Supporting information
for
Mild acid elution and MHC immunoaffinity chromatography reveal similar albeit not identical profiles of the HLA immunopeptidome

Theo Sturm1,2,3,4,*, Benedikt Sautter5, Tobias P. Wörner2,3, Stefan Stevanović5, Hans-Georg Rammensee5, Oliver Planz2, Albert J. R. Heck2,3,=, Ruedi Aebersold1,6,=,*

1 Institute of Molecular Systems Biology, Department of Biology, ETH Zürich, 8093 Zürich, Switzerland
2 Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CH Utrecht, The Netherlands
3 Netherlands Proteomics Centre, 3584 CH Utrecht, The Netherlands
4 Philochem AG, 8112 Otelfingen, Switzerland
5 Department of Immunology, Institute for Cell Biology, University of Tübingen, 72076 Tübingen, Germany
6 Faculty of Science, University of Zurich, 8057 Zürich, Switzerland
* Corresponding authors
= These authors share senior authorship.
Table of contents

Supplementary Materials and Methods

Supplementary Figure S1 (Schematic MAE versus MHC-IAC workflow)
Supplementary Figure S2 (Overlap and MS1 intensities MHC-IAC versus MAE, THP-1)
Supplementary Figure S3 (Reproducibility of MAE)
Supplementary Figure S4 (Reproducibility of MHC-IAC)
Supplementary Figure S5 (MHC-II peptides do not contribute to the observed MAE peptides)
Supplementary Figure S6 (Peptide lengths)
Supplementary Figure S7 (MHC binding affinities in MAE and MHC-IAC, THP-1)
Supplementary Figure S8 (Percentage cysteine peptides MAE and MHC-IP in Lanoix et al.)
Supplementary Figure S9 (Robustness cysteinylation pattern HLA-A2)
Supplementary Figure S10 (Robustness cysteinylation pattern HLA-B7)
Supplementary Figure S11 (Cysteinylation and carbamidomethylation pattern from Bassani-Sternberg data)
Supplementary Figure S12 (Structural parameters HLA-A2)
Supplementary Figure S13 (Structural parameters HLA-B7)
Supplementary Figure S14 (Structure position C-4)
Supplementary Figure S15 (DMSO in MAE)
Supplementary Figure S16 (Acetonitrile in MHC-IAC)
Supplementary Figure S17 (Rinsing ultrafilter improves peptide yield, most for hydrophobic peptides)
Supplementary Figure S18 (Rinsing ultrafilter shifts the MHC allotype distribution to hydrophobic)
Supplementary Figure S19 (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 Table S11 (Peptide lists), see separate Excel file

Supplementary References
Supplementary Materials and Methods

Cell culture

To minimize biological variability, all cell culture was performed in the Rammensee laboratory and was maintained for at most 3 months after thawing. All cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin and non-essential amino acids (Biochrom, cat.-no. K0293, used as recommended by manufacturer). Furthermore, we added 1 mM sodium pyruvate to the medium for JY and LCL5 cells. For JY and THP-1 cells, the basic RPMI 1640 medium was bought as solution containing 25 mM HEPES and 2 mM L-glutamine from Life Technologies (cat.-no. 52400-041), and we refer to the derived complete medium as “RPMI ready”. For LCL5 cells, we used the “RPMI powder-based” medium instead. RPMI 1640 powder (Biochrom cat.-no. T121-10) was dissolved in 10 L water and supplemented with our additives described above. Additionally, we adjusted the “RPMI powder-based” medium to 2 g/L NaHCO₃, 10 mM HEPES, 4.8 mM L-glutamine and 50 µM 2-mercaptoethanol.

For the first passages of freshly thawed cells, we used plastic cell culture flasks with a filter-containing lid. However, all cell cultures for the immunopeptidomic analyses were finally performed in airtight 2 L glass bottles filled with a maximum of 1.2 L cell culture medium each. The glass bottles were rotated at a speed of 0.75 rpm. JY cells were fed every three to four days mixing their culture 1/5 to 1/10 with fresh medium. THP-1 cells were supplied with fresh medium every two days and LCL5 cells every two to three days diluting their culture 1/2 (50% old culture and 50% fresh medium).

For some MAE samples of a separate study (Sturm et al., manuscript in preparation) also used for Supplementary Figure S9C and Supplementary Figure S10B, 5 ∙ 10⁸ JY cells each were infected with influenza virus A/Regensburg/D6/2009(H1N1) at a multiplicity of infection (MOI) = 2. This was achieved by first washing cells once with PBS and twice with infection PBS. Infection PBS consisted of PBS supplemented with 0.9 mM CaCl₂, 0.5 mM MgCl₂, 0.2% bovine serum albumin, 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were then resuspended in infection PBS, mixed with the virions by pipetting up and down 5 times and incubated at 37 °C for 1 h. After centrifugation at 235 g for 10 min at room temperature, supernatant was discarded and cells were resuspended in RPMI 1640 medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin and non-essential amino acids. After incubation in airtight 2 L bottles at 37 °C for 12 h and 16 h, respectively, cells were subjected to MAE. Using intracellular staining against influenza A virus nucleoprotein and flow cytometry, the infection rate was determined to be 38.4% and 14.3% for the first and second biological replicate, respectively.

Structural analyses related to the inhomogeneous distribution of cysteinylation and carbamidomethylation across different peptide positions

To associate the positional distribution of carbamidomethylation and cysteinylation as well as the MHC allotype dependence of the latter with the structural features of the MHC binding groove, we strived to obtain non-redundant, high quality crystal structures that are representative of the majority of MHC peptide ligands in our LC-MS data. Therefore, we selected all crystal structure entries for UniProtKB accession numbers P01892 (HLA-A*02:01) and P01889 (HLA-B*07:02) that were available in the Protein Data Bank (PDB) in June 2017 and matched the following ten criteria.
1.) The MHC/β2m/peptide crystal structure does not contain a co-crystallized T cell receptor.  
*Note: A bound T cell receptor would shield some parts of the peptide ligand otherwise exposed.*

2.) The peptide ligand involved does not contain any post-translational modification or non-standard amino acid.  
*Note: Crystal structures of MHC peptide ligands containing cysteinylated cysteine or oxidized methionine were not available, and other post-translational modifications were not considered in our LC-MS analyses.*

3.) The peptide ligand involved has a NetMHC IC50 <500 nM.  
*Note: We also used the NetMHC IC50 cutoff <500 nM for filtering of putative MHC peptide ligands in our LC-MS data. Applying the same filter for peptides in crystal structures supports the aim of obtaining structural information reflecting the situation in our LC-MS data by excluding structures with unusual binding characteristics or putatively very weak binding peptide ligands.*

4.) If there are crystal structures with similar peptide ligands differing merely by a single amino acid variation (SAV), only use one of these structures. If several of these structures fulfill all other criteria, use the one with the peptide most similar to the natural standard sequence.  
*Note: MHC peptide ligands differing only by a SAV usually have very similar orientation in the MHC binding groove. Including all related crystal structures from SAV peptides might introduce a bias towards certain partly redundant structural information.*

5.) The peptide ligand involved does not overhang from the MHC binding groove at its N- or C-terminus.  
*Note: The occurrence of overhanging MHC-I peptide ligands is well proven, but it is rather an exception and not representative of the bulk of the MHC-I immunopeptidome.*

6.) The crystal structure does not only include MHC and β2m but also a peptide ligand.

7.) The amino acid sequence of the MHC heavy chain corresponds to the native sequence of HLA-A*02:01 or HLA-B*07:02 within the MHC binding groove.

8.) If several PDB entries are available for the same MHC peptide ligand, only use the one with the higher resolution.

9.) The resolution of the structure is ≤ 300 pm.

10.) The crystal structure is directly associated with an article in a (peer reviewed) scientific journal via a PubMed link in PDB.

From the selected PDB entries (Supplementary Table S5), we calculated the coordination number (CN), the Half-Sphere Exposure (HSE)-β-up and the solvent accessible surface (SAS) for every single amino acid of the MHC peptide ligands leveraging the DSSP and HSExposure modules of Biopython’s PDB package (Biopython 1.69; DSSP 2.0.4, http://swift.cmbi.ru.nl/gv/dssp/). For the CN and the HSE-β-up, we employed sphere radii of 1.2 nm. Therefore, the CN represents the number of surrounding Ca atoms within 1.2 nm of the Ca atom of the peptide position of interest. Accordingly, the HSE-β-up counts the number of surrounding Ca atoms within 1.2 nm of the Ca atom of the peptide position of interest using only the half sphere defined by the Ca-Cβ axis, i.e. the direction of the side chain. For SAS, the default solvent probe radius of 140 pm was employed. Downstream data processing and calculations were performed with computer scripts which we wrote in R. Relative SAS (RSAS) was calculated by dividing the SAS for each peptide position by the maximal empirical SAS observed for the respective amino acid by Tien et al. Single PDB entries often contained more than one crystal structure of one and the same MHC/β2m/peptide complex reflecting structural variability. Hence, we calculated the median values for each structural parameter across all crystal structures of a given PDB entry. Finally, means and standard deviations were calculated across all selected PDB entries of the respective MHC allotype (Supplementary Table S5) using the previously obtained median values.
PDB-derived structures were visualized using the software PyMOL from Schrödinger. To depict representative peptide backbone and Cβ orientation (Supplementary Figure S12A and Supplementary Figure S13A), displayed PDB crystal structures were first aligned to that depicted crystal structure whose peptide ligand HSE-β-up values correlated best with the means across all selected 33 HLA-A*02:01 and four HLA-B*07:02 PDB entries respectively (Supplementary Table S5). Hence, for HLA-A*02:01, alignment was performed to the PDB crystal structure no. “1i7r” chains D-F, and for HLA-B*07:02, we aligned to “4u1h” chains A-C. The depicted MHC protein structure in Supplementary Figure S12A correspondingly represents PDB “1i7r” chain D, and the one in Supplementary Figure S13A is PDB “4u1h” chain A.

