

## Review Article

# Soluble HLA Technology as a Strategy to Evaluate the Impact of HLA Mismatches

**Heike Kunze-Schumacher, Rainer Blasczyk, and Christina Bade-Doeding**

*Institute for Transfusion Medicine, Hannover Medical School, Medical Park, Feodor-Lynen-Straße 5, 30625 Hannover, Germany*

Correspondence should be addressed to Christina Bade-Doeding; [bade-doeding.christina@mh-hannover.de](mailto:bade-doeding.christina@mh-hannover.de)

Received 14 April 2014; Revised 15 July 2014; Accepted 5 August 2014; Published 1 September 2014

Academic Editor: Trevor Huyton

Copyright © 2014 Heike Kunze-Schumacher et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

HLA class I incompatibilities still remain one of the main barriers for unrelated bone marrow transplantation (BMT); hence the molecular understanding of how to mismatch patients and donors and still have successful clinical outcomes will guide towards the future of unrelated BMT. One way to estimate the magnitude of polymorphisms within the PBR is to determine which peptides can be selected by individual HLA alleles and subsequently presented for recognition by T cells. The features (structure, length, and sequence) of different peptides each confer an individual pHLA landscape and thus directly shape the individual immune response. The elution and sequencing of peptides by mass spectrometric analysis enable determining the *bona fide* repertoire of presented peptides for a given allele. This is an effective and simple way to compare the functions of allelic variants and make a first assessment of their degree of permissivity. We describe the methodology used for peptide sequencing and the limitations of peptide prediction tools compared to experimental methods. We highlight the altered peptide features that are observed between allelic variants and the need to discover the altered peptide repertoire in situations of “artificial” graft versus host disease (GvHD) that occur in HLA-specific hypersensitive immune responses to drugs.

## 1. Introduction

Human leucocyte antigen (HLA) class I incompatibilities represent the major barrier for successful outcome in haematopoietic stem cell transplantation (HSCT), a therapeutic strategy for the treatment of hematologic malignancies. The best clinical outcomes for unrelated HSCT can be achieved by an 8/8 match, meaning high resolution matching on HLA-A, -B, -C, and -DRB1 levels [1]. Certain HLA class II mismatches have also gained considerable attention in the selection of unrelated donors; most notably several studies on HLA-DP permissivity have been performed [2–5]. Such a perfect HLA match of a given donor:recipient pair is in most cases impossible to find; therefore, mobilized stem cells from well-matched but unrelated donors are used for transplantation [6]. Despite great care being taken to select the best possible match these transplants may be associated with significant risks of graft versus host disease (GvHD), graft failure, or transplant related mortality [7]. Clinical data have demonstrated that the risk of GvHD correlates with

the number of mismatches within the HLA molecule, where both the type of amino acid (AA) substitution and its location can directly influence the degree of histocompatibility [8]. For these reasons, a priority for certain unrelated mismatched donors is also given to those whose polymorphisms do not affect peptide binding. If there are differences between two allelic HLA subtypes, then this is likely to alter their presented peptide binding features; thus alloreactivity is expected [9]. If the peptide binding features are not altered, then alloreactive immune reactions are more unlikely; however, they cannot be fully excluded since the genetic variability in the population may give rise to certain minor histocompatibility antigens. The impact of polymorphisms located within the peptide binding region and the risk of severe acute GvHD is well known [10]; furthermore the residues described to impact the risk of GvHD when mismatched (AA 9, 99, 116, 156) have been directly associated with the peptide anchor p2 or p $\Omega$  [11] or indirectly with peptide binding through differential interaction with the peptide loading complex and peptide selection [12, 13].

The feasibility to find the best acceptable mismatch and have successful clinical outcomes depends on our understanding of the impact of a certain HLA polymorphism. It would be beneficial to have a fast and reliable method that would allow a prediction of individual mismatch magnitudes to help clinicians in deciding which donor: recipient pair has the best chance of successful transplantation. However, the impact of a given mismatch might differ for every single allelic variant and is furthermore affected by the proteomic content that is available for presentation in an HLA molecule, since peptides undergo several selection criteria before being presented in an HLA molecule on the cell surface (Figure 1). Alongside the functional allelic differences conferred by AA exchanges that influence the interaction of the HLA heavy chain with the peptide loading complex, the peptide cargo itself has to fulfill certain criteria. Proteasomally digested peptides are transported into the endoplasmic reticulum via the transporter associated with antigen processing (TAP) and loaded onto immature HLA class I molecules with the assistance of the peptide loading complex (PLC) [14]. Here, it has to be considered that certain HLA alleles select and present their peptides through an alternative pathway, since particular polymorphism can influence which allogenic peptides are selectively bound through conformational alteration of the HLA heavy chain [12]. Peptide selection that did not undergo selection and optimization of the PLC might lead to immunological reactions since it enables the presentation of poorly tolerated peptides. This might therefore mimic foreign proteomic content to immune effector cells and lead to rejection episodes when transplanted.

HLA molecules are highly variable; accepting a mismatched donor means that a comprehensive knowledge should be allocatable. An HLA molecule that is displayed to the immune system consists of three modules, the heavy chain, beta 2 microglobulin ( $\beta 2m$ ), and the individual peptide. While  $\beta 2m$  is invariant, the heavy chain as well as the peptide is variable, and there are thousands of possible peptide/HLA (pHLA) molecules that can be displayed to immune effector cells. Variations within the AA sequence of the heavy chain may therefore alter substantially the whole pHLA landscape, for that reason the knowledge about the magnitude of a distinct mismatch within the heavy chain is important to understand and/or predict the molecules function. These polymorphisms can affect (i) the individual peptide binding motif [15–18], (ii) peptide selection [12], and/or (iii) the individual pHLA landscape [19, 20] that will be scanned by immune effector cells. Due to the enormous degree of HLA class I polymorphisms ( $n$ ) and the quantity of possible mismatches [magnitude =  $1/2 \cdot n \cdot (n - 1)$ ] that can occur there is still no strategy available for the selection of the less permissive class I mismatch when no identical donor but multiple mismatched donors are available. A mean for a comprehensive measure of histocompatibility can in the context of the extensive polymorphism only take place as a systematic study of key alleles with distinct polymorphism. Certain mismatches as, for example, position 156 in the B\*44 allelic group have been studied extensively; clinical data [21] as well as *in vitro* data [12, 22–24] are available. However, for most allelic groups, there is a lack of clinical data or *in vitro*

