ELSEVIER

Journal of Chromatography A, 924 (2001) 359–368

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Automated signature peptide approach for proteomics

Larry Riggs, Cathy Sioma, Fred E. Regnier*

Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA

Abstract

This paper addresses the issue of automating the multidimensional chromatographic, signature peptide approach to proteomics. Peptides were automatically reduced and alkylated in the autosampler of the instrument. Trypsin digestion of all proteins in the sample was then executed on an immobilized enzyme column and the digest directly transferred to an affinity chromatography column. Although a wide variety of affinity columns may be used, the specific column used in this case was a Ga(III) loaded immobilized metal affinity chromatography (IMAC) column. Ga(III)–IMAC is known to select phosphorylated peptides. Phosphorylated peptides selected by the affinity column from tryptic digests of milk were automatically transferred to a reversed-phase liquid chromatography (RPLC) column. Further fractionation of tryptic peptides on the RPLC column was achieved with linear solvent gradient elution. Effluent from the RPLC column was electrosprayed into a time-of-flight mass spectrometer. The entire process was controlled by software in the liquid chromatograph. With slight modification, it is possible to add multiple columns in parallel at any of the single column positions to further increase throughput. Total analysis time in the tandem column mode of operation was under 2 h.

Keywords: Automation; Signature peptide approach; Immobilized metal affinity chromatography; Peptides; Proteins

1. Introduction

Proteomics and the issue of how to rapidly identify thousands of proteins in biological extracts is currently a subject of great interest. Initially the strategy was to (1) separate the protein components by two-dimensional (2D) gel electrophoresis, (2) locate protein spots in the gel by staining, (3) mechanically excise spots of interest, (4) reduce, alkylate, and tryptic digest each excised protein, and (5) analyze the tryptic digests either by matrix-assisted laser desorption ionization (MALDI) or electrospray ionization (ESI) mass spectrometry [1–4]. Although this process has been automated through a series of robotic operations, it is still very complex and does not interface easily with mass spectrometry.

Recent effort in a number of laboratories is showing that multidimensional chromatography is a viable alternative to 2D gel electrophoresis as the separation engine for proteomics [5–9]. In addition, the analyte transport to the mass spectrometer from separation columns is much more efficient than from gels [10]. The most general strategy in the chromatographic approach is to initially tryptic digest all proteins in the sample then separate the peptide fragments by reversed-phase chromatography [11,12] or capillary electrophoresis [13]. Whereas 2D electrophoresis is particularly powerful in protein separations, liquid chromatography (LC) and capillary electrophoresis (CE) are unparalleled in peptide

*Corresponding author. Tel.: +1-765-494-1648; fax: +1-765-494-0359.

E-mail address: fregnier@purdue.edu (F.E. Regnier).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)00900-1

separations. In either case, the total number of peptides generated is the same.

One of the problems with the LC or CE approach described above is the large number of peptides that must be separated simultaneously as a prelude to mass spectrometry. The ideal case would be to reduce the number of peptides that must be analyzed to identify proteins. Generally, there are more peptides in tryptic digests than needed for protein identification. A small number of peptides are sufficient to identify most proteins [14]. Identification of proteins from a small number of their constituent peptides is also known as the signature peptide approach. Moreover, certain types of peptides may be of much greater interest, as in the case of post-translational modifications. These too may serve as a signature of their parent. This idea of ''signature peptides'' has led to the introduction of a number of affinity methods for the selection of a subset of peptides from proteins [7]. The general protocol is to affinity select peptides with specific structural features from a proteolytic digest then transfer them to a reversed-phase chromatography column where they are further fractionated prior to mass spectrometry. One approach is the selection of cysteine-containing peptides [15,16]. Searches of genomic databases as of November 1999 indicate that 89.3% of all the proteins in man and 85.7% of those in *Escherichia coli* will contain cysteine [17]. Histidine is another low abundance amino acid that has been used to select peptides from proteolytic digests [18]. According to genomic databases, the distribution of histidine in the proteins of a variety of organisms, such as yeast and *E. coli*, varies less than cysteine and is uniformly greater than 97% [17]. Peptides containing post-translational modifications have also been selected, as in the cases of glycopeptides [19,20] and phosphopeptides [21–23].

The question with multidimensional chromatographic approaches is whether they are easier to automate than the 2D gel electrophoresis strategy. Early studies with immunoselection of hemoglobins indicate that instrumentation designed for integrated multidimensional chromatography can automatically provide tryptic maps of selected proteins in slightly more than an hour [24]. The primary objective of this study was to examine similar systems for proteomics. Complex protein extracts were automatically reduced, alkylated, trypsin digested, specific classes of peptides affinity selected from the digests, selected peptides directly transferred to a reversed-phase chromatography column and further fractionated and the chromatographic effluent directly analyzed by ESI-MS.

2. Materials and methods

2.1. Materials

Bovine milk was obtained from the Purdue University Dairy (West Lafayette, IN, USA). Acetonitrile (ACN) and urea were purchased from Mallinckrodt. Calcium chloride ($CaCl_2$) and ammonium hydroxide (NH_4OH) were purchased from Fisher Scientific. Poros AL 20, Poros MC 20 column (100×4.6 mm I.D.) and a PepMap C_{18} reversed-phase column of 250×4.6 mm I.D. dimensions filled with 5 μm particles were purchased from PE Biosystems (Framingham, MA, USA). Trypsin, tris-(hydroxymethyl)aminomethane (Tris-base), tris-(hydroxymethyl)aminomethane hydrochloride (Tris-acid), Triton X-100, sodium phosphate monobasic, iodoacetic acid, dithiothreitol (DTT), sodium cyanoborohydride, glacial acetic acid (AA) and HPLC-grade formic acid were purchased from Sigma (St. Louis, MO, USA).

