The handle http://hdl.handle.net/1887/28890 holds various files of this Leiden University dissertation.

Author: Hombrink, Pleun
Title: Identification of minor histocompatibility antigens by reverse immunology
Issue Date: 2014-09-23
SUMMARY AND GENERAL DISCUSSION
**Summary**

T-cell recognition of MiHA plays an important role in the GVT effect of allo-SCT. Selective infusion of T-cells reactive for hematopoiesis-restricted MiHA presented in the context of HLA-class I molecules may help to separate the beneficial GVT effects from GVHD after allo-SCT. To date, only a few MiHA that form attractive targets for adoptive immunotherapy have been characterized and the number of patients that can be treated with such MiHA-selective cell therapy remains limited. In this thesis we focused on “reverse immunology” as an attractive strategy to identify clinically relevant MiHA and other T-cell epitopes. In this approach peptide predictions are the starting point and peptide candidates are subsequently screened for their capacity to induce a T-cell response. We investigated the feasibility of computational genome-wide prediction of hematopoietic MiHA and alternatively implemented mass spectrometry based HLA-peptidomics as source for candidate peptides. T-cells that reacted with these antigens were collectively isolated by MHC-tetramer pull down. Subsequently, the composition of MHC-tetramer positive T-cell populations was characterized and tested for reactivity against any of the predicted epitopes that was included in the initial MHC-tetramer panel. We generated an algorithm that could be exploited to selectively target T-cells specific for clinically relevant MiHA.

In chapter 2 we demonstrated the technical feasibility of high-throughput analysis of antigen-specific T-cell responses in small patient samples. Using three different peptide-HLA binding algorithms a set of 973 peptides that contained a SNP and were theoretically expressed by hematopoietic stem cells was produced and screened for HLA-A*0201 binding. After synthesis all peptides were tested for true peptide-HLA affinity in two separate assays resulting in a set of 333 high affinity HLA-A2 peptides that was used to generate MHC-tetramers. Subsequently, post-transplantation samples from allo-SCT patients were screened for T-cell populations reactive to any of these predicted peptides. MiHA-specific T-cell lines were generated by incubating patient PBMC samples, regardless of patient SNP status, with the collective set of 333 MHC-tetramers, followed by enrichment of MHC-tetramer positive cells on a magnetic column. Isolated T-cells were expanded and analyzed for MHC-tetramer reactivity by multi-color flow cytometry. Although a promising number of 71 peptide-reactive T-cell populations were generated by FACS sorting of MHC-tetramer positive cells, with the exception of the defined MiHA HMHA1, none of the other T-cell populations that were generated demonstrated recognition of endogenously MiHA expressing target cells, even though recognition of peptide-loaded targets was often observed. On basis of the result that the high avidity HMHA1 specific T-cell population was isolated from a transplantation setting in which both donor and recipient were homozygous negative for the MiHA phenotype we assumed that these cells were derived from the naïve T-cell repertoire. We concluded that the unbiased screening for T-cell reactivity against large sets of predicted MiHA that are solely based on epitope binding, irrespective of peptide processing data and SNP status of donor and recipient, did not result in the identification of
a novel biological relevant MiHA and we speculated that a more reliable source of naturally processed and presented peptide candidates could overcome this problem.

In chapter 3 we optimized our MHC-tetramer pull down approach to isolate antigen-specific T-cells derived from a naïve T-cell repertoire. For this purpose we took CMV as a model and aimed to generate functional CMV-specific T-cell lines from seronegative individuals within four weeks. Generated T-cell lines consisted of a variety of immunodominant CMV-epitope specific oligoclonal T-cell populations restricted to various HLA-molecules. Although all CMV-specific T-cells were isolated based on their reactivity towards a specific peptide-MHC complex, we observed a large peptide avidity variation within the isolated T-cell populations. Specifically, in addition to high-avidity T-cells, clones were also present that lacked functional reactivity. In addition, MHC-tetramer staining was not always predictive for the observed T-cell reactivity and we demonstrated that this was caused by altered MHC-peptide-TCR interaction dynamics due to the multimerized nature of MHC-tetramer complexes. Surprisingly, while the frequent isolation of naïve T-cells specific for the immunodominant A1-pp65-YSE epitope correlated with the frequent detection of a memory T-cell response in seropositive individuals, this correlation was absent for the immunodominant A2-pp65-NLV epitope, for which no high avidity T-cells could be isolated from any seronegative individuals. Based on the variable functionality of the isolated T-cell populations we put a critical note to the direct monitoring of the shape of the peripheral T-cell repertoire independent of the analysis of functional activity.

