

Article

Not peer-reviewed version

Efficient Quality Control of Peptide Pools

Gaby Bosc-Bierne , Shireen Ewald , Oliver J. Kreuzer , [Michael G. Weller](#) *

Posted Date: 15 April 2024

doi: 10.20944/preprints202404.0992.v1

Keywords: peptide pools; quality control; UHPLC-UV-HRMS; relative quantification; substance confirmation; structure; cost efficiency; CEF; infectious diseases; orbitrap



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Efficient Quality Control of Peptide Pools

Gaby Bosc-Bierne ¹, Shireen Ewald ¹, Oliver J. Kreuzer ² and Michael G. Weller ^{1,*}

¹ Federal Institute for Materials Research and Testing (BAM), Division 1.5 Protein Analysis, Richard-Willstätter-Strasse 11, 12489 Berlin, Germany.

² Peptides & Elephants GmbH, Neuendorfstraße 20b, 16761 Hennigsdorf, Germany

* Correspondence: michael.weller@bam.de; Tel.: +49-30-8104-1150

Abstract: Peptide pools composed of short amino acid sequences have proven to be versatile tools in various research areas and clinical applications. They are powerful tools for epitope mapping, immunotherapy, and vaccine development. Their importance lies in their ability to map complex protein structures, enabling a comprehensive understanding of immune responses and facilitating the identification of potential therapeutic agents. The application of peptide pools also extends to the field of personalized medicine, offering tailored solutions for diseases such as cancer and infectious diseases. Peptide pools are complex mixtures of immunostimulatory antigens primarily intended for T-cell stimulation. They are commercially available in many different compositions and variants. However, unlike other reagents that consist of only one or a few compounds, peptide pools are highly complex products with limited stability. This makes their quality control a major challenge. Quantitative peptide analysis usually requires sophisticated methods, in most cases isotope-labeled standards and reference materials. None of these are routinely available for these products. Synthesis and purification of all labeled peptides might be required. Usually, this approach would be prohibitively laborious and expensive. Therefore, an approach is needed to provide a practical and feasible method for quality control of peptide pools. With insufficient quality control, the use of such products could lead to incorrect experimental results, which would worsen the well-known reproducibility crisis in the biomedical sciences. Here we propose the use of ultra-high performance liquid chromatography (UHPLC) with two detectors, a standard UV detector (at 214 nm) for quantitative analysis and a high-resolution mass spectrometer (HRMS) for identity confirmation. To be cost-efficient and fast, quantification and identification are performed in one chromatographic run. An optimized protocol is shown, and different peak integration methods are compared and discussed. This work was performed using a peptide pool known as CEF, which consists of 32 peptides derived from cytomegalovirus (CMV), Epstein-Barr virus (EBV) and influenza virus, ranging from 8 to 12 amino acids in length.

Keywords: peptide pools; quality control; UHPLC-UV-HRMS; relative quantification; substance confirmation; structure; cost efficiency; CEF; infectious diseases; orbitrap

1. Introduction

Peptide pools are mixtures of short peptides that are mainly used in experimental immunology. Important applications include monitoring cellular immune responses, such as those following infection or vaccination. It is also used for in-vitro T-cell activation and expansion in adoptive cancer immunotherapies, as well as for epitope discovery in vaccine development. It is extremely important that the composition of the peptide pools corresponds exactly to the design and purpose of the peptide pool. For example, a peptide pool for personalized cancer immunotherapy must safely contain all peptide neoantigens. In addition, no other peptides should be present in the pool. These can otherwise lead to incorrect results [1]. Some manufacturers address this problem by double-checking the composition of the pools. One employee adds the peptide and a second employee checks that it is the correct peptide. Despite careful production methods, it is desirable to have an independent method for the correct composition of the peptide pool. In addition, some peptides in a

peptide pool can also change over time. The quality of peptide pools may be compromised by synthesis errors, degradation, adsorptive losses, dimerization, and other unwanted chemical modifications. Consequently, these issues can lead to misleading experimental results and poor experiment reproducibility [2]. Therefore, the development of improved methods for quality control (QC) of peptide pools is highly desirable. However, in many cases, routinely performed quality control protocols rely mainly on pre-characterizing individual peptides. Peptide quantification is commonly performed using HPLC-UV or the use of stable isotopically labeled (SIL) internal standards in HPLC-MS. However, this approach is not easily applicable to complex peptide pools. Quantification using HPLC-UV is difficult in complex samples, such as mixed peptide pools, due to a frequent lack of baseline separation. This issue is aggravated by the potential presence of an unknown number of synthesis byproducts and degradation products. A common analytical method to resolve complex samples is chromatography combined with mass-spectrometry and in particular HPLC-MS/MS. The use of tandem mass spectrometry (MS/MS), usually as a triple quadrupole (QqQ) system with standard resolution, achieves high selectivity by using specific mass transitions. This approach leads to excellent selectivity and can be considered the gold standard for quantification. However, these methods usually require isotopically labelled peptides for each of the peptides in the pool. This is a huge cost factor for QC considering the many different peptide pools on the market. During the last years, high-resolution mass spectrometry (HRMS) has become more and more popular, due to its excellent mass precision and accuracy. It achieves high selectivity by utilizing a narrow mass extraction window (MEW). Also very attractive is the ability to computationally derive the elemental composition of ions, which leads to a strong support of the analyte's identity and structure without the need of independent standard compounds or isotopically labelled compounds. Both methods have their advantages and disadvantages, but they both face common challenges in mass spectrometric quantification: Compound-dependent ionization efficiency and ion suppression in the ion source due to co-eluting substances, which may lead to strong matrix effects [3].

For the latter problem, various solutions exist, which can be categorized as either reduction or compensation strategies. Reduction of matrix effects can be achieved through sample clean-up (e.g., extraction) [4], separation (chromatography) [5], or sample dilution (*dilute and shoot*) [6]. Compensation for matrix effects can be achieved through matrix-matched calibration [7], standard addition [8], the echo peak technique [9], and internal standardization [10]. When ion suppression is caused by co-eluting analytes, neither sample clean-up nor dilution may be effective. Chromatographic separation of such complex mixtures may be limited, even when UHPLC is applied. In addition, some peptides are poorly ionized in MS, and hence a quantitation is not possible without a standard compound or other independent determination of the response factors of each peptide. Some peptides are fully absent in MS for various reasons. A complete mass spectrometric coverage of a protein digest was therefore deemed unrealistic [11]. In the context of a peptide pool, this would mean that some peptides would be missing and therefore could not be qualified at all. Among the mentioned compensation methods, matrix-matched calibration and standard addition are considered impractical due to calibration effort. The echo peak technique is quite difficult for multi-analyte analysis. Finally, internal standardization remains, which is a highly effective compensation method when using stable isotope-labeled analogues of the analyte. However, these are often either not commercially available or prohibitively expensive, so that they are usually not economically viable in the case of multi-analyte methods. Consequently, this study aims to explore cost- and time-efficient options for quality control of synthetic peptide pools. The hyphenated technique UHPLC-UV-HRMS, as shown in **Figure 1**, was used for this purpose. However, unlike conventional approaches, quantification and identification were performed separately using two different detectors.

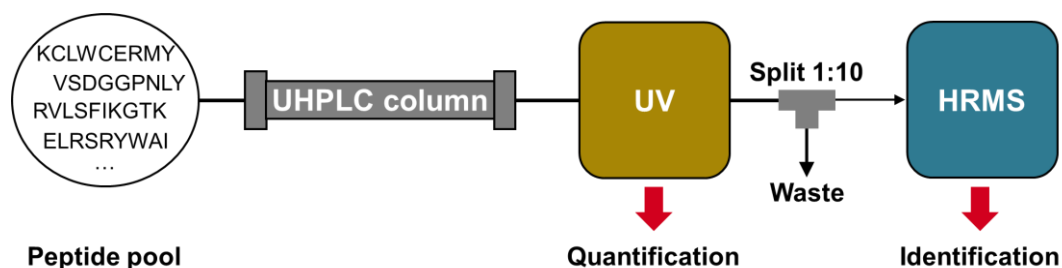


Figure 1. Simplified experimental setup for a cost- and time-efficient analysis of a peptide pool. Each of the two detectors serves a specific purpose and is used in combination. Quantification is performed using an ultraviolet (UV) detector, here at 214 nm, while identification is performed using high-resolution mass spectrometry (HRMS). Expensive internal standards labeled with stable isotopes, as well as calibration and identification runs are usually not required. However, it should be noted that the sensitivity of the UV and MS detector may need to be adjusted once during the setup, for example by using different splitting ratios, MS conditions and/or online dilution (not shown).