2.2. Protocol for trypsin immobilization

Trypsin was coupled to Poros AL 20 containing an aldehyde functional group through the amine groups on the surface of the protein. One gram of Poros AL 20 bulk media in 100 m*M* sodium phosphate (pH 7.2) was incubated in a round-bottom flask with 19 m*M* of benzamidine. 15 mg of trypsin, 66 m*M* sodium cyanoborohydride, and 100 m*M* sodium phosphate, pH 7.2, and 1.5 *M* sodium sulfate were added to the slurry in the round-bottom flask and incubated at room temperature overnight while rotating slowly. Sodium cyanoborohydride in 0.2 *M* Tris (pH 8.0) was added to the mixture and the slurry rotated for an additional 30 min. After filtering and washing with water, the immobilized enzyme media was suspended in 50 m*M* sodium phosphate buffer (pH 7.2) and packed in a 100×4.6 mm I.D. column

after which it was washed with 10 column volumes (CVs) of the packing buffer. Finally, the column was washed with 10 CVs of 1 *M* NaCl and 10 CVs of digestion buffer (50 m*M* Tris pH 8.0, 10 m*M* $CaCl_2$), respectively.

2.3. Preparation of the Ga(III)–IMAC column

The immobilized metal affinity chromatography (IMAC) column was washed and loaded with Ga(III) as follows: 15 CVs of 500 m*M* EDTA and 1 *M* NaCl were passed through the column followed by 20 CVs of double deionized (ddI) water. Next, 50 mL of 100 m*M* Ga(III)Cl_3 was passed through the column at 0.5 mL/min. Finally, a three CVs wash of ddI water and a three CVs wash with 0.1% acetic acid completed the loading of the gallium onto the IMAC column.

2.4. Instrumentation

Reduction, alkylation, digestion, affinity selection and RPLC analyses were automated on an Integral Micro-analytical Workstation (PE Biosystems). Analyte detection was achieved by monitoring absorbance at 214 nm. Bovine milk was reduced and alkylated using the instrument auto-sampler. On-column digestions were performed using a tandem column configuration. The first column consisted of a 100×4.6 mm I.D. Porozyme immobilized trypsin sorbent. Peptides generated in the immobilized enzyme column were eluted to the Ga(III)–IMAC 100×4.6 mm I.D. with 0.1% acidic acid. Phosphorylated peptides selected from digest by the IMAC column were eluted from the column with 0.2 *M* phosphate buffer (pH 8.4) and directly transferred to a 250×4.6 mm I.D. PepMap C_{18} silica column (PE Biosystems) for further purification.

The volume of each column in the analytical system along with flow-rate can be designated in the software of the Integral Micro-analytical Workstation. This allows elution volumes of columns to be defined in column volumes. It is for this reason that elution volumes in the Methods section of this paper are widely defined in terms of column volumes.

Mass spectral analyses were performed using a Mariner Biospectrometry Workstation (PE Biosystems) equipped with an atmospheric pressure ESI source. All spectra were obtained in the positive ion mode. The LC, as well as the MS, monitored the RPLC separation. Three seconds acquisition time was used to obtain each spectrum.

2.5. Automated reduction and alkylation of milk

Prior to automation, the milk was de-lipidated by centrifugation at 2500 *g* for 30 min. The lipid layer was skimmed from the surface. Milk, thus treated, was referred to as skimmed milk. Samples were then transferred to the auto-sampler of the Integral for the remainder of the analysis. The reduction step required the addition of 4 *M* urea, 10 m*M* DTT and digestion buffer for a final volume of 1 mL. To facilitate reduction, the sample was heated at 50°C for 15 min. Alkylation was then achieved through the addition of 10 m*M* iodoacetic acid (IAA) to the solution with 3 s of mixing.

2.6. Automatic on-column trypsin digestion of milk

Reduced and alkylated milk solution (500 μL) was injected automatically and passed through the trypsin column, either continuously at 50 μL/min or with stopped flow for a fixed period of time ranging up to an hour. Digestion of complex mixtures is generally complete after 15 min in the immobilized enzyme column at 37°C [25]. The resulting peptides were directly transferred to the Ga(III)–IMAC affinity column for selection of phosphorylated peptides.

2.7. De-phosphorylation of signature peptides

De-phosphorylation of the skimmed milk peptides was achieved in two ways. One was by incubating the mixture overnight at pH 11 and 37°C. This process is known to cleave the phosphate ester moiety from peptides phosphorylated at serine and threonine residues by a β-elimination process [26]. The second mode of de-phosphorylation was with 10^{-4} *M* alkaline phosphatase at pH 8 for 8 h. In both cases the resulting de-phosphorylated peptides were acidified to a pH of 3.5 and analyzed with the Ga(III)–IMAC column.

2.8. Automated affinity section

On-line acidification of the sample to pH 3.5 was achieved by introducing 30% acetic acid into the effluent from the trypsin column through a T-junction prior to the Ga(III)–IMAC column. Addition of acidic acid was achieved with a syringe pump in the Integral. The sample flow onto the IMAC column was at 0.2 mL/min for two CVs. This was followed with sequential washes with two CVs of 0.1% AA, three CVs of 30% ACN–0.1% AA and four CVs of 0.1% AA before elution of the phosphorylated peptides into the PepMap C_{18} column with three CVs of 0.2 *M* phosphate buffer, pH 8.4.

