Milaim Pepaj[1]
Anders Holm[2]
Burkhard Fleckenstein[2]
Elsa Lundanes[1]
Tyge Greibrokk[1]

[1]Department of Chemistry, University of Oslo, Oslo, Norway
[2]Institute of Immunology, Rikshospitalet University Hospital, Oslo, Norway

Original Paper

Fractionation and separation of human salivary proteins by pH-gradient ion exchange and reversed phase chromatography coupled to mass spectrometry

In the present work, a 2-D capillary liquid chromatography method for fractionation and separation of human salivary proteins is demonstrated. Fractionation of proteins according to their p*I* values was performed in the 1-D employing a strong anion exchange (SAX) column subjected to a wide-range descending pH gradient. Polystyrene-divinylbenzene (PS-DVB) RP columns were used for focusing and subsequent separation of the proteins in the 2-D. The SAX column was presaturated with a high pH buffer (A) consisting of 10 mM amine buffering species, pH 9.0, and elution was performed with a low pH elution buffer (B) having the same buffer composition and concentration as buffer A, but pH 3.5. Isoelectric point fractions eluting from the 1-D column were trapped on PS-DVB trap columns prior to back-flushed elution onto the PS-DVB analytical column for separation of the proteins. The 1-D fraction eluting at pH 9.0–8.7 was chosen for further analysis. After separation on the RP analytical column, nine RP protein fractions were collected and tryptic digested for subsequent analyses by MALDI TOF MS and column switching capillary LC coupled to ESI TOF MS and ESI QTOF MS. Eight proteins and two peptides were identified in the pH 9.0–8.7 fraction using peptide mass fingerprinting and uninterpreted MS/MS data.

Keywords: pH-gradient ion exchange chromatography / p*I* / Proteins / Two dimensional capillary liquid chromatography

Received: September 17, 2005; revised: November 10, 2005; accepted: November 10, 2005

DOI 10.1002/jssc.200500346

1 Introduction

Human saliva performs a wide variety of biological functions that are critical for the maintenance of oral health [1]. Saliva is generally described as a heterogeneous fluid composed of proteins, glycoproteins, electrolytes, and small organic compounds [2]. There is growing interest in saliva as a diagnostic fluid due to its relatively simple and minimally invasive collection [3–5]. Recently, the whole saliva proteome has been studied using 2-D gel electrophoresis (2-D GE) to separate the protein components and MS to identify the peptides produced from the in-gel digests of the proteins of interest [6–8]. Although it has the greatest protein resolving power of any available separation technique, 2-D GE has some limitations such as difficulties encountered in observing large or small molecular weight proteins, highly acidic or basic proteins, and hydrophobic proteins. Due to these limitations and the recent advances in several types of MS, 2-D LC coupled to MS methods are being increasingly adopted for separation of complex protein and protein digest samples [9–14]. The unique feature of the 2-D methods based on LC compared to 2-D GE is the relative ease with which they can be interfaced to MS analysis, especially when capillary LC columns are used. The most widely used methods involve two orthogonal steps, with the second being RP LC directly coupled with MS. In these methods, RP LC is almost always preceded by IEC.

IEC is a versatile separation principle for proteins because of its applicability to several classes of proteins and its high capacity, not to forget the possibility of performing the separation under nearly non-denaturing conditions. Generally, salt-gradient IEC is used for separation of proteins. In comparison with that technique, however, pH-gradient IEC has two outstanding features: (i) The proteins are eluted roughly in order of their iso-

Correspondence: Milaim Pepaj, Department of Chemistry, University of Oslo, P.O. Box 1033 Blindern, 0315 Oslo, Norway.
E-mail: milaimp@kjemi.uio.no.
Fax: +47-22-84-40-77.

electric points (p*I*), and (ii) the proteins are focused in narrow bands during the separation, which give better peak shapes than are usually observed in salt-gradient IEC. By using a gradient pump to generate a pH gradient over the ion exchange column, the individual proteins will focus in narrow bands where the pH is more or less equal to their p*I* values, *i.e.* they are retained when pH > p*I* and non-retained when pH ≤ p*I*, and thus finally eluted in order of their p*I* values. In other words, pH-gradient IEC combines the characteristics of isoelectric focusing and the simplicity of chromatographic techniques to provide a separation technique capable of relatively high resolution of proteins. Additionally, the inherent focusing effect of pH-gradient IEC allows injection of large sample volumes [15].

Protein identification by MS can be performed using sequence-specific peptide fragmentation or peptide mass fingerprinting (PMF) [16]. The standard approach to identify proteins includes separation of proteins by 2-D GE or 2-D LC and digestion into peptides with trypsin. The resulting peptides are extracted and PMF spectra can then be obtained by using MALDI MS or ESI MS to measure the molecular masses of the peptides. The experimentally obtained masses are then compared with the theoretical peptide masses of proteins stored in databases by means of mass search programs. Although PMF is an effective and simple tool for the identification of relatively pure proteins, it often fails to identify proteins in mixtures. Consequently, sequence-specific peptide fragmentation, obtained by LC MS/MS or MALDI MS/MS, is necessary if no protein is identified unambiguously with PMF.

In the present work, an on-line 2-D capillary LC method using a wide pH-gradient IEC in the first dimension followed by RP in the second dimension is used to fractionate and separate human saliva proteins. In addition, the p*I* fraction eluting from the first dimension column at pH 9.0–8.7 was chosen for further analysis. Following second dimension separation on the RP analytical column, nine RP protein fractions were collected and tryptic digested into peptides for subsequent MS analysis by MALDI TOF and column switching capillary LC coupled to ESI TOF and ESI QTOF. PMF and MS/MS data were used for protein identification using MASCOT software to search protein databases.

