Journal of proteome research

pubs.acs.org/jpr

Article

Mild Acid Elution and MHC Immunoaffinity Chromatography Reveal Similar Albeit Not Identical Profiles of the HLA Class I Immunopeptidome

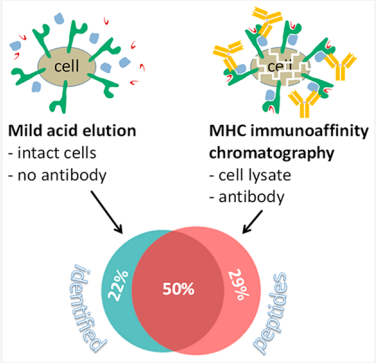
Theo Sturm,* Benedikt Sautter, Tobias P. Wörner, Stefan Stevanović, Hans-Georg Rammensee, Oliver Planz, Albert J. R. Heck, and Ruedi Aebersold*

Cite This: *J. Proteome Res.* 2021, 20, 289−304

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: To understand and treat immunology-related diseases, a comprehensive, unbiased characterization of major histocompatibility complex (MHC) peptide ligands is of key importance. Preceding the analysis by mass spectrometry, MHC class I peptide ligands are typically isolated by MHC immunoaffinity chromatography (MHC-IAC) and less often by mild acid elution (MAE). MAE may provide a cheap alternative to MHC-IAC for suspension cells but has been hampered by the high number of contaminating, MHC-unrelated peptides. Here, we optimized MAE, yielding MHC peptide ligand purities of more than 80%. When compared with MHC-IAC, obtained peptides were similar in numbers, identities, and to a large extent intensities, while the percentage of cysteinylated peptides was 5 times higher in MAE. The latter benefitted the discovery of MHC-allotype-specific, distinct cysteinylation frequencies at individual positions of MHC peptide ligands. MAE revealed many MHC ligands with unmodified, N-terminal cysteine residues which get lost in MHC-IAC workflows. The results support the idea that MAE might be particularly valuable for the high-confidence analysis of post-translational modifications by avoiding the exposure of the investigated peptides to enzymes and reactive molecules in the cell lysate. Our improved and carefully documented MAE workflow represents a high-quality, cost-effective alternative to MHC-IAC for suspension cells.



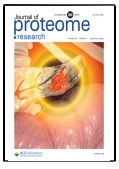
KEYWORDS: *MHC bound peptides, HLA ligandome, MHC immunoaffinity chromatography, mild acid elution, cysteinylation, post-translational modification*

INTRODUCTION

The identification and prediction of MHC class I (MHC-I) peptide ligands have gained increasing attention in recent years due to their central role in numerous diseases, including viral infections and cancer. Several novel tumor immunotherapies critically rely on MHC-I peptide ligands[1−3] because these peptides mediate the specificity of cytotoxic CD8 T cell responses against altered cells. Most MHC-I peptide ligands arise as proteasomal degradation products of endogenous proteins and are loaded on the MHC-I proteins in the endoplasmic reticulum (ER). This peptide loading depends on the presence of an MHC-allotype-specific MHC-I binding motif; i.e., the peptide needs to have a certain length (usually 8−11 amino acids) and to contain certain amino acid residues ("anchors") at usually two defined positions.[4] Peptide amino acids apart from the anchor residues also influence the affinity for a particular MHC allotype.[5] Following assembly in the ER, MHC-I peptide complexes are transported via the Golgi apparatus to the cell surface to function as cellular signatures for immunosurveillance. From the plasma membrane, they are finally targeted to endosomes for following degradation in lysosomes or for recycling to the cell surface.[6] The sum of all MHC peptide ligands from a given biological entity is referred to as the MHC peptide ligandome or the MHC immunopeptidome. MHC alleles and the corresponding MHC (protein) allotypes of humans are specifically addressed as human leukocyte antigens (HLAs).

Unintentionally, many MHC immunopeptidome analyses are biased due to technical limitations.[7,8] For example, extremely hydrophilic and extremely hydrophobic peptides are often lost during sample processing, and it has been speculated that the W6/32 antibody used in MHC-IAC might introduce a bias toward certain MHC allotypes.[9] Although it is not always practical or even possible to overcome these and other biases, it is important to be aware of them to avoid mis- or over-interpretation of data. One approach to estimate the bias introduced by a specific method is to investigate exactly the same

Received: June 1, 2020
Published: November 3, 2020



sample with at least two (ideally strongly) differing methods. MHC-IAC and MAE are the two most widely used methods for purification of cellular MHC peptide ligands for liquid chromatography mass spectrometry (LC-MS) analyses, being also quite orthogonal to each other. Therefore, a direct comparison of the peptides isolated by either method should help in estimating the bias introduced by each approach.

MHC-IAC employs bead-coupled antibodies targeting MHC protein complexes to enrich the latter from detergent-solubilized cells. MHC-IAC columns can be eluted at very low pH, resulting in the release of peptide ligands from both MHC-I and MHC class II (MHC-II) proteins.[4,10] In contrast, MAE is independent of antibodies and elutes peptides from MHC-I, but putatively not from MHC-II, at pH 3 or pH 3.3 directly from intact cells without lysing them.[10−13] Therefore, whereas MAE samples MHC peptide ligands only from MHC-I molecules at the cell surface, MHC-IAC in its standard format also extracts peptides associated with intracellular MHC molecules residing, for example, in endosomes solubilized during cell lysis.

Despite its independence of antibodies and the resulting low cost, only a few laboratories have pioneered MAE for LC-MS-based analyses.[14−18] This is mainly due to three limitations. First, the current MAE protocol is not applicable to solid tissue; second, it appears unsuitable for frozen cells;[18] and third, MAE has suffered from a very high amount of peptide and non-peptide contaminants.[10] About 60% of all peptides identified with MAE were reported to not be MHC-associated.[16] Only very recently, Lanoix et al. published an MAE data set where the fraction of putatively non-MHC-associated peptides is almost as low as that in MHC-IAC.[18] The work by Lanoix et al. clearly represents an important step in making MAE a competitive method for the more laborious and expensive MHC-IAC, but the authors did not publish a detailed description of their MAE workflow nor do they specify potential reasons for the dramatic decrease of putative contaminants as compared to previous data sets from their and other groups.

Due to the higher peptide purities obtained with MHC-IAC and its applicability to frozen cells and solid tissues, most laboratories use this method for LC-MS-based characterization of the MHC peptide ligandome and the establishment of MHC peptide binding motifs.[19−29] Therefore, MHC-IAC-derived data also prevails in public immunopeptidomic databases. One of the oldest prediction algorithms for MHC binding, Syfpeithi, is fully based on MHC-IAC-derived LC-MS data.[30] Although alternative prediction algorithms relying on measurements of peptides' MHC binding affinity have also been very successful,[31−33] the former approach has recently experienced a renaissance by fully[27,34] or partly[35,36] training the predictions on immunopeptidomic LC-MS data. Also, in this context, the applied LC-MS data has been exclusively or largely been MHC-IAC-derived. The presence of predicted MHC peptide ligands in MHC-IAC-derived data has even been used to comparatively evaluate the performance of prediction algorithms for MHC presentation.[27] However, the yield of MHC-IAC has been reported to be only about 0.5−3%.[37] If this was true, there would be a considerable risk of strong sampling biases in current data sets as well as the dependent MHC binding prediction algorithms because it is not sure that losses of ≥97% affect all MHC peptide ligands equally. Is the high fraction of putatively non-MHC binding peptides in current MAE data sets partly a result of MHC-IAC-biased determination of MHC peptide binding motifs? In 2006, a comparison of MHC-IAC with MAE identified only 85 peptides of which merely 3 were shared between both extraction methods.[14] These low numbers of identifications do not allow a solid comparison. Therefore, we set out here to perform a very rigorous comparison of advanced MHC-IAC and MAE extraction methods to estimate potential biases of both protocols comparing our data also with the most recent literature.[18] Our optimized MHC-IAC and MAE extraction methods perform almost on par considering the numbers and MS1 intensities of identified MHC peptide ligands, providing solid evidence that MAE may represent a competitive alternative. Although very good qualitative and quantitative congruency is observed between the two extraction methods, we also discuss observed biases, such as in the detection of post-translationally modified MHC peptide ligands.

METHODS

Cells

The Epstein−Barr virus (EBV) transformed lymphoblastoid B cell line JY was derived from the European Collection of Authenticated Cell Cultures (no. ECACC 94022533) and expresses HLA-A*02:01, HLA-B*07:02, and HLA-C*07:02. We purchased the human monocytic leukemia cell line THP-1[38] from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (no. DSMZ ACC 16). LCL5 is a B lymphoblastoid cell line generated in the Rammensee laboratory by transfection of human peripheral blood mononuclear cells with the B95-8 strain of EBV.[39] Both THP-1 and LCL5 were subjected to next generation sequencing of selected MHC regions at the University Hospital Tübingen by Claudia A. Müller. Thereby, THP-1 was found to contain the HLA class I alleles A*02:01, A*24:02, B*15:11, B*35:01 (or B*35:42), and C*03:03 (and/or C*03:12/13/43/55/96/103), and LCL5 was found to encode for the HLA class I alleles A*24:02, A*29:02, B*44:02, B*44:03, C*05:01 (or C*05:07N), and C*16:01. For further information about the cell culturing, see the Supporting Information.

MAE

For MAEs from JY cells, we always used 5×10^8 cells per sample. MAEs from THP-1 cells were performed with 1.25×10^8, 5×10^8, or 1.5×10^9 cells per sample, as specified in Supplementary Table S9. All steps from harvesting of the cells until ultracentrifugation were performed at 4 °C or on ice. Cells were harvested and washed three times with PBS at 211*g* using 10 min centrifugations. They were then gently pipetted up and down for 1 min in MAE buffer. We used 1, 4, or 12 mL of MAE buffer for 1.25×10^8, 5×10^8, or 1.5×10^9 cells per MAE, respectively. The MAE buffer consisted of 131 mM citric acid, 66 mM Na_2HPO_4, 150 mM NaCl, 1 μM aprotinin, and 25 mM iodoacetamide adjusted to pH 3.3 with NaOH. For some experiments pinpointed in Supplementary Table S9, the MAE buffer was supplemented with further additives that could be beneficial in theory but were found in our pretests to result in no considerable effect or a slight benefit at best. These additional additives included the protease inhibitors leupeptin at 1 mg/L, pepstatin at 0.7 mg/L, and EDTA at 2 mM, 10% of the organic solvent dimethyl sulfoxide (DMSO) as well as the glycyl-methionine (GM) dipeptide at 10 mM.[40] All tubes used during or after the treatment of cells with MAE buffer were either certified to be free of plasticizers (Eppendorf), or they were extensively washed with acid and water before usage. Following the 1 min treatment with MAE buffer, cells were immediately removed by centrifugation at 285−300*g* for 5 min. The supernatant was then centrifuged at 339−350*g* for 10 min,

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

further cleared at 3345−3406*g* for 15 min, subjected to ultracentrifugation at 257,000*g* for 1 h, and frozen at −80 °C.

The obtained peptide solution was further purified on Oasis HLB columns (barrel size 1 cm^3, 30 mg of sorbent; Waters, product no.: WAT094225) prewashed with 80% acetonitrile (CH_3CN)/0.2% TFA and 100% CH_3CN. After equilibration with 2% CH_3CN/0.1% TFA, sample loading, and washing with 2% CH_3CN/0.1% TFA, peptides were eluted with 35 or 60% CH_3CN/0.1% TFA (see Supplementary Table S9). The eluate was filtered through a prewashed 4 mL Amicon ultrafilter device with 3 kDa molecular weight cutoff (Merck Millipore, Cat.-No. UFC800324) at 4 °C. As specified in Supplementary Table S9, the ultrafilter was rinsed with a CH_3CN-rich solution in the more recent experiments recovering sticky peptides (compare parts C−E of Supplementary Figure S18). After vacuum centrifugation of the ultrafiltrates, peptides were finally cleaned using C_{18}-ZipTips (Merck Millipore) prewashed with 80% CH_3CN/0.1% TFA. We equilibrated and washed the ZipTips with 0.1% TFA and eluted using 35, 60, or 80% CH_3CN in 0.1% TFA (see Supplementary Table S9) followed by a second vacuum centrifugation. All MAEs were performed in the Rammensee laboratory to minimize interlaboratory variation. Our MAE workflow is schematically depicted in Supplementary Figure S1.

