

A Platform for High-throughput and Ultrasensitive Immunopeptidomics

Authors

Adillah Gul, Laura Van Moortel, Patrick Willems, Ilke Aernout, Laura Pedró-Cos, Kia C. Ferrell, Katie Boucher, An Staes, Simon Devos, Ine Lentacker, Bart Vandekerckhove, Caroline Demangel, Fabien Thery, and Francis Impens

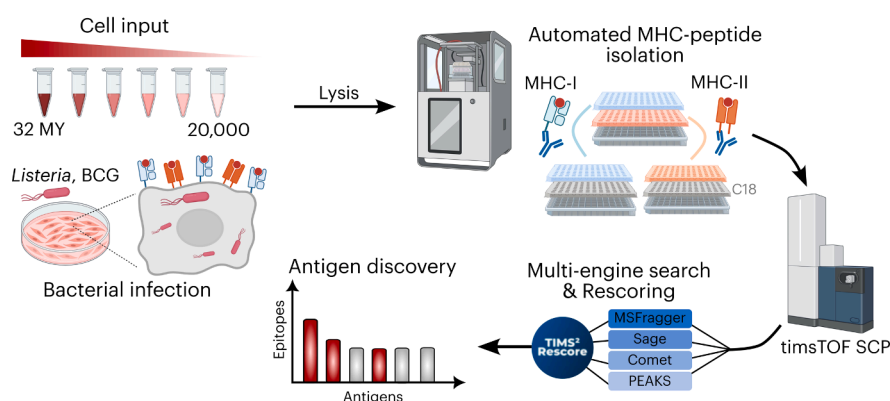
Correspondence

fabien.thery@vib-ugent.be;
francis.impens@vib-ugent.be

Graphical Abstract

In Brief

Immunopeptidomics enables untargeted discovery of peptides presented by MHC molecules, informing vaccine and immunotherapy development. Existing workflows focus either on throughput or sensitivity, while ideally both are combined. We describe a semiautomated, 96-well platform for high throughput MHC class I and II peptide isolation from low-input samples. The workflow identifies hundreds of MHC peptides from as few as 20,000 cells and successfully detects bacterial immunopeptides from *Listeria*- and BCG-infected macrophages, enabling biological applications where sample availability is limited.



Highlights

- A semiautomated 96-well platform for low-input isolation of MHC-I and -II peptides.
- Identification of hundreds of MHC-I peptides from 20,000 cells.
- Quantitative detection of bacterial immunopeptides from infected macrophages.

A Platform for High-throughput and Ultrasensitive Immunopeptidomics

Adillah Gul^{1,2,†}, Laura Van Moortel^{1,2,†}, Patrick Willems^{1,2,†}, Ilke Aernout^{3,4}, Laura Pedró-Cos⁵, Kia C. Ferrell⁵, Katie Boucher^{1,2,6}, An Staes^{1,2,6}, Simon Devos^{1,2,6}, Ine Lentacker^{3,4}, Bart Vandekerckhove^{4,7,8}, Caroline Demangel⁵, Fabien Thery^{1,2,*}, and Francis Impens^{1,2,*}

Mass spectrometry-based immunopeptidomics is a powerful approach for untargeted discovery of peptides presented on MHC molecules, which can guide the selection of vaccine antigens and immunotherapy targets. First-generation immunopeptidomics workflows require processing of hundreds of millions of cells using lengthy, manual procedures. More recent approaches focus on increasing either sensitivity or throughput, but rarely combine both aspects. Here, we describe a semiautomated immunopeptidomics platform that combines high sensitivity with high throughput by implementing highly optimized conditions for isolation of MHC class I and II peptides in a 96-well positive pressure device. Lysis in a small volume of 100 μ l allows efficient MHC capture in a 96-well filter plate with optimal pore size for automated washing, elution, and C18 purification steps. Upon analysis of 25% of the eluate from 16 million cells, our workflow identified over 13,500 MHC I and 6000 MHC II peptides on a timsTOF single-cell proteomics mass spectrometer, operating in data-dependent acquisition-parallel accumulation-serial fragmentation mode. Exploring the sensitivity limits of our platform, we identified up to 1000 MHC I peptides, including hundreds of predicted binders, from as few as 20,000 JY cells. Validating the platform's performance for quantitative biological discovery, we report the identification of known and novel bacterial immunopeptides from U937 macrophages infected with *Listeria monocytogenes* or bacillus Calmette-Guérin. Together, our optimized immunopeptidomics platform enables robust immunopeptide detection from lower-input samples in a high-throughput fashion, enabling its use for biological applications where sample amounts are limiting.

The immunopeptidome is the repertoire of peptides presented by major histocompatibility complex (MHC) molecules, also known as human leukocyte antigens (HLAs) in

humans, on the surface of cells (1). These MHC-bound peptides, called immunopeptides, are derived from self and nonself proteins processed inside the cell (2). Human MHC class I molecules present peptides that are 8- to 12-amino acids in length that result from proteasomal degradation of intracellular proteins. MHC class II molecules typically present 13- to 18-mer peptides generated by the uptake and lysosomal degradation of extracellular proteins. Peptides presented on MHC-I and MHC-II are recognized by CD8⁺ and CD4⁺ T cells, respectively. Therefore, the immunopeptidome plays a critical role in adaptive immunity, facilitating recognition of peptides derived from pathogens, tumor antigens, and aberrantly expressed self-proteins by T cells (3–6). Over the last decade, mass spectrometry (MS)-based immunopeptidomics has emerged as a powerful analytical tool for comprehensive profiling of the immunopeptidome (2, 7), enabling the discovery of bacterial antigens presented during infection (8–16) or tumor-specific epitopes that can be used for immunotherapy (17–21), among other applications.

A typical MS-based immunopeptidomics workflow consists of three steps. It starts with sample preparation, which comprises the pulldown of MHC-peptide complexes and elution of purified immunopeptides. This is followed by LC-MS/MS analysis, where mass spectra of the intact and fragmented immunopeptides are recorded. Finally, the acquired MS data is searched by dedicated software tools to assign spectra to peptide sequences (22–25). Although all three steps influence the quality of the final results, robust sample preparation is key as it directly influences the net peptide yield and overall reproducibility. This step is the most time-consuming due to the complexity of immunopeptide isolation and is often considered a major bottleneck for the clinical implementation of immunopeptidomics pipelines (26). Sample preparation begins with mild lysis of cells or tissues to

From the ¹VIB Center for Medical Biotechnology, VIB, Ghent, Belgium; ²Department of Biomolecular Medicine, Ghent University, Ghent, Belgium; ³Ghent Research Group on Nanomedicines, Ghent University, Ghent, Belgium; ⁴Cancer Research Institute Ghent (CRIG), Ghent, Belgium; ⁵Institut Pasteur, Université Paris Cité, Inserm U1224, Immunobiology and Therapy Unit, Paris, France; ⁶VIB Proteomics Core, VIB, Ghent, Belgium; ⁷Department of Diagnostic Sciences, Ghent University, Ghent, Belgium; ⁸GMP Unit Cell Therapy, Ghent University Hospital, Ghent, Belgium

[†]These authors contributed equally to this work.

*For correspondence: Francis Impens, francis.impens@vib-ugent.be; Fabien Thery, fabien.thery@vib-ugent.be.

extract MHC-peptide complexes in their native state. Immunoprecipitation (IP) is the preferred method to isolate these complexes, using specific antibodies for MHC-I and MHC-II. Once isolated, immunopeptides are eluted from the MHC molecules under mild acidic conditions, followed by a desalting step to purify the peptides for LC-MS/MS analysis (22, 27, 28). An alternative, antibody-free approach for MHC-I peptide isolation is mild acid elution (MAE) directly from the cell surface (29–31).

In recent years, technological advances have revolutionized the field of immunopeptidomics, allowing high-throughput extraction and analysis of immunopeptides. Innovations in MS instrumentation and data acquisition have greatly improved the sensitivity of immunopeptidomics workflows. This includes for instance trapped ion mobility spectrometry (TIMS) (28, 32, 33), data-independent acquisition (29, 34–36), and advanced data analysis by incorporating multiple search engines and/or data-driven rescoring (32, 33, 37, 38). Consequently, this has facilitated the evolution of sample preparation to higher throughput workflows on lower input samples. Initial sample preparation workflows, despite showing excellent peptide enrichment, required hundreds of millions of cells as starting material and depended on manual IP of immunopeptides in tubes or columns (15, 22, 25, 39). These workflows were laborious, time-consuming, and prone to variability, limiting their utility for large sample sets or scarce primary specimens. Since then, IP procedures have significantly improved and upgraded from individual column with large sample amounts to multi-well plate format compatible with lower input (26, 40–42). Together with increased instrumental and computational sensitivity, such miniaturized sample preparation workflows have overcome the challenge of performing immunopeptidomics on samples with limited material such as scarce tumor sections, lymph node biopsies, and rare cell populations.

At the same time, automation in sample preparation has remarkably enhanced overall throughput, reproducibility and sensitivity by reducing sample loss and contamination risks (26, 28, 40–46). The use of 96-well plates in immunopeptidomics sample preparation was first introduced by Chong *et al.* (26) in 2018, using stacked 96-well filter plates in combination with a positive pressure device for sequential MHC-I and MHC-II IP on tissue and cell lysates (15, 25). Since then, several other advancements in immunopeptidomics sample preparation have been reported (19, 27, 29–32), focusing on further optimizing throughput in 96-well plate workflows (28, 44) or maximizing sensitivity for low-input samples, for instance, using microfluidics (42, 43, 45, 46). However, most workflows today focus on either throughput or sensitivity, while ideally, the combination of both should be achieved. To fill this gap, we have developed a refined sample preparation workflow that is fully optimized to process low-input and crude samples in 96-well plates, aiming to combine high throughput with maximal sensitivity. We implemented a miniaturized and automated workflow in a 96-well format using a

Tecan Resolvex A200 liquid handling positive pressure device. Different from previous 96-well procedures, MHC pulldown occurs in a small lysate volume of 100 μ l under static conditions, directly in a 96-well filter plate with optimized pore size. We demonstrate sensitive detection of immunopeptidomics from 32 million to 0.5 million human cells, with the ability to recover nearly 300 predicted MHC-I binders from just 20,000 JY cells. We showcase how this miniaturized platform facilitates bacterial antigen discovery for *Listeria monocytogenes* and the tuberculosis vaccine strain bacillus Calmette-Guérin (BCG) in a quantitative manner.

EXPERIMENTAL PROCEDURES

Cell Lines

The Epstein–Barr virus-immortalized human B cell line JY (ECACC 94022533), human HeLa cells (ECACC 93021013), and human U937 (ATCC) cells were cultured at 37 °C in a humidified atmosphere at 5% CO₂. JY and U937 cells were maintained in RPMI 1640 medium with GlutaMax (#61870036, Thermo Fisher Scientific), supplemented with 10% heat-inactivated fetal bovine serum (FBS, #10270106, Thermo Fisher Scientific) without antibiotics. U937 cells were differentiated using phorbol 12-myristate 13-acetate (PMA) at 200 nM final concentration for 48 h. HeLa cells were grown without antibiotics in DMEM high glucose medium with GlutaMax and pyruvate (#12077549, Thermo Fisher Scientific) supplemented with 10% FBS. All the cells were tested negative for *mycoplasma* contamination. Cells were counted and volumes corresponding to the number of cells in the serial dilution experiment were collected by centrifugation at 300 $\times$ g, washed twice with ice-cold PBS and stored at –80 °C.