In chapter 4 we established a method for the identification of naturally processed MiHA by implementing mass spectrometry based HLA-peptidomics into a reverse immunology approach. For this purpose HLA-class I binding peptides were eluted from transformed B-cells, analyzed by mass spectrometry and matched with a database dedicated to identify polymorphic peptides. This peptide source guarantees identification of naturally processed and presented peptides. MiHA-specific T-cell lines were generated by incubating large numbers of PBMC from healthy donors with a specific set of MHC-tetramers, followed by enrichment of MHC-tetramer positive cells on a magnetic column. To increase the isolation frequency of high-affinity T-cell populations, the set of MHC-tetramers was specifically adjusted to each donor to cover only those MiHA for which the encoding SNP was screened homozygous negative in the respective donor. The identification of the LB-NISCH-1A MiHA demonstrated the feasibility of this approach. Based on these results we generated an algorithm that could be exploited to efficiently identify T-cells specific for MiHA and used this in chapter 5 to identify 2 new MiHA: LB-CLYBL-1Y and LB-TEP1-1S. Detailed analysis of the cytokine secretion profiles of the isolated T-cells demonstrated that these cells were not fully differentiated towards high IFN-γ secreting T-cells after several weeks of in vitro culture and that GM-CSF could be used as readout for T-cell reactivity to further improve the efficiency of the approach. Although the clinical relevance was only demonstrated for LB-CLYBL-1Y by the detection of MHC-multimer
positive T-cells in a patient with hematologic malignancy after allo-SCT, we underline that the identification of LB-TEP1-1S demonstrated the feasibility of our reverse immunology approach to identify MiHA that are not frequently induced \textit{in vivo}. We speculate that these so called “subdominant” MiHA can still be of interest for immunotherapeutic strategies.
**General Discussion**

Donor T-cells contribute to the success of allo-SCT and both MiHA specific CD8 and CD4 positive T-cells play a crucial role in the eradication of malignant leukemic cells. Although preclinical adoptive T-cell approaches directed against single MiHA antigens provided promising results\(^1\)\(^-\)\(^3\), the proof of principle for the effectiveness of MiHA specific immunotherapy should be provided by currently ongoing and future clinical studies. Although since the first molecular characterization of MiHA, 21 years ago\(^4\), 50 MiHA have been characterized\(^5\), only a few of them demonstrated to be expressed on hematopoietic cells, including leukemic stem cells, but not on tissues affected by GvHD. The tissue restricted expression of MiHA is crucial to enable the separation of GvL from GvHD upon adoptive transfer of T-cells specifically recognizing these MiHA to treat leukemia. Currently, there are 15 MiHA identified that are stated to be selectively expressed in hematopoietic cells\(^6\)\(^-\)\(^7\). In reality, this number is lower as T-cells reactive to several of these MiHA, recognized at least some non-hematopoietic cells. The low number of validated hematopoietic-restricted MiHA limits the broad clinically application of MiHA specific immunotherapy. Due to unbalanced population MiHA allele frequencies and HLA-restriction elements, the number of patients that could potentially be treated with MiHA based immunotherapy remains low. As a consequence the discovery of hematopoietic MiHA remains a high priority.