To determine the width of the MHC peptide binding groove in the vicinity of peptide position C-4, we calculated the distance between the Arg65 Cζ atom and the Gln155 Cδ atom of the HLA-A*02:01 protein as well as the distance between the Arg62 Cζ atom and the Gln155 Cδ atom of the HLA-B*07:02 protein (see Supplementary Figure S14A and B) harnessing Biopython’s PDB package and PyMOL. Some of the used crystal structures contained more than one position for these atoms. In this case, we calculated the distance for all positions separately and then used the mean as the distance for this crystal structure. If one PDB entry contained several crystal structures for the same MHC peptide ligand, we used the median distance from the mean crystal structure distance as the Arg65/62 Cζ - Gln155 Cδ distance for this PDB entry. Finally, we calculated the mean (±standard deviation) distance across the 33 selected HLA-A*02:01 PDB entries and the four selected HLA-B*07:02 PDB entries (compare Supplementary Table S5).
Peptides and proteins from cell culture medium

**MAE**

PBS washing 3x

elution
pH = 3.3, 1 min

centrifuge 4x
300 g, 5 min
350 g, 10 min
3400 g, 15 min
257 000 g, 1 h

Supernatant

Oasis HLB desalting

ultrafiltration

C\textsubscript{18} desalting

LC-MS

**MHC-IAC**

lysis

cell lysate

immunocapture

washing

elution 0.2% TFA

β\textsubscript{2}m

MHC peptide ligands

MHC heavy chain

Supplementary Figure S1 | Schematic of MHC peptide ligand isolation by MAE and MHC-IAC. The major steps of our optimized standard MAE (left panel) and our optimized standard MHC-IAC (right panel) workflow are outlined.
Supplementary Figure S2 | The observations presented in Figure 1 for JY cells can be reproduced using THP-1 cells — MAE and MHC-IAC identify a big overlapping set of peptides with remarkably correlated MS1 intensities. 1.25 · 10⁶ THP-1 cells each for MAE and MHC-IAC were harvested from the same cell culture. A, Most of the peptides obtained by MAE are also obtained by MHC-IAC. Almost half (455 out of 1016) of MHC-IAC peptides were reproduced with MAE. 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 IC₅₀ ≥ 500 nM) are highlighted with red circles in the dot plot.
**Supplementary Figure S3 | MAE shows good reproducibility.** Fluctuations in LC-MS performance appear to be responsible for more than the half of the variability observed in biological replicates. Hence, the contribution of fluctuations in the MAE sample preparation workflow seems to be less important (also compare Supplementary Figure S15). Absolute peptide numbers are indicated in italics. A and B, Reproducibility of biological replicates. In different months, four MAEs were performed and measured from 5 · 10^8 JY cells each. A, To calculate the overlap of identified peptides, each of the four biological replicates was compared against the other three biological replicates resulting in six comparisons. The means (± standard deviation) of the overlap and of the peptides exclusive for each biological replicate are indicated. B, D and F, MHC allotype distributions of identified peptides. Each peptide was assigned to that MHC allotype of the respective cell line yielding the highest NetMHC binding affinity. Peptides with a NetMHC IC\textsubscript{50} ≥ 500 nM were considered as non-binders. C and D, Reproducibility of MAEs using different cell numbers. The three MAEs were performed in parallel from the same THP-1 cell culture suspension applying 1.25 · 10^8 cells for MAE 1, 5 · 10^8 cells for MAE 2 and 1.5 · 10^9 cells for MAE 3. In MAE 3, only one third of the Oasis HLB eluate was further processed whereas the total Oasis HLB eluate was applied in the case of the other MAEs. Therefore, the number of cells corresponding to each LC-MS run is the same for MAE 2 and MAE 3. E and F, Reproducibility of LC-MS replicates. One MAE was performed from 5 · 10^8 JY cells (exceptionally, no ultracentrifugation was performed for this sample), and eight LC-MS replicates were measured within less than three days. Three LC-MS replicates were compared against three other ones mimicking the standard nested LC-MS triplicate data acquisition scheme routinely applied for comparisons of peptide extraction and purification workflows.
Supplementary Figure S4  |  MHC-IAC shows good reproducibility. 7.5 ml LCL5 cell pellet was split into three aliquots (approximately 1.25 · 10⁹ cells each) that were subjected to three parallel MHC-IACs. The eluate from the MHC-IAC column was passed through an ultrafilter yielding the first ultrafiltrate (UF). Then, 1 ml of 80 % acetonitrile / 0.2 % TFA was passed through the ultrafilter resulting in the rinse UF. For MHC-IAC 1 and MHC-IAC 2, the first UF and the rinse UF were processed separately after ultrafiltration, but the MS raw data of the first UF and the rinse UF were combined for data processing with Proteome Discoverer for Supplementary Figure S4. For MHC-IAC 3, the first UF and the rinse UF were physically combined directly after ultrafiltration. Note that peptide numbers are higher than in Supplementary Figure S3. This possibly increased reproducibility and might explain why the overlap of MHC-IAC replicates (A) appears to be higher than the overlap of LC-MS replicates (Supplementary Figure S3E). A, Most peptides identified in one MHC-IAC are also identified in the replicate MHC-IAC. Absolute peptide numbers are indicated in italics. B, The MHC allotype distribution of identified peptides is stable across the three MHC-IACs. MHC allotype assignment was performed as described in Supplementary Figure S3. C, Quantitative reproducibility of MHC-IAC is also high. Mean MS1 intensities of individual peptides are derived from averaging the three LC-MS measurements per MHC-IAC.
Supplementary Figure S5 | MHC-II peptide ligands do not notably contribute to the 8 to 12-mer peptide set identified by MAE. All displayed data are derived from only two different cell culture harvestings — one each for JY and THP-1 cells. Numbers in italics indicate the absolute number of peptides in each bin of the Euler diagram. A, Peptides obtained by MAE and MHC-II-IAC from 5 · 10⁸ JY cells each. For MAE, only peptides not predicted to bind to the relevant MHC-I molecules (NetMHC IC₅₀ ≥ 500 nM and relative Syfpeithi-score < 50 %) are considered. Here, the MHC-II-IAC peptide set was filtered for 8 to 12-mers to apply the same filter used for MAE. B, Same overlap scheme and data filtering as in A, but using peptides obtained from 1.25 · 10⁸ THP-1 cells each. The two overlapping peptides are PEPAKSAPAPK (11-mer from histone H2B) and PRKIEEIKDF (10-mer from 60 S ribosomal protein L38).
Supplementary Figure S6 | Peptides isolated by MAE and MHC-IAC have similar length distributions. All displayed data are derived from only two different cell culture harvestings — one each for JY and THP-1 cells. Note that the m/z range used for MS1 survey scans and MS2 precursor selection was set to 400 – 650 Th to boost peptide identifications. This mass window introduces a bias against peptides shorter than 8 amino acids and against doubly charged peptide precursors longer than 11 amino acids. However, this MS-associated bias is equal for both MAE and MHC-IAC. The green shaded areas indicate the \emph{in silico} length filter of 8 to 12 amino acids that was applied in the present study (except for this figure). \textbf{A}, Peptides obtained by MAE and MHC-IAC from $5 \cdot 10^8$ JY cells each. The summed fraction of 8- to 12-mers among all peptides is 95.6 \% and 99.2 \% for MAE and MHC-IAC respectively. \textbf{B}, Peptides obtained by MAE and MHC-IAC from $1.25 \cdot 10^8$ THP-1 cells each. The summed fraction of 8- to 12-mers among all peptides is 90.2 \% and 98.0 \% for MAE and MHC-IAC respectively.
Supplementary Figure S7 | Putative MHC binding peptides from MAE and MHC-IAC of THP-1 cells show a similar distribution of predicted MHC binding affinities. All displayed data are derived from the same cell culture harvesting using $1.25 \cdot 10^8$ THP-1 cells each for MAE and MHC-IAC. Gray bars represent the median of the NetMHC-IC$_{50}$ and the median of the corresponding peptide MS1 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.
Supplementary Figure S8 | For a given total peptide count, the percentage of cystine peptides among all identified peptides is higher in MAE as compared to MHC immunopurification (MHC-IP) also in the B-LCL data set of Lanoix et al.. MS raw data from Lanoix et al.18 were reprocessed with Proteome Discoverer 1.4 and Sequest HT using cysteinylation as a variable modification. A, In the B-LCL data set of Lanoix et al., the percentage of cystine peptides among all identified peptides increases with increasing total peptide count per experiment. The blue data points represent the three biological replicates of MAEs of 10^8 B-LCLs each. The red data points indicate the MHC-IPs from B-LCLs using 2 ∙ 10^6, 2 ∙ 10^7 or 10^8 cells per experiment. Data from both fresh and frozen B-LCLs were included for MHC-IP as there were no statistically significant differences in total peptide count or percentage of cystine peptides between fresh and frozen cells. The logarithmic regression line interpolates the percentage of cystine peptides as a function of the total peptide count after MHC-IP. B, Comparing MAE and MHC-IP experiments with equal total peptide counts, the data of Lanoix et al. show a 2.6-fold higher percentage of cystine peptides in MAE as compared to MHC-IP (P = 0.03 performing a two-tailed, paired Student’s t-test). Values of the data points from MAE depicted in (A) are printed in blue. The interpolated percentage of cystine peptides in MHC-IP is derived from the logarithmic regression line shown in (A) and is printed in red.
Supplementary Figure S9 | Robustness of the pattern of cysteine modifications observed for HLA-A*02:01 peptide ligands by MAE. The patterns of carbamidomethylation and cysteinyl modification described in Figure 5A for HLA-A*02:01 motif peptides (NetMHC IC_{50} < 500 nM) are supported by additional MAE analyses varying biological and technical parameters of the experimental setup. A few deviations of modification frequencies also occurred, but were most prominent in (B) and (D) where only lower numbers of HLA-A*02:01 motif peptides containing cysteine \((n)\) were identified due to measurements at the Orbitrap XL. These low numbers will enhance random effects when sampling the immunopeptidome, and therefore, this less perfect congruence in modification frequencies should not be considered contradictory. A, Similar modification pattern using a different peptide fragmentation method in MS. One of the MAEs from uninfected JY cells included in Figure 5A was additionally analyzed using HCD instead of EThcD. \(n = 66.\) B, Largely overlapping modification pattern using a different peptide fragmentation method and a different MS instrument with additional biological replicates. Four biological replicates of JY cells, that were distinct from the ones used for Figure 5A, were subjected to MAE as well as peptide purification at different days, and analyzed using CID at an Orbitrap XL. \(n = 51, 42, 44,\) and 37, in total 81 non-redundant peptides. C, High stability of the modification pattern in altered cellular state. The columns represent means from two MAEs of JY cells 12 h post infection (p.i.) and two MAEs of JY cells 16 h p.i. with influenza virus. Peptides were analyzed by EThcD in the same way as in Figure 5A. \(n = 107, 118, 83,\) and 97, in total 159 non-redundant peptides. Note for comparison: Figure 5A comprises three MAEs of uninfected JY cells and two MAEs of JY cells 0 h post mock infection. D, Consistent modification pattern using HLA-A*02:01 motif peptides derived from a different cell line. Four MAEs were performed in parallel from THP-1 cells and analyzed using CID at an Orbitrap XL. \(n = 26, 29, 13,\) and 14, in total 34 non-redundant peptides.
