

Contents lists available at ScienceDirect Journal of Mass Spectrometry and Advances in the Clinical Lab

journal homepage: www.sciencedirect.com/journal/journal-of-massspectrometry-and-advances-in-the-clinical-lab

Research Article

Cheaper, faster, simpler trypsin digestion for high-throughput targeted protein quantification



Christopher M. Shuford^{*}, Russell P. Grant

Labcorp, Center for Esoteric Testing, Burlington, NC, USA

ARTICLE INFO	A B S T R A C T
ARTICLEINFO Keywords: Trypsin Mass spectrometry Liquid chromatography Bottom-up proteomics Protein quantification	Introduction: LC-MS-based methods for protein quantification have a stigma of being relatively expensive and low-throughput. This is partly due to the cost and speed of trypsin digestion, which has primarily focused on advancements in research-based biomarker discovery applications that rely on protein/peptide identifications rather than clinical biomarker quantification. However, there is a need for simple, fast, and reproducibly efficient surrogate peptide recovery in clinical biomarker quantification. <i>Methods:</i> Multiple methodologies were evaluated to enhance tryptic digestion for the analysis of thyroglobulin, a prototypical serum protein biomarker. The main criteria for assessment were the yield and speed of formation of surrogate peptides. Various factors such as different additives, types of trypsin, microwave- and pressure-assisted systems, and enzyme concentration were considered as key variables, in addition to digestion time. <i>Results:</i> It was observed that digestion additives/denaturants had a significant impact on the speed and yield of digestion for each surrogate peptide. Increasing the concentration of trypsin alone was found to accelerate digestions appreciably for most surrogate peptides, without affecting the yield. However, the use of sequencing-grade trypsins and microwave/pressure-assisted systems did not offer significant advantages over the use of 'standard-grade' TPCK-treated trypsin in combination with a conventional incubator, once digestion time and additive had been optimized. <i>Conclusion:</i> We have dispelled the notion that trypsin digestion is inherently slow and expensive for targeted quantification of serum proteins. Additionally, we have established a groundwork for experimentation that can pave the way for the creation of efficient trypsin digestion protocols, aiming to optimize yield, speed, and cost. It is our hope that these advancements will promote the wider incorporation of such assays in clinical laboratories.

1. Introduction

Efficiency and, more importantly, specificity are the two aspects of trypsin digestion that have the most impact on protein and peptide identification in shotgun proteomics experiments. It is commonly assumed that trypsin has high specificity in these experiments [1] and it is a crucial factor in database search algorithms [2,3]. Consequently, optimizing trypsin digestion specificity for shotgun proteomics involves using small amounts of proprietary 'sequencing grade' trypsin with prolonged digestion times to minimize autolysis and non-specific proteolysis, which can confound protein and peptide identification [4–7]. The prevalence of such protocols has understandably given the impression that trypsin digestion is a slow and costly process, especially

as these slow and relatively expensive protocols have been adopted for targeted protein quantification in clinical testing [8,9]. While overnight digestion may be suitable for research environments, it is not ideal in clinical laboratories due to the resulting delay in providing test results.

However, costly, complex, and lengthy digestion protocols have likely been perpetuated for targeted protein quantification simply out of precedent rather than necessity. For targeted applications where the surrogate peptide sequence is already known, the specificity of trypsin digestion becomes less critical, and digestion efficiency (speed and yield) becomes paramount. Few studies have shown a tangible benefit of using proprietary enzymes or additives for targeted protein quantification compared to conventional optimization. In contrast, some researchers have recently shown that digestion can be simplified

* Corresponding author.

https://doi.org/10.1016/j.jmsacl.2023.11.002

Received 20 January 2023; Received in revised form 27 October 2023; Accepted 21 November 2023 Available online 23 November 2023

Abbreviations: 1-PrpOH, 1-propanol; ACN, acetonitrile; AAA, amino acid analysis; DOC, deoxycholate; DTT, dithiothreitol; GnHCl, guanidine; MeOH, methanol; NAT, natural surrogate peptide; SIL, stabile isotope-labeled peptide; TFE, trifluoroethanol; TPCK, tosyl phenylalanyl chloromethyl ketone.

E-mail address: shuforc@labcorp.com (C.M. Shuford).

²⁶⁶⁷⁻¹⁴⁵X/© 2023 THE AUTHORS. Publishing services by ELSEVIER B.V. on behalf of MSACL. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

dramatically for targeted applications without novel enzyme formulations or specialized equipment. Zheng and DeMarco demonstrated that heat denaturation alone of diluted serum, followed by a 20-minute digestion with 'standard' TPCK-treated trypsin, was sufficient for analysis of apolipoprotein A-1, retinol-binding protein 4, transthyretin, complement component 9 and C-reactive protein [10,11]. Similarly, Hoofnagle and colleagues showed that diluted serum heat denatured with trifluoroethanol (TFE) was suitable for analysis of vitamin Dbinding protein [12] and retinol binding protein [13] following digestion with 'standard' TPCK-treated trypsin for 90 min or less.

However, the simplified digestion protocols described above have only been applied to small volumes of highly diluted serum/plasma (≤10 µL) due to the high protein content of serum. Significant predilution of serum limits the practical specimen volume that can be analyzed in a high-throughput (i.e., 96-well plate) format, as well as the sensitivity of the method, which is often directly proportional to the input specimen volume. In this study, we systematically evaluated which additives, type of trypsin, and mode of incubation would provide the most practical and cost-effective means to efficiently digest large volumes of serum (>200 µL) with minimal dilution using thyroglobulin as a prototypical serum protein biomarker [14–17]. To facilitate these studies, thyroglobulin was spiked into serum at supraphysiological levels to allow for analysis without further enrichment following digestion. In all cases, digestion efficiency was measured not only in terms of the amount of surrogate peptide yielded, but also the speed of digestion (time required to achieve maximum yield) to enable a highthroughput digestion protocol.