2 Materials and methods

2.1 Chemicals and materials

The buffers piperazine, Tris, ethanolamine, diethanolamine, and imidazole were purchased from Sigma–Aldrich (Steinheim, Germany), while *N*-methylpiperazine was purchased from Acros Organics (Geel, Belgium). HPLC grade ACN was purchased from Rathburn Chemicals (Walkerburn, UK). HPLC gradient grade water, ethanol, and formic acid were obtained from Fluka (Buchs, Switzerland). HPLC gradient grade trifluoroacetic acid (TFA) for peptide and protein analysis was obtained from J.T. Baker (Phillipsburg, NJ, USA). Polyimide-coated fused silica capillaries were obtained from Polymicro Technologies (Phoenix, AZ, USA). The 1.0 mm ID × 15 cm strong anion exchanger PL-SAX column (10 μm, 1000 Å), the 0.3 mm ID × 15 cm analytical PLRP columns (5 μm, with pore size of 300 Å and 4000 Å), and the 1.0 mm ID × 5 mm PLRP Tracy trap columns (5 μm, with pore size of 300 Å and 4000 Å) were purchased from G&T SepTech (Kolbotn, Norway).

2.2 Preparation of buffers and saliva samples

A 500 mM stock solution of Tris, ethanolamine, diethanolamine, piperazine, *N*-methylpiperazine, and imidazole was prepared by dissolving appropriate amounts of each buffer in 500 mL of HPLC gradient grade water. The stock solution was filtered through 0.45-μm Minisart-plus filters (Sartorius, Göttingen, Germany) and stored at 4–5°C for no more than two weeks. Working buffer solutions were prepared by appropriate dilution of the stock solution with gradient grade water to desired buffer concentrations, and then titrated with 1.0 M formic acid to obtain the desired pH. The working buffer solutions were degassed with 99.998% helium (AGA, Oslo, Norway) before use. Whole saliva (approx. 5 mL) was obtained from a healthy non-smoking male subject in the morning at least ten hours after eating and after rinsing the mouth with water. In order to minimize the degradation of the proteins the sample was kept on ice during the collection process. Immediately after collection, the sample was centrifuged at 11 000 rpm for 15 min to remove insoluble material. The resulting supernatant was stored in 1 mL aliquots at −20°C. Prior to injection, the saliva samples were 5-fold diluted with start buffer to obtain a pH equal to that of the start buffer.

2.3 pH-gradient IEC

A non-splitting Hot Sep Micro-Pro binary gradient pump (Eldex Laboratories, Napa, CA, USA) was used to deliver the mobile phase. The 1.0 mm ID × 15 cm PL-SAX column was initially presaturated for 25 min with the start buffer (A) (pH 9.0) containing a mixture of 6 amines, *i.e.* piperazine, *N*-methylpiperazine, diethanolamine, ethanolamine, Tris, and imidazole at 10 mM concentrations. The elution step was performed with a low pH elution buffer (B), adjusted to pH 3.5 with formic acid, having the same amine composition and concentration as the start buffer. In order to enhance protein solubility, 10% of ethanol, was added to both start and elution buffer. Manual injections were performed using a 6-port Valco Cheminert C2 (Valco Instruments, Houston, TX, USA)

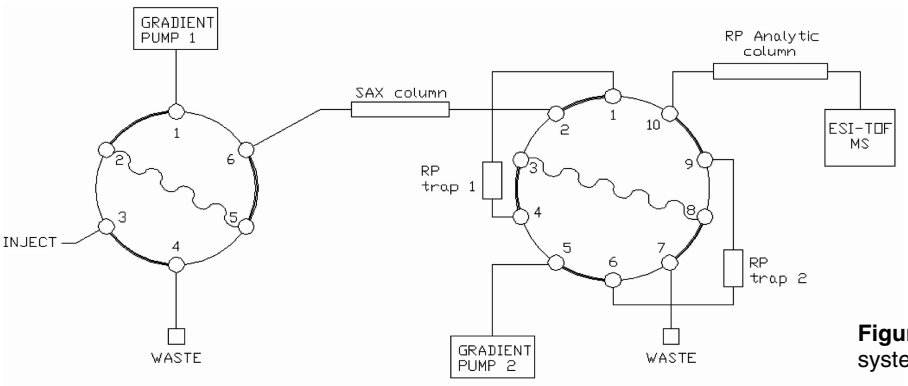

Figure 1. Schematic description of the 2-D system.

injection valve with a 100 μL external loop which was mounted inside the column oven operated at 25°C. The column oven was a Mistral 880 oven (Spark Holland Instrument, Emmen, The Netherlands). The first dimension profile was monitored at 280 nm using a Wellchrom K-2600 UV-detector (Knauer Wissenschaftlicher Gerätebau, Germany) equipped with a 44 nL capillary flow cell (75 μm capillary ID and 10 mm path length), while the pH-gradient was monitored using a low volume pH flow cell equipped with a round tip double junction flow-through pH electrode (Amersham Biosciences AB, Uppsala, Sweden). Approximately once a week, the pH electrode was detached from the housing and calibrated, while the housing was flushed with HPLC gradient grade water. The signal from the pH electrode was translated with a Cl-6507 pH amplifier and monitored with Science Workshop and Data Studio software (both from Pasco, Roseville, CA, USA). This software allowed simultaneous monitoring of the signal from the pH probe and the UV-detector.

2.4 pH-gradient IEC/RP capillary LC with UV detection

The on-line 2-D system (Fig. 1) used for the separation of proteins consisted of two non-splitting Hot Sep Micro-Pro binary gradient pumps. The valves used were a 6-port Valco Cheminert C2 injection valve with external loop for sample introduction, and a 10-port Valco Cheminert C2 switching valve to interface the two dimensions. All the components of the 2-D system were connected to the valves with 50 μm ID fused silica capillaries. One hundred μL of 5-fold diluted saliva was initially loaded onto a 1.0 mm ID × 15 cm SAX column (first dimension) and eluted fractions were trapped alternately on the one of the two 1.0 mm ID × 5 mm (4000 Å) trap columns. Back flushing from trap columns was also performed alternately with the same gradient as the one used during the separation on the 0.3 mm ID × 15 cm (4000 Å) analytical column (second dimension). Finally, on-line detection of the eluent from the RP analytical column separation was performed by UV at 280 nm. The mobile phases were (solvent A) HPLC gradient grade water/ACN (95:5) with 0.1% TFA and (solvent B) ACN/water (95:5) with 0.1% TFA at a flow rate of 15 μL/min. The gradient profile used for solvent B was as follows: 0–30% B in 3 min, 30–50% B in 10 min, 50–80% B in 2 min, and 80% B for 4 min followed by 7 min equilibration with 100% solvent A.