Bacterial Culture

A defrosted aliquot of BCG Pasteur strain was grown in Difco Middlebrook 7H9 broth Medium (#271310, Becton Dickinson) supplemented with 10% oleic albumin dextrose catalase, 0.2% glycerol, and 0.05% Tween 80 (Sigma-Aldrich, Merck), shaking at 37 °C. *L. monocytogenes* EGD (BUG600 strain) was grown in brain heart infusion broth (#10462498, Thermo Fisher Scientific), shaking at 37 °C. All the BCG and *Listeria* work was carried out aseptically in a biosafety level 2 facility.

BCG and *Listeria* Infection of U937 Cells

U937 cells were differentiated into phagocytic macrophage-like cells by stimulation with 200 nM PMA for 48 h, followed by 24 h recovery in PMA-free media. BCG was cultured into the log phase and pelleted by centrifugation at 3900 $\times$ g for 7 min. Bacteria were washed once with PBS, resuspended in PBS, and filtered through 25, 26, and 27 gauge needles to acquire a single-cell suspension. Differentiated U937 cells were infected at a multiplicity of infection of 20 for 3 h, and the medium was aspirated to remove extracellular BCG and after which fresh media was added. Infected and uninfected cells were harvested with a cell scraper after 24 h, washed two times with PBS, and dry cell pellets were stored at –80 °C.

Listeria was grown overnight and diluted in brain heart infusion broth to a density of 1e9 bacteria/ml, washed twice with PBS and resuspended in U937 growth medium without FBS. U937 cells were washed two times with PBS prior to 1 h infection with *Listeria* at an multiplicity of infection of 25. Following infection, the cells were washed twice with PBS and further cultured for 23 h in cell culture medium supplemented with 10% FBS and 40 μ g/ml gentamicin

(#G1397, Sigma-Aldrich, Merck) to eliminate extracellular bacteria. Cells were harvested using a cell scraper, washed two times with PBS, and dry cell pellets were stored at -80°C .

Generation of Immunoaffinity Columns for MHC Class I and II Pulldown

W6/32 monoclonal pan-MHC class I and PdV5.2 monoclonal pan-MHC class II antibodies were purified from HB95 (HB-95, ATCC) and PdV5.2 (kindly provided by Bart Vandekerckhove, Ghent University (47)) hybridoma cells, respectively. To prepare the W6/32 immunoaffinity column, 3 mg of antibody was diluted in 5 ml tris-buffered saline and added to 1 ml pre-washed protein-A Sepharose 4B packed beads (#101041, Thermo Fisher Scientific) in glass Econo-columns (#7374150, Bio-Rad). To prepare the PdV5.2 immunoaffinity column, 1.5 mg of antibody was diluted in 5 ml tris-buffered saline and added to protein-G Sepharose 4B packed beads (#101242, Invitrogen). Antibodies were incubated at room temperature for 1 h on an lka rolling tube mixer device. Each column was washed with 0.2 M sodium tetraborate buffer (pH 9, #B3545, Sigma-Aldrich, Merck) followed by 40 min chemical crosslinking using 20 mM dimethylpimelimidate (#D8388, Sigma-Aldrich, Merck), freshly dissolved in sodium tetraborate buffer. After crosslinking, antibody-bead complexes were washed and incubated for 2 h on an lka mixer with 5 mL 0.2 M ethanolamine (pH 8, #149582500, Thermo Fisher Scientific) to quench the crosslinking reaction. A more detailed description of the procedure can be found in the [Supplemental Methods](#).

High-throughput Isolation and Purification of MHC-I and MHC-II Peptides

Cell Lysis—Cell pellets ranging from 0.5 to 32 million cells were lysed in 0.1 ml mild lysis buffer containing 1% octyl β -D-glucopyranoside (#O9882, Sigma-Aldrich, Merck), 0.25% sodium deoxycholate (#1065040250, Millipore, Merck), 1.25 $\times$ cOmplete protease inhibitor cocktail (#4693159001, Roche), 1 mM PMSF (#52332, Sigma-Aldrich, Merck), 0.2 mM iodoacetamide (#11149, Sigma-Aldrich, Merck), and 1 mM EDTA (#EDS, Sigma-Aldrich, Merck) in 150 mM NaCl, 50 mM Tris-HCl (pH 8.0). Cells were lysed for 1 h on ice while resuspending every 15 min, then centrifuged at $20,000 \times g$ at 4°C for 10 min to remove cell debris, followed by a second centrifugation at $20,000 \times g$ at 4°C for 30 min to further clear the lysate.

Preparation of Filter Plates for MHC-I and II Pulldown—We utilized a Resolvex A200 positive pressure processor (Tecan) to automate the washing and elution steps. Subsequent IPs of MHC-I and MHC-II complexes were performed using a single-use 96-well filter microplate (0.7 μm glass fiber membrane, 2 ml/well, long drip Agilent [part #201719-100]). Approximately 5 to 10% low positive pressure was applied at each step to maintain a uniform liquid flow through the plates. A more detailed description of the procedure can be found in the [Supplemental Methods](#).

Before IP, each well was activated with 1 ml 100% acetonitrile (ACN) (Sigma-Aldrich) twice, followed by three washes with 1 ml 0.1% TFA and finally three times with 1 ml 150 mM NaCl 50 mM Tris-HCl (pH 8.0). Next, the MHC-I and MHC-II crosslinked beads were loaded on their respective plates (further referred to as “MHC-I filter” and “MHC-II filter” plates, [Fig. 1A](#)). Beads were washed five times with 1 ml 150 mM NaCl 50 mM Tris-HCl (pH 8.0).

Subsequent MHC-I and MHC-II IP—Lysates were added to the MHC-I filter plate and incubated for 1 h at 4°C to capture MHC-I complexes. Next, the unbound lysate fraction of the MHC-I IP was collected using the Tecan and loaded on the MHC-II filter plate, followed by 1 h incubation at 4°C . Meanwhile, the MHC-I filter plate was washed five times with 1 ml 150 mM NaCl in 50 mM Tris-HCl

(pH 8.0) using the Tecan. After 1 h incubation, the MHC-II filter plate was washed similarly.

Immunopeptide Desalting and Elution—Sep-Pak tC18 plates (#186002321, Waters) were activated with two times 1 ml 100% ACN and three times 1 ml Milli-Q water containing 0.1% TFA, using approximately 1 to 3% high positive pressure on the Resolvex A200 to maintain a uniform liquid flow throughout the plates. Each MHC filter plate was stacked on top of a Sep-Pak plate ([Fig. 1A](#)) and MHC-peptide complexes were eluted with 0.2 ml 10% acetic acid five times. Next, we removed the filter plate and washed the Sep-Pak plate three times with 1 ml of 0.1% TFA. Thereafter, a collection plate (#201240-100, Agilent) was mounted under the Sep-Pak tC18 plate. MHC-I peptides were eluted three times with 25% ACN/0.1% TFA, while MHC-II peptides were eluted with 40% ACN/0.1% TFA. The eluted immunopeptides were then transferred to protein LoBind tubes (#0030108450, Eppendorf), vacuum dried and stored at -20°C until LC-MS/MS injection.

LC-MS/MS Analysis

Dried peptides were dissolved in 50 μl 0.1% ACN/0.1% TFA from which 12.5 μl was injected for LC-MS/MS analysis on a Ultimate 3000 RSLC nanoLC in-line connected to a timsTOF single-cell proteomics (SCP) mass spectrometer (Bruker). Of note, the dried peptides from low input dilution series ([Fig. 2](#)) were dissolved in 17 μl 0.1% ACN/0.1% TFA from which 15 μl was injected for LC-MS/MS analysis. Trapping was conducted at a flow rate of 20 $\mu\text{l}/\text{min}$ for 2 min in loading solvent A on a 5 mm trapping column (Thermo Fisher Scientific, Pepmap, 300 μm internal diameter [I.D.], 5 μm beads). The sample was then separated on a reverse-phase column (Aurora elite 75 $\mu\text{m} \times 150$ mm 1.7 μm particles, IonOpticks) following elution from the trapping column. Separation was achieved using a linear gradient, starting at 0.5% MS solvent B (0.1% FA in water/ACN 20:80, v/v) at 250 nl/min for 30 min, increasing to 37.5% MS solvent B, then reaching 55% at 38 min, and further rising to 70% at 40 min. This was followed by a wash for 5 min and reequilibration with 99.5% MS solvent A (0.1% FA in water). The flow rate was adjusted from 250 nl/min to 100 nl/min at 20 min and restored to 250 nl/min at 40 min. A 10 parallel accumulation-serial fragmentation (PASEF)/MSMS scan acquisition method was employed in data-dependent acquisition (DDA)-PASEF mode with a precursor signals intensity threshold of 500 arbitrary units. Precursors were isolated with a 2 Th window below m/z 700 and 3 Th above, with active exclusion for 0.4 min upon reaching a target intensity threshold of 20,000 arbitrary units. In case of MHC-I peptides, an extended IM range (0.75–1.65 Vs/cm²) was applied to include HLA-I singly charged precursors, whereas a more narrow IM range (range 0.7–1.25 Vs/cm²) was used for MHC-II peptides ([Supplemental Fig. S1](#)). The mass spectrometer was operated in nonsensitive mode with an accumulation and ramp time of 100 ms, analyzing in MS from 100 to 1700 m/z , applying collision energy according to the IM range specified in [Supplemental Table S1](#).

