

High sensitivity class I immunopeptidomics on the timsTOF SCP mass spectrometer

Mass spectrometric characterization of peptide antigens presented by the major histocompatibility complex is essential to understanding infection, cancer and autoimmunity.

Abstract

Due to their low abundance, non-tryptic nature and high-complexity, comprehensive LC-MS analysis of Immunopeptidomes purified from surface HLA molecules remains highly challenging. Many peptides ionize as singly charged species hampering their identification. In trapped ion mobility spectrometry (TIMS), these peptides are readily discriminated from multiply charged ion signals by their Collisional Cross Section, reducing the interference from co-isolations that usually occurs in pure m/z analyses. Here, we use the timsTOF SCP system as a highly sensitive instrument with PASEF® for in depth analysis and show unprecedented coverage on low sample input of immunopeptide-like standard samples and real immunopeptidome.

Keywords:
Immunopeptidomics, high
sensitivity, timsTOF SCP,
PEAKS

Introduction

Immunopeptidomics aims to resolve the composition and dynamics of endogenous peptides presented by major histocompatibility complex (MHC) proteins. In humans the MHC locus encode a vital arm of adaptive immunity, where a series of highly polymorphic cell surface Human Leukocyte Antigen (HLA) proteins act to present endogenous peptides for recognition by cognate T-lymphocytes. This peptide display-recognition system enables the elimination of infected, cancerous and damaged cells by cell-mediated immunity, alongside supporting antibody production and innate immune responses. The identification of disease and tumour-related HLA presented peptides in cell lines and biopsies will progress our understanding of disease pathogenesis, immune responses and aid the preparation of targeted vaccines and immunotherapies [1].

There are two distinct classes of HLA proteins, class I and class II. Class I peptide presentation is found on all nucleated cells in the vertebrate body, peptides are typically 8-14 amino acids in length and sourced from endogenous protein production and degradation machinery. Class II peptides can originate from extracellular protein sources and are processed through endosomes, to produce clusters of peptides 15-25 amino acids in length.

Currently, mass spectrometry (MS) provides the only methodology able to reveal the exact nature of the immunopeptidome. However, analytical challenges in analysis arise from sample complexity, low abundance and the distinct peptide sequence heterogeneity found in the immunopeptidome of different people. Often these factors mean that existing proteomic methodologies adapted to target tryptic peptides and ultimately quantitate a protein, not a peptide are particularly unsuitable. For example, class I peptides are often short, consist of hydrophobic residues and unlike tryptic peptides do not always possess an K/R residue. This results in a large proportion of singly charged species and poorer fragmentation spectra after LC-MS analysis [2].

Trapped ion mobility spectrometry (TIMS) in the timsTOF series of time of flight (TOF) MS instruments can reduce spectrum complexity prior to MS analysis. In TIMS ions are trapped in an electric field according to their three-dimensional size and charge in the gas phase, resulting in separation by mobility. This allows for improved signal to noise and better sampling efficiency of m/z features by MS and improved peptide identification rates. More recently, to empower single cell proteomics (SCP) the timsTOF SCP has been developed. This instrument is equipped with a larger transfer capillary bore for up to 5-fold higher ion transfer, additional high-pressure ion transfer optics for more effective ion collection and two orthogonal deflections to maintain robustness [3].

Whilst the timsTOF SCP was developed for single cell proteomics applications, it is highly suitable to any study where analytes are present at exceptionally low concentrations. This makes this instrument ideally suited to immunopeptidomic research. Here we present optimised MS parameters for the detection and identification of class I immunopeptides. Furthermore, we validated our methods by investigating the application on a synthetic standard containing a predefined diverse repertoire of immunopeptides. Finally, we demonstrate this application on a real immunopeptidomic sample.

Material and Methods

Peptide standards for immunopeptidomics

Two immunopeptidome-like peptide standards and one endogenous immunopeptidome were prepared and measured for optimisation and evaluation of instrument sensitivity.

- HeLa protein lysate was reduced, alkylated and digested with elastase (cleaves at C-terminal of Ala, Val, Ser, Gly, Leu and Ile) at 37°C for 3 h. Peptides were purified by SEP-PAK reversed phase SPE [4].
- Synthetic peptide library consisting of 2000 synthetic peptides which were produced by solid-phase synthesis on cellulose membranes. Peptides were cleaved from the membrane into HLA specific pools of 250 peptides each, pools were combined at equimolar amounts (20 fmol/peptide) [4].

Immunopeptidomic sample preparation

The Jurkat immunopeptidome was prepared from 2.5×10^9 cells, which were lysed and immunoprecipitated with W632-Protein A beads, washed, eluted in acetic acid, and purified by 5 kDa MWCO filtration and C18 stage-tip purification.