2. Materials & methods

2.1. Materials

Unless otherwise specified, all reagents, including thyroglobulin, were purchased from MilliporeSigma (St. Louis, MO). Formic acid, methanol, and acetonitrile were obtained from Fisher Scientific (Hampton, NH), and the water used was Type-I (Millipore Milli-Q, >18 $\mbox{M}\Omega\mbox{)}.$ The sequencing grade trypsins used were Sequencing-grade Modified Trypsin, Mass Spec-Grade Trypsin/Lys-C Mix, and Mass Spectrometry-Grade Trypsin Gold from Promega Corporation (Madison, WI), as well as Dimethylated Proteomics Grade Porcine Trypsin from MilliporeSigma. Stable isotope-labeled (SIL) peptides, containing [¹³C₆,¹⁵N]-leucine (L[^]), [¹³C₆,¹⁵N]-isoleucine (I[^]), [¹³C₆,¹⁵N₄]-arginine (R[^]), and/or [¹³C₆,¹⁵N₂]-lysine (K[^]) were purchased from Biosynth (Gardner, MA) with > 98 % peptide purity (Table 1). The concentrations of stock solutions for each SIL peptide and thyroglobulin were confirmed by triplicate amino acid analysis (AAA) performed at Biosynth. The purity of the thyroglobulin standard, as verified by the vendor (MilliporeSigma), was > 99 % by SDS-PAGE.

2.2. Human serum samples

Remnant, de-identified serum samples submitted to Labcorp for

Table	1
Curron	oto

testing were used without informed consent, as determined suitable by an external Institutional Review Board (WIRB®). All serum samples had undetectable levels of thyroglobulin (<0.1 ng/mL) and thyroglobulin autoantibodies (<1 U/mL), as measured by the Access Thyroglobulin and Access Thyroglobulin Antibody II chemiluminescent immunoassays (Beckman Coulter, Brea, CA).

2.3. Sample digestion

Endogenous levels of thyroglobulin (low ng/mL) require significant, targeted enrichment of surrogate peptides to enable detection by LC-MS/MS. Therefore, all studies conducted here used pooled human serum spiked to 480 nmol/L thyroglobulin (~160 μ g/mL – approximately 10,000-fold higher than physiologically normal levels) to allow for the detection of the thyroglobulin surrogate peptides and optimization of digestion conditions without further enrichment. The serum was diluted 4-fold at the time of digestion, resulting in a theoretical maximum concentration of 160 nmol/L for surrogate peptides in the digested sample.

Trypsin digestions were performed in a 96-well plate by combining 200 μ L of sera (containing 96 pmol thyroglobulin) with 200 μ L of the indicated denaturing solution containing 25 mM DTT. The mixture was then incubated at 56 °C for 30 min. Subsequently, the denatured samples were combined 1:1 (v/v) with a trypsin solution freshly prepared in 100 mM Tris-HCl, 20 mM CaCl2, pH 8.75, and the reagent/trypsin concentrations mentioned throughout the study refer to their respective concentrations during the digestion step. Unless stated otherwise, digestions were carried out at 37 °C for the specified time period. After digestion, the reaction was quenched by adding formic acid to a final concentration of 1 %, followed by the addition of 7.5 pmol of each SIL peptide.

All digestions and time points were performed in triplicate, and standard incubations were performed with a Thermomixer® C (Eppendorf AG, Hauppauge, NY) at 2000 rpm, unless otherwise specified. Microwave-assisted digestions were performed using the Rapid Enzyme Digestion System (REDS; Hudson Surface Technology Inc.) or MARS6 system (CEM Corporation). Digestions in the REDS were carried out at 37 °C and 400 W, using a pre-heated water bath. Digestions performed in the MARS6 system involved an initial temperature ramp to 35 °C over 20 min with a maximum power of 800 W, followed by incubation at 35 °C and 400 W. Pressure-assisted digestions were performed in the Barocycler® NEP2320 (Pressure Biosciences Inc.) at 56 °C, 20 kPsi, and 120 s cycle times (10 s on / 110 s off) in PCT MicroTubes (Pressure Biosciences Inc., Cat. No. MC150-96).

2.4. LC-MS/MS analysis

Each digestion was analyzed by liquid chromatography-selected reaction monitoring (LC-SRM) using an Aria® Transcend® HPLC system (Thermo Scientific, San Jose, CA) coupled to a SCIEX Triple QuadTM 5500 mass spectrometer (Sciex, Redwood City, CA). In each run, 10 μ L of sample was injected onto a 4.6 \times 100 mm Zorbax® Eclipse XDB-C18

urrogate Peptides.							
Abbreviation	Surrogate Peptide*	a.a.	SIL Internal Standard				
LED	VTWKSR LEDIPVASLPDLHDIER ALVGKD	1372–1388	LEDIPVASLPDL^HDIER^				
TFP	LHLDSK TFPAETIR FLQGDH	1416–1423	TFPAETI^R^				
VIF	KVPESK VIFDANAPVAVR SKVPDS	1579–1590	VI^FDANAPVAVR^				
VIL	ALFRKK VILEDK VKNFYT	1955–1960	VIL^EDK^				
SQA	GRLLGR SQAIQVGTSWKQVDQFL	2213-2223	SQAI^QVGTSWK^				
GGA	SLAADRGGADVASIHLLTARATNSQL	2387-2400	GGADVASIHL [^] LTAR [^]				
EFS	GGENYK EFSELLPNR QGLKKA	2699–2707	EFSEL^LPNR^				
FSP	FYQRRR FSPDDSAGASALLR SGPYMP	1010–1023	FSPDDSAGASAL^LR^				

*Bold amino acid (a.a.) residues indicate the surrogate peptide sequence being measured while the six amino acids flanking the trypsin cleavage sites are shown only for context.

column (Agilent Technologies, Santa Clara, CA) with an identical 4.6 \times 5 mm guard column. Mobile phases A and B consisted of water and acetonitrile, respectively, each with 0.1 % formic acid. Samples were loaded for 30 s at 0 % B and 0.5 mL/min, after which the peptides were eluted by ramping to 40 % B over 5 min at the same flow rate. The column was washed with 95 % B for 3 min at 0.75 mL/min and then reequilibrated with 0 % B for 2 min at 0.5 mL/min prior to the next injection.