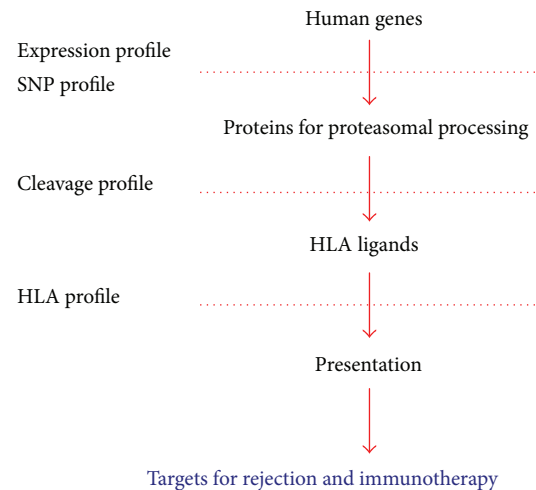


FIGURE 1: Assessment of targetable allogeneic epitopes. Every ligand has to pass several filters, before being presented in an HLA molecule. The first is the individual expression profile. Proteins then undergo further proteasomal processing, where the individual cleavage profile determines the sequence of possible HLA ligands. The last filter is the individual HLA profile that differs between allelic variants.

data that could be translated to histocompatibility. To enable a comprehensive and intelligent mismatching strategy, the fast and reproducible analysis of the peptide repertoire is required since it represents the most variable part of a HLA molecule.

## 2. Determination of Mismatch Impact Based on Peptide Sequences

HLA class I allelic peptide motifs have been described for the most frequent variants [26–29], but few studies have addressed the question to which extent variants of the same allelic HLA group differ in their peptide motifs and/or peptide repertoires [30–33].

HLA-B\*44:02<sup>Asp156</sup> and B\*44:03<sup>Leu156</sup> differ by a single AA exchange located in the PBR and are well-studied examples for strong alloreactive immunological episodes. Fleischhauer et al. described in 1990 a strong alloreactive response mediated by T cells that resulted in transplant rejection [21]. To explain this unexpected occurrence, peptide profiles of the mismatched variants have been analysed by the same author in 1994 [22]. Since no differences in peptide profiles could be detected, the unexpected alloreactivity could only be explained in 2003, where Macdonald et al. solved their structures [23] and found subtle alterations in the peptide binding. This example shows the distinct connection between peptide binding profiles and immunogenicity and started a new era of histocompatibility rating.

The nature of a given AA mismatch plays a significant role in deciding for histocompatibility; in that regard, it is thought that AA exchanges from the same chemical group are more permissive than those AAs from different chemical groups. The allelic variants B\*44:02<sup>Asp156</sup> and B\*44:35<sup>Glu156</sup> differ at a single mismatch; both of the AAs exchanged are

polar acidic and thus not expected to confer alteration of the PBR feature. However, the comparison of the peptides from B\*44:02<sup>Asp156</sup> and B\*44:35<sup>Glu156</sup> showed an unexpected alteration in the binding motif of the peptide's pΩ [34]. B\*44:35<sup>Glu156</sup> was found to bind and present a significantly high number of peptides of extraordinary length and to disfavor the binding and presentation of canonical peptides. The peptide data derived from B\*44:35<sup>Glu156</sup> highlighted a strong shift in the binding preference for Lys at the C-terminus, while this residue was completely absent in peptides eluted from B\*44:02<sup>Asp156</sup>. The observation of such a change in the peptide anchor in B\*44:35<sup>Glu156</sup> as a result of a highly conserved substitution at position 156 was unexpected. Based on the structure of B\*44:02<sup>Asp156</sup> (PDB 1M6O) [23] we were able to model B\*44:35<sup>Glu156</sup> bound to a peptide with the distinct binding motif and therefore explain the structural alteration. Cellular studies using alloreactive T cells supported the findings of the peptide diversities. In this case, the preliminary peptide data were essential to explain the strong T-cell mediated alloreactivity.

Further studies defined the differential impact of single mismatches on peptide binding and the paradox that an accumulation of multiple polymorphism located in the PBR does not alter the designated peptide binding motif.

Most class I alleles show the highest sequence polymorphism in the α1 and/or α2 domain that form the PBR and are encoded by exons 2 and 3, which therefore constitute the region of interest for histocompatibility matching. Currently, allelic variants with mismatches outside of the α1 or α2 domains are thought to have the same immune function and thus would represent a permissive mismatch in HSCT. HLA-B\*44:02:01:01 and B\*44:27 are considered to be functionally identical since they differ by the micropolymorphism Val199Ala, located in the α3 domain. To validate the theory that B\*44:02<sup>199Val</sup> and B\*44:27<sup>199Ala</sup> represent functionally identical alleles, we compared peptides and their features derived from both B\*44:02<sup>199Val</sup> and B\*44:27<sup>199Ala</sup>. The mismatch at residue 199 did not alter significantly the peptide motif, the peptide features, nor the peptide repertoire [35]; for that reason a single 199 mismatch for B\*44 variants might be considered as a permissive mismatch, but must be certainly validated by cellular assays. According to our finding Bettens et al. investigated in 2013 the alloreactivity of CD8+/CD137+ T cells in a mixed lymphocyte culture reaction and could demonstrate a complete lack of allorecognition between the Val199Ala mismatched variants [36]. These findings supported our conclusion that in the case of B\*44:02:01:01/B\*44:27 incompatibility, we could define histocompatibility in terms of the features of the bound peptides.