Data Analysis and Processing

Immunopeptidomics Searches and Analysis—Raw proteomics data were searched against the human reference proteome (August 2024, UP00005640, 20,666 proteins) using a multi-search engine workflow with follow-up rescoring as described in Willems *et al.* (33). In case of U937 infected cell experiments, *L. monocytogenes* strain EGD (UP000016703, 2847 proteins) and *Mycobacterium bovis* strain BCG/Pasteur (UP000001472, 3891 proteins) proteomes were appended to the human proteome for database searching. In brief, timsTOF (.d) data was searched by four search engines (MSFragger (48), Comet (49), Sage (50), and PEAKS Studio 12 (51)) after which search results for each engine were rescored in parallel using TIMS²Rescore (52). Peptides identified below a 1% peptide q-value were reconciled at the

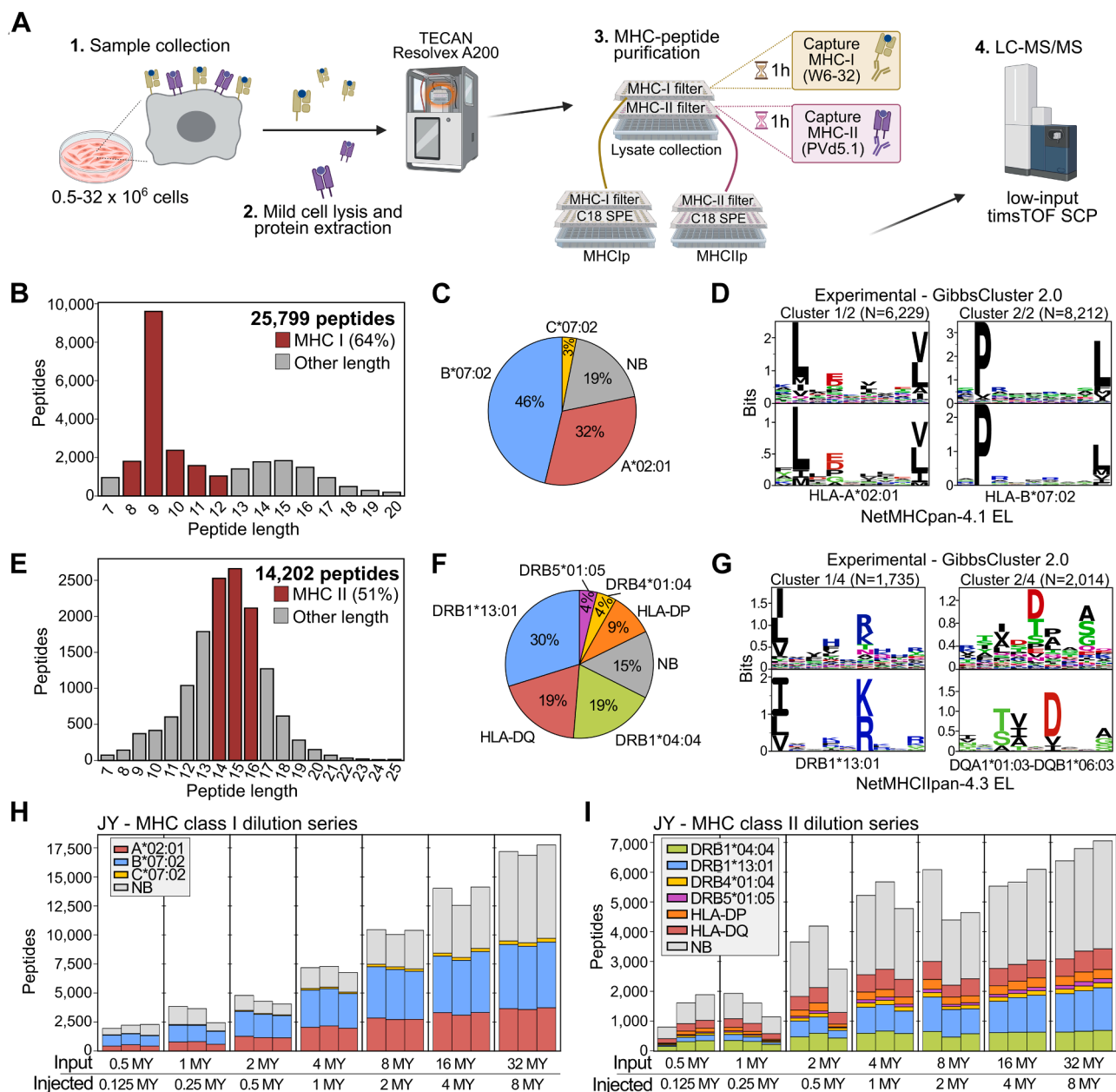


FIG. 1. A sensitive semiautomated workflow for high-throughput immunopeptidomics. A, general overview of the workflow. Cells are lysed in mild conditions, after which proteins are extracted and loaded on 96-well plates with pre-washed anti-MHC immunoaffinity beads. Sequential MHC-I and MHC-II IP, elution and desalting are performed in a semiautomated fashion on a Resolvex A200 (Tecan), after which peptides are collected and 25% is injected for LC-MS/MS on a timsTOF SCP instrument. Figure created with [BioRender.com](#). B, histogram showing the length distribution of peptides identified after MHC class I pulldown (peptide Q-value ≤ 1%). Peptides compatible with MHC-I binding (length 8–12) are indicated in red, nonbinding (NB) peptides in gray. C, proportion of identified 8 to 12 mers predicted to bind to the indicated HLA alleles by NetMHCpan-4.1 (53). D, unsupervised Gibbs clustering (55) of identified 8 to 12 mers resulting in sequence logos that match MHC class I eluted ligand (EL) motifs of NetMHCpan-4.1 (53). E, histogram showing the length distribution of peptides identified after MHC class II pulldown and elution. Peptides compatible with MHC-II binding (length 14–16) are indicated in red, NB peptides in gray. F, proportions of identified 14 to 16 mers predicted to bind to the indicated HLA alleles by NetMHCIIpan-4.3 (54). For HLA-DQ and HLA-DP, the possible α-β chain pairings (e.g. HLA-DQA10103-DQB10603) were grouped together. G, unsupervised Gibbs clustering (55) of identified 14 to 16 mers resulted in sequence logos that match MHC class II eluted ligand (EL) motifs of NetMHCIIpan-4.3 (54). H and I, number of identified peptide sequences for individual replicates, grouped by cell input amount. Samples from each amount were processed and searched separately for these graphs. Peptides are colored according to binding prediction for JY MHC class I alleles (H) and class II alleles (I), and NB peptides in gray. HLA, human leukocyte antigen; IP, immunoprecipitation; MS, mass spectrometry; NB, nonbinding; SCP, single-cell proteomics.

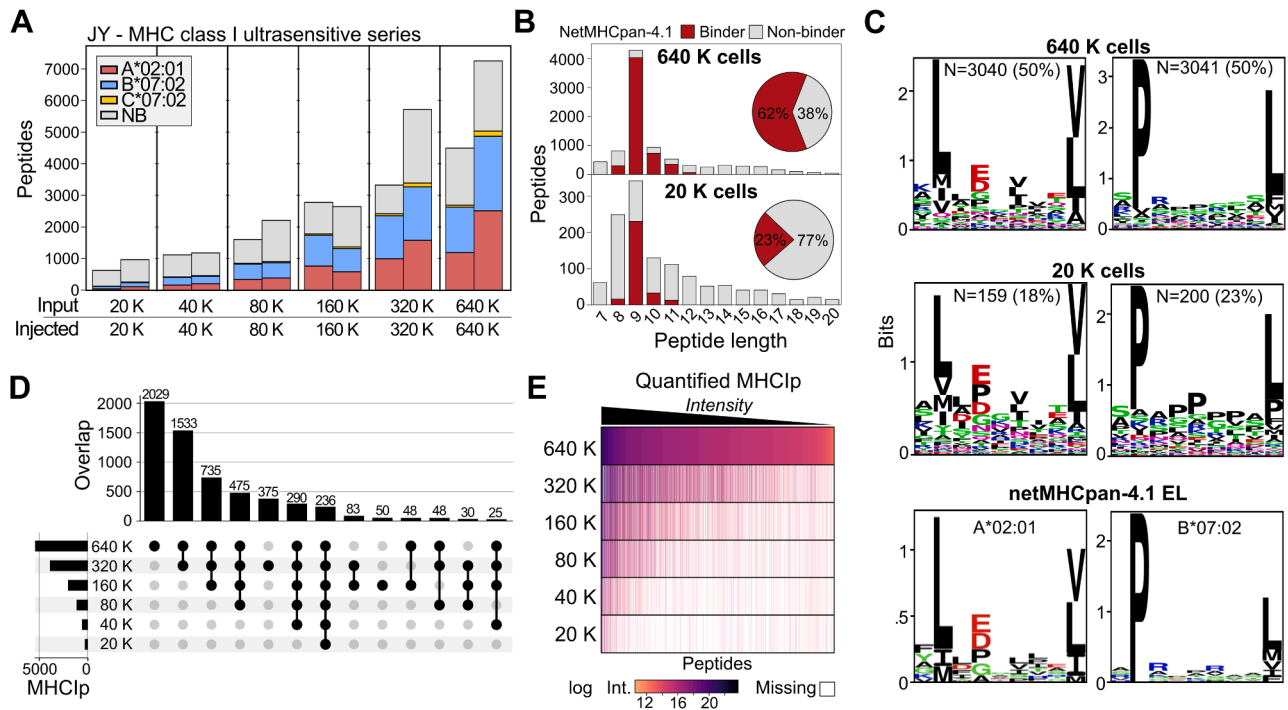


FIG. 2. Toward ultrasensitive immunopeptidomics. A, number of identified peptide sequences per replicate, grouped per cell input amount from 640,000 (640 K) down to 20,000 cells (20 K). Samples from each input amount were processed and searched independently, and 100% of the isolated immunopeptides were analyzed by LC/MS-MS. Peptides are colored according to MHC binding prediction by NetMHCpan-4.1 (53). B, histogram showing the length distribution of peptides identified after MHC class I pulldown and elution from 640 K and 20 K cells. Peptides compatible with MHC-I binding (length 8–12) are indicated in red. C, unsupervised Gibbs clustering (55) of identified class I peptides from 640 K and 20 K cells (top and middle) resulted in sequence logos that match NetMHCpan-4.1 eluted ligand (EL) sequence logos for HLA-A*02:01 and HLA-B*07:02 (bottom). D, UpSet plot showing the overlap of identified predicted MHC-I peptides (MHCip, NetMHCpan-4.1 %Rank < 2) across the different input amounts. Only intersections of more than 10 peptides are displayed. E, heatmap showing $\log_2$ intensities of all peptides identified in the 640 K dilution, sorting peptides from highest to lowest intensity in this dilution (left to right). Missing values are colored in white. Cell input amounts were searched separately using the FragPipe “Nonspecific-HLA” workflow. HLA, human leukocyte antigen; MS, mass spectrometry.

spectrum-level, excluding the rare cases where a spectrum was matched to distinct peptides by different search engines (33). Downstream immuno-informatic analyses included NetMHCpan-4.1 (53) and NetMHCippan-4.3 (54) prediction of binding strength for 8–12-mers and 14–16-mers, respectively, and unsupervised Gibbs clustering (55). Default NetMHCpan %Rank score thresholds were used to define strong and weak MHC class I binders (below 0.5 and 2, respectively) and MHC class II binders (below 2 and 5, respectively).

Differential Analysis—To analyze differential immunopeptide abundances between infected and uninfected samples, we used FragPipe (version 22.0) including MSBooster (38) and IonQuant quantification (56). Resulting peptide intensities (from “combined peptides.tsv” reports) were used for peptide-level differential statistics. To assess differential peptide abundance, we resorted to an empirical Bayes test in limma (57) following recommendations of a statistical benchmark (58). Significantly regulated peptides were filtered at an adjusted p value ≤ 0.05 and absolute fold change ≥ 2 . This differential test does not impute missing values, with the drawback of not reporting peptides exclusively detected in one of the conditions (infected/uninfected). Therefore, we performed an additional differential detection analysis where we discerned peptides/proteins uniquely identified in one condition (three out of four replicates) and absent in the other condition.

Gene Set Enrichment Analysis—For the differential immunopeptides identified during infection, their corresponding UniProtKB

accessions were used for protein-level gene set enrichment analysis. In the case a peptide matched multiple proteins, the first UniProtKB accession was used as representative. Accessions were entered in the online webtool g:Profiler using default settings (59).