**Future prospects for the identification of clinically relevant MiHA by reverse immunology**

Although only 3 out of the 50 characterized MiHA were discovered by reverse immunology\(^8\)\(^-\)\(^10\), in theory this is the most efficient methodology towards identification of hematopoietic MiHA. Typical for this approach is the screening for T-cells reactive against predicted antigens. In theory this approach circumvents the need to screen isolated T-cells for unwanted reactivity against non-hematopoietic screening and allows focusing on SNP encoding MiHA with a favorable allele frequency and presented in a HLA molecule with high population coverage. Unfortunately, the historical low sensitivity and yield of every prediction step required a compensatory up-scaling of the initial numbers of candidate sequences to be screened. This resulted in a rather complex approach and several attempts failed to identify biologically relevant epitopes\(^11\)\(^-\)\(^14\). The efficacy of reverse immunology approaches is dependent on the quality of peptide predictions and the capacity to screen for peptide specific T-cell reactivity. As the prediction of MiHA critically relies on the identification of polymorphisms, new molecular developments that allow robust gene expression profiling and high resolution genome sequencing will help to separate reliable polymorphisms from sequence errors and increase the efficacy of valid MiHA candidate predictions. In addition, characterization of proteosomal cleavage sites and improvement of peptide-HLA binding prediction algorithms enables to better understand the process...
of antigen processing and presentation. Due to the emerging availability of these bioinformatics data an increasing number of MiHA candidates may be predicted. However, only when these predictions are linked to high quality gene expression profiling of virtually all organs and tissue types to select for MiHA candidates that are exclusively expressed by the hematopoietic compartment, MiHA reactive T-cell libraries can be established that may broaden the clinically application of MiHA specific immunotherapy. Ultimately this may result in “off the shelf” personalized MiHA-based therapies.

When the predicted MiHA library that was used in chapter 2 was compared with the eluted peptide libraries used in chapter 4 and 5 that yield a significant part of the B-cell HLA-ligandome, the minimal amount of coverage was striking. Length variants of fewer than 5 out of the 1000 predicted peptides were found after sequencing of eluted ligands. Although we only used genomic sequence data, gene expression profiling and peptide-HLA prediction algorithms for MiHA prediction, no significant improvement was observed when we screened a putative MiHA library that included proteosomal cleavage and TAP translocation predictions (unpublished data). These low yields indicate that either the current knowledge of antigen processing and presentation is insufficient to efficiently predict cell-surface presentation of specific peptide-MHC complexes or that the analysis of HLA-ligandomes is far from perfect. HLA-ligandome analysis by mass spectrometry may be strongly biased towards the most abundantly expressed genes (e.g. over presented housekeeping genes) and genes with a lower abundant but favorable gene expression profile may be lost in background noise. In addition, the efficiency of HLA-ligandome analysis drastically decreases when low numbers of cells, in the range of millions, are used as starting material. Ideally, large primary tumor samples should be analyzed and healthy tissues of the patient, for example, peripheral blood cells, can serve as a source of reference material at least for excluding abundant normal self-peptides from consideration15.

Although the identified HLA-eluted peptides may be a fraction of the total naturally processed peptide pool, we demonstrate in chapter 4 and 5 that this may serve as a reliable source for MiHA identification, when successfully implemented into a reverse immunology approach. As the correlation between the transcriptome and HLA peptidome has been demonstrated to be low16, it remains questionable how much efficiency gain may be expected when new methods to generate genome, gene or transcript sequence data are used to predict relevant T-cell epitopes. Although demonstrated to be technically feasible, even the state-of-the-art high resolution exome sequencing (only coding regions of the genome are sequenced) of patient hematopoietic or malignant cells alone may not be sufficient to identify large amounts of clinically relevant T-cell epitopes17-19. Alternatively, strategies that focus on ribosome bound mRNA profiling, also called polysome profiling, may boost the efficiency of T-cell epitope predictions one step further as it maps the translation of proteins and thereby underscoring measurements of mRNA expression levels only20,21. Polysome profiling revealed thousands of translational pause sites and
unannotated translation products indicating an unanticipated complexity to mammalian proteomes. In yeast, comparing the rate of translation with mRNA abundance from the same samples revealed a roughly 100-fold range of translation efficiency (as measured by the ratio of ribosome footprints to mRNA reads). In addition, translation was demonstrated for approximately 85% of genes for which mRNA abundance was measured suggesting that a 15% efficiency gain may be expected when epitope predictions are based on polysome profiling compared to total mRNA sequencing. Establishing hematopoietic cell or tumor specific polysome profiling sequencing libraries and merge them with a well established HLA-peptide affinity prediction algorithm may theoretically boost the efficiency of reverse immunology based MiHA prediction.