2.5 Second dimension separation of the pH 9.0–8.7 fraction and tryptic digestion of the collected RP fractions

The on-line 2-D system (Fig. 1) used to enrich and separate proteins present in the first dimension fraction eluting at pH 9.0–8.7 (time width of 10 min) was the same as the one used in Section 2.4, except that the 4000 Å trap and analytical reversed-phase columns were replaced with 300 Å columns. Enriched proteins were back-flushed from the trap column and subsequently separated on the analytical column using the same solvents as in Section 2.4 and a long gradient of solvent B with the following profile: 0–30% B in 8 min, 30–50% B in 28 min, 50–80% B in 9 min, and 80% B for 5 min. Protein fractions eluting from the RP analytical column were manually collected, based on peaks, in 1.5 mL polypropylene tubes for the purpose of tryptic digestion and subsequent MS analysis using MALDI TOF, ESI TOF, and ESI QTOF. Shortly after separation, fractions collected from the RP analytical column were evaporated under a gentle nitrogen stream and redissolved in 80 μL 50 mM ammonium bicarbonate buffer, pH approx. 7.8. Twenty μL of 0.2 μg/μL proteomics grade porcine trypsin (Sigma–Aldrich) was added to the 80 μL protein samples, making the final sample volume up to 100 μL. The mixtures were stored at a temperature of 37°C for 24 h.

2.6 MALDI TOF MS of tryptic peptides

Approximately one-mL volumes of the tryptic digests were purified over GELoader tips (Eppendorf) packed with the polystyrene-divinylbenzene (PS-DVB) based POROS R2 material from Applied Biosystems (Foster City, CA, USA). The retained peptides were washed twice with

30 μL of H_2O containing 0.1% TFA prior to elution onto a stainless steel target plate from Bruker Daltonics (Bremen, Germany) using ACN/H_2O (70:30) containing 10 mg/mL α-cyano-4-hydroxy-cinnamic acid (α-CHCA). The samples were allowed to crystallize at ambient temperature and pressure prior to analysis by an Ultraflex MALDI TOF MS instrument from Bruker Daltonics. The MALDI TOF MS was operated in the positive reflectron mode with a delayed ion extraction of 40 ns. Instrument calibration was performed using a tryptic digest of bovine serum albumin (BSA). The obtained spectra were processed using FlexAnalysis v2.2 and BioTools v2.2 from Bruker Daltonics and searched against the Swiss-Prot and NCBInr databases using MASCOT software from Matrix Science (London, UK).

2.7 Column switching capillary LC ESI TOF MS of tryptic peptides

Each RP fraction that was analyzed by MALDI TOF MS was also analyzed by column switching capillary LC ESI TOF MS. The column switching system used for enrichment and separation of tryptic peptides consisted of an LC isocratic loading pump (L-7100, Hitachi, Tokyo, Japan) and a non-splitting Eldex Micro-Pro binary gradient pump. The valves used were a 6-port Valco Cheminert C2 injection valve with external loop for sample introduction, and a 6-port Valco Cheminert C2 switching valve with a precolumn for sample concentration. Approx. 50 μL (from approx. 99 μL) of each tryptic digested fraction was injected using a 75 μL sample loop. Prior to injection, 5 μL of a 1.0% TFA solution was added to the tryptic peptides in order to enhance trapping of the peptides on the 1.0 mm ID × 5 mm (300 Å) trap column. The remaining (approx. 50 μL) sample was stored at −20°C for further ESI QTOF analyses. The loading mobile phase consisted of water and 0.1% TFA. Back flushing from the trap column was performed with the same gradient as was used during the separation on the 0.3 mm ID × 15 cm (300 Å) analytical column. The mobile phases used were (solvent A) HPLC gradient grade water with 0.1% formic acid and 0.01% TFA and (solvent B) ACN with 0.1% formic acid and 0.01% TFA at a flow rate of 5 μL/min. The gradient profile used for solvent B was as follows: 3–15% B in 3 min, 15–30% B in 15 min, 30–70% B in 5 min, and 70% B for 4 min followed by 15 min equilibration with 3% solvent B. Peptide analysis were performed with a Micromass LCT TOF MS (Micromass, Manchester, UK). The TOF MS was equipped with a Z-spray atmospheric pressure ionization source for ESI, which was modified to handle flow rates in the low μL/min range. Ionization was performed in positive mode. The following voltages were used: 3.0 kV on the capillary, 17 V on the sample cone, and 3 V on the extraction cone. In order to obtain a stable spray performance and aid solvent vaporization, nebulizer gas and desolvation gas were applied at 50 L/h and 350 L/h, respectively. The TOF MS instrument was controlled by MassLynx v4.0 software, and mass spectra were acquired in the *m/z* range 200–2000. The molecular weights (M_r) of the tryptic fragments were measured as monoisotopic masses. Doubly and triply charged peptides were manually deconvoluted to obtain singly protonated molecular ions $[MH]^+$ for database searching.

2.8 Column switching capillary LC ESI QTOF MS of tryptic peptides

Each RP fraction that was analyzed by MALDI TOF MS and ESI TOF MS was also analyzed by column switching capillary LC ESI QTOF MS. The column switching system and LC conditions were the same as in Section 2.7. Sequence information by MS/MS was obtained with the QTOF (Micromass) operated in the positive mode with a capillary voltage of 3.0 kV, a cone voltage of 17 V, and 3 V on the extraction cone. The source and desolvation temperatures were set to 100°C and 120°C, respectively. Mass spectra were obtained over a *m/z* ratio of 200–2000 and the resolution was typically within the range of 5000–10000. 0.5-second positive MS survey scans were performed to monitor the separation until MS/MS was triggered when the intensity of the observed peptides exceeded a predefined threshold value of 10 counts/s. MS/MS acquisition was completed and switched back to survey scan when the precursor intensity fell below a predefined threshold (5 counts/s). Argon was employed as collision gas and collision energies used were in the range of 20 to 60 eV. The QTOF MS instrument was controlled by MassLynx v4.0 software, and MS/MS data files appropriate for MASCOT (SEQUEST dta files) were created using BioLynx software that smoothes and centroids the data.