MHC-IAC

We used the W6/32 monoclonal antibody[41] to purify HLA-A, -B, and -C peptide complexes. MHC-II peptide complexes were extracted harnessing a 1:1 mixture of L243 (ref 42, anti HLA-DR) and Tü39 (ref 43, anti HLA-DR, -DQ, and -DP) monoclonal antibodies. A 40 mg portion of CNBr-activated Sepharose 4B (GE Healthcare) per 1 mg of antibody was washed with 1 mM HCl for 30 min. The antibodies were then coupled to the sepharose in 0.5 M NaCl/0.1 M $NaHCO_3$ at pH 8.3 for 2 h at room temperature. Remaining reactive groups were blocked with 0.2 M glycine for 1 h at room temperature followed by two washes with PBS.

The amount of cells used per MHC-IAC sample is indicated in the figures showing the respective data and additionally compiled in Supplementary Table S10. Cells were washed three times in PBS before freezing. The frozen cell pellet was thawed in an equal volume of 2-fold concentrated lysis buffer consisting of cold PBS, cOmplete Protease Inhibitor Cocktail from Roche, and 1.2% CHAPS. All following steps before the vacuum concentration were performed at 4 °C or on ice. After slow rotation for 1 h, the cell lysate was sonicated three times for 20 s followed by another 1 h of slow rotation. The supernatant from a centrifugation for 99 min (LCL5) or 45 min (JY, THP-1) at approximately 18,000*g* was passed through a 0.2 μm (LCL5) or a 5 μm (JY, THP-1) filter and loaded onto the MHC-IAC columns prewashed with lysis buffer. For each sample, we used two connected Econo-Columns from Bio-Rad, one filled with the W6/32-sepharose and the other filled with L243/Tü39-sepharose employing 1 mg of antibody per 1 mL of cell pellet per column. The samples were cyclically pumped through the columns overnight. We then linearly pumped PBS for 30 min and water for 1 h for washing followed by shortly pumping air for drying. All pumping was performed at a flow rate of 1 to 2 mL/min. The sepharose matrix was acidified with one drop of 10% TFA per 1 mg of antibody, fully covered with 0.2% TFA, shaken at 300 rpm for about 15 min, and eluted by air pressure. The incubation and elution with 0.2% TFA was repeated seven times. To test the effect of organic solvents during elution, dedicated samples were eluted differently, as specified in Supplementary Figure S16 and Supplementary Table S10.

The eluate was filtered through prewashed 0.5 or 4 mL Amicon ultrafilter devices with a 3 or 10 kDa molecular weight cutoff, as specified in Supplementary Table S10. The ultrafilter was rinsed with 80% CH_3CN/0.2% TFA (LCL5) or 50% CH_3CN/0.2% TFA (JY, THP-1), and the rinse ultrafiltrate was combined with the first ultrafiltrate except for the two dedicated samples of LCL5 cells (compare Supplementary Table S10). The combined ultrafiltrate of LCL5 cells was subjected to lyophilization, whereas all other samples were concentrated by vacuum centrifugation. Final cleaning of peptides was performed with C_{18}-ZipTips, as described for the MAE samples eluting peptides with 80% CH_3CN in 0.1% TFA followed by a second vacuum centrifugation. All MHC-IACs were performed in the same laboratory in Tübingen to minimize technical variability. Our MHC-IAC workflow is schematically depicted in Supplementary Figure S1.

LC-MS Measurements

The volume of the concentrated samples was adjusted for LC-MS injections of 5 μL per run using 1% CH_3CN/0.05% TFA. Doing so, we ensured that each LC-MS run from JY cells corresponded to about 5×10^7 cells. The fraction of the original sample injected per LC-MS run is specified in Supplementary Table S9 and Supplementary Table S10.

For this study, we applied two different LC-MS systems—one employing an LTQ Orbitrap XL and the other an Orbitrap Fusion (both from Thermo Scientific). The measurements at the Orbitrap Fusion were used here merely for a better characterization of the frequencies of cysteine modifications (see Supplementary Table S9 for the assignment of MS raw files to the mass spectrometer). All other measurements including all MHC-IAC-derived samples were performed on the LTQ Orbitrap XL, and therefore, this setup is described first. The LTQ Orbitrap XL was coupled online to a nano ultra high performance LC (UltiMate 3000 RSLCnano, Dionex) via a nanoelectrospray ion source (Nanospray II, Thermo Scientific). Samples were loaded onto a 2 cm × 75 μm C_{18} trapping column (Acclaim PepMap100, Dionex) at 4 μL/min for 5.75 min. Peptides were then separated on a 25 cm × 50 μm column with C_{18} beads of 2 μm diameter (Acclaim PepMap RSLC, Dionex) at 50 °C using a flow rate of 175 nL/min and a 90 min gradient ranging from 2.4 to 32% CH_3CN in 0.1% HCOOH. One LC-MS run (see Supplementary Table S10) at the LTQ Orbitrap XL was exceptionally not performed with the 25 cm C_{18} column but with a 50 cm C_{18} column that was otherwise identical (Acclaim PepMap RSLC, Dionex). For the 50 cm column, the gradient duration was 140 min, while the other settings were the same as those for the 25 cm column. MS1 scans were acquired at a resolution of 60,000 at 400 Th, and the "lock mass" option[44] was enabled using the 445.120025 ion for real time internal calibrations. For all samples derived from MAE or MHC-IAC with W6/32 antibody, the measured MS1 m/z range was 310−650 Th (LCL5 cells) or 400−650 Th (JY and THP-1 cells), and only 2+ and 3+ charged precursors were allowed for fragmentation. For MHC-IAC with L243/Tü39 antibody, MS1 spectra were acquired between 300 and 1500 Th, and all precursors with ≥2 positive charges were eligible for fragmentation. The five most abundant precursors from every MS1 scan were selected for collision-induced dissociation (CID) using an isolation width of 2 Th, and they were then dynamically excluded from repeated fragmentation for 3 s. We

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

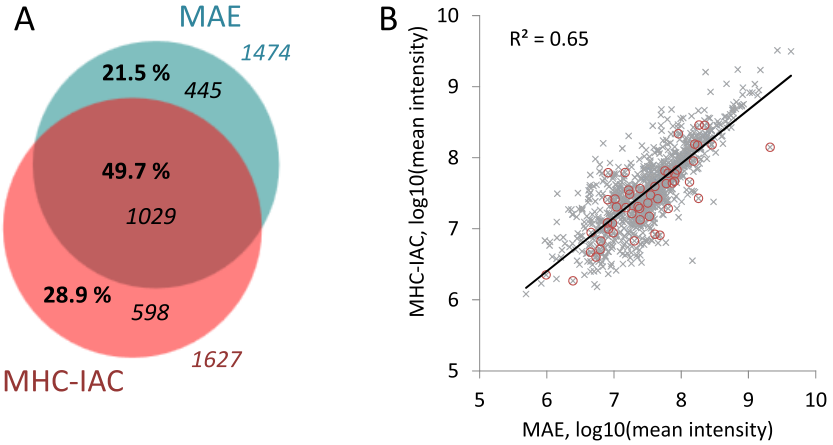


Figure 1. MAE and MHC-IAC identify a large overlapping set of peptides with correlating MS1 intensities. Data are derived from exactly the same cell culture using 5×10^8 JY cells each for MAE and MHC-IAC. (A) Most of the peptides obtained by MAE are also obtained by MHC-IAC and vice versa. The numbers in italics indicate absolute numbers of peptides. (B) Quantitative MS1 information derived from MAE versus MHC-IAC is similar for most peptides. The mean area of MS1 intensities was calculated for all peptides shared between MAE and MHC-IAC based on the three LC-MS replicates per extraction method. Predicted MHC non-binders (NetMHC $IC_{50} \geq 500$ nM) are highlighted with red circles in the dot plot.

applied a normalized collision energy of 35% and an activation time of 30 ms for CID, and we recorded MS2 spectra in the linear trap quadrupole. Automatic Gain Control targets were 500,000 ions to be reached in a maximum of 500 ms for MS1 and 10,000 ions to be reached in a maximum of 200 ms for MS2.

The Orbitrap Fusion was coupled online to an ultra high performance LC (1290 Infinity, Agilent) via a nanoelectrospray ion source (Nanospray Flex, Thermo Scientific). Samples were loaded at a flow rate of 5 μL/min for 5 min onto a 2 cm × 100 μm C_{18} trapping column packed in-house with C_{18} beads of 3 μm diameter (ReproSil-Pur C18-AQ, Dr. Maisch GmbH). Peptides were then separated on a 50 cm × 75 μm column in-house filled with C_{18} beads of 2.7 μm diameter (Poroshell 120 EC-C18, Agilent). Separation was run at 21 °C using a flow rate of approximately 300 nL/min and a 90 min gradient ranging from 5.6% CH_3CN to 32% CH_3CN in 0.1% HCOOH. MS1 scans were acquired at a resolution of 60,000 at 200 Th. The measured MS1 m/z range was 400−650 Th, and only 2+ and 3+ charged precursors were allowed for fragmentation. The Top Speed method was applied with cycle times of 3 s to fragment the most abundant precursor ions either by HCD (only LC-MS runs for Supplementary Figure S9A and Supplementary Figure S10A) or by EThcD (all other LC-MS runs at Orbitrap Fusion; compare Supplementary Table S9). Selected precursors were isolated at an isolation width of 1.6 Th and dynamically excluded from repeated fragmentation for 18 s. We applied a collision energy of 35% for HCD and a supplemental activation collision energy of 40% for EThcD, and we recorded MS2 spectra at a resolution of 15,000 at 200 Th in the Orbitrap mass analyzer. Automatic Gain Control targets were 400,000 ions to be reached in a maximum of 50 ms for MS1 and 50,000 ions to be reached in a maximum of 250 ms for MS2.

LC-MS Data Analysis

MS data were processed with Proteome Discoverer 1.4 from Thermo Fisher. We utilized Mascot version 2.2.4 from Matrix Science as a search engine for our Orbitrap XL data and the Q Exactive data,[22] whereas Sequest[45] HT was used for our Orbitrap Fusion measurements as well as for the Q Exactive HF data[18] and the comparative reprocessing of our data in Supplementary Table S2. For searching Orbitrap XL data, the allowed MS1 mass error was usually 5 ppm, but for some experiments, it had to be homogeneously increased to up to 15 ppm due to the occasional loss of the lock mass[44] at specific retention times. For linear trap quadrupole MS2 spectra, 0.5 Da mass error was accepted. The Q Exactive data[22] was reprocessed with 6 ppm MS1 and 0.02 Da MS2 mass tolerance. Our Orbitrap Fusion data and the Q Exactive HF data[18] were searched with 10 ppm MS1 and 0.02 Da MS2 mass tolerance. For EThcD spectrum matching, c and z ions obtained a weight of 1, while the weight of b and y ions was reduced to 0.5. Oxidation of methionine and cysteinylation of cysteine were set as variable modifications. Carbamidomethylation of cysteine was considered in database searches regarding peptide extractions applying iodoacetamide, i.e., for all of our MAEs and the MHC-IACs from Bassani-Sternberg et al.[22] The target protein database consisted of the human Swiss-Prot database (release September 27, 2013; 20,279 proteins; www.uniprot.org) supplemented with contaminant protein sequences. In the case of the influenza virus experiments, proteins of influenza virus A/Regensburg/D6/2009(H1N1) were also included. Our peptide spectral match false discovery rate (PSM FDR) cutoff of ≤5% was calculated by Percolator[46] using the reversed target protein sequences as a decoy database. Except for the dedicated processing for Supplementary Figure S6, only peptides of 8−12 amino acids in length were allowed. Additionally, we filtered the obtained peptides to be identified with search engine rank 1 and to have a Mascot ion score of ≥20 or a Sequest XCorr of ≥2. For the Euler diagrams as well as for the bar charts of MHC allotype distributions, peptides differing only in the modification status of methionine or cysteine were considered as equal. Euler diagrams were generated using BioVenn (www.biovenn.nl). MS1 intensities were derived from MS1 area calculations of Proteome Discoverer 1.4. Hydrophobicity indices were assigned by predicting the LC retention times of peptides using the Sequence Specific Retention Calculator (SSRCalc[47]) available at http://hs2.proteome.ca/SSRCalc/SSRCalcQ.html. We selected the SSRCalc model "100 Å C18 column, 0.1% Formic Acid 2015".

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

RESULTS

Peptides Extracted by MAE and MHC-IAC Largely Overlap

Although several published[48−51] and non-published laboratory-specific variations in MHC peptide ligand preparations might strongly influence the results (see the section on acetonitrile in ultrafiltration below and refs 8 and 52), we reasoned that some general features of identified MHC peptide ligandomes should be shared within MAE approaches and within MHC-IAC approaches, respectively, due to the fundamental characteristics of each of the two extraction methods (Supplementary Figure S1).

