ELSEVIER

Journal of Chromatography A, 924 (2001) 345–357

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Proteomics based on selecting and quantifying cysteine containing peptides by covalent chromatography

Shihong Wang, Fred E. Regnier*

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

Abstract

This paper describes a procedure in which cysteine containing peptides from tryptic digests of complex protein mixtures were selected by covalent chromatography based on thiol–disulfide exchange, identified by mass spectrometry, and quantified by differential isotope labeling. Following disruption of disulfide bridges with 2,2′-dipyridyl disulfide, all proteins were digested with trypsin and acylated with succinic anhydride. Cysteine containing peptides were then selected from the acylated digest by disulfide interchange with sulfhydryl groups on a thiopropyl Sepharose gel. Captured cysteine containing peptides were released from the gel with 25 m*M* dithiothreitol (pH 7.5) containing 1 m*M* (ethylenedinitrilo)tetraacetic acid disodium salt and alkylated with iodoacetic acid subsequent to fractionation by reversed-phase liquid chromatography (RPLC). Fractions collected from the RPLC column were analyzed by matrix-assisted laser desorption ionization mass spectrometry. Based on isotope ratios of peptides from experimental and control samples labeled with succinic and deuterated succinic anhydride, respectively, it was possible to determine the relative concentration of each peptide species between the two samples. Peptides obtained from proteins that were up-regulated in the experimental sample were easily identified by an increase of the relative amount of the deuterated peptide. The results of these studies indicate that by selecting cysteine containing peptides, the complexity of protein digest could be reduced and database searches greatly simplified. When coupled with the isotope labeling strategy for quantification it was possible to determine proteins that were up-regulated in plasmid bearing *Escherichia coli* when expression of plasmid proteins was induced. Up-regulation of several proteins of *E. coli* origin was also noted. © 2001 Published by Elsevier Science B.V.

Keywords: Peptides; Cysteine; Proteins; Dipyridyl disulfide

1. Introduction

Efforts directed at the analysis of the proteome have gained impetus within the last few years [1–6]. The most widely used approach to proteomics is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) where proteins are separated in the first dimension by isoelectric focusing and in the second by sodium dodecyl sulfate–PAGE [7]. Proteins thus separated are commonly visualized through staining or fluorescent labeling, mechanically excised from the gel slabs, exposed to trypsin while still in the gel slices, and identified by mass fingerprinting the tryptic peptides obtained from gel slices [8–12]. 2D-PAGE can resolve up to 6000 spots and has been used in a wide variety of studies ranging from localizing individual proteins at a subcellular level [13] and recognizing changes in protein expression [14] to identifying target genes of regulatory proteins [15]. The problem with 2D-PAGE is that quantification is poor, it is labor intensive, difficult to auto-

*Corresponding author. Tel.: +1-317-4941-648; fax: +1-317-4940-359.

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

0021-9673/01/$ – see front matter
PII: S0021-9673(01)00961-X

mate, and does not couple readily to mass spectrometry (MS).

Promising new techniques are emerging that replace 2D-PAGE. Most involve some combination of high-performance liquid chromatography (HPLC) or capillary electrophoresis (CE) coupled with MS to either create a ''virtual 2D gel'' [16–23] or go directly to the peptide level of analysis by tryptic digesting all the proteins in samples as the initial step of analysis. A recent report shows that in a direct comparison of 2D-PAGE and chromatographic methods, almost twice as many proteins were identified with chromatographic methods in a shorter time [24].

But there is still the issue of quantification. Measuring either the relative abundance of proteins or changes in protein concentration remains a major challenge in proteomics. Heavy isotope labeling is emerging as an alternative to staining and dye labeling methods for quantification. Isotope labeling tactics generally involve differential labeling of proteins or peptides from control and experimental samples with different isotopic forms of a derivatizing agent. Subsequent to proteolysis, quantification is achieved by comparing the relative abundance of peptide species in the two samples. The protein parents of peptides are identified from the mass and often some sequence analyses. Peptide sequence is obtained by mass analysis of sequence ladders generated during collision-induced dissociation (CID) in the mass spectrometer. Two different isotope-labeling methods have recently been described that exploit isotope ratio quantification. One is the isotope-coded affinity tag (ICAT) method that labels and selects cysteine containing peptides from proteolytic digests [23]. The second is based on a global internal standard technology (GIST) designed to generate sample identifiable forms of peptides for all components of tryptic digests [25]. In both of these methods, isotopically labeled peptide mixtures from control and experimental samples are mixed before subsequent fractionation and mass analysis.

Peptide selection following isotope labeling is an important issue in chromatographic approaches to proteomics. The enormous advantage of using peptide selectors is that specific features of proteins may be targeted for study. Moreover, the complexity of tryptic digests can be greatly reduced by selecting specific classes of peptides. Targeted analyses based on antibody selection of specific proteins or types of post-translational modification [26–28], lectin selection of specific types of glycosylation [29,30], and avidin selection of biotinylated peptides [31–33] have already been described. Immobilized metal affinity chromatography provides another type of selection, particularly in the case of histidine containing peptides [25,34] and those that are phosphorylated [35]. But more selection methods are likely to emerge because these methods do not cover all the classes of modification nature provides. In addition to selectivity, simplicity and ease of use are other important criteria for high throughput screening methods.

Covalent chromatography has not been described in the context of targeted peptide selection although it has been used as a means of protein separation since 1973 [36]. It has long been known that thiol bearing support matrices may be used to isolate proteins through thiol–disulfide interchange [36–42]. The problem in using this method for selection of cysteine containing peptides from tryptic digests is that sulfhydryl groups are generally alkylated before proteolysis. This precludes disulfide exchange.

This paper explores the development of a covalent chromatography and isotope labeling method for selecting cysteine containing peptides from proteolytic digests and quantifying protein dynamics in *Escherichia coli*.

2. Experimental

2.1. Materials

DL-Dithiothreitol (DTT), iodoacetic acid, L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin, urea, isopropyl β-thiogalactopyranoside (IPTG), protamine sulfate, guanidine hydrochloride, *N*-hydroxylamine, α-cyano-4-hydroxycinnamic acid, tris(hydroxymethy)aminomethane (Tris base), tris(hydroxymethyl)aminomethane hydrochloride (Tris acid), angiotensin and insulin B chain were purchased from Sigma (St. Louis, MO, USA). Bacto Agar, tryptone peptone and yeast extract were from DIFCO (Sparks, MD, USA). Sodium phosphate, sodium chloride, calcium chlo-

ride, sodium acetate, (ethylenedinitrilo)tetraacetic acid disodium salt (EDTA), sodium hydrogencarbonate, potassium sulfate, ethylene glycol, isopropanol, acetic acid and HPLC-grade acetonitrile (ACN) were from Mallinckrodt Baker (Phillipsburg, NJ, USA). 2,2′-Dipyridyl disulfide, bromoacetic acid ($BrC^1H_2CO_2^1H$), deuterated bromoacetic acid ($BrC^2H_2CO_2^2H$), succinic anhydride ($C^1H_2C^1H_2C_2O_3$) and deuterated succinic anhydride ($C^2H_2C^2H_2C_2O_3$) were obtained from Aldrich (Milwaukee, WI, USA). HPLC-grade trifluoroacetic acid (TFA) was purchased from Pierce (Rockford, IL, USA). Poros 50 R2 media and the PepMap C_{18} column were from Applied Biosystems (Framingham, MA, USA). Thiolpropyl Sepharose 6B was from Amersham Pharmacia Biotech (Piscataway, NJ, USA).