LC-MS and data analysis

Chromatographic separation was performed using a nanoElute equipped with an Aurora column (pre-fitted nanoZero™ connection and integrated emitter tip, 25 cm x 75 μm , 1.7 μm , C_{18} ; IonOpticks, Australia) with gradient times of 20 min, 36 min and 66 min (2-37% acetonitrile in 0.1% formic acid; 150 nL/min) coupled to a captive spray ionisation source on a timsTOF SCP. The instrument was operated in dda-PASEF mode with default settings consisting of 10 ramps of 166 ms each. MS/MS spectra were acquired with a threshold intensity of 500 and target intensity of 20,000 and active exclusion was done for 0.4 min. Collision energy was dependent on ion mobility values, using a slope of $1K_0 = 0.6$ at 20 eV to $1K_0 = 1.59$ at 59 eV. Charge states 0-5 were

fragmented. Ion mobility- m/z windows and alterations to several of these default settings are described in the results. The raw data were processed in PEAKS Studio Xpro, searched against a human Uniprot database of 20,606 reviewed entries. General search parameters: 10 ppm MS tolerance; 0.05 Da MS/MS tolerance; enzyme: none; digestion mode: non-specific; database: human (uniprot reviewed). For the HeLa lysate and synthetic library fixed modifications: Carbamidomethylation (C) was included. For the Jurkat immunopeptidome variable modifications: Oxidation (M), Oxidation or Hydroxylation (DKNPY, C-term), Deamidation (NQ). Peptides were reported at a FDR of <1% cut off determined using a target decoy strategy.

Results and Discussion

Inclusion of singly charged peptides

To create a readily available standard that simulates class I peptides HeLa lysate was digested with elastase and analysed by the timsTOF SCP using default settings. Figure 1 is a heat map of ion mobility and m/z dimensions for all precursors eluted during analysis. Using this plot a clear charge-based separation is resolvable. The force from the field gradient in the tims tunnel is qE , so +1 ions "feel" only half of the force of +2 ions. Typically, these ions are excluded from the precursor selection in PASEF bottom-up proteomics, as they originate from background signals or have no further influence on protein identification as multiply charged versions are readily available.

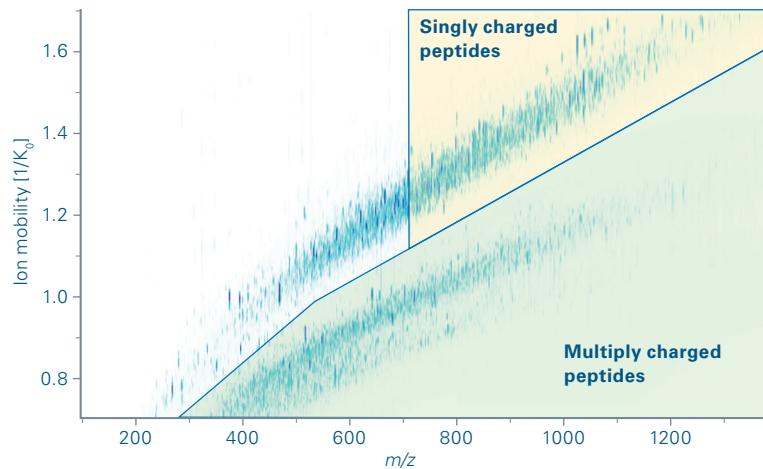


Figure 1
Precursor selection for dda-PASEF. Multiply charged precursors and singly charged precursors with $m/z \geq 700$ are selected.

To adapt the default method for class I immunopeptides, the inclusion of singly charged precursors ($m/z \geq 700$) was tested by adapting the ion mobility window to sample this region for MS/MS analysis (Figure 1). Figure 2A shows the number of sequences identified with amino acid lengths ranging from 7 to 30 from a 40 ng HeLa lysate digested with elastase. Generally, most sequences identified have an amino acid length of 7-14 amino acids indicating that our standard provides a reasonable surrogate to the class I immunopeptidome. When singly charged peptides ($m/z \geq 700$) are included in the precursor selection window shorter

