Published in final edited form as:

J Proteomics. 2020 September 30; 228: 103921. doi:10.1016/j.jprot.2020.103921.

An input-controlled model system for identification of MHC bound peptides enabling laboratory comparisons of immunopeptidome experiments

Martin G. Klatt^a, Zita E. H. Aretz^{a,b}, Michael Curcio^a, Ron S. Gejman^{a,c}, Heather F. Jones^{a,d}, David A. Scheinberg^{a,d,*}

^aMolecular Pharmacology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, USA. 10065

^bPhysiology Biophysics and Systems Biology Program, Weill Cornell Medicine, New York, USA. 10065

^cTri-Institutional MD-PhD Program, Sloan Kettering Institute, Weill Cornell Medicine and Rockefeller University, New York, USA. 10065

^dPharmacology Department, Weill Cornell Medicine, New York, USA. 10065

Abstract

Characterization of MHC-bound peptides by mass spectrometry (MS) is an essential technique for immunologic studies. Many efforts have been made to quantify the number of MHC-presented ligands by MS and to define the limits of detection of a specific MHC ligand. However, these experiments are often complex and comparisons across different laboratories are challenging. Therefore, we compared and orthogonally validated quantitation of peptide:MHC complexes by radioimmunoassay and flow cytometry using TCR mimic antibodies in three model systems to establish a method to control the experimental input of peptide MHC:complexes for MS analysis. Following isolation of MHC-bound peptides we identified and quantified an MHC ligand of interest with high correlation to the initial input. We found that the diversity of the presented ligandome, as well as the peptide sequence itself affected the detection of the target peptide.

Furthermore, results were applicable from these model systems to unmodified cell lines with a tight correlation between HLA-A*02 complex input and the number of identified HLA-A*02 ligands. Overall, this framework provides an easily accessible experimental setup that offers the

*Corresponding author.

Author contributions

M.G.K., Z.A., M.C., H.F.J and R.S.G. performed and analyzed experiments. M.G.K. and D.A.S. designed experiments. M.G.K., Z.A. wrote the original draft of the manuscript. D.A.S. provided funding, edited the manuscript, and supervised the project. All authors reviewed and contributed to the manuscript.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of interest

D.A. Scheinberg has potential conflicts of interest, defined by *Elsevier Publishing Group* by ownership in, income from, or research funds from: Pfizer, Sella Life Sciences, Iovance, Eureka Therapeutics, and Bristol Myers Squibb. DAS, MGK and RSG are inventors of technology owned by MSK that has been filed for patent protection.

opportunity to control the peptide:MHC input and in this way compare immunopeptidome experiments not only within but also between laboratories, independent of their experimental approach.

Identification of peptides presented on major histocompatibility complex (pMHC) molecules by immunopurification and subsequent mass spectrometry (MS) is a crucial technique for understanding T cell immunity and design of immunotherapies [1–2]. Additionally, MS offers not only a qualitative, but also a quantitative approach to measure the amount of an HLA ligand of interest on the cell surface [3]. However, the isolation processes, as well as LC-MS/MS analyses of MHC-presented peptides, are complex procedures with numerous variables that can differ between laboratories. Furthermore, it has been reported that sample handling leads to losses of up to 97–99% of pMHCs during MHC ligand isolation [4], and therefore the isolation method is an important factor that can substantially bias experimental procedures. Therefore, there has been a call for “standardized protocols leading to comparable results between different laboratories” independently of their isolation method or the analytical approach used [5].

To address the problem of inter-laboratory comparability we investigated the sensitivity and robustness of input-controlled MHC ligand isolations and subsequent untargeted MS to mimic standard MHC ligand identification experiments and to demonstrate that the investigated model systems can provide the basis for comparable inter-laboratory model experiments.