The optimization of our MAE method was based on a protocol from the Thibault group from 2008.[16] We hypothesized that many peptide and non-peptide contaminants might originate from cells bursting during the centrifugations after the application of MAE buffer (pH 3.3) or from extracellular vesicles and remaining cell debris exposed to acetonitrile and TFA during the first desalting step. Hence, after resuspending the cells in MAE buffer, we performed additional centrifugations that gradually increased in strength and time, and we finally ultracentrifuged the supernatant. Furthermore, we introduced additional protease inhibitors into the procedure and added GM dipeptide[40] to the MAE elution buffer. Our reference MHC-IAC method employed for comparison with MAE was the standard MHC-IAC method applied in the Rammensee laboratory using the anti HLA-A, -B, -C antibody W6/32 and including the acetonitrile-based ultrafilter wash (see the Methods section for details; compare ref 48).

We used both the B lymphoblastoid cell line JY and the monocytic leukemia THP-1[38] cell line to carry out a direct comparison of MAE and MHC-IAC starting from the same cell culture and using LC-MS of the extracted peptidome as the readout (Supplementary Figure S1). MAE did result in similar peptide numbers when compared to MHC-IAC. Specifically, MAE yielded 9% (JY cells) to 30% (THP-1 cells) fewer peptides than MHC-IAC but also revealed MHC peptide ligands not detected by MHC-IAC. Of the 2072 peptides obtained in total from the JY cells, 50% were discovered with both MAE and MHC-IAC (Figure 1A). For THP-1 cells, 36% of the identified 1268 peptides were found with both extraction methods (Supplementary Figure S2A). Notably, the overlap was higher when considering only the peptides predicted as strong MHC binders. Such peptides with a NetMHC[31] IC_{50} of <50 nM showed 55% (JY) and 41% (THP-1) overlap, respectively, comparing MAE with MHC-IAC. In contrast, putative contaminants, assigned by having a NetMHC IC_{50} of ≥500 nM and a relative Syfpeithi[30] score of <50%, were rarely identical in both MAE and MHC-IAC. They displayed only 17% (JY, 32 shared peptides of 189 contaminants) and 7% (THP-1, 12 shared peptides of 168 contaminants) overlap, respectively.

At first glance, the fraction of peptides shared between MAE and MHC-IAC might not appear very high. However, when performing MAEs from biological replicates of JY cells in different months, the overlap between the MAE samples was in the same range (46%, Supplementary Figure S3A). Comparing MAE versus MAE or MHC-IAC versus MHC-IAC from a *single* cell culture harvesting (like in the comparison of MAE and MHC-IAC), the overlap was higher (Supplementary Figure S3C and Supplementary Figure S4A; compare Supplementary Figure S15A and E as well as Supplementary Figure S16A and D), but even LC-MS triplicates of one and the same sample provided only 64% overlap due to the stochastic nature of the applied data dependent acquisition (DDA) MS scheme (Supplementary Figure S3E). Hence, more than half of the 50% (JY) and 64% (THP-1) of the peptides that are not shared between MAE and MHC-IAC can already be due to the LC-MS peptide identification variability. Please note that LC-MS runs of different samples were processed separately, i.e., independent from each other, in Proteome Discoverer. Peptide identifications were also not matched between runs based on precursor m/z and retention time. Taking into account the variability that is inherent to each peptide extraction method, the systematic qualitative difference between MAE and MHC-IAC appears to be quite limited.

Efficient elution of peptides from MHC-II proteins is putatively not feasible in MAE,[11−13] and therefore, we put no effort into analyzing the MHC-II immunopeptidome by MAE. We tested whether MHC-II peptide ligands might contaminate our MAE data set and contribute to the putatively not MHC-I binding peptide repertoire discovered by us. However, this did not seem to be an appreciable interference (Supplementary Figure S5).

Besides the qualitative similarities, we observed a remarkably good correlation of quantitative MS1 information on the shared peptides identified with MAE and MHC-IAC (R^2 = 0.65 and R^2 = 0.62, respectively; Figure 1B and Supplementary Figure S2B). This did not change significantly when considering only shared putative MHC non-binders (NetMHC IC_{50} ≥ 500 nM, R^2 = 0.65 and R^2 = 0.68, respectively). Although the correlation between replicate MAEs (R^2 ≥ 0.83, compare Supplementary Figure S15D and Supplementary Figure S15H) and between replicate MHC-IACs (R^2 = 0.91, Supplementary Figure S4C) was even better, the variation inherent to each of the two extraction methods and the LC-MS triplicate analyses again contributes a considerable part of the observed total variation between MAE and MHC-IAC. In line with the observed correlations, 69.3% (data set of Figure 1B) and 63.2% (data set of Supplementary Figure S2B), respectively, of peptides shared between MAE and MHC-IAC differed by less than a factor of 2 in mean MS1 intensities, while this percentage was ≥79.1% in the MAE versus MAE comparisons (data sets of Supplementary Figure S15D and Supplementary Figure S15H) and 87.7% in the MHC-IAC versus MHC-IAC comparison (data set of Supplementary Figure S4C).

Peptides Extracted by MAE and MHC-IAC Share Generic Features

We next asked whether peptides isolated by MAE and MHC-IAC differ in their immunobiochemical properties. Regarding peptide lengths, both extraction methods yielded very similar results with 9-mers being the dominant group followed by 10-mers (Supplementary Figure S6). This is in agreement with the known lengths preferred by the MHC-I allotypes of JY and THP-1 cells (www.syfpeithi.de).[30] Please note that we excluded singly charged precursor ions from fragmentation for MS2 which might have introduced a small bias against short peptides. In MAE, the summed fraction of 8- to 12-mers among all identified peptides was 96% (JY) and 90% (THP-1), respectively, which is much better than the 52%[14] or almost 60%[16] reported for MAE in previous years and on par with recent data from Lanoix et al.[18] As expected, MHC-IAC yielded an even higher fraction of peptides of the appropriate length (99% for JY and 98% for THP-1, Supplementary Figure S6). According to this finding and common habit in MHC immunopeptidomics, we considered only 8- to 12-mers for all

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

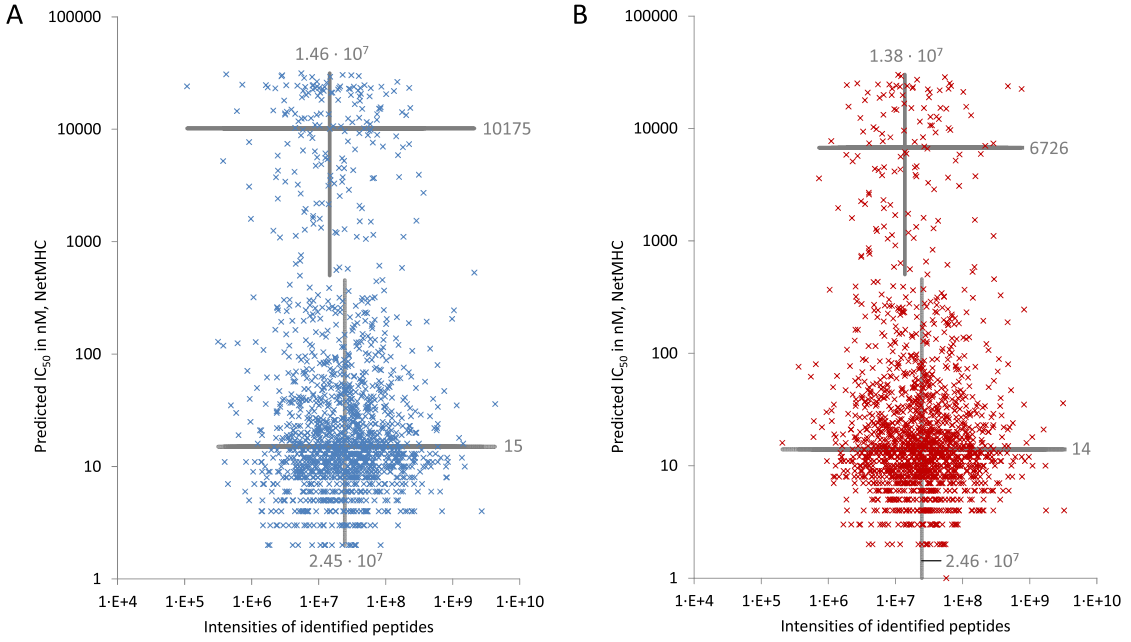


Figure 2. Peptides isolated by MAE and MHC-IAC show a similar distribution of predicted MHC binding affinities. All displayed data originate from the same cell culture harvesting using 5×10^8 JY cells each for MAE and MHC-IAC. Each data point is derived from the mean area of MS1 intensities of a single peptide as measured in triplicate LC-MS injections. The gray bars represent the median of the NetMHC-IC_{50} and the median of the corresponding peptide intensities, respectively. Medians were calculated separately for predicted MHC binders (IC_{50} < 500 nM) and non-binders ($IC_{50} \geq$ 500 nM) and are indicated with gray numbers. (A) Peptides obtained by MAE. (B) Peptides obtained by MHC-IAC.

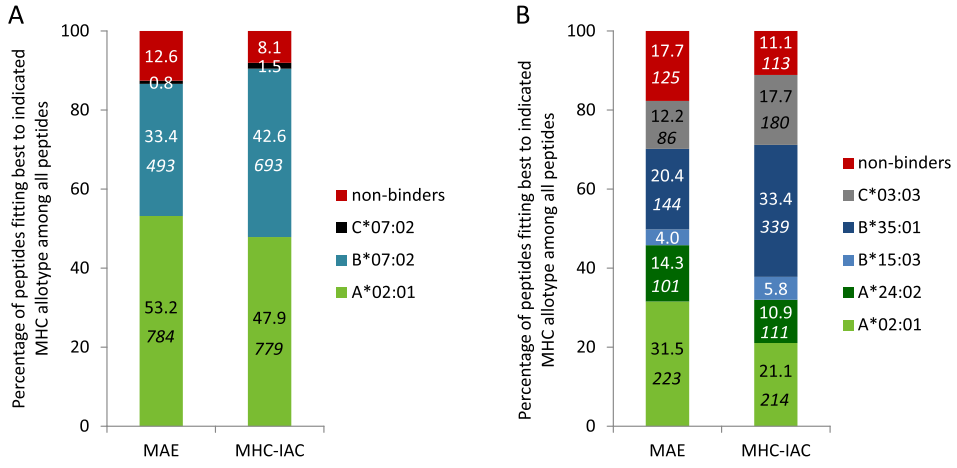


Figure 3. The modestly higher peptide yield of MHC-IAC as compared to MAE was mainly due to a better performance for HLA-B allotypes. Each peptide was assigned to that MHC allotype of the respective cell line yielding the highest NetMHC binding affinity. Peptides with a NetMHC $IC_{50} \geq$ 500 nM were considered as non-binders. Absolute peptide numbers are indicated in italics. (A) Peptides obtained from exactly the same cell culture using 5×10^8 JY cells each for MAE and MHC-IAC. (B) Peptides obtained from exactly the same cell culture using 1.25×10^8 THP-1 cells each for MAE and MHC-IAC.

following analyses, as peptides longer than 12 amino acids are expected to mostly represent contaminants.

Notably, the pH and time in which peptides are eluted from MHC-I proteins in MAE appear to not be low enough and long enough, respectively, for the (complete) denaturation of MHC-II peptide complexes.[11,12] We therefore checked whether very-high-affinity MHC-I peptide ligands might resist the elution at pH 3.3 in MAE. However, we did not find any evidence for a strong underrepresentation of such peptides in MAE as compared to MHC-IAC. Several high-affinity peptides with predicted NetMHC IC_{50} values as low as 3 nM and even 2 nM were recovered by the MAE approach (Figure 2, Supplementary Figure S7).