The knowledge of allele specific peptide binding data enables an isolated and individual assessment of every AA mismatch. Through comparative analysis of the peptides derived from HLA-B\*44:02 and B\*44:08, differing at residues 41, 45, and 46, we were able to exclusively assign position 45 of the heavy chain to significantly affect the peptide binding motif while residues 41 and 46 had no influence on the bound peptides [16]. While HLA-B\*44:02 has a peptide anchor motif

of Glu at p2, B\*44:08 binds at p2 Gln and Leu. The peptide data provide evidence that mismatches at position 45 for B\*44 variants should be avoided, since it could be assumed that an alteration of the peptide features would most likely lead to a conversion of the accessible pHLA surface and consequently affect T-cell recognition. The comparison of peptide data for, HLA-B\*44:02 and B\*44:06, both of these subtypes are being distinguished by 7 residues polymorphic (24, 32, 41, 45, 63, 67, and 80), showed a differential impact on peptide binding and allowed for assignment of each residue within the heavy chain to a certain residue within the peptide [11]. Residues 45, 63, and 67 are highly polymorphic and show in all pHLA structures analysed abundant contact frequencies to the anchoring p1 and p2 position of peptides [11], highlighting their role in peptide specificity. By soluble HLA technology we could demonstrate the differences in peptide anchoring. While B\*44:02 is not anchored by p1, but by Glu at p2 and Phe, Tyr, and Trp at the C-term, B\*44:06 is anchored by Asp and Glu at p1, Pro and Ala at p2, and Tyr and Trp at the C-term. Moreover, MALDI-TOF analysis (a tool for the mass analysis of peptides) of low or high affinity peptide pools (Figure 2) suggest that in contrast to B\*44:02, B\*44:06 preferably selects high affinity peptides and binds peptides tightly. Additionally, B\*44:02 had a tendency to bind extraordinary long peptides whereas peptides bound to B\*44:06 were restricted to the canonical length of 8–10 AAs. The knowledge of peptide data for B\*44:06 allowed detailed structural investigation of the mismatch impact. Therefore we could demonstrate that in both alleles Ser167 allows the nonpolymorphic Arg170 to hydrogen bond with the peptide p1 residue. For B\*44:06, molecular modeling suggests that the Glu63 > Asn63 polymorphism enhances the preference of the P1 pocket for acidic AAs [11]. Translating these observations into histocompatibility, residue 63 should be considered as a nonpermissive mismatch. This is further supported by the observation that its frequency in contacting the peptide at main positions (p1, p2) could be observed in >95% of pHLA structures reflecting a distinct constraint of peptide motifs for B44 variants [11]. This data allowed “alloreactivity ranking” of the mismatched residues within the B\*44:06 heavy chain.

Peptide data from HLA-B\*44:09 showed how the 5 polymorphic residues (77, 80, 81, 82, and 83) distinguishing B\*44:09 from B\*44:02 affected the peptide binding motif. Both alleles illustrate an E at p2. In contrast to the C-terminal peptide binding motif of B\*44:02, Trp, Phe, Tyr, or Leu, B\*44:09 derived peptides are restricted predominantly to Leu or Phe. A small percentage of peptides were shared between the allotypes, those peptides contained the restricted B\*44:09 anchor motif of Phe or 10 Leu at the pΩ position [17]. The modeled structure of B\*44:09 reveals that residues 77 and 80 contact the peptide main chain at the peptide's pΩ; however, residue 81 contacts the peptide's pΩ side chain and appears to be the main cause of sequence specificity at the peptide's C-term. Similarly as described for B\*44:06, residue 81 in B\*44:09 could be demonstrated to have a high impact on peptide specificity.

It could also be demonstrated that certain alleles sharing the same peptide motif with their mismatched allelic variant show a higher affinity for peptides of an extraordinary length,

