ELSEVIER

Journal of Chromatography A, 949 (2002) 153–162

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Quantitative proteomics strategy involving the selection of peptides containing both cysteine and histidine from tryptic digests of cell lysates

Shihong Wang, Xiang Zhang, Fred E. Regnier*

Department of Chemistry, 3164A Brown Building, Purdue University, West Lafayette, IN 47907-1393, USA

Abstract

This paper describes a procedure for quantitative proteomics that selects peptides containing both cysteine and histidine residues from tryptic digests of cell lysates. Cysteine-containing peptides were selected first by covalent chromatography using thiol disulfide exchange. Following the release of cysteine-containing peptides from the covalent chromatography column with reductive cleavage, histidine-containing peptides were captured by passage through an immobilized metal affinity chromatography column loaded with copper. Quantification was achieved in a four-step process involving (i) differential labeling of control and experimental samples with isotopically differing forms of succinic anhydride, (ii) mixing the two globally labeled samples, (iii) fractionating the labeled peptides by reversed-phase liquid chromatography, and (iv) determining the isotope ratio in individual peptides by mass spectrometry. The results of these studies indicate that by selecting peptides containing both cysteine and histidine, the complexity of protein digests could be substantially reduced. Up-regulated proteins from plasmid bearing *Escherichia coli* that had been induced with isopropyl β-thiogalacto-pyranoside were identified and quantified by the global internal standard technology (GIST) described above. Database searches were greatly simplified because the number of possible peptide candidates was reduced more than 95%.

Keywords: Immobilized metal affinity chromatography; *Escherichia coli*; Isotope analysis; Peptides; Cysteine; Histidine

1. Introduction

Proteomics aims to identify all the proteins expressed by an organism at a certain time and under certain conditions. The most widely used approach to accomplish this has been two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). The excellent resolving power of two-dimensional electrophoresis is achieved through the separation of proteins on the basis of two different physicochemical properties—isoelectric point (p*I*) and molecular mass [1–3]. 2D-PAGE can resolve up to 6000 spots. Subsequent to separation, proteins are visualized by staining or fluorescent labeling, spots are excised from the gel, enzymatically digested, and the peptide fragments analyzed by mass spectrometry [4–8]. 2D-PAGE has been used in studies ranging from localizing individual proteins at a subcellular level [9] to recognizing changes in protein expression [10] and identifying target genes of regulatory proteins [11]. Among the serious limitations of 2D-PAGE are the facts that quantification is poor, it is labor intensive, automation is difficult, and it cannot be coupled directly to mass spectrometry.

*Corresponding author. Tel.: +1-317-494-1648; fax: +1-317-494-0359.

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

0021-9673/02/$ – see front matter
PII: S0021-9673(01)01509-6

Promising new techniques are emerging that replace 2D-PAGE. Most involve some combination of HPLC or CE with mass spectrometry to either create a ''virtual two-dimensional gel'' [12–15] 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 [16].

The most common approach in the chromatographic methods is to enzymatically digest all the proteins in the sample first and then fractionate the mixture by multidimensional chromatography. Ideally, the chromatographic modes should be orthogonal in separation mechanism to maximize the resolution of components. Although a variety of combinations have been used, the last step before mass spectrometry is generally reversed-phase liquid chromatography (RPLC). Ion-exchange chromatography [17,18] and affinity chromatography have been most frequently reported as the first dimension separation mode. The affinity methods that have been reported generally target a particular type of amino acid, such as histidine [19,20] or cysteine [21,22], or specific post-translational modifications.

The sequence of peptides is closely related to the parent proteins from which they are derived. Often proteins can be identified from a single, or a small number of peptides. This means there is redundant information in the proteolytic digest. Only a small number of peptides in a digest are actually needed to identify all the proteins in a sample. In the ideal case, only one peptide would be needed per protein, so long as that peptide is suitable for mass spectrometry. This single peptide could be used as both a quantitative and qualitative signature for identification of its parent protein. Several studies have been reported on this signature peptide approach to proteomics [19,23,24]. Complex mixtures were fractionated to generate peptides that were then used as analytical surrogates for the proteins from which they were derived. Selection of surrogates targeted low abundance amino acid residues such as histidine [19,20] or cysteine [21,22], or peptides carrying specific post-translational modifications such as glycosylation [23,24] or phosphorylation sites [25]. These studies found that by selecting signature peptides it is possible to greatly simply complex samples while at the same time still identify their parent proteins [19,23,24].