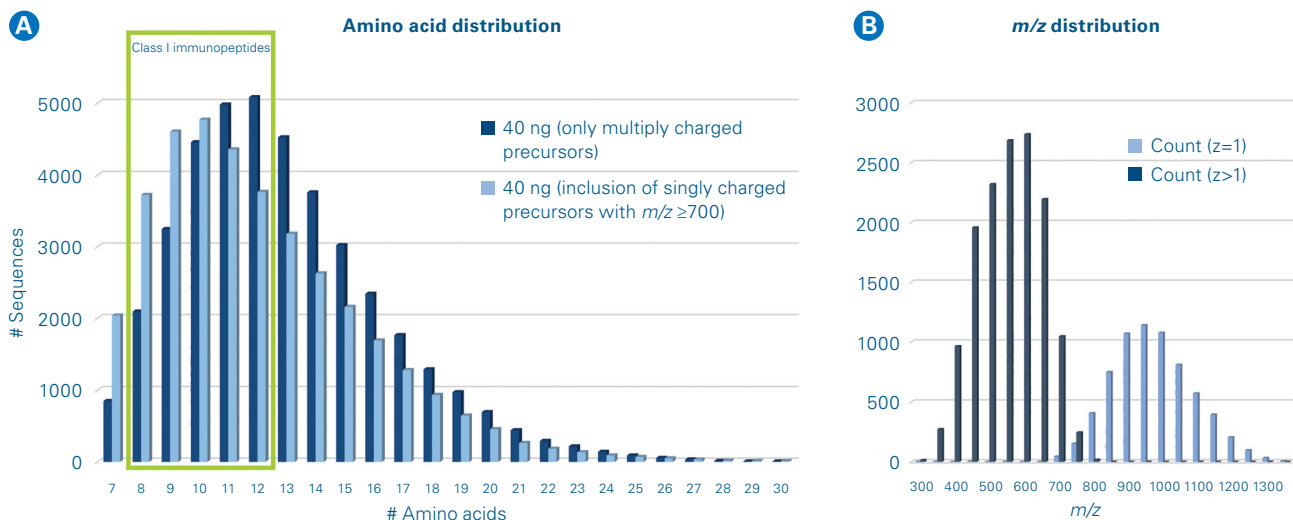


Figure 2

(A) Amino acid distribution of HeLa lysate digested with elastase with and without inclusion of singly charged precursors. **(B)** m/z distribution of peptides with 8-12 amino acids with inclusion of singly charged precursors.

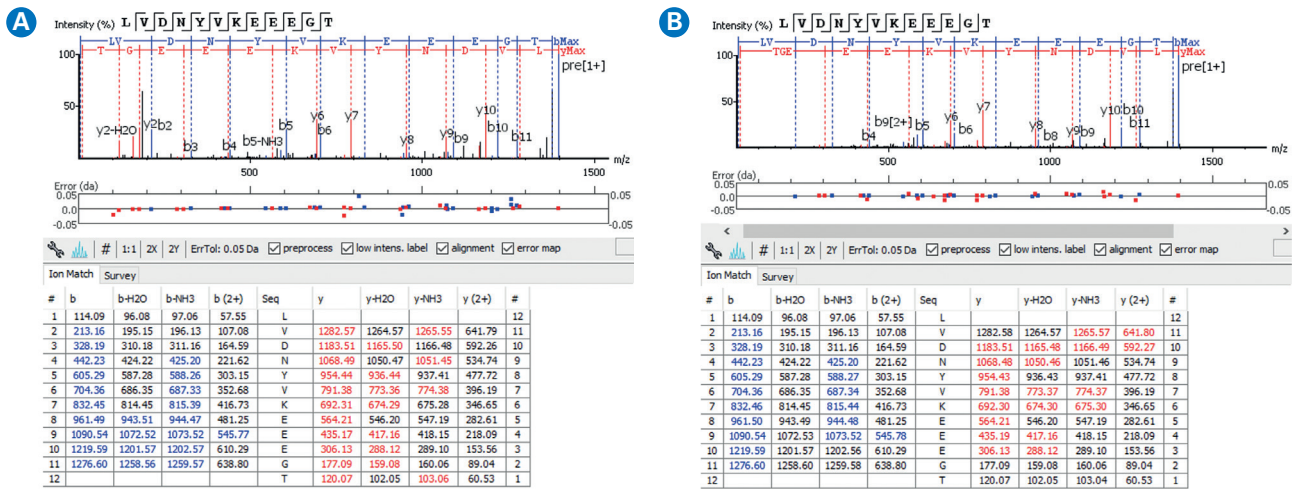


Figure 3 MS/MS spectra of LVDNYVKEEEGT with TIMS Stepping enabled (A) and without TIMS Stepping (B).

(7-10 amino acids) peptides are identified (Figure 2A). It was also evident that the majority of singly charged short peptides (8-12 amino acids) identified derived from singly charged precursors ($m/z \geq 700$) (Figure 2B). These results indicate that using a specific ion mobility window to include only high m/z singly charged features shifts the focus of the MS/MS sequencing, and improves the identification of peptides, that at least in length, resemble the class I immunopeptidome.