Visualization—Plots were generated using matplotlib (60) and seaborn (61) in Python (version 3.8.10). Sequence logos were generated using Logomaker (62). UpSet plots were made using the Python implementation module upsetplot (63).

Experimental Design and Statistical Rationale

Immunopeptidomics optimization experiments (Supplemental Figs. S2–S6) were conducted in technical IP replicates, with a single 25% injection of each sample on a timsTOF SCP instrument. Immunopeptidomics benchmarking experiments using JY, HeLa, and U937 cells (Fig. 1 and Supplemental Figs. S7–S11) were performed with biological triplicates, starting from individual cell pellets for each replicate and cell number. 25% of each sample was injected once on the timsTOF SCP, going from low to high input to minimize sample carry-over. Ultrasensitive immunopeptidomics benchmarking experiments using JY cells (Fig. 2 and Supplemental Figs. S12–S16) were performed in two biological replicates, starting from individual cell pellets for each replicate and each cell number. 100% of each sample was injected on the timsTOF SCP instrument, going from low to high input. *L. monocytogenes* infection experiments were performed in

three technical replicates (Fig. 3 and Supplemental Fig. S17), while BCG infection experiments were conducted in four biological replicates (Fig. 3 and 4 and Supplemental Figs. S17–18). For these experiments, 25% of each sample was injected once on the timsTOF SCP using a randomized block design, and significantly regulated peptides were defined at an adjusted p value ≤ 0.05 and absolute fold change ≥ 2 using an empirical Bayes test in limma (57) following recommendations of a statistical benchmark (58). Also peptides reproducibly detected in three out of four biological replicates, but absent in the other condition were defined as significantly regulated.

RESULTS

An Automated Workflow for High-throughput Immunopeptidomics

Our previous immunopeptidomics workflow required approximately 500 million cells as input for manual IP in large columns (8, 15), thus requiring large sample input amounts with limited throughput. Here, we aimed to miniaturize and automate sample preparation by reducing the input amount 10- to 50-fold while increasing sample throughput by using a 96-well format. We initially tested optimal conditions for low-input cell lysis, IP and elution on JY cells, an Epstein–Barr virus-immortalized human B cell line, and found that static IP at high protein concentration (small lysate volume) with elution in 10% acetic acid led to the highest number of identified MHC-I and II peptides (Supplemental Figs. S2–S5). This optimized procedure was then automated on a Tecan Resolvex A200 positive pressure device, using 96-well filter plates with a pore size of 0.7 μm preventing flow-through by gravity, but allowing a controlled and consistent flow through the wells when positive pressure is applied. In the resulting workflow, cell pellets are lysed in 100 μl and cleared lysates are loaded on 96-well filter plates for sequential pulldown of MHC-I and II complexes, directly on the plate under static conditions. Following MHC capture, the flow-through is pushed through and multiple wash steps are applied by automated buffer dispensing by the Tecan. Immunopeptides are eluted under acidic conditions and collected in a 96-well Sep-Pak C18 plate for further purification, also on the Tecan (Fig. 1A). Performing MHC-II after MHC-I pulldown was necessary to avoid coenrichment of MHC-I peptide complexes during MHC-II IP (Supplemental Fig. S6). We also found that the pan-MHC-II PdV5.2 antibody shows similar performance to other MHC class II antibodies (TU39 or IVA12, Supplemental Fig. S6). Finally, isolated immunopeptides are analyzed by LC-MS/MS analysis using a timsTOF SCP mass spectrometer operating in DDA-PASEF mode, followed by immunopeptide identification using four different search engines with data-driven rescoring as described (33).

To test the sensitivity of the immunopeptidomics platform, we enriched MHC class I and II peptides from JY cells, using input amounts ranging from 0.5 to 32 million cells. For each input amount, triplicate samples were processed side by side, and from each sample, 25% of the isolated immunopeptides was injected onto the timsTOF SCP, thus corresponding to

0.125 to 8 million cell equivalents injected. First, we assessed the global quality of the immunopeptidome by searching all samples together. Over all input amounts, a total of 25,799 peptides were identified in the MHC-I pulldown, including 16,404 (64%) 8- to 12-mers (Fig. 1B, Supplemental Data S1). The majority of these 8- to 12-mers (81%) were predicted to bind at least one JY HLA allele (Fig. 1C) and clustered according to the anticipated HLA binding motifs (Fig. 1D). For MHC-II, we identified 14,202 peptides in JY cells, including 7304 (51%) 14- to 16-mers (Fig. 1E, Supplemental Data S2), of which 85% were predicted binders (Fig. 1F) and yielded the expected MHC sequence motifs (Fig. 1G).

Next, we searched the samples from each input amount separately for fair comparison. For both MHC class I and II pulldowns, we observed a progressive decline in the number of identified peptides with decreasing input amount (Fig. 1, H and I). For instance, from 0.5 million cells (0.125 million injected) on average 2158 MHC-I peptides were identified, while from 32 million cells (8 million injected) on average 17,266 peptides were identified, approximately eight times more. Both the usage of multiple search engines and data-driven rescoring boosted immunopeptide identification, particularly at lower cell inputs (Supplemental Fig. S7). Approximately 50 to 60% of MHC-I peptides and 30 to 50% of MHC-II peptides were identified in all replicates in ≥ 2 million cell inputs (0.5 million injected), with MHC-I peptide intensities showing ~ 0.9 pairwise Pearson correlation between replicates (Supplemental Figs. S8 and S9), demonstrating quantitative reproducibility of the workflow. Interestingly, the proportion of NetMHCpan-predicted MHC-I binders is slightly higher at lower input numbers. With 0.5 million cells (0.125 million injected) as input, 66% of the peptides were predicted to bind JY HLA alleles on average, while for 32 million cells (8 million injected), 55% were predicted binders. Closer inspection revealed this to be partially explained by reported unspecific coisolation of MHC-II binders (64) particularly at higher cell input amounts where MHC-II binders represent up to $\sim 50\%$ of the nonbinding MHC-I peptides (Supplemental Data S1). As such, an increase in cell number input is not linearly correlated with the number of identified MHC class I-binding peptides. This is exemplified by the fact that the number of predicted MHC class I binders starts to plateau around 16 million cells (4 million injected) (Fig. 1H), while relative gains in MHC class II peptide identifications decrease around 4 million cells (1 million injected) (Fig. 1J). Using fixed and rather high amounts of antibody across all input amounts (300 μg W6/32 and 150 μg PdV5.2), likely these plateaus reflect a saturation in cell lysis or MHC solubilization rather than antibody capture. We also observed that for both MHC-I and -II peptides, the median binding strength increased with decreasing input amounts, indicating that the identification of strong-binding peptides is favored over weak-binding peptides in low-input samples (Supplemental Fig. S10).

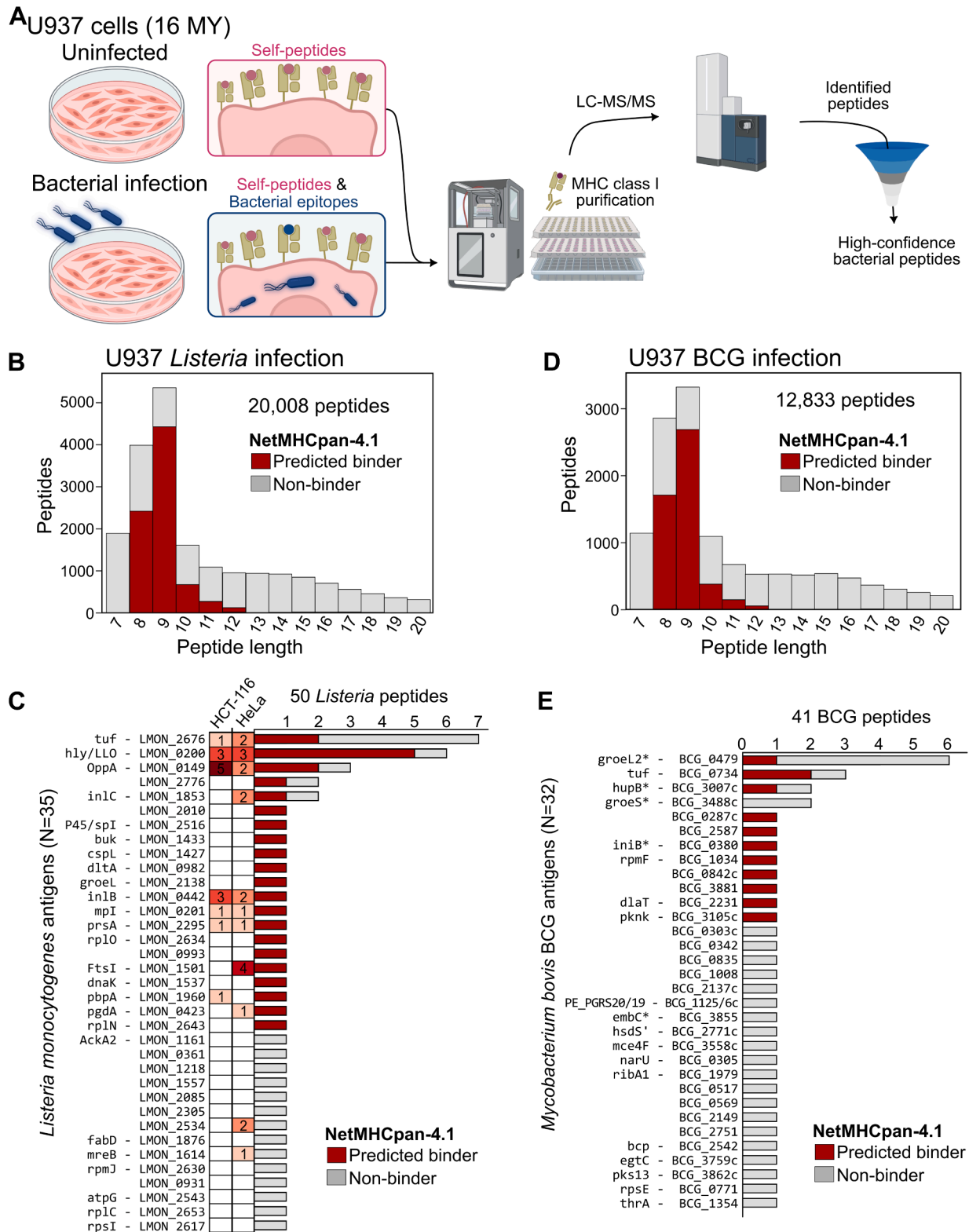


FIG. 3. Bacterial antigen detection. A, bacterial immunopeptidomics workflow scheme. Differentiated U937 cells infected with bacteria and uninfected counterparts are subjected to the immunopeptidomics workflow (16 million cell input, 25% injected). This enables the detection of bacterial epitopes among numerous host self-peptides. After purification and LC-MS/MS of immunopeptides, data was filtered to retain only high-confidence bacterial immunopeptides. Figure created with [BioRender.com](https://www.biorender.com). B and D, histogram showing the length distribution of peptides identified after MHC class I pulldown. Peptides predicted to bind to U937 MHC-I alleles by NetMHCpan-4.1 (53) (%Rank < 2) are