Supplementary Figure S10 | Robustness of the pattern of cysteine modifications observed for HLA-B*07:02 peptide ligands by MAE. The patterns of carbamidomethylation and cysteinylation described in Figure 5B for HLA-B*07:02 motif peptides (NetMHC IC_{50} < 500 nM) are supported by additional MAE analyses varying biological and technical parameters of the experimental setup. A, Similar modification pattern using a different peptide fragmentation method in MS. One of the MAEs from uninfected JY cells included in Figure 5B was additionally analyzed using HCD instead of EThcD. n = 43. B, High stability of the modification pattern in altered cellular state. The columns represent means from two MAEs of JY cells 12 h post infection (p.i.) and two MAEs of JY cells 16 h p.i. with influenza virus. Peptides were analyzed by EThcD in the same way as in Figure 5B. n = 39, 37, 31, and 26, in total 56 non-redundant peptides. Note for comparison: Figure 5B comprises three MAEs of uninfected JY cells and two MAEs of JY cells 0 h post mock infection.
Supplementary Figure S11 | 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. Same as Figure 5 but using reprocessed LC-MS data from Bassani-Sternberg et al.\textsuperscript{22} obtained by MHC-IAC from three biological replicates of JY cells; the cell lysate contained 0.2 mM iodoacetamide. \textbf{A}, Frequency of modified cysteine residues in cysteine-containing peptides with a NetMHC IC\textsubscript{50} < 500 nM for HLA-A*02:01; numbers of HLA-A*02:01 motif peptides containing modified cysteine (n) = 77, 41 and 53 respectively, in total 97 non-redundant peptides. \textbf{B}, Frequency of modified cysteine residues in cysteine-containing peptides with a NetMHC IC\textsubscript{50} < 500 nM for HLA-B*07:02; numbers of HLA-B*07:02 motif peptides containing modified cysteine (n) = 33, 13, and 32 respectively, in total 48 non-redundant peptides.
Supplementary Figure S12 | The highest cysteinylation frequency is observed for those positions of HLA-A*02:01 peptide ligands that are partly but not maximally exposed. 33 crystal structure entries of HLA-A*02:01 were retrieved from PDB, and structural parameters for exposure of individual peptide positions were calculated. A, Overlay of five representative HLA-A*02:01 crystal structures showing 9-mer peptide backbone and Cβ atom orientations. As positions 4 and C-4 were of particular interest in this study, we selected crystal structures from peptides not identical in the amino acid species at these positions for representative visualization. Considering this primary criterion, we selected those crystal structures whose HSE-β-up values correlated best with the mean HSE-β-up values calculated across all 33 HLA-A*02:01 crystal structures (see Supplementary Table S5). B, The coordination number (CN) underscores that peptide positions 1 and 3 (highlighted in green) showing the highest percentage of carbamidomethylation in MAE (see Figure 5A) are deeply buried in the MHC binding groove. C, The Half-Sphere Exposure (HSE)-β-up as well as (D) the relative solvent accessible surface (RSAS) demonstrate for example that peptide positions C-4 and C-1 (highlighted in red) showing the highest cysteinylation frequency in MAE (see Figure 5A) are partly but not very exposed. The mean relative standard deviation was only 21 % for HSE-β-up but 59 % for RSAS highlighting the better robustness of the former. B-D, Assignment of positions as “(partly) buried” or “(partly) exposed” is based on arbitrary thresholds aiming at an easy classification. Columns represent means of the 33 crystal structure PDB entries with error bars visualizing standard deviations. Corresponding P values are provided in Supplementary Table S6. E, Plotting the frequency of cysteinylated cysteine (cysteine) residues from Figure 5A against the mean HSE-β-up from (C) demonstrates that cysteinylation is not only rare at the buried MHC anchor positions but also at the highly exposed positions 4 and C-5, whereas the partly exposed positions C-4 and C-1 show the highest cysteinylation frequency. F and G, In agreement with the very low cysteinylation frequency at the most exposed positions 4 and C-5, big and hydrophobic amino acids are also strongly underrepresented at positions 4 (F) and C-5 (G) of HLA-A*02:01 peptide ligands (green bars) both in comparison to positions 4 and C-5 respectively of HLA-B*07:02 peptide ligands (blue bars) and in relation to their frequency in the human proteome. The ratio (observed amino acid frequency) / (amino acid frequency in the human proteome) = 1, that would be expected if neither the MHC processing pathway nor binding to the MHC protein would influence amino acid frequencies in MHC peptide ligands, is highlighted with a dotted red line. Same MAE data as in Figure 5. Amino acids are displayed in the order of their hydrophobicity index. Note that positions 4 and C-5 are identical for 9-mer peptides.
Supplementary Figure S13 | A coarse insight into the average orientation of peptide ligands in the HLA-B*07:02 binding groove. Four crystal structure entries of HLA-B*07:02 were retrieved from PDB, and structural parameters for exposure of individual peptide positions were calculated. A, Overlay of three HLA-B*07:02 crystal structures (see Supplementary Table S5) showing 9-mer peptide backbone and Cβ atom orientations. The peptide RPQVPLRPM containing three proline residues is depicted in bright blue, and the orientation of some of its side chains obviously differs from that of the other two 9-mers, especially at position C-4. B, The coordination number (CN) underscores that peptide positions 1 and 3 showing very low cysteinylation frequencies (see Figure 5B) are deeply buried in the MHC binding groove. Columns represent means of the four crystal structure PDB entries. C, Because of the high diversity of the four HLA-B*07:02 peptide ligands underlying the crystal structures, we represent the mean Half-Sphere Exposure (HSE)-β-up for every position with two distinct means; whereas the columns with error bars represent the means of all four PDB entries, the second column for every position (without error bars) represents the mean HSE-β-up of only the more typical 9-mer sequences TPQDLNTML and RPMTYKGAL while excluding the peptide RPQVPLRPM containing three proline residues (compare A) as well as the 11-mer HLA-B*07:02 peptide ligand RPHERNGFTVL. HSE-β-up demonstrates for example that peptide position C-1 (highlighted in dark red) showing the highest cysteinylation frequency in MAE (see Figure 5B) is partly but not very exposed. Positions 4, C-5 and C-4 appear to be the most exposed ones, but they showed much lower cysteinylation frequencies than position C-1 (see Figure 5B). B and C, Assignment of positions as “(partly) buried” or “(partly) exposed” is based on the same arbitrary thresholds used in Supplementary Figure S12 B and C respectively for easy comparison and classification. Error bars visualize standard deviations. Corresponding P values are provided in Supplementary Table S7.
Supplementary Figure S14 | At peptide position C-4, the MHC peptide binding groove is narrower in HLA-B*07:02 than in HLA-A*02:01, and this might contribute to force amino acid side chains into a more solvent exposed orientation. A and B, Both in HLA-A*02:01 (A) and in HLA-B*07:02 (B), the peptide position C-4 is in close proximity to an exposed arginine residue in the α1 helix (Arg65 and Arg62 respectively) and the exposed Gln155 in the α2 helix. However, the distance between the Arg65/62 and the Gln155 is usually shorter in HLA-B*07:02 than in HLA-A*02:01. The yellow arrow indicates the distance between the Arg65/62 Cζ atom and the Gln155 Cδ atom, which is 1.61 (±0.18) nm for HLA-A*02:01 using the mean (±s.d.) from the 33 HLA-A*02:01 PDB crystal structure entries and 1.28 (±0.15) nm for HLA-B*07:02 as calculated from the four HLA-B*07:02 PDB crystal structure entries (P = 0.012, heteroskedastic Student’s t-test; Supplementary Table S5 and Supplementary Methods). A to D, Crystal structures with MHC peptide ligands containing leucine at position C-4 are displayed to support a more homogeneous comparison between HLA-A*02:01 and HLA-B*07:02. Note that the hydrophobic leucine side chain has some similarity with the hydrophobic region at the basis of cystine side chains. A and C, The crystal structure of ILSALVGIL in complex with HLA-A*02:01 from PDB entry “1eez” chains A and C. B and D, The crystal structure of TPQDLNTML in complex with HLA-B*07:02 from PDB entry “4u1h”. C, In almost all of the crystal structures from the seven (out of the 33 preselected) HLA-A*02:01 PDB entries with leucine at peptide position C-4, the leucine is accommodated in a relatively solvent protected sideward orientation within the wide HLA-A*02:01 peptide binding groove. D, The leucine side chain at position C-4 of the HLA-B*07:02 peptide ligand TPQDLNTML sticks out of the narrow HLA-B*07:02 peptide binding groove resulting in a high solvent exposure (compare Supplementary Table S8). E, At peptide position C-4, the amino acid tyrosine having a big hydrophobic region at the basis of its side chain as well as the big hydrophobic amino acids phenylalanine, leucine and isoleucine occur less frequently in HLA-B*07:02 peptide ligands as compared to HLA-A*02:01 peptide ligands. This is in line with the assumption that the energetically unfavorable exposure of hydrophobic side chains to aqueous solvent contributes to underrepresentation of (partly) hydrophobic amino acid residues including cystine at highly exposed peptide positions. Same MAE data as in Figure 5. Amino acids are displayed in the order of their hydrophobicity index. Methionine frequencies at position C-4 showed no difference between HLA-A*02:01 and HLA-B*07:02 possibly due to oxidation of methionine in the MHC peptide binding groove which would make methionine more hydrophilic than suggested by its hydrophobicity index. F, P values for the comparison of HLA-A*02:01 and HLA-B*07:02 peptide ligands in (E) derived from heteroskedastic Student’s t-tests with Benjamini-Hochberg correction for five comparisons.