Peptides with low predicted MHC binding affinities are usually assigned as putatively non-MHC-related, contaminating peptides. In agreement with the original recommendation of NetMHC version 3, we here used the IC_{50} 500 nM threshold for the discrimination between MHC binders and MHC non-binders.[31] Among the putative MHC binders, the median predicted IC_{50} was very similar for MAE (15−17 nM) and MHC-IAC (14 nM). Remarkably, the overall median over the MS1 intensities of all identified MHC peptide ligands differed by less than 2% between the two extraction methods (Figure 2, Supplementary Figure S7), which is less than the difference observed in peptide numbers (Figure 1A and Supplementary Figure S2A) and suggests a very similar total peptide yield. As

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

expected, MS1 intensities of MHC non-binders were higher in MAE as compared to MHC-IAC, but the difference was on average only 6% (JY, Figure 2) and 14% (THP-1, Supplementary Figure S7), respectively. Most importantly, the fraction of contaminating, MHC non-binding peptides among all identified peptides was only 5 and 7%, respectively, higher in MAE as compared to MHC-IAC (Figure 3; refer to Supplementary Figure S3D and Supplementary Figure S4B for reproducibility). In all of our MAE samples, this fraction was lower than 18%, which is a big improvement to the 60%[16] and approximately 30−40%[17] reported for MAE in previous years and similar to the MAE data reported recently by Lanoix et al.[18]

It cannot be ruled out that the anti HLA-A, -B, -C monoclonal antibody W6/32 frequently applied in MHC-IAC has a bias toward certain MHC-I allotypes or MHC-I peptide complexes, and this has not been investigated comprehensively so far.[9] Here, we compared the MHC allotype assignments of peptides obtained with MAE versus W6/32-based MHC-IAC. While the two extraction methods yielded virtually equal numbers for the investigated HLA-A ligands, MHC-IAC resulted in a higher number of ligands for the respective HLA-B and HLA-C allotypes (Figure 3). This difference was significant ($P < 0.01$) for HLA-B*07:02 (calculating the confidence interval using MAE-based biological replicates depicted in Supplementary Figure S3B) as well as for HLA-B*35:01 and HLA-C*03:03 (calculating the confidence intervals using the MAE replicates shown in Supplementary Figure S3D that were performed from the same THP-1 cell harvesting as the MHC-IAC). Note that the fraction of certain MHC allotypes differed by up to 4% even between a pair of triplicate LC-MS measurements (Supplementary Figure S3F).

Better Recovery of Cysteinylated Peptides by MAE than MHC-IAC

The underrepresentation of cysteine-containing peptides is a well-known bias of immunopeptidomic data sets. One reason for this is that cysteine residues of MHC peptide ligands can convert to cyst*i*ne residues by forming disulfide bridges with free cysteine originating, e.g., from the cell culture medium.[53] Accordingly, the bias against cysteine-containing peptides in immunopeptidomic data sets can be reduced by considering cysteinylation of cysteine as a variable modification in MS data processing.[27,34] Cysteinylated cysteine can synonymously be referred to as cyst*i*ne. The relative numbers of cysteine-containing peptides in our MHC-IAC data varied between <1 and 3% when comparing samples of different cell lines and cell cultures, and they also differed by more than a factor of 4 in the subfractions of peptides with cysteinylated and unmodified cysteine residues, respectively. However, such variations were very small if the samples were derived from the same cell culture (Supplementary Table S1). In contrast, even when extracting peptides from exactly the same cell culture, MAE yielded on average more than 5 times higher relative numbers of peptides containing cyst*i*ne residues than MHC-IAC (Table 1).

To further validate the higher percentage of cyst*i*ne residues in MAE as compared to MHC-IAC, we reprocessed the LC-MS data from MAE and MHC immunopurification performed by Lanoix et al.[18] Thereby, we set cysteinylation as a variable modification for database searches. Comparing MAE and MHC-IAC data that share equal total peptide counts, we observed a 2.6-fold higher percentage of cysteinylated peptides in Lanoix's MAE data compared to their MHC immunopurification data (Supplementary Figure S8). However, it should be noted that

Table 1. Peptide Extracts Following MAE Show a Much Higher Relative Number of Cysteinylated Peptides than Those from MHC-IAC[a]

		cysteinylated peptides in %	
cell line	MAE and MHC-IAC from exact same cell culture	MAE	MHC-IAC
JY	yes	3.26	0.43
	no	2.95	0.57
		3.21	n.a.
		2.37	n.a.
THP-1	yes	2.83	0.59
		2.30	n.a.
		3.42	n.a.
mean		**2.91**	**0.53**

[a]Data points for JY cells represent biological replicates, while all percentage values for THP-1 are derived from parallel extractions from the same cell culture. To improve comparability between MAE and MHC-IAC, only samples measured at the LTQ Orbitrap XL are considered in this table. The observed difference between MAE and MHC-IAC has a P value of 2.6×10^{-6} performing a two-tailed, heteroskedastic Student's t-test.

MAE yielded a much lower total peptide count than MHC immunopurification in the study of Lanoix et al. (Supplementary Table S2).

Most immunopeptidomic studies either do not consider cysteine modifications[18,22,52,54−56] or use cysteinylation as the only cysteine modification allowed in database searches.[7,24,27,57] In contrast, Trujillo et al. reported glutathionylation and mono-, di-, and trioxidation as the most abundant cysteine modifications in their MHC-IAC data sets generated without an alkylating reagent like iodoacetamide.[58] We therefore subjected our MS raw data from Figure 1 (MAE and MHC-IAC from JY cells) to additional, separate database searches allowing these variable modifications. However, for each of the four modifications, we obtained only between 0 and 7 peptide hits per extraction method. We then repeated the searches using (non-existing) variable decoy modifications replacing the cysteine SH group with SC_4, SH_4, or SN_4. For each of these decoy modifications, we obtained 1−4 peptide hits per extraction method, implying that such low peptide numbers cannot be used for reliable comparisons. Our same data set yielded 48 and 7 cysteinylated peptides in MAE and MHC-IAC, respectively, demonstrating that cysteinylation was much more prominent than glutathionylation and mono-, di-, and trioxidation of cysteine.

MAE Reveals a Distinct Pattern of Cysteinylation for HLA-A*02:01 and HLA-B*07:02 Peptide Ligands

We next harnessed the superior relative and absolute recovery of cysteinylated peptides in our MAE to obtain new insights into the immunobiochemistry of cysteinylation by investigating these peptides' frequencies in different MHC allotypes and by exploring the frequencies of cysteinylation at the distinct positions of the MHC peptide ligands. Also, to improve the statistical power of these analyses, we performed five additional MAEs from JY cells and measured the samples with EThcD[23,59] at the Orbitrap Fusion, yielding 3633, 4956, 4925, 5149, and 5064 peptides, respectively.

Importantly, the percentage of putative MHC binders among MAE-derived peptides was the same or even higher for cysteinylated as compared to non-cysteinylated peptides, indicating that the better recovery of cysteinylated peptides in

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

MAE as compared to MHC-IAC is not due to contaminants. In both JY and THP-1 cells, the fraction of putative HLA-A*02:01 binders was significantly higher for cysteinylated versus non-cysteinylated peptides (Figure 4A). In contrast, putative HLA-B*07:02 ligands were underrepresented in the set of cysteinylated peptides (Figure 4B), demonstrating that the occurrence of cystine residues is MHC allotype dependent.

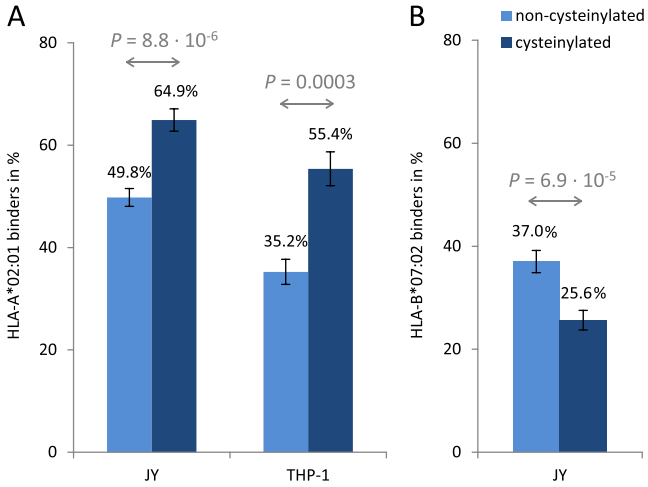


Figure 4. Cysteinylated peptides isolated by MAE are enriched in HLA-A*02:01 binders and diminished in HLA-B*07:02 binders. Data for JY cells are derived from five MAEs of JY cells that were analyzed employing EThcD. Data from THP-1 cells represent four MAEs performed in parallel and measured using CID. Identified peptides of each sample were grouped into two bins: non-cysteinylated (pale blue) and cysteinylated (dark blue) peptides. The mean number of non-cysteinylated peptides per MAE was 4640 for JY cells and 969 for THP-1 cells. The mean number of cysteinylated peptides per MAE amounted to 105 for JY cells and 32 for THP-1 cells. Peptides with a NetMHC IC_{50} of <500 nM were considered as binders. Columns represent means, and error bars indicate standard deviation. Depicted *P* values are derived from two-tailed, heteroskedastic Student's *t*-tests with Bonferroni correction for three comparisons. (A) Proportion of predicted HLA-A*02:01 binders among all identified peptides. (B) Proportion of predicted HLA-B*07:02 binders among all identified peptides.

Looking closer into these data, we observed highly distinct frequencies of cysteinylation at different positions of putative MHC peptide ligands, and these position-specific frequencies also differed between MHC allotypes. As expected, we observed only a very low cysteinylation frequency at position 2 and at the C-terminal position of HLA-A*02:01 and HLA-B*07:02 peptide ligands, i.e., at the MHC anchor positions. Interestingly, similarly low cysteinylation frequencies also occurred at position 4 of HLA-A*02:01 peptide ligands and at positions 1 and 3 of HLA-B*07:02 peptide ligands (Figure 5). Note that the identification of a few peptides with cysteinylation at a given position might simply be a random result of the false positive rate inherent to MS statistics. Hence, low observed cysteinylation frequencies at a given position do not yet prove that cysteinylation occurs there, but they clearly show that cysteinylation is not common at these positions. The low numbers of (cysteinylated) cysteine residues at the MHC anchor positions were not merely a result of NetMHC-based filtering for putative HLA-A*02:01 and HLA-B*07:02 ligands. This is evident from the fact that, even when including putative MHC non-binders with cysteine modification in the statistics, per sample at most four peptides (2.5%) were cysteinylated at position 2 and at most three (1.8%) at their C-terminus.

Cystine residues showed a strong preference for the penultimate C-terminal position (C-1) of both HLA-A*02:01 and HLA-B*07:02 peptide ligands. Additionally, higher frequencies of cystine residues were observed at position C-4 of HLA-A*02:01 but not HLA-B*07:02 peptide ligands (Figure 5). The observed cysteinylation patterns also contained a multitude of other statistically significant differences between individual positions (Supplementary Table S3 and Supplementary Table S4) and were largely stable using different biological conditions and MS settings (Supplementary Figure S9 and Supplementary Figure S10).

It should be noted that the observed preference of cystine for certain positions does not imply that cystine is enriched relative to other amino acids at that position. Although position C-1 is that position of HLA-A*02:01 peptide ligands showing the highest percentage of (cysteinylated, carbamidomethylated, or unmodified) cysteine, peptides containing cysteine at this position still constituted only 0.9% of the total number of HLA-A*02:01 peptide ligands. If the occurrence of amino acids was completely random and if there were no methodological biases against cysteine, the percentage of such peptides should be around 2.3%—the percentage of cysteine in the human proteome.[60]

Cysteinylation Pattern Obtained by MHC-IAC

Having characterized the positional cysteinylation pattern of MAE-derived peptides, we asked whether results obtained by MHC-IAC would differ. Measuring peptides from MAE or MHC-IAC at the LTQ Orbitrap XL, we demonstrated that the relative frequency of cysteinylated peptides was much lower in MHC-IAC as compared to MAE (Table 1). In combination with the only moderate total number of peptides obtainable at the LTQ Orbitrap XL, our overall numbers of MHC-IAC-derived cysteinylated peptides were too low for robust statistics on positional cysteinylation frequencies. Therefore, we reprocessed a more comprehensive MHC-IAC data set of JY cells measured at a Q Exactive by Bassani-Sternberg et al.[22] using cysteinylation as a variable modification in our database searches. Although the preference of cystine residues for position C-2 appeared more pronounced in this MHC-IAC data set as compared to our MAE data set, the other main features of the MAE-derived cysteinylation pattern were shared (Supplementary Figure S11).

