Available online at www.sciencedirect.com
ScienceDirect
International Journal of Mass Spectrometry 259 (2007) 174–183
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Identification of peptidase substrates in human plasma by FTMS based differential mass spectrometry

Nathan A. Yates [a], Ekaterina G. Deyanova [a], Wayne Geissler [b], Matthew C. Wiener [c], Jeffrey R. Sachs [c], Kenny K. Wong [b], Nancy A. Thornberry [b], Ranabir Sinha Roy [b], Robert E. Settlage [a], Ronald C. Hendrickson [a,*]

[a] *Department of Molecular Profiling, Proteomics, Merck Research Laboratories, 126 E Lincoln Ave, Rahway NJ 07065, United States*
[b] *Department of Metabolic Disorders, Merck Research Laboratories, 126 E Lincoln Ave, Rahway NJ 07065, United States*
[c] *Department of Applied Computer Science and Mathematics, Merck Research Laboratories, 126 E Lincoln Ave, Rahway NJ 07065, United States*

Received 16 June 2006; received in revised form 11 September 2006; accepted 12 September 2006
Available online 23 October 2006

Abstract

Approximately 2% of the human genome encodes for proteases. Unfortunately, however, the biological roles of most of these enzymes remain poorly defined, since the physiological substrates are typically unknown and are difficult to identify using traditional methods. We have developed a proteomics experiment based on FTMS profiling and differential mass spectrometry (dMS) to identify candidate endogenous substrates of proteases using fractionated human plasma as the candidate substrate pool. Here we report proof-of-concept experiments for identifying *in vitro* substrates of aminopeptidase P2, (APP2) and dipeptidyl peptidase 4 (DPP-4), a peptidase of therapeutic interest for the treatment of type 2 diabetes. For both proteases, previously validated peptide substrates spiked into the human plasma pool were identified. Of note, the differential mass spectrometry experiments also identified novel substrates for each peptidase in the subfraction of human plasma. Targeted MS/MS analysis of these peptides in the complex human plasma pool and manual confirmation of the amino acid sequences led to the identification of these substrates. The novel DPP-4 substrate EPLGRQLTSGP was chemically synthesized and cleavage kinetics were determined in an *in vitro* DPP-4 enzyme assay. The apparent second order rate constant (k_{cat}/K_M) for DPP-4-mediated cleavage was determined to be $2.3 \times 10^5\ M^{-1}\ s^{-1}$ confirming that this peptide is efficiently processed by the peptidase *in vitro*. Collectively, these results demonstrate that differential mass spectrometry has the potential to identify candidate endogenous substrates of target proteases from a human plasma pool. Importantly, knowledge of the endogenous substrates can provide useful insight into the biology of these enzymes and provides useful biomarkers for monitoring their activity *in vivo*.

Keywords: Proteomics; Differential mass spectrometry; Human plasma; Biomarker; FTMS

1. Introduction

Proteases play important roles in diverse physiological processes ranging from cellular metabolism and immune surveillance to cell survival and apoptosis. Although proteases are recognized as important enzymes, they represent a largely untapped family of potential therapeutic targets. The biological roles of most human proteases remain poorly defined, since their physiological substrates are often unknown and are difficult to identify using traditional methods such as HPLC [1] and zymography [2]. Positional scanning techniques [3] and combinatorial library strategies [4] are often effective in determining subsite specificity near the sessile bond; however, these approaches require prior knowledge of the substrate and may be limited by the diversity of the peptide library.

Comparative proteomic profiling, via MALDI-TOF (matrix assited laser desorption ionization) or LC–MS (liquid chromatography–mass spectrometry), provides an alternative approach to identify novel physiological substrates. Such approaches have been used in the analysis of complex samples to identify peptides whose ion abundances are quantitatively

Abbreviations: dMS, differential mass spectrometry; FTMS, Fourier transform mass spectrometry

* Corresponding author at: Merck & Co., Inc., 126 E Lincoln Ave, RY800-B301, Rahway NJ 07065, United States. Tel.: +1 732 594 5940; fax: +1 732 594 3371.
E-mail address: ronald_hendrickson@merck.com (R.C. Hendrickson).

1387-3806/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2006.09.020

different between two well-controlled experimental conditions. Examples include the identification of class I epitopes from HIV infected B cells [5], substrates processed by the multi-enzyme proteosome complex [6], and even virulence factors [7].

LC–MS based methods are well suited to the analysis of complex biological mixtures and provide three-dimensional profiles that allow analytes to be differentiated on the basis of mass-to-charge ratio (m/z), retention time, and ion intensity. High resolution Fourier transform mass spectrometers (FTMS) offer new opportunities for the analysis of increasingly complex mixtures and the recent development of ion trap FTMS hybrid instruments has made it possible to acquire high quality LC–FTMS data [8–11]. These instruments have the capacity to generate data with mass accuracy better than 5 ppm, resolving power in excess of 100,000, large dynamic range, and absolute sensitivity in the attomole range (10^{-18} moles). Such improved instrument performance and the need to identify statistically meaningful quantitative differences in peptide and protein abundance without the use of stable isotope labels has heightened the importance for LC–MS based profiling methods.

In practice, difficulties associated with the acquisition and analysis of LC–MS profiles have limited the utility of comparative proteomic profiling methods. Retention time stability, LC separation consistency and the reliable mass spectrometer response are requirements to generate reproducible LC–MS data for a large number of samples. In addition, the large volume and complexity of LC–MS profile data make manual analysis increasingly difficult. Both of these challenges are becoming less of an issue as instrument vendors strive to increase the tolerance of their products and as automated software tools that reliably analyze LC–MS profile data become available [12,13].

Here we describe the development of an LC–FTMS based comparative protein profiling method for the detection of protease substrates in biological fluids such as plasma and cerebral spinal fluid (CSF). We evaluated reproducible sample preparation methodologies that could generate pools of candidate substrates for interrogation against any target protease of interest. Fractionation technology from Digilab Biovision GmbH, Hanover, Germany, was adopted for the isolation of low molecular weight peptides (typically <15,000 Da) from human plasma ultrafiltrate [14]. Incubation of processed human plasma samples with target proteases followed by reproducible LC–FTMS profiling and automated data analysis allowed us to establish an experimental paradigm for detecting protease substrates in human plasma, which we call differential mass spectrometry (dMS) [13]. In these experiments, complex mixtures of potential peptide substrates are incubated with a target protease and the hydrolysis of peptides is monitored over time, which is quite similar to that of a routine enzymatic assay that measures the rate of proteolysis for a single substrate. The unique characteristic, however, is that differential mass spectrometry provides a multiplexed assay wherein virtually all peptides signals detected in the mass spectrometer are monitored, hence facilitating the analysis of complex and "undefined" biological mixtures. Ions that show a statistically significant change in abundance over time that is characteristic of a substrate or product peptide, are distinguished from those peptides, which are unchanged. Targeted MS/MS analysis is used to acquire product ion spectra from which the amino acid sequence is deduced. Here, we describe two proof-of-concept differential mass spectrometry experiments to identify substrates of DPP-4 [15], an enzyme for which highly selective small molecule inhibitors are available, and aminopeptidase P2 (APP2) [16], an enzyme for which no selective chemical inhibitor has been reported.