Briefly, our model consisted of three major steps, see Fig.1: First, we used three cell line models expressing a pMHC of interest. Two of the cell lines are *transporter associated with peptide presentation* (TAP)-deficient which were genetically engineered with the PresentER genetic system [6] to express only a single pMHC (H2-Kb:SIINFEBL or HLAA*02:RMFPNAPYL) on the cell surface of each cell; the third cell line was EL4-OVA, which presents the SIINFEBL epitope, among other peptides normally found on the cell. The amount of pMHC presented on the cell surface was determined with two orthogonal methods, radioimmunoassay and flow cytometry, the latter being the technique we envision to be used when inter-laboratory comparisons are performed, because of its widespread availability. We were able to quantitate the number of cell surface ligands precisely because we also had available TCR mimic monoclonal antibodies that bind to pMHC epitope. The measurement of the exact numbers of pMHC on a single cell in combination with cell number and transduction efficiency (ranging between 16.6 and 94.6%) then allowed us to define the numerical input of pMHC molecules per experiment. Second, immunopurification of pMHC, elution of MHC ligands, and MS analyses were performed using a merged protocol from two different publications [7–8]. However, for inter-laboratory comparisons, the parameters of this multistep procedure can be varied, and are not of interest for the final readout of the experiment.

In the last step, acquired untargeted MS data were analyzed with whole proteome databases to mimic standard MHC ligand isolation experiments. Precursor intensities were used for relative quantitation and to determine the detection threshold for the model MHC ligands. Of note, only MS1 data were used to define the limit of detection, MS2 spectra were used to

validate that the respective precursor signal actually corresponds to the target sequence and to identify other MHC ligands present in the sample. Finally, these quantitative, input-controlled results, which include the detection limit as well as the p-values for peptide spectrum matches (PSMs), can then be used for intra-laboratory comparisons if identical processing software and parameters are utilized.

As it is crucial for our approach to ensure a reliable technique for the standardized pMHC input we characterized and compared two orthogonal assays in three model systems. Radioimmunoassay and flow cytometry with commercially available quantitation beads were utilized to determine the copy number of pMHC complexes on the cell surface of two TAP-deficient cell lines (RMA-S and T2) that were transduced with PresentER construct expressing H2-Kb:SIINFEKL or A*02:RMFPNAPYL, respectively. Additionally, EL4-OVA cells were analyzed as they also present H2-Kb:SIINFEKL on the cell surface; as this cell line is not TAP-deficient, many other MHC ligands are displayed simultaneously, as might be expected in a typical immunopeptidome experiment. For quantitation of H2-Kb:SIINFEKL, the antibody 25-D1.16 was used; for A*02:RMFPNAPYL, the TCR mimic antibody ESK1 was used [9]. All cell lines were repeatedly tested by radioimmunoassay and flow cytometry-based quantitation for the total amount of MHC class I complexes displaying the ligand of interest (H2-Kb or HLAA*02) and the copy number of the specific pMHC (Fig. 2A.) The observed much higher number of MHC molecules on T2 cells versus RMA-S cells can be explained by the bigger cell size of T2 cells compared to RMA-S cells as both cell lines are TAP-deficient. EL4 cells on the other hand express more MHC molecules on the cell surface compared to RMA-S cells as EL4 are TAP-sufficient. Moreover, for the two presentER transduced cell lines the fraction of target peptide:MHC to total MHC was much higher for SIINFEKL compared to RMFPNAPYL. This difference might occur as the presentER system cannot perfectly mimic the physiological peptide:MHC loading which might result in these differences or through different levels of TAP-deficiency in these two cell lines.

Although the radioimmunoassay systematically detected 3-fold higher copy numbers of whole MHC and specific pMHC, the two quantification methods were highly correlated ($R^2=0.994$) (Fig. 2B.) This difference in absolute numbers of cell surface molecules read by these two methods is likely because the number of fluorescent moieties on each antibody might not be exactly 1, and is not disclosed by the manufacturer, whereas the specific activity of the radioimmunoassay can be calculated to determine the exact number of CPM per antibody molecule units. Another explanation for the discrepancy between the two methods might be that for flow cytometry Fc receptor blocking reagents were used whereas for the RIA no such reagents were used which could result in more unspecific Fc receptor mediated binding.

However, as flow cytometry is more broadly available to most labs, we chose this method as the standard quantitation assay to determine the input of pMHC in our MS experiments.