Cysteinylation Occurred Most Frequently at Positions with Intermediate Side Chain Exposure

The first report on cysteinylated MHC peptide ligands already speculated that cysteinylation was a "highly probable modification of any free cysteine residues in peptides that point out of the MHC-binding pocket".[53] One problem for testing this hypothesis is the limited number of available MHC crystal structures combined with the fact that the side chain accessibility of MHC peptide ligands is often strongly dependent on the MHC allotype and the sequence of the presented peptide. However, we reasoned that averaging structural measures of peptide ligand side chain exposure across many different crystal structures representing the same MHC allotype might reveal some positions that are on average more exposed than others and might therefore be, again on average, more prone for cysteinylation. Using 10 criteria specified in the Supporting Information, we compiled a set of crystal structures from the Protein Data Bank (PDB) for our side chain exposure calculations. Of the 231 PDB submissions of HLA-A*02:01

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

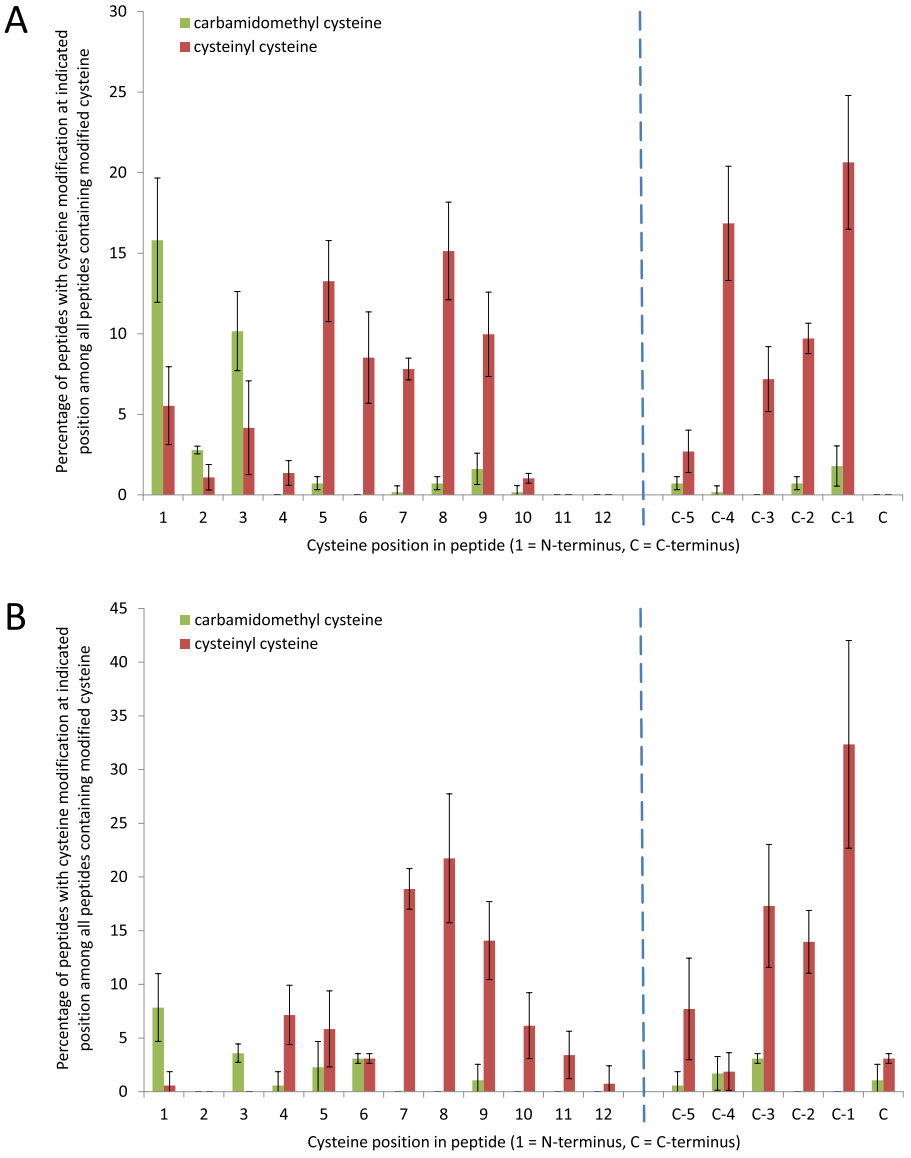


Figure 5. Cysteinylated cysteine residues are preferentially located at position C-1 of both HLA-A2 and HLA-B7 peptide ligands, whereas a preference for position C-4 occurs in the context of HLA-A2 but not HLA-B7. The columns represent means from five MAEs of JY cells that were analyzed employing EThcD. To adjust for different peptide lengths, more C-terminal amino acid positions were also counted relative to the C-terminus, so, e.g., "C-4" refers to the position four amino acids N-terminal of the C-terminus, whereas "C-1" represents the position adjacent to the C-terminus. The error bars indicate the standard deviation. Accompanying *P* values are given in Supplementary Tables S3 and S4. (A) Frequency of modified cysteine residues in cysteine-containing peptides with a NetMHC IC_{50} of <500 nM for HLA-A*02:01; numbers of HLA-A*02:01 motif peptides containing modified cysteine (n) = 114, 114, 64, 102, and 117, respectively, in total 173 non-redundant peptides. (B) Frequency of modified cysteine residues in cysteine-containing peptides with a NetMHC IC_{50} of <500 nM for HLA-B*07:02; numbers of HLA-B*07:02 motif peptides containing modified cysteine (n) = 32, 31, 27, 35, and 40, respectively, in total 66 non-redundant peptides.

available in June of 2017, we thereby ended up with 33 representative entries, and from the available 5 HLA-B*07:02 crystal structure submissions, 4 passed our criteria.

Because not only the solvent accessible surface (SAS) but also the relative SAS (RSAS) is strongly dependent on the type of amino acid residue at a given peptide position, we took advantage of the more robust Half-Sphere Exposure (HSE)-β-up measure of side chain accessibility. Whereas the coordination number (CN) counts the number of all surrounding Cα atoms within a given distance (1.2 nm in our study) from the peptide's Cα atom of interest, HSE-β-up only counts the Cα atoms from the CN that are situated in the half sphere defined by the Cα−Cβ axis, i.e., in the direction of the side chain.[61] Calculating the mean CN, HSE-β-up, and RSAS values for each peptide position across the 33 selected HLA-A*02:01 PDB entries as specified in the Supporting Information, we obtained representative average measures for exposure (Supplementary Figure S12) that differed significantly between the peptide positions (Supplementary Table S6).

Visual inspection of representative HLA-A*02:01 crystal structures as well as calculated CN, HSE-β-up, and RSAS values highlight that positions 1 and 3 together with the HLA-A*02:01 anchor positions 2 and C-terminus are on average the most buried positions of HLA-A*02:01 peptide ligands, whereas positions 4 and C-5 (identical for 9-mer peptides) are typically the most exposed residues (Supplementary Figure S12A−D). As

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

expected, the buried positions 1 and 3 were associated with a low cysteinylation frequency apparently due to steric hindrance. Remarkably, the strongly exposed positions 4 and C-5 showed similarly low cysteinylation frequencies (Figure 5A and Supplementary Figure S12E), whereas the most frequently cysteinylated positions C-4 and C-1 turned out to be only partly exposed on average. Hence, steric accessibility is not the only prerequisite for abundant cysteinylation at a certain position of MHC peptide ligands.

Is the low frequency of cysteinylation at the strongly exposed positions 4 and C-5 possibly a result of the chemical properties of cystine? Cystine residues are much bigger than the standard amino acids, and they possess a long hydrophobic region between their two Cα atoms. If this long hydrophobic arm is exposed to the aqueous solution, this is energetically not optimal. In support of this notion, we found that not only very big but also hydrophobic amino acids are strongly underrepresented at positions 4 and C-5 of HLA-A*02:01 peptide ligands (Supplementary Figure S12F and G).

We next asked whether the observed relation between positional cystine frequencies in HLA-A*02:01 peptide ligands and average positional residue exposure is MHC-allotype-specific or whether it might represent a more general association. To address this question, we compared cystine frequencies of HLA-B*07:02 peptide ligands with the structural exposure parameters derived from the four respective PDB entries. However, this was not straightforward; for SAS and RSAS, we could not observe any significant difference between the individual positions, and for CN and HSE-β-up, such differences were much less abundant than in the case of HLA-A*02:01 peptide ligands (Supplementary Table S7). This is not surprising given the low number of four PDB entries for HLA-B*07:02 and the fact that two of them are putatively not very representative of average HLA-B*07:02 peptide ligands (Supplementary Figure S13). Nevertheless, some of our basic findings obtained for HLA-A*02:01 peptide ligands were similarly observed for HLA-B*07:02 peptide ligands. Also, for HLA-B*07:02, peptide ligand positions 1 and 3 are deeply buried, and this was associated with very low cysteinylation frequencies. Again, those positions that seemed most exposed on average, i.e., positions 4, C-5, and C-4 of HLA-B*07:02 peptide ligands, showed much lower cysteinylation frequencies than the most abundantly cysteinylated position C-1, with the latter being characterized by partly but not fully exposed amino acid side chains (Figure 5B and Supplementary Figure S13).

The most prominent difference in the cysteinylation patterns of HLA-A*02:01 versus HLA-B*07:02 peptide ligands occurred at position C-4. Whereas this position was one of the two preferred sites of cystine residues for HLA-A*02:01 peptide ligands, it was seldom cysteinylated in HLA-B*07:02 peptide ligands. We realized that the MHC peptide binding groove in the vicinity of peptide position C-4 is usually narrower in HLA-B*07:02 as compared to HLA-A*02:01 (Supplementary Figure S14A−D). HSE-β-up values of putatively typical HLA-B*07:02 peptide ligands have already suggested that peptide position C-4 is on average more strongly exposed in HLA-B*07:02 than in HLA-A*02:01 (Supplementary Figure S12C and Supplementary Figure S13C). Confining the comparison to crystal structures containing the amino acids leucine, phenylalanine, or tyrosine (which share some size and hydrophobicity features with cystine) at peptide position C-4, it is also obvious by RSAS and by visual inspection that HLA-A*02:01 tends to accommodate these amino acids deeper in its binding groove than HLA-B*07:02, resulting in lower, energetically more favorable, solvent exposure (Supplementary Table S8 and Supplementary Figure S14C−F). Therefore, it is tempting to speculate that HLA-A*02:01 is also superior to HLA-B*07:02 in shielding the hydrophobic basis of the cystine side chain at position C-4 from the aqueous solvent supporting higher cysteinylation frequencies at this side. However, the generation of crystal structures with cysteinylated MHC peptide ligands would be necessary to prove this hypothesis.

Efficient Detection of Unmodified Cysteine Residues

Robust identification of MHC peptide ligands with unmodified cysteine residues requires the application of a cysteine protecting reagent such as iodoacetamide and appropriate processing of LC-MS data. We suppose that the absence of this step in the analytical process might have contributed to the underrepresentation of cysteine residues in current data sets of MHC peptide ligands.[27,34] Indeed, most immunopeptidomic laboratories do not routinely add a cysteine protecting reagent during cell lysis for MHC-IAC.[7,8,18,24,27,49,54−57] Accordingly, for better representativeness of this comparative study, we did not protect free cysteines in our MHC-IAC procedure. In contrast, iodoacetamide is routinely included in the MAE elution buffer as an inhibitor of cysteine proteases,[16] and we therefore also applied it in the present MAE experiments. For the first time (as far as we are aware), we included carbamidomethylation in the bioinformatic processing of MAE-derived LC-MS data to further reduce the bias against cysteine-containing peptides.

In MAE, the lion's share of carbamidomethylation in putative HLA-A*02:01 peptide ligands occurred at the three N-terminal peptide residues (Figure 5 and Supplementary Figure S9). For comparison with MHC-IAC, we took advantage of a data set of JY cells from Bassani-Sternberg et al.[22] where iodoacetamide had been employed, and we reprocessed it using carbamidomethylation as a variable modification. For this MHC-IAC data, we found that carbamidomethylation was more equally distributed across the length of HLA-A*02:01 peptide ligands, and it was missing at positions 1 and 2 (Supplementary Figure S11).

The partly contradictory carbamidomethylation patterns of HLA-A*02:01 peptide ligands in MAE and MHC-IAC might be explained by the different biochemical conditions in MAE and MHC-IAC. In MHC-IAC, the iodoacetamide was added during cell lysis and had already been washed off when peptides were eluted from MHC proteins.[22] Therefore, the unmodified cysteine residues that had reduced accessibility to iodoacetamide while sitting in the MHC binding groove were likely to go undetected in MHC-IAC. In contrast, it is conceivable that reduced, i.e., unmodified, cysteine residues originate *de novo* during cell lysis for MHC-IAC because the MHC peptide ligands are artificially exposed to the reducing components of the cytosol, or simply because of the chemical equilibrium of the redox reaction (e.g., cysteinylation and decysteinylation). If iodoacetamide is present, such artificially unmodified cysteine residues are irreversibly carbamidomethylated, and this could account for the higher proportion of carbamidomethylation in the more solvent exposed middle and C-terminal regions of HLA-A*02:01 peptide ligands (compare Supplementary Figure S12A−D) observed in MHC-IAC as compared to MAE. In contrast to MHC-IAC,[22] our MAE protocol ensured the presence of iodoacetamide during peptide elution from the MHC proteins. Buried, unmodified cysteine residues with reduced accessibility to iodoacetamide (and other modifying

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

molecules) while sitting in the MHC binding groove therefore got a chance to be carbamidomethylated and detected by MAE. Accordingly, the buried position 1 of HLA-A*02:01 peptide ligands was the dominant position for carbamidomethylation in MAE (Figure 5 and Supplementary Figure S9), whereas no carbamidomethylation was observed at this position in MHC-IAC (Supplementary Figure S11). MAE inherently avoids the exposure of MHC peptide ligands to reducing cytosolic molecules and consequently limits the possibility for *de novo* generation of unmodified cysteine residues during sample processing. In summary, taking the application of a cysteine protecting reagent for granted, we therefore hypothesize that MAE reveals the native distribution of unmodified cysteine residues much better than MHC-IAC.