2. Experimental design and methods

2.1. DPP-4—prototypic peptidase for which a selective chemical inhibitor is available

Human blood was collected in EDTA vacutainer tubes, chilled on ice and processed to separate plasma via standard laboratory protocols. Human plasma ultrafiltrate was prepared from these samples [14] by Digilab Biovision GmbH. In a typical experiment, 0.5 ml plasma equivalents of ultrafiltrate were re-suspended in 55 μl assay buffer (10 mM HEPES, pH 7.5, 50 μg/ml BSA) and 250 nM each of control peptides GLP-1(7-36)-amide, GHRH(1-44), and PACAP38, corresponding to validated *in vitro* substrates of DPP-4 [17]. Samples were sonicated for 5 min and centrifuged at 15,000 × g to remove insoluble material. The supernatant was incubated with 10 nM recombinant soluble DPP-4 (1 μl of 550 nM stock) [18]. The experiment consisted of two arms: one where the fractionated plasma samples in assay buffer were spiked with control peptides and treated with DPP-4 as described, the other where identical samples were also treated with 10 μM of compound A, (2*R*)-4-oxo-4-[3-(trifluoromethyl)-5,6-dihydro[1,2,4]triazolo [4,3-a]pyrazin-7(8H)-yl]-1-(2,5-difluorophenyl)butan-2-amine fumarate, a potent and highly selective inhibitor of DPP-4 [19]. Both samples were incubated at 37 °C and 20 μl aliquots were removed at 0 min and 60 min time points and quenched with HCl to a final concentration of 100 mM for LC–MS analysis.

2.2. Aminopeptidase P2—proof of concept where no potent selective chemical inhibitor is available

Aminopeptidase P2 enzyme was purchase from (R&D Systems, Minneapolix, MN (cat #2490-ZN)) and incubations were performed essentially as described above but without the addition of GLP-1(7-36)-amide, GHRH(1-44), and PACAP38 or a small molecule inhibitor. The experiment consisted of a negative control arm comprising only fractionated plasma sample in assay buffer, and a positive arm, corresponding to an identical sample treated with 10 nM aminopeptidase P2.

2.3. Mass spectrometry—high resolution LC–MS profiling on the LTQ-FTMS

Each sample was profiled by performing replicate injections (APP2 $n = 5$, DPP-4 $n = 10$) using a commercially available LTQ-FT mass spectrometer (Thermo Electro Corp., San Jose CA) equipped with a micro-capillary HPLC system (Agilent, San Mateo CA) and a nano-elecrospray ionization source (Proxeon, Odense Denmark). The micro-capillary HPLC apparatus

consisted of a FAMOS micro autosampler (Dionex Corp., Sunnyvale CA) and a 1100 series capillary pump (Agilent Technologies, Palo Alto CA). An HPLC solvent system composed of solvent A (0.1 M HOAc) and 90% MeCN in 0.1 M HOAc vs. (MecN in the article) was used to produce a linear gradient that increased the organic composition from 0%B to 30%B at a rate of 3%B μl^{-1}. The microliter pickup mode was used to inject 1.0 μl of each sample (∼10 μl plasma equivalent) onto the head of a 75 μm × 15 cm fused silica column packed with POROS R1 (Applied Biosystems) material. A flow rate of 1.0 μL/min was used to elute peptides from the column and introduce them into the nano-electrospray source. The primary instrument parameters that were used to ionize the peptides and transmit the resulting gas phase ions into the LTQ-FT were: 3.5 kV source voltage, 240 °C capillary temperature, 25.0 V capillary potential, tube lens potential is 85 V vs. (60 V in the article), and no sheath or auxiliary gas flow. An AGC cut off was used to limit the maximum ion current introduced into the mass spectrometer to 1.0e6 and 1.0e4 (arbitrary units) vs. (1.5e4 in the article) for Fourier transform (FT) and linear ion trap (IT) full scan spectra, respectively. An AGC target value of 5.0e3 was used for IT MSn spectra. Full scan FT specta were recorded from *m/z* 300 to 2000 at a resolution setting of 50,000 and mass accuracy <5 ppm. Coincident with the acquisition of the FT spectra, three low resolution IT tandem mass spectra were recorded using data dependent acquisition. A total of eight samples (five injections each) were analyzed for the APP2 experiment with an injection-to-injection time of 45 min. All samples were maintained at 4 °C for the duration of the LCMS analysis.

2.4. Differential mass spectrometry, targeted MS/MS analysis, and peptide identification

To identify candidate substrates and products of a target protease, we utilized differential mass spectrometry, a label free LC–MS method for detecting significant differences in complex peptide and protein mixtures, as previously described [13]. Fig. 1 shows an overview of the general workflow that was used to assay for APP2 substrates and products. Enzyme was added to one of two technical replicates of fractionated plasma (the "positive" arm), and aliquots were removed from both arms and quenched at 0, 10, 20, and 30 min time points. LCMS profiles were recorded using five technical replicates for each sample and the resulting data images were grouped by assay condition (±enzyme). To facilitate comparisons between LC–MS runs, *m/z*, and time measurements were rounded to the nearest (0.01) for *m/z* and to the nearest 0.05 min (3 s) for time. For DPP-4 experiments, enzyme was added to both samples and a selective DPP-4 inhibitor was added or absent in the positive and negative arms, respectively.

A peak alignment algorithm was used to compensate for peak retention time shift. Briefly, at each binned *m/z*, we searched the

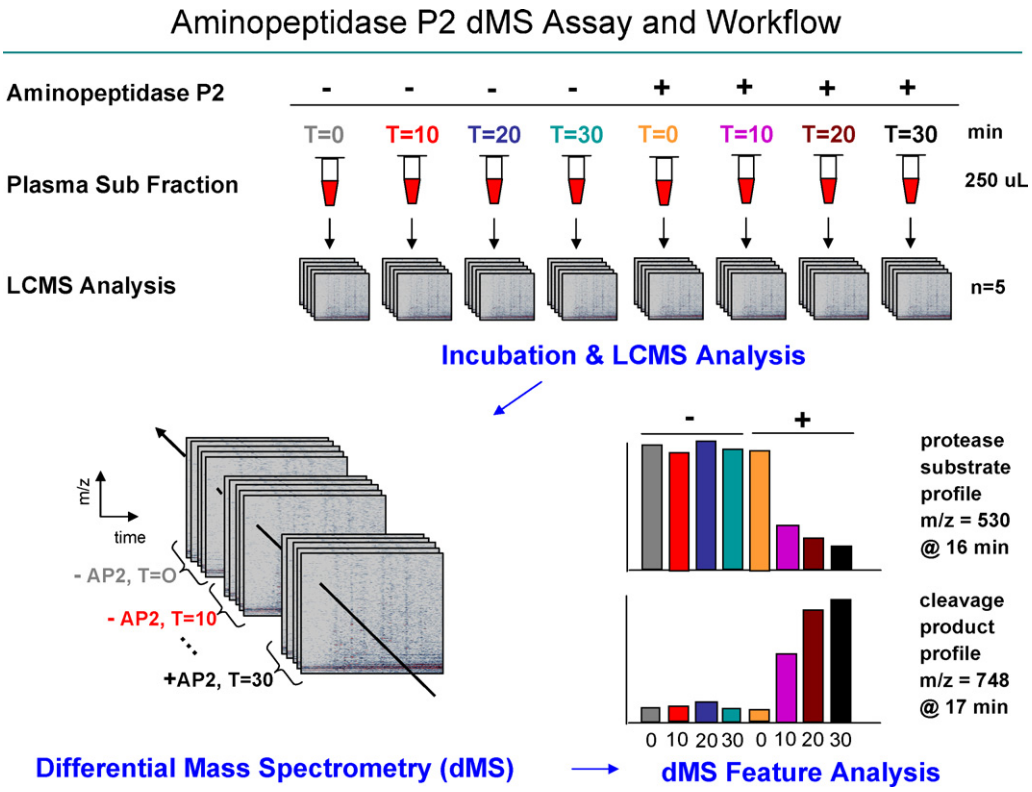

Fig. 1. Schematic representation of the differential mass spectrometry assay used to identify endogenous substrates of aminopeptidase P2 in human plasma. Two technical replicates of a human plasma peptide sub fraction were incubated ±APP2 and time points were collected at 0, 10, 20, and 30 min intervals. Each of the resulting eight samples was analyzed by LC–MS ($n = 5$) on an LTQ-FTMS that was equipped with a micro-capillary LC column and nano-spray ionization source. Differential mass spectrometry (dMS) was used to detect ion signals ("features") that exhibited a statistically significant change in abundance between assay conditions (±enzyme) and over time. Histograms for each dMS feature were visually inspected to select features the exhibit the temporal profiles of protease substrates and products. Tandem mass spectrometry (not shown) was used to identify the amino acid sequence of select peptides.