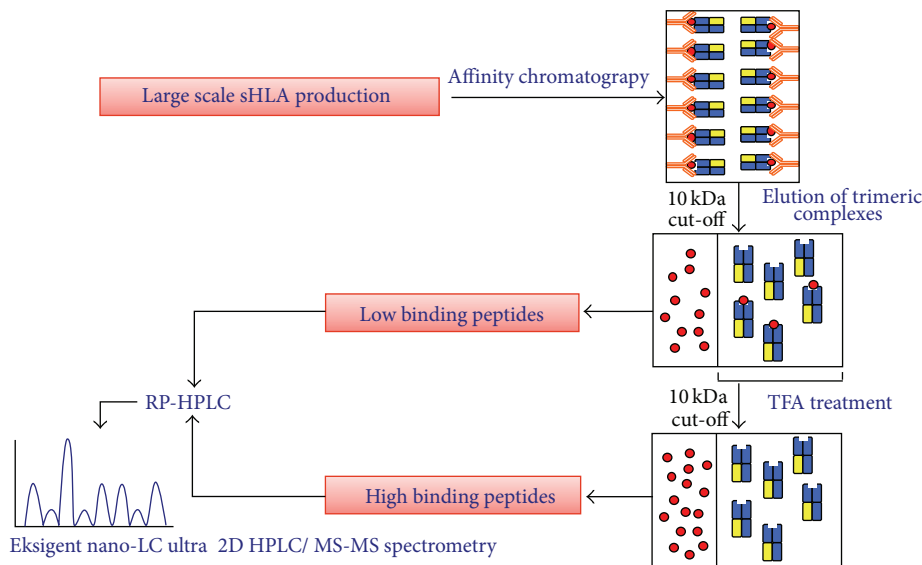


FIGURE 2: Generation of peptide sequences from sHLA molecules. This figure gives a schematic overview of the steps towards the generation of peptide sequences obtained from sHLA molecules. The HLA heavy chain is given in blue;  $\beta 2m$  is shown in yellow; and the peptide is shown in red. Cell culture supernatants containing sHLA molecules are passed over an N-hydroxysuccinimide- (NHS-) activated HiTrap column coupled to mAb w6/32. Trimeric complexes (class I heavy chain,  $\beta 2m$  and peptide) are eluted using a pH 2.7 elution buffer. Here, peptides from sHLA complexes can be differentiated into low and high binding peptides. The trimeric elution fractions are filtered through a 10 kDa cut-off membrane and the peptides detected in the flow through are considered to be of low affinity. The retentate containing dimeric (heavy chain and  $\beta 2m$ ) as well as trimeric complexes is then treated with 0.1% trifluoroacetic acid (TFA) to elute high binding peptides that can finally be separated by filtration through an additional 10 kDa cut-off membrane. Flow through fractions containing the low or high affinity peptides are subjected to mass spectrometric analysis using an Eksigent NanoLC Ultra 2D HPLC coupled to an orbitrap ion trap. Database queries can finally be carried out using Mascot software [25] incorporating the IPI human and the respective decoy databases.

an observation that is only possible through extended peptide sequencing. This phenomenon could be demonstrated for peptides bound to subtypes of the HLA-B\*41 group, HLA-B\*41:03, and B\*41:04 [20]. Further examination of these variants bound to previous sequenced peptides by high resolution X-ray crystallography demonstrated that polymorphism at positions 97 and 114 in B\*41 variants have a strong impact on the steric and electrostatic properties of the PBR and explained their divergent peptide confirmations.

### 3. Individual Peptide Repertoires Explain the Basis of GvHD-Like Drug-Mediated Hypersensitivity Reactions

Another example how a single mismatch within the HLA heavy chain influences the immunological self and guides towards biased T-cell responses is the alteration of presented peptides through the interference of small drug molecules, resulting in the phenomena of HLA-associated drug hypersensitivity. Here, a single AA mismatch within the PBR might change the susceptibility of an allele to allow a small drug molecule to bind and modulate the PBR as it could be demonstrated for the abacavir sensitive allele HLA-B\*57:01 [37]. Here, abacavir is able to occupy part of the PBR and modifies the selected peptide repertoire thereby causing a strong T-cell-mediated immune response that is resolved upon withdrawal of medication.

For the carbamazepine sensitive allelic variants HLA-A\*31:01 and -B\*15:02 we could recently demonstrate distinct functional differences through sequencing of their bound peptides (Kunze-Schumacher et al., 2014, manuscript in preparation).

Understanding the mechanistic basis of drug-mediated hypersensitivity reactions can only take place as a measurement of the allele specific individual peptide repertoire.

### 4. Soluble HLA Technology

Peptide sequencing by the use of soluble HLA (sHLA) technology is a powerful and convenient tool towards the understanding of histocompatibility and high throughput screening of the individual allele specific immunopeptidome. Another possibility to obtain peptides is the classical way of isolating membrane bound HLA molecules from cells; however, this method necessitates high numbers of cellular material. Besides, treatment of cells using lysis buffers might lead to failure of obtaining low affinity peptides, a problem that is not of any concern when using sHLA technology. It could be demonstrated that peptides obtained by soluble technology show the same repertoire as peptides obtained from membrane bound molecules [38]. Furthermore sHLA molecules select their peptides through the same loading pathway and are associated with the peptide loading complex in the same way than membrane bound molecules [12]. Soluble HLA technology thus represents a simple and easy instrument for the determination of mismatch magnitude.

For the use of sHLA molecules exon 1 through 4 from a given HLA variant is cloned into an appropriate expression vector and transferred preferentially into an HLA class I negative cell line, for example, LCL 721.221 cells. When the question has to be answered is what peptides are presented during a viral infection or in certain tumor cell, it is as well possible to clone a recombinant C-terminal tag (e.g., V5, c-Myc) for further purification, which would enable the exclusive purification of the desired recombinant HLA molecule without contaminations of native HLA molecules. The highest rate of sHLA expressing cells is achieved through lentiviral transduction of the desired HLA subtype in target cells as described previously [12]. For large scale sHLA production cells are transferred into bioreactors and sHLA molecules can be directly purified from the cell culture supernatant using an appropriate antibody, for example, W6/32 (Figure 2) bound to an affinity column. Peptides can be directly isolated from trimeric sHLA complexes and differentiated into low or high affinity peptides.

Soluble HLA technology offers the flexibility to achieve high amounts of proteins that are secreted in the cell culture supernatant without the need of high cell numbers. To quantify the protein concentration a sandwich enzyme-linked immunosorbent assay (ELISA) can be used as a simple and easy method for high throughput screening. Here, the anti-HLA-A-B-C W6/32 [39, 40] monoclonal antibody is employed as capture antibody; HRP-conjugated anti- $\beta$ 2m mAb serves as the detection antibody [12].

## 5. Peptide Data Are the Basis for Bioinformatic Prediction Tools

Numerous peptide binding prediction algorithms are available; they can be categorised into three major groups: motif and scoring matrix based methods, hidden Markov models, and artificial neural networks [41–45]. However, all those prediction algorithms are based on the limited knowledge of available peptide binding data that means prediction is only possible for a small fraction of known MHC proteins. The growing number of HLA alleles makes it impossible to provide prediction tools for variants or certain polymorphism where no experimental data exist. Structure based methods for predicting the allogenicity of a certain mismatch can only take place for alleles with known peptide sequences. Experimental peptide binding data and allele specific peptide motifs allow for the prediction and ranking of T-cell epitopes by certain prediction algorithms such as SYFPEITHI [46, 47], NetMHC [48], RANKPEP [49], PeptideCheck, [50] or BIMAS [51]. Here it has to be taken into account that *in silico* peptide binding prediction such as for vaccine development can only take place when peptide sequences and binding specificities for an allele of interest have been identified previously. If the allele specific peptide binding features, the motif, or the preferred length is known, it is possible to predict peptides and peptide binding data from desired proteins such as virus epitopes, which might be tested by *in vitro* systems [52–54]. However, all these programs are not able to address the binding and presentation of noncanonical