Supplementary Figure S15 | The addition of 10 % DMSO to the MAE elution buffer applied to the cells has no major effect on the observed peptide pattern. The comparison was performed both with JY (A – D) and THP-1 (E – H) cells using 5 · 10⁸ cells per MAE. For each comparison, cells were harvested and washed together and split into two populations just before the application of the MAE elution buffer containing no or 10 % DMSO. A and E, Both MAE elution buffers yield very similar peptide numbers (indicated in italics), and the identified peptides strongly overlap. The addition of 10 % DMSO to the MAE elution buffer furthermore has no major impact on the observed MHC allotype distribution as assigned by NetMHC (<500 nM IC₅₀ threshold for binders, B and F), the cumulative peptide hydrophobicity profile (visualised separately for the three LC-MS replicates, C and G) and the MS1 intensities (mean of three LC-MS replicates applied for every sample, depicting shared peptides, D and H). C and G, Note for interpretation: the higher the SSRCalc hydrophobicity index, the higher is the predicted hydrophobicity of the peptide. The steeper the curve, the more peptides were identified in this hydrophobicity range.
Supplementary Figure S16 | There is at most a minor fraction of MHC peptide ligands that does not elute from the MHC with 0.2 % TFA but can be recovered by the addition of acetonitrile. The comparison was performed both with JY (A – C) and THP-1 (D – F) cells using 5 · 10⁸ cells per MHC-IAC derived from the same cell culture harvesting of JY or THP-1, respectively. The MHC-IAC columns of all samples were first eluted four times with 0.2 % TFA. For the standard samples, four additional elutions with 0.2 % TFA followed as usual. For the acetonitrile elution samples, the first four elutions with 0.2 % TFA were followed by four elutions with 80 % (JY) or 60 % (THP-1) acetonitrile in 0.2 % TFA. A and D, The elution including acetonitrile yielded slightly lower peptide numbers (indicated in italics) than the elution based on 0.2 % TFA only, but the overlap of peptides obtained with both elution strategies was high. Minor changes of the MHC allotype distribution (B and E, peptide numbers in italics) and the cumulative peptide hydrophobicity profile (visualised separately for the three LC-MS replicates, C and F) were observed in the JY but not the THP-1 samples. C and F, Note for interpretation: the higher the SSRCalc hydrophobicity index, the higher is the predicted hydrophobicity of the peptide. The steeper the curve, the more peptides were identified in this hydrophobicity range.
Supplementary Figure S17 | Rinsing the ultrafilter with an acetonitrile-rich solution improves overall peptide yield and is particularly beneficial for hydrophobic peptides. 7.5 ml LCL5 cell pellet was split into three aliquots (approximately $1.25 \times 10^9$ cells each) that were subjected to three parallel MHC-IACs. The eluate from the MHC-IAC column was passed through an Amicon ultrafilter device with 10 kDa molecular weight cut-off yielding the first ultrafiltrate (UF). Then, 1 ml of 80 % acetonitrile / 0.2 % TFA was passed through the ultrafilter resulting in the rinse UF. For MHC-IAC 1 and MHC-IAC 2, a 4 ml Amicon was used, and the first UF and rinse UF were processed separately after ultrafiltration, whereas for MHC-IAC 3, we applied a 0.5 ml Amicon and combined the first UF and the rinse UF. A, Samples containing the rinse UF yielded about 2 to 3 times more peptide identifications than the first UF and approached saturation regarding obtainable peptide identifications at the LC-MS system employed. The better performance of the rinse UF was particularly pronounced for hydrophobic peptides. Note for interpretation: the higher the SSRCalc hydrophobicity index, the higher is the predicted hydrophobicity of the peptide. The steeper the curve, the more peptides were identified in this hydrophobicity range. B, The boost in peptide yield caused by acetonitrile-based filter rinsing is even more obvious in view of MS1 intensities as compared to peptide numbers. For a better comparability of peptide abundance, we only included the MS1 intensities of the 445 most intense peptides per LC-MS run as 445 was the minimum number of peptides identified in any of the underlying 15 LC-MS runs. The mean MS1 intensity of the three LC-MS measurements per sample is represented by the columns with error bars indicating standard deviation. The gain in mean MS1 intensity for the rinse UF relative to the first UF was factor 41 for MHC-IAC 1 and factor 27 for MHC-IAC 2.
Supplementary Figure S18 | Rinsing the ultrafilter with an acetonitrile-rich solution shifts the distribution of observed peptides towards those MHC allotypes that favor hydrophobic ligands. Peptides were assigned to the MHC allotype with the best NetMHC IC_{50}. Peptides with a predicted IC_{50} ≥ 500 nM were classified as non-binders. Numbers in italics indicate absolute peptide numbers. The total peptide number per sample is indicated above each column. For dominant peptide groups, the absolute peptide number is also indicated below the percentage numbers within the columns. A and B, Same triplicate MHC-IAC experiment as in Supplementary Figure S17. A, In relative terms, peptide ligands of HLA-A*29:02 and A*24:02 profited more from the acetonitrile-based ultrafilter rinsing than the HLA-B*44:02 and B*44:03 peptide ligands. B, HLA-A*29:02 and A*24:02 derived peptides are on average more hydrophobic than those from HLA-B*44:02 and B*44:03 as demonstrated by the fact that they dominate at the end of the C_{18} reverse phase LC gradient whereas the ligands of the B*44:02 and B*44:03 allotypes prevail at earlier elution times. The data plotted in B is derived from the combined UF of MHC-IAC 3. C, D and E, Even in the MAE procedure (where, in contrast to MHC-IAC, the first UF already contained 60 % acetonitrile), the typically hydrophobic HLA-A*02:01 peptide ligands profited most from rinsing the ultrafilter with 2 ml 60 % acetonitrile / 0.1 % TFA. C, Data obtained by one MAE of 5 · 10^8 JY cells. The corresponding peptide elution profile of the first UF is depicted in Supplementary Figure S19. D and E, Data obtained by one MAE of 1.5 · 10^9 THP-1 cells. Both the first UF and the rinse UF were split in three equal aliquots. One aliquot each was processed separately while another aliquot each of the first UF and the rinse UF were pooled resulting in the combined UF. E, The elution profile of the combined UF reveals that HLA-A*02:01 and HLA-A*24:02 derived peptides are on average more hydrophobic than the peptide ligands assigned to HLA-B*15:03 and B*35:01.
Supplementary Figure S19 | Observed and predicted hydrophobicity of peptide ligands of different MHC allotypes. Peptide ligands of different MHC allotypes have distinct hydrophobicity properties, and this is sufficiently factored in by NetMHC but not by Syfpeithi prediction algorithms for HLA-A*02:01 and HLA-B*07:02. A, B, D, F, and G. Results obtained by MAE of 5 \cdot 10^6 JY cells. C, E, and H, HLA-A*02:01 and HLA-B*07:02 nonamer binders were predicted from the SwissProt human protein database, and for each MHC allotype, 5000 peptides were selected for display. A, Observed elution profile of peptides in LC-MS. Peptides were assigned to the MHC allotype with the best NetMHC IC_{50}. Peptides with a predicted IC_{50} ≥ 500 nM were classified as non-binders. B, The SSRCalc hydrophobicity index is a good surrogate parameter for measured LC retention times. Same as (A), but with observed retention times replaced with SSRCalc hydrophobicity index (higher values correspond to higher hydrophobicity / late elution times in reversed phase nanoHPLC). C, The hydrophobicity pattern of NetMHC-predicted HLA-A*02:01 and HLA-B*07:02 peptide ligands from SwissProt shown here resembles the hydrophobicity pattern experimentally observed in (A) and (B) respectively. D, Same as (B), but using Syfpeithi instead of NetMHC for MHC allotype assignment. Peptides were assigned to the MHC allotype with the best relative Syfpeithi score. Peptides with a predicted relative Syfpeithi score <50 % were classified as non-binders. E, The hydrophobicity pattern of Syfpeithi-predicted HLA-A*02:01 and HLA-B*07:02 peptide ligands from SwissProt shown here reveals clear deviations from the MHC allotype ratios experimentally observed in (D) for hydrophobic peptides. Too little hydrophobic peptides from SwissProt were selected for HLA-A*02:01 assignment and too many of them for HLA-B*07:02 assignment. F, The anchor amino acids themselves contribute only little to the observed hydrophobicity pattern. Same as (B), but with anchor amino acid positions (i.e. position 2 and C-term) replaced by alanin in silico after NetMHC prediction but before the calculation of SSRCalc hydrophobicity indices. G, Same as (B), but solely using anchor amino acids for MHC allotype assignment instead of NetMHC predictions. Peptides were assigned to HLA-A*02:01 if Leu or Met occurred at position 2 and Val or Leu at the C-terminus. Prolin at position 2 combined with Leu at the C-terminus was considered as indicative of an HLA-B*07:02 ligand. H, Predicted peptides from SwissProt that contain the appropriate anchor amino acids for HLA-A*02:01 are on average more hydrophobic than those containing the anchor amino acids for HLA-B*07:02. However, this explains only a small fraction of the big difference in hydrophobicity between HLA-A*02:01 and HLA-B*07:02 ligands observed in (G). Anchor amino acids were defined as in (G).
Supplementary Table S1 | MHC-IAC peptide extracts from different cell cultures show a variable proportion of cysteinylated peptides and an even less constant fraction of peptides with unmodified cysteine. Data for JY cells and THP-1 cells are derived from two different cell cultures each. The three samples of LCL5 cells represent parallel MHC-IAC extractions of the same cell culture. Numbers in italics correspond to cysteinylation percentages obtained with cell culture medium “M1” (see legend) and are also depicted in Table 1. The reasons for the high variability in the occurrence of cysteinylated and unmodified cysteine residues remain to be defined, but some possible explanations are mentioned in the following. The presence of different MHC allotypes in each cell line has an impact (compare Figure 4). Possibly, the concentration of cystine in the medium makes a difference. Additionally, cysteine modifications can be assumed to be influenced by the redox potential of the cell culture medium also changed by antioxidative additives like 2-mercaptoethanol. The medium’s redox potential might vary according to the availability of oxygen which is dependent on cell density, cell growth rate and the time span between cell harvesting and the previous feeding of cells in the airtight roller bottles. Oxidations of cysteine might vary with storage times of frozen cell pellets or with the duration of sample storage prior to LC-MS measurement. As can be seen from the LCL5 data, the percentages of cysteinylated and unmodified cysteine residues were very reproducible if the samples were derived from the same cell culture.

