze MT. Detection of naturally processed and HLA-A1-presented melanoma T-cell epitopes defined by CD8(+) T-cells' release of granulocyte-macrophage colony-stimulating factor but not by cytolysis. Clin Cancer Res. 1996;2(1):87-95.

12. Storkus WJ, Zeh HJ 3rd, Salter RD, Lotze MT. Identification of T-cell epitopes: rapid isolation of class I-presented peptides from viable cells by mild acid elution. J Immunother Emphasis Tumor Immunol. 1993;14(2):94-103.

13. Zeh HJ 3rd, Leder GH, Lotze MT, Salter RD, Tector M, Stuber G, et al. Flow-cytometric determination of peptide-class I complex formation. Identification of p53 peptides that bind to HLA-A2. Hum Immunol. 1994;39(2):79-86.

14. Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature. 1991;351(6324):290-6.

15. Lamberth K, Roder G, Harndahl M, Nielsen M, Lundegaard C, Schafer-Nielsen C, et al. The peptide-binding specificity of HLA-A*3001 demonstrates membership of the HLA-A3 supertype. Immunogenetics. 2008;60(11):633-43.

16. Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, et al. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. Immunome Res. 2008;4:2.

17. Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom. 1984;11(11):601.

18. Hildebrand WH, Domena JD, Shen SY, Lau M, Terasaki PI, Bunce M, et al. HLA-B15: a widespread and diverse family of HLA-B alleles. Tissue Antigens. 1994;43(4):209-18.

19. Chelvanayagam G. A roadmap for HLA-A, HLA-B, and HLA-C peptide binding specificities. Immunogenetics. 1996;45(1):15-26.

20. Wang Q, Canutescu AA, Dunbrack RL Jr. SCWRL and MolIDE: computer programs for side-chain conformation prediction and homology modeling. Nat Protoc. 2008;3(12):1832-47.

21. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997;18(15):2714-23.

22. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994;50(Pt 5):760-3.

23. DeLuca DS, Khattab B, Blaszczyk R. A modular concept of HLA for comprehensive peptide binding prediction. Immunogenetics. 2007;59(1):25-35.

24. DeLuca DS, Blaszczyk R. Implementing the modular MHC model for predicting peptide binding. Methods Mol Biol. 2007;409:261-71.

25. Benichou G. Direct and indirect antigen recognition: the pathways to allograft immune rejection. Front Biosci. 1999;4:D476-80.

26. Garrett TP, Saper MA, Bjorkman PJ, Strominger JL, Wiley DC. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. Nature. 1989;342(6250):692-6.

27. Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. J Mol Biol. 1991;219(2):277-319.

28. Parker KC, Bednarek MA, Hull LK, Utz U, Cunningham B, Zweerink HJ, et al. Sequence motifs important for peptide binding to the human MHC class I molecule, HLA-A2. J Immunol. 1992;149(11):3580-7.

29. Falk K, Rotzschke O, Takiguchi M, Grahovac B, Gnau V, Stevanovic S, et al. Peptide motifs of HLA-A1, -A11, -A31, and -A33 molecules. Immunogenetics. 1994;40(3):238-41.

30. Kubo RT, Sette A, Grey HM, Appella E, Sakaguchi K, Zhu NZ, et al. Definition of specific peptide motifs for four major HLA-A alleles. J Immunol. 1994;152(8):3913-24.

31. DiBrino M, Tsuchida T, Turner RV, Parker KC, Coligan JE, Biddison WE. HLA-A1 and HLA-A3 T cell epitopes derived from influenza virus proteins predicted by peptide binding motifs. J Immunol. 1993;151(11):5930-5.

32. DiBrino M, Parker KC, Shiloach J, Turner RV, Tsuchida T, Garfield M, et al. Endogenous peptides with distinct amino acid anchor residue motifs bind to HLA-A1 and HLA-B8. J Immunol. 1994;152(2):620-31.

33. Malcherek G, Falk K, Rotzschke O, Rammensee HG, Stevanovic S, Gnau V, et al. Natural peptide ligand motifs of two HLA molecules associated with myasthenia gravis. Int Immunol. 1993;5(10):1229-37.

34. Sutton J, Rowland-Jones S, Rosenberg W, Nixon D, Gotch F, Gao XM, et al. A sequence pattern for peptides presented to cytotoxic T lymphocytes by HLA B8 revealed by analysis of epitopes and eluted peptides. Eur J Immunol. 1993;23(2):447-53.

35. Arnett KL, Huang W, Valiante NM, Barber LD, Parham P. The Bw4/Bw6 difference between HLA-B*0802 and HLA-B*0801 changes the peptides endogenously bound and the stimulation of alloreactive T cells. Immunogenetics. 1998;48(1):56-61.

36. Singh-Jasuja H, Scherer HU, Hilf N, Arnold-Schild D, Rammensee HG, Toes RE, et al. The heat shock protein gp96 induces maturation of dendritic cells and down-regulation of its receptor. Eur J Immunol. 2000;30(8):2211-5.

37. Alvarez I, Collado J, Daura X, Colome N, Rodriguez-Garcia M, Gallart T, et al. The rheumatoid arthritis-associated allele HLA-DR10 (DRB1*1001) shares part of its repertoire with HLA-DR1 (DRB1*0101) and HLA-DR4 (DRB*0401). Arthritis Rheum. 2008;58(6):1630-9.

38. Wenzel T, Eckerskorn C, Lottspeich F, Baumeister W. Existence of a molecular ruler in proteasomes suggested by analysis of degradation products. FEBS Lett. 1994;349(2):205-9.