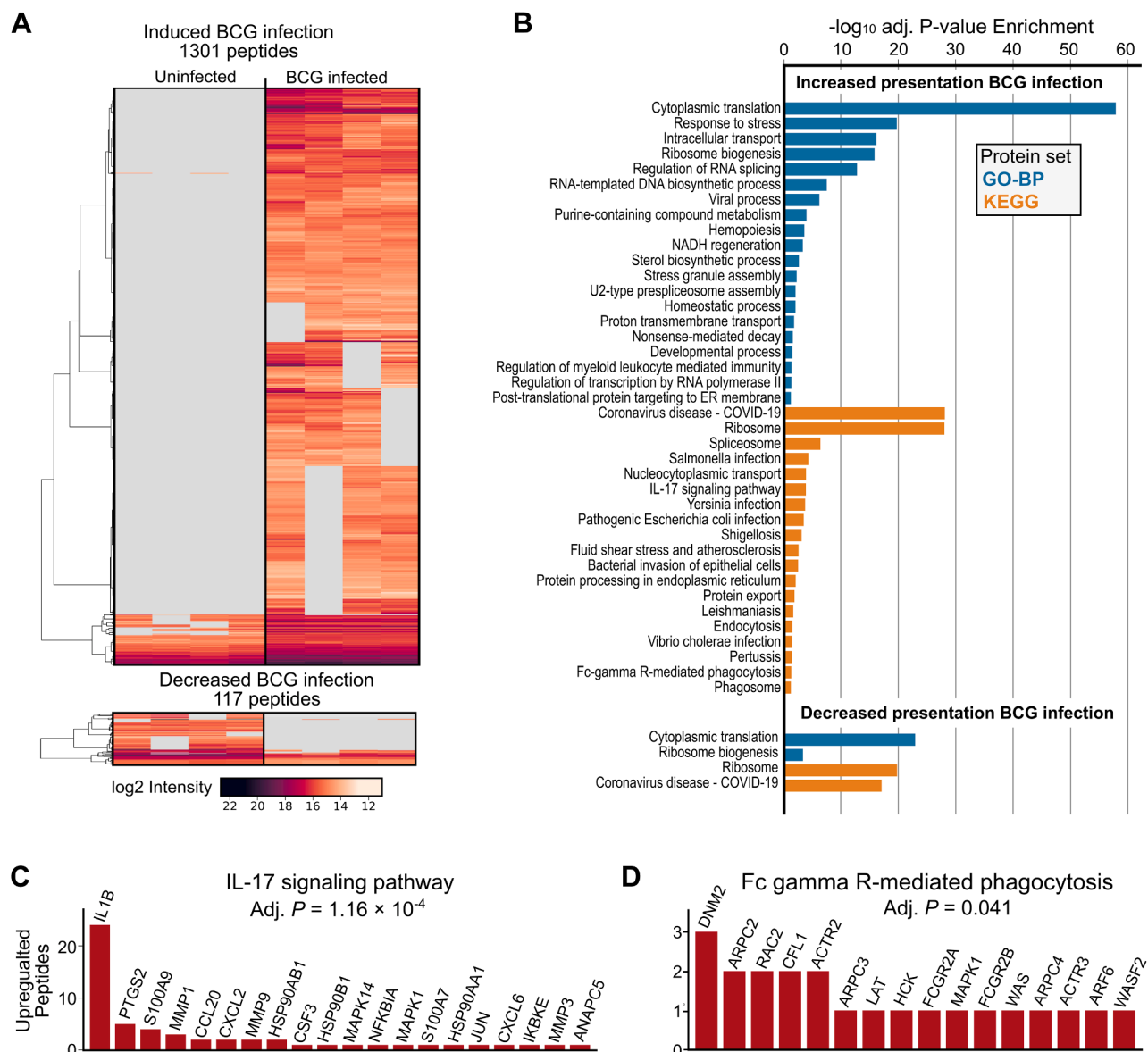


FIG. 4. Host immunopeptide remodeling in response to BCG infection. A, heatmap showing log₂ intensities of significant differentially regulated immunopeptides (adj. *p* value ≤ 0.05 and absolute fold change ≥ 2) as well as immunopeptides uniquely detected in three out of four replicates in infected or uninfected condition. B, functional enrichment analysis of UniProtKB protein accessions with up- or downregulated immunopeptides under infection using g:Profiler (59). C and D, number of significantly upregulated peptides in the enriched protein sets “IL-17 signaling pathway” (C) and FcγR-mediated phagocytosis (D). BCG, bacillus Calmette-Guérin.

To test whether these observations generalize beyond JY cells, we also isolated MHC-I peptides from different amounts of HeLa and U937 cells, representing human epithelial and macrophage cells, respectively. The detected peptides were also of high quality and showed preferred binding patterns to

the anticipated HLA alleles, with a higher proportion of 8-mers in U937 cells due to immunopeptide presentation on HLA-B*18:01 and HLA-B*51:01 (Supplemental Fig. S11, A–F, Supplemental Data S3 and S4) (65). Although in general fewer peptides were identified compared to JY cells, for both cell

indicated in red. C and E, the number of unique immunopeptides identified per *Listeria* protein (C) or *Mycobacterium bovis* BCG protein (E) is shown in a histogram. Immunopeptides predicted as MHC class I binders by NetMHCpan-4.1 (53) (%Rank < 2) are indicated in red, other peptides in gray. C, the number of identified peptides per cell line in our previous *Listeria* study (8), which used 350 to 540 million cells as input per replicate, is displayed in a heatmap. E, BCG proteins identified in previous immunopeptidomics studies (9, 10, 68, 69) are indicated by an asterisk (*). BCG, bacillus Calmette-Guérin; MS, mass spectrometry.

types the number of identified immunopeptides also plateaued with a sample size of 16 million cells (4 million injected), which we therefore defined as the optimal input amount for maximum detection of immunopeptides (Fig. 1, *H* and *I*, Supplemental Fig. S11, *H* and *I*). Together, these experiments established an automated workflow for immunopeptidomics sample preparation in 96-well format, compatible with a wide range of input amounts.

Toward Ultrasensitive Immunopeptidomics

Given the successful identification of more than 2000 MHC class I peptides from 0.5 million cells (0.125 million injected) (Fig. 1*H*), we tested the capability of our platform to identify immunopeptides from very low (< 1 million) cell input amounts. Starting from 0.64 million cells, we generated a two-fold serial dilution series down to 20,000 cells as input. The same automated sample preparation workflow was used, but now the full sample amount (100%) was injected on the timsTOF SCP. Analyzing both samples with 640,000 cells, we identified 8858 unique peptides of which 5482 (62%) were predicted MHC class I binders (Fig. 2, *A* and *B*, Supplemental Data S5). At a 30-fold lower input of only 20,000 cells, we still detected 1244 peptides with 292 (23%) predicted binders. A peak of 9-mers was still discernible, and unsupervised Gibbs clustering recovered the leading JY HLA sequence motifs of HLA-A*02:01 and HLA-B*07:02 (Fig. 2, *B* and *C*, Supplemental Fig. S12), indicating successful immunopeptide isolation at these ultra-low inputs. Also here, multi-engine rescoring aided in recovering peptides from these limited cell inputs (Supplemental Fig. S13). In addition, we repeated this experiment but now with application of the MS-compatible nonionic surfactant *n*-dodecyl- β -D-maltoside (DDM), which is used to increase peptide identification in single-cell experiments by decreasing peptide adsorption to plasticware (66). The same trend and number of identified peptides was observed without DDM, suggesting no beneficial effect of DDM in this setup (Supplemental Fig. S14, Supplemental Data S6).

Comparing identified MHC-binding peptides across input amounts revealed a common subset of class I peptides that progressively narrows with each dilution, eventually reaching 236 peptides that were identified across all dilutions (Fig. 2*D*). Thus, 236 out of 292 (81%) MHC-I-presented peptides in the 20,000-cell samples were present in all dilutions. To address why these particular immunopeptides were retained at very low input, we explored multiple biological and technical reasons. To explore the role of peptide intensity, we ran a FragPipe HLA peptidome analysis with IonQuant (56) quantification separately for each dilution (Supplemental Data S7). This revealed MS intensity to be a determining factor, reflecting high immunopeptide abundance and/or favorable MS ionization or LC elution, with high-abundant immunopeptides in the 640,000 cell input more likely to be identified in lower dilutions (Fig. 2*E*). The binding strength of predicted MHC-I binders was similarly distributed across dilutions (Supplemental Fig. S15). We also

inspected the potential sources of the relatively higher proportion of predicted nonbinding (NB) peptides at very low cell dilutions. Although the origin of most of these peptides remains unclear, we observed that 25 to 30% of these peptides belong to typical proteomic contaminants (e.g. keratins) that were reproducibly identified across dilutions (Supplemental Fig. S16, *A* and *B*). NB peptides showed comparable search engine scores to predicted peptide binders, despite exhibiting lower intensities, suggesting these represent *bona fide* identifications present at lower abundance (Supplemental Fig. S16*C*). Column carryover was negligible, though a few dozen peptides identified in the 20,000 cell sample likely carried over from a preceding tryptic quality control and wash run (Supplemental Fig. S16*D*). Overall, this suggests NBs to originate from common MS contaminants introduced during low-input sample preparation in our setup.

Detection of Bacterial Immunopeptides Presented by Infected Cells

When applied to infected cells or tissues, immunopeptidomics allows the untargeted detection of bacterial immunopeptides, making it a powerful technology for antigen discovery and bacterial vaccine development (8, 67). We previously reported the detection of bacterial immunopeptides presented by cells infected with *L. monocytogenes* (8, 15, 33, 36) using several hundred million cells as input. Here, we aimed to test whether our optimized workflow would achieve similar performance in bacterial antigen detection using lower cell input amounts. To this end, we infected 16 million U937 cell samples with *L. monocytogenes* EGD (further referred to as *Listeria*) and isolated MHC class I immunopeptides from three infected and three uninfected samples for LC-MS/MS analysis (Fig. 3*A*). In total, 20,008 peptides were identified, including 12,995 8- to 12-mers of which 7922 (61%) were predicted binders of U937 HLA alleles (Fig. 3*B*, Supplemental Data S8). The detected peptides showed preferred binding patterns to the anticipated HLA alleles, with again a higher proportion of 8-mers resulting from HLA-B*18:01- and HLA-B*51:01-presented peptides (Supplemental Fig. S17*A*). Next, we stringently filtered high-confidence bacterial peptides as described (33). Although most peptides were of human origin, we detected 50 high-confidence bacterial peptides derived from 35 *Listeria* proteins. Of these, elongation factor Tu (*tuf*), virulence factors *hly/LLO* and *inlC*, periplasmic oligopeptide-binding protein *OppA* and cell wall carboxypeptidase *LMON_2776* were represented by multiple immunopeptides (Fig. 3*C*). This observation is in line with the previously reported immunodominance of these *Listeria* antigens, where *tuf*, *hly/LLO* and *OppA* were also identified by multiple immunopeptides in infected HeLa and HCT-116 cell cultures (8, 33, 36). In addition to *Listeria*, we applied our workflow to U937 cells infected with BCG. Here, 12,833 MHC class I peptides were identified, including 8480 8- to 12-mers with 4992 (59%) predicted binders (Fig. 3*D*, Supplemental Fig. S17*B*, Supplemental Data S9). We detected

41 high-confidence bacterial peptides derived from 32 BCG proteins. Again, several BCG proteins were represented by multiple immunopeptides, including elongation tuf, groEL2, groES and hupB (Fig. 3E). Interestingly, peptides from these three proteins were also identified by a previous immunopeptidomics study that mapped MHC class II peptides on BCG-infected cells (9, 10), and the same holds true for embC (68). Also IniB was identified in *Mycobacterium tuberculosis*-infected macrophages (69). Together, these data demonstrate the ability of our workflow to effectively capture relevant bacterial immunopeptides from limited amounts of infected cells.