selected ion chromatograms for peaks existing in all samples within a condition. Due to small changes in chromatography from run to run, the common peaks will probably not occur at exactly the same time in the data for each sample. For example, a peak at *m/z* 601.23, might appear at 28.2 min in the reference sample (which we choose arbitrarily) and at 28.4 min in another sample. From this we can calculate the shift that will align the peak with the corresponding peak in a reference sample. In the example above, the time 28.2 would need to be shifted forward 0.2 min. Due to estimation issues, different peaks may give rise to different shifts at the same or nearby times. We make the assumption that the actual shift required depends only on elution time, and so smooth the shifts arising from different peaks to find a "consensus" function to align a particular sample to the reference sample for the appropriate condition. Similarly, peaks existing in the mean data for each condition can be used to align the reference samples for different conditions. Applying the within-condition and then the between-condition alignments puts all samples into a common time frame for analysis.

Differences were identified by performing a non-parametric test (Kruskal test) at each point and requiring those features that reached significance above a point wise *p*-value threshold of $p \leq 0.05$ to persist in time. The *p*-value for a feature is taken to be the geometric mean of pointwise *p*-values. This is not strictly correct, because intensities as adjacent time points are not independent. Nonetheless, it is a useful guide for feature prioritization. Features believed to arise from a single analyte (as different isotopes and at different charge states) are grouped and given a dMS group number (referred to as grouped dMS features). Visual inspection of the dMS features was used to select features with temporal profiles characteristic of substrates and products (monotonic decreases and increases in ion abundance, respectively, over time). Targeted MS/MS analysis was performed to generate collision activated dissociation (CAD) mass spectra for dMS features of interest using a two *m/z* isolation width and an excitation setting of 35 (no units). Peptide sequences were obtained by manual *de novo* sequencing of MS/MS spectra and candidate sequences were searched against a human protein database in silico.

3. Results and discussion

A large family of proteases cleave circulating peptides and hormones in the vasculature. For example, angiotensin converting enzyme (ACE) processes Angiotensin I to release the vasoconstrictor Angiotensin II, and DPP-4 cleaves and inactivates the incretin hormone GLP-1 [15]. To recapitulate the complex plasma peptide pool as a source of potential protease substrates *in vitro*, we needed to reduce the complexity of human plasma and obtain samples that were specifically enriched in peptides. We therefore employed an ultrafiltration method developed by Digilab Biovision GmbH to isolate a low molecular weight (<15 kDa) sub fraction (peptidomics^{TM}) of plasma. Whereas the plasma concentrations of peptides present in this sub fraction

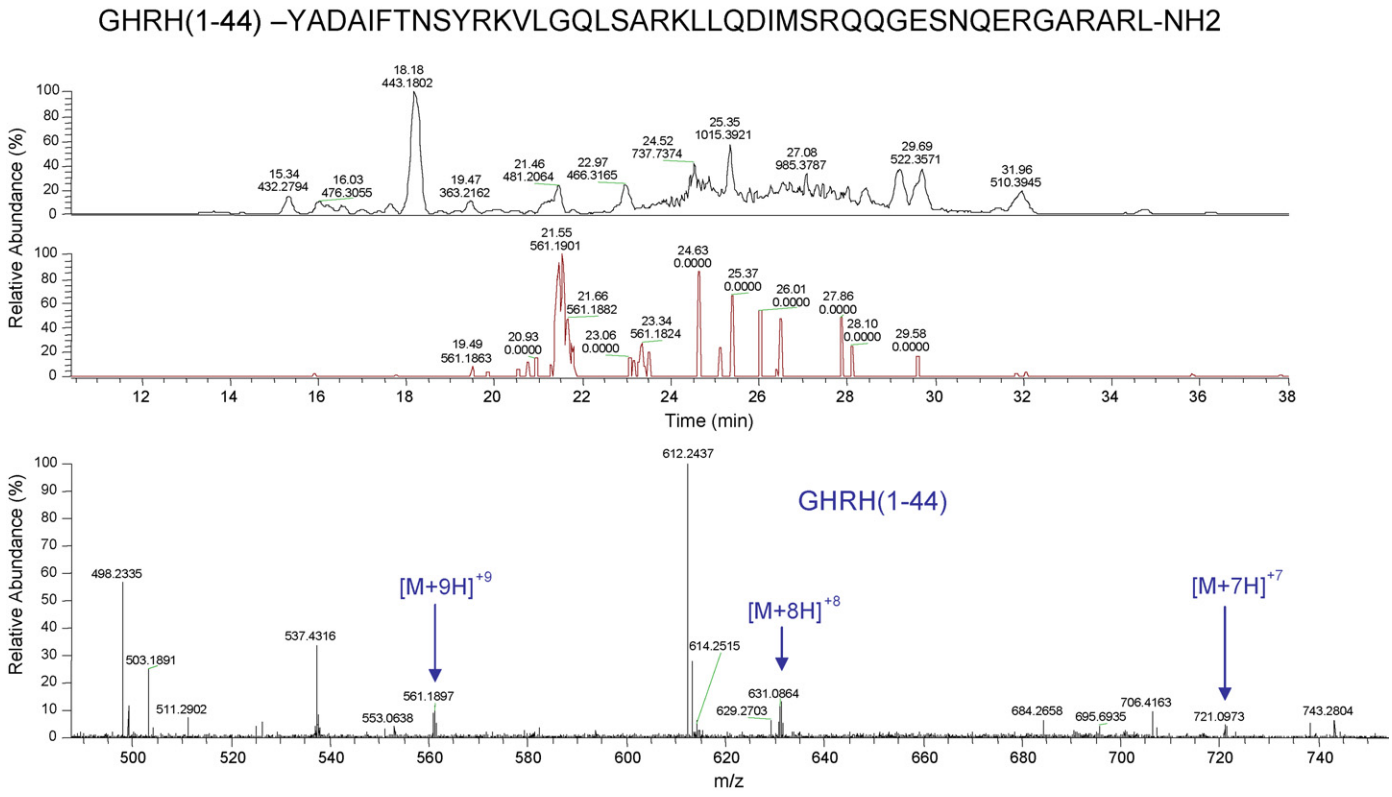

Fig. 2. Representative base peak chromatogram, selected ion chromatogram (*m/z* 561.1807–561.1967), and average full scan mass spectrum (21.35–21.69 min) for a single LC–MS analysis of a human plasma subfraction. The sample was spiked with 250 nM (250 fmol on column) of the peptide hormone GHRH(1-44) (MW_{avg} = 5039.7) that is observed as distribution of positively charged ions with +9, +8, and +7 charge states and an nominal *m/z* equal to 561, 631, and 721, respectively.

and the absolute concentrations of all peptides in the samples are unknown, the removal of high molecular weight proteins (>90% depletion) and the selective enrichment of low molecular weight peptides make the ultrafiltration procedure pragmatic for preparing a complex pool of plasma peptides from which protease substrates may be identified. Separate experiments have indicated that the Biovision methodology typically enables a 1000-fold enrichment of known peptides spiked into plasma (data not shown).