Acidic Elution of MHC Molecules Appears to Be Quite Complete in MAE and MHC-IAC

Despite some striking differences with regard to the identification of unmodified cysteine residues, overall peptide identifications and peptide features were similar between MAE and MHC-IAC. We next asked whether this reflects a rather representative recovery of the immunopeptidome by either method or whether the similarity might just be a result of a potential bias shared between MAE and MHC-IAC, i.e., the use of acid for the dissociation of peptide ligands from the MHC.

We addressed the question with a dual approach and first tested the effect of organic solvents on MHC peptide complex dissociation, assuming that these could promote the elution of hydrophobic peptides unspecifically bound to denatured MHC proteins. However, the inclusion of 10% DMSO in the MAE elution buffer did not notably change the overall peptide yield, MHC allotype distribution, recovery of hydrophobic peptides, and MS1 peptide intensity profiles (Supplementary Figure S15). Likewise, the inclusion of 60 or 80% acetonitrile in the elution solution for MHC-IAC columns had no major effect on the observed peptide pattern, and it did not increase the peptide yield (Supplementary Figure S16).

In the second approach, we performed a repeated MAE from JY cells 3 h after the first MAE. The number of identified peptides was only 2.1% (first biological replicate) and 3.5% (second biological replicate), respectively, in the repeated MAE when referenced to the first MAE of exactly the same cells. Hence, the acidic elution of MHC molecules appears to be quite complete in MAE and MHC-IAC, and we found no evidence that it would introduce a major bias.

Our results also indicate that the application of repeated MAEs[9,10] for experiments on the short-term kinetics of MHC peptide ligands seems not beneficial. Importantly, although 79−82% of JY cells were still alive after the first and directly before the repeated MAE, less than 2% survived the repeated MAE taking place 3 h after the first MAE (two biological replicates; staining with trypan blue). Obviously, the MAE procedure results in substantial cellular stress limiting the biological value of potential kinetic studies employing repeated MAEs.

Usage of Acetonitrile in Ultrafiltration Can Markedly Change the Results

Finally, we compared the variability between MAE and MHC-IAC with the variability originating from different ways of performing MHC-IAC. One critical step that has been differently performed in immunopeptidomic studies since the early 1990s is the separation of MHC peptide ligands from accompanying proteins after elution from the MHC-IAC column.[8] While some investigators have applied reversed-phase LC for this purpose (e.g., refs 22 and 62−65), many others have preferred ultrafiltration instead (e.g., refs 19, 23, 24, and 55−57).

We hypothesized that MHC peptide ligands, especially if they are hydrophobic, are partly lost during the ultrafiltration step due to non-covalent adsorption to the large inner surface of the ultrafiltration membrane. To test and overcome this potential shortcoming, we passed an acetonitrile-rich solution through the filter after complete filtration of the MHC-IAC eluate. This rinse ultrafiltrate yielded 2−3 times more peptide identifications and boosted MS1 peptide intensities by 27-fold and 41-fold, respectively, when compared to the first ultrafiltrate. The benefits were especially pronounced for hydrophobic peptides. The combined ultrafiltrate, consisting of the mixture of the first and the rinse ultrafiltrate, resulted in the same superior performance as the rinse ultrafiltrate (Supplementary Figure S17). Therefore, the acetonitrile-based filter rinse and the subsequent analysis of the combined ultrafiltrate have been implemented as a standard in the Rammensee laboratory since 2014 and were also used for the MHC-IACs in the present study. Remarkably, the overlap of identified peptides between the first ultrafiltrate (two MHC-IAC replicates) and the combined ultrafiltrate was only 22 and 26%, respectively, and thereby considerably lower than the overlap between MAE and MHC-IAC.

We noted that the overall peptides' MHC allotype assignment was considerably shifted in the rinse ultrafiltrate and the combined ultrafiltrate, as compared to the first ultrafiltrate by itself. The magnitude of this alteration was similarly high as in the comparison of MAE versus MHC-IAC (Supplementary Figure S18A; compare to Figure 3). The shift introduced by acetonitrile-based filter rinsing appeared to be the result of the distinct average hydrophobicities of peptide ligands from different MHC allotypes. We benchmarked the MHC-allotype-specific hydrophobicity patterns of identified MHC peptide ligands by plotting the percentage contribution of each MHC allotype across the retention time of our LC-MS analyses. This way, it was experimentally demonstrated, for example, that HLA-A*24:02 peptide ligands are on average more hydrophobic than HLA-B*44:03 peptide ligands (Supplementary Figure S18B). It proved that peptides derived from MHC allotypes with a preference for hydrophobic ligands benefitted the most from the acetonitrile-based rinse. Because the acetonitrile-based filter rinsing strongly boosted peptide yield and reduced the bias against hydrophobic peptides, it had been included in the MHC-IAC protocol used for the comparisons with MAE (see the Methods section for details).

Given the huge improvement in peptide yield achieved by acetonitrile-based filter rinsing in MHC-IAC, we also tested the effect of an acetontrile-based filter rinsing in MAE. As expected from the fact that the MAE peptides are already contained in an acetonitrile-rich solution when passed through the ultrafilter for the first time, the benefit of the rinsing was only moderate here. However, in agreement with the results from MHC-IAC, peptides derived from MHC allotypes preferring hydrophobic ligands (e.g., HLA-A*02:01) benefitted most strongly from the rinsing, and hence, their percentage was higher in the rinse ultrafiltrate and the combined ultrafiltrate, as compared to the first ultrafiltrate alone (Supplementary Figure S18C−E).

The distinct hydrophobicity profiles observed for peptide ligands of different MHC allotypes were not merely an effect of allotype assignment by the NetMHC algorithm[31] which indeed appropriately takes hydrophobicity into account in its binding

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

predictions (Supplementary Figure S19A−C). Similar hydrophobicity profiles were also obtained when assigning our identified peptides using either the Syfpeithi algorithm[30] or solely predefined anchor amino acid residues (Supplementary Figure S19D and G). However, we noticed that the Syfpeithi algorithm does not sufficiently consider peptide hydrophobicity in its MHC binding predictions (Supplementary Figure S19D, E, and H).

DISCUSSION

In this manuscript, we introduce, describe, and evaluate an optimized protocol for MAE. It resulted in a markedly lower percentage of non-MHC-related peptide contaminants compared to MAE studies published in earlier years.[16,17] In all samples, more than 82% of identified 8- to 12-mer peptides could be assigned to an appropriate MHC allotype using the NetMHC threshold of IC_{50} < 500 nM recommended in the original publication by Lundegaard et al.[31] Notably, this threshold is much more stringent than the thresholds of IC_{50} ≤ 1250 nM[17] or even IC_{50} < 5000 nM[66] that were applied in earlier studies employing MAE. Only recently, Lanoix et al. obtained similarly low numbers of non-MHC-related peptide contaminants in MAE as we did.[18] However, unlike us, Lanoix et al. do not state possible reasons for the much better purity as compared to (their) previous MAE studies, and they do not report their detailed, putatively improved protocol. Considering the Minimal Information About an Immuno-Peptidomics Experiment (MIAIPE) guidelines,[67] we make our detailed MAE and MHC-IAC protocols available for the public.

Contradictory to previous assumptions,[10,37] Lanoix et al. observed much lower peptide numbers in MAE as compared to MHC immunopurification.[18] In contrast, the number of identified peptides and especially their median MS1 intensities were on par between our MAE and our MHC-IAC, benefitting a better matched comparison of peptide properties between the two extraction methods. Considering that we measured our direct comparison of MAE and MHC-IAC two years before Lanoix et al. using a much less advanced mass spectrometer and that we applied only 10% instead of 100% of the sample per LC-MS injection, our MAE protocol seems to perform very favorably, and the higher peptide numbers in the MHC immunopurification of Lanoix et al. are readily explained. However, an accurate comparison of peptide yields between Lanoix et al. and our protocols is hampered by the use of different cell lines and LC-MS equipment (Supplementary Table S2).

If indeed ≥97% of MHC peptide ligands are lost during MHC-IAC as reported by Hassan et al., this loss seems to be surprisingly equally distributed across the different peptide species, and it does not seem to preclude an overall quite representative recording of the MHC immunopeptidome. Hassan et al. observed that a main source of peptide losses was the immunoaffinity purification step, which is totally absent in MAE.[37] We found no evidence that the acidic nature of the elution step common to both MAE and MHC-IAC would represent a considerable limitation for efficient recovery of MHC peptide ligands because the use of organic solvents during elution had no major effect on peptide yield or characteristics of obtained peptides. A relatively complete and efficient elution is also suggested by the facts that, first, a repeated MAE from MAE-eluted cells was unable to recover a substantial amount of peptide and that, second, the median MS1 intensity of strong MHC binders (NetMHC IC_{50} < 50 nM) was similar to the one of weak MHC binders (50 ≤ NetMHC IC_{50} < 500 nM; compare Figure 2 and Supplementary Figure S7). Therefore, it remains to be clarified where exactly and for what reason the potentially large but similar peptide losses in both MHC-IAC and MAE originate. This question has already been highlighted as the first "main challenge in the field" of immunopeptidomics by the Human Immuno-Peptidome Project[68] and will need thorough additional investigations.

We showed that the majority of peptides detected by MAE can also be identified by MHC-IAC, further verifying the reliability of the results obtained by our optimized MAE workflow. The strong overlap between MAE and MHC-IAC data boosts the general confidence in our current picture of the MHC immunopeptidome, and it also confirms that standard MHC peptide binding algorithms generally perform rather well, even when they are fully built on MHC-IAC data as Syfpeithi.[30] Indeed, the overlap in MAE versus MHC-IAC data was higher than that in the comparison of MHC-IAC with versus MHC-IAC without acetonitrile-based ultrafilter rinse (Supplementary Figure S17), and it was very similar to the reported overlaps for MHC-IACs that merely differed in the separation of peptides from MHC proteins and β_2m after elution from the MHC-IAC matrix (Figure 2E in Nicastri et al.[8]). The remaining imperfectness in the overlaps implies that the parallel use of different isolation methods can still boost the number of identified peptides strongly (Figure 1A, Supplementary Figure S2A).[52]

The moderate differences in the MHC allotype distribution of peptides obtained by MAE and MHC-IAC are in agreement with the hypothesis that the antibodies applied in MHC-IAC might possibly introduce a bias toward certain MHC allotypes.[9] However, we demonstrated that antibody independent variations in the MHC-IAC workflow, namely, the manner of performing ultrafiltration, can shift the MHC allotype distribution of observed peptides similarly strong as MHC-IAC does in comparison to MAE. Therefore, the relative contribution of W6/32 antibody bias versus sample processing bias is not deducible from our study, but at least the overall bias was not huge with regard to the MHC allotypes investigated here. Additional studies will be needed to pinpoint the potential preference of individual MHC-IAC antibodies toward certain MHC allotypes.

Our data show that, in the MHC-IAC workflow, rinsing of the ultrafilter with an acetonitrile-rich solution can boost peptide yields and will thereby reduce biases (e.g., against hydrophobic peptides) associated with this method. However, the magnitude of the rinsing benefit can be assumed to strongly depend on the amount of the sample as well as on the size and type of the ultrafilter applied. In agreement with our observation that most peptides can stick to the ultrafiltration device under aqueous conditions, Ritz et al.[63] reported that the separation of MHC heavy chain, β_2m, and peptides by reversed-phase LC[4,22,62,65] yielded more peptides than an ultrafiltration without organic solvent. Nevertheless, even if it was possible to eliminate all biases specific for the MAE or the MHC-IAC workflows, important general biases will remain as, for example, the bias introduced by C_{18} desalting,[8] the loss of very hydrophobic peptides on plastic equipment, and the more prominent detection of peptides with preferable MS ionization properties.