2.9 Protein identification

Protein identification was accomplished using MASCOT software from Matrix Science. PMF was performed using experimental monoisotopic masses from both MALDI TOF MS and ESI TOF MS experiments to search the Swiss-Prot or NCBInr protein databases. Most of the MASCOT search parameters were held constant for all of the data generated in this study. These include the variable modification, *i.e.* oxidation of methionine, cleavage by trypsin, one missed cleavage, *Homo sapiens* for the species, and no restriction on the p*I* value or molecular mass. For MALDI TOF MS and ESI TOF MS, the peptide mass tolerance was set to 0.1 Da. For the MS/MS data obtained with ESI QTOF, ESI QTOF was chosen as the instrument type, and the peptide and fragment mass tolerances were set to 0.5 Da. Positive protein identification was based on standard MASCOT criteria for statistical analysis of LC MS/MS data and PMF.

3 Results and discussion

The main goal of this study was to explore the capabilities of the 2-D capillary LC system for fast fractionation and separation of saliva proteins using pH-gradient IEC in the first dimension and reversed-phase LC in the second dimension. The second goal was to explore the use of different MS techniques for identification of saliva proteins present in the first dimension fraction eluting at pH 9.0–8.7. The selected fraction was enriched on the RP trap column and separated into fractions on the second dimension RP analytical column with subsequent tryptic digestion and MALDI TOF MS, LC ESI TOF MS, and LC ESI QTOF MS analysis. PMF and uninterpreted MS/MS spectra were used in the identification process by searching the Swiss-Prot and NCBInr protein databases using MASCOT software.

3.1 Fractionation and separation of saliva proteins by 2-D capillary LC with UV-detection using large loading volumes

Fractionation techniques such as salt-gradient IEC and chromatofocusing (CF) are most widely used as first dimension in a 2-D LC system for protein separation. While CF generally exploits the buffering capacity of the ion exchanger to obtain a retained intra-column pH gradient, the opposite is true for pH-gradient IEC [17]. A wide-range pH gradient using a mixture of 6 amine buffering species at 10 mM concentrations and 10% ethanol in both start and elution buffer was employed to fractionate saliva proteins into three p*I* fractions in the first dimension as shown in Fig. 2a. The 1.0 mm ID × 15 cm SAX column was initially presaturated with the start buffer (A) at pH 9.0 prior to elution with elution buffer (B) at pH 3.5 using a linear gradient from 0–100% B in 45 minutes. The start and elution buffer are composed of low molecular weight amine buffer components having pK_a values which are approximately equally spaced throughout the gradient pH range. The use of wide-range pH gradients offers some advantages compared to the more common narrow pH gradients. The former allows fractionation of highly basic or acidic proteins in p*I* fractions during the same run. Moreover, when the p*I* fraction of interest is isolated, the remaining p*I* fractions or highly acidic proteins can be washed from the column with the low pH elution buffer without any additional washing step. This is particularly important when dealing with complex samples containing both highly acidic and basic proteins, such as saliva.

An important aspect of any multidimensional separation system is its ability to improve the detection of analytes present in low quantities during the analyses of complex samples. The inherent focusing effect of pH-gradient IEC, which allows injection of large sample volumes to facilitate detection of low abundance proteins, was demonstrated in our previous work [15]. Consequently, 100 μL diluted saliva samples were loaded onto the first dimension SAX column. Thus, the use of pH-gradient IEC in the first dimension provided both sample concentration and p*I* fractionation. Additionally, a larger ID (1.0 mm) col-

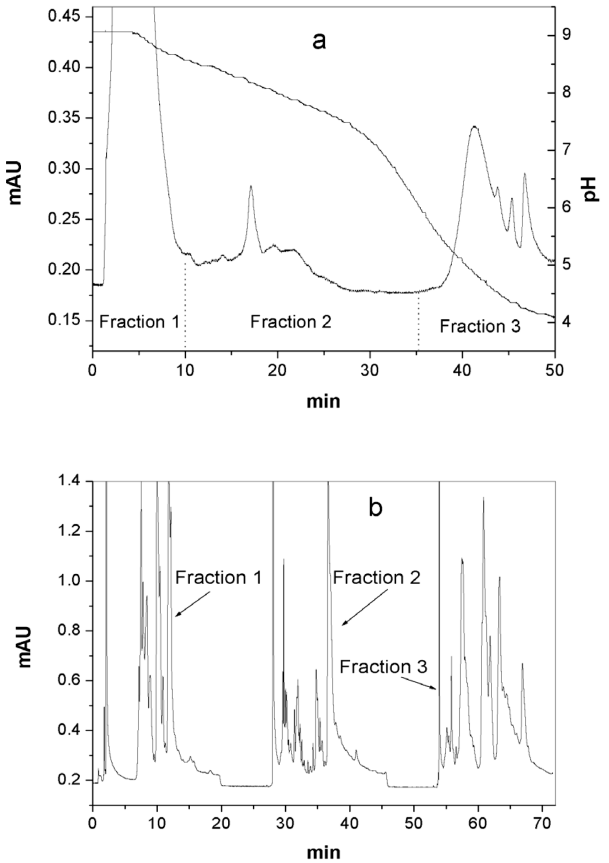

Figure 2. a) 1-D fractionation of 5-fold diluted saliva in three p*I* fractions (time fractions 1: 10 min, 2: 26 min, and 3: 14 min) by pH-gradient IEC on a 1.0 mm ID × 15 cm PL-SAX column. The start buffer (A) consisted of 10 mM piperazine, *N*-methylpiperazine, imidazole, Tris, diethanolamine, and ethanolamine and ethanol (90:10)) (pH 9.0), and the elution buffer (B) was the same as start buffer, but of pH 3.5. The gradient program was as follows: 0–100% B in 45 min, then 100% B in 10 min. The injection volume was 100 μL. The flow rate was 50 μL/min and the detection was performed at 280 nm. **b)** 2-D RP separation of the three p*I* fractions on a 0.3 mm ID × 15 cm 4000 Å PS-DVB analytical column after fractionation (1-D), trapping of two fractions on 1.0 mm ID × 5 mm 4000 Å PS-DVB trap columns, and back-flushing to the analytical column. The elution was performed as follows: mobile phase A consisted of HPLC gradient grade water/ACN (95 : 5) with 0.1% TFA and mobile phase B consisted of ACN/water (95 : 5) with 0.1% TFA. The gradient profile used for solvent B was as follows: 0–30% B in 3 min, 30–50% B in 10 min, 50–80% B in 2 min, and 80% B for 4 min followed by 7 min equilibration with 0% solvent B. The flow rate was 15 μL/min and the UV-detection was performed at 280 nm.