The separation of quantification and identification between the UV detector and mass spectrometer provides several benefits. Firstly, it eliminates the need for stable isotope-labeled internal standards. MS-based identification eliminates the need for additional HPLC identification runs to determine the retention time (t_R) of each peptide, which might require all peptides in pure form, and enables non-target analysis of any other compounds, such as synthesis byproducts or degradation products, if necessary. The method was developed using a common peptide pool of 32 peptides (CEF peptide pool [12,13]). The optimal chromatographic results were obtained using a shallow gradient, elevated temperature, moderate flow rate, and the ion-pairing reagent TFA at low concentrations of 0.05% and 0.04% to flatten the baseline and to reduce ion suppression by TFA. Relative quantification could be achieved from the 214 nm UV signal, using the integration by the perpendicular drop method combined with the formation of a limited number of peak groups, or alternatively by using peak fitting algorithms. High-resolution mass-spectrometry (*Orbitrap*) delivered a peak and compound assignment by comparison with the expected peptide mass (to charge ratio). As a bonus of the HRMS method, a total of at least 45 unexpected compounds could be tentatively identified.

2. Materials and Methods

2.1. Chemicals

CEF advanced peptide pool (purity >90%), containing 32 peptides selected from specific HLA class I-restricted T-cell epitopes of cytomegalovirus (CMV), Epstein-Barr virus (EBV), and influenza virus, was obtained from peptides & elephants GmbH (Hennigsdorf, Germany). LC-MS grade trifluoroacetic acid (TFA) >99.5% (85183) was purchased from Thermo Fisher Scientific, Life Technologies GmbH (Darmstadt, Germany). LC-MS grade acetonitrile (ACN) was obtained from Th. Geyer GmbH & Co. KG (Renningen, Germany). Helium 6.0 (10100530) was purchased from Linde AG (Berlin, Germany). LC-MS grade LTQ ESI positive ion calibration solution >99.5% (88322) was obtained from Thermo Fisher Scientific GmbH (Dreieich, Germany). Water was purified by an Ultra-Pure Water System from Millipore Co., (Burlington, MA, USA), with a resistivity of 18.2 M Ω ·cm.

2.2. Sample Preparation

The lyophilized CEF advanced peptide pool sample (0.8 mg) was stored at -20°C and equilibrated at room temperature (RT) for 30 minutes before use. Afterwards, the sample was dissolved in 50 μ L of ACN/TFA 0.05% (v/v) and then in 50 μ L of water/TFA 0.05% (v/v) to obtain a stock solution of 8 μ g/ μ L (approximately 0.25 μ g/ μ L per peptide). Before use, both solvents were degassed by sonication (45 kHz, 15 min, RT). After 15 minutes of equilibration, the stock solution was divided into 20 μ L aliquots and stored in 0.5 mL polypropylene micro tubes at -20°C. Before each LC-

MS analysis, the 20 μL aliquots were thawed at room temperature and diluted with 180 μL of water/TFA 0.05% (v/v) to a final concentration of 0.8 $\mu\text{g}/\mu\text{L}$ (approximately 0.025 $\mu\text{g}/\mu\text{L}$ per peptide). After centrifugation (6000 rpm, 5 min, RT), 180 μL of the supernatant was transferred to an amber HPLC glass vial with a 200 μL glass insert.

The solvents used to dissolve the peptide pool were also utilized to prepare blank samples. After diluting 100 μL of a 1:1 mixture of water/TFA 0.05% (v/v) and ACN/TFA 0.05% (v/v) with 900 μL of water/TFA 0.05% (v/v), a blank solution with an ACN concentration of 5% was obtained. After centrifugation (6000 rpm, 5 min, RT), 800 μL of the supernatant was transferred to an amber HPLC glass vial and stored at -20°C .

2.3. UHPLC-HRMS Analysis

Peptide pool samples were analyzed on a Thermo Scientific UltiMate 3000 RSLCnano UHPLC System hyphenated to a Thermo Fisher Scientific Exactive Orbitrap High-Resolution Mass Spectrometer. The autosampler temperature was set to 4°C . Chromatographic separation was conducted using an Acclaim PepMap RSLC C18 column (100 \AA , 2 μm , 0.3 mm \times 150 mm) with an Acclaim PepMap RSLC C18 trap column (100 \AA , 5 μm , 0.3 mm \times 5 mm), both from Thermo Scientific. The mobile phases for chromatography consisted of 0.05% (v/v) TFA in water (A) and 0.04% (v/v) TFA in acetonitrile (B). The mobile phases were degassed by purging with helium for 5 minutes. The peptides were eluted using the following gradient: 4% B isocratic for 4 min (6 $\mu\text{L}/\text{min}$), linear increase to 44% B over 100 min (6 $\mu\text{L}/\text{min}$), linear increase to 95% B over 0.1 min (6 $\mu\text{L}/\text{min}$), held at 95% B for 6 min (10 $\mu\text{L}/\text{min}$), returned to initial conditions within 0.1 min (10 $\mu\text{L}/\text{min}$), and maintained at a flow rate of 10 $\mu\text{L}/\text{min}$ for 11 min and 6 $\mu\text{L}/\text{min}$ for 1 min. The column oven temperature was set at 55°C . The injection volume was 1 μL , conducted in full loop mode, with a flush volume of 5 μL , a flush volume 2 of 3 μL , and a loop overflow of 2 μL . UV detection was performed at 214 nm. Before analyzing the peptides in the mass spectrometer, the eluate was split in a 1:10 ratio using a T-piece. Subsequent electrospray ionization was conducted in positive mode (ESI+). The mass spectrometer settings were adjusted as follows: spray voltage at 4.5 kV, capillary voltage at 30 V, capillary temperature at 320°C , tube lens voltage at 80 V, skimmer voltage at 24 V, and the scan range from 200 to 1800 m/z. The MS scans were acquired with ultra-high resolution (100,000 at 200 m/z) at a scan rate of 1 Hz, balanced automatic gain control (AGC) target (1×10^6), and a maximum injection time (IT) of 500 ms. External mass calibration using LTQ ESI positive ion calibration solution provided a mass accuracy of <5 ppm. Data acquisition was performed using Xcalibur 2.2 with Dionex Chromatography Mass Spectrometry Link (DCMS^{Link}).

2.4. Data Processing

Analysis of mass spectrometry data was conducted using Xcalibur 2.2. Theoretical exact masses and m/z values of the peptides were calculated using the ChemCal [14]. Experimental m/z values were confirmed with a maximum mass tolerance of 5 ppm.

UV spectrometry data was analyzed either using Chromeleon 7.2.10 or PeakFit 4.12. In Chromeleon, baseline correction was performed by subtracting a blank run injection. A baseline with a defined start and end point was added. For peak detection and integration, a minimum area was defined. Peak areas were measured by the perpendicular drop method. In PeakFit, prior to peak fitting, a baseline is obtained using PeakFit's automated baseline fitting option (2nd Deriv Zero algorithm, BEST, tolerance 10%). The active baseline points were selected manually. To identify and fit peaks, the AutoFit Peaks I Residuals option was selected. The data for peak detection was smoothed using the Savitzky-Golay algorithm, and the optimal smoothing level was automatically determined by the AI Expert option (0.05%). Peaks were fitted using the EMG+GMG peak function. The amplitude rejection threshold was set at 2.5%, and residuals were added. Graphical adjustments were made alongside automatic peak placement. Peak fit preferences were adjusted with the following settings: Maximum iterations = 10,000, convergence to significant digits in Chi-Square = 6, built-in peak functions constraints at $a_0 = 99$, $a_1 = 100$, $a_3 = 140$, $a_4 = 140$, fit extent = 1/1. As a minimization procedure, the least-squares method was chosen. For constructing the curvature

matrix, a sparse root-finding procedure was selected. The fits were rerun until the coefficient of determination (r^2) did not change in the second decimal place.

Sample peak capacities were calculated using the equation PC_s (sample peak capacity) = t_r (retention time of the last peak) - t_i (retention time of the first peak)/FWHM (peak width at 50% height). Peak widths were calculated for well-resolved peptides: peptide 3 for gradient optimization, peptide 31 for TFA concentration optimization, and peptide 1 for temperature optimization.

3. Results and Discussion

3.1. Optimized UHPLC Parameters

For the quantification of the peak areas obtained by a UV detector, baseline separation of all individual components would be desirable. For a fast check, the peak capacity of a separation can be calculated [15,16]. Peak capacity is considered the preferential measure of the performance of a gradient separation [17]. As the most general definition, the peak capacity is the number of peaks that can be separated within a retention window. In practical terms, the peak capacity is the gradient run time divided by the average peak width. In a sample with a given number of components, the larger the peak capacity, the more likely it is for all components to be separated without critical overlap [18]. Accordingly, in optimizing the general UHPLC method, the goal was to achieve a high peak capacity. Since changes in separation conditions affect both peak width and the retention window [19], it is preferable to use the sample peak capacity (PC_s) related to the retention window as defined by Dolan et al. [20] rather than the peak capacity related to gradient time. The concept of peak capacity is based on the idealized assumption that all peaks occur at regular intervals. However, usually peaks are more or less randomly distributed. Consequently, the highest peak capacity does not necessarily result in the best separation of a specific sample.