Electrospray ionization was conducted using the Turbo VTM source set to 500 °C source temperature, 5500 V ESI voltage, and 70 units for both gas 1 and gas 2. Scheduled SRM was performed using an acquisition window of 45 s, cycle time of 500 ms, CAD setting of 7 units, and unit resolution settings for both Q1 and Q3. Three transitions were monitored for both the unlabeled/natural (NAT) and SIL peptides (Supplemental Material, **Table S1**). All peak detection and integration were performed by uploading data into Skyline [18], and the resulting peak area information was exported into Excel® 2010 (Microsoft, Redmond, WA) for further analysis.

2.5. Data analysis

Given that the SIL peptides were added after digestion, the concentration of each NAT surrogate peptide yielded through digestion of the intact thyroglobulin protein can be calculated using the internal reference method [19]. This calculation can be done using the following equation:

$$\frac{[NAT]}{[SIL]} = \frac{A_{NAT}}{A_{SIL}} \text{ or } [NAT] = \frac{A_{NAT}}{A_{SIL}} \times [SIL]$$

 A_{NAT} and A_{SIL} are the summed peak areas measured for matched SRM transitions, respectively (Supplemental Material, Table S1). [NAT] is the unknown amount of the unlabeled surrogate peptide being determined, while [SIL] is the amount of the SIL peptide added post-digestion, which is derived from the AAA-assigned concentration of the top-stock.

During time-course studies, the rate of peptide formation/production and degradation was modeled using pseudo first order kinetics, as previously described [20]. This model was then used to calculate the expected digestion time and maximum amount of surrogate peptide yielded in an unbiased manner, rather than qualitatively assessing when the digestion was "complete". In cases where a plateau or maximum in peptide formation was not observed, the digestion time was reported as "> N hrs", where N represents the maximum time point tested.

It is important to note that all studies herein considered peptide concentration instead of just peptide response (i.e., signal). This was done to ensure that ionization suppression/enhancement did not affect the ability to discern digestion recovery. Additionally, internal calibration was performed using SIL peptides, so the reported peptide concentration in these studies is not expected to reflect the thyroglobulin protein concentration, unless 100 % yield of the surrogate peptide was achieved through digestion. Therefore, all results for peptide concentrations are presented as the final concentration in the digested serum sample, rather than the unprocessed serum sample. This is done to emphasize that the primary goal of the study was to improve digestion yield (and speed) in order to enable high sensitivity LC-MS/MS measurements downstream, rather than improving the accuracy of digestion-based protein measurements without the use of a protein calibrant.

3. Results

Four different digestion time course experiments were conducted using a standard digestion procedure. The four variables assessed in each experiment were: 1) digestion additive, 2) trypsin type, 3) incubator/reactor, and 4) trypsin concentration. Detailed characteristics of these variables are described below as well as summarized in the Supplemental Material (Table S2).

3.1. Additive screening

Nine different additives or denaturants were evaluated to determine their effectiveness in improving the speed and/or yield of tryptic digestion. This evaluation was conducted using a common serum pool that was spiked with exogenous thyroglobulin. Each additive was tested at two different concentrations to assess if the results were concentration dependent. Additionally, three time points (1, 4, and 20 hrs) were evaluated to (coarsely) examine the time required to achieve maximum peptide yield and to detect any potential peptide degradation or loss. The data obtained from this evaluation is presented for the VIL peptide in Fig. 1, as well as for other surrogate peptides in the Supplemental Material (Figure S1–S7).

The highest amount of VIL peptide yielded across all conditions was 57.8 nmol/L, which was achieved after a 20-hour digestion using 5 % TFE. In comparison, near-maximum yield (over 85 % of the maximum) was obtained using various other additives, although this level of recovery was not achieved under all conditions. Only the concentration of 1 M urea, not 0.25 M urea, resulted in near-maximum recovery after 4- and 20-hour digestion. Both concentrations of DOC resulted in near-maximum recovery, but the higher concentration required 20 h of digestion, while the lower concentration only needed 4 h. In two instances, namely 0.5 % CHAPS and 1.25 % TFE, near-maximum yield was achieved with a 4-hour digestion, but not with the longer 20-hour digestion due to apparent peptide degradation.

While many interesting observations could be made from a detailed analysis of the data mentioned above, the main objective of this study was to identify a single denaturing condition that would achieve nearmaximum yield for all 8 surrogate peptides in the shortest time possible. To achieve this, heat maps were created, which displayed the highest peptide concentration obtained under each digestion condition along with the minimum time required to reach that concentration (Fig. 2). While none of the conditions resulted in complete digestion or a 100 % yield of any surrogate peptide (i.e., 'complete digestion'), nearmaximum yield was achieved on average using 8 of the 18 conditions tested. These conditions included 0.25 M Urea, 1 M Urea, 1 M GnHCl, 0.25 M Thiourea, 0.2 % DOC, 0.5 % CHAPS, 1.25 % TFE, and 5 % TFE. Among these eight, five conditions had an average digestion time of less than 5 h for the evaluated surrogate peptides: 0.25 M Urea, 1 M Urea, 0.2 % DOC, 0.5 % CHAPS, and 1.25 % TFE. Only the LED peptide required more than four hours to achieve near-maximum vield under these conditions, indicating that its primary sequence posed a significant challenge to efficient digestion rather than the higher-order protein structure. Notably, among all eight surrogate peptides, only the LED peptide contained acidic residues known to inhibit trypsin digestion [21] near both its N- and C-terminal cleavage sites (see Table 1). While all five of these conditions are viable options to facilitate rapid and efficient digestion, it was ultimately decided to use the condition containing 0.2 % DOC due to its lack of evidence for peptide degradation during prolonged digestion times and its practical advantage of being removable through acid precipitation for subsequent sample processing, if needed [22].