2.9. Reversed-phase liquid chromatography

Samples were chromatographed on a PepMap C_{18} column equilibrated with mobile phase 1A, containing 0.1% ACN and 0.1% formic acid in ddI water, at a flow-rate of 1.00 mL/min for two CVs. Gradient elution of the analyte was achieved using 100% mobile phase 1A to 100% mobile phase 2A (95% ACN–0.1% formic acid in ddI water) over 60 min at a flow-rate of 1 mL/min. The gradient was then held at 100% mobile phase B for an additional CV. Throughout the analysis, peptide elution was monitored with an on-line UV detector set at 214 nm. The peptides were simultaneously monitored by ESI-MS directly coupled to LC.

2.10. Automation of the three-column system for the analysis of phosphorylated proteins

A schematic of the automated system is shown in Fig. 1a. Solvents used are as follows: 1A=0.1% ACN–0.1% formic acid, 1B=0.1% acetic acid, 1C=0.2 *M* phosphate buffer, 2A=95% ACN–0.1% formic acid, 2B=0.1% acetic acid–30% ACN, 2C=digestion buffer, syringe pump 30% acetic acid. The analytical reversed-phase column was equilibrated in solvent 1A at 1 mL/min for two CVs prior to use. The system can be purged with a mobile phase at any of the LC processing steps with the following settings, as shown in Fig. 1a: injector valve, load position; valve 1, counterclockwise, positions 1-2, 3-4, etc. connected; valve 1, clockwise, positions 1-10, 9-8, etc. connected; valve 2, clockwise, positions 1-10, 9-8, etc. connected. The immobilized trypsin and gallium IMAC columns were equilibrated in 100% 2C and 100% 1B, respectively.

Step 1. The injection valve was in the counterclockwise position 1-2, 3-4, etc. when a 500 μL reduced and alkylated sample was loaded by the auto-sampler into the sample loop. The injection valve and valve 1 were moved to the clockwise position 1-10, 9-8, etc. allowing the pumps to deliver buffer 2C, pushing the sample to the immobilized trypsin column at a rate of 5 μL/min for 0.5 CV. This was enough volume to place the sample into the trypsin column. The pumps were then stopped and the sample was allowed to incubate on the column for up to 15 min to ensure complete digestion (Fig. 1b).

Step 2. The injection valve, valves 1 and 2 were switched to the counterclockwise position 1-2, 3-4, etc. The syringe pump injected two CVs of 30% acetic acid through the immobilized trypsin column, carrying the peptide mixture through a mixing tee and onto the gallium IMAC column (Fig. 1c). Those peptides that did not bind to the column were eluted to waste.

Step 3. The injection valve was switched to the clockwise position 1-10, 9-8, etc. The Ga–IMAC column was washed with three CVs of buffers 1B and 2B followed by four CVs of buffer 1B again.

Step 4. Valves 1 and 2 were switched to the clockwise position 1-10, 9-8, etc. Buffer 1C was pumped through the Ga–IMAC causing the phosphorylated peptides to elute from the column and concentrate onto the head of the PepMap C_{18} column (Fig. 1d).

Step 5. Valves 1 and 2 were then switched to the counterclockwise position 1-2, 3-4, etc. and the PepMap C_{18} column was equilibrated with buffer 1A. Valve 1 was switched to the clockwise position 1-10, 9-8, etc. connecting the ESI to the outlet of the C_{18} column. A gradient of 100% 1A buffer to 100% 2A buffer was passed over the column for a duration of 60 min (Fig. 1e).