Although of low abundance, multiple histidine and cysteine peptides are generally selected from each protein. The question to be examined here is whether it is possible to be even more selective and obtain peptides of even lower abundance. Analyses of DNA databases suggest that peptides containing both cysteine and histidine are of even lower abundance than either alone, but that up to 80% of all proteins would have one such peptide. This paper will focus on database analyses, the development of a double selection method for cysteine and histidine, and a few examples of the efficacy of this method.

2. Experimental

2.1. Materials

DL-Dithiothreitol (DTT), iodoacetic acid, L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin, urea, isopropyl β-thiogalacto-pyranoside (IPTG), protamine sulfate, guanidine hydrochloride, *N*-hydroxylamine, α-cyano-4-hydroxycinnamic acid, tris(hydroxymethyl)aminomethane (Tris base), tris(hyroxymethyl) 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 chloride, 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, 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 PepMap C_{18} column were from Applied Biosystems (Framingham, MA, USA). Thiopropyl Sepharose 6B was from Amersham Pharmacia Biotech (Piscataway, NJ, USA). A

TSK gel Chelate-5PW column (75 mm×7.5 mm I.D.) with a particle size of 10 μm and average pore size of 1000 Å was obtained from Tosohaas (Montgomeryville, PA, USA).

2.2. Escherichia coli cell culture

The expression plasmid pC/PAD with the phenylacetaldehyde 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 phenylacetaldehyde 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. E. coli lysate derivatization followed by tryptic digestion

A 10–30-mg sample of control or experimental *E. coli* lysate was individually reduced with DTT at a 40 molar excess in 6 *M* guanidine·HCl at alkaline pH [26]. The reduced sample was applied to an RPLC column (250×4.6 mm I.D.) packed with Poros 50 R2 (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 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 denaturation 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. Acylation of peptides by succinic anhydride

After proteolysis of derivatized proteins, the digestion product was buffer exchanged into phosphate buffer, pH 7–8. A 50-fold molar excess of succinic anhydride and deuterated succinic anhydride were added individually to experimental and control samples. 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. 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 for covalent chromatography afterwards.

2.6. Capture of activated thiol peptides on Thiopropyl Sepharose 6B

Dry Thiopropyl 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 digests (15–20 ml) of control and experimental samples were combined together and 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 thiopyridone (λ_{max}=343 nm) reduced to zero.

2.7. 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.

2.8. Immobilized metal affinity chromatography

A TSK gel Chelate-5PW column (75 mm×7.5 mm I.D.) with a particle size of 10 μm and average pore size of 1000 Å (Tosohaas) was used for immobilized metal affinity chromatography experiments. All chromatographic steps were performed using a BioCAD Micro-Analytical Workstation from Applied Biosystems. Before each separation, the column was washed successively with ten column volumes each of water, 50 m*M* EDTA in 50 m*M* phosphate with 500 m*M* NaCl, pH 7.5, water, 200 m*M* NaCl, 0.5 *M* NaOH, 200 m*M* NaCl and again water. Thereafter, the stationary phase was loaded with the metal by perfusing the column with 20 m*M* solution of copper sulfate in water until saturation. Subsequently, the column was rinsed with ten column volumes of 200 m*M* NaCl, water and 20 column volumes of elution buffer (50 m*M* phosphate, 500 m*M* NaCl, pH 3.9) to remove excess of copper sulfate. After the column had been re-equilibrated with 20 column volumes of initial buffer (50 m*M* phosphate, 500 m*M* NaCl, pH 7.5), sample collected after previous covalent chromatography was applied. After washing the column with more than 20 column volumes of initial buffer, the product was eluted by elution buffer. The effluent was monitored by continuously recording the absorbance at 215 nm until the absorbance decreased to baseline level.