TIMS stepping can improve b/y ion series

De novo sequencing scores can be used to support database search results and improve confidence in immunopeptidomics data. Efficient *de novo* sequencing requires high quality MS/MS spectra with a near full ion series. In the timsTOF the ion transfer parameters determine which m/z range is efficiently detected, by default <200 m/z are usually not detected in MS/MS spectra. If TIMS stepping is activated, two ramps with different ion transfer settings are measured and combined resulting spectra consist of both low and high mass fragment ions with sufficient intensity able to improve *de novo* sequencing. We investigated if TIMS stepping could improve peptide identification in the elastase standard and therefore be applicable to immunopeptidomics. Figure 3 shows the assignment made by PEAKS for MS/MS spectra acquired with the default

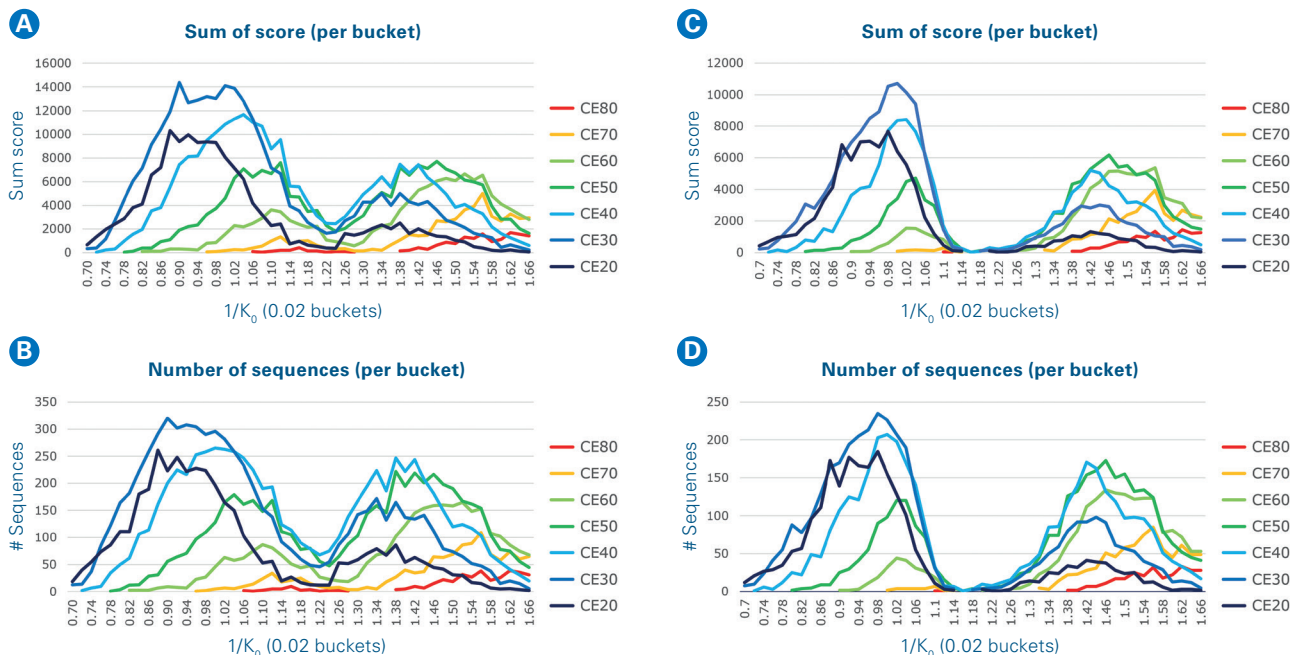


Figure 4 (A) Summed score of all identified sequences and (B) number of identified sequences per 0.02 bucket ($1/k_0$). (C) Summed score of identified sequences with an amino acid length of 8-12 and (D) corresponding number of identified sequences per 0.02 bucket ($1/k_0$).

settings for ion transfer and with TIMS stepping activated (Figure 3 A,B). With TIMS stepping activated ions in the lower m/z range are now also detected as expected (Figure 3 A,B). However, the overall score of the identified sequences could not be improved by TIMS stepping (data not shown) and the number of b and y fragments in the higher m/z range seem to be sufficient for the identification of peptides. When comparing the results with and without TIMS stepping, many more peptides were identified without TIMS stepping. This is likely due to the cycle time, which doubles when TIMS stepping is activated, so the number of PASEF ramps has to be halved, which reduces the number of different precursors selected in the cycle. In summary, the standard ion transfer settings for bottom-up proteomics are also acceptable for immunopeptidomic analyses.

Collision energy optimization with focus on singly charged precursors

Next, we focused on the optimization of collision energies especially for singly charged precursors. In PASEF the collision energies are set according to the ion mobility values of precursors. By using set CE values (20, 30, 40, 50, 60, 70, 80 eV) and multiple injections of 40 ng HeLa lysate with fast 20 min gradients the effect of CE on peptide identification rates and PEAKS score at different ion mobility buckets was determined. Plots with the summed score of identified sequences per bucket and the number of sequences per bucket are shown in (Figure 4 A-D). To evaluate the best performing collision energies, two collision energy gradients were created. The first was based on the highest summed score per bucket, the second one on the highest summed score and average score per bucket, with default collision energies set for bottom-up proteomics as a reference (Table 1). Using standard injection of 40 ng over the standard 66-minute gradient best summed score strategy gave the highest overall number of peptide identifications, this effect was pronounced at higher ion mobility values where singly charged species are found (Figure 4 A-D). Overall CE optimization resulted in a modest 8% improvement over default strategy developed for proteomics (Figure 5). As the number of identifications and their summed score in the range of $1/k_0 = 1.10$ -1.22 is comparable in all collision energy gradients, the final gradient has been flattened in this range (Table 1 in brackets).