2.2. E. coli cell culture

The expression plasmid pC/PAD with the phenyacetaldehyde dehydrogenase (PAD) gene inserted was a generous gift from Dr. Henry Weiner's laboratory at Purdue University. The expression vector was transformed into *E. coli* strain JM109 for specifically expressing PAD. The vector has an Ipp promoter and a lac operator so that expression of phenyacetaldehyde dehydrogenase can be induced with IPTG. The transformed *E. coli* was plated on a Luria–Bertani (LB) plate and incubated at 37°C overnight. A colony was selected and incubated in 5 ml LB medium overnight and then added to LB medium at 1:100 ratio and cultured at 37°C. IPTG (1 m*M*) was used to induce the culture. Cells were harvested 2 h after induction.

2.3. Preparation of E. coli lysate

E. coli cells were resuspended in 50 m*M* phosphate buffer, pH 7.5 containing 500 m*M* NaCl. The suspended culture was passed through a French Press at 1200 p.s.i. and the effluent centrifuged at 24 000 *g* for 30 min at 4°C (1 p.s.i. = 6894.76 Pa). Protamine sulfate was added to the supernatant at the level of 1.0 g/100 ml and the solution stirred in a 4°C cold room for 30 min. The sample was again centrifuged at 24 000 *g* for 30 min at 4°C and the supernatant was collected. This supernatant will subsequently be referred to as the *E. coli* lysate. The uninduced cell lysate was used as a control and the induced cell lysate was used as the experimental sample.

2.4. Protein derivatization followed by tryptic digestion

A sample of 10–30 mg of protein was first reduced with DTT at a 40 molar excess in 6 *M* guanidine·HCl at alkaline pH [43]. The reduced protein was applied to a reversed-phase chromatography column (250×4.6 mm I.D.) packed with POROS 50 R2 media (from Applied Biosystems). After thoroughly washing the column with water containing 0.1% TFA, 1–2 m*M* 2,2′-dipyridyl disulfide in 0.1 *M* sodium acetate (pH 5.0) containing 50 m*M* EDTA was passed through the column continuously for more than 2 h. Excess 2,2′-dipyridyl disulfide was then eluted from the column with water containing 0.1% TFA until the absorbance at both 280 nm and 343 nm decreased to a baseline level. This process was monitored at 343 nm. Activated proteins were eluted from the reversed-phase column with 70% acetonitrile containing 0.1% TFA in water. Solvent was evaporated and the proteins were used for tryptic digestion in the following step.

Activated proteins (10–30 mg) were denatured in digestion buffer containing 8 *M* urea in 50 m*M* Tris (pH 8.0) with 10 m*M* $CaCl_2$ and after denaturing the solution was diluted with Tris buffer to a urea concentration of 2 *M*. Trypsin was added in a 1:50 (w/w) ratio and the mixture allowed to digest overnight at 37°C.

2.5. Capture of activated thiol peptides on Thiolpropyl Sepharose 6B

Dry Thiolpropyl Sepharose 6B (2–3 g) was solvated and reduced for more than an hour in 10 ml solution containing 1% (w/v) DTT, 0.3 *M* sodium hydrogencarbonate and 1 m*M* disodium EDTA (pH 8.0). The reduced gel was then washed thoroughly on a fritted glass filter with a solution of 0.1 *M* acetic acid, 0.5 *M* NaCl, and 1 m*M* disodium EDTA (pH 2–3) until DTT was no longer detectable.

The tryptic digest (15–20 ml) was buffer exchanged into 0.1 *M* acetate (pH 5.0), 50 m*M* EDTA with 0.5 *M* K_2SO_4 and 6 *M* urea and then mixed

with 5–8 ml of reduced Thiopropyl Sepharose 6B. The mixture was then gently shaken at room temperature for more than 2 h. Aliquots of sample were monitored at 343 nm until absorption from thiolpyridone (λ_{max} = 343 nm) reduced to zero.

2.6. Selection of cysteinyl peptides

After peptide capture, the gel was filtered sequentially on a glass filter with 50 ml 0.1 *M* acetate buffer (pH 5.0) containing 0.5 *M* K_2SO_4 and 6 *M* urea and 50 ml of 0.1 *M* acetate buffer (pH 5.0) with 1.0 *M* NaCl. Tris–HCl (0.1 *M*, pH 7.5) was then used to adjust the gel to alkaline pH and the gel was washed with 50 ml of 40% ethylene glycol in 0.1 *M* Tris–HCl (pH 7.5) and then 50 ml of 30% isopropanol in 0.1 *M* Tris–HCl (pH 7.5). Immobilized cysteine containing peptides were released from the gel with 10 ml of 25 m*M* DTT in 0.1 *M* Tris–HCl (pH 7.5) and 1 m*M* EDTA. The released peptides were alkylated with iodoacetic acid before they were fractionated by reversed-phase chromatography.

2.7. Alkylation of cysteinyl peptides with bromoacetic acid and deuterated bromoacetic acid

An aliquot of peptides isolated from lysozyme were released from Sepharose 6B by 25 m*M* DTT and was alkylated in the darkness for an hour at room temperature with 50 m*M* bromoacetic acid containing an equimolar mixture of bromoacetic acid and deuterated bromoacetic acid. The reaction product was applied to reversed-phase chromatography and the fractions were examined with matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS).

2.8. Acylation of peptides by succinic anhydride and deuterated succinic anhydride

A 50-fold molar excess of succinic anhydride and deuterated succinic anhydride were added individually to experimental and control samples in phosphate buffer, pH 7–8. Labeling reagent was added in small aliquots over the course of the first hour and the reaction allowed to proceed for another 2 h. *N*-Hydroxylamine was then added in excess and the pH was adjusted to 11–12. Incubation with hydroxylamine was allowed to proceed for 10 min. The function of the hydroxylamine reaction was to hydrolyze esters that might have been formed during the acylation reaction.

Acylation after proteolysis of derivatized proteins by 2,2′-dipyridyl disulfide was also examined. After proteolysis of derivatized proteins, the acylation reaction was executed in the same manner described above except that the product was buffer exchanged into the binding buffer containing 0.1 *M* acetate (pH 5.0), 50 m*M* EDTA with 0.5 *M* K_2SO_4 and 6 *M* urea. Capture on Thiolpropyl Sepharose 6B was as described above.

2.9. Reversed-phase chromatography

All chromatographic steps were performed using a BioCAD Micro-Analytical Workstation from Applied Biosystems. Reversed-phase chromatography on a 250×4.6 mm I.D. PepMap C_{18} column (Applied Biosystems) was used to separate peptides attained from cysteine selection. Either a 60 min or 2 h linear gradient from 1% ACN, 0.1% TFA in water to 70% ACN, 0.1% TFA in water was used to fractionate peptide mixtures. Fractions were manually collected and lyophilized before analysis by MALDI-MS.

2.10. MALDI time-of-flight (TOF) MS

MALDI-MS was performed using a Voyager DE-RP BioSpectrometry workstation from Applied Biosystems. Samples were prepared by mixing a 1 μl sample with 1 μl of matrix solution. The matrix solution was ACN–water (50:50) containing 1% TFA saturated with α-cyano-4-hydroxycinnamic acid. A 1 μl matrix volume was spotted into a well of the MALDI sample plate on top of samples and allowed to air-dry before being placed in the mass spectrometer. All peptides were analyzed in the reflector, positive ion mode by delayed extraction using an accelerating voltage of 20 kV. External calibration was achieved using a mixture of standard peptides containing angiotensin I (M_r 1296.68) and insulin B chain (M_r 3494.65).