2.11. ESI-MS analysis

A PE Biosystems Mariner Biospectrometry Work-

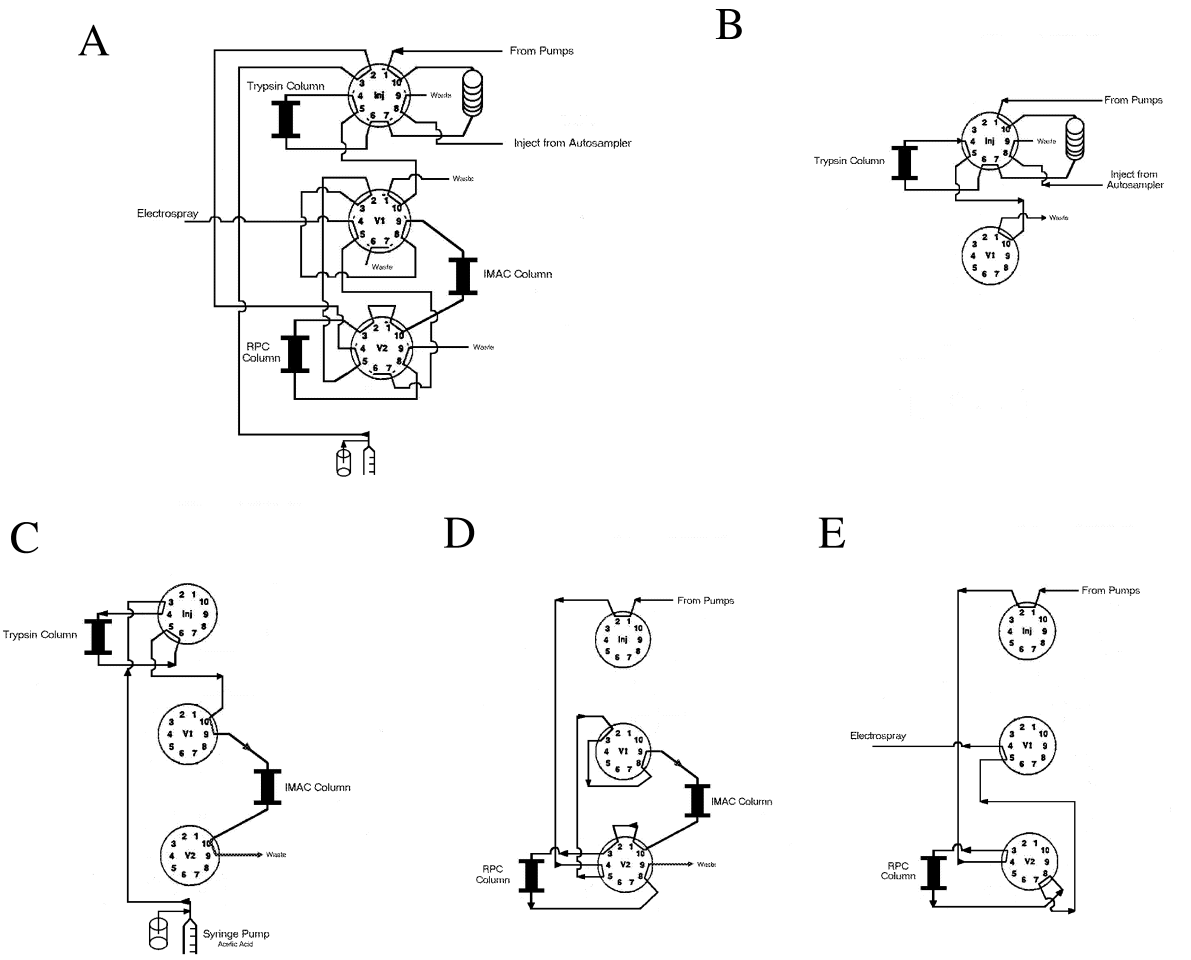

Fig. 1. Three-column chromatography system for the isolation and characterization of phosphorylated milk proteins. (A) Overall plumbing system layout for the automatic multidimensional LC–MS (trypsin→affinity→RP chromatography→MS) process. (B) Step 1, reduced and alkylated sample is injected from autosampler and flowed through immobilized perfusion trypsin column. (C) Syringe pump delivers 30% acetic acid solvent to the trypsin column eluting tryptic peptides to the IMAC column. The pH of the tryptic peptides is reduced to pH 3.5 allowing only phosphorylated peptides to bind to the gallium(III)–IMAC column. The column was washed with 20% acetonitrile, 0.1% acetic acid and then washed with 0.1% acetic acid. (D) Elution of the phosphopeptides to the RPLC column with 0.2 *M* phosphate buffer pH 8.4. (E) The RPLC column was equilibrated with 0.1% formic acid before elution of the peptides to the MS via a gradient of 100% 0.1% formic acid to 100% 95% acetonitrile, 0.1% formic acid.

station with a sampling rate of three spectra/s was used for automatic electrospray ionization mass spectrometric analysis. During the coupled ESI-MS acquisition, masses were scanned from m/z 300 to 3000 at 140 V nozzle potential. The instrument was tuned and mass calibrated by manual injection of Mariner Calibration mixture (PE Biosystems) at 3.0 μL/min.

3. Results and discussion

3.1. General system design

The signature peptide approach chosen for these studies involves at least six steps: (i) reduction, (ii) alkylation, (iii) proteolysis, (iv) affinity selection of signature peptides, (v) reversed-phase chromato-

graphy and (vi) electrospray mass spectrometry. In a completely automated system, peptide identification would be automated as well. Automated peptide identification has already been reported [17] and will not be a concern in this work. Based on previous experience of our laboratory, automation of the above steps was achieved (i) in a tandem column system [17], (ii) chemical reactions requiring reagent addition were achieved in the auto-sampler of the commercial LC instrumentation [27], (iii) proteolytic digestion was executed on an immobilized enzyme column to increase the rate of hydrolysis [23], (iv) analytes were directly transferred between columns through valves and (v) the sequencing of analytical operations was controlled with the LC data system. Systems that meet these specifications can be easily assembled from a number of commercial instruments. Fluid handling in the studies described below required two types of pumps: a high-pressure pumping system capable of delivering mobile phases for the elution of both the affinity and reversed-phase columns and a low-pressure syringe pump for reagent additions. Valves for column coupling may be of either the six- or 10-port type, but 10-port valves are preferred, allowing more complex configurations. The system configuration and location of the various control components used for the experiments described in this paper are shown in Fig. 1. An additional design incorporates the off-line use of a MALDI-time-of-flight (TOF) MS for detection.

3.2. Proteolysis

Reduction and alkylation are essential prerequisites for proteolysis of many proteins. These reactions were achieved in the auto-sampler of the LC instrument as described previously [28]. The position of all sample vials, reagent vials, and reaction vials were pre-set in the computer using a sample table layout. A detailed method was written using the instrument software package indicating that an aliquot of sample was to be withdrawn from a particular sample vial and injected into a specified reaction vial, containing a reducing agent, such as DTT. The reducing agent had been obtained by the auto-sampler and placed in the reaction vial prior to the addition of the sample. The sample and reducing agent were then mixed and heated in the incubator. Iodoacetic acid was added to the same reaction vial and mixed to achieve alkylation. Reduction and alkylation were achieved in parallel with the proteolysis, affinity selection and reversed-phase chromatography steps. The auto-sampler was programmed to advance to the reduction and alkylation of a new sample while the LC system was digesting and analyzing a previous sample.