Immunopurification experiments were performed for all three model systems after copy numbers of specific MHC ligands and transduction efficiency were determined by flow cytometry to define the initial amounts of pMHCs that were used in every experiment. All

experiments were carried out in technical duplicates or triplicates. For the identification of the target MHC ligands we used the Byonic software [10] searching against the complete human or mouse proteome (plus ovalbumin) and for calculating the precursor intensities skyline software [11] was utilized. Importantly, for all model systems the MHC ligand of interest could be detected with p-values for peptide spectrum matches (PSM) <0.01 and precursor intensity by peak are always correlated well with the pMHC input. For the RMFPNAPYL peptide immunoprecipitated from the PresentER T2 cells, 4.8 fmol of pMHC input were sufficient for detection whereas for the MHC ligand SIINFEKL, that was also displayed and isolated from the PresentER system, 10 amol of pMHC was the limit of detection when only detection of the precursor was considered. Of note, for samples with 100 and 10 amol H2-Kb:SIINFEKL input target sequence identification via MS2 spectra through Byonic was only feasible with p-values for PSM of <0.05 and 0.83 respectively, see Fig. 2C, 2D and Table 1. Consistent with the reports of T-cell epitopes associated with impaired peptide processing (TEIPPs) [12] we detected a small set of peptides presented by H2-Kb and H2-Db complexes in RMA-S cells and by HLA-A*02 molecules in T2 cells which showed the known characteristics of TEIPPs such as above average peptide length and suboptimal affinity to MHC complexes (Suppl. Table 1).

In contrast to RMA-S cells, for high confidence detection of the SIINFEKL peptide in the EL4-OVA setting, a much higher input of 727 fmol was needed, see Fig. 2E and Table 1. This required increase in input of over 700-fold is most likely to be the consequence of higher background in the EL4-OVA model system that continues to present cytoplasmic peptides, compared to the TAP-deficient RMA-S cells. To illustrate this possibility, we identified all MHC binders that were detected with p-values <0.01 and that were assigned to either H2-Kb, H2-Db or HLA-A*02 alleles by netMHCpan 4.0 [13] algorithm with a standard 2% rank cutoff for MHC binders. On EL4-OVA cells, over 150 times more background MHC ligands were identified compared to RMA-S PresentER SIINFEKL cells. These background MHC ligands could lower the sensitivity for detection of any particular pMHC either during the immunopurification or the LC-MS/MS analysis steps (Suppl. Table 1). As the correlation between peptide:MHC complex input measured by flow cytometry and identification of MHC ligands seemed robust, we also asked if these results could be transferred to unmodified HLA-A*02 positive cell lines used in classic immunopeptidome experiments. We quantified the HLA-A*02 surface levels from 6 different cell lines using the BB7.2 antibody by flow cytometry and subsequently performed immunoprecipitation and MS analysis on 10 million cells per cell line so that the input of HLA complexes was only affected by the different HLA-A*02 surface levels. Strikingly, there was a very strong correlation ($R^2=0.993$) between HLA-A*02 complexes per cell and the number of identified unique HLA-A*02 ligands (Fig. 2F and 2G, Suppl. Table 1). Additionally, a calculation of the average number of HLA-A*02 complexes per peptide necessary for detection in the 6 different cell lines showed that for 5 out of the 6 cell lines, molecule concentrations ranged between 1.8 and 5.1 fmol, which represents the thresholds seen in our presentER model system very well. This theoretical number is then thought to benchmark the complete immunopeptidome isolation process including all variables in one specific laboratory and will allow interlaboratory comparison.