2.9. Reversed phase liquid chromatography

RPLC on a 250×4.6 mm I.D. PepMap C_{18} column (Applied Biosystems) was used to separate peptides obtained from cysteine selection. A 60-min 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 matrix-assisted laser desorption ionization mass spectrometry (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).

3. Results and discussion

3.1. In silico analysis of the proteome in common organisms

E. coli and yeast are popular targets for

Table 1
The number and percentage of proteins and tryptic peptides from *E. coli*, yeast and humans that contain cysteine, histidine, or both residues according to genomic analyses

	Proteins						Completely digested tryptic peptides					
	E. coli		Yeast		Human		*E. coli*		Yeast		Human	
Total numbers	4289		6200		69 798		132 768		333 247		2 641 532	
His-containing	4136	96.4%	6013	97.0%	57 574	82.5%	22 935	17.3%	49 846	15.0%	444 450	16.8%
Cys-containing	3674	85.7%	5681	91.6%	62 309	89.3%	12 465	9.4%	31 008	9.3%	391 659	17.2%
His- and Cys-containing	2524	58.8%	5551	89.5%	53 945	77.3%	4695	3.5%	9386	2.8%	128 888	5.7%

proteomics studies. *E. coli* is theoretically capable of producing 4289 proteins according to genomic analyses (Table 1). Of these proteins, 96.4% would contain at least one histidine, 85.7% would have at least one cysteine, and roughly 58.8% of the proteins expressed by this organism could produce at least one tryptic peptide with both cysteine and histidine residues. Moreover, these 4289 proteins would produce 132 768 tryptic peptides if digestion were complete. Among these peptides, 17.3% contain histidine and 9.4% contain cysteine residues. Only 3.5% contain both histidine and cysteine. Thus, most of the proteins coded by the *E. coli* genome would contain either a histidine residue, a cysteine residue, or both.

Yeast is slightly more complicated, theoretically being capable of producing 6200 proteins (Table 1). Approximately 97% would yield a tryptic peptide bearing one or more histidine residues, 91.6% would have one or more cysteine residues, and 89.5% would carry both of these amino acids in at least one tryptic peptide. Only 2.8% of all peptides in a complete tryptic digest of a yeast extract would have both amino acids. It is seen that the number of proteins containing both amino acids is almost the same as the number carrying cysteine or histidine alone.

Obviously, the genome of higher animals is much more complicated. Databases suggest that humans produce at least 69 798 proteins and that tryptic digests of these proteins will be an order of magnitude more complex (Table 1). In silico tryptic digestion yields 2 641 532 peptides of which 16.8% contain histidine, 17.2% contain cysteine, and only 5.7% of the predicted peptides (i.e. 128 888) contain both amino acids. Again it is seen that the number of proteins in human samples containing both cysteine and histidine is roughly equivalent to the number of proteins found in the whole proteome (Table 1).

The fact that (i) the number of proteins in yeast and humans which yield peptides with both histidine or cysteine in the same tryptic fragment does not differ substantially from those containing only histidine or cysteine alone and (ii) the number of fragments containing both histidine and cysteine is much lower than those containing either amino acid alone, can be exploited in proteomics. It has already been demonstrated that selection of peptides containing low abundance amino acids greatly reduces sample complexity and still gives broad coverage of all the proteins in a sample [19,23,24,27]. There is the potential that double selection for peptides that contain both amino acids would reduce sample complexity by 97.2% in yeast and give 89.5% coverage of all proteins. Reduction of sample complexity and coverage are slightly lower in the case of human samples where there is a 94.3% reduction with 77.3% coverage. Finally, a 96.5% reduction in sample complexity is expected with *E. coli* and would give 58.8% coverage.

3.2. Selection strategy

Cysteine and histidine selection have already been described [19–22], but not in a double selection mode. There is the question of which methods described in the literature would be most suitable for double selection and quantification of doubly selected peptides.