An immobilized trypsin column was chosen for digesting proteins because (i) samples of a few microliters can be digested and transported through the enzyme bed with little dilution, (ii) a 10–1000-fold excess of enzyme can be used to accelerate reaction rates by several orders of magnitude, (iii) contamination of the sample with peptides from human skin and trypsin autodigestion is eliminated, (iv) reaction time is easily controlled with flow-rate, (v) the immobilized enzyme column can process many samples, and (vi) the system is easy to automate. Immobilized enzyme columns of 1–25 cm in length were examined. The concentration of bound trypsin in the 10 cm column on which most of the studies were based was in the range of 10 mg/mL. When a 1 mg/mL solution of protein with a relative molecular mass of 50 000 was swept through this reactor the enzyme-to-substrate ratio after dilution is roughly 15. Enzyme excess is even greater in dilute solutions. In the case of a 1 μg/mL solution, the molar excess of trypsin would be 15 000. Obviously, such an enormous excess of enzyme greatly accelerates protein digestion. Another advantage of the large concentration of enzyme on an immobilized column is that trypsin inhibitors in a sample such as lysozyme in milk will have little effect on the activity of the column.

Digestion reaction times, being controlled automatically by flow-rate, were varied from less than a minute to hours. In cases of more complex mixtures, stopped-flow incubation was incorporated. One of the advantages of stopped-flow incubation is that the system can be executing the reversed-phase chromatographic separation of a previously digested sample during proteolysis of a new sample. Recent studies in 25 cm columns have shown that proteolysis of an *E. coli* lysate can be achieved in 15 min [24]. This brings proteolysis in trypsin columns into the same time frame as the reversed-phase

Table 1
Time comparison of the reduction, alkylation and digestion of a protein mixture in solution vs. automation of the process

Reduction 30 min	Reduction 15 min
Alkylation 20 min	Alkylation 3 s
Digestion 8 h	Digestion 30 min
IMAC 20 min	IMAC 20 min
RPLC 30 min	RPLC 30 min
MS off-line 1 h	MS on-line 3 s
Total time 10 h, 40 min	Total time 1 h, 35 min

separation step. Proteolysis in free solution takes much longer (Table 1).

3.3. Affinity selection

It has been noted above that affinity selection of peptides is used in proteomics and that signature peptides from post-translationally modified proteins are being used to study regulation of this process. Phosphorylation is a type of post-translational modification (PTM) that can be easily studied with this approach. Phosphorylation can occur at one or more sites in the same protein. It generally occurs on serine, threonine or tyrosine, and is thought to have multiple functions [29]. IMAC columns loaded with a variety of metals have the ability to capture and release phosphorylated peptides, but Ga(III) loaded columns are best in terms of percent recovery and avoidance of non-specific binding [20].

A problem of executing Ga(III)–IMAC selection immediately after proteolysis is that effluent from the enzyme column is at pH 8.0 and the binding of phosphopeptides to the IMAC column is best achieved at slightly acidic pH. Therefore, as the sample was being transferred from the immobilized trypsin column to the IMAC column the effluent was adjusted to roughly pH 3.5 via a T-junction with acetic acid. The phosphorylated peptides were bound and most non-phosphorylated peptides were eluted to waste. A wash of 30% ACN and 0.1% AA removed any remaining non-specifically bound peptides, leaving only phosphorylated peptides on the column. While there are several methods of eluting phosphopeptides from the IMAC column, 0.2 *M* phosphate was chosen for its high recovery of peptides. EDTA may also be used as a desorbing agent but is problematic in that it can interfere with interpretation during mass spectrometry. A third method of elution is to use a weak base, however poor recovery is reported [20].

3.4. Reversed-phase chromatography

Affinity selected peptides were eluted from the IMAC column directly onto the reversed-phase column, enhancing speed and recovery while minimizing the potential of contamination. Further resolution of the selected peptides is desirable to reduce the complexity of the mixture. Reversed-phase chromatography of peptides in this study was completed by standard methods using a PepMap C_{18} silica-based octadecylsilane bonded phase column. Comparisons between the IMAC selected peptides and the peptides resolved by reversed-phase chromatography are shown in Figs. 2 and 3, respectively.

3.5. Mass spectrometry

Complete automation of sample delivery to a mass spectrometer is most easily achieved with ESI-MS. Automation can be achieved with MALDI-MS, but fractions must be collected then aliquoted onto MALDI sample plates along with the MALDI matrix. Although equipment was available for automation of MALDI-MS, it was not used.

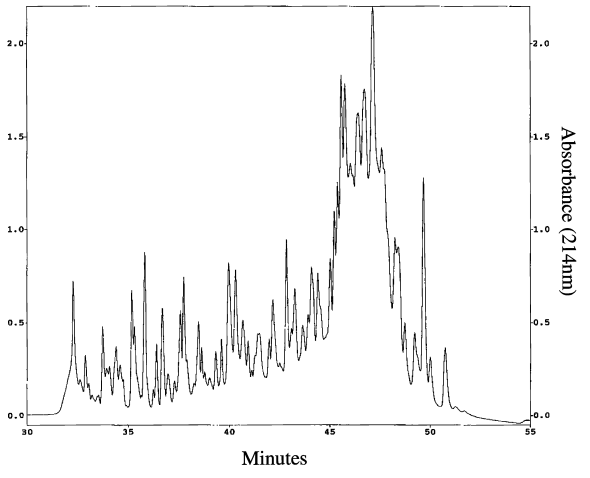

Fig. 2. Phosphorylated milk peptides Ga(III)–IMAC selected and resolved by RPLC and MS.