Defining easy to implement model systems and approaches that allow comparison of immunopeptidomic experiments between laboratories is a problem that has not been addressed by the field in a satisfactory manner. Using identical cell numbers or cell pellets shared between laboratories can still be biased by shipment or numerous other factors as well as variable expression and presentation of MHC ligands. In this study, we demonstrated that using cell lines that are engineered to display an MHC ligand of choice either exclusively or in combination with other MHC ligands can form the basis to solve this problem. Flow cytometry-based quantitation of the target MHC ligand, e.g. with the H2Kb:SIINFEKL specific antibody 25-D1.16 has been shown through orthogonal validation by radioimmunoassay to be a reliable method to provide a measurement of exact input of pMHCs which allows an intra- and interlaboratory standardization of immunopeptidomic experiments. Immunopurification and analysis of the isolated MHC ligands by LC-MS/MS demonstrated a robust correlation with the pMHC input, but also illustrated the strong effect of the diversity of the immunopeptidome for the detection of a defined MHC ligand. In EL4 cells, the detection of the MHC ligand SIINFEKL required over 700 times higher input of peptide:MHC complexes than in the presentER system, in which presentation of cytoplasmic peptides is reduced. These differences might have been overcome by performing a targeted MS experiment; in order to mimic more realistic experimental procedures, we assumed in all experiments to be agnostic of the peptide sequence of interest. However, in a single peptide MHC displaying system, detection of the SIINFEKL and the RMFPNAPYL peptide was feasible with p-values for PSM <0.01 with an input of 1 and 5 fmol of pMHC, respectively. If higher p-values were accepted SIINFEKL would be identified with only 10 amol of pMHC input. Immunopeptidomic experiments are usually performed in systems with hundreds to thousands of different pMHCs displayed and therefore a single epitope model does not reflect the complexity of a biological sample. On the contrary, having background-free model systems might allow better optimization of the isolation protocol as there are fewer confounding variables. However, when we applied the idea to control the input of an immunopeptidome experiment by flow cytometry to 6 HLA-A*02 positive cell lines we observed a tight correlation between the amounts of HLA-A*02 complexes per cell and the numbers of identified A*02 HLA ligands. Furthermore, the theoretical amount of HLA complex input sufficient for HLA ligand detection was around 5 fmol in 5 out of 6 samples which overlaps well with the thresholds of detection seen in the presentER model systems. Calculation of this theoretical parameter can then be used to characterize the complete immunopeptidome isolation and identification setup from cell lysis to peptide identification in a certain laboratory and can furthermore be compared to other laboratories independent of their isolation methods or mass spectrometry instruments used. Therefore, reporting the amount of input material by flow cytometry quantification of surface MHC may help to make experiments comparable within and between laboratories.

Altogether, studies of both the presentER and EL4-OVA system provided a highly controlled model to either optimize protocols within laboratories or to compare performance of isolation techniques between different laboratories independently of all variables in the immunopurification and peptide analysis workflow. The method also provides an even more broadly applicable system for HLA-A*02 positive cell lines and could be further expanded to other common HLA alleles in the future; in the case of A*02, the input of total A*02

complexes can be easily determined and the number of identified HLA-A*02 binders thus allows average estimation of a laboratory- or method-specific threshold of detection for comparison between laboratories.

Materials and Methods

Cell lines

T2, MCF-7, MDA-MB231 and EL4-OVA cell lines were purchased from American Type Culture Collection (ATCC). T2 cells were maintained in IMDM media, MCF-7 and MDA-MB231 in DMEM media and EL4-OVA cells in RPMI media adjusted to contain 1.5 g/L sodium bicarbonate, 4.5 g/L glucose, 10 mM HEPES and 1.0 mM sodium pyruvate and supplemented with 0.05 mM 2-mercaptoethanol and 0.4 mg/ml G418. RMA-S cells were a generous gift from the Schietinger Lab at MSKCC. BV173 was provided by H. J. Stauss (University College London, London, United Kingdom), AML14 was a gift from Ross Levine, SUDHL4 was provided by Ana Younes and all the aforementioned cells were maintained in RPMI media. All media were supplemented with 10% FBS and 2mM glutamine.

Cloning the PresentER cassette and different PresentER constructs

All cloning was performed as described previously [6]. The PresentER cassette as well as the SIINFEKL and RMFPANPYL minigenes can be found at Addgene (Cat. #102942 and #102944)

Production of retrovirus and transduction

HEK293T Phoenix amphoteric cells were transfected with polyethylenimine (PEI) and PresentER plasmid (15µg DNA: 45µg PEI) in 10cm TC plates. Virus was harvested every 12 hours, pooled, concentrated with Clontech Retro-X and frozen. T2 cells were spinoculated with virus in non-TC treated 6-well plates at 32°C × 2,000xg for 2 hours with 4µg/ml of polybrene. RMA-S cells were spinoculated with virus in non-TC treated 6-well plates at 32°C × 1,000xg for 2 hours with 4µg/ml of polybrene. The volume of viral supernatant that led to 1/3 maximal transduction efficiency was established for each batch of virus produced by transduction with serial dilutions of viral supernatant. Transduced cells were selected with 1µg/ml (T2) or 4µg/ml (RMA-S) of puromycin for 2–3 days and transduction efficiency determined by GFP levels.