| Cell line | Total number of identified peptides | Relative number of cysteine-containing peptides among all identified peptides in % | Remarks on sample processing |
|-----------|-----------------------------------|-----------------------------------|-----------------------------|
| JY        | 1627                              | 0.43                              | F1, M1, NP                  |
|           | 2292                              | 0.57                              | F2, M1, NP                  |
| THP-1     | 1016                              | 0.59                              | F1, M1, P                   |
|           | 1563                              | 1.86                              | F2, M2, P                   |
| LCL5      | 2117                              | 0.94                              | F3, M3, P                   |
|           | 2134                              | 1.03                              | F3, M3, P                   |
|           | 2010                              | 1.19                              | F3, M3, P                   |

F1: cell pellets were processed by MHC-IAC a few days after freezing;
F2: cell pellets were processed by MHC-IAC several months after freezing;
F3: cell pellets were processed by MHC-IAC about three years after freezing;
M1: cell culture medium “RPMI ready” (see Supplementary Materials and Methods for details); cystine concentration in “RPMI ready” is 20 g/L;
M2: cell culture medium “RPMI ready modified” differing from “RPMI ready” as follows: 30 µM 2-mercaptoethanol, 377 µM sodium pyruvate, 60 mM glutamine and 30 mM HEPES;
M3: cell culture medium “RPMI powder-based” (see Supplementary Materials and Methods for details) containing 50 g/L cystine and 50 µM 2-mercaptoethanol;
NP: the second indicated JY sample was desalted and measured by LC-MS 5 weeks after the first indicated JY sample;
P: the indicated samples of this cell line were desalted and measured with nested replicates within the same group of LC-MS runs within a few days.
Supplementary Table S2 | Experimental setup and peptide yield of our MHC immunopeptidomic analyses in comparison with the data from Lanoix et al.\textsuperscript{18}

| Method       | MAE                              | MHC-IAC                           |
|--------------|----------------------------------|-----------------------------------|
| Data set     | Our EThcD data                   | Our THP-1 data                    | Lanoix et al.                      | Our THP-1 data | Lanoix et al. |
| Cell line    | JY                               | THP-1                             | B-LCL                             | THP-1          | B-LCL         |
| Number of cells | 5 \times 10^8     | 1.25 \times 10^8                  | 10^8 / rep.                       | 1.25 \times 10^8  | 10^8 / rep.  |
| Type of replicate | LC-MS injection | LC-MS injection                  | Biological                        | LC-MS injection | Biological   |
| Percentage of sample injected per LC-MS run | 10%                             | 10%                               | 100%                              | 10%            | 100%          |
| Mass spectrometer | Orbitrap Fusion | LTQ Orbitrap XL                   | Q Exactive HF                     | LTQ Orbitrap XL | Q Exactive HF |
| Peptide number (rep. 1) | 3623                             | 476                               | 596                               | 801            | 2221          |
| Peptide number (rep. 2) | 3249                             | 403                               | 369                               | 879            | 2148          |
| Peptide number (rep. 3) | 3254                             | 429                               | 794                               | 848            | 2043          |

Due to the different cell types and the different LC-MS setups, an accurate comparison of MAE and MHC-IAC performance between the two studies is not possible. For a rough comparative estimation, we selected representative LC-MS data from each study and reprocessed it with identical settings whenever appropriate using Proteome Discoverer 1.4 and Sequest HT as detailed in the methods section. Note, that we based the comparison of our MAE with our MHC-IAC on THP-1 cells derived from splitting of exactly the same cell culture suspension whereas Lanoix et al.\textsuperscript{18} applied separate biological replicates.
**Supplementary Table S3 | Statistical analysis of the cysteine modification pattern depicted in Figure 5A (HLA-A*02:01 motif peptides).**

| Position | 1     | 2     | 3     | 4     | C-5   | C-4   | C-3   | C-2   | C-1   | C    |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
| 1        | 0.0194| 0.4524| 0.0231| 0.0749| **0.0016**| 0.2953| 0.0226| **0.0013**| 0.0130|      |
| 2        | **0.0032**|       | 0.0925| 0.6014| 0.0694| **0.0013**| **0.0032**| **1.1E-05**| 0.0013| 0.0499|
| 3        | 0.0418| **0.0046**|       | 0.1092| 0.3637| **0.0013**| 0.1092| 0.0174| **0.0013**| 0.0461|
| 4        | **0.0018**| **0.0001**| **0.0018**| 0.0248| 0.1092| **0.0013**| **0.0035**| **1.1E-05**| 0.0013| 0.0238|
| C-5      | **0.0018**| **0.0003**| **0.0018**| 0.3830| 0.0711| **0.0013**| **0.0086**| **0.0002**| **0.0013**| 0.0174|
| C-4      | **0.0018**| **0.0001**| **0.0018**| 0.3830| 0.0711| **0.0013**| **0.0086**| **0.0002**| **0.0013**| 0.0174|
| C-3      | 0.0018| 0.0001| 0.0018|       | n.a. | 0.0248| 0.3830|       | 0.0612| **0.0019**| **0.0032**|
| C-2      | **0.0018**| **0.0003**| **0.0018**| 0.0248| 1.0000| 0.0711| 0.0248|       | **0.0069**| **0.0002**|
| C-1      | **0.0018**| 0.1658| **0.0018**| 0.0428| 0.1468| 0.0527| 0.0428| 0.1468|       | **0.0013**|
| C        | **0.0018**| **0.0001**| **0.0018**| n.a. | 0.0248| 0.3830| n.a. | 0.0248| 0.0428|      |

Frequencies of carbamidomethylation were compared between all indicated positions in a pairwise manner (42 comparisons), and the resulting *P* values are listed with green shading. Accordingly, we also compared the frequencies of cysteinylation using all indicated pairs of positions (45 comparisons) and depict the obtained *P* values with red shading. All comparisons were performed with two-tailed, heteroskedastic Student’s *t*-tests using separate Benjamini-Hochberg corrections for carbamidomethylation and cysteinylation. *P* values ≤ 0.01 are printed in bold and highlighted with more intense shading. n.a. = not applicable.
Supplementary Table S4 | Statistical analysis of the cysteinylation pattern depicted in Figure 5B (HLA-B*07:02 motif peptides).

| Position | 2     | 3     | 4     | C-5    | C-4    | C-3    | C-2    | C-1    | C     |
|----------|-------|-------|-------|--------|--------|--------|--------|--------|-------|
| 1        | 0.3826| 0.3826| 0.0080| 0.0354 | 0.2449 | 0.0068 | 0.0015 | 0.0068 | 0.0162|
| 2        | n.a.  | 0.0093| 0.0313| 0.0907 | 0.0068 | 0.0032 | 0.0068 | 0.0015 |       |
| 3        | 0.0093| 0.0313| 0.0907| 0.0068 | 0.0032 | 0.0068 | 0.0015 |        |       |
| 4        | 0.8284| 0.0162| 0.0205| 0.0103 | 0.0077 | 0.0392 | 0.0068 | 0.1095 |       |
| C-5      |       | 0.0610| 0.0313| 0.0541 | 0.0068 |        |        |        |       |
| C-4      |       |       | 0.0068| 0.0015 | 0.0068 | 0.2252 |        |        |       |
| C-3      |       |       |       | 0.3098| 0.0313 | 0.0100 |        |        |       |
| C-2      |       |       |       |       | 0.0183 | 0.0061 |        |        |       |
| C-1      |       |       |       |       |       | 0.0068 |        |        |       |

Frequencies of cysteinylation were compared between all indicated positions in a pairwise manner (44 comparisons), and the resulting $P$ values are listed with red shading. Due to the very low numbers of identified HLA-B*07:02 motif peptides containing carbamidomethylated cysteine, we do not perform statistical analyses for carbamidomethylation here. All comparisons were performed with two-tailed, heteroskedastic Student’s t-tests using Benjamini-Hochberg correction. $P$ values ≤ 0.01 are printed in bold and highlighted with more intense shading. n.a. = not applicable.
Supplementary Table S5 | PDB crystal structures used for the analyses.

| HLA allotype | PDB accession no. | Peptide ligand | No. of amino acid residues | PDB chain name | Correlation ($R^2$) of HSE-β-up with mean |
|--------------|------------------|----------------|---------------------------|----------------|-----------------------------------------|
| A*02:01     | 1b0g             | ALMGFFPVL      | 9                         | C              | 0.963                                   |
|              | 1duz             | LLFGYPVYV      | 9                         | C              | 0.831                                   |
|              | 1eez             | ILSALVGIL      | 9                         | C              | 0.923                                   |
|              | 1hhg             | TLTSCNTSV      | 9                         | F              | 0.664                                   |
|              | 1hhb             | FLPSDFPSV      | 10                        | C              | 0.778                                   |
|              | 1i4f             | GYVDREHTV      | 10                        | C              | 0.833                                   |
|              | 1i7r             | FAPGFFPVL      | 9                         | C              | 0.963                                   |
|              | 1tvb             | ITDOVPFSV      | 9                         | C              | 0.882                                   |
|              | 2c7u             | SLFNTIAVL      | 9                         | F              | 0.894                                   |
|              | 2v2x             | SLFNTVATL      | 9                         | C              | 0.736                                   |
|              | 2x4o             | KLTPLCVTL      | 9                         | F              | 0.914                                   |
|              | 2x4s             | AMDSNTLEL      | 9                         | C              | 0.843                                   |
|              | 2x4u             | ILKEPVHG      | 9                         | C              | 0.808                                   |
|              | 3d25             | VLDDELLEA      | 9                         | F              | 0.885                                   |
|              | 3gso             | NLVPVWA5V      | 9                         | P              | 0.885                                   |
|              | 3h7b             | MLWGYLQYV      | 9                         | C              | 0.909                                   |
|              | 3hpj             | RMFPNAPYL      | 9                         | C              | 0.867                                   |
|              | 3i6g             | GLMWSYYVF      | 9                         | C              | 0.750                                   |
|              | 3kla             | SLLMNITQL      | 9                         | C              | 0.844                                   |
|              | 3mgo             | RLYQNPPTYI     | 10                        | C              | 0.842                                   |
|              | 3mre             | GLCTLNWAML     | 9                         | P              | 0.790                                   |
|              | 3npn             | CLNGCWCYV      | 9                         | P              | 0.910                                   |
|              | 3nrr             | KLVAGINA       | 9                         | P              | 0.850                                   |
|              | 3pwn             | LLYGFVNYI      | 9                         | C              | 0.929                                   |
|              | 3rew             | CLGGLLTMV      | 9                         | C              | 0.911                                   |
|              | 3cv5d            | KVAELVHF       | 9                         | C              | 0.916                                   |
|              | 4gks             | FLTGIGITV      | 10                        | C              | 0.824                                   |
|              | 4i4w             | ILAKFLHRE      | 9                         | C              | 0.770                                   |
|              | 4jfo             | ALAGIGILTV     | 10                        | C              | 0.918                                   |
|              | 5c0e             | YQFGPDSPFA     | 10                        | C              | 0.793                                   |
|              | 51ny             | KAVGPDPLYV     | 10                        | C              | 0.791                                   |
|              | 3vcl             | RHERNGFTVL     | 11                        | C              | 0.980                                   |
|              | 4u1h             | TPQDINTML      | 9                         | C              | 0.966                                   |
|              | 4u1k             | RQPQVRPM       | 9                         | F              | 0.861                                   |
|              | 5eo1             | RPMTYKGAL      | 9                         | C              | 0.908                                   |

Note that many PDB accession numbers point to more than one crystal structure of the same MHC-β₂m-peptide complex and hence contain the same MHC peptide ligand with more than one PDB chain name. $R^2$ values are derived from the Pearson correlation of HSE-β-up of each individual peptide ligand structure with that of the mean across all listed ligands of the corresponding MHC allotype. Hence, the higher the $R^2$ value of the structure, the more likely the structure represents typical peptide backbone and side chain orientations. The crystal structures used for the visualization in Supplementary Figure S12A and Supplementary Figure S13A are highlighted with pale blue shading.
### Supplementary Table S6 | Statistical analysis of structural parameters of HLA-A*02 peptide ligands from PDB crystal structures.