Histidine selection has been achieved exclusively by immobilized affinity chromatography (IMAC) using columns loaded with copper [19]. It has been noted with Cu-IMAC of histidine-containing peptides that primary amine groups substantially in-

crease the probability of non-histidine-containing peptides binding to the column. Acylating primary amino groups, particularly those at the amino terminus of peptides, increased the selectivity of Cu-IMAC columns for histidine [19,28]. Thus it is desirable to incorporate acylation into signature peptide proteomics based on histidine.

Another issue is the choice of a labeling strategy for quantification. The two most widely used have been the global internal standard technology (GIST) and isotope-coded affinity tag (ICAT) methods. The GIST method exploits the fact that proteolysis generates amino groups in peptides that may be acylated, either with succinic anhydride or acetate [19,29]. Control samples are derivatized with one isotopic form of the acylating agent while experimental samples are treated with another isotopic form. Control and experimental samples are then mixed and individual components of the sample selected for analysis. Subsequent to the selection and further fractionation by reversed-phase liquid chromatography, relative differences in the concentration of specific peptides between samples are measured by mass spectrometry. The ICAT method is very similar, but targets cysteine [15]. In the ICAT reagent, the alkylating agent is coupled to biotin through a polyoxyethylene bridge that is either unlabeled (d_0) or labeled with eight deuterium atoms (d_8). The process of differential labeling followed by selection and quantification is similar to that of the GIST method. Either would work for double selection. The GIST approach was chosen in these studies for two reasons. One was that acylation of amino groups in the GIST method enhances histidine selection with Cu-IMAC. The second is that the differentially labeled forms of peptides are more widely separated in ICAT derivatized peptides [30]. Peptides that are labeled with deuterated reagents often elute slightly earlier in reversed-phase liquid chromatography. Although the resolution is small, it can make a difference in the accuracy of quantification.

3.3. Chromatography

Covalent chromatography of sulfhydryl-containing proteins has been achieved by covalent bond formation with free thiol groups [31] or through thiol–disulfide interchange [32–34]. A modification of this procedure has recently been used to select cysteine-containing peptides (Fig. 1) [29]. Following initial treatment of a protein mixture with 0.1 *M* 2,2′-dipyridyl disulfide (pH 5), the protein thiols are activated for thiol–disulfide interchange by converting them into 2-pyridyl disulfide groups. Because the thiopyridone released in the disulfide interchange process absorbs strongly at 343 nm, the reaction is easily monitored. After proteins were derivatized with 2,2′-dipyridyl disulfide, the mixture was 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. The tryptic digest was then buffer exchanged into 0.1 *M* acetate (pH 5.0) containing 0.5

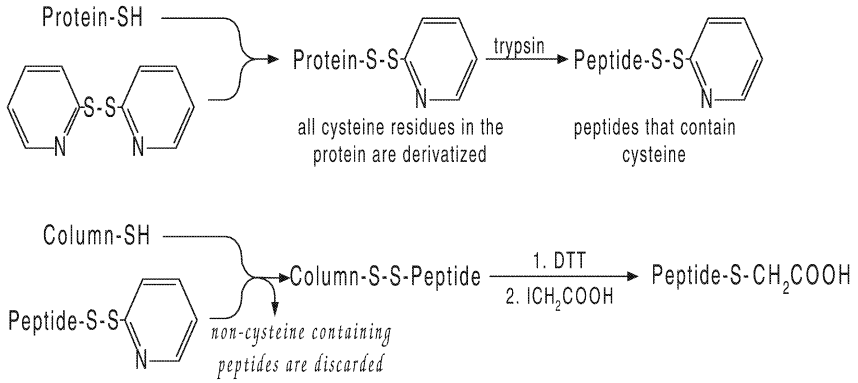

Fig. 1. Procedure for selecting cysteine-containing peptides by covalent chromatography.