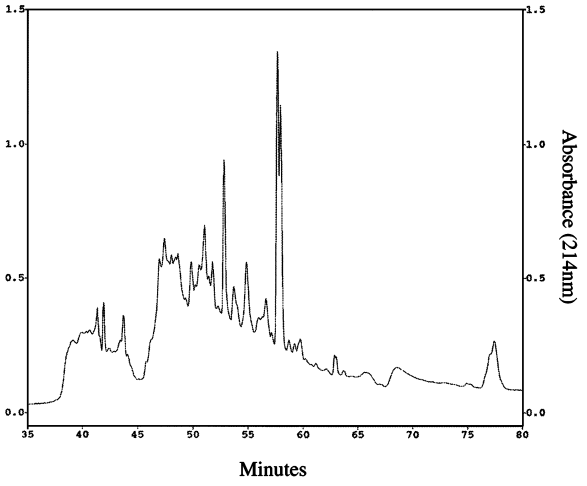

Fig. 3. Reversed-phase chromatogram of digested milk resolved on a PepMap C_{18} reversed-phase column. Elution was achieved with a linear gradient ranging from 0.1% acetonitrile (ACN) and 0.1% formic acid in double diionized water (ddI water) to 95% ACN–0.1% formic acid in ddI water at 1 mL/min over 60 min.

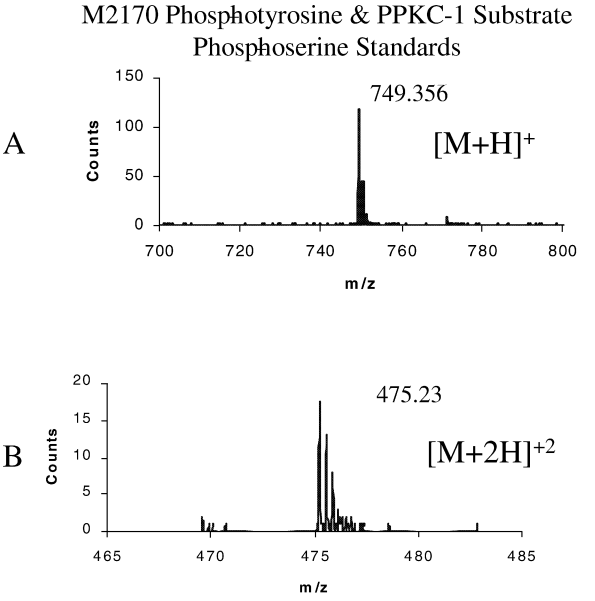

Fig. 4. Tyrosine and serine phosphorylated peptide standards selected via gallium–IMAC, resolved by RP chromatography and MS. (A) Ac–Ile–pTyr–Gly–Glu–Phe–NH_2 identified via the $[M+H]^+$ ion. (B) H–Lys–Arg–Pro–pSer–Gln–Arg–His–Gly–Ser–Lys–Tyr–NH_2 identified via the $[M+2H]^{2+}$ ion.

3.6. Analysis of a complex sample

The model system chosen for this study was milk. Milk contains four types of phosphorylated proteins called caseins: αS1, αS2, β and κ. These proteins are phosphorylated on the serine or threonine amino acids. αS1-Casein is phosphorylated at eight sites, αS2 at 11 and β at five. κ-Casein is phosphorylated at only one site. Phosphorylation at fewer than the maximum number of sites is also possible. γ-Casein and proteose peptones in milk are fragments of the original four caseins. Traces of lytic enzymes in milk degrade milk proteins. Unidentified proteins in the samples analyzed in this study probably result from these protein fragments.

To ensure that the Ga(III)–IMAC could capture phosphorylated peptides, standards containing serine or tyrosine phosphorylated amino acids were captured and identified as shown in Fig. 4.

Fig. 2 shows the chromatogram of skimmed milk digested and resolved on a reversed-phase chromatography column. At the later stage of the gradient, the figure shows a group of poorly resolved peptides. Based on the known structures of the casein components, it is estimated that over 300 peptides will be produced from a tryptic digest of skimmed milk. Fig. 3 shows the reversed-phase chromatogram of peptides captured from the same milk digest by the Ga(III)–IMAC column. Peptides released from the IMAC column were directly transferred to the RPLC column. Approximately 22 peaks were detected from the initial mixture of more than 300 peptides. Hydrolysis of phosphate esters in these samples with either alkaline phosphatase or base prior to Ga(III)–IMAC resulted in the capture of no peptides by the IMAC column (data not shown).