Quantitative Remodeling of the Human Immunopeptidome upon BCG Infection

Bacterial infection impacts many cellular processes, resulting in altered presentation of host cell immunopeptides in addition to bacterial peptides. Using a quantitative, nonspecific HLA workflow in FragPipe, we aimed to identify differential human immunopeptide presentation upon BCG infection. Firstly, we assessed the reproducibility of peptide quantification. Uninfected controls displayed a higher quantitative reproducibility (Pearson $R = 0.885 \pm 0.034$) than BCG-infected samples (Pearson $R = 0.794 \pm 0.045$), suggesting that infection introduces greater quantitative variability (Supplemental Fig. S18A). Furthermore, the multidimensional scaling plot based on immunopeptide intensities revealed a clear separation between uninfected controls and BCG-infected samples (Supplemental Fig. S18B). Next, we compared the intensities of the human immunopeptides between the BCG-infected and uninfected samples. This analysis revealed 1301 immunopeptides that were significantly upregulated (adj. $p \leq 0.05$, $\geq$ two-fold change, 117 peptides) or uniquely present in the infected samples (1184 peptides), while 117 peptides were upregulated (32 peptides) or uniquely present in the uninfected samples (85 peptides) (Fig. 4A, Supplemental Data S10). Gene set enrichment analysis revealed that many regulated peptides (both up and down) were derived from ribosomal proteins, which in general are overrepresented MHC-I source proteins (45, 70, 71), and many peptides upregulated upon infection originated from proteins linked to infection, inflammation and immune signaling (Fig. 4B). For instance, interleukin-1 beta showed a drastic upregulation of 24 immunopeptides during infection, and was annotated alongside other immune-related proteins as part of the KEGG “IL-17 signaling pathway” that was significantly enriched (adj. $P = 1.16 \times 10^{-4}$) (Fig. 4C). Interleukin-1 beta is a well-known pro-inflammatory cytokine that accumulates in monocytes and macrophages in response to pathogen infection (72, 73). Interestingly, it is produced in the cytosol as an inactive precursor protein that requires a maturation step for its release (73, 74), possibly facilitating its strong presentation by MHC-I under infection. In addition, proteins like Fcγ receptors (FCGR2A and B) and ARP2/3 complex subunits (ACTR and ARPC proteins) involved in “FcγR-mediated phagocytosis” were more presented by

infected cells (adj. $P = 0.041$) (Fig. 4D). Taken together, these data show the high reproducibility and quantitative performance of our workflow, revealing immune remodeling under infection, with increased presentation by MHC class I molecules of proteins involved in inflammation, phagocytosis, and other immune-related processes triggered by BCG infection.

DISCUSSION

We here describe a platform for immunopeptidomics sample preparation that is both sensitive and compatible with high-throughput analysis. Sequential capture of MHC class I and II immunopeptides is achieved through optimized IP in 96-well format on a positive pressure device. In combination with sensitive timsTOF SCP data acquisition and analysis (28, 32, 33) we demonstrate that the platform can comprehensively capture and identify MHC class I and II immunopeptidomes using sample sizes ranging from 32 to 0.5 million cells, with 16 million cells as optimum after which no further gain in detected binders was observed. Exploring the sensitivity limits of our platform, we could detect over 1200 peptides from only 20,000 cells, albeit challenged with lower percentages of predicted binders. To further showcase the discovery potential of the platform, we performed bacterial antigen discovery in 16 million U937 macrophages infected with *Listeria* or BCG. This led to the identification of known and novel bacterial antigens, despite a 30-fold reduction in input material compared to previous analyses (8, 9, 33, 36, 68, 69). Thus, our optimized platform allows sensitive immunopeptide detection in a high-throughput fashion, enabling its application on biological sample types where input amounts are limited.

Cell lysis, IP and elution steps are determining steps for successful immunopeptide isolation, which were all carefully optimized. Traditional IP protocols typically mix antibody-bound beads with lysate under agitation (8, 25); however, we observed that MHC-I IP under static conditions produces equivalent results (Supplemental Fig. S2, A–C). As a more important factor, we uncovered that efficient MHC-I IP heavily depends on a high protein concentration, which can be achieved by reducing the lysis buffer volume (Supplemental Fig. S2, D and E). To increase sample throughput, we tested multiple 96-well filter plates with pore sizes ranging from 25 to 0.2 μm using a positive pressure device. With larger pores causing leakage during incubation steps and smaller pores requiring higher pressure, we selected a 0.7 μm filter plate as an optimal middle ground, allowing a steady flow without leakage nor clogging. Finally, we tested IP elution and peptide purification methods (Supplemental Figs. S4 and S5), determining 10% acetic acid IP elution with Sep-Pak tC18 peptide purification as a preferred method.

Together, these optimizations led to the most important difference with previous multi-well sample preparation procedures (Supplemental Table S2): the small volume of only 100 μl used for lysis and MHC capture. Other multi-well procedures typically start from larger lysis volumes of 1 ml or higher such

as the original 96-well procedure by Chong *et al.* (26). Inspired by this protocol, we also employ static IP, however, instead of relying on gravity flow, we optimized the plate pore size to prevent leakage and to allow flow control by positive pressure. Overall, our platform reduces sample transfers, thereby avoiding sample losses and enhancing sensitivity compared to procedures involving single tubes (8, 15) or incubation of deep well plates under rotation (28). Similar to other immunopeptidomics platforms using Waters (26), Agilent (40, 41) or KingFisher (42) devices, we implemented our procedure on a Tecan Resolvex A200 to automate the MHC capture, washing and elution steps, making use of the instrument's ability to dispense, incubate and push through the necessary buffer solutions. It is important to note that we aimed to develop a workflow that is universally applicable, regardless of the number of MHC molecules present in the samples. By keeping the reagents and conditions in the workflow constant regardless of the input amount, this inherently means that the workflow will perform suboptimal for input amounts far above or below the working range established here. It is possible that the protocol for extreme low input samples could further be optimized, however, one caveat is that these optimizations would be heavily biased toward the sample type used (in this case JY cells) and not extrapolate to other types of input material. As an alternative approach for low-input immunopeptidomics, microfluidics set-ups have emerged that further miniaturize immunopeptidomics sample preparation with minimal sample transfers. However, these setups require highly specialized equipment and expertise for implementation while only processing a single sample each time, making this approach impractical for projects demanding higher throughput.

For maximal performance, low-input sample preparation should be coupled to sensitive MS data acquisition. To this end, we analyzed our samples on a timsTOF SCP instrument operating in DDA-PASEF mode, leveraging ion mobility-*m/z* peptide precursor selection polygons tailored for immunopeptidomics (28, 32, 33) (Supplemental Fig. S1). Importantly, the cell input numbers of the ultrasensitive series (20,000–640,000 cells) reported reflect the actual numbers used for lysis, more closely mimicking real-life applications compared to the cell equivalents that are sometimes used for method optimization (28, 32). In these experiments, DDM did not increase the number of identified immunopeptides, indicating no benefit of this surfactant for low-input immunopeptidomics in our setup. However, it is important to note that immunopeptide identification numbers are largely dependent on the used cell line and its MHC expression levels, making comparisons across studies and cell types difficult. To further maximize reliable immunopeptide identification, we applied advanced data analysis by integrating the results of four independent search engines that were rescored by TIMS²-Rescore (52), as recently described (33). In future work, we plan to couple our platform to alternative MS acquisition methods,

including data-independent acquisition-PASEF (30, 36, 75), or to other acquisition schemes such as Slice-PASEF (76) or Synchro-PASEF (77). Alternatively, the high sensitivity and speed offered by Orbitrap Astral instruments holds great promise for low-input immunopeptidomics (29). In terms of sample preparation, we envision to improve the sensitivity of our platform even further by using magnetic beads (42, 46) along with biotinylated MHC antibodies (35, 45) for IP, to work with even smaller sample volumes and to prevent endogenous antibody binding to protein-A or -G beads when processing tissue samples. The latter could also be achieved by introducing a standard preclearing step, which, together with more stringent wash steps, might further reduce the observed background of non-MHC binding peptides in our experiments. Interestingly, MAE has regained interest and was shown to offer specific and sensitive enrichment of MHC class I peptides as a cost-effective alternative for antibody-based IP (29–31). However, due to its need for live cells in suspension, MAE has limited applicability to solid tissue samples and frozen cells (31, 78), though it was successfully applied to clinical tissue fragments (29). In addition, so far it has only been reported for MHC-I, limiting its applicability to samples where the immunopeptidome of both MHC classes are of interest.

In this study, we have also demonstrated the ability of our immunopeptidomics platform to identify bacterial antigens in a lower-input setting using 16 million infected macrophages. This input downscaling substantially decreases the upfront cell culturing and hands-on time, which makes the procedure compatible with higher biosafety level pathogens for which the maximal number of cells that can be infected is limited. The lower input also allows quantitative immunopeptidomics analysis on multiple replicates and conditions. Firstly, we performed antigen discovery for *L. monocytogenes*, a food-borne model pathogen for which we previously gained a comprehensive view on its presented antigens, allowing us to cross-reference the *Listeria* antigens identified by the low-input platform. Despite a 30-fold reduction in input amount, we identified 50 MHC-I peptides from 35 *Listeria* proteins in infected U937 macrophages. Of these 35 antigens, 12 were identified in previous *Listeria* immunopeptidomics studies with high cell number inputs (8, 33, 36), including previously detected immunodominant antigens and virulence factors such as LLO, OppA, InlC, and InlB. After this verification, we applied our platform on U937 cells infected with BCG, an attenuated *M. bovis* strain. This analysis uncovered 41 MHC-I peptides from 32 BCG proteins, of which groEL2, groES, hupB and embC were also detected by previous BCG antigen discovery studies in higher input analyses (9, 10, 68), while InlB was picked up in *M. tuberculosis*-infected U937/A2 cells (69). Aside from the identification of pathogen-derived immunopeptides, bacterial infection can remodel the host immunopeptidome (14, 15). We previously reported how *Listeria* infection promoted presentation of immune-related

self proteins in spleens of infected mice (15). Similarly, we here report how BCG infection of U937 macrophages leads to increased MHC-I presentation of host proteins partaking in immune pathways related to immune signaling and phagocytosis (Fig. 4), pathways that were previously reported to be upregulated in BCG-infected macrophages (79). From a technical perspective, these analyses demonstrated the excellent quantitative reproducibility of our immunopeptidomics platform, with an average Pearson correlation of 0.88 between immunopeptide intensities in uninfected biological replicate samples. Together, these data validate our platform for biological immunopeptidomics applications where a combination of high sensitivity, high throughput, and quantitative performance are required.