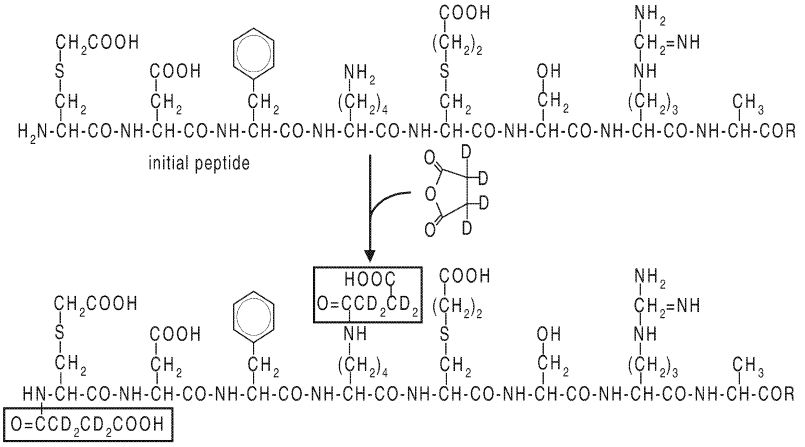

Fig. 2. Acylation of a peptide at primary amino residues by deuterated succinic anhydride.

M K_2SO_4 and 6 *M* urea and then 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 (λ_{max}=343 nm) during this batch process provides a convenient way to monitor binding. After binding was finished, the gel was washed thoroughly to eliminate non-specific binding. 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 then alkylated with iodoacetic acid before they were fractionated by reversed-phase liquid chromatography.

Quantitation of protein expression was achieved with a GIST strategy in which tryptic peptides were labeled at their amino termini and lysine residues with succinic anhydride (Figs. 2 and 3). The labeled control and experimental samples were then combined and cysteine selected. After the covalent chromatography step described above, the cysteine selection product was applied to a Cu(II) loaded IMAC column for histidine selection [13].

Although Cu(II)-IMAC is capable of resolving peptides based on histidine content, all peptides

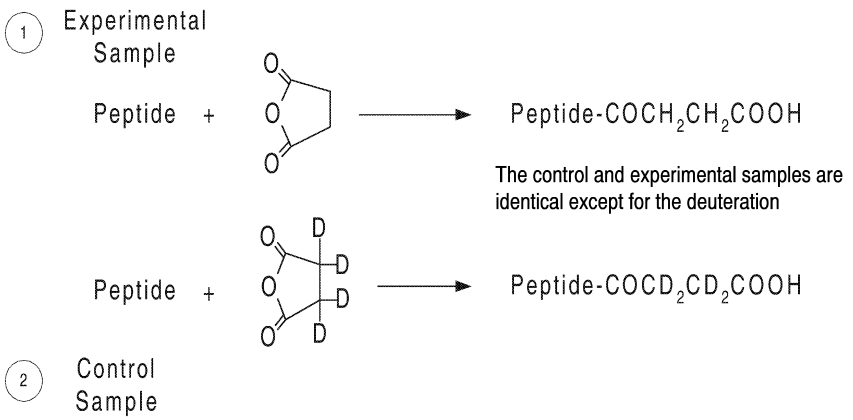

Fig. 3. Differential acylation of peptides from control and experimental samples with succinic anhydride and deuterated succinic anhydride.