Antibodies and labeling reagents

ESK1 monoclonal antibody specific for HLA-A*02:RMFPNAPYL was purified by Eureka Therapeutics and used for either radioimmunoassay or flow cytometry. H2-Kb:SIINFEKL specific antibody 25-D1.16 and H2-Kb specific Y3 were purchased from BioXcell (Cat. #BE0207 and #BE0172) and used for either radioimmunoassay or flow cytometry. The BB7.2 clone was purified at the MSKCC antibody core facility. For flow cytometry ESK1 and 25-D1.16 were labeled with Innova Biosciences Lightning Link kits (Cat. #703–0010) according to manufacturer's protocols to results in an equimolar ratio of antibody and fluorophore. For H2-Kb and HLA-A*02 quantitation Y3-PE and BB7.2-PE antibodies were purchased from BioLegend (Cat. #141603 and #343305).

Quantitation of surface molecules by flow cytometry

For flow cytometry-based quantitation experiments BD Quantibrite-PE kits were used (BDBiosciences, Cat. #340495) according to the manufacturers protocol. Flow cytometry was performed with a standard protocol and a dilution of the above mentioned antibodies at 1:100 for 30 min on ice. Flow cytometry data were collected on a Fortessa (BD Biosciences) and analyzed with FlowJo V9.9.

Radioimmunoassay

ESK1, 25-D1.16, BB7.2 and Y-3 was labeled with ^{125}I (PerkinElmer) using the chloramine-T method. Antibody (100 μg) was reacted with 1 mCi ^{125}I and 20 μg chloramine-T, quenched with 200 μg Na-metabisulfite, then separated from free ^{125}I using a 10DG column and equilibrated with 2% bovine serum albumin in PBS. Specific activities of products were in the range of 4–8 mCi/mg. PresentER or EL4-OVA cells (only cells with transduction efficiencies >95% were used) and respective wildtype variants (T2, RMA-S and EL4) were harvested. Cells were washed once with PBS and re-suspended in 2% human serum in PBS at 10^7 cells/ml on ice. Cells (10^6 per tube, in duplicate) were incubated with ^{125}I -labeled antibody (1 $\mu\text{g}/\text{ml}$) for 45 minutes on ice, then washed extensively with 1% bovine serum albumin in PBS on ice. Bound radioactivity was measured by a gamma counter, specific binding was determined, and the number of bound antibodies per cell was calculated from specific activity. For final copy numbers unspecific background binding of T2, RMA-S or EL4 cells was subtracted.

Immunopurification of HLA class I ligands

For immunopurification affinity columns were prepared as follows: 40 mg of Cyanogen bromide-activated-Sepharose® 4B (Sigma-Aldrich, Cat# C9142) was activated with 1mM hydrochloric acid (Sigma-Aldrich, Cat# 320331) for 30 min. Subsequently, 0.5 mg of W6/32 antibody (BioXCell, Cat #BE0079) and for mouse cell lines 0.5 mg M1/42.3.9.8 antibody (BioXCell, Cat #BE0077) was coupled to sepharose in presence of binding buffer (150mM sodium chloride, 50 mM sodium bicarbonate, pH 8.3; sodium chloride: Sigma-Aldrich, Cat# S9888, sodium bicarbonate: Sigma-Aldrich, Cat#S6014) for at least 2 hours at room temperature. Sepharose was blocked for 1 h with glycine (Sigma-Aldrich, Cat# 410225). Columns were equilibrated with PBS for 10 min. Suspension cells were harvested directly, adherent cell lines after incubating 15 min with CellStripper solution (Corning™, Cat# 25056CI) and washed three times in ice-cold sterile PBS (Media preparation facility MSKCC). For presentER and EL4-OVA cells number of cells used is indicated in Table 1, for all other experiments 10^7 cells were used. Afterwards, cells were lysed in 7.5 ml 1% CHAPS (Sigma-Aldrich, Cat# C3023) in PBS, supplemented with protease inhibitors (cOmplete, Cat# 11836145001) for 1 hour at 4°C. Lysate was spun down for 1 hour with 20,000 g at 4°C. Supernatant was run over column through peristaltic pumps with 1 ml/min flow rate overnight in cold room. Affinity columns were washed with PBS for 30 min, water for 30 min, then run dry, and HLA complexes subsequently eluted five times with 200 μl 1% trifluoroacetic acid (TFA, Sigma/Aldrich, Cat# 02031). For separation of HLA ligands from their HLA complexes tC18 columns (Sep-Pak tC18 1 cc Vac Cartridge, 50 mg Sorbent per Cartridge, 37–55 μm Particle Size, Waters, Cat# WAT036820) were prewashed with 80%