Although both exome and polysome sequencing methods, in contrast to HLA peptidome analysis have the benefit to be applicable on limited size malignant samples and may identify disease related mutations and genomic translocations, the unpredictability of post-translational modifications and proteosomal degradation of cellular proteins into peptides remain a bottleneck for efficient epitope prediction. As proteasomes are responsible for generation of the majority of MHC class I presented T-cell epitopes, the development of new prediction algorithms that are more accurate than those currently in use will be crucial to boost the efficacy of reverse immunology based epitope predictions. In addition, ongoing and future technical developments in the field of mass spectrometry that enable more sensitive peptide detection in limited material samples will greatly boost the efficiency in which eluted peptide libraries can be used to predict reliable T-cell epitopes. Novel developments may also enable a more efficient identification of peptides eluted from limited size samples such as specific primary cell subsets or hematopoietic malignancies.

**Rationale for the characterization of new MiHA by reverse immunology**

Based on the fact that approximately 3 million SNP were identified per individual genome, irrespective of MiHA phenotype frequencies, each allo-SCT HLA-matched donor and recipient pair is mismatched for thousands of SNP and it is likely that many MiHA have yet to be discovered. The recurrent characterization of identical MiHA specific T-cell responses in various patients after allo-SCT suggests a strong role for immunodominance. Dissecting the mechanism of immunodominance whereby the donor T-cells responds to only a few of the many possible MiHA may become very relevant for future immunotherapy as some MiHA specific T-cell responses that are frequently induced *in vivo* may induce GVL whereas others can mediate GVHD. It remains to be investigated whether this phenomenon is regulated by high to borderline presence of MiHA specific precursor T-cells or by other mechanisms such as tolerance induction to molecular mimicry epitope. The identification of LB-NISCH-1A and LB-TEP1-1S in chapter 4 and 5 illustrates that our approach allows the identification of highly immunogenic MiHA that are not frequently
induced in vivo as no MHC-tetramer positive T-cell populations were observed in the blood of patients that were positive for the MiHA and received an allo-SCT and subsequent DLI from a donor who was homozygous negative for the SNP encoding MiHA. The identification of subdominant MiHA may be restricted to reverse immunology approaches as the T-cell repertoire of patients that are screened in forward immunology approaches is skewed towards in vivo induced high immunodominant MiHA. If the identified MiHA is restricted to a favorable gene expression profile, subdominant MiHA may be of therapeutic interest as they can be exploited in potential peptide vaccination or adoptive T-cell therapies.

**Increasing the yield of MHC-tetramer pull down and FACS analysis**

In our approach the MHC-multimer based enrichment and analysis of the generated T-cell lines is crucial for the validation of MiHA and other T-cell epitopes. The absence of MHC-multimer positive T-cells however, does not necessarily indicate the non-immunogenicity of a specific peptide candidate. Precursor frequencies of antigen specific T-cells in an unchallenged donor PBMC sample vary and may be so low that the identification of these cells may be a matter of chance. In addition, potential antigen specific T-cells may be lost during the first 10 days of in vitro culture after MHC-multimer isolation. To prevent potential loss of MiHA specific T-cell populations during culture we, as an alternative, performed a single cell sort of MHC-tetramer positive T-cell population directly after MHC-tetramer pull down. Unpublished results have shown that this approach resulted in an increased number of unique MHC-tetramer positive T-cells to be screened. To allow selection of high-avidity T-cell clones in a high-throughput manner we isolated T-cells that express the surface activation marker CD137 after challenge with endogenously processed and presented antigen. This activation marker is in contrast to other activation markers upregulated by all T-cells, regardless of differentiation stage, 24 hours after in vitro antigen activation. As a consequence, recently activated in vitro primed naïve T-cells can be easily isolated based on their CD137 expression by FACS sorting.