3.2. Trypsin types

When screening the various additives, it is noteworthy that no denaturing condition or digestion time tested provided 100 % recovery (i.e., 'complete digestion') of any surrogate peptide. These studies were conducted with TPCK-treated bovine trypsin. Therefore, multiple 'sequencing' grade trypsins as well as a trypsin/LysC mixture were evaluated for improvements in digestion yield. Additionally, each enzyme was evaluated as a digestion time course to determine if there was any observable difference in the speed of digestion. Exemplar time course results for 3 of the 8 surrogate peptides are shown in Fig. 3A. For



Fig. 1. The concentration of VIL surrogate peptide was measured during the screening of 9 additives at 2 different concentrations and 3 digestion time points each. The 9 additives tested were urea, guanidine (GnHCl), thiourea, deoxycholate (DOC), CHAPS, trifluoroethanol (TFE), acetonitrile (ACN), methanol (MeOH), and 1-propanol (1-PrpOH). All digestions were performed using TPCK-treated bovine trypsin at an enzyme to protein ratio of 1:200 (w/w). The dotted red line is plotted at 49.1 nmol/L, which represents the cut-off for 'near-maximum' yield. This value was defined as 85 % of the maximum yield observed among all conditions for this surrogate peptide. The maximum yield achieved was 57.8 nmol/L, obtained after a digestion time of 20 h with 5 % TFE. It's important to note that this cut-off is not based on the theoretical maximum yield of 160 nmol/L, which is calculated based on the starting amount of protein. Results exceeding the near-maximum cut-off are colored green, indicating success. Results colored orange represent the maximum observed concentration for a specific additive without exceeding the near-maximum cut-off. All results presented are averages from triplicate digestion experiments, and error bars are shown as +/- 1 standard deviation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2 of the 3 surrogate peptides, it is apparent that the 'sequencing grade' enzymes provided no improvement in the yield or speed of digestion compared to TPCK-treated bovine trypsin. However, for the EFS peptide, there were obvious differences in the speed of digestion between the different enzyme formulations. While all enzymes provided approximately the same yield of the EFS peptide after 18 h, the trypsin/LysC mixture clearly achieved this yield much faster than the other enzyme formulations.

However, in general, the 'sequencing grade' enzymes did not provide faster digestion than TPCK-treated bovine trypsin or yield more surrogate peptide. By modeling each digestion time course as previously described [20], it was possible to estimate the maximum amount of peptide yielded for each time course and the required digestion time (Fig. 3B). Among all the enzymes tested, TPCK-treated bovine trypsin yielded either the most surrogate peptide or at least 90 % of the maximum amount of peptide yielded across all enzyme studies for all 8 signature peptides. Furthermore, only proteomics-grade dimethylated porcine trypsin decreased the digestion time by more than 30 min for any surrogate peptide relative to TPCK-treated bovine trypsin (peptides: TFP, VIF, and SQA). The one exception to this was the EFS peptide example using the trypsin/LysC mixture (Trypsin 5 in Fig. 3), which decreased the digestion time by more than 8 h relative to TPCK-treated bovine trypsin without sacrificing yield. The preference of the EFS peptide for the Trypsin/LysC mixture was likely due to the presence of an acidic residue in the P1' position of the N-terminal lysine (see **Table 1**). Both trypsin and LysC are inhibited by acidic residues near the scissile bond; however, the influence has been shown to be less pronounced for LysC [23].

3.3. Enzyme reactors

Two commercially available microwave reactors and one commercially available pressure reactor were assessed. This was done through a digestion time course to determine if there were any improvements in

Relative Yield	Additive	LED	TFP	VIF	VIL	SQA	GGA	EFS	FSP	Average Relative Yield	Average Digestior Time (hrs
100%	0.25 M Urea	88%	93%	88%	77%	83%	99%	89%	81%	87%	4.1
95%	1.0 M Urea	95%	100%	90%	91%	83%	87%	92%	83%	90%	3.8
90%	0.25 M GnHCl	65%	96%	89%	92%	82%	58%	92%	67%	80%	10.9
85%	1.0 M GnHCl	37%	78%	87%	92%	100%	100%	94%	88%	85%	13.3
80%	0.0675 M Thiurea	75%	90%	88%	74%	80%	90%	86%	84%	83%	4.5
70%	0.25 M Thiourea	95%	98%	86%	96%	84%	86%	97%	83%	91%	6.9
60%	0.2% DOC	86%	99%	96%	90%	85%	84%	100%	85%	91%	4.1
50%	0.8% DOC	24%	92%	98%	85%	73%	66%	92%	100%	79%	13.3
40%	0.125% CHAPS	45%	83%	87%	46%	78%	77%	62%	59%	67%	5.3
30%	0.5% CHAPS	100%	91%	82%	87%	78%	74%	85%	80%	85%	4.1
20%	1.25% TFE	100%	100%	89%	96%	78%	88%	96%	95%	93%	4.5
10%	5.0% TFE	82%	91%	87%	100%	74%	86%	92%	99%	89%	9.3
0%	10% ACN	48%	92%	100%	37%	89%	89%	97%	61%	77%	6.9
	37.5% ACN	4%	16%	71%	7%	40%	23%	8%	16%	23%	18.0
Digestion	10% MeOH	3%	9%	57%	2%	31%	8%	5%	11%	16%	9.6
Time (hrs)	37.5% MeOH	23%	77%	86%	70%	51%	47%	67%	46%	59%	16.0
1	10% 1-PrpOH	30%	25%	83%	77%	70%	68%	63%	66%	60%	2.9
4	37.5% 1-PrpOH	1%	1%	61%	4%	24%	10%	3%	16%	15%	2.1
20											
	Max Yield (nmol/L)	69.8	83.4	93.4	57.8	113.4	63.7	66.1	58.2		
	Max Recovery	44%	53%	59%	37%	72%	41%	42%	37%		

Fig. 2. The 'relative yield' for each surrogate peptide is shown for all tested additives. This was determined by normalizing the highest peptide concentration measured at the three time points for each condition to the maximum concentration yielded across all conditions for that peptide. Results in bold indicate near-maximum yield, defined as > 85 % of the maximum yield for a given peptide. It is important not to confuse this with percent recovery, which was calculated based on a theoretical maximum peptide concentration (160 nmol/L) derived from the initial amount of protein. The maximum percent recovery achieved for each peptide is summarized at the bottom. The speed of digestion for a specific peptide is indicated by superimposed bars. This represents either the minimum time at which near-maximum yield was observed for a given additive concentration or, in cases where near-maximum yield was not attained, the time at which the highest peptide concentration was obtained for the given additive concentration.