2.11. MS–MS

All MS–MS experiments, using electrospray ioni-

zation analyses, were carried out on a FinniganMAT LCQ (Thermoquest, San Jose, CA, USA) mass spectrometer system. The electrospray needle voltage was set at 3.5 kV, the heated capillary voltage was set to 10 V and the capillary temperature 225°C. Typical background source pressure was $1.2 \cdot 10^{-5}$ Torr (1 Torr = 133.322 Pa). The sample flow-rate was approximately 10 μl/min. The drying gas was nitrogen. The LCQ was scanned to 2000 u for these experiments. The sample was dissolved in acetonitrile and water.

Helium was introduced into the system to an estimated pressure of 1 millitorr to improve trapping efficiency and also act as the collision gas during the CID experiments. The collision energy was set to 32.5% of the maximum available from the 5 V tickle voltage, with an 18 mass unit isolation window.

3. Results and discussion

3.1. Covalent chromatography of cysteine containing peptides

Covalent chromatography of sulfhydryl containing proteins has been achieved by covalent bond formation with free thiol groups [44] or through thiol–disulfide interchange [37–42]. As noted above, neither of these procedures can be used to select cysteine containing peptides when their sulfhydryl groups were alkylated before proteolysis. An alternative strategy is required for covalent chromatography of cysteine containing tryptic peptides from complex mixtures of proteins.

One solution would be to tryptic digest proteins captured through either of the procedures noted above and then recover immobilized cysteine peptides by reductive cleavage of the disulfide bridge coupling them to the column [45,46]. This approach has many disadvantages. One is that immobilized proteins are more difficult to digest than those in solution. Another is that digestion will be less complete because steric limitations will prevent complete proteolysis. Still another issue is that the structure of some proteins will cause immobilization to occur more readily on one side of the protein than another. Digestion is more easily achieved in solution or an immobilized enzyme column than when protein analytes are immobilized at a surface. There is also the problem that not all the cysteine residues in a protein will be involved in covalent immobilization, and those that are will likely be random. There is a high probability this procedure would not give a representative sample of cysteine containing tryptic peptides from all the proteins in a sample.

3.1.1. Disulfide bond cleavage

Obviously, the most suitable procedure would be one in which proteins were both derivatized and predigested to facilitate covalent chromatography. The procedure outlined in Fig. 1. was chosen for this reason. This procedure has the advantage that disulfide bridges in proteins are disrupted by 2,2′-dipyridyl disulfide to facilitate proteolysis without precluding covalent capture through disulfide exchange later as demonstrated in the cases of serum albumin and ribonuclease [47]. The fact that 2,2′-dipyridyl disulfide is sparingly soluble and disrupts the structure of digestive enzymes later in proteolysis is an issue that must be confronted in this method. A solid-phase extraction technique was used to deal with this problem. Following initial treatment of protein mixtures with 0.1 *M* 2,2′-dipyridyl disulfide (pH 5), samples were adsorbed on a reversed-phase packing material, Poros 50 R2 in a 250×4.6 mm I.D. column in this case, and washed continuously with fresh 1–2 m*M* 2,2′-dipyridyl disulfide (pH 5). This allows samples to be treated with a large quantity of the sparingly soluble reagent to ensure complete activation of thiols and excess reagent to be washed away at the end of the process. Because the thiopyridone released in the disulfide interchange absorbs strongly at 343 nm, the reaction is easily monitored. Proteins activated in this way were eluted

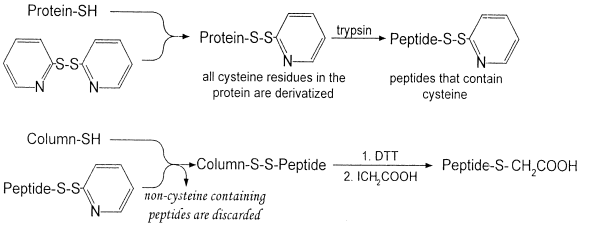

Fig. 1. Procedure for selecting cysteine-containing peptides by covalent chromatography. Details of the procedure are described in the Experimental section.

from the reversed-phase column with acetonitrile–water (70:30) containing 0.1% TFA and the solvent evaporated.

3.1.2. Proteolysis

Proteins activated with 2,2′-dipyridyl disulfide were dissolved in a Tris buffer (pH 8.0) containing 2 *M* urea and trypsin was added at a 1:50 (w/w) ratio relative to protein content. Digestion was allowed to proceed overnight.

3.1.3. Selection of cysteinyl peptides

The tryptic digest was buffer exchanged into 0.1 *M* acetate (pH 5.0) containing 0.5 *M* K_2SO_4 with 6 *M* urea and the solution added to reduced thiopropyl Sepharose gel. Because of the large excess of thiol groups in the gel, cysteine containing peptides were captured by the solid phase through disulfide interchange. The release of thiopyridone ($\lambda_{max} = 343$ nm) during this batch process provides a convenient way to monitor binding. Washing was continued until absorbance at 343 nm returned to background.

Immobilized cysteine containing peptides were released from the thiopropyl Sepharose gel with 25 m*M* DTT (pH 7.5) containing 1 m*M* EDTA. The released peptides were alkylated with iodoacetic acid before they were fractionated by reversed-phase chromatography.

3.1.4. Alkylation of cysteinyl peptides

Sulfhydryl groups of cysteine containing peptides released from the column were alkylated to preclude oxidation and the associated disulfide bond formation between peptides. As with proteins, iodoacetate, iodoacetamide, or vinyl pyridine may be used. Iodoacetate was used in all the studies outlined below.

3.2. Assessing capture efficiency

An issue in a procedure designed to capture specific peptides is selectivity. Capture selectivity with the procedure described above was accessed using isotopic labeling. MALDI mass spectra of cysteine containing peptides alkylated with a mixture of bromoacetic acid and deuterated bromoacetic acid appear as doublets separated by 2 u in the case of a peptide with a single cysteine residue. Addition of each cysteine residue to a peptide increases the mass difference in the doublet by 2 u. The need for exact mass analysis is diminished with this strategy because all doublets in the MALDI mass spectrum of the same peak height ratio will have come from cysteine containing peptides. Bromoacetic acid was used instead of the more efficient alkylating agent, iodoacetic acid, because of its commercial availability.

Lysozyme was selected as the model protein for studying capture specificity because it is small (14 000 molecular mass) and the proportion of cysteine containing peptides is high. Of the tryptic peptides in this proteins having a molecular mass greater than 500 Da, four out of 10 contain cysteine. Immediately after release from the thiol sorbent, peptides were alkylated with a mixture of bromoacetic acid and deuterated bromoacetic acid and the reaction mixture was fractionated by reversed-phase chromatography (Fig. 2). Each fraction was analyzed by MALDI-MS (Table 1). The objective here was to show that all peptides captured by the column contained cysteine. Three out of the four expected cysteine containing peptides were selected. There were also some peptides produced by trypsin miscleavage. Doublets, as well as single peaks were identified by matching their molecular mass with theoretical values. Mass spectra of some of the detected double peaks are shown in Fig. 3A–C.