DATA AVAILABILITY

The MS proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository (80) with the dataset identifier PXD062352. Annotated MS/MS spectra of JY cell input experiments and bacterial infection experiments and used protein databases are accessible at the Zenodo repository under <https://doi.org/10.5281/zenodo.19368946>.

Supplemental data—This article contains supplemental data (8, 26, 28, 40–46, 51–55).

Acknowledgments—We thank the VIB Proteomics Core for LC-MS/MS analysis.

Author Contributions—A. G., L. V. M., P. W., K. C. F., A. S., S. D., I. L., C. D., F. T., and F. I. writing—review and editing; A. G., L. V. M., P. W., F. T., and F. I. writing—original draft; A. G., L. V. M., I. A., L. P. C., K. C. F., K. B., and A. S. investigation; A. G., L. V. M., and P. W. visualization; A. G., L. V. M., A. S., and F. T. methodology; P. W. software; P. W. formal analysis; P. W. data curation; A. S. and F. I. conceptualization; S. D., B. V., C. D., and F. I. resources; F. T. and F. I. supervision.

Funding and Additional Information—F. I., C.M., B. V., and I. L. acknowledge support from the Horizon Europe Project BAXERNA 2.0 [101080544]. F. I., I. L., and B. V. acknowledge support from Ghent University Concerted Research Action grant BOF21/GOA/033, and F. I. from Starting Grant BOF/STA/202209/011, as well as from the European Research Council (ERC Consolidator Grant #101089193). A. G. is supported by a PhD fellowship from the Higher Education Commission (HEC) Pakistan. This work was supported by the Research Foundation–Flanders (FWO) through a Junior Postdoctoral fellowship to F. T. (12AN524N) and a PhD fellowship for Strategic Basic Research to I. A. (1S40923N).

Conflict of Interest—The authors declare no competing interests.

Abbreviations—The abbreviations used are: ACN, acetonitrile; BCG, bacillus Calmette–Guérin; DDA, data-dependent acquisition; DDM, n-dodecyl- β -D-maltoside; FBS, fetal bovine serum; HLA, human leukocyte antigen; IP, immunoprecipitation; MAE, mild acid elution; MS, mass spectrometry; PASEF, parallel accumulation–serial fragmentation; PMA, phorbol 12-myristate 13-acetate; SCP, single-cell proteomics; TIMS, trapped ion mobility spectrometry; tuf, elongation factor Tu.

Received December 3, 2025, and in revised form, April 5, 2026
Published, MCPRO Papers in Press, May 18, 2026, <https://doi.org/10.1016/j.mcpro.2026.101590>

REFERENCES

1. Pishesha, N., Harmand, T. J., and Ploegh, H. L. (2022) A guide to antigen processing and presentation. *Nat. Rev. Immunol.* **22**, 751–764
2. Thibault, P., and Perreault, C. (2022) Immunopeptidomics: reading the immune signal that defines self from nonself. *Mol. Cell Proteomics* **21**, 100234
3. Bassani-Sternberg, M., Pletscher-Frankild, S., Jensen, L. J., and Mann, M. (2015) Mass spectrometry of human leukocyte antigen class I peptides reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell Proteomics* **14**, 658–673
4. Caron, E., Vincent, K., Fortier, M. H., Laverdure, J. P., Bramouille, A., Hardy, M. P., et al. (2011) The MHC I immunopeptidome conveys to the cell surface an integrative view of cellular regulation. *Mol. Syst. Biol.* **7**, 533
5. Shapiro, I. E., and Bassani-Sternberg, M. (2023) The impact of immunopeptidomics: from basic research to clinical implementation. *Semin. Immunol.* **66**, 101727
6. Gfeller, D., and Bassani-Sternberg, M. (2018) Predicting antigen presentation—what could we learn from a million peptides? *Front Immunol.* **9**, 1716
7. Flender, D., Vilenne, F., Adams, C., Boonen, K., Valkenburg, D., and Baggerman, G. (2025) Exploring the dynamic landscape of immunopeptidomics: unravelling posttranslational modifications and navigating bioinformatics terrain. *Mass Spectrom. Rev.* **44**, 599–629
8. Mayer, R. L., Verbeke, R., Asselman, C., Aernout, I., Gul, A., Eggermont, D., et al. (2022) Immunopeptidomics-based design of mRNA vaccine formulations against *Listeria monocytogenes*. *Nat. Commun.* **13**, 6075
9. Bettencourt, P., Muller, J., Nicastrì, A., Cantillon, D., Madhavan, M., Charles, P. D., et al. (2020) Identification of antigens presented by MHC for vaccines against tuberculosis. *NPJ Vaccin.* **5**, 2
10. Almuji, S. S., Stylianou, E., Nicastrì, A., Satti, I., Korompis, M., Li, S., et al. (2025) MetE: a promising protective antigen for tuberculosis vaccine development. *Front Immunol.* **16**, 1593263
11. Leddy, O., Ogongo, P., Huffaker, J., Gan, M., Milligan, R., Mahmud, S., et al. (2025) Immunopeptidomics can inform the design of mRNA vaccines for the delivery of *Mycobacterium tuberculosis* MHC class II antigens. *Sci. Transl. Med.* **17**, eadw9184
12. Leddy, O., White, F. M., and Bryson, B. D. (2023) Immunopeptidomics reveals determinants of *Mycobacterium tuberculosis* antigen presentation on MHC class I. *Elife* **12**, e84070
13. Leddy, O., Yuki, Y., Carrington, M., Bryson, B. D., and White, F. M. (2025) Targeting infection-specific peptides in immunopeptidomics studies for vaccine target discovery. *J. Exp. Med.* **222**, e20250444
14. Cormican, J. A., Medfal, L., Wawrzyniuk, M., Pasen, M., Afrache, H., Fourny, C., et al. (2025) PEPSeek-Mediated identification of novel epitopes from viral and bacterial pathogens and the impact on host cell immunopeptidomes. *Mol. Cell Proteomics* **24**, 100937
15. Gul, A., Pewe, L. L., Willems, P., Mayer, R., Thery, F., Asselman, C., et al. (2024) Immunopeptidomics mapping of *Listeria monocytogenes* T cell epitopes in mice. *Mol. Cell Proteomics* **23**, 100829