3.1.1. Optimizing Gradient Slope

In general, a lower gradient slope is expected to result in higher gradient retention factors, leading to increased resolution and peak capacity. Therefore, the initial gradient slope of 1.7 %B/min was reduced to 0.4 %B/min. The results of the different gradient slopes are shown in **Figure 2**:

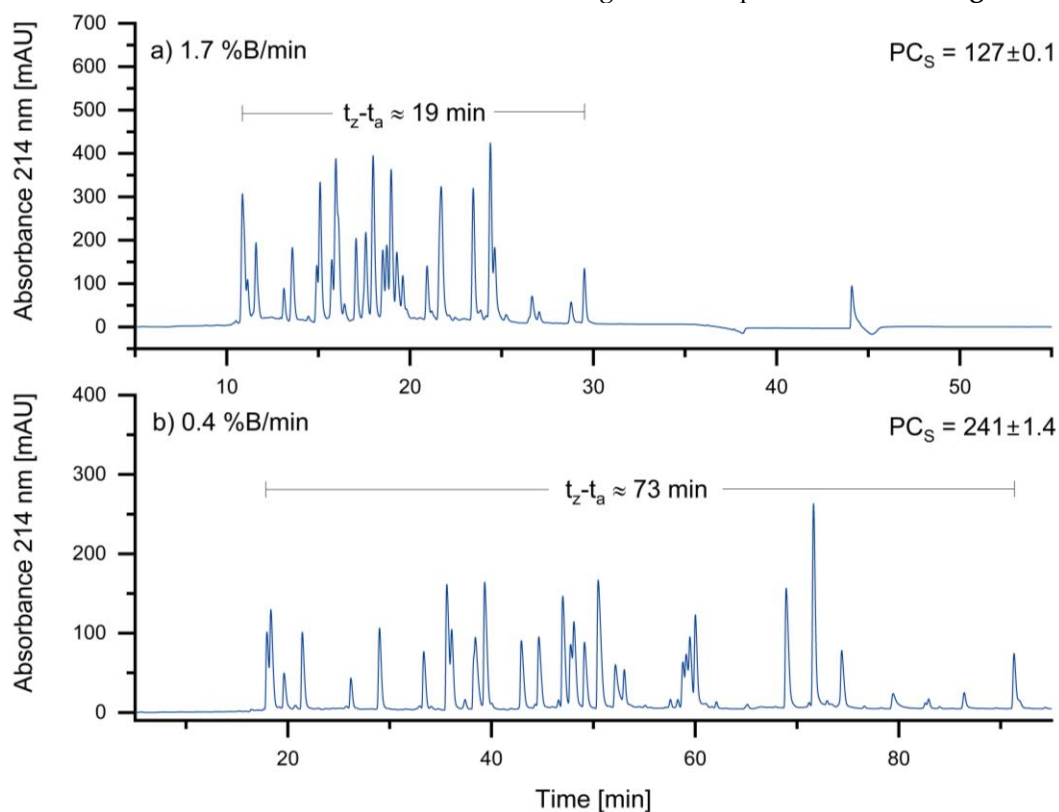


Figure 2. Impact of decreasing gradient slope on the separation (sample peak capacity, PCs) of the CEF advanced peptide pool, consisting of 32 peptides. a) 1.7 %B/min; b) 0.4 %B/min. The sample peak capacity increased as the gradient slope decreased. The reported values represent the average of technical duplicates (n=2). The peptide sequences are listed in **Table 2**. The column oven temperature was set at 40°C, with all other UHPLC conditions as described in section 2.3.

As expected, the sample peak capacity increases as the slope of the gradient decreases. Based on these results, a gradient of 0.4 %B/min was selected for the standard method. An even lower gradient was not considered due to the associated long run times.

3.1.4. Optimizing Concentration of the Ion-Pairing Reagent TFA

In these experiments, the concentration of TFA was varied. Typically, TFA is used for peptide separations at concentrations ranging from 0.05% to 0.1% (v/v). However, Chen et al. [21] argued that this concentration range is not ideal for most peptide applications. Hence, TFA concentrations of 0.05%, 0.1%, and 0.2% were tested in eluent A. In order to counteract baseline drift, the TFA concentration in eluent B was reduced by 20%. The results are displayed in **Figure 3**:

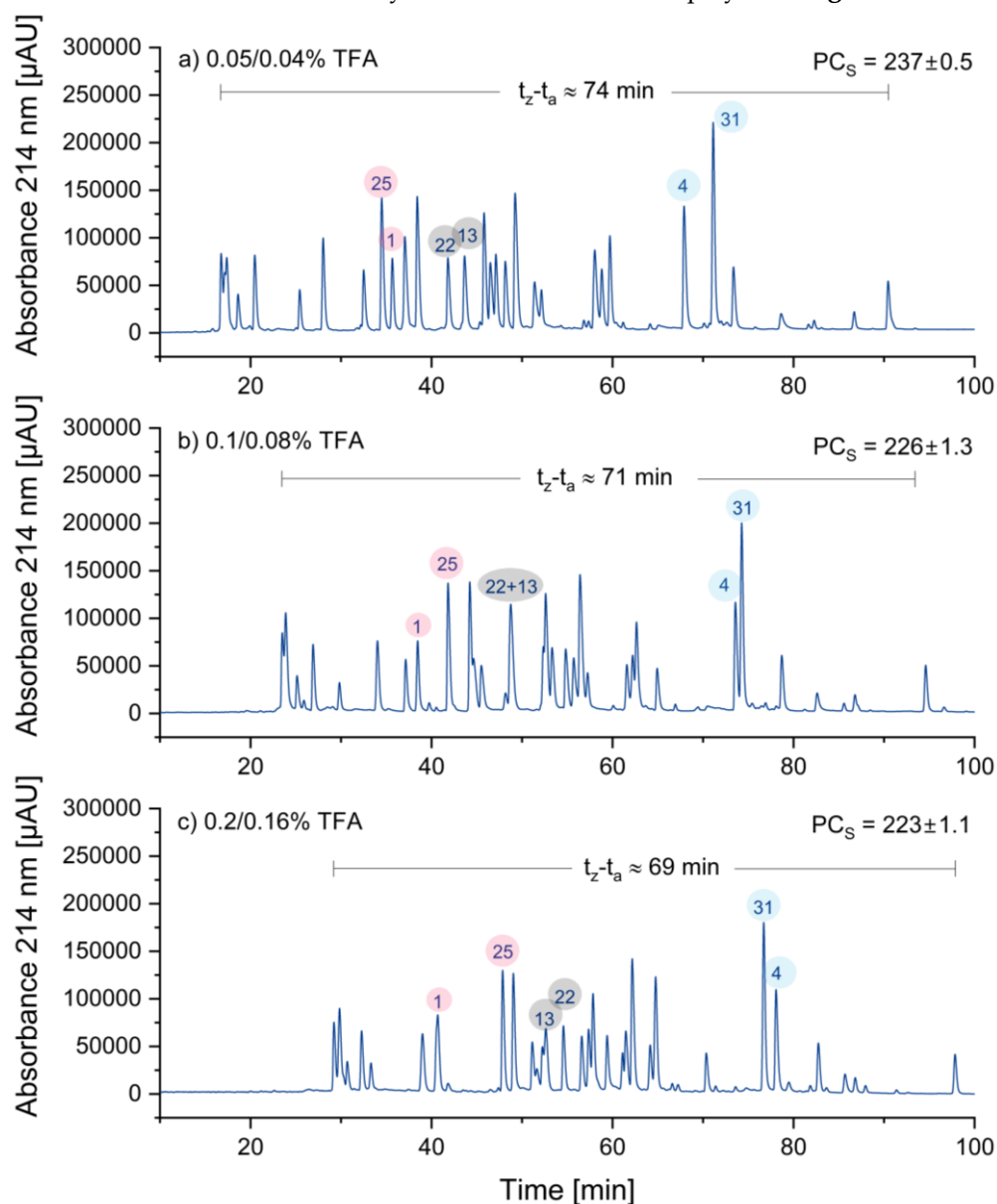


Figure 3. Influence of increasing TFA concentration on the UHPLC separation (sample peak capacity, PCs) of the peptide pool containing 32 peptides. a) 0.05/0.04%; b) 0.1/0.08%; c) 0.2/0.16%. The

concentration values pertain to eluent A/eluent B. Higher concentrations of TFA resulted in longer retention times, and a shift in selectivity, as evidenced by the altered elution order of peptide pairs 1/25, 22/13, and 4/31. The reported values represent the average of technical triplicates (n=3). The peptide sequences are listed in **Table 2**. The column oven temperature was set at 40°C, with all other UHPLC conditions as described in section 2.3.