Fig. 3. Digestion was conducted using five different trypsins and at eight different time points ranging from 0.5 to 18 h. The same procedure was followed for all digestions, which consisted of 0.2% DOC, an enzyme-to-protein ratio of 1:200 (w/w), and incubation on a Thermomixer. In Figure (A) the concentrations of the VIL, GGA, and EFS surrogate peptides are plotted for each trypsin and time point tested. Each data point represents the mean result of triplicate digestions, with error bars indicating +/- 1SD. The solid line overlaying the data represents the pseudo-first order model for each digestion. In Figure (B), the maximum concentration obtained for each surrogate peptide is derived from the model, along with the digestion time required to achieve maximum yield. The trypsins used in the study were as follows: Trypsin 1 - TPCK-treated bovine trypsin (MilliporeSigma); Trypsin 2 - proteomics-grade dimethylated porcine trypsin (MilliporeSigma); Trypsin 3 - sequencing-grade modified porcine trypsin (Promega); Trypsin 4 - mass spectrometry-grade trypsin gold (Promega); Trypsin 5 - mass spectrometry-grade trypsin/LysC mix (Promega).

the speed of digestion or amount of surrogate peptide produced compared to standard incubation on a Thermomixer. Examination of the digestion time course results for three representative surrogate peptides clearly demonstrated that pressure-assisted digestion was considerably faster than conventional incubation (Fig. 4A). On the other hand, the speed of microwave-assisted digestion varied depending on the system used, and both systems yielded significantly lower amounts of peptide compared to conventional incubation.

Modeling the time course data confirmed that both microwave systems yielded less than half the amount produced by conventional incubation for the three slowest-forming surrogate peptides (LED, VIL, and EFS). This suggests that formation may have ceased due to enzyme inactivation within the microwave reactor, rather than reaching a 'dead end' [24]. Despite providing lower peptide yield, the MARS6 system significantly reduced digestion time for all surrogate peptides. Similarly, pressure-assisted digestion showed lower yield for the three slowforming peptides (LED, VIL, and EFS), but resulted in markedly reduced digestion. In fact, maximum yield with pressure-assisted digestion was achieved in under two hours for all peptides and in 30 min or less for 5 of 8 peptides.

3.4. Trypsin concentration

The above studies were all conducted with a relatively low enzyme: protein ratio (w/w) of approximately 1:200 for trypsin digestion. However, despite the low ratio, each analysis consumed a significant amount of trypsin (60 μ g) due to the large volume (200 μ L) and high protein concentration in serum (~60 mg/mL). Nevertheless, because TPCK-treated trypsin is relatively inexpensive, higher amounts were evaluated in a time course study to assess their impact on yield and digestion speed. As expected, increasing the amount of trypsin resulted in higher rates of peptide formation (Fig. 5A); however, the concentration of trypsin had little effect on the amount of surrogate peptide produced. For each surrogate peptide, the maximum yield observed in each condition was within \pm 20 % of the overall average (Fig. 5B). Additionally, for most peptides, maximum yield was achieved in less than two hours. The exception was the LED peptide, which showed slower formation even when trypsin concentration was increased by 10-fold. This could be attributed to the presence of acidic residues at both termini known to inhibit trypsin digestion [21].

The only adverse effect of increasing the amount of trypsin occurred with the GGA peptide, which exhibited degradation at higher concentrations of TPCK-treated trypsin. This phenomenon has been previously reported [25] and is suspected to be caused by non-specific cleavage of hydrophobic residues (F, Y, W, M, L, and H) by autolyzed trypsin and/or contaminating chymotrypsin. Although the GGA peptide contains hydrophobic leucine residues that are susceptible to this non-specific cleavage, it does not appear predictive given that the EFS and FSP peptides also contain pairs of leucine residues (see Table 1), yet do not show evidence of degradation at higher concentrations of TPCK-treated trypsin.

4. Discussion

Bottom-up, shotgun proteomics seeks to identify and quantify the maximum number of proteins feasible, in order to increase the likelihood of obtaining biologically or clinically meaningful insights. Given the extensive research in shotgun proteomics, many aspects of trypsin digestion have been optimized with this goal in mind. However, relatively few studies have focused on optimizing trypsin digestion for targeted protein analysis, which only requires enough surrogate peptide to detect and quantify the associated protein. Clinical laboratories prioritize speed and yield as key metrics for digestion in targeted protein assays, while also considering simplicity and affordability [26].

Based on this understanding of digestion efficiency, we conducted an



Fig. 4. Digestion was conducted in four different 'enzyme reactors' at eight different time points. These time points ranged from 0.5 to 18 h for the Thermomixer, 8 min to 4 h for REDS and MARS6, and 4 min to 2 h for the Barocycler. Apart from these variations, all digestions followed the same procedure, which involved 0.2 % DOC and an enzyme:protein ratio of 1:200 (w/w) of TPCK-treated bovine trypsin. (A) The concentrations of the VIL, GGA, and EFS surrogate peptides were measured and plotted for each 'enzyme reactor' and time point tested. Each data point represents the mean result of triplicate digestions, with error bars indicating +/-1SD. However, data obtained after > 4 h for the Thermomixer are not shown on the plots to allow for better visualization of the other 'enzyme reactors'. The solid line overlaid on the data represents the pseudo-first order model of each digestion. (B) The maximum concentration obtained for each surrogate peptide was determined from the model, along with the digestion time required to achieve maximum yield.