Peptides in these samples were identified by mass alone based on known structures in the literature. β-Casein peptides observed in this sample were as follows. The peptide containing residues 33–48 with one phosphorylation site produced ions at 1031.99 $[M+2H]^{2+}$ and 688.33 $[M+3H]^{3+}$ (Fig. 5A). An αS2-casein peptide of amino acid residues 1–21 with one phosphorylation site produced an ion at 1375.66 $[M+2H]^{2+}$ (Fig. 5A). A β-casein tryptic peptide of amino acid residues 1–25 containing four phosphorylation sites was identified by the spectral ion at 1563.49 $[M+2H]^{2+}$, 1043.16 $[M+3H]^{3+}$ and

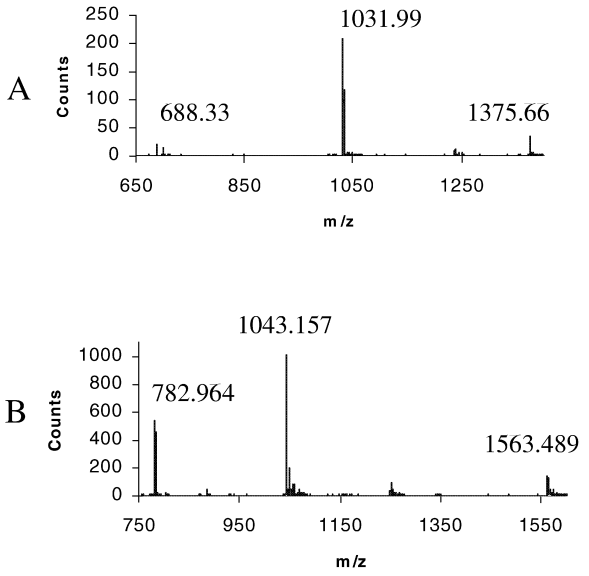

Fig. 5. αS2- and β-casein peptides from milk were selected by gallium–IMAC, resolved on reversed-phase chromatography and identified with ESI-MS. (A) β-Casein peptide aa33–48 identified by $[M+3H]^{3+}$ ion at 688 and $[M+2H]^{2+}$ ion at 1032; αS2-casein peptide aa1–21 identified by $[M+2H]^{2+}$ ion at 1375. (B) β-Casein peptide aa1–25 identified by $[M+4H]^{4+}$ ion at 783, $[M+3H]^{3+}$ ion at 1043 and $[M+2H]^{2+}$ ion at 1563.

782.96 $[M+4H]^{4+}$ (Fig. 5B). Two peptides from αS2-casein were identified, each containin four phosphorylation sites (Fig. 6). One peptide from amino acid residues 16–35 produced a distinguishing ion at 434.87 $[M+5H]^{5+}$. The other peptide from residues 20–35 was observed at m/z 441.14 $[M+6H]^{6+}$. αS1 Peptides coming from amino acid residues 48–69 and 56–69, with three phosphorylation sites, were identified by ions 448.12 $[M+3H]^{3+}$, 437.14 $[M+4H]^{4+}$ and 444.11 $[M+6H]^{6+}$, respectively. A peptide from κ-casein with one phosphorylation site is shown in Fig. 7. This peptide originates from amino acid residues 22–39 and is identified by ions at 1156.78 $[M+2H]^{2+}$ and 771.99 $[M+3H]^{3+}$. Several other peptides in the spectra above could not be identified, even when considering the possibility of partial cleavage by trypsin or the presence of chymotrypsin in the digestive enzyme. These peptides are suspected to be the result of naturally occurring lytic enzymes in milk. The identification of the unknown peptides could be identified through MS–MS sequencing, but doing so was beyond the objectives of this study.

Table 1 presents a step-by-step time comparison of a manual process versus the automated process described in this study. One advantage of the automated system is the total elapsed time to analyze the sample. Identifying an unknown protein mixture using the automated system described above reduces total analysis time to one-tenth that of a manual process. Furthermore, with the exception of affinity selection and resolution by RPLC, each step in the automated process is much less than the manual process. Another advantage enables the analysis of multiple samples in parallel, i.e. reduction, alkylation and digestion of a protein mixture while affinity selection and further identification of a separate

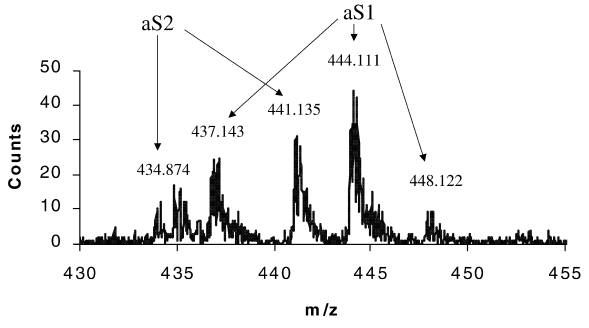

Fig. 6. αS1- and αS2-casein peptides from milk were selected from gallium–IMAC, resolved with reversed-phase chromatography and identified with ESI-MS. αS1 peptides aa48–69 were identified with the $[M+6H]^{6+}$ ion at 444, aa56–69 with the $[M+4H]^{4+}$ ion at 437 and aa125–135 with $[M+3H]^{3+}$ at 448. αS2 peptides aa20–35 were identified with the $[M+5H]^{5+}$ ion at 433 and aa16–35 with the $[M+6H]^{6+}$ ion at 441.

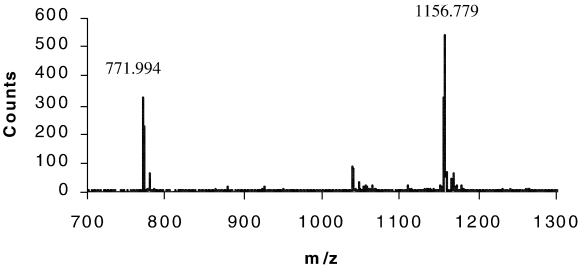

Fig. 7. κ-Casein peptide from milk was selected by gallium–IMAC, resolved on reversed-phase chromatography and identified with ESI-MS. κ-Casein peptide aa22–39 was identified from the $[M+3H]^{3+}$ ion at 771 and the $[M+2H]^{2+}$ ion at 1155.

sample. This allows the analysis of three samples simultaneously.