The data showed that covalent chromatography selected most cysteine containing peptides, although

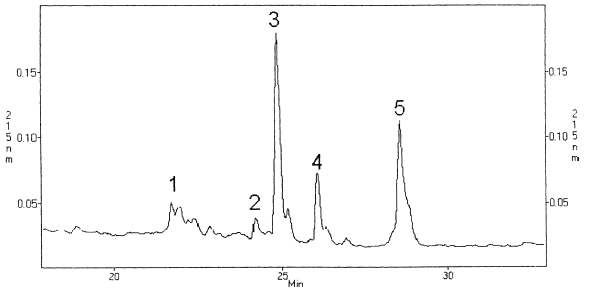

Fig. 2. Reversed-phase chromatogram of cysteine containing peptides isolated from lysozyme after covalent chromatography. Peptides were eluted from a 250×4.6 mm I.D. C_{18} column with a 60 min linear gradient ranging from 1% ACN in 0.1% aqueous TFA to 70% ACN in 0.1% aqueous TFA at a flow-rate of 1 ml/min. Five fractions were collected and analyzed by MALDI-TOF-MS.

Table 1
MALDI sequence analysis of tryptic peptides derived from lysozyme after covalent chromatography

Fraction	Theoretical mass	Mass observed	Error		Sequences from database
			Δ	(%)	
1	1050.5	1050.6[a], 1052.6[b]	0.1	(0.01)	CELAAAMKR[d]
2	994.4	994.6[a], 996.6[b]	0.2	(0.02)	WWCNDGR
3	936.4	936.7[c]	0.3	(0.03)	WWCNDGR
	1434.6	1434.4[c]	−0.2	(0.01)	WWCNDGRTPGSR[d]
4	1334.6	1334.6[a], 1336.6[b]	0	(0)	CKGTDVQAWIR[d]
5	1268.6	1268.6[c]	0	(0)	GYSLGNWVCAAK
	1326.6	1326.6[a], 1328.6[b]	0	(0)	GYSLGNWVCAAK

Δ indicates the difference between experimental (m/z value) and theoretical masses, including % error.
[a] Peptides alkylated by bromoacetic acid.
[b] Peptides alkylated by deuterated bromoacetic acid.
[c] Non-alkylated peptides due to incomplete reaction of bromoacetic acid.
[d] Incomplete tryptic digestion product.

not all of them. It should be noted that this was achieved with peptides containing one cysteine. In the digest of cell lysates, there will also be peptides with two or more cysteines, they could be more difficult to release. The existence of single peaks in the MALDI mass spectra is due to low reactivity of bromoacetic acid towards cysteine. However, the masses of these peptides matched non-alkylated peptides containing cysteine. No non-cysteine containing peptides were detected.

3.3. Quantification by MALDI-MS

Internal standards are frequently used in quantification. When a standard substance very similar to an analyte, but distinguishable in some manner, is added in a known amount to a complex mixture, analyte concentration can readily be determined by comparison to that of the internal standard. It is important that the internal standard be as similar to the analyte in chemical properties as possible so they will behave the same way in all steps of the analysis. Application of this approach to tryptic peptide based proteomics presents a dilemma in that thousands of analytes of unknown structure are being examined. The question is how to create internal standards for thousands of unknowns. When one is only interested in the relative difference in analyte concentration between two samples, the task becomes much simpler. A similar problem arises in the analysis of expression at the mRNA (cDNA) level. In this case, the dilemma was solved by labeling all the species in control and experimental samples with similar, but distinguishable fluorophors, mixing the two samples, and determining the fluorophor ratio after purification of each polynucleotide specie [48,49]. Similar strategies are now being described in proteomics using heavy isotope labeling of peptides [23,25]. An objective of the work described below was to examine the efficacy of deuterated succinic anhydride in creating internal standards for peptide quantification.

All tryptic peptides, except those that are N-terminally blocked, have primary amino groups. Based on the fact that primary amino groups react readily with succinic anhydride, it is possible to derivatize peptides containing either an N-terminal group, lysine with a free ε-amino group, or both with succinic anhydride (Fig. 4). Peptides in experimental samples were derivatized with succinic anhydride whereas peptides from control samples were treated with deuterated succinic anhydride as a means to examine the relative ratio of peptide species between samples (Fig. 5).

Peptides containing single arginine and lysine residues, respectively were split into lots labeled ‘‘control’’ and ‘‘experimental’’. The experimental sample was treated with succinic anhydride and the control sample with deuterated succinic anhydride in both cases. The MALDI mass spectrum of the arginine containing peptide appears as a doublet, separated by 4 u as seen in Fig. 6. Succinic anhydride reacted with the amino terminus of the peptide but not arginine. In contrast, the mass

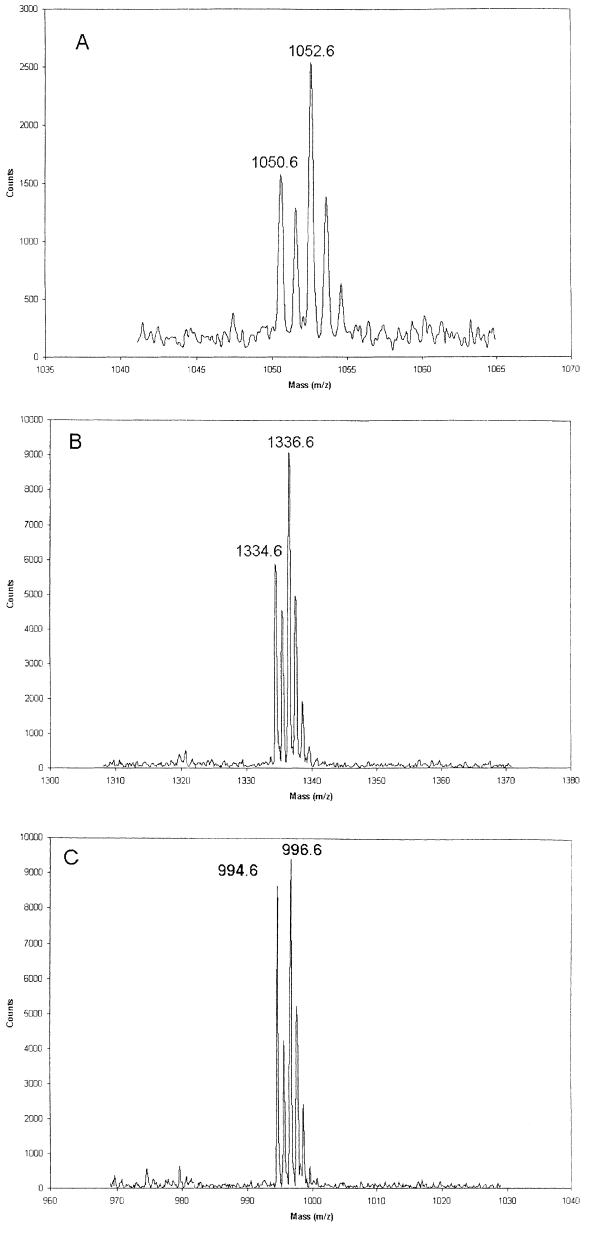

Fig. 3. MALDI-TOF mass spectra of tryptic peptides from lysozyme. These peptides were isolated by covalent chromatography and alkylated with an equimolar mixture of bromoacetic acid and deuterated bromoacetic acid. The labeled peaks were derived principally from the carbon-12 isotopic forms of the peptide derivatized with bromoacetic acid and deuterated bromoacetic acid, respectively. The sequence of the peptide in panel A is CELAAAMKR, that in B is CKGTDVQAWIR, and the peptide in panel C is WWCNDGR.