Fig. 5. Trypsin digestion was conducted using four different concentrations of trypsin at eight time points ranging from 0.5 to 18 h. All digestions followed the same procedure, which included 0.2% DOC, TPCK-treated bovine trypsin, and incubation on a Thermomixer. (A) Measured concentrations of VIL, GGA, and EFS surrogate peptides are plotted against each trypsin concentration and time point analyzed. Each data point represents the mean value obtained from triplicate digestions, with the error bars indicating +/- 1SD. The solid line displayed in the graph represents the pseudo-first order model applied to each digestion. (B) The maximum concentration achieved for each surrogate peptide, as determined by the model, along with the corresponding digestion duration required to achieve this maximum yield.

evaluation of various mechanisms for altering tryptic activity. Specifically, we focused on 8 surrogate peptides for thyroglobulin, a clinical biomarker. Two of these surrogate peptides (VIF and FSP) are already utilized in clinical tests [14-17]. Previous studies have highlighted the diverse impact of different additives used for protein digestion in serum [27], underscoring the importance of evaluating multiple additives – at varying concentrations - to identify those that enable efficient digestion of a specific protein or surrogate peptide. Our results also demonstrate the advantages of examining multiple time points during early development to select additives that not only yield an ample amount of surrogate peptides but also promote rapid digestion without false maxima caused by peptide degradation. Other studies have suggested that heat denaturation alone may suffice for digestion of certain serum proteins [10,11]; however, these studies employed significantly higher temperatures to denature highly diluted serum samples, which was not feasible in our work using large volumes of minimally diluted plasma that resulted in a semi-solid specimen impeding further processing.

While additive selection had a significant impact, the use of 'sequencing grade' trypsin had minimal positive or negative effects on digestion speed or yield despite their reported benefits for protein and peptide sequencing [28]. Although one peptide (EFS) exhibited considerably faster digestion when using a LysC/Trypsin mixture (<2 h) compared to TPCK-treated bovine trypsin (>10 h), similar rapid digestion (<1 h) could be achieved for the same peptide by increasing the amount of TPCK-treated bovine trypsin 10-fold. As noted by others [7], this is particularly pertinent because 'sequencing grade' enzymes used in these studies are expected to cost 1000 times more (by dry weight) than the TPCK-treated bovine trypsin employed here. Therefore, while it is conceivable that further optimization studies specific to each trypsin could improve yield or speed of digestion, the potential benefits are unlikely to justify the higher cost.

Notably, all optimization and comparative studies performed herein used a relatively low enzyme:protein ratio of 1:200 (w/w) compared to most proteomic studies. This was done due to the amount of serum

proteins being digested per analysis (~12 mg total protein per digestion of 200 µL serum) and the high cost of sequencing-grade trypsins. Although increasing the concentration of trypsin is expected to speed up digestion for the various conditions and trypsin types evaluated here, it did not improve peptide yield as long as sufficient digestion time was given to achieve maximum yield. However, one potential benefit of 'sequencing grade' trypsins was highlighted by the GGA peptide, which exhibited instability when exposed to higher concentrations of TPCKtreated trypsin for prolonged periods. This phenomenon has been reported previously and attributed to cleavage by contaminating chymotrypsin and/or non-specific cleavage by trypsin following autolysis, which is largely mitigated with 'sequencing grade' trypsin [20,25]. While this behavior is not desirable and would likely require using another surrogate peptide where possible, it should not inherently prevent the use of a surrogate peptide. Peptide instability simply means that digestion time needs to be well controlled in order to manage the sensitivity of the surrogate peptide's measurement and that internal standards should be added before starting the digestion step so that a steady-state light: heavy peptide ratio can be achieved for quantification after reaching maximum digestion [20].

Pressure-assisted and microwave-assisted digestion also showed potential to speed up trypsin digestion for almost all surrogate peptides with little adverse effect on their overall yield. Further improvement may be possible if additives were optimized for each enzyme reactor – for example, 1-PrpOH may provide better digestion efficiency than DOC with pressure-assisted digestion [29]; however, further optimization of the enzyme reactors was not attempted as similar, rapid digestion speeds were achievable with conventional incubation by simply increasing the amount of trypsin. In fact, increasing the concentration of trypsin 10fold with conventional incubation was feasible due to the relatively low cost of TPCK-treated bovine trypsin and was more favorable for a high-throughput clinical lab environment compared to using specialized reactors and associated consumables.

One limitation of this study was that it only evaluated 8 surrogate

peptides from a single protein biomarker. It has been shown that trypsin digestion efficiency can vary between proteins [27] and is strongly influenced by acidic residues or repeat arginine/lysine residues surrounding cleavage sites [21], making the selection of surrogate peptides critical when trying to find universally optimal conditions for multiplexed protein/peptide assays. Although the conclusion of using more trypsin to accelerate digestion should be generally applicable based on standard kinetic principles, it is possible that other proteins or surrogate peptides may benefit to a greater extent in terms of yield or speed of digestion when using sequencing grade trypsin or specialized enzyme reactors compared to the 8 surrogate peptides of thyroglobulin studied here. Additionally, different matrix types (e.g., serum versus tissue) may also prefer other enzymes or reactors. Nonetheless, this study provides a framework for empirically comparing variables that influence digestion for targeted quantification of any protein biomarker(s), with a focus on digestion metrics important to the clinical laboratory, such as speed and vield.

It should also be noted that immobilized trypsins have the potential to increase the speed and efficiency of digestion, but this was not evaluated in this study. Specifically, thermally-stable trypsin immobilized onto solid resin or freely in solution has been shown to facilitate rapid and efficient digestion of serum proteins without the need for additional denaturation steps [30]. This would be ideal for targeted applications in a high-throughput clinical lab environment, however, the cost of these engineered trypsin is still substantial compared to the TPCK-treated bovine trypsin used in this study. Additionally, online digestion using immobilized trypsin has been shown to provide rapid digestion for individual samples [31], but it is relatively expensive on a per sample basis, has limited throughput due to sequential digestions instead of parallel digestions, and cannot process large volumes of unfractionated serum proteins.

To address this issue, we aimed to find the simplest and most costeffective way to accelerate digestion of large volumes of serum. This was achieved by increasing the amount of TPCK-treated bovine trypsin after identifying the additive, and the additive concentration, that generally provided the highest yield of each surrogate peptide in the shortest amount of time. We found that at a 1:20 (w/w) trypsin:protein ratio, maximum yield was achieved in less than 2 h for 7 of the 8 surrogate peptides for thyroglobulin. For 4 out of 8 surrogate peptides, maximum yield was achieved in less than 1 h using the optimized additive - in this case, 0.2 % DOC. It may be possible to further increase digestion speed by increasing the amount of trypsin further (i.e., using a 1:10 or 1:5 (v/v) trypsin:protein ratio) or by slightly increasing the incubation temperature [32] without adversely affecting yield.