To estimate the number of peptides that are typically screened against a target protease in a single experiment, we examined the number of peptide ions detected in the LC–FTMS profiles in the Biovision sub fraction of human plasma. Fig. 2 shows a representative base peak chromatogram, selected ion chromatogram (*m/z* 561.1807–561.1967), and average full scan mass spectrum (21.35–21.69 min) for a single LC–MS analysis of a sample. In this example we spiked the sample with 250 nM (250 fmol on column) of the peptide hormone GHRH(1-44) (MW_{avg} = 5039.7), which is a validated DPP-4 substrate *in vitro* [17]. Note that the peptide is observed as a distribution of positively charged ions $(M + nH)^{+n}$ with +9, +8, and +7 charge states and corresponding *m/z* values of 561, 631, and 721, respectively. The relative abundance of the +8 GHRH signal observed at *m/z* 631 is approximately 50× below the most abundant ion species in the sample and it is estimated that as many as 500 peptides are present at relative abundances equal to or greater than the signal observed for GHRH(1-44). The relative abundance of GHRH is at least 20 times greater than the minimum signal that can be detected in our LTQ-FTMS system. On the basis of these data and the sample volumes described in Section 2, we estimate that up to 5000 peptides can be detected and screened in a single experiment.

To establish prototypic assay conditions and methodology we screened DPP-4, a peptidase that cleaves several bioactive peptides, against the fractionated plasma samples. The peptide hormones GLP-1(7-36)-amide, GHRH(1-44), and PACAP38, which are validated *in vitro* substrates of DPP-4 were spiked into the sample as positive controls for the initial experiments. Following incubation of the samples and acquisition of the LC–MS profiles, dMS was used to detect peptide signals or "dMS features" that increased or decreased in abundance with incubation time and were sensitive

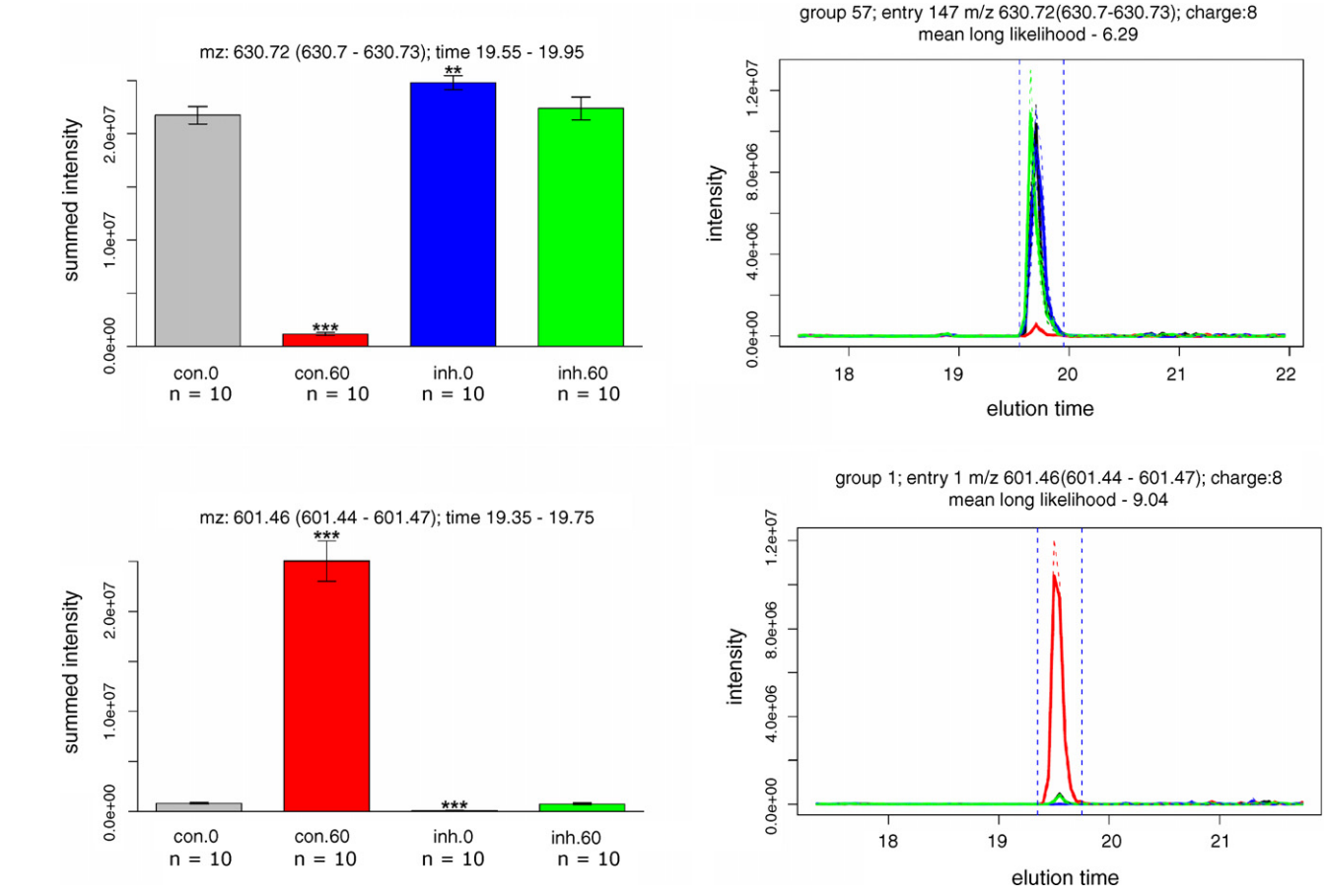

Fig. 3. Representative data for two dMS features (*m/z* 630.72 @ 19.75 min and *m/z* 601.46 @ 19.55 min) corresponding to the turnover of GHRH(1-44) by DPP-4 in an *in vitro* substrate profiling assay. The histograms and selected ion chromatograms shown correspond to the average signals obtained for ten replicate analyses of each sample (uninhibited vs. inhibited enzyme). The ion abundances shown were quantitated at 0 min (−inhibitor), 60 min (−inhibitor), 0 min (+inhibitor), and 60 min (+inhibitor), after addition of enzyme. The dMS feature at *m/z* 631.72, 21.3 min corresponds to the +8 charge state of intact GHRH(1-44) spiked into the sample as a positive control. The corresponding DPP-4 metabolite GHRH[3-44] is represented by the +9 charge state dMS feature at *m/z* 601.46, 19.55 min. The intact substrate is stabilized in the presence of the inhibitor, with concomitant suppression of product formation, confirming that the processing of GHRH[1-44] is DPP-4 dependent.

to the presence of a selective DPP-4 inhibitor, compound A [19]. Shown in Fig. 3 are representative data for two dMS features from this experiment. The histograms and selected ion chromatograms shown correspond to the average signals obtained for ten replicate analyses of each sample at 0 min (−inhibitor), 60 min (−inhibitor), 0 min (+inhibitor) and 60 min (+inhibitor) after addition of the enzyme. The control peptide GHRH(1-44) has the amino acid sequence YADAIFTNSYRKVLGQLSARKLLQDIMSRQQGESNQERGARARL-NH_2 and is cleaved by DPP-4, which releases the N-terminal dipeptide YA. The dMS features shown in Fig. 3 (*m/z* 630.72 @ 19.75 min and *m/z* 601.46 @ 19.55 min) correspond to +8 charge states of the GHRH(1-44) substrate and the corresponding des-dipeptidyl product ions, respectively. As expected, the latter elutes from the HPLC column before the full-length substrate. Inhibition of substrate disappearance and product formation in the presence of a DPP-4 chemical inhibitor confirmed that the cleavage of GHRH(1-44) was DPP-4 dependent. The comprehensive dMS analysis of inhibited and uninhibited samples generated a ranked list dMS features that are grouped on the basis of charge state and the observation of multiple isotopes to represent an individual analyte. The GHRH(3-44) product peptide ion at *m/z* 601.46 was the most significant difference detected between the two arms of the experiment (mean log likelihood = −9.04), while the corresponding intact GHRH(1-44) substrate ion at *m/z* 631.72 was the 57th ranked dMS feature group (Supplementary data table).