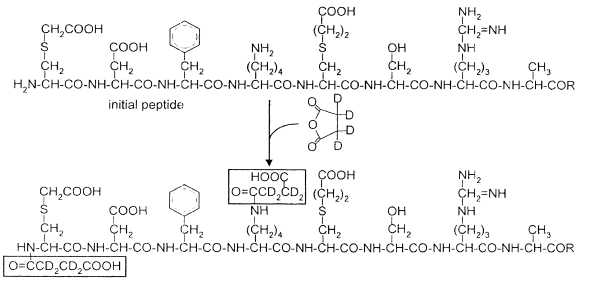

Fig. 4. Acylation of a peptide at primary amino residues by deuterated succinic anhydride.

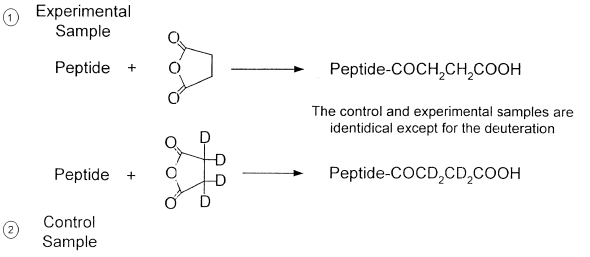

Fig. 5. Differential acylation of peptides from control and experimental samples with succinic anhydride and deuterated succinic anhydride. Details of the reaction are described in the Experimental section.

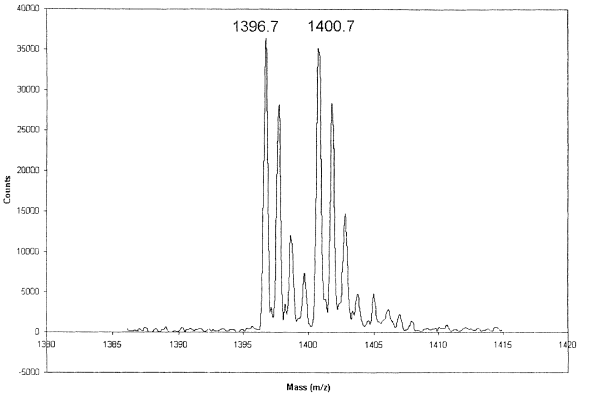

Fig. 6. MALDI-TOF mass spectrum of an arginine containing peptide NH_2–Asp–Arg–Val–Tyr–Ile–His–Pro–Phe–His–Leu–OH labeled with succinic anhydride and deuterated succinic anhydride. The labeled peaks were derived principally from the carbon-12 isotopic forms of the peptide derivatized with succinic anhydride and deuterated succinic anhydride, respectively.

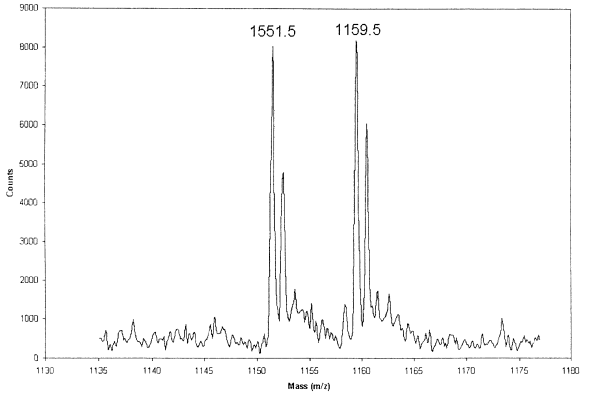

Fig. 7. MALDI-TOF mass spectrum of a lysine containing peptide NH_2–Pro–Thr–His–Ile–Lys–Trp–Gly–Asp–OH labeled with succinic anhydride and deuterated succinic anhydride. The labeled peaks were derived principally from the carbon-12 isotopic forms of the peptide derivatized with succinic anhydride and deuterated succinic anhydride, respectively.

spectrum of the lysine containing peptide appears as a doublet separated by 8 u (Fig. 7). Clearly, the ϵ-amino group of lysine reacts with the alkylating agent. A synthetic peptide containing both lysine and serine was also examined. The mass spectrum of the reaction product showed two doublets separated by 8 and 12 u, respectively (Fig. 8). Treatment of the derivatized peptide with hydroxylamine at pH 11–12

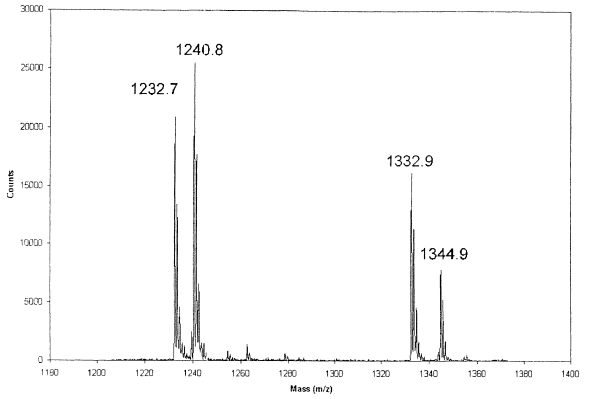

Fig. 8. MALDI-TOF mass spectrum of a lysine and serine containing peptide NH_2–Lys–Arg–Pro–Pro–Gly–Phe–Ser–Pro–Phe–OH labeled with succinic anhydride and deuterated succinic anhydride. The labeled peaks were derived principally from the carbon-12 isotopic forms of the peptide derivatized with succinic anhydride and deuterated succinic anhydride, respectively. The reaction mixture was not treated with hydroxylamine.

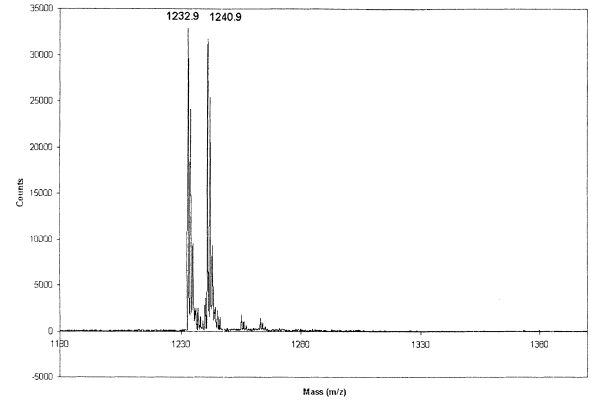

Fig. 9. MALDI-TOF mass spectrum of a lysine and serine containing peptide NH_2–Lys–Arg–Pro–Pro–Gly–Phe–Ser–Pro–Phe–OH labeled with succinic anhydride and deuterated succinic anhydride. The labeled peaks were derived principally from the carbon-12 isotopic forms of the peptide derivatized with succinic anhydride and deuterated succinic anhydride, respectively. Note that this peptide was used in Fig. 8. The reaction mixture was treated with hydroxylamine.

for a few minutes eliminated the higher mass doublet separated by 12 u (Fig. 9). These results are interpreted to mean that the hydroxyl group on serine can be esterified by succinic anhydride but that the ester is readily hydrolyzed by hydroxylamine. Phenolate and imidazolyl groups will also be partially acylated and the product hydrolyzed by the treatment noted above.