However, it is worth noting that the LED peptide continued to form slowly, requiring over 10 h to reach maximum yield at the highest trypsin:protein ratio tested. This highlights the importance of selecting appropriate surrogate peptides when developing a targeted assay for protein quantification. The specificity of the surrogate peptide sequence is crucial to ensure the correct measurand is considered, and many "rules" for surrogate peptide selection focus on optimizing sensitivity based on predicted digestion efficiency and stability. Nevertheless, this study emphasizes the importance of considering pragmatic considerations, such as digestion speed and/or ease (requiring fewer steps or less specialized/costly reagents) when selecting surrogate peptides, especially for tests performed in the clinical lab.

5. Conclusions

These studies demonstrate that empirical optimization can make trypsin digestion efficient in terms of speed, yield, and cost. While 'sequencing grade' trypsins and enzyme reactors offered some advantages, their higher cost or increased complexity may not be justified for a high throughput clinical lab. Our research emphasizes the importance of optimizing the additives used for denaturation and digestion, followed by titration of 'standard' TPCK-treated bovine trypsin. This approach can result in a highly efficient and stable digestion process for the majority of surrogate peptides. Therefore, we recommend that digestion efficiency and stability be considered as key design criteria for any high throughput clinical assay. In such cases, slow-forming or unstable surrogate peptides should be excluded in favor of those that respond well to the high concentrations of trypsin necessary to facilitate rapid digestion.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Martin K. Green for his contribution to this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmsacl.2023.11.002.

References

- J.V. Olsen, S.E. Ong, M. Mann, Trypsin cleaves exclusively C-terminal to arginine and lysine residues, Mol. Cell. Proteomics 3 (2004) 608–614, https://doi.org/ 10.1074/mcp.T400003-MCP200.
- [2] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data, Electrophoresis 20 (1999) 3551–3567, https://doi.org/10.1002/(SICI)1522-2683 (19991201)20:18<3551::AID-ELPS3551>3.0.CO(2-2.
- [3] J. Cox, N. Neuhauser, A. Michalski, R.A. Scheltema, J.V. Olsen, M. Mann, Andromeda: A peptide search engine integrated into the MaxQuant environment, J Proteome Res. 10 (2011) 1794–1805, https://doi.org/10.1021/pr101065j.
- [4] J.M. Burkhart, C. Schumbrutzki, S. Wortelkamp, A. Sickmann, R.P. Zahedi, Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics, J Proteomics. 75 (2012) 1454–1462, https://doi.org/10.1016/j.jprot.2011.11.016.
- [5] B. Niu, M. Martinelli, Y. Jiao, C. Wang, M. Cao, J. Wang, E. Meinke, Nonspecific cleavages arising from reconstitution of trypsin under mildly acidic conditions, PLoS One 15 (2020) 1–18, https://doi.org/10.1371/journal.pone.0236740.
- [6] M. Hu, Y. Liu, K. Yu, X. Liu, Decreasing the amount of trypsin in in-gel digestion leads to diminished chemical noise and improved protein identifications, J Proteomics. 109 (2014) 16–25, https://doi.org/10.1016/j.jprot.2014.06.017.
- [7] J. Bunkenborg, G. Espadas, H. Molina, Cutting edge proteomics: Benchmarking of six commercial trypsins, J Proteome Res. 12 (2013) 3631–3641, https://doi.org/ 10.1021/pr4001465.
- [8] X.J. Li, C. Hayward, P.Y. Fong, M. Dominguez, S.W. Hunsucker, L.W. Lee, M. McLean, S. Law, H. Butler, M. Schirm, O. Gingras, J. Lamontagne, R. Allard, D. Chelsky, N.D. Price, S. Lam, P.P. Massion, H. Pass, W.N. Rom, A. Vachani, K. C. Fang, L. Hood, P. Kearney, A blood-based proteomic classifier for the molecular characterization of pulmonary nodules, Sci Transl Med. 5 (2013), https://doi.org/ 10.1126/scitranslmed.3007013.
- [9] C. Bradford, R. Severinsen, T. Pugmire, M. Rasmussen, K. Stoddard, Y. Uemura, S. Wheelwright, M. Mentinova, D. Chelsky, S.W. Hunsucker, P. Kearney, D. Hickok, T.C. Fleischer, I. Ichetovkin, J.J. Boniface, G.C. Critchfield, J.M. Peltier, Analytical validation of protein biomarkers for risk of spontaneous preterm birth, Clinical Mass Spectrometry. 3 (2017) 25–38, https://doi.org/10.1016/j. clinms.2017.06.002.
- [10] Y.Z. Zheng, M.L. DeMarco, Manipulating trypsin digestion conditions to accelerate proteolysis and simplify digestion workflows in development of protein mass spectrometric assays for the clinical laboratory, Clinical Mass Spectrometry. 6 (2017) 1–12, https://doi.org/10.1016/j.clinms.2017.10.001.
- [11] M.L.D. Junyan Shi, Y.Z. Zheng, D.D. Sin, A Streamlined Method for Quantification of Apolipoprotein A1 in Human Plasma by LC-MS/MS, Clin Chem. 64 (2018) 1782–1784, https://doi.org/10.1016/j.nurpra.2014.09.017.
- [12] C.M. Henderson, P.L. Lutsey, J.R. Misialek, T.J. Laha, E. Selvin, J.H. Eckfeldt, A. N. Hoofnagle, Measurement by a Novel LC-MS/MS Methodology Reveals Similar Serum Concentrations of Vitamin D-Binding Protein in Blacks and Whites, Clin Chem. 62 (2016) 179–187, https://doi.org/10.1373/clinchem.2015.244541.
- [13] W.S. Phipps, D.N. Greene, H. Pflaum, T.J. Laha, J.A. Dickerson, J. Irvine, A. E. Merrill, P. Ranjitkar, C.M. Henderson, A.N. Hoofnagle, Small volume retinol binding protein measurement by liquid chromatography-tandem mass spectrometry, Clin Biochem. 99 (2022) 111–117, https://doi.org/10.1016/j. clinbiochem.2021.10.005.
- [14] A.N. Hoofnagle, J.O. Becker, M.H. Wener, J.W. Heinecke, Quantification of Thyroglobulin, a Low-Abundance Serum Protein, by Immunoaffinity Peptide Enrichment and Tandem Mass Spectrometry, Clin Chem. 54 (2008) 1796–1804, https://doi.org/10.1373/clinchem.2008.109652.