peptides (>8–10 AAs) [55] that might elicit strong alloreactive T-cell responses [56, 57] and it is well known to date that certain HLA variants bind peptides of unusual length [20, 58, 59]. Several long peptides for HLA class I molecules have been predicted by extending known shorter epitopes or by screening peptide libraries, including overlapping 16-mer [60] or 15-mer peptides [61, 62]. Even the binding of peptides with a length up to 25 AAs could be observed [58], suggesting that the length limitation of naturally processed HLA bound peptides is primarily controlled by their availability following antigen processing. The knowledge of the allele specific self peptidome will help to update current programs and bypass these limitations.

The prerequisite for comprehensive prediction tools is the knowledge about functional differences of HLA key alleles. Two of these tools, namely, *histocheck* [63] and *pocketcheck* [11], will be explained in the following. *Histocheck* is a tool to support an estimation of the allogenic potential between mismatched HLA variants. The rating score between two mismatched variants is based on the dissimilarity of the mismatched AAs. Furthermore, *Histocheck* designates and visualizes TCR and/or peptide binding AAs and directs towards estimation of histocompatibility. These data are based on available structures of pHLA complexes adapted from the classical pocket definitions, specified in 1991 by Saper et al. [64] and updated in 1996 by Chelvanayagam [65]. There have been several attempts to define which positions in the HLA binding groove influence the specificity of bound AAs at each position in the peptide based on X-ray crystallography. Structure databases (e.g., RCSB protein data bank) are the basis for molecular modeling of alleles whose structures have not been solved by X-ray crystallography, yet [16, 17], but have been investigated on the basis of their presented peptides. The generation of these structural data is only possible in the context of known allele specific peptides. Over the years more than 100 crystal structures became available that facilitated a broader definition of pockets, *pocketcheck* [11]. X-ray crystallographic determination of pHLA class I complexes provided valuable information for understanding how peptides bind to individual HLA class I molecules. Precisely, the allogenicity of a given mismatch is based on the information which residues within the HLA PBR would contact which peptide residue.

Several attempts have been taken to predict peptide binding to MHC molecules [49, 51, 66–71] for the assessment of permissivity or for vaccine development. For cellular therapeutics, CTLs that target infected cells require the identification of epitopes that distinguish healthy and infected cells. The knowledge about the self-proteomic content is therefore imperative for comprehensive and updated prediction tools.

If predicted epitopes of pathogenic origin would ever be presented *in vivo* remains questionable, before they have not been isolated and determined. Therefore “native” peptide data represent the prerequisite for all further prediction tools.

Successful prediction programs may significantly enlarge the donor pool for cellular, peptide based vaccination therapies and for certain patients by enabling a ranking of mismatches, building a bridge towards intelligent mismatching for better clinical outcomes.

## 6. Conclusion

The knowledge about the impact of distinct HLA class I mismatches in the context of experimental data such as peptide data and functional T-cell studies along with clinical data, if available, is appreciated by the feasibility to mismatch patients and still have successful clinical outcomes. Our approach in dividing peptides generated by sHLA technology into low or high affinity also allows a prediction of the relative half-life time and thus relative immunogenicity of a given pHLA complex.

Given the fact that every HLA allele has an individual peptide repertoire, it becomes obvious how many single pHLA molecules are available for an antigenic T cell. For that reason it has been proposed that it should be assessed whether GvHD and GvL can be related to the recipient specific peptide repertoire and its set of shared peptides by the donor HLA molecule [72]. While the risks associated with HLA-mismatches for unrelated bone marrow transplantation are clear, it becomes important to understand how certain peptides are selected by distinct alleles and thus how a given mismatch can influence the peptide selection and presentation process.

Towards that line, soluble HLA technology is a simple and powerful instrument for the detection of peptide binding motifs from allelic HLA variants. As a tool for high throughput analysis of polymorphism on HLA function, peptide sequencing enables evaluation of the best match and to highlight the potential risks of unmatched transplantation. Furthermore, peptide data represent the prerequisites for successful antitumor therapies. However, the high cost of peptide sequencing limits the acquisition of binding data. For that reason, it is imperative to select key alleles that enable the prediction of different allelic subtypes. Peptide motif based ranking of allelic mismatches can contribute to the donor selection process when no HLA-identical donor is available to prevent mismatching from being a matter of chance.