The retention times in **Figure 3** increased as the TFA concentration increased. Since, the increase in retention times was more significant in the front region of the chromatograms compared to the rear region the retention window became smaller, which is a disadvantage. Consequently, the sample peak capacity decreased somehow with increasing TFA concentration. Most interestingly, three peptide peak pairs (25/1 in red, 22/13 in grey, and 4/31 in blue) are highlighted, showing significant changes in selectivity, resulting in a reversal in elution order. Due to potentially lower ion suppression and the wider retention window, the decision was made to choose the lowest concentration of TFA, which is 0.05/0.04%.

3.1.5. Optimizing Temperature

A change in temperature influences chromatographic separations in different ways. As the temperature increases, the viscosity of the liquid mobile phase decreases, which leads to higher diffusion coefficients and thus to somewhat narrower peaks. This is expected to increase peak capacity, and hence it is suggested to choose the highest temperature, which is feasible in a specific system [19]. 60°C represented the maximum temperature specified by the UHPLC column manufacturer. As a compromise between separation power and column stability, we have finally chosen 55°C. **Figure 4** shows the transition from 35°C to 55°C. This change led at least to a small improvement of the peak width (FWHM) from 0.31 min to 0.28 min.

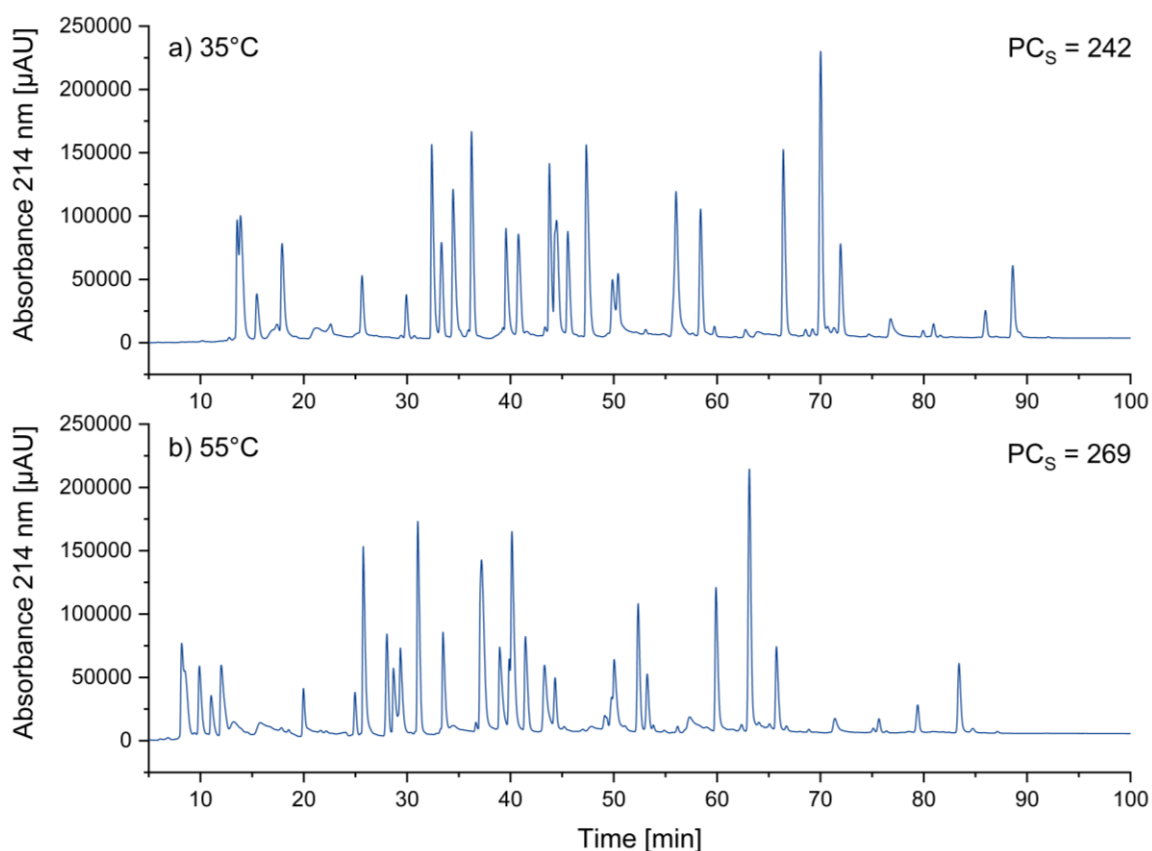


Figure 4. Influence of increasing temperature on the separation (sample peak capacity, PC_s) of the CEF advanced peptide pool containing 32 peptides. a) 35°C; b) 55°C. A small increase of the sample peak capacity was obtained. Additionally, retention times decreased. The peptide sequences are listed in **Table 2**. For UHPLC conditions, please refer to Section 2.3.

3.1.6. Final UHPLC Method

The optimization of gradient slope, concentration of ion-pairing reagent TFA, and temperature led to the UHPLC method outlined in **Table 1**:

Table 1. Optimized UHPLC method for the analysis of the CEF advanced peptide pool. For complete UHPLC conditions, please refer to Section 2.3.

UHPLC Parameter	Optimized Value
Gradient slope	0.4 %B/min
Initial %B	4
Final %B	44
Trifluoroacetic acid (TFA)	0.05/0.04% (v/v) in eluent A/B
Temperature	55°C
Flow rate	6 μ L/min

Altogether, the optimized parameters in **Table 1** led to the separation of 30 of the 32 peptides in the CEF advanced peptide pool. Two out of the 32 peptides exhibited poor peak shapes and were excluded from quantification. Both affected peptides possess a N-terminal cysteine. This might have led to undesired interactions between the N-terminus and active sites in the chromatographic system, resulting in significant peak broadening and potential analyte loss. Six out of the 32 peptides highly overlapped (peak to valley ratio of less than 2), which had to be considered in the integration step.

3.2. Relative UV Quantification

The primary use of the developed method was for relative quantification. This enables a cost-efficient evaluation of whether a sample has degraded during storage or transportation. Additionally, it can be utilized to compare samples from various batches or manufacturers for any discrepancies. Relative quantification does not necessarily require calibration, eliminating the need for highly characterized standard peptides and the labor-intensive determination of calibration functions. Only the peak areas of peptides from the compared samples are required. When comparing two sets of measurements, a relative deviation in percent from the reference mean is calculated for each peptide. This mean deviation \pm confidence interval can be compared to the method's precision (relative standard deviation, RSD) for the respective peptide. If the mean deviation of a peptide exceeds the method precision in a relevant way, it indicates a potential change in composition of the peptide pool. If the mean deviations \pm confidence interval are all less than or equal to the corresponding method precision, the peptide pool can be considered unchanged based on the result of this method.

However, it is important to note that a significant deviation does not necessarily mean that the product is out of specification. Independent criteria must be used to define what deviation from the previous composition can be tolerated from a practical point of view. This is especially true for peptide pools, as it has been shown that in respect of T cell activation, some peptides are present in a large excess, while others are present in a more limited concentration [13], which can make any losses more critical. This assessment can also depend heavily on the application, which can vary greatly from customer to customer.

To determine peak areas, a classical perpendicular drop method [22] was used as the default. However, due to the considerable overlap of some peptide peaks, some peak areas could not be reliably determined. When slight variations in analysis conditions cause minor changes in the separation of closely overlapping peaks, using perpendicular drop integration can lead to significant variations in peak areas and, consequently, inaccurate results. This problem is illustrated by peptide 19 in **Figure 5a**, showing two injections from different series (injection #10 on the left and injection #20 on the right) of peptide pool measurements. There was a 42-hour interruption between these series. As a result, the separation (peak-to-valley ratio) between peptide 19 and the subsequent peptide 18 was altered, perhaps due to a slight change of the mobile phase.