Our study provides important new insights regarding the detection and biochemistry of cysteine-containing MHC peptide ligands. Cysteinylation is of critical immunological importance because T cells can discriminate cysteinylated from

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

unmodified cysteine residues.[53] Chen et al. found that CD8+ T cells induced against the non-cysteinylated form of an influenza virus peptide were about 10,000-fold less sensitive to the cysteinylated form of the same peptide.[69] T cell assays investigating cysteine-containing MHC peptide ligands can be confounded by cysteinylation of the peptides as a consequence of free cysteine or its reduced dimer cystine occurring in the cell culture medium.[53,69] We are not aware of studies investigating intracellular cysteinylation of MHC peptide ligands. Nevertheless, there is compelling evidence for the immunological relevance of cysteinylation of MHC-I peptide ligands *in vivo*; Chen et al. demonstrated CD8+ T cells induced by influenza virus infection of mice that were about 10 times more sensitive to the cysteinylated as compared to the unmodified form of an influenza virus peptide, suggesting that these T cells encountered the cysteinylated peptide form *in vivo*.[69] Meadow et al. hypothesized that *in vivo* extracellular cysteinylation of MHC peptide ligands might occur as a result of free cysteine occurring in the blood serum,[53] and in agreement with this assumption, we identified cysteinylated peptides in urine.[70]

Our results suggest the following four measures to obtain immunopeptidomic data sets that suffer less from underrepresentation of cysteine residues:[27,34] (1) Cysteinylation should be included as a variable modification in LC-MS database searches.[24,27,53] (2) MAE should more often be considered as an alternative to MHC-IAC because it can recover up to about 5 times more cysteinylated MHC peptide ligands than MHC-IAC. (3) A cysteine protecting reagent like iodoacetamide should be applied during MAE and MHC-IAC to support the identification of MHC peptide ligands with previously unmodified cysteine residues, and if iodoacetamide is used, (4) carbamidomethylation of cysteine should be added as a modification during bioinformatic processing of LC-MS data. The latter two measures are standard in the field of proteomics. However, a cysteine protecting reagent has not been used in most MHC immunopeptidomics studies until now, and the appropriate consideration of carbamidomethylation in processing of resulting LC-MS data has usually been ignored for MHC peptide ligands.

The much more frequent observation of cysteinylated MHC peptide ligands in MAE versus MHC-IAC supports our hypothesis that MAE preserves post-translational modifications of MHC peptide ligands better than MHC-IAC. However, it remains to be studied for more modifications whether MAE is indeed superior to MHC-IAC regarding most or even all post-translational modifications. Currently, this hypothesis is mainly built on theoretical considerations. MAE rapidly separates the peptides of interest from the cells and thereby minimizes the spatial possibility and time during which enzymes or other molecular components of cells can change the originally presented ligands. In contrast, MHC-IAC first produces a cell lysate where MHC peptide complexes float around for hours providing plenty of physical contact and time for non-native alterations by intracellular molecules, including possible removal of post-translational modifications from MHC peptide ligands. This reasoning is in line with the fact that the much lower frequency of cysteinylated peptides in MHC immunopurification as compared to MAE was less pronounced in the reprocessed data from Lanoix et al.[18] (Supplementary Figure S8) as compared to our study (Table 1). Note that MHC peptide ligands were exposed to the cell lysate overnight in our MHC-IAC study, whereas Lanoix et al. reduced this time to about 5 h.

Of course, MHC-IAC remains the only option to characterize post-translationally modified MHC peptide ligands in solid tissues. This can still be expected to provide very valuable information if appropriate measures are taken, for example, the application of phosphatase inhibitors when studying phosphorylation[54] or the addition of iodoacetamide for the characterization of cysteine modifications (Supplementary Figure S11).

The cysteinylation pattern observed in MAE can be assumed to be the native pattern that is seen by T cells in *in vitro* assays like ELISPOTs. Four reasons substantiate this notion: (1) MAE, in contrast to MHC-IAC, avoids the recovery of MHC peptide ligands from intracellular compartments that are inherently invisible to T cells. (2) Neither the PBS used for washing of cells nor the MAE buffer contain cysteine, cysteine donors, or modifying enzymes. (3) Cells are resuspended in the cold MAE elution buffer for just 1 min, and the eluted MHC peptide ligands are immediately separated from the cells by repeated centrifugations, leaving little time for potential modifications caused by MAE-stressed cells. In contrast to MHC-IAC, the MHC peptide ligands are not exposed to the reducing components of the cytosol that might result in decysteinylation. (4) Free cysteine potentially released from rare burst cells during the MAE elution step is hampered to form new disulfide bonds due to the acidic pH of 3.3 in the MAE elution buffer and due to the presence of inactivating iodoacetamide.

We exemplarily measured the frequencies of cystine residues at the individual positions of HLA-A*02:01 and HLA-B*07:02 peptide ligands and demonstrated MHC-allotype-specific as well as positional differences. Characterization of the overall and positional cysteinylation frequencies for a larger range of MHC allotypes as well as the less biased recovery of unmodified (carbamidomethylated) cysteines in MAE should provide valuable information to be included in MHC binding prediction algorithms to assign cysteine-containing peptides more confidently to certain MHC allotypes and to better discriminate MHC binders from MHC non-binders.

Interestingly, the occurrence of originally unmodified (detected as carbamidomethylated) versus cysteinylated cysteine residues was not consistently complementary to each other. For example, at the highly exposed position 4 of HLA-A*02:01 peptide ligands, our MAE data revealed almost no unmodified (carbamidomethylated) cysteine residues and only a very low frequency of cysteinylation (Figure 5). Unmodified and cysteinylated cysteine residues both share a common hydrophobic (part of the) side chain, and both might therefore be underrepresented at this position along with other hydrophobic amino acid residues (Supplementary Figure S12F). However, it is also conceivable that potentially present cysteine residues are modified differently, possibly in many different substoichiometric forms and thereby escape detection. If, e.g., a plethora of diverse very short cysteine-containing peptides would form a disulfide bond with the exposed cysteine residues at position 4 of HLA-A*02:01 peptide ligands, the latter would go undetected in the current immunopeptidomic workflows.

In summary, we developed, described, and compared strongly improved protocols for both MHC-IAC involving ultrafiltration and MAE. Besides the well established advantages of MAE, we demonstrated four new major advantages for our optimized MAE workflow: (1) Peptide numbers obtained by MAE were almost as high as those in MHC-IAC. (2) Peptide identities and their MS1 intensities were similar between MAE and MHC-IAC. (3) The percentage of putatively MHC binding peptides was only slightly lower in MAE as compared to MHC-IAC and

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

by far outperformed almost all previous MAE immunopeptidomic studies. (4) The recovery of cysteinylated MHC peptide ligands was much better in MAE as compared to MHC-IAC, reducing the bias against cysteine-containing peptides inherent to immunopeptidomic studies.

It will be interesting to characterize additional post-translational modifications of MHC peptide ligands by MAE and to compare the results with current MHC-IAC data. MAE might also be particularly valuable for studying the kinetics of MHC presentation, for example, after viral infection; MAE, in contrast to MHC-IAC, only recovers MHC peptide ligands from the cell surface excluding those still residing in the Golgi apparatus or reinternalized into endosomes. For sure, the costly MHC-IAC will remain a crucial method in immunopeptidomic laboratories, and it is unrivaled for solid tissues and frozen cells. However, we also expect that the optimized, cheap, and reproducible MAE workflow that generates high-quality peptide samples will encourage many laboratories to use it, and we expect that it will contribute to an easy and reliable characterization of MHC peptide ligands.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00386.

Supplementary Materials and Methods; Supplementary Figure S1, overlap and MS1 intensities MHC-IAC versus MAE, THP-1; Supplementary Figure S2, reproducibility of MAE; Supplementary Figure S3, reproducibility of MHC-IAC; Supplementary Figure S4, MHC-II peptides do not contribute to the observed MAE peptides; Supplementary Figure S5, peptide lengths; Supplementary Figure S6, MHC binding affinities in MAE and MHC-IAC, THP-1; Supplementary Figure S7, percentage cystine peptides MAE and MHC-IP in Lanoix et al.; Supplementary Figure S8, robustness cysteinylation pattern HLA-A2; Supplementary Figure S9, robustness cysteinylation pattern HLA-B7; Supplementary Figure S10, cysteinylation and carbamidomethylation pattern from Bassani-Sternberg data; Supplementary Figure S11, structural parameters HLA-A2; Supplementary Figure S12, structural parameters HLA-B7; Supplementary Figure S13, structure position C-4; Supplementary Figure S14, DMSO in MAE; Supplementary Figure S15, acetonitrile in MHC-IAC; Supplementary Figure S16, rinsing ultrafilter improves peptide yield, most for hydrophobic peptides; Supplementary Figure S17, rinsing ultrafilter shifts the MHC allotype distribution to hydrophobic; Supplementary Figure S18, observed and predicted hydrophobicity of different MHC allotype ligands; Supplementary Table S1, percentage of cysteinylated and unmodified cysteine in MHC-IAC; Supplementary Table S2, total peptide counts Lanoix et al.; Supplementary Table S3, *P* values Cys modifications HLA-A2; Supplementary Table S4, *P* values Cys modifications HLA-B7; Supplementary Table S5, PDB crystal structures used for the analyses; Supplementary Table S6, statistics structural PDB parameters HLA-A*02; Supplementary Table S7, statistics structural PDB parameters HLA-B*07; Supplementary Table S8, RSAS at position C-4; Supplementary Table S9, MAE setups and raw files; Supplementary Table S10, MHC-IAC setups and raw files; Supplementary references (PDF)

Supplementary Table S11 showing lists of peptides (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

Theo Sturm − *Institute of Molecular Systems Biology, Department of Biology, ETH Zürich, 8093 Zürich, Switzerland; Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands; Netherlands Proteomics Centre, 3584 CH Utrecht, The Netherlands; Philochem AG, 8112 Otelfingen, Switzerland;* orcid.org/0000-0002-7994-2599; Email: Theo.Sturm@philochem.ch

▽**Ruedi Aebersold** − *Institute of Molecular Systems Biology, Department of Biology, ETH Zürich, 8093 Zürich, Switzerland; Faculty of Science, University of Zurich, 8057 Zürich, Switzerland*; Email: aebersold@imsb.biol.ethz.ch

Authors

Benedikt Sautter − *Department of Immunology, Institute for Cell Biology, University of Tübingen, 72076 Tübingen, Germany*

Tobias P. Wörner − *Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands; Netherlands Proteomics Centre, 3584 CH Utrecht, The Netherlands*

Stefan Stevanović − *Department of Immunology, Institute for Cell Biology, University of Tübingen, 72076 Tübingen, Germany*

Hans-Georg Rammensee − *Department of Immunology, Institute for Cell Biology, University of Tübingen, 72076 Tübingen, Germany*

Oliver Planz − *Department of Immunology, Institute for Cell Biology, University of Tübingen, 72076 Tübingen, Germany*

▽**Albert J. R. Heck** − *Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands; Netherlands Proteomics Centre, 3584 CH Utrecht, The Netherlands;* orcid.org/0000-0002-2405-4404

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00386

Author Contributions

T.S., A.J.R.H., and R.A. conceived the project. All authors designed the experiments. T.S. performed the MHC-IAC. T.S. and B.S. performed the MAE. T.S. performed the LC-MS analyses. T.S. and T.P.W. performed the *in silico* structural analyses related to cysteinylation. T.S. analyzed the data. T.S. wrote the manuscript with contributions from A.J.R.H. and R.A.

Notes

The authors declare the following competing financial interest(s): T.S. has been an employee of Philochem AG since July 1, 2018 (LC-MS measurements had been completed before). H.-G.R. is shareholder of Immatics Biotechnologies GmbH and Curevac AG. O.P. is a shareholder and consultant of Atriva Therapeutics GmbH.

▽A.J.R.H., R.A.: These authors share senior authorship.

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

The LC-MS raw files and the associated Proteome Discoverer msf-files of this study are publicly available at the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[71] with the data set identifiers PXD012771 for Orbitrap XL data and PXD012437 as well as PXD012498 for Orbitrap Fusion data. Lists of identified peptides are provided as an Excel file (Supplementary Table S11) for all samples which are considered in Figures 1−5 or in Table 1.

■ ACKNOWLEDGMENTS

We would like to thank the group of Pierre Thibault and Etienne Caron for sharing a basic protocol for MAE from 2008/2014 which was the basis for our optimizations. We are grateful for discussions with Heiko Schuster, Daniel J. Kowalewski, Moreno Di Marco, Etienne Caron, Henk van den Toorn, and Juliane Schneider. Linus Backert and Henk van den Toorn provided valuable bioinformatic support. Ana Marcu provided technical assistance, and Claudia Falkenburger produced the antibodies used in MHC-IAC. Heike Friedrich cultured LCL5 cells. T.S., T.P.W., and A.J.R.H. were funded by the large-scale proteomics facility Proteins@Work (Project 184.032.201) embedded in the Netherlands Proteomics Centre and supported by the Netherlands Organization for Scientific Research (NWO). A.J.R.H. acknowledges support by the Spinoza Prize of NWO (Project SPI.2017.028) and by the European Union Horizon 2020 program FET-OPEN project MSmed, Project 686547. R.A. acknowledges the following grant support: TBVAC2020 (2-73838-14), ERC grant Proteomics v3.0 (ERC-2008-AdG_20080422), ERC Proteomics 4D (670821), and the Swiss National Science Foundation (3100A0-688 107679).