umn in the first dimension not only increased the overall loading capacity of the 2-D system, but also provided enhanced mass sensitivity when a smaller ID (0.3 mm) column was used in the second dimension coupled to a concentration sensitive detection method, such as UV or ESI. Initially, three time fractions eluting from the first dimension SAX column were enriched on trap columns, back flushed from the trap columns, and separated on a 0.3 mm ID × 15 cm (4000 Å) polystyrene-divinylbenzene (PS-DVB) analytical column using a fast RP gradient as shown in Fig. 2b. The first (basic) fraction eluted in the pH range from 9.0 to 8.7, *i.e.* 10 min fraction (Fig. 2a), and was chosen for further analysis. The reversed-phase separation of this fraction obtained with the wide pore (4000 Å) analytical column in the second dimension was, however, not sufficient to permit collection of pure intact proteins for subsequent tryptic digestion and MS analysis. Ideally, the collected fractions should contain as few proteins as possible prior to tryptic digestion, when MS analyses are combined with PMF. Hence, several gradients of solvent B were performed in order to improve the separation and to facilitate fraction collection. It was observed that fast gradients resulted in improved separation compared to slow gradients, indicating that the low capacity wide pore (4000 Å) column appears to be better suited for separation of less complex protein mixtures with fast gradients. Consequently, other columns were examined, and it was found that the smaller pore size (300 Å) PS-DVB column provided the best separation when long elution gradients were used. Although the best separation was obtained using a 50 min long gradient (Fig. 3), the column was not able to baseline separate all the proteins and coelution of proteins was inevitable. Even longer gradients were also examined, resulting in poor peak shapes and a significant increase in peak broadening. Thus, the 50 min gradient was chosen and nine fractions based on peaks (Fig. 3) were manually collected for the purpose of tryptic digestion.

In addition to its practical feasibility, the current on-line 2-D protein separation platform also has some other prominent features: (a) the inherent focusing effect of the pH-gradient IEC in the first dimension enables the enrichment of low-level proteins by large volume injections; (b) the capability for selectively peeling off selected p*I* fractions by controlling the pH slope of the elution buffer provides a method for readily interfacing pH-gradient IEC to a second dimension for separation, and (c) the use of a wide pH gradient allows the fractionation of both highly basic and acidic proteins in p*I* fractions during the same run. In its present form, the on-line 2-D platform is highly suitable for fractionation and separation of less complex samples or for the enrichment and separation of few fractions eluting from the first dimension.

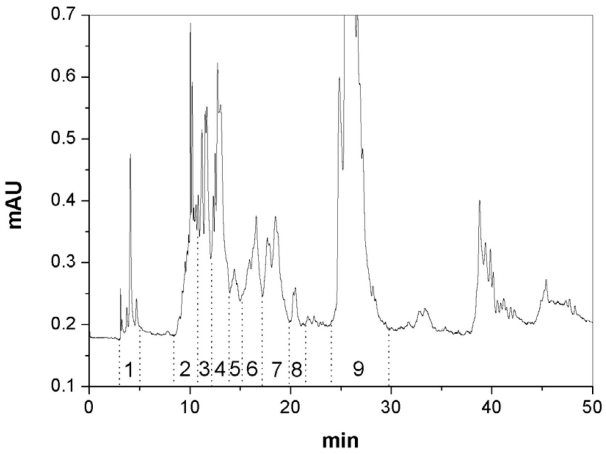

Figure 3. 2-D RP separation of the 1-D fraction eluting at pH 9.0–8.7. The elution conditions were the same as in Fig. 2b, except for the gradient profile, which was as follows: 0–30% B in 8 min, 30–50% B in 28 min, and 50–80% B in 9 min and 80% B for 5 min. Nine RP fractions (1–9) based on peaks were manually collected.

However, when dealing with complex samples containing hundreds of proteins and peptides, the number of fractions eluting from the first dimension must be increased. This is, however, determined by the separation speed of the second dimension and the number of trap columns used in the 2-D platform. An increase in the number of fractions in the first dimension will significantly decrease the complexity of the sample reinjected in the second dimension, which in turn will facilitate protein separation, fraction collection, and identification. This can be performed on-line or off-line. In the on-line approach, a minimum of 10 trap columns are needed to enrich fractions eluting from the first dimension, and the enrichment process must be independent of the separation speed in the second dimension. In the off-line approach, as many fractions as possible can be collected, and, what is very important, they can be analyzed at any time independently of the second dimension. Thus, our emphasis in future work will be directed towards exploring the use of both on-line 2-D systems with 10 trap columns and off-line 2-D systems.

3.2 Identification of proteins by MS

The proteins present in the first dimension fraction eluting at pH 9.0–8.7 were separated in the second dimension and nine reversed-phase (RP) fractions based on peaks were manually collected. The collected RP fractions were tryptic digested and analyzed by MALDI TOF MS, column switching capillary LC ESI TOF MS, and column switching capillary LC ESI QTOF MS. Each RP fraction was analyzed by all three different MS techniques.