3.4. Sequencing by MS–MS

Peptide identification is greatly facilitated by sequencing. The objective of the experiment described below was to examine the utility of succinate labeling in MS–MS based sequencing. It has been noted above that peptides containing lysine are derivatized at their C-termini with succinic anhydride whereas those with arginine are not. Peptides with these amino acids at the C-termini were chosen for analysis to analyze the impact of this difference on their spectra. Following the same approach as in previous experiments, samples were differentially labeled with either succinic anhydride or deuterated succinic anhydride. After reaction, the control and experimental samples were combined and sequenced by MS–MS as seen in Figs. 10 and 11.

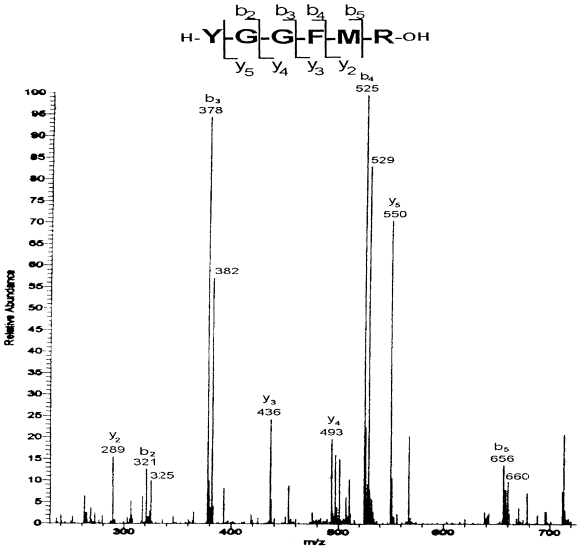

Fig. 10. Tandem mass spectrum of the peptide NH_2–Tyr–Gly–Gly–Phe–Met–Arg–OH derivatized by succinic anhydride and deuterated succinic anhydride, respectively.

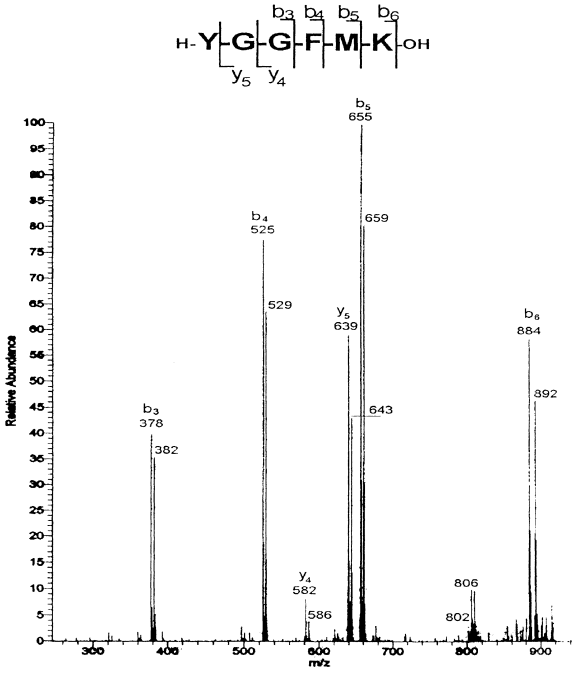

Fig. 11. Tandem mass spectrum of the peptide NH_2–Tyr–Gly–Gly–Phe–Met–Lys–OH derivatized by succinic anhydride and deuterated succinic anhydride, respectively.

The mass spectrum of the peptide chosen with arginine at its C-terminus had an initial molecular mass of 729.85. After reaction with succinic anhydride, either 100 or 104 u was added depending on the deuterium content of the labeling reagent. An 18 u window was chosen for parent ion selection so that both the deuterated and non-deuterated parents would be selected. It is seen in Fig. 10 that both b and y ions were detected and that all the b ions appear as doublets, such as m/z 321 and 325 for b_2, m/z 378 and 382 for b_3, m/z 525 and 529 for b_4 and m/z 656 and 660 for b_5. In contrast, y ions appear as single peaks. This is as would be predicted. Since succinic anhydride labels the amino terminus of the peptide and do not react with arginine, all ions carrying the amino terminus will appear in spectra as doublets separated by 4 u. The sequence of the peptide in Fig. 10 can easily be seen to be NH_2–Tyr–Gly–Gly–Phe–Met–Arg–OH. Because of the absence of the amino terminus in y ions, they will appear as singlets. This experiment confirms that beyond not interfering with fragmentation patterns in MS–MS, isotope labeling can assist in assigning specific structural elements to ions.

The mass spectrum of the peptide chosen with a lysine bearing C-terminus had an initial molecular mass of 701.84 (Fig. 11). Because peptides with a free amino terminus and a single lysine residue have two primary amino groups, labeling with succinic anhydride will cause mass to increase by either 200 or 208 u depending on the deuterium content of the anhydride. Again a parent ion selection window of 18 u was used. It is seen that the mass spectrum contains a doublet of peaks at 902 and 910 when using a full mass scan. As in Fig. 10, most of the ions in this spectrum were b type ions appearing as doublets separated by 4 u. But unlike b ion doublets, those from y type ions are of lower intensity, except for the y_5 ion at m/z 639 and 643. The sequence of the peptide in this case can be seen to be NH_2–Tyr–Gly–Gly–Phe–Met–Lys–OH. The tandem mass spectrometry experiments confirm both the specificity of the labeling and the value of labeling in spectral interpretation.

3.5. Protein expression in E. coli

Bacteria with inducible plasmids provide good

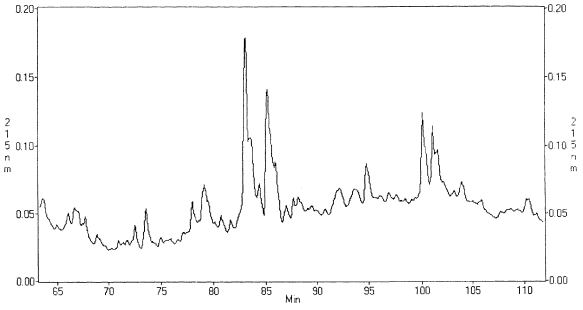

Fig. 12. Reversed-phase chromatogram of acylated cysteine containing peptides isolated from *E. coli* tryptic digest using covalent chromatography. Peptides were eluted from a 250×4.6 mm I.D. C_{18} column with a 120 min linear gradient ranging from 1% ACN in 0.1% aqueous TFA to 70% ACN in 0.1% aqueous TFA at a flow-rate of 0.5 ml/min.

systems to test the efficacy of the methods described above. A variety of plasmids (vectors) are available in *E. coli* that may be induced by a small molecule to synthesize a specific protein, as in the production of therapeutic proteins in biotechnology. Up-regulation of this single protein must then be detected in the background of all the other proteins produced by the cell. The protein background will likely be a thousand or more proteins based on the fact that *E. coli* can potentially produce 4288 different proteins.