| Position | 1       | 2       | 3       | 4       | C-5      | C-4      | C-3      | C-2      | C-1      | C       |
|----------|---------|---------|---------|---------|----------|----------|----------|----------|----------|---------|
| 1        | 1.8E-25 | 1.3E-09 | 3.1E-26 | 2.8E-18 | 7.1E-10  | 0.0030   | 1.6E-10  | 1.4E-27  | 1.1E-26  |
| 2        | 5.0E-06 | 1.0E-12 | 9.1E-34 | 5.1E-25 | 3.5E-17  | 1.0E-11  | 4.0E-30  | 6.7E-43  | 0.2308   |
| 3        | 2.2E-15 | 1.2E-19 | 4.1E-31 | 2.8E-22 | 1.2E-13  | 3.8E-07  | 2.4E-20  | 4.6E-34  | 2.1E-14  |
| 4        | 1.4E-40 | 1.3E-45 | 2.7E-31 | 0.6349  | 0.0005   | 2.4E-10  | 1.5E-19  | 7.8E-09  | 2.9E-34  |
| C-5      | 9.5E-39 | 1.1E-43 | 1.3E-31 | 0.4936  | 0.0038   | 4.3E-09  | 4.5E-13  | 4.0E-05  | 1.8E-25  |
| C-4      | 3.1E-31 | 6.0E-35 | 1.3E-27 | 0.9791  | 0.5844   | 0.0007   | 7.3E-05  | 0.9720   | 1.3E-17  |
| C-3      | 4.1E-19 | 2.5E-21 | 4.8E-12 | 1.5E-07 | 3.4E-08  | 4.6E-07  | 0.7440   | 4.0E-05  | 3.2E-12  |
| C-2      | 3.7E-34 | 7.6E-39 | 1.6E-20 | 6.5E-14 | 1.6E-14  | 1.0E-10  | 3.8E-07  | 2.4E-31  |
| C-1      | 5.5E-42 | 1.2E-45 | 1.3E-22 | 3.3E-13 | 1.0E-13  | 1.7E-09  | 0.3304   | 0.2000   | 2.4E-43  |
| C        | 3.2E-18 | 2.5E-23 | 0.0161  | 3.3E-38 | 5.7E-37  | 3.7E-29  | 3.7E-14  | 7.2E-28  | 2.2E-33  |

Values of four structural parameters were calculated for 33 HLA-A*02 crystal structure PDB entries (Supplementary Figure S12 and Supplementary Table S5) and compared between all indicated peptide ligand amino acid positions in a pairwise manner (45 comparisons per parameter) using two-tailed, heteroskedastic Student’s t-tests and separate Benjamini-Hochberg corrections for each of the four parameters. The resulting P values are listed with red shading for HSE-β-up, green shading for CN, blue shading for SAS and violet shading for RSAS. P values ≤ 0.01 are printed in bold and highlighted with more intense shading.
### Supplementary Table S7 | Statistical analysis of structural parameters of HLA-B*07 peptide ligands from PDB crystal structures.

| Position | 1 | 2 | 3 | 4 | C-5 | C-4 | C-3 | C-2 | C-1 | C |
|----------|---|---|---|---|-----|-----|-----|-----|-----|----|
| 1        |   |   |   |   | 0.0098 | 0.0139 | 0.0029 | 0.0024 | 0.0800 | 0.3846 | 0.0430 | 0.0308 | 0.0090 |
| 2        | 0.1094 | 0.0434 | 0.0002 | 0.0007 | 0.0160 | 0.0276 | 0.0002 | 7.8E-07 | 0.2320 |
| 3        | 0.0309 | 0.0137 | 4.0E-05 | 0.0002 | 0.0191 | 0.0568 | 0.0002 | 0.0004 | 0.1818 |
| 4        | 6.3E-06 | 1.1E-05 | 0.0013 | 0.7067 | 0.3614 | 0.0038 | 0.0028 | 0.0090 | 4.0E-05 |
| C-5      | 0.0020 | 0.0009 | 0.0008 | 0.4468 | 0.3072 | 0.0024 | 0.0009 | 0.0217 | 0.0002 |
| C-4      | 0.0008 | 0.0002 | 0.0002 | 0.2178 | 0.8499 | 0.0430 | 0.3846 | 0.5550 | 0.0160 |
| C-3      | 0.0137 | 0.0085 | 0.0309 | 0.0424 | 0.0229 | 0.0176 | 0.0286 | 0.0259 | 0.0308 |
| C-2      | 0.0005 | 6.2E-05 | 0.0007 | 0.0434 | 0.0943 | 0.0309 | 0.1065 | 0.2481 | 8.9E-05 |
| C-1      | 0.0006 | 7.9E-05 | 0.0007 | 0.0607 | 0.1011 | 0.0338 | 0.1011 | 0.8990 |
| C        | 0.0264 | 0.0060 | 0.2917 | 0.0002 | 0.0011 | 0.0002 | 0.0216 | 0.0001 | 0.0001 |

Values of four structural parameters were calculated for four HLA-B*07 crystal structure PDB entries (Supplementary Figure S13 and Supplementary Table S5) and compared between all indicated peptide ligand amino acid positions in a pairwise manner (45 comparisons per parameter) using two-tailed, heteroskedastic Student’s t-tests and separate Benjamini-Hochberg corrections for each of the four parameters. The resulting \(P\) values are listed with red shading for HSE-\(\beta\)-up, green shading for CN, blue shading for SAS and violet shading for RSAS. \(P\) values ≤ 0.01 are printed in bold and highlighted with more intense shading.
Supplementary Table S8 | At peptide position C-4, amino acid residues sharing some size and hydrophobicity features with cysteinylated cysteine are more solvent exposed in HLA-B*07:02 peptide ligands than in HLA-A*02:01 peptide ligands.

| Amino acid at position C-4 | HLA-B*07:02 PDB accession number | HLA-B*07:02 peptide ligand in PDB structure | HLA-A*02:01 PDB accession numbers | RSAS at position C-4 | Side chain orientation at position C-4 |
|---------------------------|---------------------------------|-------------------------------------------|-----------------------------------|----------------------|--------------------------------------|
| L                         | 4u1h                            | TPQDLNTML                                  | 1eez, 2x4o, 3i6g, 3mre, 3rew, 3v5d | 0.59                 | HLA-B*07:02 0.59  HLA-A*02:01 0.25 (±0.09) |
|                           |                                 |                                           |                                   |                      | Upwards                          |
| F                         | 5eo0                            | RPMTFKCAL                                   | 1b0g, 1hhh, 1hhi, 1i7r, 3pwn, 4i4w | 0.78                 | HLA-B*07:02 0.78  HLA-A*02:01 0.23 (±0.18) |
|                           |                                 |                                           |                                   |                      | Upwards                          |
| Y                         | 5eo1                            | RPMTYKCAL                                   | 1duz, 3h7b                       | 0.81                 | HLA-B*07:02 0.81  HLA-A*02:01 0.54 and 0.67 |
|                           |                                 |                                           |                                   |                      | Upwards                          |

Whereas HSE-β-up is more robust than RSAS against different amino acid types when estimating the exposure at a given position, RSAS has the advantage that it really calculates the surface area that is exposed to solvent. Therefore, RSAS is more closely linked to the energetic forces driving the hydrophobic effect than HSE-β-up. Out of the totally available five PDB crystal structure entries for HLA-B*07:02, three have amino acid residues at peptide position C-4 that share some size and hydrophobicity features with cysteine. Here, we directly compare the RSAS and side chain orientation at position C-4 in these three structures with those HLA-A*02:01 crystal structures (out of the 33 preselected ones, see Supplementary Table S5) that have the same amino acid in peptide position C-4. For the sake of completeness, we separately and exceptionally include PDB entry “5eo0” into the comparison although the involved peptide ligand is merely a single amino acid variation (SAV) version of the peptide ligand of PDB entry “5eo1”. For an illustration of the side chain orientation at position C-4 refer to Supplementary Figure S14 C and D. Numbers in brackets indicate standard deviation, and the number before the bracket represents the mean across the used HLA-A*02:01 PDB entries.
Supplementary Table S9 | MAE setups varying between experiments and assignment of MAE results to MS raw files. For all MAEs from JY cells, we used $5 \times 10^8$ cells per sample. The sign (==) indicates that an aliquot of the same cell culture was also subjected to MHC-IAC (compare Supplementary Table S10). The asterisk (*) denotes that ZipTip eluates from two biological replicates of $5 \cdot 10^8$ JY cells each were pooled. If not mentioned otherwise, the first ultrafiltrate (UF) and the rinse UF was combined for each sample. Raw files derived from the Orbitrap Fusion are highlighted in violet; all other raw files were acquired at the LTQ Orbitrap XL. Note that some figures and tables are listed in several rows of the first column.

| Figure / Table / main text | Cells | Additives to MAE elution buffer | Percentage of CH$_3$CN for elution of Oasis HLB column | Percentage of CH$_3$CN for elution of C$_18$ ZipTip | Percentage of sample injected per LC-MS run | MS raw files |
|-----------------------------|-------|---------------------------------|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|---------------|
| Figure 1, Figure 2A, Figure 3A, Supplementary Figure S3A and B, Supplementary Figure S5A, Supplementary Figure S6A, Supplementary Figure S9B, Supplementary Figure S18C, Supplementary Figure S19, Table 1 | JY (==) | GM | 60% | Rinse UF not used for indicated MS raw files | 35% | 10% | 150425_ThSt_C12_MAE_Probe1_10%_Rep1_25cm90min3s_msms29.raw | 150425_ThSt_C12_MAE_Probe1_10%_Rep2_25cm90min3s_msms35.raw | 150425_ThSt_C12_MAE_Probe1_10%_Rep3_25cm90min3s_msms41.raw |
| Figure 3B, Figure 4, Supplementary Figure S2, Supplementary Figure S3D, Supplementary Figure S5B, Supplementary Figure S6B, Supplementary Figure S7A, Supplementary Figure S9D, Table 1, Supplementary Table S2 | THP-1 (1.25 - $10^8$ cells) (==) | Leupeptin, pepstatin, EDTA, GM | 60% | 60% CH$_3$CN / 0,1% TFA | 60% | 10% | 150808_ThSt_C12-W1_MAE_Probe3_10%_Rep1_25cm90min3s_msms5.raw | 150808_ThSt_C12-W1_MAE_Probe3_10%_Rep2_25cm90min3s_msms15.raw | 150808_ThSt_C12-W1_MAE_Probe3_10%_Rep3_25cm90min3s_msms26.raw |
| Figure 4, Figure 5, Supplementary Figure S12E-G, Supplementary Figure S14 E and F, Supplementary Table S3, Supplementary Table S2, Supplementary Table S4 | JY | Leupeptin, pepstatin, GM | 35% | 35% CH$_3$CN / 0,1% TFA | 35% | 10% | F1_Agilent6_20170204_TST_N3W3_0h_uninf_10pc_Rep1_MAE_EThcD_OT.raw | F1_Agilent6_20170204_TST_N3W3_0h_uninf_10pc_Rep2_MAE_EThcD_OT.raw | F1_Agilent6_20170204_TST_N3W3_0h_uninf_10pc_Rep3_MAE_EThcD_OT.raw |
| Figure 4, Figure 5, Supplementary Figure S12E-G, Supplementary Figure S14 E and F, Supplementary Table S3, Supplementary Table S4 | JY (*) | Leupeptin, pepstatin, GM | 60% | 60% CH$_3$CN / 0,1% TFA | 60% | 5% | F1_Agilent6_20161122_TST_N3_0h_uninf_10pc_Rep1_MAE_EThcD_OT.raw | F1_Agilent6_20161122_TST_N3_0h_uninf_10pc_Rep2_MAE_EThcD_OT.raw | F1_Agilent6_20161122_TST_N3_0h_uninf_10pc_Rep3_MAE_EThcD_OT.raw |