3.2.1 MALDI TOF MS

MALDI MS is known to be the most commonly used technique for performing PMF. MALDI TOF MS is fast, robust, easy to perform, sensitive (low fmol range), and accurate (low ppm range). The predominant detection of singly charged peptides by MALDI MS greatly facilitates the evaluation of PMFs. Although PMF is relatively simple and an effective tool for the identification of relatively pure proteins, it often fails to identify proteins in mixtures. Slightly more complicated protein mixtures containing two or three proteins might also be reliably analyzed, but the method is not suitable for more complex samples. As discussed in Section 3.1, the fractions analyzed contained more proteins. This was also confirmed by the mass spectra obtained by MALDI TOF MS. Despite that, seven proteins were unambiguously identified (Table 1) using MALDI TOF MS when tryptic peptide masses of each RP protein fraction were submitted to Swiss-Prot or NCBInr using MASCOT. Fig. 4 shows the MALDI TOF MS analysis of peptide fragments digested from RP fraction 9. Most of the abundant peaks (labeled with an asterisk) were matched with the partial sequence of salivary α-amylase. Proteins identified in each RP fraction, using MALDI TOF MS, are summarized in Table 1. Three proteins, immunoglobulin kappa chain (fraction 2), PRO0433 and zinc finger protein 40 (both in fraction 7) were identified only by MALDI, while immunoglobulin heavy chain variable region (fraction 3), cystatin D (fraction 4), cystatin SN (fraction 6), and α-amylase (fraction 9) were identified by both MALDI and ESI. In addition, cystatin SN and α-amylase were also identified by ESI QTOF MS. In all MALDI cases, more than 3 peptide masses matched, and more than 36% sequence coverage

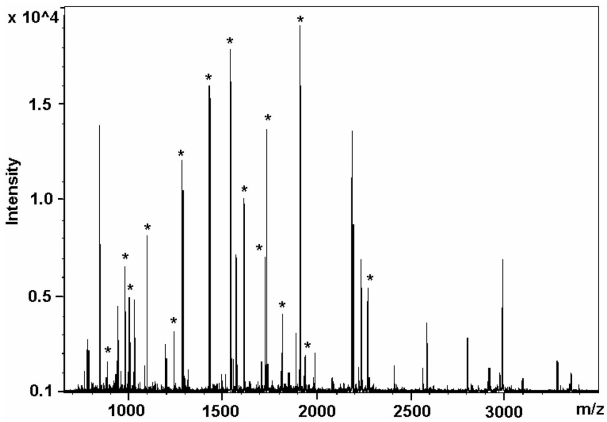

Figure 4. MALDI TOF MS analysis of peptide fragments digested from RP fraction 9. Most of the abundant peaks (labeled with an asterisk) were matched with the partial sequence of salivary α-amylase.

was achieved, with immunoglobulin heavy chain (variable region) having the highest sequence coverage of 72%, and α-amylase having the highest number of matched peptides (18). Furthermore, the actual difference between the molecular mass of the experimental peptide and the corresponding theoretical peptide (expressed in ppm), which is a useful guide when used in conjunction with sequence coverage, was less than 54 ppm.

3.2.2 Column switching capillary LC ESI TOF MS

In order to facilitate detection of low abundancy peptides resulting from low abundancy proteins, sample enrichment has to be performed by injecting large

Table 1. Salivary proteins/peptides identified by PMF and MS/MS in the 1-D fraction eluting at pH 9.0–8.7.

RP fraction	Accession number	Protein/peptide name	M_r	p*I*	Matched peptides	Sequence coverage (%)	MASCOT score
1	gi|229138	Kallidin II (peptide)	1188	12.0	1[c)]	90[c)]	36[c)*)]
	gi|350217	Peptide PA, saliva	3796	8.59	2[b)], 1[c)]	100[b)]	73[b)], 46[c)*)]
2	gi|33570054	Immunoglobulin kappa chain variable region	9833	7.71	3[a)]	58[a)]	69[a)]
3	gi|56675804	Immunoglobulin heavy chain variable region	8535	9.46	5[a)], 3[b)]	72[a)], 41[b)]	91[a)], 78[b)]
	gi|21311281	Immunoglobulin heavy chain	15475	9.03	3[b)]	24[b)]	71[b)]
4	P28325	Cystatin D	16123	7.63	5[a)], 3[b)]	46[a)], 28[b)]	87[a)], 72[b)]
6	P01037	Cystatin SN precursor	16351	6.82	4[a)], 3[b)], 1[c)]	43[a)], 31[b)]	81[a)], 75[b)], 118[c*)]
7	gi|14189954	PRO0433	11019	9.76	4[a)]	51[a)]	74[a)]
	gi|45160100	Zinc finger protein 40	21245	9.52	7[a)]	36[a)]	80[a)]
9	P04745	α-amylase, salivary precursor	57731	6.47	18[a)],13[b)],1[c)]	37[a)], 22[b)]	95[a)],125[b)], 90[c*)]

a) MALDI TOF MS.
b) ESI TOF MS.
c) QTOF MS/MS.
* Ion score.

volumes. By using a column switching system, fast enrichment of the peptides is achieved as the peptides are focused on a short pre-column prior to back-flushed elution onto the analytical column. Consequently, the peptides were efficiently focused on the pre-column using a non-eluting loading solvent consisting of water and 0.1% TFA at 100 μL/min allowing injection of the total sample volume (approx. 50 μL) of the tryptic peptides. The optimal ESI and LC conditions do not coincide and compromises between the optimum separation and detection are required. For optimum separation, 0.1% TFA is commonly added to the mobile phase in RP LC of peptides and proteins. Since it is well known that the combination of TFA and ESI MS results in signal reduction and spray instability [18–20], a combination of formic acid (0.1%) and TFA (0.01%) was used as additive in the mobile phase. Several different solvent gradients were evaluated in order to optimize the separation of the tryptic peptides. However, coelution of the peptides was inevitable due to the complexity of the sample arising from the high number of peptides generated by tryptic digestion, and thus indicating the presence of more than one protein in the collected RP fractions. This is illustrated in Fig. 5, which shows the total ion chromatograms (TIC) obtained from three selected tryptic digested RP fractions. When triply and doubly charged tryptic peptides were deconvoluted to singly charged peptides and searched against protein databases, 5 proteins were unambiguously identified (Table 1) using PMF. Proteins identified in each RP fraction using ESI TOF MS are summarized in Table 1. In all cases, more than 3 peptide masses matched, and more than 22% sequence coverage was achieved, with immunoglobulin heavy chain (variable region) having the highest sequence coverage of 31% and α-amylase having the highest number of matched peptides (13). The mass accuracy was within 65 ppm, which is slightly less than that obtained with MALDI. Additionally, the sequence coverage and the number of matched peptides were less than those obtained with MALDI (Table 1). The presence of several proteins in the analyzed fractions significantly complicates the PMF identification process, as discussed in the MALDI MS section. Post-translational modification (PTM) of peptide residues is another prominent parameter which complicates the PMF identification process. However, manual inspection of mass spectra searching for PTM was beyond the scope of this work.