acetonitrile (ACN, Sigma-Aldrich, Cat# 34998) in 0.1% TFA and equilibrated with two washes of 0.1% TFA. Samples were loaded, washed again with 0.1% TFA and eluted in 400 μ l 30% ACN in 0.1% TFA. Sample volume was reduced by vacuum centrifugation for mass spectrometry analysis.

LC-MS/MS analysis of HLA ligands

Samples were analyzed by high resolution/high accuracy LC-MS/MS (Lumos Fusion, Thermo Fisher). Peptides were desalted and concentrated prior to being separated using direct loading onto a packed- in-emitter C18 column (75 μ m ID/12cm, 3 μ m particles, Nikkyo Technos Co., Ltd. Japan). The gradient was delivered at 300nl/min increasing linear from 2% Buffer B (0.1% formic acid in 80% acetonitrile) / 98% Buffer A (0.1% formic acid) to 30% Buffer B / 70% Buffer A, over 70 minutes. MS and MS/MS were operated at resolutions of 60,000 and 30,000, respectively. Only charge states 1, 2 and 3 were allowed. 1.6 Th was chosen as isolation window and collision energy was set at 30%. For MS/MS, maximum injection time was 100ms with an AGC of 50,000.

Mass spectrometry data processing

Mass spectrometry data was processed using Byonic software (version 2.7.84, Protein Metrics, Palo Alto, CA) through a custom-built computer server equipped with 4 Intel Xeon E5-4620 8-core CPUs operating at 2.2 GHz, and 512 GB physical memory (Exxact Corporation, Fremont, CA). Mass accuracy for MS1 was set to 6 ppm and to 20 ppm for MS2, respectively. Digestion specificity was defined as unspecific and only precursors with charges 1,2, and 3 and up to 2 kDa were allowed. Protein FDR was disabled to allow complete assessment of potential peptide identifications. Oxidization of methionine was set as variable modifications for all samples. T2 samples were searched against UniProt Human Reviewed Database (20,349 entries, <http://www.uniprot.org>, downloaded June 2017) and RMA-S cells as well as EL4-OVA cells were searched against UniProt Mouse Reviewed Database (17,033 entries, <http://www.uniprot.org>, downloaded August 2019) supplemented with the ovalbumin protein sequence. All databases included target decoys derived from the reversed sequences of all included proteins. Peptides were selected with a minimal log prob value of 2 corresponding to p-values from PSMs<0.01. We added false positive hits that were identified after selection for peptides of 8–12 amino acids in length and assigned to HLA-A*02 complexes via netMHCpan (Suppl. Table 1). For AML14 false positive peptide identifications without HLA assignment were also included in the Suppl. Table 1). The calculated FDR based on these false and true positive identifications was <1% for all analyzed samples reassuring our selection for the logprob filter value of 2. For relative quantitation skyline software (version 3.1, MacCoss Lab Software) was used. Masses of precursors of peptide target sequences were searched in all relevant .raw files and peak areas of all replicates compared. We enabled matching mass and retention times of target precursors to other samples to ensure highest sensitivity and detection in all relevant samples independently of selection for MS2 fragmentation. For quantitation only time windows 30 seconds before and after the identification of the correct MHC ligand sequence were considered.

Assignment of peptide sequences to HLA alleles

To assign peptides which passed the above-mentioned MS quality filters to their HLA complexes which they most likely bind to we used the netMHCpan 4.0 algorithm [13] with default settings. No binding affinity predictions were enabled. Therefore, all peptides with affinity %ranks below 2 were considered binder.

Software

All graphs were drawn with Graphpad Prism 7.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Alex Kentsis for access to the Byonic Software and the Proteome Resource Center at Rockefeller University for the performance of all LC-MS/MS experiments.