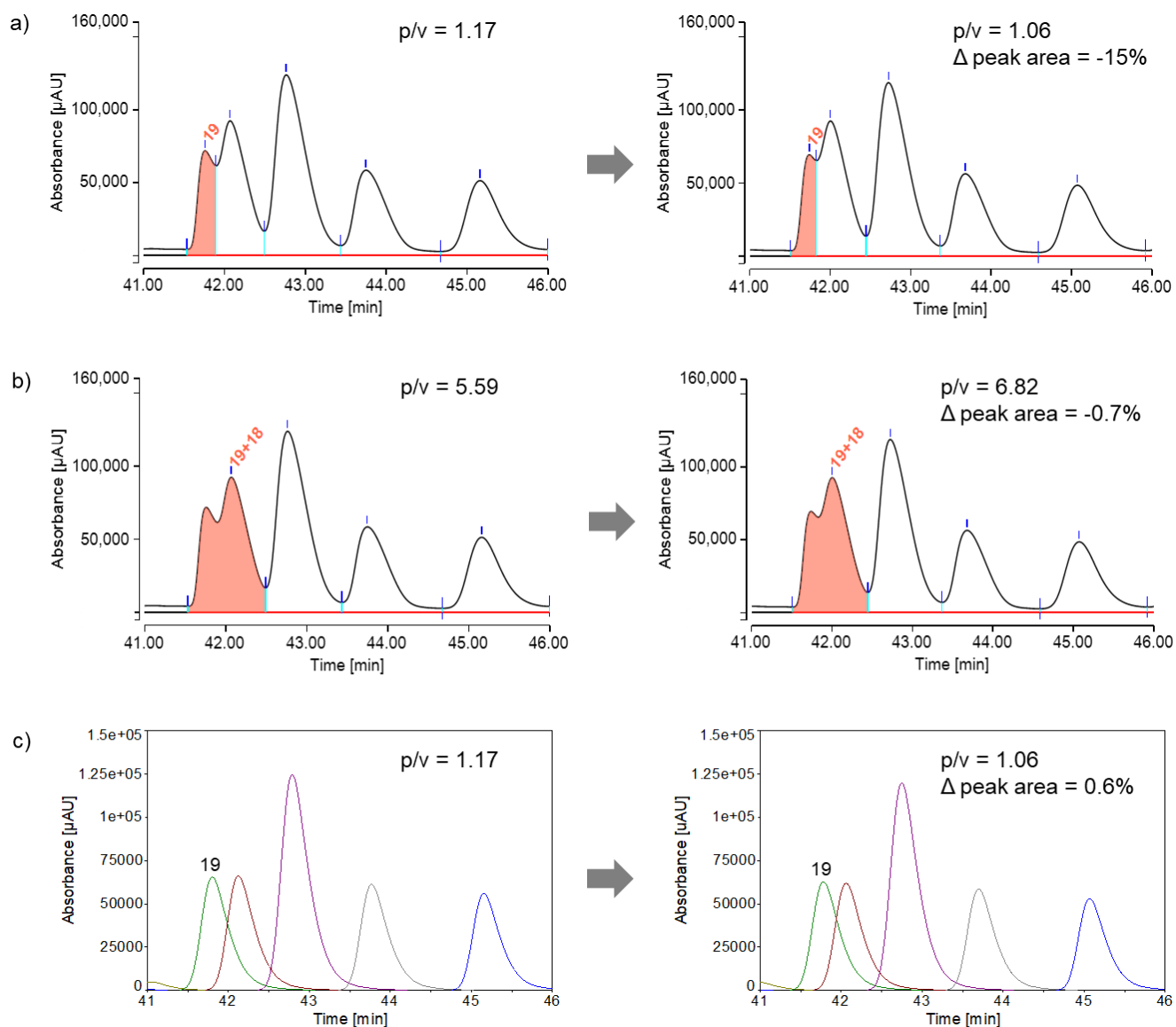


Figure 5. Variations in peak areas (Δ peak area) due to altered peak-to-valley ratio (Δ p/v) of strongly overlapping peaks and potential solutions. a) A change in the peak-to-valley ratio of highly overlapping peptide 19 between injection #10 (left side) and #20 (right side) led to an integration error of -15% when using the perpendicular drop method. b) By grouping the peaks of peptides 19 and 18, the integration error of the vertical drop method could be reduced from -15% to only -0.7%. c) Peak fitting enabled a precise determination of the peak area even for the highly overlapping peptide 19. The peak area deviation was 0.6%. For detailed instructions on peak fitting, refer to section 2.4. Peak-to-valley ratios were calculated for injection #10. The percentage values also refer to injection #10.

One potential solution was to group peaks as shown in **Figure 5b**. Grouping peaks lead to significantly reduced integration errors. When peptide 19 was evaluated as part of a peak group using the perpendicular drop method, the peak area deviation decreased from -15% to less than -1%. Another potential solution was peak fitting as depicted in **Figure 5c**, which was considered more robust against changes in the overlap of poorly resolved peaks. Even for the strongly overlapping peptide 19, the fitted peaks (green) showed only minimal differences in peak area. The reproducibility (relative standard deviation, RSD) of peak areas for all peptides in the CEF advanced peptide pools from seven consecutive measurements is shown in **Table 2**, comparing the perpendicular drop method (perp. drop) with the peak fitting method.

Table 2. Reproducibility (relative standard deviation, RSD) of peak areas for all peptides of the CEF advanced peptide pool using perpendicular drop method (perp. drop) or peak fitting. High relative standard deviation in peak area and low peak-to-valley ratios ($p/v < 2$) of some peptides indicate the need for peak groups when using the perpendicular drop method. Peak fitting results in low relative standard deviations in peak areas, even for highly overlapping peptides, eliminating the necessity for peak grouping.

Peptide name	Sequence	p/v ¹	Peak area perp. drop ² [$\mu\text{AU} \cdot \text{min}$]	RSD peak area perp. drop [%]	Peak area peak fitting ² [$\mu\text{AU} \cdot \text{min}$]	RSD peak area peak fitting [%]
8*	KTGGPIYKR	n.a.	/	/	1.412	0.6
14*	AVFDRKSDAK	n.a.	/	/	0.501	0.9
10	ILRGSVAHK	24.1	0.939	0.2	0.979	0.6
12	RLRAEAQVK	11.9	0.514	0.1	0.563	0.3
11	RVRAYTYSK	24.5	1.224	0.1	1.261	0.4
32	TPRVTGGGAM	27.6	0.685	0.2	0.737	0.2
28	YPLHEQHGM	60.2	1.358	0.8	1.527	0.2
16	ATIGTAMYK	31.7	0.960	0.8	1.028	0.2
25	SRYWAIRTR	51.9	1.922	0.1	2.144	0.2
1	VSDGGPNLY	22.0	1.000	0.2	1.000	0.5
21	RAKFKQLL	3.5	0.721	0.3	0.819	0.1
9	RVLSFIKGTK	4.1	0.897	0.2	0.898	0.2
29	IPSINVHHY	20.1	2.021	0.1	2.228	0.3
22	FLRGRAYGL	35.8	1.208	0.4	1.194	0.1
13*	SIIPSGPLK	1.5	1.181	0.4	1.464	0.2
23*	QAKWRLQTL	1.6	1.749	0.3	1.456	0.3
19*	RPPIFIRRL	1.2	0.691	2.0	1.168	0.3
18*	LPFDKTTVM	1.5	1.461	0.8	1.187	0.2
20	ELRSRYWAI	7.5	2.270	0.2	2.229	0.3
15	IVTDFSVIK	8.8	1.162	0.2	1.099	0.2
17	DYCNVLNKEF	13.0	1.102	0.2	0.995	0.3
26	ASCMGLIY	11.0	1.343	0.2	1.123	0.3
27	RRIYDLIEL	8.3	0.818	0.4	0.788	0.2
24	SDEEEAIVAYTL	19.4	1.374	0.3	1.342	0.1
7	NLVPMVATV	13.1	0.989	0.2	0.918	0.2
4	FMYSDFHFI	52.2	2.184	0.1	2.106	0.2
31	EFFWDANDIY	40.7	2.924	0.2	2.804	0.2
30	EENLLDFVRF	10.8	0.867	0.1	0.876	0.2
6	GLCTLVAML	17.4	0.552	0.4	0.410	0.2
3	GILGFVFTL	45.0	0.895	1.0	0.825	0.8
2**	CTELKLSDY	/	/	/	/	/
5**	CLGGLTMV	/	/	/	/	/
8+14	s.o.	32.2	1.891	0.2	1.912	0.4
13+23	s.o.	22.1	2.947	0.2	2.920	0.2
19+18	s.o.	5.6	2.152	0.1	2.355	0.2

¹ Peak-to-valley ratio, calculated for peaks resulting from injection #10 using perpendicular drop method. ² Peak areas were normalized using peptide 1 as an internal standard and averaged over seven injections (injections #4-10). *Additional evaluation as part of a peak group. **Peptide area was not analyzed due to poor peak shape. n.a. (not available): Peak-to-valley ratio (p/v) could not be determined.