## Abbreviations

HLA: Human leucocyte antigen  
 HSCT: Hematopoietic stem cell transplantation  
 GvHD: Graft versus host disease  
 pHLA: Peptide-HLA complex  
 PBR: Peptide binding region  
 AA: Amino acid.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] S. J. Lee, J. Klein, M. Haagenson et al., "High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation," *Blood*, vol. 110, no. 13, pp. 4576–4583, 2007.
- [2] K. Fleischhauer, F. Locatelli, M. Zecca et al., "Graft rejection after unrelated donor hematopoietic stem cell transplantation for thalassemia is associated with nonpermissive HLA-DPB1 disparity in host-versus-graft direction," *Blood*, vol. 107, no. 7, pp. 2984–2992, 2006.
- [3] E. Zino, L. Vago, S. Di Terlizzi et al., "Frequency and targeted detection of HLA-DPB1 T cell epitope disparities relevant in unrelated hematopoietic stem cell transplantation," *Biology of Blood and Marrow Transplantation*, vol. 13, no. 9, pp. 1031–1040, 2007.
- [4] R. Crocchiolo, E. Zino, L. Vago et al., "Nonpermissive HLA-DPB1 disparity is a significant independent risk factor for mortality after unrelated hematopoietic stem cell transplantation," *Blood*, vol. 114, no. 7, pp. 1437–1444, 2009.
- [5] B. Mierzejewska, P. M. Schroder, C. E. Baum et al., "Early acute antibody-mediated rejection of a negative flow crossmatch 3rd kidney transplant with exclusive disparity at HLA-DP," *Human Immunology*, vol. 75, no. 8, pp. 703–708, 2014.
- [6] T. Zuckerman and J. M. Rowe, "Alternative donor transplantation in acute myeloid leukemia: which source and when?" *Current Opinion in Hematology*, vol. 14, no. 2, pp. 152–161, 2007.
- [7] E. W. Petersdorf, "Risk assessment in haematopoietic stem cell transplantation: histocompatibility," *Best Practice and Research: Clinical Haematology*, vol. 20, no. 2, pp. 155–170, 2007.
- [8] E. W. Petersdorf, "Optimal HLA matching in hematopoietic cell transplantation," *Current Opinion in Immunology*, vol. 20, no. 5, pp. 588–593, 2008.
- [9] M. A. Fernandez-Vina, T. Wang, S. J. Lee et al., "Identification of a permissible HLA mismatch in hematopoietic stem cell transplantation," *Blood*, vol. 123, no. 8, pp. 1270–1278, 2014.
- [10] J. Pidala, T. Wang, M. Haagenson et al., "Amino acid substitution at peptide-binding pockets of HLA class I molecules increases risk of severe acute GVHD and mortality," *Blood*, vol. 122, no. 22, pp. 3651–3658, 2013.
- [11] T. Huyton, N. Ladas, H. Schumacher, R. Blasczyk, and C. Bade-Doeding, "Pocketcheck: updating the HLA class I peptide specificity roadmap," *Tissue Antigens*, vol. 80, no. 3, pp. 239–248, 2012.
- [12] S. Badrinath, P. Saunders, T. Huyton et al., "Position 156 influences the peptide repertoire and tapasin dependency of human leukocyte antigen B2.2 microglobulin allotypes," *Haematologica*, vol. 97, no. 1, pp. 98–106, 2012.
- [13] D. Zernich, A. W. Purcell, W. A. Macdonald et al., "Natural HLA class I polymorphism controls the pathway of antigen presentation and susceptibility to viral evasion," *Journal of Experimental Medicine*, vol. 200, no. 1, pp. 13–24, 2004.
- [14] P. A. Wearsch and P. Cresswell, "The quality control of MHC class I peptide loading," *Current Opinion in Cell Biology*, vol. 20, no. 6, pp. 624–631, 2008.
- [15] C. Bade-Doeding, T. Huyton, B. Eiz-Vesper, and R. Blasczyk, "The composition of the F pocket in HLA-A\*74 generates C-terminal promiscuity among its bound peptides," *Tissue Antigens*, vol. 78, no. 5, pp. 378–381, 2011.
- [16] S. Badrinath, T. Huyton, H. Schumacher, R. Blasczyk, and C. Bade-Doeding, "Position 45 influences the peptide binding motif of HLA-B\*44:08," *Immunogenetics*, vol. 64, no. 3, pp. 245–249, 2012.
- [17] T. Huyton, H. Schumacher, R. Blasczyk, and C. Bade-Doeding, "Residue 81 confers a restricted C-terminal peptide binding motif in HLA-B\*44:09," *Immunogenetics*, vol. 64, no. 9, pp. 663–668, 2012.