The inclusion of a DPP-4 selective inhibitor, compound A, in one arm of the DPP-4 experiment provided a means of differentiating between the general cleavage of peptides by residual plasma proteases in the fractionated samples, and specific DPP-4-mediated substrate cleavage. Fig. 4 shows histograms and selected ion chromatograms for dMS features (*m/z* 504.73 @ 17.8 min and *m/z* 567.94 @ 17.8 min) corresponding to the +9 and +8 charge states of the DPP-4 substrate PACAP38 [17]. In this experiment, PACAP38 decreased in abundance at the 60 min time point regardless of the presence or absence of the DPP-4 inhibitor. Additional experiments performed in the absence of DPP-4 enzyme confirmed that the *in vitro* cleavage of PACAP38 was due to background protease activity present in this sub fraction of human plasma. Hence, the unbiased detection of significant changes in peptide abundance by differential mass spectrometry, and the ability to compare dMS features across multiple treatment groups (±inhibitor or ±enzyme) provides a means of observing and differentiating the activity of multiple proteases in a single assay.

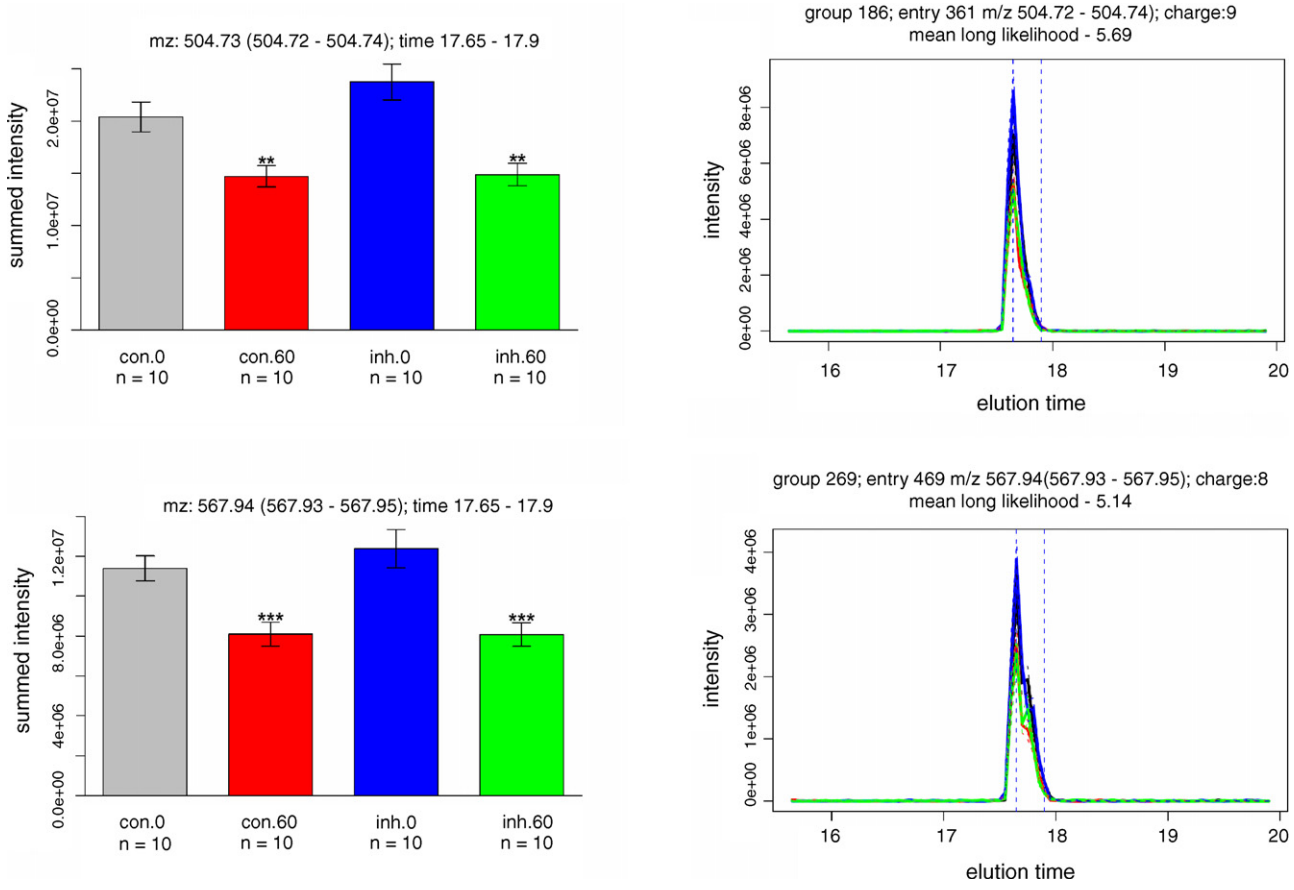

Fig. 4. Histograms and selected ion chromatograms for dMS features (*m/z* 504.73 @ 17.8 min and *m/z* 567.94 @ 17.8 min) corresponding to the +9 and +8 charge states of the known DPP-4 substrate PACAP38 spiked into the fractionated plasma samples. Features were quantitated at 0 min (−inhibitor), 60 min(−inhibitor), 0 min (+inhibitor), and 60 min (+inhibitor) after addition of the enzyme. Intact PACAP38 was observed to decrease in abundance at the 60 min time point regardless of the presence or absence (±) of the DPP-4 inhibitor, indicating that this *in vitro* cleavage was due to background protease activity present in the BioVision sub fraction of human plasma.