3.2.3 Column switching capillary LC ESI QTOF MS

In addition to MALDI TOF MS and ESI TOF MS analyses, each RP digested fraction was also analyzed by ESI QTOF MS with collision-induced dissociation (CID) to generate peptide fragmentation. Protein identification is achieved by searching the protein databases with product ions

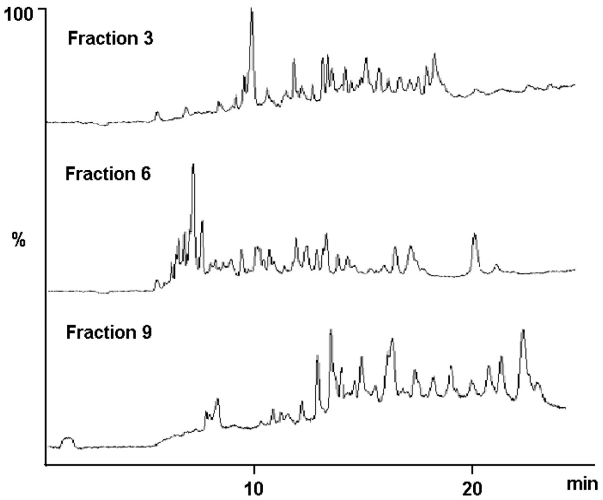

Figure 5. Total ion chromatogram (TIC) of tryptic peptides from the RP fractions 3, 6, and 9 obtained with column switching capillary LC ESI TOF MS. Following back-flushing from the 1.0 mm ID × 5 mm (300 Å) trap column, the tryptic peptides were separated on a 0.3 mm ID × 15 cm (300 Å) analytical column. The mobile phases used were (solvent A) HPLC gradient grade water with 0.1% formic acid and 0.01% TFA and (solvent B) ACN with 0.1% formic acid and 0.01% TFA with a flow rate of 5 μL/min. The gradient profile used for solvent B was as follows: 3–15% B in 3 min, 15–30% B in 15 min, 30–70% B in 5 min, and 70% B for 4 min followed by 15 min equilibration with 3% solvent B.

combined with the original peptide mass. The partial sequence information obtained from MS/MS has more discriminating power than PMF and may allow the identification of proteins based on a single peptide. When protein databases were searched with uninterpreted MS/MS spectra of peptide fragments, only 2 peptides and 2 proteins were identified (Table 1). The most likely explanation for this is that the MS/MS spectrum obtained (for the different peptides) is of insufficient quality for matching a sequence database. A possible reason for the poor quality of the MS/MS spectrum is fragmentation of intact peptide ions in the ion source prior to the first stage of mass analysis. These in-source fragment ions can trigger a data-dependent MS/MS spectrum, but will not be identified when searching for peptides derived solely from tryptic cleavages. Another possible explanation for the low number of identified proteins could be that a number of precursor peptides were not selected for CID due to coelution of peptides. Nevertheless, the significant probability MASCOT ion score obtained for the identified peptides and proteins using uninterpreted MS/MS spectra was higher than 36. The doubly charged peptide (*m/z* 958.10) derived from cystatin SN precursor had the highest ion score of 118. The MS spectra of this doubly charged peptide obtained from fraction 6, is shown in Fig. 6. Panel A shows the precursor ion of this doubly-charged peptide, while panel B shows the MS/MS spec-

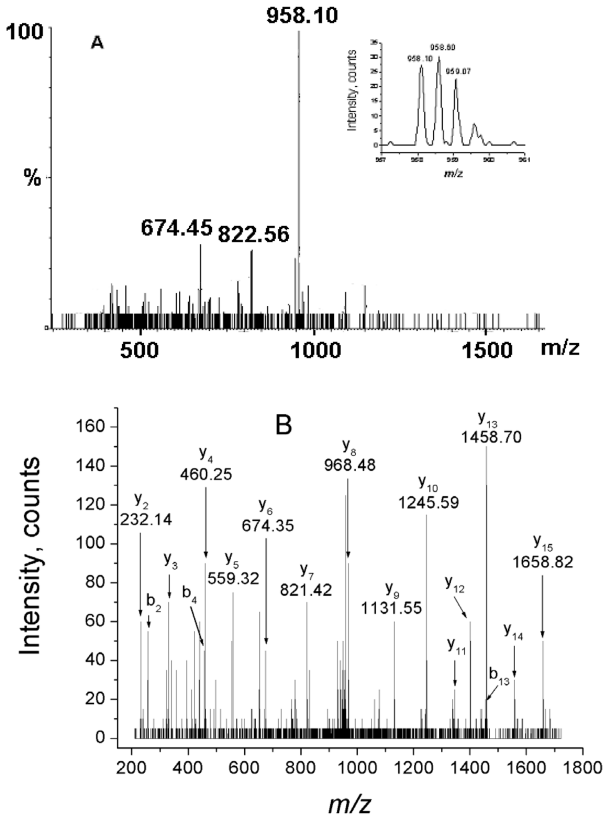

Figure 6. Column switching capillary LC ESI QTOF mass spectrum (A) and MS/MS (B) of a doubly charged tryptic peptide, QQTVGGVNYFFDVEVGR (*m/z* 958.10), originated from salivary cystatin SN precursor.

trum. The MS/MS spectra represent the amino acid sequence of a tryptic peptide, QQTVGGVNYFFDVEVGR. The spectra shown in Fig. 6 exemplify that unambiguous protein identification using uninterpreted MS/MS spectra relies totally on the quality of the MS/MS spectrum.

As already mentioned, the primary goal of this work was to develop a fast and highly practical 2-D platform where proteins elute roughly in the order of their p*I* values in the first dimension. The fraction analyzed eluted in the pH range of 9.0–8.7, and thus should ideally contain proteins with a p*I* value higher than 8.7. This was in good agreement with four identified proteins and two peptides with theoretical p*I* values higher than 8.7 (Table 1). However, four other proteins with theoretical p*I* values below 8.7, *i.e.* alpha-amylase (pI_{theo} 6.47), cystatin SN (pI_{theo} 6.82), cystatin D (pI_{theo} 7.63), and immunoglobulin kappa chain (pI_{theo} 7.71) also eluted in the same fraction. There are most likely several reasons for such behavior. For example, during chromatofocusing (CF), only part of a protein binds to the stationary phase through specific areas of the protein or individual charge sites. Therefore, elution of proteins at pH values higher than their isoelectric points (p*I*) is commonly observed [21]. Most likely the same phenomena will also occur during pH-gradient IEC elution. Another reason may be found in the presence of protein isoforms and post transitional modification (PTM) of the proteins. Regarding α-amylase, the PTM phenomenon was observed by Loo *et al.* [22] in 2-D gels. They observed that α-amylase gave a horizontal streaking in a p*I* range of 7.8–5.2.