16. Karunakaran, K. P., Yu, H., Jiang, X., Chan, Q., Goldberg, M. F., Jenkins, M. K., *et al.* (2017) Identification of MHC-Bound peptides from dendritic cells infected with *Salmonella enterica* Strain SL1344: implications for a nontyphoidal salmonella vaccine. *J. Proteome Res.* **16**, 298–306
17. Zhang, B., and Bassani-Sternberg, M. (2023) Current perspectives on mass spectrometry-based immunopeptidomics: the computational angle to tumor antigen discovery. *J. Immunother. Cancer* **11**, e007073
18. Bassani-Sternberg, M., Braunlein, E., Klar, R., Engleitner, T., Sinitcyn, P., Audehm, S., *et al.* (2016) Direct identification of clinically relevant neo-epitopes presented on native human melanoma tissue by mass spectrometry. *Nat. Commun.* **7**, 13404
19. Shapiro, I. E., Huber, F., Michaux, J., and Bassani-Sternberg, M. (2025) Sensitive neoantigen discovery by real-time mutanome-guided immunopeptidomics. *Nat. Commun.* **16**, 7269
20. Apavaloai, A., Zhao, Q., Hesnard, L., Cahuzac, M., Durette, C., Larouche, J. D., *et al.* (2025) Tumor antigens preferentially derive from unmutated genomic sequences in melanoma and non-small cell lung cancer. *Nat. Cancer* **6**, 1419–1437
21. Ely, Z. A., Kulstad, Z. J., Gunaydin, G., Addepalli, S., Verzani, E. K., Casarubios, M., *et al.* (2025) Pancreatic cancer-restricted cryptic antigens are targets for T cell recognition. *Science* **388**, eadk3487
22. Purcell, A. W., Ramarathnam, S. H., and Ternet, N. (2019) Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nat. Protoc.* **14**, 1687–1707
23. Marino, F., Chong, C., Michaux, J., and Bassani-Sternberg, M. (2019) High-Throughput, fast, and sensitive immunopeptidomics sample processing for mass spectrometry. *Methods Mol. Biol.* **1913**, 67–79
24. Mohan, S. V., Datta, K. K., Ziegman, R., Smith, C., and Gowda, H. (2021) Protocol for purification and identification of MHC class I immunopeptidome from cancer cell lines. *STAR Protoc.* **2**, 100385
25. Kowalewski, D. J., and Stevanovic, S. (2013) Biochemical large-scale identification of MHC class I ligands. *Methods Mol. Biol.* **960**, 145–157
26. Chong, C., Marino, F., Pak, H., Racle, J., Daniel, R. T., Muller, M., *et al.* (2018) High-throughput and sensitive immunopeptidomics platform reveals profound Interferon-gamma-Mediated remodeling of the Human Leukocyte Antigen (HLA) ligandome. *Mol. Cell Proteomics* **17**, 533–548
27. Wacker, M., Bauer, J., Wessling, L., Dubbelaar, M., Nelde, A., Rammensee, H. G., *et al.* (2023) Immunoprecipitation methods impact the peptide repertoire in immunopeptidomics. *Front Immunol.* **14**, 1219720
28. Phulphagar, K. M., Ctordecka, C., Jacome, A. S. V., Klaeger, S., Verzani, E. K., Hernandez, G. M., *et al.* (2023) Sensitive, High-Throughput HLA-I and HLA-II immunopeptidomics using parallel accumulation-serial fragmentation Mass Spectrometry. *Mol. Cell Proteomics* **22**, 100563
29. Bathini, M., Bocaniciu, D., Johnson, F. D., de Jong, R. C. P., Yu, F., Aloï, V. D., *et al.* (2026) MHC1-TIP enables single-tube multimodal immunopeptidome profiling and uncovers intratumoral heterogeneity in antigen presentation. *Commun. Biol.* **9**, 296
30. [preprint] Beyrle, J., Yamasaki, Y., Ford, M., Conche, C., Chen, Y., Arnold, A., *et al.* (2026) MAETi: mild acid elution in a tip enables immunopeptidome profiling from 1E4 cells. *bioRxiv*. <https://doi.org/10.1101/2024.12.20.628848>
31. Sturm, T., Sautter, B., Wörner, T. P., Stevanovic, S., Rammensee, H. G., Planz, O., *et al.* (2021) Mild acid elution and MHC immunofluorescence chromatography reveal similar albeit not identical profiles of the HLA class I immunopeptidome. *J. Proteome Res.* **20**, 289–304
32. Gomez-Zepeda, D., Arnold-Schild, D., Beyrle, J., Declercq, A., Gabriels, R., Kumm, E., *et al.* (2024) Thunder-DDA-PASEF enables high-coverage immunopeptidomics and is boosted by MS(2)Rescore with MS(2)PIP timsTOF fragmentation prediction model. *Nat. Commun.* **15**, 2288
33. Willems, P., Thery, F., Van Moortel, L., De Meyer, M., Staes, A., Gul, A., *et al.* (2025) Maximizing Immunopeptidomics-Based bacterial epitope discovery by multiple search engines and rescoring. *J. Proteome Res.* **24**, 2141–2151
34. Pak, H., Michaux, J., Huber, F., Chong, C., Stevenson, B. J., Muller, M., *et al.* (2021) Sensitive immunopeptidomics by leveraging available large-scale Multi-HLA spectral libraries, data-independent acquisition, and MS/MS prediction. *Mol. Cell Proteomics* **20**, 100080
35. Wahle, M., Thielert, M., Zwiebel, M., Skowronek, P., Zeng, W. F., and Mann, M. (2024) IMBAS-MS discovers organ-specific HLA peptide patterns in plasma. *Mol. Cell Proteomics* **23**, 100689
36. Willems, P., Staes, A., Miret-Casals, L., Demichev, V., Devos, S., and Impens, F. (2025) Data-Independent immunopeptidomics discovery of low-abundant bacterial epitopes. *J. Proteome Res.* **24**, 6295–6304
37. Declercq, A., Bouwmeester, R., Hirschler, A., Carapito, C., Degroove, S., Martens, L., *et al.* (2022) MS(2)Rescore: Data-Driven rescoring dramatically boosts immunopeptide identification rates. *Mol. Cell Proteomics* **21**, 100266
38. Yang, K. L., Yu, F., Teo, G. C., Li, K., Demichev, V., Ralser, M., *et al.* (2023) MSBooster: improving peptide identification rates using deep learning-based features. *Nat. Commun.* **14**, 4539
39. Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., *et al.* (1992) Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255**, 1261–1263
40. Zhang, L., McAlpine, P. L., Heberling, M. L., and Elias, J. E. (2021) Automated ligand purification platform accelerates immunopeptidome analysis by mass spectrometry. *J. Proteome Res.* **20**, 393–408
41. Pollock, S. B., Rose, C. M., Darwish, M., Bouziat, R., Delamarre, L., Blanchette, C., *et al.* (2021) Sensitive and quantitative detection of MHC-I displayed neoepitopes using a semiautomated workflow and TOMAHQA mass spectrometry. *Mol. Cell Proteomics* **20**, 100108
42. Lim Kam Sian, T. C. C., Goncalves, G., Steele, J. R., Shamekhi, T., Bramberger, L., Jin, D., *et al.* (2023) SAPrim, a semi-automated protocol for mid-throughput immunopeptidomics. *Front Immunol.* **14**, 1107576
43. Li, X., Pak, H. S., Huber, F., Michaux, J., Taillandier-Coidard, M., Altamiras, E. R., *et al.* (2023) A microfluidics-enabled automated workflow of sample preparation for MS-based immunopeptidomics. *Cell Rep. Methods* **3**, 100479
44. Abelin, J. G., Bergstrom, E. J., Rivera, K. D., Taylor, H. B., Klaeger, S., Xu, C., *et al.* (2023) Workflow enabling deepscale immunopeptidome, proteome, ubiquitylome, phosphoproteome, and acetylome analyses of sample-limited tissues. *Nat. Commun.* **14**, 1851
45. Feola, S., Haapala, M., Peltonen, K., Capasso, C., Martins, B., Antignani, G., *et al.* (2021) PeptiCHIP: a microfluidic platform for tumor antigen landscape identification. *ACS Nano* **15**, 15992–16010
46. Tanuwidjaya, E., Lim Kam Sian, T. C. C., Steele, J. R., Goncalves, G., Woodhouse, I. B., Chang, J., *et al.* (2025) SAPrim 2.0: a semi-automated protocol for mid-throughput soluble HLA immunopeptidomics. *Front Immunol.* **16**, 1546629
47. Rijkers, G. T., Roord, J. J., Koning, F., Kuis, W., and Zegers, B. J. (1987) Phenotypic and functional analysis of B lymphocytes of two siblings with combined immunodeficiency and defective expression of major histocompatibility complex (MHC) class II antigens on mononuclear cells. *J. Clin. Immunol.* **7**, 98–106
48. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D., and Nesvizhskii, A. I. (2017) MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520
49. Eng, J. K., Jahan, T. A., and Hoopmann, M. R. (2013) Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24
50. Lazear, M. R. (2023) Sage: an open-source tool for fast proteomics searching and quantification at Scale. *J. Proteome Res.* **22**, 3652–3659
51. Zhang, J., Xin, L., Shan, B., Chen, W., Xie, M., Yuen, D., *et al.* (2012) PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol. Cell Proteomics* **11**, M111.010587
52. Declercq, A., Devreese, R., Scheid, J., Jachmann, C., Van Den Bossche, T., Preikschat, A., *et al.* (2025) TIMS(2)Rescore: a data dependent acquisition-parallel accumulation and serial fragmentation-optimized data-driven rescoring pipeline based on MS(2)Rescore. *J. Proteome Res.* **24**, 1067–1076
53. Reynisson, B., Alvarez, B., Paul, S., Peters, B., and Nielsen, M. (2020) NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454
54. Nilsson, J. B., Kaabinejad, S., Yari, H., Kester, M. G. D., van Balen, P., Hildebrand, W. H., *et al.* (2023) Accurate prediction of HLA class II antigen presentation across all loci using tailored data acquisition and refined machine learning. *Sci. Adv.* **9**, ead6367

55. Andreatta, M., Alvarez, B., and Nielsen, M. (2017) GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Res.* **45**, W458–W463
56. Yu, F., Haynes, S. E., and Nesvizhskii, A. I. (2021) IonQuant enables accurate and sensitive label-free quantification with FDR-Controlled match-between-runs. *Mol. Cell Proteomics* **20**, 100077
57. Smyth, G. K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3
58. van Ooijen, M. P., Jong, V. L., Eijkemans, M. J. C., Heck, A. J. R., Andeweg, A. C., Binai, N. A., et al. (2018) Identification of differentially expressed peptides in high-throughput proteomics data. *Brief Bioinform* **19**, 971–981
59. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198
60. Hunter, J. D. (2007) Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95
61. Waskom, M. L. (2021) Seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021
62. Tareen, A., and Kinney, J. B. (2020) Logomaker: beautiful sequence logos in Python. *Bioinformatics* **36**, 2272–2274
63. Lex, A., Gehlenborg, N., Strobelt, H., Vuilleumot, R., and Pfister, H. (2014) UpSet: visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph* **20**, 1983–1992
64. Partridge, T., Nicastrì, A., Kliszczak, A. E., Yindom, L. M., Kessler, B. M., Ternette, N., et al. (2018) Discrimination between human leukocyte antigen class I-Bound and Co-Purified HIV-Derived peptides in immunopeptidomics workflows. *Front Immunol.* **9**, 912
65. Guasp, P., Alvarez-Navarro, C., Gomez-Molina, P., Martin-Esteban, A., Marcilla, M., Barnea, E., et al. (2016) The peptidome of Behcet's disease-associated HLA-B*51:01 includes two subpeptidomes differentially shaped by endoplasmic reticulum aminopeptidase 1. *Arthritis Rheumatol.* **68**, 505–515
66. Tsai, C. F., Zhang, P., Scholten, D., Martin, K., Wang, Y. T., Zhao, R., et al. (2021) Surfactant-assisted one-pot sample preparation for label-free single-cell proteomics. *Commun. Biol.* **4**, 265
67. Aernout, I., Verbeke, R., Thery, F., Willems, P., Elia, U., De Smedt, S. C., et al. (2025) Challenges and opportunities in mRNA vaccine development against bacteria. *Nat. Microbiol.* **10**, 1816–1828
68. Karunakaran, K. P., Yu, H., Jiang, X., Chan, Q. W. T., Sigola, L., Millis, L. A., et al. (2024) Immunoproteomic discovery of *Mycobacterium bovis* antigens, including the surface lipoprotein Mpt83 as a T cell antigen useful for vaccine development. *Vaccine* **42**, 126266
69. Flyer, D. C., Ramakrishna, V., Miller, C., Myers, H., McDaniel, M., Root, K., et al. (2002) Identification by mass spectrometry of CD8(+) T-cell *Mycobacterium tuberculosis* epitopes within the Rv0341 gene product. *Infect Immun.* **70**, 2926–2932
70. Perreault, C. (2010) The origin and role of MHC class I-associated self-peptides. *Prog. Mol. Biol. Transl. Sci.* **92**, 41–60
71. Hickman, H. D., Luis, A. D., Buchli, R., Few, S. R., Sathiamurthy, M., VanGundy, R. S., et al. (2004) Toward a definition of self: proteomic evaluation of the class I peptide repertoire. *J. Immunol.* **172**, 2944–2952
72. Cooper, A. M., Mayer-Barber, K. D., and Sher, A. (2011) Role of innate cytokines in mycobacterial infection. *Mucosal Immunol.* **4**, 252–260
73. Ford, S. G., Caswell, P., Brough, D., and Seoane, P. I. (2025) The secretion of interleukin-1beta. *Cytokine Growth Factor Rev.* **84**, 101–113
74. Monteleone, M., Stow, J. L., and Schroder, K. (2015) Mechanisms of unconventional secretion of IL-1 family cytokines. *Cytokine* **74**, 213–218
75. Oliinik, D., Gurung, H. R., Zhou, Z., Leskoske, K., Rose, C. M., and Klaeger, S. (2025) diaPASEF analysis for HLA-I peptides enables quantification of common cancer neoantigens. *Mol. Cell Proteomics* **24**, 100938
76. [preprint] Sinn, L. R., Szyrwiel, L., Grossmann, J., Lau, K., Faisst, K., Qin, D., et al. (2025) Slice-PASEF: maximising ion utilisation in LC-MS proteomics. *bioRxiv*. <https://doi.org/10.1101/2022.10.31.514544>
77. Skowronek, P., Krohs, F., Lubeck, M., Wallmann, G., Itang, E. C. M., Koval, P., et al. (2023) Synchro-PASEF allows precursor-specific fragment ion extraction and interference removal in data-independent acquisition. *Mol. Cell Proteomics* **22**, 100489
78. Lanoix, J., Durette, C., Courcelles, M., Cossette, E., Comtois-Marotte, S., Hardy, M. P., et al. (2018) Comparison of the MHC I immunopeptidome repertoire of B-Cell lymphoblasts using two isolation methods. *Proteomics* **18**, e1700251
79. Schaefer, Z., Iradukunda, J., Lumngwena, E. N., Basso, K. B., Blackburn, J. M., and Parker, I. K. (2024) Multilevel proteomics reveals epigenetic signatures in BCG-Mediated macrophage activation. *Mol. Cell Proteomics* **23**, 100851
80. Perez-Riverol, Y., Bandla, C., Kundu, D. J., Kamatchinathan, S., Bai, J., Hewapathirana, S., et al. (2025) The PRIDE database at 20 years: 2025 update. *Nucleic Acids Res.* **53**, D543–D553