The detection efficiency of MHC-multimer positive T-cells by combinatorial coding FACS analysis is depending on the number of fluorescent labels with sufficient intrinsic fluorescent capacity. The quantum dots that were used in chapter 2, 3, 4 and 5 allowed the use of two-color combinations. The competition for cell surface peptide-MHC complex resulted in a decreased intensity for all MHC-multimers combinations. As a consequence, the MHC-tetramer staining intensity of antigen-specific T-cell populations with a low affinity for the MHC-multimer complex may drop below threshold, especially when combinations of the lowest intensity Qdots were used. Although this may serve as an additional selection step to bias our selection strategy towards T-cell populations with a moderate to high affinity for the MHC-tetramer complex we show in chapter 3 that MHC-tetramer staining was not always predictive for T-cell avidity and therefore potential T-cell populations of interest may also be lost. The recent availability of new fluorochromes with increased fluorescent
intensities and smaller emission spectra, such as brilliant (ultra)violet dyes\textsuperscript{33}, may be used to further expand the complexity of combinatorial coding MHC-tetramer analysis by either adding or replacing to the number of fluorochromes used in two-color combinations or by designing a three- or four-color combination matrix\textsuperscript{34}. Increasing the complexity of combinatorial coding MHC-tetramer analysis will only serve purpose if T-cell recognition of large sets of potential clinically relevant peptides needs to be screened in either unbiased T-cell repertoires or limited size patient PBMC samples. In the case of dissecting patient-specific T-cell reactivity towards tumor associated neoantigens, T-cell responses were demonstrated to be biased towards a few epitopes derived from highly expressed genes, suggesting that it may be feasible to analyze patient-specific T-cell reactivity towards these type of antigens with relatively small peptide sets\textsuperscript{17}. If only a few peptide candidates needs to be screened, the complexity of combinatorial coding MHC-tetramer analysis may be downgraded towards a few optimal fluorescent labels to facilitate analysis of larger groups of patient and donor samples and thereby increasing the change of finding rare but useful T-cell clones.

**Isolation of naïve antigen-specific T-cells by MHC-tetramer pull down**

Characterization of antigen specific T-cells within the naïve repertoire can be of clinical relevance for various reasons. Detection of naïve T-cell populations may contribute to optimize rational vaccine design or to select cord blood samples with pre-existing MiHA or tumor specific T-cells prior to umbilical cord transplantation. However, due to extreme low frequencies, characterization of naïve T-cell populations requires enrichment of MHC-tetramer positive cells prior to analysis, in contrast to memory T-cell populations that expanded in vivo after antigen encounter. In chapter 2, 3, 4 and 5 we demonstrate the successful isolation of high-avidity MiHA and CMV specific T-cells from an unprimed setting. Although direct phenotypic characterization to demonstrate their naïve phenotype was lacking due to the low frequency of these cells directly after MHC-tetramer isolation, we demonstrate that these cells were not terminally differentiated towards high IFN-γ producing T-cells, unlike memory T-cells. The in vitro generated T-cell clones and lines were solely selected based on TCR-MHC-tetramer with no subsequent selection for T-cell reactivity. As the TCR diversity in the naïve repertoire has been reported to be at least 100-fold higher compared to the memory T-cell repertoire and high avidity T-cells are selectively enriched in the memory subset a substantial part of the isolated T-cells will be of low avidity\textsuperscript{35}. The composition of the isolated T-cell populations most likely reflected the broad MHC-tetramer positive T-cell repertoire before antigen driven T-cell selection.

In chapter 3 we demonstrate that no strict correlation between MHC-tetramer reactivity and T-cell avidity exist by studying various A1-pp65-YSE and A1-pp50-VTE specific T-cell clones. Although controversial, this discrepancy may be explained by the staining with multimerized MHC-peptide complexes. Multimerization of MHC-peptide
complexes alter the TCR-MHC-peptide dissociation on- and off-rate kinetics and may result in increased binding affinity of the multimerized MHC-peptide complex to surface TCR. This discrepancy may however not be observed when solely monitoring the memory T-cell repertoire which is dominated by a few antigen-experienced immunodominant high avidity T-cell clones.

Currently, the numbers of MiHA that are characterized by forward immunology exceed those that are characterized by reverse immunology, but our efforts doubled the number of MiHA identified by reverse immunology in only a small time period. Although our results did not meet the expectations that were set at the start of this project in 2007, they indicate that T-cell epitope predictions are of unexpected complexity and that new technical implementations can turn reverse immunology into a more successfully approach. In the coming years, it will be interesting to assess how new bioinformatics information can be used to improve the efficiency of reverse immunology strategies. If successful, reverse immunology based strategies can be of great value to make “off the shelf” personalized MiHA-based therapies a realistic option.
References


13. Popovič J, Li LP, Kloetzel PM et al. The only proposed T-cell epitope derived from the TEL-AML1 translocation is not naturally processed. Blood 2011;118:946-954.


Summary and General Discussion