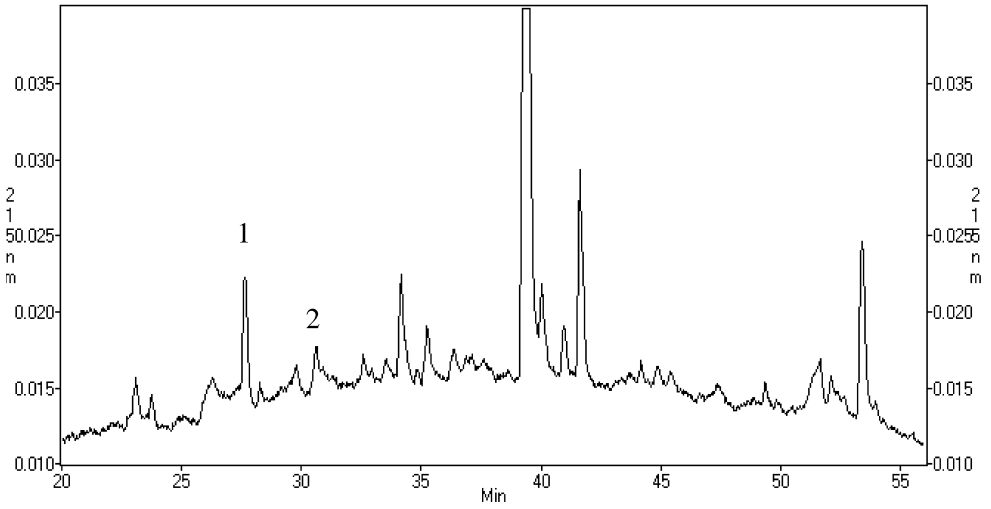

Fig. 4. Reversed-phase liquid chromatogram of acylated peptides from *E. coli* containing both cysteine and histidine tryptic peptides. The sample used in this separation was obtained using a combination of covalent chromatography and immobilized metal affinity chromatography on a copper loaded column. The protocol for this double section procedure is described in the Experimental section. 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 0.5 ml/min. Peaks 1 and 2 were analyzed by MALDI–TOF-MS, respectively, and identified through database search.

captured by the column were released and collected in a single fraction that was further fractionated by RPLC (Fig. 4). Even though this double selection procedure is expected to simplify the mixture at least six-fold and only a fraction of the theoretically possible peptides were probably expressed, the chromatogram is still complicated.

3.4. Applications

Several effluent fractions from the reversed-phase column (Fig. 4) were collected and analyzed by MALDI–MS. Fig. 5A and B shows the examples of analyzed fractions from peaks 1 and 2 in Fig. 4. The doublet in Fig. 5A showing an 8-u difference between the peaks obviously came from a peptide with a lysine residue at the C-terminus. A peak height ratio comparison between control and experimental samples indicated that the parent protein in the experimental sample was up-regulated 3.3-fold. Based on the assumption that (i) the peptide had one or more cysteine and histidine residues, (ii) it had a C-terminal lysine, and (iii) the mass was determined with an accuracy of at least 1 u, the only tryptic peptide from *E. coli* matching these criteria resides in the protein elongation factor EF-Tu.

Fig. 5B shows a doublet with the peaks varying by 4 u, thus it has a C-terminal arginine. Peak height ratio analysis suggests the parent protein was up-regulated only 1.4-fold. Assuming that (i) the peptide had one or more cysteine and histidine residues, (ii) a C-terminal arginine, and (iii) that the mass was determined with an accuracy of at least 1 u, four peptides were found matching these criteria. The four possible peptide candidates were LHVHDENNECGIGDVVEIR from 30 S ribosomal subunit protein S17, WQGQCSACHAWNTITEVR from ATP dependent protease, TLVFCVTNAHADMVVEELR from host restriction endonuclease R, and DQAMGFCFEAGADEDTHER from an activator for hydrogen peroxide-inducible genes. Differentiation between these four peptides requires more data, such as sequence data or chromatographic properties.

Without double selection, 45 possible peptide candidates matched the mass of the first peptide and 40 candidates matched the mass of the second. It is seen that in these experiments, double selection of peptides identified the peptide by mass alone in one