4. Conclusion

It may be concluded that the multiple steps required in the signature peptide approach demonstrates the ability to simplify complex protein mixtures through multidimensional chromatography, presenting an addressable mixture to the mass spectrometer. The automated system being used demonstrated that the entire process of reduction, alkylation, digestion, affinity selection, reversed-phase chromatography and MS takes only one-tenth of the time that the same process would take by solution digest and manual analysis. It can also be imagined that this automation process clearly reduces the loss of material, and contamination brought into an analysis by manual manipulation. Furthermore, it can be concluded that this automation process is highly reproducible and amenable to many affinity selections and multidimensional chromatographic analyses.

Acknowledgements

The authors gratefully acknowledge support from the National Institute of Health (grant number NIH I ROI GM59996-01) and Applied Biosystems.

References

[1] J. Yates, J. Mass Spectrom. 33 (1998) 1.
[2] D. Figeys, A. Ducret, J.R. Yates, R. Aebersold, Nature Biotechnol. 14 (1996) 1579.
[3] U. Hellman, C. Wernstedt, J. G'Onez, C.H. Heldin, Anal. Biochem. 224 (1995) 451.
[4] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850.
[5] G.T. Nepom, J.D. Lippolis, F.M. White, S. Masewicz, J.A. Marto, A. Herman, C.J. Luckey, B. Faulk, J. Shabanowitz, D.F. Hunt, Proc. Natl. Acad. Sci. USA 98 (2001) 1763.
[6] R.S. Taylor, C.C. Wu, L.G. Hays, J.K. Eng, J.R. Yates III, K.E. Howell, Electrophoresis 21 (2000) 3441.
[7] F.E. Regnier, A. Amini, A. Chakraborty, M. Geng, J. Ji, L. Riggs, C. Sioma, S. Wang, X. Zhang, LC–GC 19 (2001) 200.
[8] T.J. Griffin, S.P. Gygi, B. Rist, R. Aebersold, A. Loboda, A. Jilkine, W. Ens, K.G. Standing, Anal. Chem. 73 (2001) 978.
[9] M.T. Davis, J. Beierle, E.T. Bures, M.D. McGinley, J. Mort, J.H. Robinson, C.S. Spahr, W. Yu, R. Leuthy, S.D. Patterson, J. Chromatogr. B 752 (2001) 281.
[10] K. Johansen, P.O. Edholm, J. Chromatogr. 506 (1990) 471.
[11] R.S. Annan, M.J. Huddleston, R. Verma, R.J. Deshaies, S.A. Carr, Anal. Chem. 73 (2001) 393.
[12] J.K. Eng, A.L. McCormack, J.R. Yates, J. Am. Soc. Mass Spectrom. 5 (1994) 976.
[13] A.J. Link, L.G. Hays, E.B. Carmack, J.R. Yates, Electrophoresis 18 (1997) 1314.
[14] M. Geng, X. Xiang, F. Regnier, J. Chromatogr. A, submitted for publication.
[15] D.R. Goodlett, J.E. Bruce, G.A. Anderson, B. Rist, L. Pasa-Tolic, O. Fiehn, R.D. Smith, R. Aebersold, Anal. Chem. 72 (2000) 1112.
[16] S. Wang, F. Regnier, J. Chromatogr. A, in press.
[17] X. Zhang, Personal communication, 2001.
[18] J. Ji, A. Chakraborty, M. Geng, X. Zhang, A. Amini, M. Bina, F. Regnier, J. Chromatogr. B 745 (2000) 197.
[19] M. Geng, J. Ji, F.E. Regnier, J. Chromatogr. A 870 (2000) 295.
[20] R. Cramer, W.J. Richter, E. Stimson, A. Burlingam, Anal. Chem. 70 (1998) 4939.
[21] M.C. Posewitz, P. Tempst, Anal. Chem. 71 (1999) 2883.
[22] W. Zhou, A.B. Merrick, M.G. Khaledi, K.B. Tomer, J. Am. Soc. Mass Spectrom. 11 (2000) 273.
[23] M. Zachariou, I. Traverso, T.W. Hearn, J. Chromatogr. 646 (1993) 107.
[24] F.Y. Hsieh, H. Wang, C. Elicone, J. Mark, S.A. Martin, F.E. Regnier, Anal. Chem. 68 (1996) 455.
[25] S. Wang, F.E. Regnier, J. Chromatogr. A, in press.
[26] H. Jaffe, C. Veeranna, H.C. Pant, Biochemistry 37 (1998) 16211.
[27] Y.F. Hsieh, N. Gordon, F. Regnier, N. Afeyan, S.A. Martin, G.J. Vella, Mol. Divers. 2 (1996) 189.
[28] T. Nadler, C. Blackburn, J. Mark, N. Gordon, F. Regnier, G. Vella, J. Chromatogr. A 743 (1996) 91.
[29] K. Kivirikko, R. Myllyla, T. Philajaniemi, in: J. Harding, M. Crabbe (Eds.), Post-translational Modification of Proteins, CRC Press, Boca Raton, FL, 1992, p. 1.