Funding

This study was supported by Leukemia and Lymphoma Society, NIH P30CA 008748, NIH R01 CA55349, P01 CA23766, Solomon Foundation, and Tudor Funds. M.G.K. is supported by the German Research Foundation with the individual research grant KL 3118/1-1. RSG is supported by NCI F30 CA200327 and NIGMS T32GM07739. HFJ is supported by NIH T32 CA062948. The content of this study is solely the responsibility of the authors and does not necessarily represent official views of the German Research Foundation.

References

- [1]. Bassani-Sternberg M, Coukos G, Mass spectrometry-based antigen discovery for cancer immunotherapy, *Curr Opin Immunol* 41 (2016) 9–17. [PubMed: 27155075]
- [2]. Hillen N, Stevanovic S, Contribution of mass spectrometry-based proteomics to immunology, *Expert Rev Proteomics* 3(6) (2006) 653–64. [PubMed: 17181480]
- [3]. Tan CT, Croft NP, Dudek NL, Williamson NA, Purcell AW, Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring, *Proteomics* 11(11) (2011) 2336–40. [PubMed: 21598389]
- [4]. Hassan C, Kester MG, Oudgenoeg G, de Ru AH, Janssen GM, Drijfhout JW, Spaapen RM, Jimenez CR, Heemskerk MH, Falkenburg JH, van Veelen PA, Accurate quantitation of MHC-bound peptides by application of isotopically labeled peptide MHC complexes, *J Proteomics* 109 (2014) 240–4. [PubMed: 25050860]
- [5]. Caron E, Kowalewski DJ, Chiek Koh C, Sturm T, Schuster H, Aebersold R, Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes Using Mass Spectrometry, *Mol Cell Proteomics* 14(12) (2015) 3105–17. [PubMed: 26628741]
- [6]. Gejman RS, Jones HF, Klatt MG, Chang AY, Oh CY, Chandran SS, Korontsvit T, Zakahleva V, Dao T, Klebanoff CA, Scheinberg DA, Identification of the targets of T cell receptor therapeutic agents and cells by use of a high throughput genetic platform, *Cancer Immunol Res* (2020).
- [7]. Nelde A, Kowalewski DJ, Stevanovic S, Purification and Identification of Naturally Presented MHC Class I and II Ligands, *Methods Mol Biol* 1988 (2019) 123–136. [PubMed: 31147937]
- [8]. Bassani-Sternberg M, Braunlein E, Klar R, Engleitner T, Sinitcyn P, Audehm S, Straub M, Weber J, Slotta-Huspenina J, Specht K, Martignoni ME, Werner A, Hein R, D HB, Peschel C, Rad R, Cox J, Mann M, Krackhardt AM, Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry, *Nat Commun* 7 (2016) 13404. [PubMed: 27869121]

- [9]. Dao T, Yan S, Veomett N, Pankov D, Zhou L, Korontsvit T, Scott A, Whitten J, Maslak P, Casey E, Tan T, Liu H, Zakhaleva V, Curcio M, Doubrovina E, O'Reilly RJ, Liu C, Scheinberg DA, Targeting the intracellular WT1 oncogene product with a therapeutic human antibody, *Sci Transl Med* 5(176) (2013) 176ra33.
- [10]. Bern M, Kil YJ, Becker C, Byonic: advanced peptide and protein identification software, *Curr Protoc Bioinformatics* Chapter 13 (2012) Unit13 20.
- [11]. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, *Bioinformatics* 26(7) (2010) 966–8. [PubMed: 20147306]
- [12]. Seidel UJ, Oliveira CC, Lampen MH, Hall T, A novel category of antigens enabling CTL immunity to tumor escape variants: Cinderella antigens, *Cancer Immunol Immunother* 61(1) (2012) 119–25. [PubMed: 22116347]
- [13]. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M, NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data, *J Immunol* 199(9) (2017) 3360–3368. [PubMed: 28978689]

Highlights

- Comparisons of peptide:MHC surface quantitation via flow cytometry and radioimmunoassay shows MHC ligand isolation correlates with quantified input of unlabeled peptide:MHC complexes.
- Unbiased detection of target MHC ligands in three model cell lines with a minimal input of 10 attomole of peptide:MHC complex.
- The diversity of presented peptide:MHC complexes affects detection of target MHC ligands by MS.
- Results are applicable to human A*02 cell line experiments with similar limits of detection
- Flow cytometry-based quantitation of MHC complexes provides a reliable and easily applicable approach to compare MHC ligand isolation techniques within and among labs.