C.M. Shuford and R.P. Grant

- [15] M.M. Kushnir, A.L. Rockwood, W.L. Roberts, D. Abraham, A.N. Hoofnagle, A. W. Meikle, Measurement of thyroglobulin by liquid chromatography - Tandem mass spectrometry in serum and plasma in the presence of antithyroglobulin autoantibodies, Clin Chem. 59 (2013) 982–990, https://doi.org/10.1373/ clinchem.2012.195594.
- [16] N.J. Clarke, Y. Zhang, R.E. Reitz, A Novel Mass Spectrometry-Based Assay for the Accurate Measurement of Thyroglobulin From Patient Samples Containing Antithyroglobulin Autoantibodies, J. Invest. Med. 60 (2012) 1157–1163, https:// doi.org/10.2310/JIM.0b013e318276deb4.
- [17] B.C. Netzel, S.K.G. Grebe, A. Algeciras-Schimnich, Usefulness of a Thyroglobulin Liquid Chromatography-Tandem Mass Spectrometry Assay for Evaluation of Suspected Heterophile Interference, Clin Chem. 60 (2014) 1016–1018, https://doi. org/10.1373/clinchem.2014.224816.
- [18] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966–968, https://doi.org/10.1093/bioinformatics/btq054.
- [19] C.M. Shuford, J.J. Walters, P.M. Holland, U. Sreenivasan, N. Askari, K. Ray, R. P. Grant, Absolute Protein Quantification by Mass Spectrometry: Not as Simple as Advertised, Anal Chem. 89 (2017) 7406–7415, https://doi.org/10.1021/acs. analchem.7b00858.
- [20] C.M. Shuford, R.R. Sederoff, V.L. Chiang, D.C. Muddiman, Peptide Production and Decay Rates Affect the Quantitative Accuracy of Protein Cleavage Isotope Dilution Mass Spectrometry (PC-IDMS), Mol. Cell. Proteomics 11 (2012) 814–823, https:// doi.org/10.1074/mcp.0112.017145.
- [21] T. Šlechtová, M. Gilar, K. Kalíková, E. Tesařová, Insight into Trypsin Miscleavage: Comparison of Kinetic Constants of Problematic Peptide Sequences, Anal Chem. 87 (2015) 7636–7643, https://doi.org/10.1021/acs.analchem.5b00866.
- [22] Y. Lin, J. Zhou, D. Bi, P. Chen, X. Wang, S. Liang, Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins, Anal Biochem. 377 (2008) 259–266, https://doi.org/10.1016/J.AB.2008.03.009.
- [23] P.D. Gershon, Cleaved and missed sites for trypsin, Lys-C, and Lys-N can be predicted with high confidence on the basis of sequence context, J Proteome Res. 13 (2014) 702–709, https://doi.org/10.1021/pr400802z.

- [24] P. Brownridge, R.J. Beynon, The importance of the digest: Proteolysis and absolute quantification in proteomics, Methods 54 (2011) 351–360, https://doi.org/ 10.1016/j.ymeth.2011.05.005.
- [25] P.L. Loziuk, J. Wang, Q. Li, R.R. Sederoff, V.L. Chiang, D.C. Muddiman, Understanding the role of proteolytic digestion on discovery and targeted proteomic measurements using liquid chromatography tandem mass spectrometry and design of experiments, J Proteome Res. 12 (2013) 5820–5829, https://doi.org/ 10.1021/pr4008442.
- [26] C64: Quantitative Measurement of Proteins and Peptides by Mass Spectrometry, 1st ed., Clinical and Laboratory Standards Institute, 2021.
- [27] M.A. Kuzyk, D.B. Hardie, J. Yang, D.S. Smith, A.M. Jackson, C.E. Parker, C. H. Borchers, A Quantitative Study of the Effects of Chaotropic Agents, Surfactants, and Solvents on the Digestion Efficiency of Human Plasma Proteins by Trypsin, J Proteome Res. 9 (2010) 5422, https://doi.org/10.1021/pr100656u.A.
- [28] T. Glatter, C. Ludwig, E. Ahrné, R. Aebersold, A.J.R. Heck, A. Schmidt, Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion, J Proteome Res. 11 (2012) 5145–5156, https://doi.org/10.1021/ pr300273g.
- [29] Proteolysis-PrEP: Optimization of Pressure Cycling Conditions for Protein Digestion with Trypsin in Solution, Pressure Biosciences. https://www. pressurebiosciences.com/documentsbytype/trypsin-prep-91815-v1-pdf/ viewdocument/22, (accessed 13 March 2022).
- [30] V. Shah, M.E. Lassman, Y. Chen, H. Zhou, O.F. Laterza, Achieving efficient digestion faster with Flash Digest: potential alternative to multi-step detergent assisted in-solution digestion in quantitative proteomics experiments, Rapid Commun. Mass Spectrom. 31 (2017) 193–199, https://doi.org/10.1002/rcm.7774
- [31] F.E. Regnier, J. Kim, Accelerating trypsin digestion: the immobilized enzyme reactor, Bioanalysis 6 (2014) 2685–2698, https://doi.org/10.4155/bio.14.216.
- [32] J.L. Nickerson, A.A. Doucette, Maximizing Cumulative Trypsin Activity with Calcium at Elevated Temperature for Enhanced Bottom-Up Proteome Analysis, Biology (basel). 11 (2022), https://doi.org/10.3390/biology11101444.