When using the perpendicular drop method, it is evident that the peak area of peptide 19 fluctuates significantly more, with a relative standard deviation of 2.0%, compared to the other peaks. Therefore, peptide 19 was grouped with the highly overlapping peptide 18, significantly improving peak area reproducibility (peptides 19+18: 0.1%) compared to individual peaks. When using the peak fitting method, even this highly overlapping peptide exhibited low relative standard deviation in peak area (peptide 19: 0.3%), eliminating the need for a peak group. Peptides 13 and 23 also showed low peak-to-valley ratios ($p/v < 2$). Therefore, when using the perpendicular drop method, it is recommended to group these peaks, which is unnecessary for the peak fitting method. Peptides 8 and 14 were unresolved in some replicates using the perpendicular drop method, requiring them to be assessed as a peak group from the beginning. When applying peak fitting to peptides 8 and 14, two distinct peaks were identified, with a relative standard deviation in peak area below 1%, indicating acceptable results. Peptides 2 and 5 could neither be analyzed with the perpendicular drop method nor with the peak fitting method, as they exhibited poor peak shapes. As already mentioned, one possible reason for this could be undesirable interactions between the analyte and the non-inert surfaces of the chromatographic system. **Figure 6** provides an overview of the two integration options.

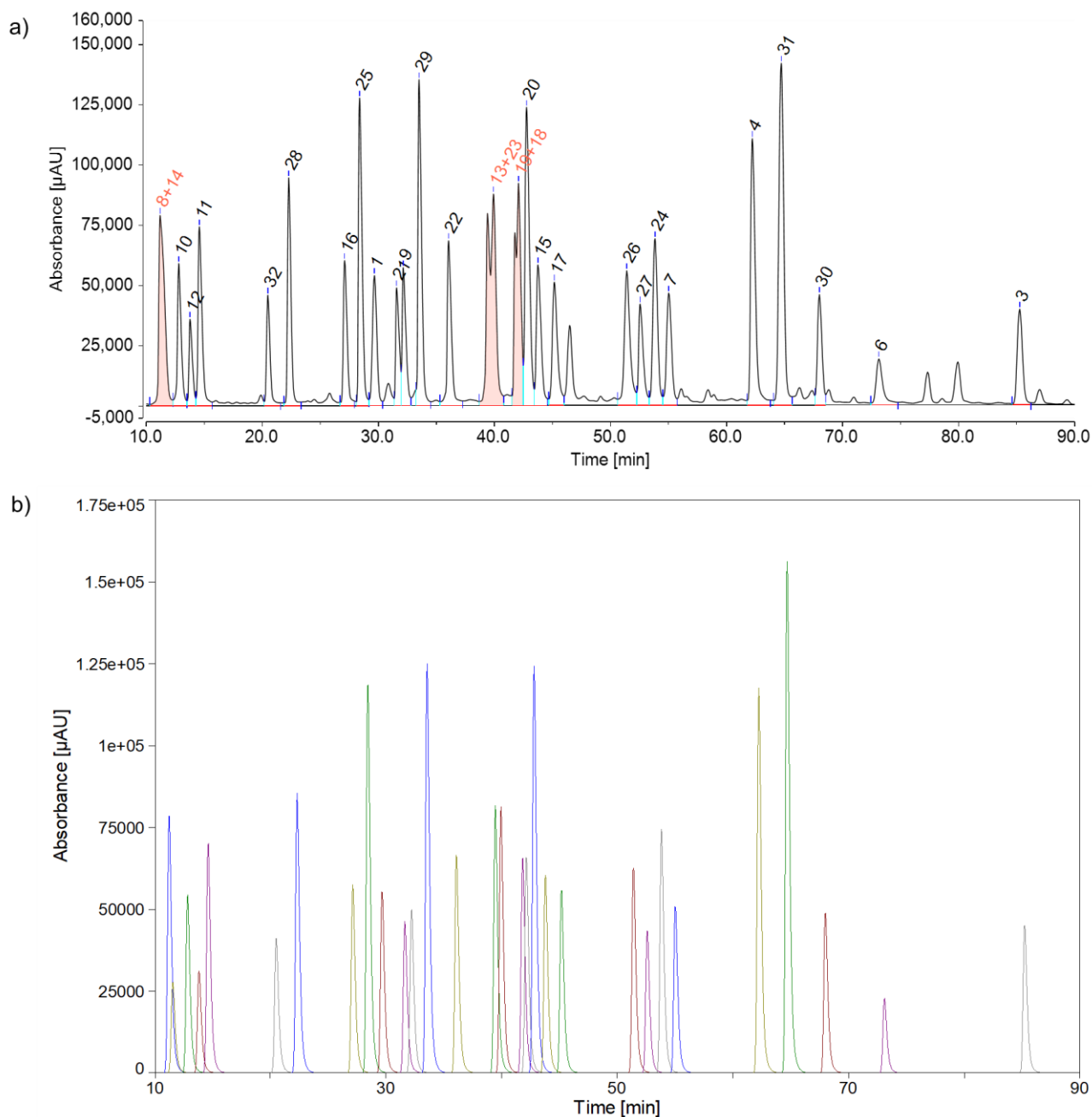


Figure 6. Examination of the CEF advanced peptide pool consisting of 32 peptides, listed in **Table 2**
a) Integration by a routine software based on a perpendicular drop algorithm. Some peaks (peptides 8+14, 13+23 and 19+18) have been combined due to a lack of separation. b) Integration by a peak fitting software. Due to the excellent reproducibility of the peak areas by peak fitting, it was not necessary to form peak groups.

When comparing two sets of measurements, it is crucial to consider that the resolution of individual peaks may vary significantly. This variation may result from differences in concentration or column separation efficiency. In this case, it might be necessary to modify the criteria for creating peak groups. One alternative approach could involve forming peak groups until a minimum resolution of 1.5 (almost baseline separation) is achieved (refer to **Table S1** and **Figure S1** in the Supplementary Materials). Although quality problems cannot be immediately traced back to a single peptide and some minor sensitivity issues may occur, defining peak groups seems to be a straightforward and efficient method for detecting degradation of poorly resolved peaks without excessive effort. It is also important to note that the used PeakFit software had some limitations when compared to commonly used chromatography data software (CDS). The raw data had to be initially reduced because it exceeded the maximum number of data points. Additionally, the software has an upper limit of 100 peaks to be fitted. Furthermore, the user has to select a peak fitting method, including the Residuals method, Second Derivative method, and Gaussian Deconvolution method. Also, a suitable peak function needs to be identified, which may be dependent on the sample and the experimental system. A good understanding of nonlinear curve fitting is necessary for these aspects. Hence, for routine analytical purposes, it is recommended to use a CDS, possibly in combination with peak grouping, due to its user-friendliness and simplicity, as long as automated peak fitting methods are not included in most standard chromatography software.

3.3. High Resolution Mass Spectrometry

High-resolution mass spectrometry is a powerful technique for accurately identifying and quantifying compounds in complex samples by measuring molecule masses (or m/z) precisely. The high mass accuracy enables reliable identification of compounds. Using the high-resolution Orbitrap mass spectrometer in this study, narrow mass extraction windows (MEW) of 5 ppm were applied to confirm the peptide structure, as shown for peptide 30 in **Figure 7**.