- [18] K. R. Prilliman, D. Crawford, H. D. Hickman, K. W. Jackson, J. Wang, and W. H. Hildebrand, "Alpha-2 domain polymorphism and HLA class I peptide loading," *Tissue Antigens*, vol. 54, no. 5, pp. 450–460, 1999.
- [19] F. E. Tynan, D. Elhassen, A. W. Purcell et al., "The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation," *Journal of Experimental Medicine*, vol. 202, no. 9, pp. 1249–1260, 2005.
- [20] C. Bade-Döding, A. Theodossis, S. Gras et al., "The impact of human leukocyte antigen (HLA) micropolymorphism on ligand specificity within the HLA-B\*41 allotypic family," *Haematologica*, vol. 96, no. 1, pp. 110–118, 2011.
- [21] K. Fleischhauer, N. A. Kernan, R. J. O'Reilly, B. Dupont, and S. Y. Yang, "Bone marrow-allograft rejection by T lymphocytes recognizing a single amino acid difference in HLA-B44," *The New England Journal of Medicine*, vol. 323, no. 26, pp. 1818–1822, 1990.
- [22] K. Fleischhauer, D. Avila, F. Vilbois et al., "Characterization of natural peptide ligands for HLA-B\*4402 and -B\*4403: implications for peptide involvement in allorecognition of a single amino acid change in the HLA-B44 heavy chain," *Tissue Antigens*, vol. 44, no. 5, pp. 311–317, 1994.
- [23] W. A. Macdonald, A. W. Purcell, N. A. Mifsud et al., "A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T cell recognition," *Journal of Experimental Medicine*, vol. 198, no. 5, pp. 679–691, 2003.
- [24] J. Herman, V. Jongeneel, D. Kuznetsov, and P. G. Coulie, "Differences in the recognition by CTL of peptides presented by the HLA-B\*4402 and the HLA-B\*4403 molecules which differ by a single amino acid," *Tissue Antigens*, vol. 53, no. 2, pp. 111–121, 1999.
- [25] M. Hirosawa, M. Hoshida, M. Ishikawa, and T. Toya, "MAS-COT: multiple alignment system for protein sequences based on three-way dynamic programming," *Computer Applications in the Biosciences*, vol. 9, no. 2, pp. 161–167, 1993.
- [26] K. Falk, O. Rotzschke, S. Stevanovic et al., "Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules," *Nature*, vol. 351, no. 6324, pp. 290–296, 1991.
- [27] R. T. Kubo, A. Sette, H. M. Grey et al., "Definition of specific peptide motifs for four major HLA-A alleles," *Journal of Immunology*, vol. 152, no. 8, pp. 3913–3924, 1994.
- [28] R. Maier, K. Falk, O. Rotzschke et al., "Peptide motifs of HLA-A3, -A24, and -B7 molecules as determined by pool sequencing," *Immunogenetics*, vol. 40, no. 4, pp. 306–308, 1994.
- [29] M. DiBrino, K. C. Parker, J. Shiloach et al., "Endogenous peptides with distinct amino acid anchor residue motifs bind to HLA-A1 and HLA-B8," *Journal of Immunology*, vol. 152, no. 2, pp. 620–631, 1994.
- [30] T. Sudo, N. Kamikawaji, A. Kimura et al., "Differences in MHC class I self peptide repertoires among HLA-A2 subtypes," *The Journal of Immunology*, vol. 155, no. 10, pp. 4749–4756, 1995.
- [31] P. Krausa, C. Münz, W. Keilholz et al., "Definition of peptide binding motifs amongst the HLA-A\*30 allelic group," *Tissue Antigens*, vol. 56, no. 1, pp. 10–18, 2000.
- [32] K. R. Prilliman, K. W. Jackson, M. Lindsey, J. Wang, D. Crawford, and W. H. Hildebrand, "HLA-B15 peptide ligands are preferentially anchored at their C termini," *Journal of Immunology*, vol. 162, no. 12, pp. 7277–7284, 1999.
- [33] C. Bade-Doeding, D. S. DeLuca, A. Seltsam, R. Blasczyk, and B. Eiz-Vesper, "Amino acid 95 causes strong alteration of peptide position P $\Omega$  in HLA-B\*41 variants," *Immunogenetics*, vol. 59, no. 4, pp. 253–259, 2007.
- [34] S. Badrinath, T. Huyton, H. Kunze-Schumacher, H. Andreas Elsner, R. Blasczyk, and C. Bade-Doeding, "Differential impact of HLA-B\*44 allelic mismatches at position 156 on peptide binding specificities and T-cell diversity," *Journal of Stem Cell Research & Therapy*, vol. 4, no. 4, 2014.
- [35] C. Bade-Doeding, P. Cano, T. Huyton et al., "Mismatches outside exons 2 and 3 do not alter the peptide motif of the allele group B\*44:02P," *Human Immunology*, vol. 72, no. 11, pp. 1039–1044, 2011.
- [36] F. Bettens, U. Schanz, and J. M. Tiercy, "Lack of recognition of HLA class I mismatches outside  $\alpha 1/\alpha 2$  domains by CD8+ alloreactive T lymphocytes: the HLA-B44 paradigm," *Tissue Antigens*, vol. 81, no. 6, pp. 414–418, 2013.
- [37] P. T. Illing, J. P. Vivian, N. L. Dudek et al., "Immune self-reactivity triggered by drug-modified HLA-peptide repertoire," *Nature*, vol. 486, no. 7404, pp. 554–558, 2012.
- [38] K. E. Scull, N. L. Dudek, A. J. Corbett et al., "Secreted HLA recapitulates the immunopeptidome and allows in-depth coverage of HLA A\*02:01 ligands," *Molecular Immunology*, vol. 51, no. 2, pp. 136–142, 2012.
- [39] C. J. Barnstable, W. F. Bodmer, G. Brown et al., "Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens—new tools for genetic analysis," *Cell*, vol. 14, no. 1, pp. 9–20, 1978.
- [40] F. M. Brodsky, P. Parham, and C. J. Barnstable, "Monoclonal antibodies for analysis of the HLA system," *Immunological Reviews*, vol. 47, pp. 3–61, 1979.
- [41] Y. Kim, J. Sidney, C. Pinilla, A. Sette, and B. Peters, "Derivation of an amino acid similarity matrix for peptide:MHC binding and its application as a Bayesian prior," *BMC Bioinformatics*, vol. 10, article 394, 2009.
- [42] H. Mamitsuka, "Predicting peptides that bind to MHC molecules using supervised learning of hidden Markov models," *Proteins*, vol. 33, no. 4, pp. 460–474, 1998.
- [43] M. Nielsen and O. Lund, "NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction," *BMC Bioinformatics*, vol. 10, article 296, 2009.
- [44] C. Zhang, M. G. Bickis, F. Wu, and A. J. Kuslik, "Optimally-connected Hidden Markov models for predicting MHC-binding peptides," *Journal of Bioinformatics and Computational Biology*, vol. 4, no. 5, pp. 959–980, 2006.
- [45] Y. Zhao, B. Gran, C. Pinilla et al., "Combinatorial peptide libraries and biometric score matrices permit the quantitative analysis of specific and degenerate interactions between clonotypic TCR and MHC peptide ligands," *Journal of Immunology*, vol. 167, no. 4, pp. 2130–2141, 2001.
- [46] H. Rammensee, J. Bachmann, N. P. N. Emmerich, O. A. Bachor, and S. Stevanović, "SYFPEITHI: database for MHC ligands and peptide motifs," *Immunogenetics*, vol. 50, no. 3-4, pp. 213–219, 1999.
- [47] M. M. Schuler, M. Nastke, and S. Stevanović, "SYFPEITHI: database for searching and T-cell epitope prediction," *Methods in Molecular Biology*, vol. 409, pp. 75–93, 2007.
- [48] C. Lundegaard, K. Lamberth, M. Harndahl, S. Buus, O. Lund, and M. Nielsen, "NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11," *Nucleic Acids Research*, vol. 36, pp. W509–W512, 2008.
- [49] P. A. Reche, J. Glutting, H. Zhang, and E. L. Reinherz, "Enhancement to the RANKPEP resource for the prediction of peptide