The *E. coli* vector pC/PAD used in this study carried a genetic construct for the expression of PAD that could be induced with IPTG. Lysate from *E. coli* grown in the absence of IPTG served as the control whereas lysate taken from a culture shortly after induction with IPTG provided the experimental

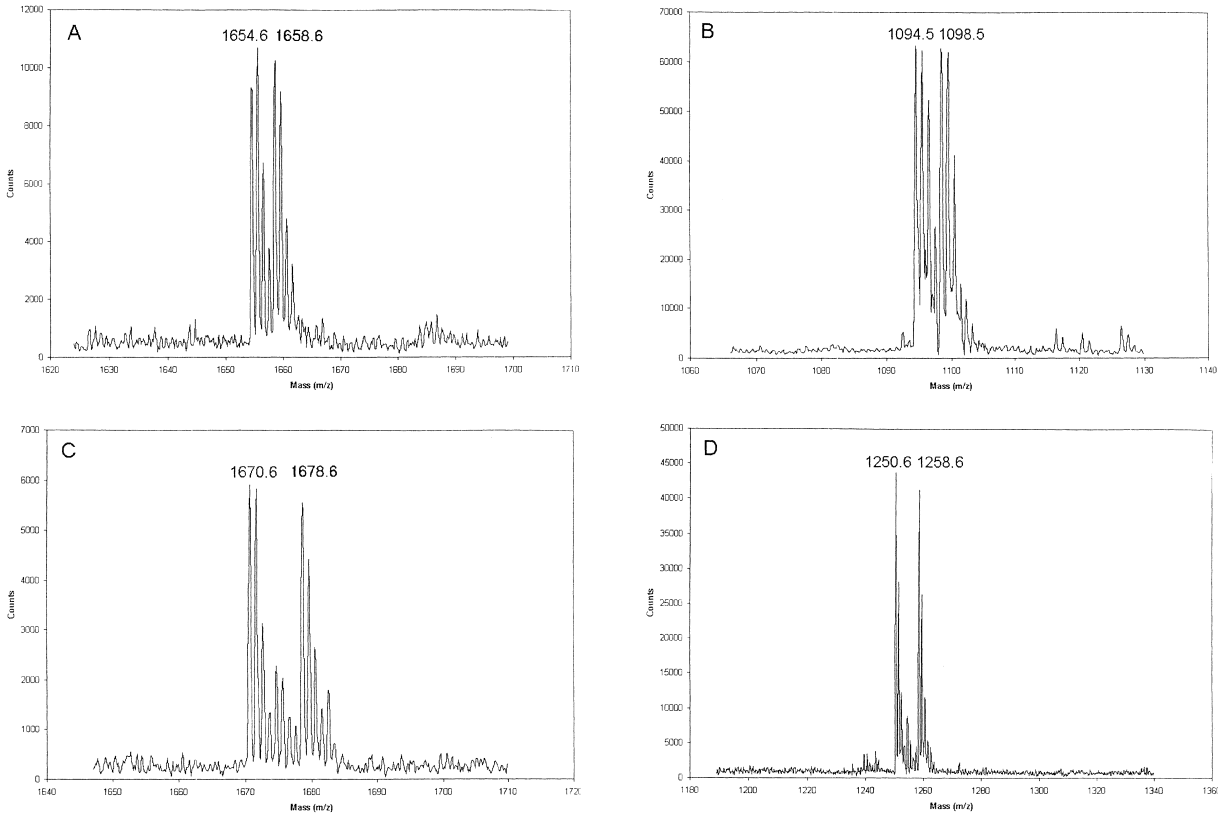

Fig. 13. MALDI-TOF mass spectra of acylated peptides isolated from an *E. coli* tryptic digest after selection by covalent chromatography. The fact that the ratio of non-deuterated to deuterated peptide is roughly 1 in analytes A–D indicates their concentration did not change between the control and experimental samples.

sample. Sampling time after induction was kept short (2 h) to assure that the amount of induced protein was low and detection would be more difficult. Experimental and control samples were treated separately through the steps of activation with 2,2′-dipyridyl disulfide, tryptic digestion, and derivatization with succinic anhydride and deuterated succinic anhydride, respectively. Subsequent to labeling the samples were mixed and separated by a combination of covalent and reversed-phase chromatography. The reversed-phase chromatogram after selection by covalent chromatography is seen in Fig. 12.

Fractions collected from the reversed-phase column were analyzed by MALDI-MS (Fig. 13A–D). The expected doublets varying by 4 u (Fig. 13A and B) and 8 u (Fig. 13C and D) are clearly seen in the spectra. It should also be seen that in many of these doublets the peak height ratio is roughly one. These are proteins whose synthesis is directed by nuclear DNA of the organism, not the plasmid. Spectra from two of the collected fractions showed doublet peak height ratios that would indicate the peptides came from protein(s) that had been up-regulated (Fig. 14A and B). Mass analysis indicates that their masses match those of cysteine containing tryptic peptides from phenyacetaldehyde dehydrogenase. Peak height ratios from the peptides in Fig. 14A and B obtained by integrating the total areas in the isotopic peak clusters suggest 2.4- and 2.5-fold induction, respectively. It is expected that the degree of regulatory change of peptides from the same protein should be the same.

This experiment demonstrates the efficacy of the selection and labeling methods described above for quantification of proteins in complex mixtures, such as cellular lysates.

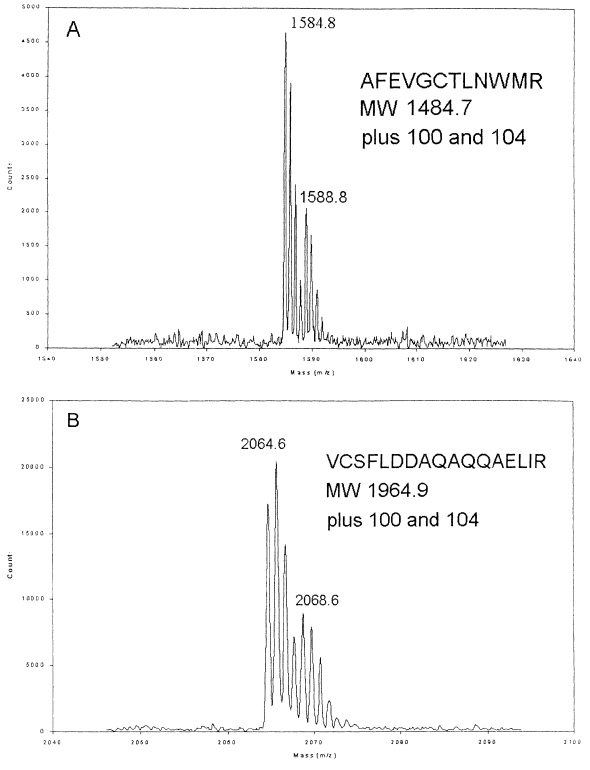

Fig. 14. (A, B) MALDI-TOF mass spectra of two acylated cysteine containing peptides isolated from an *E. coli* tryptic digest using covalent chromatography. Note that the ratio of non-deuterated to deuterated peptide in both cases is 2.4. This indicates they were both up-regulated in the experimental sample. MW= Molecular mass.

4. Conclusion

Based on the data presented above, it is concluded that cysteine containing peptides can be selected from tryptic digests of complex protein mixtures using covalent chromatography. By combining the steps of covalent chromatography, reversed-phase chromatography and mass spectrometry, tryptic digests of complex protein mixtures were greatly simplified. Also by coupling isotopically labeled internal standard analysis with signature peptide selection, quantification of proteins expressed in *E. coli* was demonstrated.

It is concluded that procedures designed to select small numbers of peptides from each protein based on unique structural elements of the peptides are a great asset in reducing the complexity of both samples and database searches. When coupled with isotope labeling to quantify changes of protein concentration between samples, the procedures described in this paper will provide a powerful new technique for the study of cellular dynamics in cases of stress from the environment and disease.