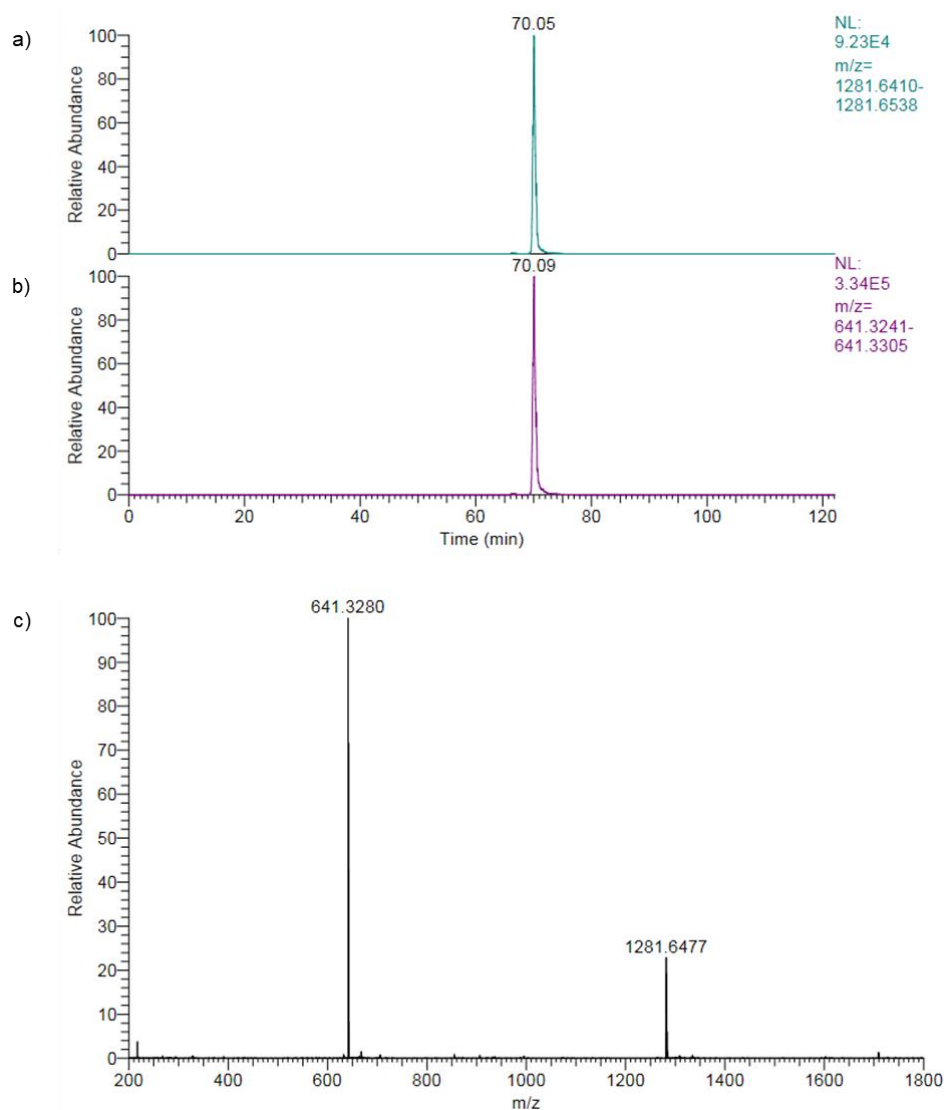


Figure 7. Peptide confirmation using high-resolution mass spectrometry (HRMS). As an example, extracted ion chromatograms (XIC) with a 5 ppm mass extraction window and a high-resolution mass spectrum of peptide 30 are shown. Calculated $[M+H]^+ = 1281.6474$ and $[M+2H]^{2+} = 641.3273$. a) XIC of $[M+H]^+$; b) XIC of $[M+2H]^{2+}$; c) Mass spectrum corresponding to the peak maximum in b).

Retention times (UV), the most intense ion, and the deviation between experimental and calculated m/z values are listed for each peptide in the CEF advanced peptide pool in **Table 3**.

Table 3. Exact mass measurements of the 32 peptides in the CEF advanced peptide pool using high-resolution mass spectrometry (HRMS). For most peptides, the doubly charged ion showed the highest intensity. Minimal deviations between theoretical and observed m/z values provide strong confidence in confirming the peptide structure.

Peptide	tr (UV) [min]	Ion	Calc. m/z [14]	Exp. m/z	$\Delta m/z$	Mass error [ppm]
8	11.18	$[M+2H]^{2+}$	510.3035	510.3034	0.0001	-0.1
14	11.18	$[M+2H]^{2+}$	568.8066	568.8066	0.0000	0.1
10	12.79	$[M+2H]^{2+}$	490.8036	490.8042	0.0006	1.2
12	13.77	$[M+2H]^{2+}$	535.8251	535.8254	0.0003	0.6
11	14.57	$[M+2H]^{2+}$	572,3171	572.3171	0.0000	0.0

32	20.47	[M+2H] ²⁺	473.7424	473.7426	0.0002	0.5
28	22.27	[M+2H] ²⁺	556.2531	556.2537	0.0006	1.1
16	27.09	[M+2H] ²⁺	478.2495	478.2496	0.0001	0.2
25	28.37	[M+2H] ²⁺	604.8360	604.8366	0.0006	1.0
1	29.65	[M+H] ⁺	921.4312	921.4310	-	-0.3
					0.0002	
21	31.57	[M+2H] ²⁺	502.3242	502.3245	0.0003	0.6
9	32.14	[M+2H] ²⁺	574.8611	574.8615	0.0004	0.6
29	33.49	[M+2H] ²⁺	540.2853	540.2858	0.0005	1.0
22	36.04	[M+2H] ²⁺	526.8036	526.8040	0.0004	0.7
13	39.40	[M+2H] ²⁺	456.2817	456.2823	0.0006	1.4
23	39.91	[M+2H] ²⁺	572.3353	573.3356	0.0003	0.5
19	41.76	[M+2H] ²⁺	584.3773	584.7387	0.0014	2.4
18	42.07	[M+2H] ²⁺	526.2783	526.2786	0.0003	0.6
20	42.76	[M+2H] ²⁺	597.3249	597.3254	0.0005	0.8
15	43.75	[M+2H] ²⁺	511.3001	511.3008	0.0007	1.5
17	45.16	[M+2H] ²⁺	622.7844	622.7853	0.0009	1.4
26	51.41	[M+H] ⁺	857.3896	857.3901	0.0005	0.6
27	52.55	[M+2H] ²⁺	595.8482	495.8488	0.0006	1.0
24	53.83	[M+H] ⁺	1339.6264	1339.6265	0.0001	0.1
7	55.02	[M+2H] ²⁺	943.5281	943.5295	0.0014	1.5
4	62.22	[M+2H] ²⁺	603.7681	603.7694	0.0013	2.2
31	64.71	[M+H] ⁺	1319.5579	1319.5580	0.0001	0.1
30	67.99	[M+2H] ²⁺	641.3273	641.3280	0.0007	1.0
6	73.13	[M+H] ⁺	920.4944	920.4958	0.0014	1.6
3	85.26	[M+H] ⁺	966.5659	966.5663	0.0004	0.4
2	*	[M+H] ⁺	1071.5027	1071.5027	0.0000	0.0
5	*	[M+H] ⁺	906.4787	906.4808	0.0021	2.3

* Retention time was not determined due to poor peak shape.

The identity of all 32 expected peptides could be confirmed with very low deviation between the observed m/z value (experimental m/z) and the theoretical m/z value (calculated m/z). In most cases, the observed m/z value matched the calculated m/z value up to the third decimal place, with a maximum ppm value of 2.4. Two out of the 32 peptides exhibited poor peak shapes, hindering the identification of their retention times. As suspected earlier, this might have been caused by unwanted interactions between the analyte and non-inert surfaces in the chromatographic system.

Confirming peptide structures in complex mixtures does not necessarily require a high-resolution mass spectrometer [23]. Common triple quadrupole mass spectrometers (QqQ-MS/MS) are known for their high selectivity using specific mass transitions (ion-fragment pairs). Additionally, the investment costs are usually lower than those of HRMS systems. However, HRMS provides some benefits, including high mass accuracy and the ability to provide isotopic patterns that assist in structural elucidation. Moreover, working in full scan mode (FS mode) is possible without a significant loss of sensitivity. Full scan mode enables non-targeted and retrospective data analysis. Hence, this feature allows users to examine degradation products at any time. Developing HRMS methods is quicker and simpler than the commonly used Multiple Reaction Monitoring mode (MRM mode) in tandem mass spectrometry. In MRM mode, detecting mass transitions involves identifying parent ions and fragment ions. In general, three mass transitions are identified, with the most intense transition serving as the quantifier and the other two as qualifiers. Collision energy is optimized for selected mass transitions. These steps must be carried out for each analyte, making MRM method development relatively time-consuming. Overall, both HRMS and MS/MS are valuable mass spectrometry tools, and the choice between the two techniques depends on the specific analytical requirements and experimental objectives.

4. Conclusions

In this study, we investigated the potential for a cost- and time-efficient UHPLC-HRMS quality control for synthetic peptide pools using the popular peptide pool CEF advanced. For the chromatographic separation of a complex peptide pool, a flat gradient, elevated temperature, and trifluoroacetic acid (TFA) at a concentration of 0.05/0.04% were beneficial. The optimized separation protocol enabled the relative quantification of 30 out of the 32 peptides in the peptide pool. Two peptides were identified based on their masses, but quantification was not performed due to poor peak shape. Since both peptides contain an N-terminal cysteine, modifying the respective peptide pool(s) should be considered to avoid potential interactions between free thiol groups and the HPLC system. In addition, using inert HPLC systems might eliminate this problem. Determining UV peak areas was challenging due to significant overlap among some peaks. When using simple perpendicular drop integration, relative quantification was possible by grouping peaks for insufficiently resolved peaks. An alternative approach to dealing with the observed resolution changes and avoiding the need for peak grouping would be to use more computationally demanding peak fitting methods.