- binding to MHC molecules using profiles," *Immunogenetics*, vol. 56, no. 6, pp. 405–419, 2004.
- [50] D. S. DeLuca and R. Blasczyk, "Implementing the modular MHC model for predicting peptide binding," *Methods in Molecular Biology*, vol. 409, pp. 261–271, 2007.
- [51] K. C. Parker, M. A. Bednarek, and J. E. Coligan, "Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains," *The Journal of Immunology*, vol. 152, no. 1, pp. 163–175, 1994.
- [52] J. J. Miles, D. Elhassen, N. A. Borg et al., "CTL recognition of a bulged viral peptide involves biased TCR selection," *The Journal of Immunology*, vol. 175, no. 6, pp. 3826–3834, 2005.
- [53] F. E. Tynan, N. A. Borg, J. J. Miles et al., "High resolution structures of highly bulged viral epitopes bound to major histocompatibility complex class I: implications for T-cell receptor engagement and T-cell immunodominance," *The Journal of Biological Chemistry*, vol. 280, no. 25, pp. 23900–23909, 2005.
- [54] F. E. Tynan, H. H. Reid, L. Kjer-Nielsen et al., "A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule," *Nature Immunology*, vol. 8, no. 3, pp. 268–276, 2007.
- [55] S. R. Burrows, J. Rossjohn, and J. McCluskey, "Have we cut ourselves too short in mapping CTL epitopes?" *Trends in Immunology*, vol. 27, no. 1, pp. 11–16, 2006.
- [56] K. J. Green, J. J. Miles, J. Tellam, W. J. M. van Zuylen, G. Connolly, and S. R. Burrows, "Potent T cell response to a class I-binding 13-mer viral epitope and the influence of HLA micropolymorphism in controlling epitope length," *European Journal of Immunology*, vol. 34, no. 9, pp. 2510–2519, 2004.
- [57] F. E. Tynan, S. R. Burrows, A. M. Buckle et al., "T cell receptor recognition of a "super-bulged" major histocompatibility complex class I-bound peptide," *Nature Immunology*, vol. 6, no. 11, pp. 1114–1122, 2005.
- [58] M. J. Bell, J. M. Burrows, R. Brennan et al., "The peptide length specificity of some HLA class I alleles is very broad and includes peptides of up to 25 amino acids in length," *Molecular Immunology*, vol. 46, no. 8-9, pp. 1911–1917, 2009.
- [59] M. J. Rist, A. Theodossis, N. P. Croft et al., "HLA peptide length preferences control CD8<sup>+</sup> T cell responses," *The Journal of Immunology*, vol. 191, no. 2, pp. 561–571, 2013.
- [60] M. Ghei, D. F. Stroncek, and M. Provenzano, "Analysis of memory T lymphocyte activity following stimulation with overlapping HLA-A\*2402, A\*0101 and Cw\*0402 restricted CMV pp65 peptides," *Journal of Translational Medicine*, vol. 3, article 23, 2005.
- [61] S. Matsumura, H. Kita, X. He et al., "Comprehensive mapping of HLA-A0201-restricted CD8 T-cell epitopes on PDC-E2 in primary biliary cirrhosis," *Hepatology*, vol. 36, no. 5, pp. 1125–1134, 2002.
- [62] S. L. Slezak, M. Bettinotti, S. Selleri, S. Adams, F. M. Marincola, and D. F. Stroncek, "CMV pp65 and IE-1 T cell epitopes recognized by healthy subjects," *Journal of Translational Medicine*, vol. 5, article 17, 2007.
- [63] H.-A. Elsner, D. DeLuca, J. Strub, and R. Blasczyk, "HistoCheck: rating of HLA class I and II mismatches by an internet-based software tool," *Bone Marrow Transplantation*, vol. 33, no. 2, pp. 165–169, 2004.
- [64] M. A. Saper, P. J. Bjorkman, and D. C. Wiley, "Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution," *Journal of Molecular Biology*, vol. 219, no. 2, pp. 277–319, 1991.
- [65] G. Chelvanayagam, "A roadmap for HLA-A, HLA-B, and HLA-C peptide binding specificities," *Immunogenetics*, vol. 45, no. 1, pp. 15–26, 1996.
- [66] S. Buus, S. L. Lauemøller, P. Worning et al., "Sensitive quantitative predictions of peptide-MHC binding by a "Query by Committee" artificial neural network approach," *Tissue Antigens*, vol. 62, no. 5, pp. 378–384, 2003.
- [67] S. Carrasco Pro, M. Zimic, and M. Nielsen, "Improved pan-specific MHC class I peptide-binding predictions using a novel representation of the MHC-binding cleft environment," *Tissue Antigens*, vol. 83, no. 2, pp. 94–100, 2014.
- [68] D. S. DeLuca, B. Khattab, and R. Blasczyk, "A modular concept of HLA for comprehensive peptide binding prediction," *Immunogenetics*, vol. 59, no. 1, pp. 25–35, 2007.
- [69] W. W. P. Liao and J. W. Arthur, "Predicting peptide binding to Major Histocompatibility Complex molecules," *Autoimmunity Reviews*, vol. 10, no. 8, pp. 469–473, 2011.
- [70] D. Rognan, S. L. Lauemøller, A. Holm, S. Buus, and V. Tschinke, "Predicting binding affinities of protein ligands from three-dimensional models: Application to peptide binding to class I major histocompatibility proteins," *Journal of Medicinal Chemistry*, vol. 42, no. 22, pp. 4650–4658, 1999.
- [71] D. Rognan, "Molecular dynamics simulation of MHC-peptide complexes as a tool for predicting potential T cell epitopes," *Biochemistry*, vol. 33, no. 38, pp. 11476–11485, 1994.
- [72] E. W. Petersdorf, M. Malkki, K. Hsu et al., "16th IHIW: international histocompatibility working group in hematopoietic cell transplantation," *International Journal of Immunogenetics*, vol. 40, no. 1, pp. 2–10, 2013.