Acknowledgements

The authors greatly acknowledge financial support from the National Institute of Health (grant number GM59996) and Applied Biosystems.

References

[1] B.E. Chong, D.B. Wall, D.M. Lubman, S.J. Flynn, J. Rapid Commun. Mass Spectrom. 11 (1997) 1900.
[2] J.A. Loo, J. Brown, G. Critchley, C. Mitchell, P.C. Andrews, R.R. Ogorzalek Loo, Electrophoresis 20 (1999) 743.
[3] G.J. Opitek, S.M. Ramirez, J.W. Jorgenson, M.A. Moseley, Anal. Biochem. 258 (1998) 349.
[4] L. Yang, C.S. Lee, S.A. Hofstadler, L. Pasa Tolic, R.D. Smith, Anal. Chem. 70 (1998) 3235.
[5] P.K. Jensen, L. Pasa Tolic, G.A. Anderson, J.A. Horner, M.S. Lipton, J.E. Bruce, R.D. Smith, Anal. Chem. 71 (1999) 2076.
[6] L. Pasa Tolic, P.K. Jensen, G.A. Anderson, M.S. Lipton, K.K. Peden, S. Martinovic, N. Tolic, J.E. Bruce, R.D. Smith, J. Am. Chem. Soc. 121 (1999) 7949.
[7] P.H. O'Farrel, J. Biol. Chem. 250 (1975) 4007.
[8] E. Mortz, O. Vorm, M. Mann, P. Roepstorff, Biol. Mass Spectrom. 23 (1994) 249.
[9] H.H. Rasmussen, E. Mortz, M. Mann, P. Roepstorff, J.E. Celis, Electrophoresis 15 (1994) 406.
[10] K.R. Clauser, S.C. Hall, D.M. Smith, J.W. Webb, L.E. Andrews, H.M. Tran, L.B. Epstein, A.L. Burlingame, Proc. Natl. Acad. Sci. USA 92 (1995) 5072.
[11] M.R. Wilkins, C. Pasquali, R.D. Appel, K. Ou, O. Golaz, J.C. Sanchez, J.X. Yan, A.A. Gooley, G. Hughes et al., Biotechnology 14 (1996) 61.
[12] J. Qin, D. Fenyõ, Y. Zhao, W.W. Hall, D.M. Chao, C.J. Wilson, R.A. Young, B.T. Chait, Anal. Chem. 69 (1997) 3995.
[13] M. Fountoulakis, H. Lagen, S. Evers, C. Gray, B. Takacs, Electroproresis 18 (1997) 1193.
[14] T. Nystrom, N.H. Albertson, K. Flardh, S. Kjellegerg, FEMS Microbiol. Ecol. 74 (1990) 129.
[15] F.C. Neidhardt, R.A. VanBogelen, Biochem. Biophys. Res. Commun. 100 (1981) 894.
[16] F. Foret, O. Mueller, J. Thorne, W. Goetzinger, B.L. Karger, J. Chromatogr. A 716 (1995) 157.
[17] W. Weinmann, C.E. Parker, L.J. Deterding, D.I. Papac, J. Hoyes, M. Przybylski, K.B. Tomer, J. Chromatogr. A 680 (1994) 353.
[18] B. Yeung, T.J. Porter, J. Vath, Anal. Chem. 69 (1997) 2510.
[19] D. Figeys, A. Ducret, R. Aebersold, J. Chromatogr. A 763 (1997) 295.
[20] K.L. Walker, R.W. Chiu, C.A. Monnig, C.L. Wilkins, Anal. Chem. 67 (1995) 4197.
[21] G.A. Valaskovic, N.L. Kelleher, F.W. McLafferty, Science 273 (1996) 199.
[22] M.D. Bauer, S. Yiping, F. Wang, in: 10th Symposium of the Protein Society, Techniques in Protein Chemistry VIII, Academic Press, San Diego, CA, 1997, p. 37.
[23] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[24] M.P. Washburn, D. Wolters, J. Yates, Nat. Biotechnol. 19 (2001) 242.
[25] J. Ji, A. Chakraborty, M. Geng, X. Zhang, A. Amini, M. Bina, F. Regnier, J. Chromatogr. B 745 (2000) 197.
[26] M. Wilcheck, T. Miron, Biochim. Biophys. Acta 278 (1972) 1.
[27] G. Barnard, E. Bayer, M. Wilchek, Y. Amir-Zaltsman, F. Kohen, Methods Enzymol. 133 (1986) 284.
[28] S. Yoshitake, Y. Yamada, F. Ishikawa, R. Masscyeff, Eur. J. Biochem. 101 (1979) 395.
[29] K. Zhu, R. Bressan, P. Hasegawa, L. Murdock, FEBS Lett. 390 (1996) 271.
[30] A. Apffel, J. Chakel, W. Hancock, C. Souders, T. Timkulu, E. Pungor, J. Chromatogr. A 732 (1996) 27.
[31] D.C. Schriemer, T. Yalcin, L. Li, Anal. Chem. 70 (1998) 1569.
[32] A. Quesnel, A. Delmas, Y. Trudelle, Anal. Biochem. 231 (1995) 182.
[33] P.M. Fischer, M.E.H. Howden, J. Immunoassay 11 (1990) 311.
[34] Y. Nakagawa, T. Yip, M. Belew, J. Porath, Anal. Biochem. 75 (1981) 168.
[35] M.C. Posewitz, P. Tempst, Anal. Chem. 71 (1999) 2883.
[36] K. Brocklehurst, J. Carlsson, M.P.J. Kierstan, E.M. Crook, Biochem. J. 133 (1973) 573.
[37] K. Brocklehurst, J. Carlsson, M.P.J. Kierstan, E.M. Crook, Methods Enzymol. 34 (1974) 531.
[38] J. Carlsson, A. Svenson, FEBS Lett. 42 (1974) 183.
[39] F. Fex, C.B. Laurell, E. Thulin, Eur. J. Biochem. 75 (1977) 181.
[40] F. Cormier, C. Charest, C. Dufresne, Biotech. Lett. 11 (1989) 797.
[41] B.S. Baines, K. Brocklehurst, P.R. Carey, M. Jarvis, E. Salih, A.C. Storer, Biochem. J. 233 (1986) 119.
[42] T.P. Bradshaw, R.B. Dunlop, Biochim. Biophys. Acta 1163 (1993) 165.
[43] A.M. Crestfield, S. Moore, W.H. Stein, J. Biol. Chem. 238 (1963) 622.
[44] N.F. Tabachnik, P. Blackburn, C.M. Peterson, A. Cerami, J. Pharmacol. Exp. Ther. 220 (1982) 243.
[45] T.A. Egorov, A. Svenson, L. Rydén, J. Carlsson, Proc. Natl. Acad. Sci. USA 72 (1975) 3029.
[46] L. Ryden, H. Norder, J. Chromatogr. 215 (1981) 341.
[47] A. Svenson, J. Carlsson, D. Eaker, FEBS Lett. 73 (1977) 171.
[48] G. Yershov, V. Barsky, A. Belgovskiy, E. Kirillov, E. Kreindlin, I. Ivanov, S. Parinov, D. Guschin, A. Drobishev, S. Dubiley, A. Mirzabekov, Proc. Natl Acad. Sci. USA 93 (1996) 4913.
[49] M. Schena, D. Shalon, R. Heller, A. Chai, P.O. Brown, R.W. Davis, Proc. Natl. Acad. Sci. USA 93 (1996) 10614.