Table 1
Collision energies optimized according to highest summed score and highest summed score in combination with highest average score.

timsTOF SCP		
	$1/K_0$	CE [eV]
CE (based on summed score)	0.70	20 (20)
	1.06	30 (30)
	1.10	48 (40)
	1.16	41 (-)
	1.34	40 (40)
	1.67	70 (70)
CE (based on summed and average score)	0.70	20
	1.06	30
	1.10	48
	1.18	48
	1.68	70
Default	0.6	20
	1.6	59

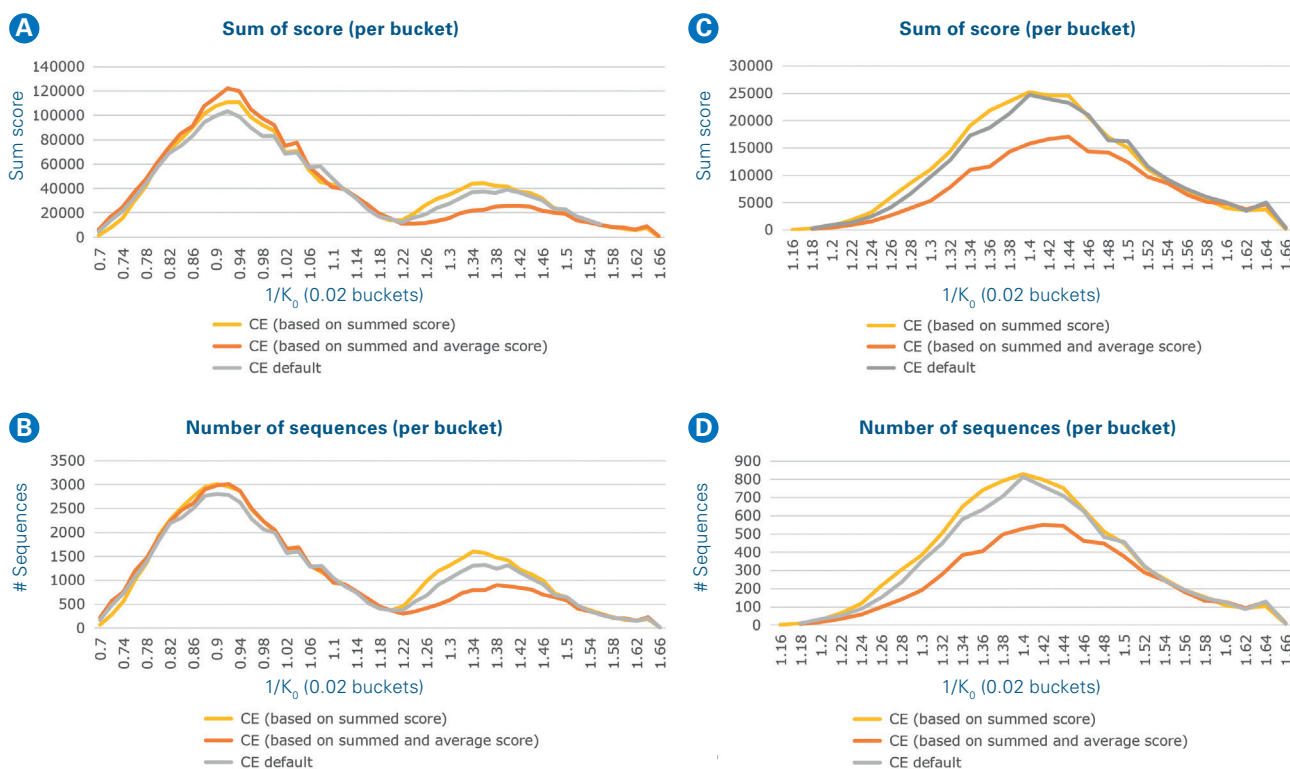


Figure 5

(A) Summed score of all identified sequences and (B) number of identified sequences per 0.02 bucket ($1/k_0$). (C) Summed score of identified sequences with an amino acid length of 8-12 and $z=1$ and (D) corresponding number of identified sequences per 0.02 bucket ($1/k_0$) – all duplicate identifications removed.