(Table continued on next page)
| Figure / Table / main text | Cells | Additives to MAE elution buffer | Percentage of CH₃CN for elution of Oasis HLB column | Rinsing of ultrafilter | Percentage of CH₃CN for elution of C₁₈ ZipTip | Percentage of sample injected per LC-MS run | MS raw files |
|---------------------------|-------|---------------------------------|---------------------------------------------------|----------------------|---------------------------------------------|--------------------------------------------|-------------|
| Figure 4, Supplementary Figure S3C and D, Supplementary Figure S9D, Supplementary Figure S15E-H, Table 1 | THP-1 (5 · 10⁸ cells) | Leupeptin, pepstatin, EDTA, GM | 60% | 60% CH₃CN / 0,1% TFA | 60% | 10% | 150808_ThSt_C12-W1_MAE_Probe1_10%_Rep1_25cm90min3s_msms3.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe1_10%_Rep2_25cm90min3s_msms11.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe1_10%_Rep3_25cm90min3s_msms20.raw |
| Figure 4, Supplementary Figure S3C and D, Supplementary Figure S9D, Supplementary Figure S15E-H, Table 1 | THP-1 (1.5 · 10⁹ cells) | Leupeptin, pepstatin, EDTA, GM | 60% | 60% CH₃CN / 0,1% TFA | 60% | 3.3% | 150724_ThSt_C12-W1_MAE_Probe4_UF-Misch-vorZT_3komma3%_Rep1_25cm90min3s_msms7.raw |
| | | | | | | | 150724_ThSt_C12-W1_MAE_Probe4_UF-Misch-vorZT_3komma3%_Rep2_25cm90min3s_msms11.raw |
| | | | | | | | 150724_ThSt_C12-W1_MAE_Probe4_UF-Misch-vorZT_3komma3%_Rep3_25cm90min3s_msms13.raw |
| Figure 4, Supplementary Figure S9D, Supplementary Figure S15E-H, Table 1 | THP-1 (5 · 10⁸ cells) | Leupeptin, pepstatin, EDTA, GM, 10% DMSO | 60% | 60% CH₃CN / 0,1% TFA | 60% | 10% | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep1_25cm90min3s_msms23.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep2_25cm90min3s_msms13.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep3_25cm90min3s_msms3.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep4_25cm90min3s_msms4.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep5_25cm90min3s_msms5.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep6_25cm90min3s_msms6.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep7_25cm90min3s_msms7.raw |
| | | | | | | | 150808_ThSt_C12-W1_MAE_Probe2_10%_Rep8_25cm90min3s_msms8.raw |
| Supplementary Figure S3A and B, Supplementary Figure S9B, Table 1 | JY | None | 60% | No | 80% | 10% | 141111_ThSt_C11-W2_MAE_Probe2_10%_Rep1_25cm90min3s_msms4.raw |
| | | | | | | | 141112_ThSt_C11-W2_MAE_Probe2_10%_Rep4_25cm90min3s_msms3.raw |
| | | | | | | | 141112_ThSt_C11-W2_MAE_Probe2_10%_Rep5_25cm90min3s_msms4.raw |
| | | | | | | | 141112_ThSt_C11-W2_MAE_Probe2_10%_Rep6_25cm90min3s_msms5.raw |
| | | | | | | | 141112_ThSt_C11-W2_MAE_Probe2_10%_Rep7_25cm90min3s_msms6.raw |
| | | | | | | | 141112_ThSt_C11-W2_MAE_Probe2_10%_Rep8_25cm90min3s_msms7.raw |
| Supplementary Figure S3E and F | JY | GM | 60% | No | 80% | 10% | 150106_ThSt_C11-W4_MAE_Probe4_10%_Rep1_25cm90min3s_msms12.raw |
| | | | | | | | 150106_ThSt_C11-W4_MAE_Probe4_10%_Rep2_25cm90min3s_msms20.raw |
| | | | | | | | 150106_ThSt_C11-W4_MAE_Probe4_10%_Rep3_25cm90min3s_msms28.raw |
| | | | | | | | 150227_ThSt_C11-W5_MAE_Probe1-Oasis1_10%_2ugZT_Rep1_25cm90min3s_msms30.raw |
| | | | | | | | 150227_ThSt_C11-W5_MAE_Probe1-Oasis1_10%_2ugZT_Rep2_25cm90min3s_msms38.raw |
| | | | | | | | 150227_ThSt_C11-W5_MAE_Probe1-Oasis1_10%_2ugZT_Rep3_25cm90min3s_msms44.raw |
| Supplementary Figure S9A, Supplementary Figure S10A | JY(*) | Leupeptin, pepstatin, GM | 60% | 60% CH₃CN / 0,1% TFA | 60% | 5% | F1_Agilent6_20161022_TST_C15_Probe1und2ausC13_MAE_HCD_OT_Rep7_18s.raw |
| | | | | | | | F1_Agilent6_20161022_TST_C15_Probe1und2ausC13_MAE_HCD_OT_Rep4_18s.raw |
| | | | | | | | F1_Agilent6_20161023_TST_C15_Probe1und2ausC13_MAE_HCD_OT_Rep9_7B_18s.raw |

(Table continued on next page)
### Supplementary Table S9 (continued)

| Figure / Table / main text | Cells | Additives to MAE elution buffer | Percentage of CH$_3$CN for elution of Oasis HLB column | Rinsing of ultrafilter | Percentage of CH$_3$CN for elution of C$_{18}$ ZipTip | Percentage of sample injected per LC-MS run | MS raw files |
|---------------------------|-------|---------------------------------|---------------------------------|--------------------------|---------------------------------|---------------------------------|----------------|
| **Supplementary Figure S9C,**  
**Supplementary Figure S10B** | JY    | Leupeptin, pepstatin, GM        | 35%                             | 35% CH$_3$CN / 0,1% TFA  | 35%                             | 10%                             | F1_Agilent6 20161122 TST N3 12h flu 10pc Rep1 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 12h flu 10pc Rep2 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 12h flu 10pc Rep3 MAE EsbD OT.raw  
F1_Agilent6 20161122 F1_Agilent6 20161122 TST N3 12h flu 10pc Rep4 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 12h flu 10pc Rep5 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 16h flu 10pc Rep1 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 16h flu 10pc Rep2 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 16h flu 10pc Rep3 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 16h flu 10pc Rep4 MAE EsbD OT.raw  
F1_Agilent6 20161122 TST N3 16h flu 10pc Rep5 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 12h flu 10pc Rep1 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 12h flu 10pc Rep2 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 12h flu 10pc Rep3 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 12h flu 10pc Rep4 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 12h flu 10pc Rep5 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep1 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep2 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep3 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep4 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep5 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep5 MAE EsbD OT.raw  
F1_Agilent6 20170204 TST N3W3 16h flu 10pc Rep5 MAE EsbD OT.raw  |
| **Supplementary Figure S15A-D** | JY    | GM                              | 60%                             | No                       | 35%                             | 10%                             | 150425 ThSt C12 MAE Probe2 10% Rep1 25cm90min3s msms31.raw  
150425 ThSt C12 MAE Probe2 10% Rep2 25cm90min3s msms36.raw  
150425 ThSt C12 MAE Probe2 10% Rep3 25cm90min3s msms40.raw  
JY    | GM,    | 10% DMSO                        | 60%                             | No                       | 35%                             | 10%                             | 150425 ThSt C12 MAE Probe4 10% Rep1 25cm90min3s msms25.raw  
150425 ThSt C12 MAE Probe4 10% Rep2 25cm90min3s msms37.raw  
150425 ThSt C12 MAE Probe4 10% Rep3 25cm90min3s msms44.raw  |
| **Supplementary Figure S18C** | JY    | GM                              | 60%                             | 60% CH$_3$CN / 0,1% TFA, only rinse UF used for indicated MS raw files | 35%                             | 10%                             | 150425 ThSt C12 MAE Probe1 Spuel UF 10% Rep1 25cm90min3s msms27.raw  
150425 ThSt C12 MAE Probe1 Spuel UF 10% Rep2 25cm90min3s msms45.raw  |
| **Supplementary Figure S18D** | THP-1  
(1.5 ∙ 10$^9$ cells) | Leupeptin, pepstatin, EDTA, GM  | 60%                             | Rinse UF not used for indicated MS raw files | 60%                             | 3.3%                            | 150724 ThSt C12 W1 MAE Probe4 1UF 3komma3% Rep1 25cm90min3s msms5.raw  
150724 ThSt C12 W1 MAE Probe4 1UF 3komma3% Rep2 25cm90min3s msms12.raw  
150724 ThSt C12 W1 MAE Probe4 1UF 3komma3% Rep3 25cm90min3s msms15.raw  
THP-1  
(1.5 ∙ 10$^9$ cells from above) | Leupeptin, pepstatin, EDTA, GM  | 60%                             | 60% CH$_3$CN / 0,1% TFA, only rinse UF used for indicated MS raw files | 60%                             | 3.3%                            | 150724 ThSt C12 W1 MAE Probe4 Spuel UF 3komma3% Rep1 25cm90min3s msms6.raw  
150724 ThSt C12 W1 MAE Probe4 Spuel UF 3komma3% Rep2 25cm90min3s msms9.raw  
150724 ThSt C12 W1 MAE Probe4 Spuel UF 3komma3% Rep3 25cm90min3s msms14.raw  |