Overall, this study shows that cost-efficient quality control of synthetic peptide pools is possible by relative quantification using UV detection and peptide confirmation by high-resolution mass spectrometry. Some limitations arise from the need to form peak groups as long as algorithms for peak fitting are not generally available in standard chromatography software. The method presented here represents a significant improvement for the quality control of peptide pools up to medium size. The applicability to much larger peptide pools has yet to be demonstrated. Further improvements to this method could include the use of the latest UHPLC systems with higher resolution and the use of longer columns and separation times to increase peak capacity. The evaluation of peak purity with a diode array detector or the parallel use of a second column with at least partially orthogonal separation properties could provide additional information. At present, the mass spectrometric evaluation was performed manually. Automated peak assignment and confirmation could be also added. It should be noted that high-resolution mass spectrometry, as used in this study, is not mandatory. In many cases, standard resolution MS, such as triple quadrupole mass spectrometry (LC-MS/MS systems), would be sufficient to confirm the presence of the respective peptide. However, HRMS can be very useful for non-target analysis to identify degradation products, synthesis by-products and other impurities. For new and larger peptide pools, the use of a chromatography optimization software and the introduction of artificial intelligence (AI) for the optimization of the separation and integration methods, as well as the examination of complex mass spectra would be attractive objectives for future developments.

Author Contributions: Conceptualization, G.B.B., and M.G.W.; methodology, G.B.B., and M.G.W.; investigation, G.B.B., and S.E.; data curation, G.B.B.; writing—original draft preparation, G.B.B.; writing—review and editing, G.B.B., O.J.K. and M.G.W.; visualization, G.B.B.; supervision, M.G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Bundesanstalt für Materialforschung und -prüfung (BAM), grant number Ideen_2016_15.

Data Availability Statement: Not applicable.

Conflicts of Interest: Dr. Oliver J. Kreuzer is founder and director of peptides & elephants GmbH and a commercial provider of peptide pools. The other authors declare no conflict of interest.

References

1. de Graaf, M.T.; de Beukelaar, J.W.; Burgers, P.C.; Luider, T.M.; Kraan, J.; Smitt, P.A.S.; Gratama, J.W. Contamination of Synthetic HuD Protein Spanning Peptide Pools with a CMV-Encoded Peptide. *Cytom Part A* **2008**, *73a*, 1079-1085, doi:10.1002/cyto.a.20636.
2. Weller, M.G. The Protocol Gap. *Method Protocol* **2021**, *4*, doi: 10.3390/mps4010012.
3. Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M. Ion suppression; A critical review on causes, evaluation, prevention and applications. *Talanta* **2013**, *115*, 104-122, doi:10.1016/j.talanta.2013.03.048.

4. Müller, C.; Schäfer, P.; Störtzel, M.; Vogt, S.; Weinmann, W. Ion suppression effects in liquid chromatography-electrospray-ionisation transport-region collision induced dissociation mass spectrometry with different serum extraction methods for systematic toxicological analysis with mass spectra libraries. *J Chromatogr B* **2002**, *773*, 47-52, Doi 10.1016/S1570-0232(02)00142-3.
5. Larger, P.J.; Breda, M.; Fraier, D.; Hughes, H.; James, C.A. Ion-suppression effects in liquid chromatography-tandem mass spectrometry due to a formulation agent, a case study in drug discovery bioanalysis. *J Pharmaceut Biomed* **2005**, *39*, 206-216, doi:10.1016/j.jpba.2005.03.009.
6. Greer, B.; Chevallier, O.; Quinn, B.; Botana, L.M.; Elliott, C.T. Redefining dilute and shoot: The evolution of the technique and its application in the analysis of foods and biological matrices by liquid chromatography mass spectrometry. *Trac-Trend Anal Chem* **2021**, *141*, doi: 10.1016/j.trac.2021.116284.
7. Gros, M.; Petrovic, M.; Barceló, D. Development of a multi-residue analytical methodology based on liquid chromatography-tandem mass spectrometry (LC-MS/MS) for screening and trace level determination of pharmaceuticals in surface and wastewaters. *Talanta* **2006**, *70*, 678-690, doi:10.1016/j.talanta.2006.05.024.
8. Ito, S.; Tsukada, K. Matrix effect and correction by standard addition in quantitative liquid chromatographic-mass spectrometric analysis of diarrhetic shellfish poisoning toxins. *J Chromatogr A* **2002**, *943*, 39-46, doi: 10.1016/S0021-9673(01)01429-7.
9. Zrostlíková, J.; Hajslová, J.; Poustka, J.; Begany, P. Alternative calibration approaches to compensate the effect of co-extracted matrix components in liquid chromatography-electrospray ionisation tandem mass spectrometry analysis of pesticide residues in plant materials. *J Chromatogr A* **2002**, *973*, 13-26, doi: 10.1016/S0021-9673(02)01196-2.
10. Berg, T.; Strand, D.H. ¹³C labelled internal standards - A solution to minimize ion suppression effects in liquid chromatography-tandem mass spectrometry analyses of drugs in biological samples? *J Chromatogr A* **2011**, *1218*, 9366-9374, doi:10.1016/j.chroma.2011.10.081.
11. Meyer, B.; Papisotiriou, D.G.; Karas, M. 100% protein sequence coverage: a modern form of surrealism in proteomics. *Amino Acids* **2011**, *41*, 291-310, doi:10.1007/s00726-010-0680-6.
12. Moldovan, I.; Targoni, O.; Zhang, W.J.; Sundararaman, S.; Lehmann, P.V. How frequently are predicted peptides actually recognized by CD8 cells? *Cancer Immunol Immun* **2016**, *65*, 847-855, doi:10.1007/s00262-016-1840-7.
13. Zhang, W.J.; Moldovan, I.; Targoni, O.S.; Subbramanian, R.A.; Lehmann, P.V. How much of Virus-Specific CD8 T Cell Reactivity is Detected with a Peptide Pool when Compared to Individual Peptides? *Viruses-Basel* **2012**, *4*, 2636-2649, doi:10.3390/v4112636.
14. Patiny, L.; Borel, A. ChemCalc: A Building Block for Tomorrow's Chemical Infrastructure. *J Chem Inf Model* **2013**, *53*, 1223-1228, doi:10.1021/ci300563h.
15. Horvath, C.G.; Lipsky, S.R. Peak Capacity in Chromatography. *Anal Chem* **1967**, *39*, 1893, doi:DOI 10.1021/ac50157a075.
16. Giddings, J.C. Maximum Number of Components Resolvable by Gel Filtration and Other Elution Chromatographic Methods. *Anal Chem* **1967**, *39*, 1027-1028, doi: 10.1021/ac60252a025.
17. Neue, U.D. Theory of peak capacity in gradient elution. *J Chromatogr A* **2005**, *1079*, 153-161, doi:10.1016/j.chroma.2005.03.008.
18. Davis, J.M.; Giddings, J.C. Statistical-Theory of Component Overlap in Multicomponent Chromatograms. *Anal Chem* **1983**, *55*, 418-424, doi: 10.1021/ac00254a003.

19. Wang, X.L.; Stoll, D.R.; Schellinger, A.P.; Carr, P.W. Peak capacity optimization of peptide separations in reversed-phase gradient elution chromatography: Fixed column format. *Anal Chem* **2006**, *78*, 3406-3416, doi:10.1021/ac0600149.
20. Dolan, J.W.; Snyder, L.R.; Djordjevic, N.M.; Hill, D.W.; Waeghe, T.J. Reversed-phase liquid chromatographic separation of complex samples by optimizing temperature and gradient time I. Peak capacity limitations. *J Chromatogr A* **1999**, *857*, 1-20, doi: 10.1016/S0021-9673(99)00765-7.
21. Chen, Y.; Mehok, A.R.; Mant, C.T.; Hodges, R.S. Optimum concentration of trifluoroacetic acid for reversed-phase liquid chromatography of peptides revisited. *J Chromatogr A* **2004**, *1043*, 9-18, doi:10.1016/j.chroma.2004.03.070.
22. Novak, J.; Petrovic, K.; Wicar, S. Efficacy of Corrections Applied in Resolution of Overlapping Chromatographic Peaks by Perpendicular Drop Method. *J Chromatogr* **1971**, *55*, 221-229, doi:10.1016/S0021-9673(01)82546-2.
23. Abbas, I.M.; Hoffmann, H.; Montes-Bayón, M.; Weller, M.G. Improved LC-MS/MS method for the quantification of hepcidin-25 in clinical samples. *Anal Bioanal Chem* **2018**, *410*, 3835-3846, doi:10.1007/s00216-018-1056-0.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.