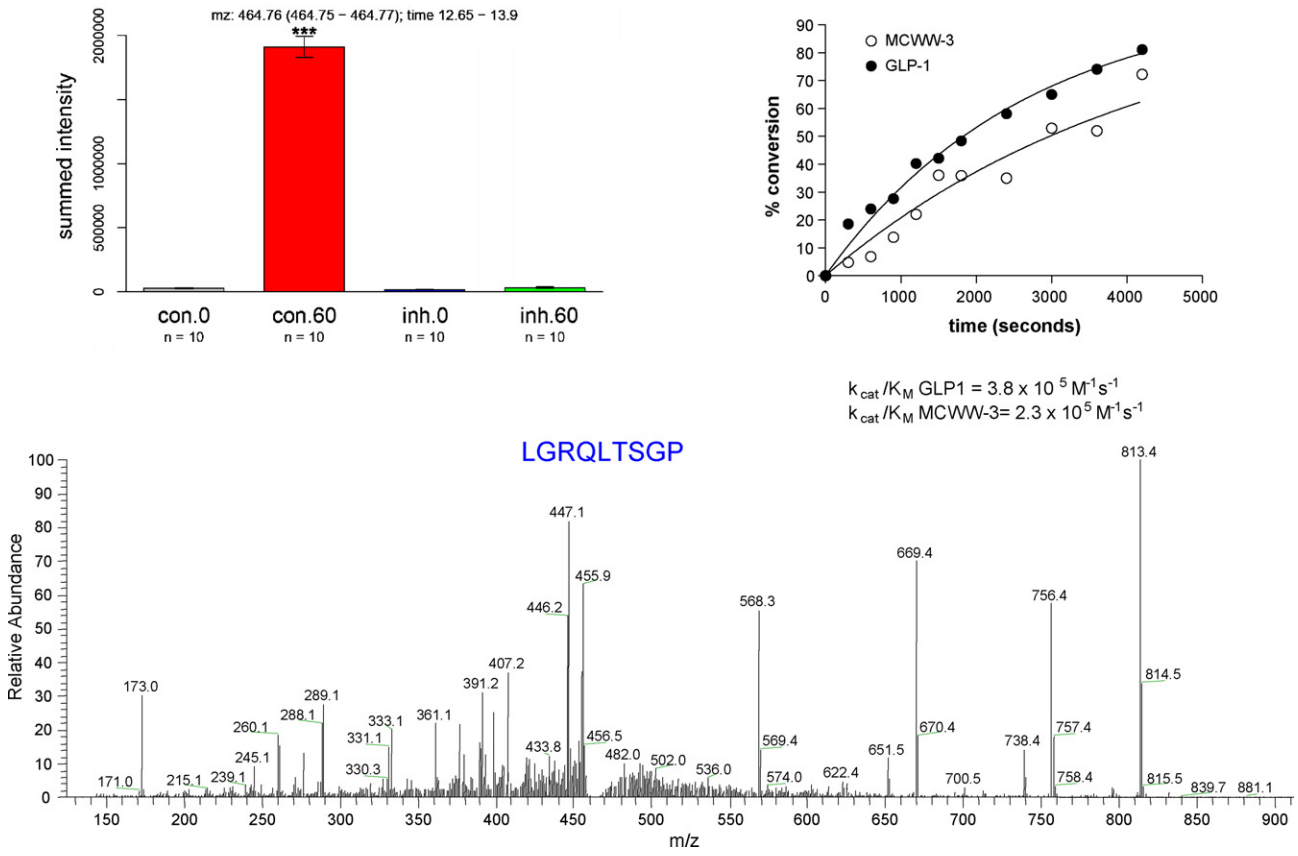

Fig. 5. Bar graph, MS/MS spectra, and k_{cat}/K_M curves for a novel DPP-4 substrate/product peptide EP/LGRQLTSGP identified in human plasma. Bar graph are for dMS features (*m/z* 464.76 @ 13.3) corresponding to the +2 charge state of the DPP-4 product peptide LGRQLTSGP. The grey, red, blue, and green colors correspond to the 0 min (−inhibitor), 60 min (−inhibitor), 0 min (+inhibitor), and 60 min (+inhibitor) treatment group, respectively. The amino acid sequesnce of the peptide was determined by manually interpreting the MS/MS spectrum using *de novo* methods. A synthetic peptide was prepared and used to determine k_{cat}/K_M.

In addition to detecting the turnover of validated DPP-4 substrates added into the samples as positive controls, the assay for DPP-4 resulted in the detection of 1093 grouped dMS features with only a few (∼1%) of the features showing characteristic substrate or product profiles (shown in Supplementary data table). Importantly, the 19th grouped feature exhibited a temporal profile corresponding to a candidate substrate and product of DPP-4-mediated proteolysis in the peptidic subfraction of human plasma. Fig. 5 illustrates a histogram corresponding to a dMS feature quantitated at 0 min (−inhibitor), 60 min (−inhibitor), 0 min (+inhibitor), and 60 min (+inhibitor), which exhibits the characteristic temporal profile of a candidate DPP-4 product peptide. The MS/MS spectra generated on the (M + 2H) + 2 ion at *m/z* 464.76 is also shown in the figure. The MS/MS data were manually interpreted to deduce the amino acid sequence of the product peptide, LGRQLTSGP. This amino acid sequence was confirmed by chemically synthesizing the peptide and demonstrating that its accurate *m/z* and experimentally determined MS/MS product spectra were identical to that of the dMS feature observed in the experiment. A BLAST [20] analysis identified this peptide as a fragment of α-2 plasmin inhibitor [21] a serpin that inhibits the activity of the plasma serine protease plasmin. Since DPP-4 is a P1-Pro/Ala-selective protease that releases N-terminal dipeptides from substrates, the corresponding intact substrate peptide was deduced to have the sequence EPLGRQLTSGP, which corresponds to the propeptide fragment of α-2 plasmin inhibitor [21]. Although DPP-4 cleavage of this peptide *in vivo* has yet to be demonstrated, it may serve as a convenient biomarker for plasma DPP-4 activity. We synthesized the peptide substrate and tested it for cleavage by recombinant soluble DPP-4 in an *in vitro* MS-based kinetics assay. As illustrated in Fig. 5, the peptide was rapidly processed by DPP-4, with cleavage efficiency (k_{cat}/K_M) similar to that of GLP-1, a known *in vivo* substrate of the enzyme.

In addition to the DPP-4 experiments using fractionated human plasma as the peptide pool, we also profiled aminopeptidase P2 using this methodology. APP2 has a well-defined site of cleavage at the amino terminus of peptides upstream of proline [16] and represents a peptidase for which no selective chemical inhibitors are available. dMS was performed and temporal profiles of dMS features were used to identify candidate substrates and products as described for DPP-4; dMS features representative of candidate substrate and product peptides show no changes in abundance for the negative control arm and a corresponding decrease or increase in relative concentration over time for the positive control arm.

The assay for APP2 resulted in the detection of 405 grouped dMS features with approximately 2% and 1% of the features showing characteristic substrate and product profiles, respectively. Background protease activity accounted for 8% of the

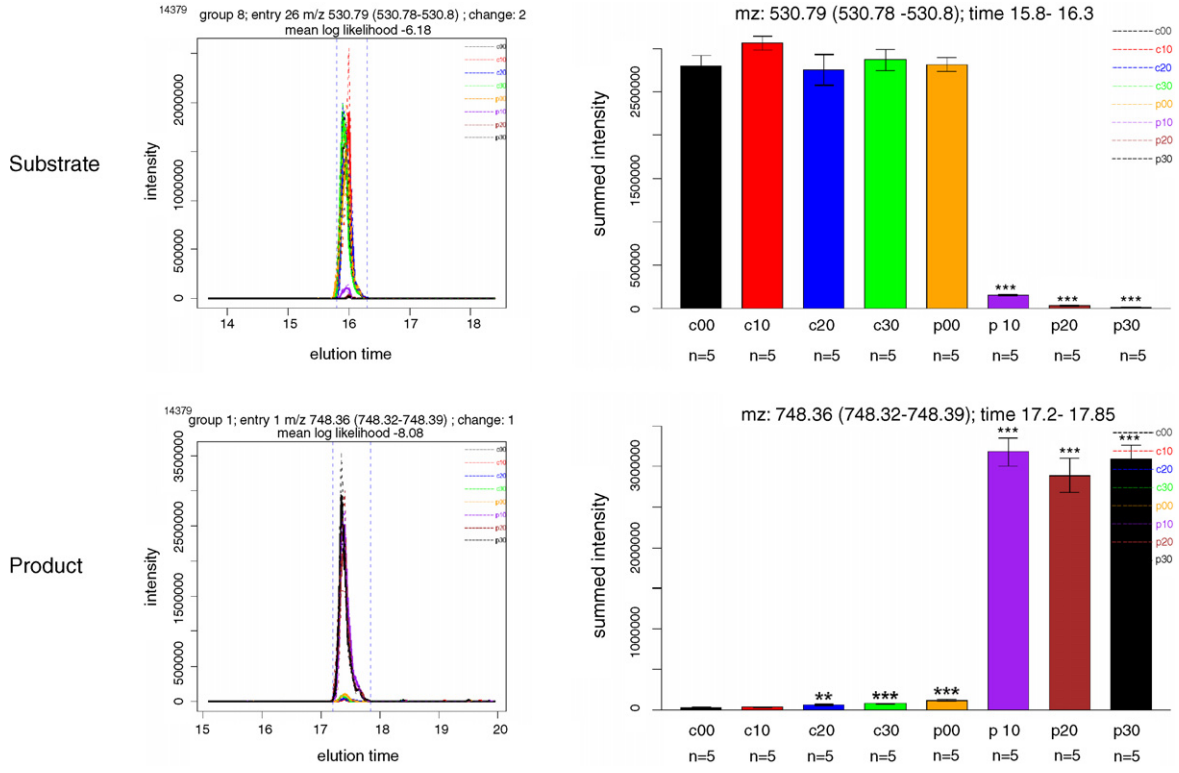

Fig. 6. Selected ion chromatograms and histograms for two dMS features (*m/z* 530.79 @ 16.1 min and *m/z* 748.36 @ 17.5 min) that exhibit characteristic substrate and product temporal profiles detected by incubating APP2 with a sub fraction of human plasma. The labels c00, c10, c20, c30, p00, p10, p20, and p30 correspond to the control (c) and peptidase (p) arms at 0, 10, 20, and 30 min time points.