Sensitivity for immunopeptidomics samples

Precursors with low abundance can be measured multiple times to increase their signal-to-noise ratio during a PASEF cycle. To improve the quality of MS/MS spectra of low-abundant precursors, the target intensity can be increased to allow more MS/MS repetitions for these signals. On the other hand, lowering the intensity threshold could lead to more low abundant precursors being included in the precursor selection, which could be useful for small sample amounts. Combinations of intensity thresholds of 100, 250, 500 with target intensities of 20,000 and 40,000 were assessed to evaluate the MS/MS spectra quality in the HeLa elastase digest (40 ng sample load, 66 min gradient).

At intensity thresholds of 100, 250 or 500 increasing the target intensity to 40,000 reduced the number of sequences identified by 6-20% (Figure 6A-C). The lower number of peptide identifications is most likely due to the higher number of repetitions per precursor required to meet the target intensity, leading to less time for selecting different precursors for fragmentation, especially in combination with a very low intensity threshold. When comparing the three intensity thresholds (100, 250 and 500) with a target intensity of 20,000, the number of sequences identified was comparable (Figure 6D). Overall, an intensity threshold of 500 and target intensity of 20,000 leads to MS/MS spectra of high quality and peptide identification rates that are optimal for immunopeptidomic analyses.

Next the sensitivity of the SCP was evaluated by a serial dilution of the elastase HeLa digest (40 ng to 0.0625 ng) using optimal values for CE, intensity threshold and target intensity. We could readily detect 600 peptides from as little as 62.5 pg of peptides (Figure 7A). Finally, we compared optimal intensity/target thresholds over this serial dilution with a combination of lower intensity threshold/target intensity (250/14000) to see if lower values could improve sensitivity whilst maintaining the number of MS/MS repetitions (Figure 7B). A small improvement in peptide identifications was observed at peptide loads 2 ng (2277 → 2415) and 10 ng (6704 → 6998) with no effect seen at higher or lower loads whilst PEAKS scores were unaffected by these adjustments.

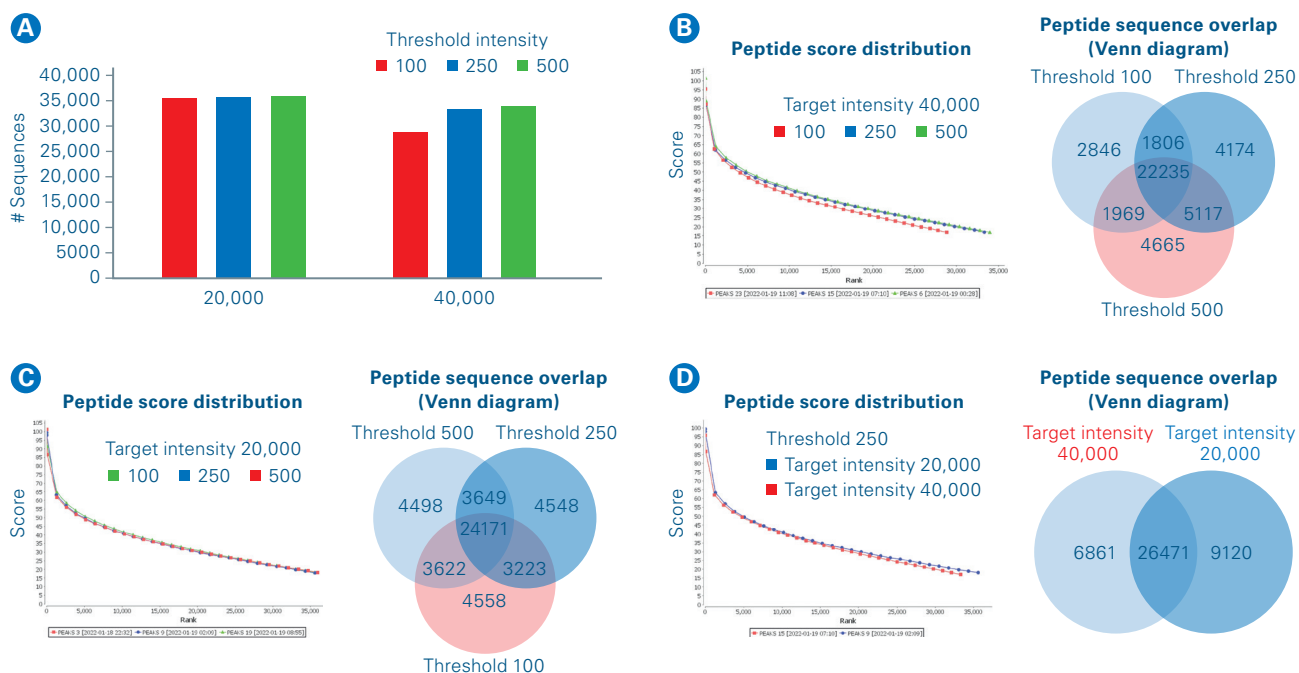


Figure 6

Total peptide sequences identified, PEAKS score distribution and overlap for the combination of different Target Intensity and Threshold Intensity tested.