Significance

Although immunopectidomics is an essential tool for the characterization of MHC bound peptides on the cell surface, there are no easily applicable established protocols available that allow comparison of immunopectidome experiments across laboratories. Here, we demonstrate that controlling the peptide:MHC input for immunopurification and LC/MS-MS experiments by flow cytometry in pre-defined model systems allows the generation of qualitative and quantitative data that can easily be compared between investigators, independently of their methods for MHC ligand isolation for MS.

I. Standardization of specific peptide:MHC input (for inter- and intra-laboratory comparison) by flow cytometry and quantitation beads

II. Experimental parameters for purification and identification of MHC ligands (variable for intra-laboratory, fixed for interlaboratory comparison)

III. Pre-defined criteria for qualitative or quantitative identification of the specific peptide:MHC ligand

Fig. 1 –. Experimental workflow for input-controlled analysis of immuno-peptidome experiments.

(I) Flow cytometry-based quantitation is performed to determine the number of molecules of a specific pMHC per cell. In conjunction with transduction efficiency and total cell number it allows to define the pMHC input. (II) Exemplifying lists of different variables of MHC ligand isolation as well as LC-MS/MS analysis, which can or can not be altered in the experimental setting. (III) Analysis of MS data with different algorithm for qualitative and quantitative detection of the MHC ligand of interest. These results can be used for experimental comparisons.

Fig. 2 – Comparison and orthogonal validation of flow cytometry-based quantitation of pMHC and radioimmunoassay.

(A) Copy numbers of specific pMHC and total MHC determined by radioimmunoassay and flow cytometry-based quantitation. (B) Visualization and linear regression of the results in (A). All data points are the mean of technical triplicates and all quantitations have been performed at least twice with similar results.

Relative quantitation of input-controlled immunopeptidomics experiments. (C-E)

Relative quantitation by precursor intensity measured by peak area with skyline software for T2-PresentER:RMFPNAPYL (C), RMA-S-PresentER:SIINFEKL (D) and EL4-OVA cells (E). Quantitation was calculated by skyline software and time window for quantitation was 30 sec before and after the detection of the target sequence in MS2 spectra.

Correlation of HLA-A*02 complex input and number of identified HLA-A*02 binding peptides in unmodified cancer cell lines. (F). Sample and input characteristics (G)

Correlation of HLA-A*02 complex input and identified A*02 assigned HLA ligands. Error bars in 2B and 2G as well as deviations in 2A and 2F indicate SD.

Table 1.

Overview of experimental input by cell number and molecules, p-values for MHC ligand identification and number of other identified MHC ligands

RMA-S PresentER SIINFEKL				EL4-OVA				T2 PresentER RMFPNAPYL			
#of cells [cells]	input [pMHC]	p-value for PSM [target]	diversity [MHC ligands]	#of cells [cells]	input [pMHC]	p-value for PSM [target]	diversity [MHC ligands]	#of cells [cells]	input [pMHC]	p-value for PSM [target]	diversity [MHC ligands]
180 M	61 fmol	<0.01	11	80 M	727 fmol	<0.01	318	85 M	29 fmol	<0.01	9
60 M	20 fmol	<0.01	2	40 M	363 fmol	n.d.	255	28 M	9.5 fmol	<0.01	6
30 M	10 fmol	<0.01	6	10 M	90.9 fmol	n.d.	117	14 M	4.8 fmol	<0.01	3
15 M	5.1 fmol	<0.01	6	3.3 M	30.0 fmol	n.d.	32	7 M	2.4 fmol	n.d.	2
7.5 M	2.5 fmol	<0.01	3	1.7 M	15.4 fmol	n.d.	22	3.5 M	1.2 fmol	n.d.	0
3 M	1.0 fmol	<0.01	2								
0.3 M	0.10 fmol	<0.05	2								
0.03 M	10 amol	0.83	2								

Transduction efficiencies for MS experiments: RMA-S:SIINFEKL=16.6%, EL4-OVA=93.7%, T2:RMFPNAPYL=94.6%. n.d.= not detected.