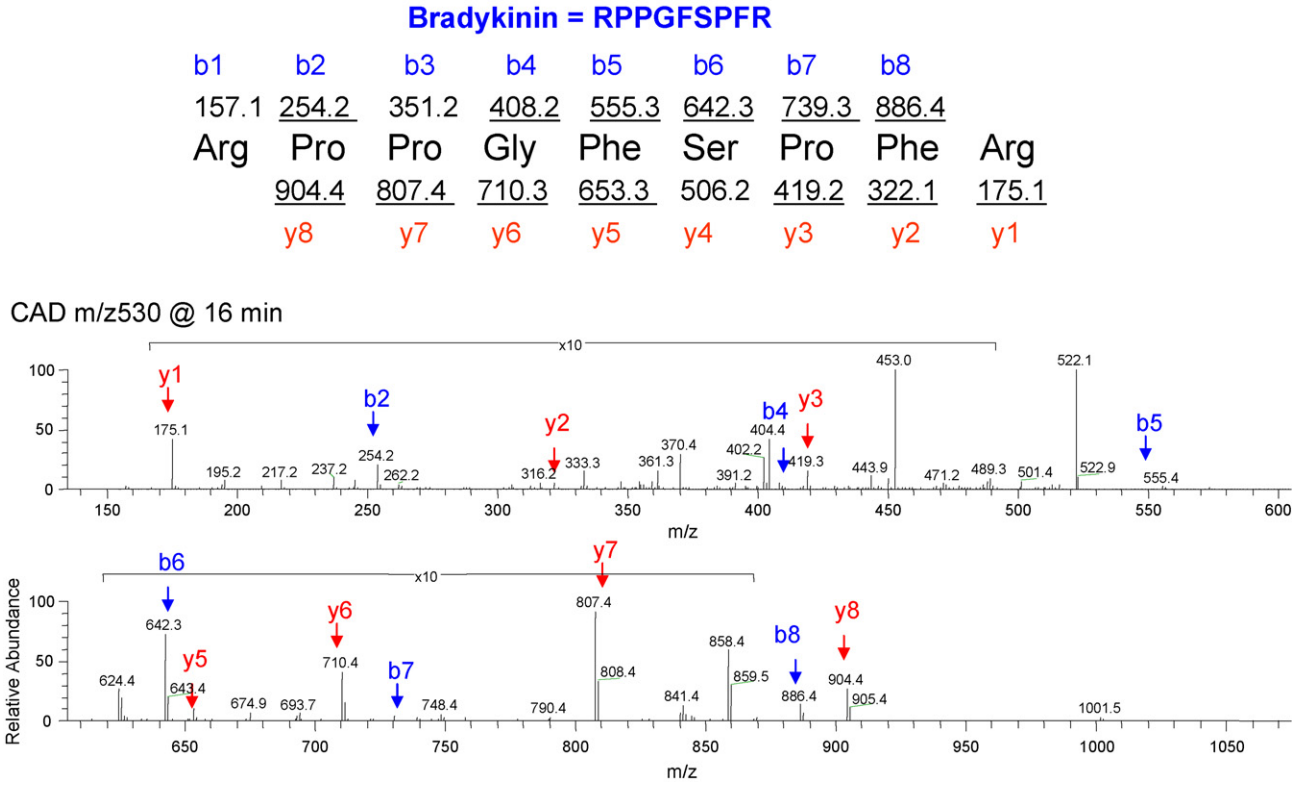

Fig. 7. Tandem mass spectrum and *de novo* sequence data obtained for a dMS substrate feature (*m/z* 530.79 @ 16.1 min) in the APP2 experiment, which was determined to be bradykinin.

Table 1
APP2 substrates and products identified in a sub fraction of human plasma by differential mass spectrometry

	Peptide sequence	Charge state	*m/z*		Mass accuracy (ppm)
			Theoretical	Experimental	
Substrates					
1	RPPGFSPFR	$M2H^+$	530.7880	530.7877	0.5
		$M3H^+$	354.1944	354.1943	0.2
2	RPPGYSPFR	$M2H^+$	538.7854	538.7855	0.1
		$M3H^+$	359.5260	359.5261	0.2
Products					
1	R.PPGFSPF.R	MH^+	748.3665	748.3666	0.1
		$M2H^+$	374.6869	374.6871	0.5
2	R.PPGYSPF.R	MH^+	764.3614	764.3618	0.5
		$M2H^+$	382.6843	382.6847	1.0
3	R.PPGFSPFR	MH^+	904.4676	904.4677	0.1
		$M2H^+$	452.7374	452.7375	0.2

dMS features and the remaining 89% of the dMS features could not be attributed to proteolysis occurring in a single arm of the study. Importantly, the most statistically significant dMS features exhibited temporal profiles expected for APP2 substrates and products. Fig. 6 shows the selected ion chromatograms and bar graphs for two dMS features (*m/z* 530.79 @ 16.1 min and *m/z* 748.36 @ 17.5 min) that exhibit characteristic substrate and product profiles, respectively. A common observation for experiments with pooled fractionated human plasma compared to proteomic profiling experiments involving comparative analysis of plasma samples from individual subjects is that the dMS features of the former are extremely pronounced and clear. This observation may be explained by the absence of biological variability between samples and the fact that peptide cleavage typically results in the quantitative turnover of one peptide species to another. Overall, the small number of substrate/product dMS features obtained for APP2 are consistent with what we would expect for a reasonably selective peptidase.

To identify the amino acid sequences of candidate APP2 substrate and product peptides found in human plasma, tandem mass spectrometry (MS/MS) of the appropriate ions was performed. Fig. 7 shows the MS/MS spectrum recorded for a substrate-like dMS feature at a retention time of 16 min and *m/z* of 530. Manual *de novo* analysis of this MS/MS spectrum resulted in the unambiguous identification of this peptide as bradykinin (RPPGFSPFR), a known substrate of APP2 present in plasma. The plasma level of bradykinin is approximately 10 pg/ml [16], but the sample corresponding to ultrafiltrate from 0.5 ml of plasma apparently contains high levels of this peptide, possibly due to degradation of high and low molecular weight kininogens during sample preparation. Table 1 contains the amino acid sequence, charge state, theoretical *m/z*, experimental *m/z*, and the accuracy of the *m/z* measurement expressed in parts-per-million for other APP2 substrates and products that were identified in this experiment. All of the substrate peptides appear to be related to bradykinin and satisfy the X-Pro N-terminus specificity requirement of APP2. APP2 substrates and products incorporating the sequence PPGYSPF were not found in the current human protein database, suggesting that the parent peptides or proteins may be novel. In addition, some of the product peptides show cleavage at both the N-termini (characteristic of APP2) and at the C-termini (due to background proteolysis).