(Table continued on next page)
### Supplementary Table S9 (continued)

| Figure / Table / main text | Cells                  | Additives to MAE elution buffer | Percentage of CH\textsubscript{3}CN for elution of Oasis HLB column | Percentage of CH\textsubscript{3}CN for elution of C\textsubscript{18} ZipTip | Percentage of sample injected per LC-MS run | MS raw files                                                                 |
|-----------------------------|------------------------|---------------------------------|---------------------------------------------------------------|---------------------------------|---------------------------------------------|--------------------------------------------------------------------------------|
| Main text about repeated MAE from same cells | JY                     | Leupeptin, pepstatin, GM        | 60%                                                          | 60% CH\textsubscript{3}CN / 0,1% TFA                              | 60%                                            | 160111_ThSt_C13_MAE_Probe1_0h_10%_Rep1_25cm90min3s_msms9.raw                   |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe1_0h_10%_Rep2_25cm90min3s_msms11.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe1_0h_10%_Rep3_25cm90min3s_msms13.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe1_3h_10%_Rep1_25cm90min3s_msms10.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe1_3h_10%_Rep2_25cm90min3s_msms12.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe1_3h_10%_Rep3_25cm90min3s_msms14.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe2_0h_10%_Rep1_25cm90min3s_msms15.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe2_0h_10%_Rep2_25cm90min3s_msms17.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe2_0h_10%_Rep3_25cm90min3s_msms19.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe2_3h_10%_Rep1_25cm90min3s_msms16.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe2_3h_10%_Rep2_25cm90min3s_msms18.raw                  |
|                             |                        |                                 |                                                               |                                 |                                              | 160111_ThSt_C13_MAE_Probe2_3h_10%_Rep3_25cm90min3s_msms20.raw                  |
Supplementary Table S10 | MHC-IAC setups varying between experiments and assignment of MHC-IAC results to MS raw files. If not mentioned otherwise, the first ultrafiltrate (UF) and the rinse UF was combined for each sample. The sign (==) indicates that an aliquot of the same cell culture was also subjected to MAE (compare Supplementary Table S9). The asterisk (*) indicates that this sample was separated on a 50 cm column instead of a 25 cm column using a 195 min gradient instead of the usual 90 min gradient as detailed in Supplementary Methods. All measurements were performed at the LTQ Orbitrap XL. Note that some figures and tables are listed in several rows of the first column. MWCO = molecular weight cut-off.

| Figure / Table | Cell line | Cell amount | Elution of MHC-IAC column | Size and MWCO of Amicon ultrafiltration unit | Rinsing of ultrafiltr | Percentage of sample injected per LC-MS run | MS raw files |
|---------------|-----------|-------------|---------------------------|---------------------------------------------|-----------------------|---------------------------------------------|-------------|
| Figure 1, Figure 2B, Figure 3A, Supplementary Figure S6A, Supplementary Figure S16A-C, Table 1, Supplementary Table S1 | JY (==) | 5 \times 10^6 cells | 8 times 0.2% TFA | 0.5 ml, 3 kDa | 50% CH3CN / 0.2% TFA | 10% | 150425_ThSt_C12_W6_Probe5_10%, Rep1_25cm90min3s_msms33.raw 150425_ThSt_C12_W6_Probe5_10%, Rep2_25cm90min3s_msms39.raw 150425_ThSt_C12_W6_Probe5_10%, Rep3_25cm90min3s_msms42.raw |
| Figure 3B, Supplementary Figure S2, Supplementary Figure S6B, Supplementary Figure S7B, Table 1, Supplementary Table S1, Supplementary Table S2 | THP-1 (==) | 1.25 \times 10^7 cells | 8 times 0.2% TFA | 0.5 ml, 3 kDa | 50% CH3CN / 0.2% TFA | 10% | 150808_ThSt_C12-W1_W6_Probe7_10%, Rep1_25cm90min3s_msms8.raw 150808_ThSt_C12-W1_W6_Probe7_10%, Rep2_25cm90min3s_msms16.raw 150808_ThSt_C12-W1_W6_Probe7_10%, Rep3_25cm90min3s_msms25.raw |
| Supplementary Figure S4, Supplementary Figure S17, Supplementary Figure S18A and B, Supplementary Table S1 | LCL5 | 2.5 ml pellet ≈ 1.25 \times 10^7 cells for each of the two samples | 8 times 0.2% TFA | 4 ml, 10 kDa | Rinse UF not used for indicated MS raw files | 20% | 140411_TS_LCL5_C9-W1-1a_W_20%, Rep9_25cm90min3s_IncludeMonomer_msms2.RAW 140411_TS_LCL5_C9-W1-1a_W_20%, Rep10_25cm90min3s_IncludeMonomer_msms12.RAW 140411_TS_LCL5_C9-W1-1a_W_20%, Rep15_25cm90min3s_IncludeMonomer_msms14.RAW 140411_TS_LCL5_C9-W1-1a_W_20%, Rep23_25cm90min3s_IncludeMonomer_msms23.RAW |
| | | | 8 times 0.2% TFA | 4 ml, 10 kDa | 80% CH3CN / 0.2% TFA, only rinse UF used for indicated MS raw files | 20% | 140411_TS_LCL5_C9-W1-1b_W_20%, Rep1_25cm90min3s_IncludeMonomer_msms3.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep9_25cm90min3s_IncludeMonomer_msms13.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep15_25cm90min3s_IncludeMonomer_msms22.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep23_25cm90min3s_IncludeMonomer_msms24.RAW |
| | | | 8 times 0.2% TFA | 4 ml, 10 kDa | 80% CH3CN / 0.2% TFA, only rinse UF used for indicated MS raw files | 20% | 140411_TS_LCL5_C9-W1-1b_W_20%, Rep9_25cm90min3s_IncludeMonomer_msms3.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep15_25cm90min3s_IncludeMonomer_msms22.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep23_25cm90min3s_IncludeMonomer_msms24.RAW |
| | | | 8 times 0.2% TFA | 4 ml, 10 kDa | 80% CH3CN / 0.2% TFA, only rinse UF used for indicated MS raw files | 20% | 140411_TS_LCL5_C9-W1-1b_W_20%, Rep9_25cm90min3s_IncludeMonomer_msms3.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep15_25cm90min3s_IncludeMonomer_msms22.RAW 140411_TS_LCL5_C9-W1-1b_W_20%, Rep23_25cm90min3s_IncludeMonomer_msms24.RAW |
| Supplementary Figure S5 | JY | 5 \times 10^6 cells | 8 times 0.2% TFA | 0.5 ml, 10 kDa | 50% CH3CN / 0.2% TFA | 10% | 150425_ThSt_C12_LTue_Probe5_10%, Rep1_25cm90min3s_msms16.raw 150425_ThSt_C12_LTue_Probe5_10%, Rep2_25cm90min3s_msms19.raw 150425_ThSt_C12_LTue_Probe5_10%, Rep3_25cm90min3s_msms22.raw |
| | | | 8 times 0.2% TFA | 0.5 ml, 10 kDa | 50% CH3CN / 0.2% TFA | 10% | 150813_ThSt_C12-W1_LTue_Probe7_10%, Rep1_25cm90min3s_msms8.raw 150813_ThSt_C12-W1_LTue_Probe7_10%, Rep2_25cm90min3s_msms12.raw 150813_ThSt_C12-W1_LTue_Probe7_10%, Rep3_25cm90min3s_msms16.raw |
| Supplementary Figure S16A-C | JY | 5 \times 10^6 cells | 4 times 0.2% TFA, 4 times 80% CH3CN / 0.2% TFA | 0.5 ml, 3 kDa | 50% CH3CN / 0.2% TFA | 10% | 150425_ThSt_C12_W6_Probe6_10%, Rep1_25cm90min3s_msms34.raw 150425_ThSt_C12_W6_Probe6_10%, Rep2_25cm90min3s_msms38.raw 150425_ThSt_C12_W6_Probe6_10%, Rep3_25cm90min3s_msms43.raw |

(Table continued on next page)
| Figure / Table       | Cell line | Cell amount | Elution of MHC-IAC column | Size and MWCO of Amicon ultrafiltration unit | Rinsing of ultrafilter | Percentage of sample injected per LC-MS run | MS raw files                                                                 |
|----------------------|-----------|-------------|---------------------------|---------------------------------------------|------------------------|-------------------------------------------|--------------------------------------------------------------------------------|
| Supplementary Figure S16D-F | THP-1     | 5 - 10⁸ cells | 8 times 0.2% TFA          | 0.5 ml, 3 kDa                               | 50% CH₃CN / 0.2% TFA   | 10%                                       | 150808_ThSt_C12-W1_W6_Probe5_10%_Rep1_25cm90min3s_msms12.raw                 |
|                      |           |             |                           |                                             |                        |                                           | 150808_ThSt_C12-W1_W6_Probe5_10%_Rep2_25cm90min3s_msms12.raw                 |
|                      |           |             |                           |                                             |                        |                                           | 150808_ThSt_C12-W1_W6_Probe5_10%_Rep3_25cm90min3s_msms19.raw                |
|                      | THP-1     | 5 - 10⁸ cells | 4 times 0.2% TFA, 4 times 60% CH₃CN / 0.2% TFA | 0.5 ml, 3 kDa                               | 50% CH₃CN / 0.2% TFA   | 10%                                       | 150808_ThSt_C12-W1_W6_Probe6_10%_Rep1_25cm90min3s_msms7.raw                 |
|                      |           |             |                           |                                             |                        |                                           | 150808_ThSt_C12-W1_W6_Probe6_10%_Rep2_25cm90min3s_msms14.raw                |
|                      |           |             |                           |                                             |                        |                                           | 150808_ThSt_C12-W1_W6_Probe6_10%_Rep3_25cm90min3s_msms22.raw                |
| Table 1, Supplementary Table S1 | JY        | 20 ml pellet ≈ 10¹⁰ cells | 8 times 0.2% TFA          | 4 ml, 3 kDa                                | 50% CH₃CN / 0.2% TFA   | 0.5% (*)                                  | 150604_ThSt_C12_W6_Probe7_0komma5%_Rep2_50cm195min3s_msms31.raw              |
| Supplementary Table S1 | THP-1     | 5 ml pellet ≈ 2.5 - 10⁹ cells | 8 times 0.2% TFA          | 4 ml, 3 kDa                                | 50% CH₃CN / 0.2% TFA   | 2%                                        | 150808_ThSt_C12-W1_W6_Probe8_2%_Rep1_25cm90min3s_msms1.raw                  |
|                      |           |             |                           |                                             |                        |                                           | 150808_ThSt_C12-W1_W6_Probe8_2%_Rep2_25cm90min3s_msms2.raw                  |
Supplementary References

72. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637 (1983).

73. Touw, W. G. *et al.* A series of PDB-related databanks for everyday needs. *Nucleic Acids Res* **43**, D364-368 (2015).

74. Hamelryck, T. & Manderick, B. PDB file parser and structure class implemented in Python. *Bioinformatics* **19**, 2308-2310 (2003).

75. Tien, M. Z., Meyer, A. G., Sydykova, D. K., Spielman, S. J. & Wilke, C. O. Maximum allowed solvent accessibilities of residues in proteins. *PLoS One* **8**, e80635 (2013).

76. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132 (1982).