A Bar plot showing the total number of peptides identified. **B-C** Line plots for PEAKS score distribution and Venn diagram for sequences identified at Target Intensity fixed to either 20,000 or 40,000 with varying Thresholds (100, 250, 500) and **D** Thresholds fixed to 250 with varying Target Intensity (20,000 and 40,000).

Validation and assessment of sequencing bias in synthetic HLA peptide standard

To begin validation of these methods we analysed a pool of 2000 synthetic HLA peptides (20 fmol/peptide, 66 min gradient). We compared the optimized methods “Pol” that included the singly charged precursors ($m/z \geq 700$) and “CE2” with additionally optimized collision energies as described above, using default proteomics as a reference (Table 2). In total 82% of the target sequences could be identified and 87% of all sequences identified were expected as either targets of synthesis or sub-sequences of these targets (Figure 8A). Inclusion of 1+ ions (Pol & CE2) gave the highest number of overall peptides including 394 unique sequences (Figure 8B). Stratification of these results by allele indicated that inclusion of 1+ ions (Pol & CE2) was particularly effective in identifying hydrophobic sequences that bind A02:01 HLA compared to other methods (Figure 8C). This effect was less pronounced in A03:01 peptides that like tryptic peptides possess a C-terminal K/R residue (Figure 8C).

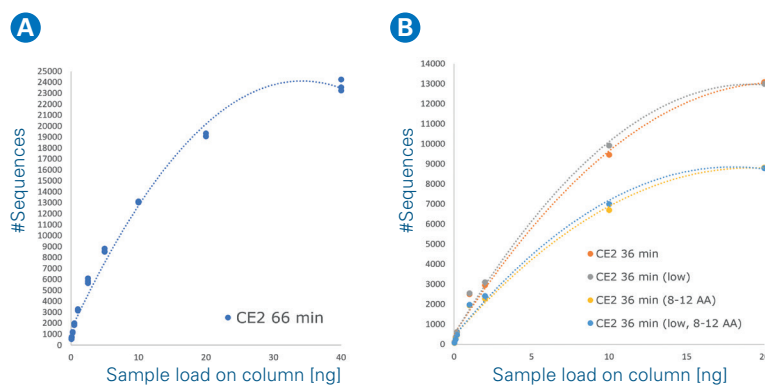


Figure 7

(A) Dilution series of HeLa lysate (0.02 ng – 40 ng sample load) measured with a 66 min gradient (3 replicates) using default thresholds (intensity threshold = 500, target intensity 20,000) and **(B)** 36 min gradient analysis of dilution series of HeLa lysate (0.02 ng – 20 ng sample load) with default (Orange, Yellow) and lower thresholds (Gray, Blue) (low, intensity threshold = 250, target intensity 14,000) stratified by peptide amino acid (AA) length (All: Gray and Orange or 8-12: Yellow and Blue).

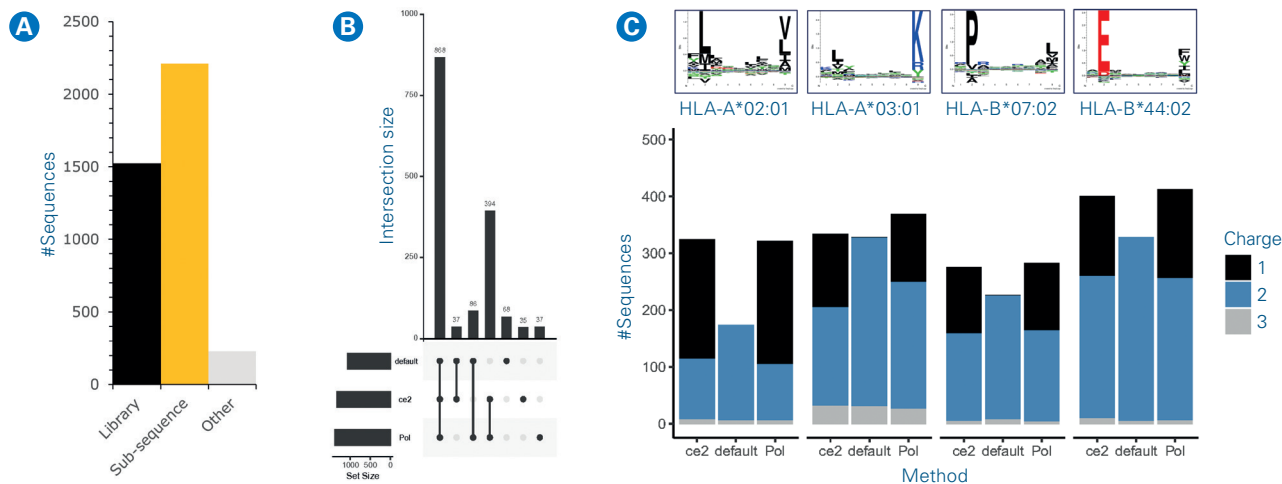


Figure 8

(A) Overview of identified peptide from the library. **(B)** Peptide sequences overlapping between methods. **(C)** Identified sequences for individual methods by allele comparison and charge state.