4. Conclusion

We have developed an LC–MS based profiling assay that offers the potential to detect and identify endogenous peptidase substrates and products present in human plasma. By using dMS to analyze samples differing only in the presence or absence of a specific active protease of interest, the method allows the activity of the target protease to be differentiated from those of any background proteases. The limited sensitivity of the assay remains an issue for low abundance substrates. Hence, under these experimental conditions we were unable to detect picomolar endogenous substrates of DPP-4 such as GIP or GLP-1(7-36)-amide, even though the latter substrate was confirmed to be present in the plasma subfraction by ELISA (data not shown). Nevertheless, using LC–MS/MS and *de novo* sequencing we have demonstrated the ability to identify high abundance protease substrates from a complex human plasma peptide pool, which could provide useful biomarkers for monitoring target protease activities *in vivo*. For example, the peptide EPLGRQLTSGP corresponding to the propeptide domain of α-2 plasmin inhibitor was identified as a novel DPP-4 substrate in our experiments, and is turned over with a specificity constant of $2.3 \times 10^5\ M^{-1}\ s^{-1}$, comparable to that of GLP-1(7-36)-amide cleavage ($k_{cat}/K_M = 4.2 \times 10^5\ M^{-1}\ s^{-1}$). If cleavage of the propeptide by DPP-4 is demonstrated *in vivo*, then the peptide may serve as a convenient circulating high abundance biomarker for plasma DPP-4 activity, amenable to detection by MS without the need for developing specific antibodies and immunoassays.

New mass spectrometry technologies such as ORBI trap mass spectrometry [22] and higher field strength magnets for FTMS promise to improve the resolution of the LC–MS profiling data while maintaining dynamic range, hence further improving the

differential mass spectrometry when profiling complex samples. Although these experiments appear not to be limited by single scan dynamic range, the high resolution FTMS ($R = 50{,}000$) data allows us to resolve features which are degenerate in a 3D ion trap mass spectrometer. State-of-the-art fragmentation techniques such as infrared multi-photon dissociation [23], electron capture dissociation (ECD) [24] and electron transfer dissociation (ETD) [25,26] should improve the ability to sequence large non-tryptic peptides which are typically generated by peptidases and are resistant to fragmentation by CAD. Differential mass spectrometry may be used to understand in greater detail the substrate specificity of key proteases involved in pathologies such as Alzheimer's disease (gamma-secretase and BACE [27]) or aid in the effort towards understanding the biology of novel therapeutic targets such as PCSK9, which has been implicated in the regulation of LDL levels in humans [28].

In summary, differential mass spectrometry is a useful approach to detect and identify novel protease substrates and products in complex biological samples and this tool for protein biochemistry holds promise in the search for new therapeutic targets and markers of target engagement and disease.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijms.2006.09.020.

References

[1] Q.C. Meng, S. Oparil, J. Chromatogr. A 743 (1996) 105.
[2] M.S. Lantz, P. Ciborowski, Methods Enzymol. 235 (1994) 563.
[3] P.R. Sleath, R.C. Hendrickson, S.R. Kronheim, C.J. March, R.A. Black, J. Biol. Chem. 265 (1990) 14526.
[4] T.A. Rano, T. Timkey, E.P. Peterson, J. Rotonda, D.W. Nicholson, J.W. Becker, K.T. Chapman, N.A. Thornberry, Chem. Biol. 4 (1997) 149.
[5] M.V. di, D. Arnott, V. Barnaba, D.J. Loftus, K. Sakaguchi, C.B. Thompson, S. Salemi, C. Mastroianni, A. Sette, J. Shabanowitz, D.F. Hunt, E. Appella, J. Exp. Med. 183 (1996) 2509.
[6] G. Niedermann, G. King, S. Butz, U. Birsner, R. Grimm, J. Shabanowitz, D.F. Hunt, K. Eichmann, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 8572.
[7] M. Hackett, L. Guo, J. Shabanowitz, D.F. Hunt, E.L. Hewlett, Science 266 (1994) 433.
[8] M.W. Senko, C.L. Hendrickson, L. Pasa-Tolic, J.A. Marto, F.M. White, S. Guan, A.G. Marshall, Rapid Commun. Mass Spectrom. 10 (1996) 1824.
[9] J.E. Syka, J.A. Marto, D.L. Bai, S. Horning, M.W. Senko, J.C. Schwartz, B. Ueberheide, B. Garcia, S. Busby, T. Muratore, J. Shabanowitz, D.F. Hunt, J. Proteome. Res. 3 (2004) 621.
[10] S.L. Wu, J. Kim, W.S. Hancock, B. Karger, J. Proteome. Res. 4 (2005) 1155.
[11] G.A. Valaskovic, N.L. Kelleher, F.W. McLafferty, Science 273 (1996) 1199.
[12] W. Wang, H. Zhou, H. Lin, S. Roy, T.A. Shaler, L.R. Hill, S. Norton, P. Kumar, M. Anderle, C.H. Becker, Anal. Chem. 75 (2003) 4818.
[13] M.C. Wiener, J.R. Sachs, E.G. Deyanova, N.A. Yates, Anal. Chem. 76 (2004) 6085.
[14] H. Tammen, I. Schulte, R. Hess, C. Menzel, M. Kellmann, T. Mohring, P. Schulz-Knappe, Proteomics 5 (2005) 3414.
[15] D.J. Drucker, Cell Metab. 3 (2006) 153.
[16] U. Hilgenfeldt, T. Puschner, U. Riester, J. Finsterle, J. Hilgenfeldt, E. Ritz, Am. J. Physiol. 275 (1998) F88.
[17] L. Zhu, C. Tamvakopoulos, D. Xie, J. Dragovic, X. Shen, J.E. Fenyk-Melody, K. Schmidt, A. Bagchi, P.R. Griffin, N.A. Thornberry, R. Sinha Roy, J. Biol. Chem. 278 (2003) 22418.
[18] B. Leiting, K.D. Pryor, J.K. Wu, F. Marsilio, R.A. Patel, C.S. Craik, J.A. Ellman, R.T. Cummings, N.A. Thornberry, Biochem. J. 371 (2003) 525.
[19] G.R. Lankas, B. Leiting, R.S. Roy, G.J. Eiermann, M.G. Beconi, T. Biftu, C.C. Chan, S. Edmondson, W.P. Feeney, H. He, D.E. Ippolito, D. Kim, K.A. Lyons, H.O. Ok, R.A. Patel, A.N. Petrov, K.A. Pryor, X. Qian, L. Reigle, A. Woods, J.K. Wu, D. Zaller, X. Zhang, L. Zhu, A.E. Weber, N.A. Thornberry, Diabetes 54 (2005) 2988.
[20] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, J. Mol. Biol. 215 (1990) 403.
[21] W.E. Holmes, L. Nelles, H.R. Lijnen, D. Collen, J. Biol. Chem. 262 (1987) 1659.
[22] A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning, Anal. Chem. 78 (2006) 2113.
[23] D.P. Little, J.P. Speir, M.W. Senko, P.B. O'Connor, F.W. McLafferty, Anal. Chem. 66 (1994) 2809.
[24] R.A. Zubarev, D.M. Horn, E.K. Fridriksson, N.L. Kelleher, N.A. Kruger, M.A. Lewis, B.K. Carpenter, F.W. McLafferty, Anal. Chem. 72 (2000) 563.
[25] J.J. Coon, B. Ueberheide, J.E. Syka, D.D. Dryhurst, J. Ausio, J. Shabanowitz, D.F. Hunt, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 9463.
[26] J.J. Coon, J. Shabanowitz, D.F. Hunt, J.E. Syka, J. Am. Soc. Mass Spectrom. 16 (2005) 880.
[27] D.J. Selkoe, D. Schenk, Annu. Rev. Pharmacol. Toxicol. 43 (2003) 545.
[28] J. Cohen, A. Pertsemlidis, I.K. Kotowski, R. Graham, C.K. Garcia, H.H. Hobbs, Nat. Genet. 37 (2005) 161.