■ REFERENCES

(1) Walter, S.; et al. Multipeptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. *Nat. Med.* **2012**, *18*, 1254−1261.
(2) Liddy, N.; et al. Monoclonal TCR-redirected tumor cell killing. *Nat. Med.* **2012**, *18*, 980−987.
(3) Gubin, M. M.; et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* **2014**, *515*, 577−581.
(4) Falk, K.; Rötzschke, O.; Stevanović, S.; Jung, G.; Rammensee, H. G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **1991**, *351*, 290−296.
(5) Deres, K.; Beck, W.; Faath, S.; Jung, G.; Rammensee, H. G. MHC/peptide binding studies indicate hierarchy of anchor residues. *Cell. Immunol.* **1993**, *151*, 158−167.
(6) Neefjes, J.; Jongsma, M. L.; Paul, P.; Bakke, O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **2011**, *11*, 823−836.
(7) Demmers, L. C.; Heck, A. J. R.; Wu, W. Pre-fractionation Extends but also Creates a Bias in the Detectable HLA Class Iota Ligandome. *J. Proteome Res.* **2019**, *18*, 1634−1643.
(8) Nicastri, A.; Liao, H.; Muller, J.; Purcell, A. W.; Ternette, N. The Choice of HLA-Associated Peptide Enrichment and Purification Strategy Affects Peptide Yields and Creates a Bias in Detected Sequence Repertoire. *Proteomics* **2020**, *20*, 1900401.
(9) Schumacher, F. R.; et al. Building proteomic tool boxes to monitor MHC class I and class II peptides. *Proteomics* **2017**, *17*, 1600061.
(10) Caron, E.; et al. Analysis of Major Histocompatibility Complex (MHC) Immunopeptidomes Using Mass Spectrometry. *Mol. Cell. Proteomics* **2015**, *14*, 3105−3117.
(11) Sugawara, S.; Abo, T.; Kumagai, K. A simple method to eliminate the antigenicity of surface class I MHC molecules from the membrane of viable cells by acid treatment at pH 3. *J. Immunol. Methods* **1987**, *100*, 83−90.
(12) Lee, J. M.; Watts, T. H. On the dissociation and reassociation of MHC class II-foreign peptide complexes. Evidence that brief transit through an acidic compartment is not sufficient for binding site regeneration. *J. Immunol* **1990**, *144*, 1829−1834.
(13) Storkus, W. J.; Zeh, H. J., 3rd; Salter, R. D.; Lotze, M. T. Identification of T-cell epitopes: rapid isolation of class I-presented peptides from viable cells by mild acid elution. *J. Immunother.* **1993**, *14*, 94−103.
(14) Gebreselassie, D.; Spiegel, H.; Vukmanovic, S. Sampling of major histocompatibility complex class I-associated peptidome suggests relatively looser global association of HLA-B*5101 with peptides. *Hum. Immunol.* **2006**, *67*, 894−906.
(15) Antwi, K.; et al. Proteomic identification of an MHC-binding peptidome from pancreas and breast cancer cell lines. *Mol. Immunol.* **2009**, *46*, 2931−2937.
(16) Fortier, M. H.; et al. The MHC class I peptide repertoire is molded by the transcriptome. *J. Exp. Med.* **2008**, *205*, 595−610.
(17) Pearson, H.; et al. MHC class I-associated peptides derive from selective regions of the human genome. *J. Clin. Invest.* **2016**, *126*, 4690−4701.
(18) Lanoix, J.; et al. Comparison of the MHC I Immunopeptidome Repertoire of B-Cell Lymphoblasts Using Two Isolation Methods. *Proteomics* **2018**, *18*, 1700251.
(19) Hunt, D. F.; et al. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **1992**, *255*, 1261−1263.
(20) Neidert, M. C.; et al. Natural HLA class I ligands from glioblastoma: extending the options for immunotherapy. *J. Neuro-Oncol.* **2013**, *111*, 285−294.
(21) Giam, K.; et al. A comprehensive analysis of peptides presented by HLA-A1. *Tissue Antigens* **2015**, *85*, 492−496.
(22) Bassani-Sternberg, M.; Pletscher-Frankild, S.; Jensen, L. J.; Mann, M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell. Proteomics* **2015**, *14*, 658−673.
(23) Mommen, G. P.; et al. Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD). *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 4507−4512.
(24) Hassan, C.; et al. The human leukocyte antigen-presented ligandome of B lymphocytes. *Mol. Cell. Proteomics* **2013**, *12*, 1829−1843.
(25) Bassani-Sternberg, M.; et al. Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18769−18776.
(26) Barnea, E.; et al. The Human Leukocyte Antigen (HLA)-B27 Peptidome in Vivo, in Spondyloarthritis-susceptible HLA-B27 Transgenic Rats and the Effect of Erap1 Deletion. *Mol. Cell. Proteomics* **2017**, *16*, 642−662.
(27) Abelin, J. G.; et al. Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. *Immunity* **2017**, *46*, 315−326.
(28) Khodadoust, M. S.; et al. Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature* **2017**, *543*, 723−727.
(29) Ternette, N.; et al. Defining the HLA class I-associated viral antigen repertoire from HIV-1-infected human cells. *Eur. J. Immunol.* **2016**, *46*, 60−69.
(30) Rammensee, H.; Bachmann, J.; Emmerich, N. P.; Bachor, O. A.; Stevanović, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **1999**, *50*, 213−219.
(31) Lundegaard, C.; et al. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8−11. *Nucleic Acids Res.* **2008**, *36*, W509−512.
(32) Andreatta, M.; Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **2016**, *32*, 511−517.
(33) Nielsen, M.; et al. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One* **2007**, *2*, e796.

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304

(34) Bassani-Sternberg, M.; et al. Deciphering HLA-I motifs across HLA peptidomes improves neo-antigen predictions and identifies allostery regulating HLA specificity. *PLoS Comput. Biol.* **2017**, *13*, e1005725.
(35) O'Donnell, T. J.; et al. MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst* **2018**, *7*, 129−132.e4.
(36) Jurtz, V.; et al. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J. Immunol.* **2017**, *199*, 3360−3368.
(37) Hassan, C.; et al. Accurate quantitation of MHC-bound peptides by application of isotopically labeled peptide MHC complexes. *J. Proteomics* **2014**, *109*, 240−244.
(38) Tsuchiya, S.; et al. Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int. J. Cancer* **1980**, *26*, 171−176.
(39) Kyzirakos, C. Häufig erkannte EBV-spezifische CD4+ T-Zellepitope für eine optimierte Immuntherapie bei EBV-assoziierten Erkrankungen. Ph.D. Dissertation, University of Tübingen, Tübingen, Germany, 2013.
(40) Saini, S. K.; et al. Dipeptides catalyze rapid peptide exchange on MHC class I molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 202−207.
(41) Barnstable, C. J.; et al. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. *Cell* **1978**, *14*, 9−20.
(42) Lampson, L. A.; Levy, R. Two populations of Ia-like molecules on a human B cell line. *J. Immunol.* **1980**, *125*, 293−299.
(43) Pawelec, G. P.; Shaw, S.; Ziegler, A.; Müller, C.; Wernet, P. Differential inhibition of HLA-D- or SB-directed secondary lymphoproliferative responses with monoclonal antibodies detecting human Ia-like determinants. *J. Immunol.* **1982**, *129*, 1070−1075.
(44) Olsen, J. V.; et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **2005**, *4*, 2010−2021.
(45) Eng, J. K.; McCormack, A. L.; Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.
(46) Käll, L.; Canterbury, J. D.; Weston, J.; Noble, W. S.; MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **2007**, *4*, 923−925.
(47) Krokhin, O. V.; Spicer, V. Predicting peptide retention times for proteomics. *Curr. Protoc. Bioinf.* **2010**, *31*, 13.14.1−13.14.15.
(48) Nelde, A.; Kowalewski, D. J.; Stevanovic, S. Purification and Identification of Naturally Presented MHC Class I and II Ligands. *Methods Mol. Biol.* **2019**, *1988*, 123−136.
(49) Dudek, N. L.; Croft, N. P.; Schittenhelm, R. B.; Ramarathinam, S. H.; Purcell, A. W. A Systems Approach to Understand Antigen Presentation and the Immune Response. *Methods Mol. Biol.* **2016**, *1394*, 189−209.
(50) Chong, C.; et al. High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferongamma-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Mol. Cell. Proteomics* **2018**, *17*, 533−548.
(51) Bassani-Sternberg, M. Mass Spectrometry Based Immunopeptidomics for the Discovery of Cancer Neoantigens. *Methods Mol. Biol.* **2018**, *1719*, 209−221.
(52) Pandey, K.; et al. In-depth mining of the immunopeptidome of an acute myeloid leukemia cell line using complementary ligand enrichment and data acquisition strategies. *Mol. Immunol.* **2020**, *123*, 7−17.
(53) Meadows, L.; et al. The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. *Immunity* **1997**, *6*, 273−281.
(54) Abelin, J. G.; et al. Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. *Nat. Protoc.* **2015**, *10*, 1308−1318.
(55) Bergmann, T.; et al. The common equine class I molecule Eqca-1*00101 (ELA-A3.1) is characterized by narrow peptide binding and T cell epitope repertoires. *Immunogenetics* **2015**, *67*, 675−689.
(56) Di Marco, M.; et al. Unveiling the Peptide Motifs of HLA-C and HLA-G from Naturally Presented Peptides and Generation of Binding Prediction Matrices. *J. Immunol.* **2017**, *199*, 2639−2651.
(57) Olsson, N.; et al. T-Cell Immunopeptidomes Reveal Cell Subtype Surface Markers Derived From Intracellular Proteins. *Proteomics* **2018**, *18*, 1700410.
(58) Trujillo, J. A.; et al. The cellular redox environment alters antigen presentation. *J. Biol. Chem.* **2014**, *289*, 27979−27991.
(59) Frese, C. K.; et al. Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal. Chem.* **2012**, *84*, 9668−9673.
(60) Shen, S.; et al. Probabilistic analysis of the frequencies of amino acid pairs within characterized protein sequences. *Phys. A* **2006**, *370*, 651−662.
(61) Hamelryck, T. An amino acid has two sides: a new 2D measure provides a different view of solvent exposure. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 38−48.
(62) Croft, N. P.; et al. Kinetics of antigen expression and epitope presentation during virus infection. *PLoS Pathog.* **2013**, *9*, e1003129.
(63) Ritz, D.; et al. High-sensitivity HLA class I peptidome analysis enables a precise definition of peptide motifs and the identification of peptides from cell lines and patients' sera. *Proteomics* **2016**, *16*, 1570−1580.
(64) Rötzschke, O.; et al. Qa-2 molecules are peptide receptors of higher stringency than ordinary class I molecules. *Nature* **1993**, *361*, 642−644.
(65) Milner, E.; et al. The effect of proteasome inhibition on the generation of the human leukocyte antigen (HLA) peptidome. *Mol. Cell. Proteomics* **2013**, *12*, 1853−1864.
(66) Granados, D. P.; et al. Proteogenomic-based discovery of minor histocompatibility antigens with suitable features for immunotherapy of hematologic cancers. *Leukemia* **2016**, *30*, 1344−1354.
(67) Lill, J. R.; et al. Minimal Information About an Immuno-Peptidomics Experiment (MIAIPE) Immuno-Peptidomics Experiment (MIAIPE). *Proteomics* **2018**, *18*, 1800110.
(68) Caron, E.; Aebersold, R.; Banaei-Esfahani, A.; Chong, C.; Bassani-Sternberg, M. A Case for a Human Immuno-Peptidome Project Consortium. *Immunity* **2017**, *47*, 203−208.
(69) Chen, W.; Yewdell, J. W.; Levine, R. L.; Bennink, J. R. Modification of cysteine residues in vitro and in vivo affects the immunogenicity and antigenicity of major histocompatibility complex class I-restricted viral determinants. *J. Exp. Med.* **1999**, *189*, 1757−1764.
(70) Sturm, T.; et al. Mouse urinary peptides provide a molecular basis for genotype discrimination by nasal sensory neurons. *Nat. Commun.* **2013**, *4*, 1616.
(71) Vizcaino, J. A.; et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **2016**, *44*, D447−456.

https://dx.doi.org/10.1021/acs.jproteome.0c00386
J. Proteome Res. 2021, 20, 289−304