Table 2

Methods used for the analysis of synthetic HLA peptides

Method	Polygon	$1/K_0$	Ramp time	PASEF ramps	Cycle time	Target intensity	Intensity threshold	$1/K_0$	CE
Pol	$z=1>700$ m/z + Multicharged	0.7-1.7	166	10	1.89	20,000	500	0.6	20
								1.6	59
CE2	$z=1>700$ m/z + Multicharged	0.7-1.7	166	10	1.89	20,000	500	0.70	20
								1.06	30
								1.10	40
								1.34	40
								1.68	70
Default	Multicharged	0.7-1.7	166	10	1.89	20,000	500	0.6	20
								1.6	59

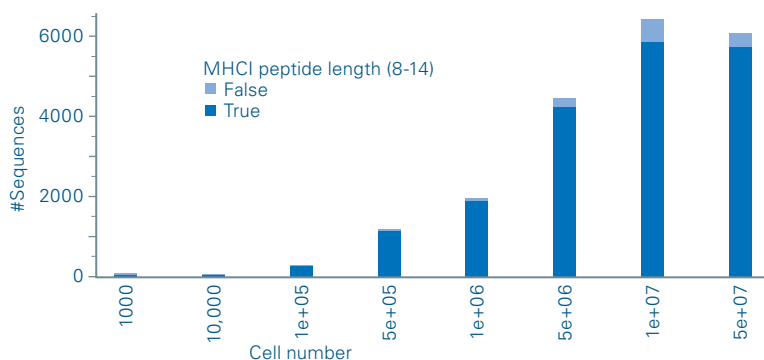


Figure 9
Overview of identified immunopeptides prepared from a Jurkat cell-line.

Finally, to translate these results into the laboratory a serial dilution (0.000936-46.8 ng) that corresponds to 1000 to 50 million cells of an immunopeptidomic sample prepared from the Jurkat cell-line was analyzed. A low number (15-20) of likely HLA-peptides (8-14 amino acids in length) could be detected with a peptide load of 0.0936 ng that was equivalent to 100,000 cells, with an exponential increase in peptide number (4075-4315) that appeared to reach saturation at 9.36 ng equivalent to 10,000,000 cells (Figure 9). The majority of identified immunopeptides had a peptide length of 8-14 amino acids.

Conclusion

- The timsTOF SCP provides a robust and sensitive platform for immunopeptidomics
- The MS method was optimized with respect to precursor selection, collision energies, intensity threshold, and target intensity to improve immunopeptide detection and identification.
 - Inclusion of singly charged species enhances peptide identification and repertoire of detectable HLA peptides
 - Collision energy optimisation enhances identification of singly charged species
- The optimized method showed high sensitivity for the standard sample, the immunopeptidome samples derived from the Jurkat cell-line and on cervical cancer cell lines and culture as shown by Peng et al [5] recently.

References

- [1] Chong C, Coukos G, Bassani-Sternberg M, 2022. *Identification of tumor antigens with immunopeptidomics*. Nat Biotechnol, **40**(2):175-188. doi: 10.1038/s41587-021-01038-8. Epub 2021 Oct 11. PMID: 34635837.
- [2] Purcell AW, Ramarathinam SH, Ternette N 2019. *Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics*. Nat Protoc, **14**(6):1687-1707. doi: 10.1038/s41596-019-0133-y. Epub 2019 May 15. PMID: 31092913
- [3] Brunner AD, Thielert M, Vasilopoulou C, Ammar C, Coscia F, Mund A, Hoerning OB, Bache N, Apalategui A, Lubeck M, Richter S, Fischer DS, Raether O, Park MA, Meier F, Theis FJ, Mann M 2022. *Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation*. Mol Syst Biol, **18**(3):e10798. doi: 10.15252/msb.202110798. PMID: 35226415; PMCID: PMC8884154.
- [4] Parker R, et al. *The Choice of Search Engine Affects Sequencing Depth and HLA Class I Allele-Specific Peptide Repertoires*. Molecular & Cellular Proteomics, **20**,100124; DOI:https://doi.org/10.1016/j.mcpro.2021.100124
- [5] Peng X, et al. *Novel Canonical and Non-Canonical Antigens That Extend Current Targets for Immunotherapy of HPV-Driven Cervical Cancer*. <http://dx.doi.org/10.2139/ssrn.4022700>

For Research Use Only. Not for use in clinical diagnostic procedures.

Bruker Switzerland AG

Fällanden · Switzerland
Phone +41 44 825 91 11

Bruker Scientific LLC

Billerica, MA · USA
Phone +1 (978) 663-3660