The first dimension basic fraction analyzed contained 8 identified proteins and two peptides. Yet a number of known basic salivary proteins (*e.g.* cystatin C, histatins, lysozyme, basic proline rich proteins, *etc.*) were not identified. Possible reasons for this could be that: (i) the large amounts of α-amylase (highly abundant protein in saliva) complicated the identification of certain low abundance proteins, such as lysozyme; (ii) the basic proline rich proteins were not detected due to their primary sequence. The proline rich regions have few typical tryptic cleavage sites, and trypsin cleavages at unusual sites have been reported in previous studies using Edman sequencing [23, 24]. Moreover, (iii) PMF, which was the primary identification tool, is not an appropriate identification method for protein mixtures, and (iv) the quality of MS/MS spectra obtained was insufficient to match sequences in databases.

4 Concluding remarks

A wide-range pH-gradient IEC coupled on-line with PS-DVB-RP has been used to fractionate (first dimension) and separate (second dimension) salivary proteins. The 2-D platform used is a fast and highly practical platform where proteins elute roughly in the order of their p*I* values in the first dimension. The inherent focusing effect of the pH-gradient IEC enables the enrichment of low-level proteins by loading large sample volumes directly onto the column. Moreover, isolated p*I* fractions are readily transferred on-line to the RP analytical column via trap columns. Thus, pH-gradient IEC has the potential of becoming an important fractionation technique in proteomic studies, not only for p*I*-fractionation and enrichment of weakly expressed proteins by injecting large volumes, but also as a highly efficient reversed-phase compatible dimension in 2-D LC systems. It is not likely that the present 2-D LC method will be able to provide the same high resolution of proteins as 2-D GE, but 2-D LC represents a much faster and more easily automated method.

The project has been supported by the EU research training network HPRN-CT-2001-00180 (COM-CHROM). The authors gratefully acknowledge John Vedde for help with the ESI QTOF MS experiments.

5 References

[1] Edgar, W. M., *Br. Dent. J.* 1992, *172*, 305 – 312.

[2] Humphrey, S. P., Williamson, R. T., *J. Prosthet. Dent*. 2001, *85*, 162 – 169.

[3] Hofman, L. F., *J. Nutr*. 2001, *131*, 1621S – 1625S.

[4] Lawrence, H. P., *J. Can. Dent. Assoc*. 2002, *68*, 170 – 174.

[5] Kaufman, E., Lamster, I. B., *Crit. Rev. Oral Biol. Med*. 2002, *13*, 197 – 212.

[6] Yao, Y., Berg, E. A., Costello, C. E., Troxler, R. F., Oppenheim, F. G., *J. Biol. Chem*. 2003, *278*, 5300 – 5308.

[7] Ghafouri, B., Tagesson, C., Lindahl, M., *Proteomics* 2003, *3*, 1003 – 1015.

[8] Vitorino, R., Lobo, M. J., Ferrer-Correira, A. J., Dubin, J. R., Tomer, K. B., Domingues, P. M., Amado, F. M., *Proteomics* 2004, *4*, 1109 – 1115.

[9] Opiteck, G. J., Jorgenson, J. W., *Anal. Chem*. 1997, *69*, 2283 – 2291.

[10] Wilmarth, P. A., Riviere, M. A., Rustvold, D. L., Lauten, J. D., Madden, Th. E., David, L. L., *J. Proteome Res*. 2004, *3*, 1017 – 1023.

[11] Zhu, K., Kachman, M. T., Miller, F. R., Lubman, D. M., Zand, R., *J. Chromatogr. A* 2004, *1053*, 133 – 142.

[12] Janini, G. M., Conrads, T. P., Veenstra, T. D., Issaq, H. J., *J. Chromatogr. B* 2003, *787*, 43 – 51.

[13] Fujii, K., Nakano, T., Hike, H., Usui, F., Bando, Y., Tojo, H., Nishimura, T., *J. Chromatogr. A* 2004, *1057*, 107 – 113.

[14] Wenner, B. R., Lovell, M. A., Lynn, B. C., *J. Proteome Res*. 2004, *3*, 97 – 103.

[15] Andersen, T., Pepaj, M., Trones, R., Lundanes, E., Greibrokk, T., *J. Chromatogr. A* 2004, *1025*, 217 – 226.

[16] Aebersold, R., Goodlett, D. R., *Chem. Rev*. 2001, *101*, 269 – 295.

[17] Kaltenbrunner, O., Tauer, C., Brunner, J., Jungbauer, A., *J. Chromatogr*. 1993, *639,* 41 – 49.

[18] Eshraghi, J., Chowdhury, S. K., *Anal. Chem*. 1993, *65*, 3528 – 3533.

[19] Miliotis, T., Ericsson, P. O., Marko-Varga, G., Svensson, R., Nilsson, J., Laurell, T., Bischoff, R., *J. Chromatogr. B* 2001, *752*, 323 – 334.

[20] Chowdhury, S. K., Chait, B. T., *Anal. Chem*. 1991, *63*, 1660 – 1664.

[21] Wagner, G., Regnier, F. E., *Anal. Biochem*. 1982, *126*, 37 – 43.

[22] Hu, S., Xie, Y., Ramachandran, P., Ogorzalek Loo, R. R., Li, Y., Loo, J. A., Wong, D. T., *Proteomics* 2005, *5*, 1714 – 1728.

[23] Saitoh, E., Isemura, S., Sanada, K., *J. Biochem*. 1983, *93*, 495 – 502.

[24] Wong, R. S., Hofman, T., Bennick, A., *J. Biol. Chem*. 1979, *254*, 4800 – 4808.