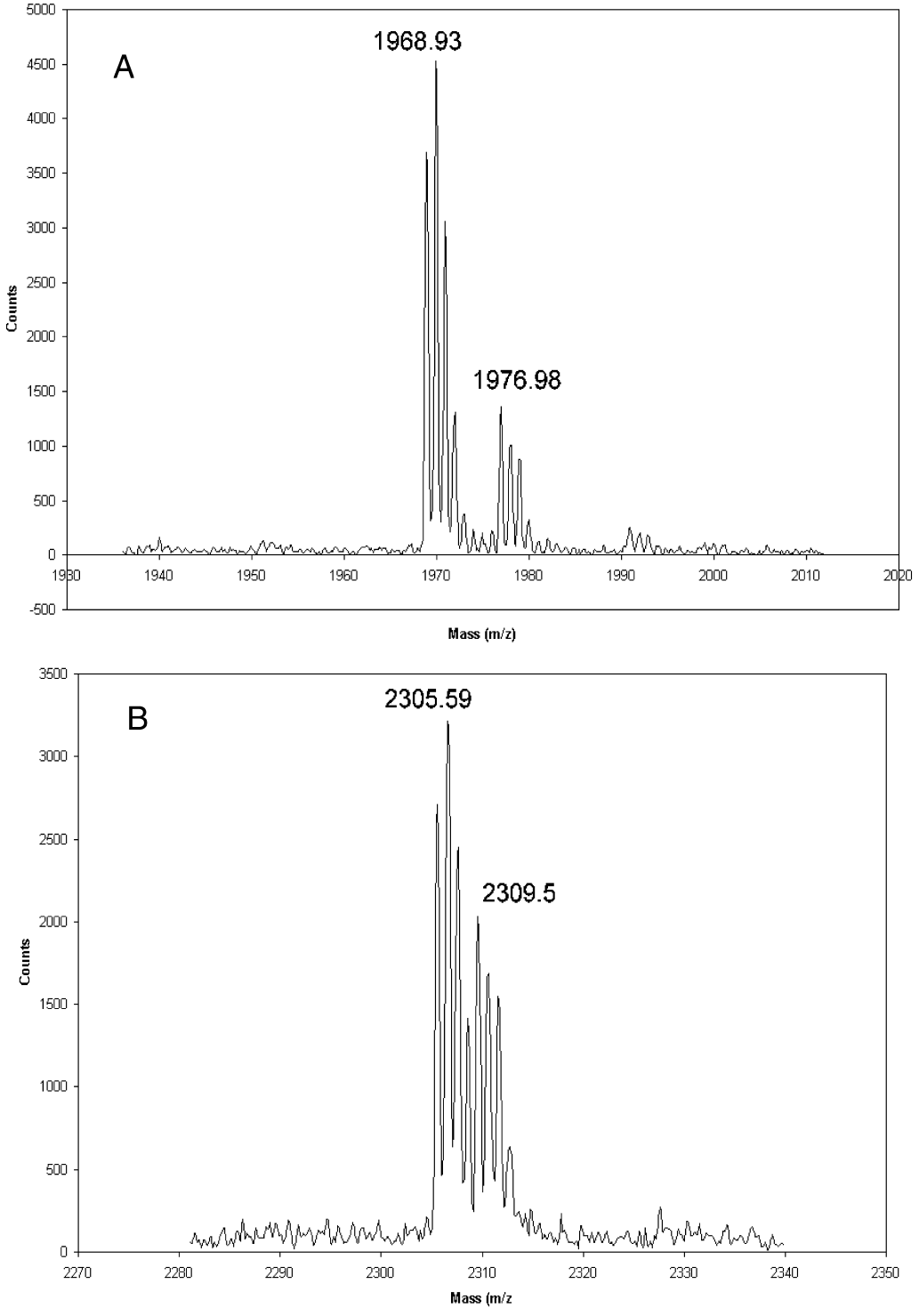

Fig. 5. (A,B) MALDI–TOF-MS spectra of two acylated cysteine- and histidine-containing peptides isolated from an *E. coli* tryptic digest using covalent chromatography and immobilized metal affinity chromatography loaded with copper. Note that the ratio of non-deuterated to deuterated peptide in (A) is 3.3, and that in (B) is 1.4. This indicates they were up-regulated in the experimental sample to a different extent.

case and reduced the complexity of the database 10-fold in another.

4. Conclusion

It is concluded that procedures for targeting surrogate peptides with two selectable features could efficiently simplify protein digests and during the course of doing so diminish the complexity of database searches and the need for sequence data. Coupled with a global internal standard labeling strategy, the double selection strategy provides an efficient method for recognizing and identifying proteins whose concentration is impacted by regulatory stimuli. These methods could be a valuable tool in the study of regulatory biology.

Acknowledgements

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

References

[1] P. O'Farrell, J. Biol. Chem. 250 (1975) 4007.
[2] P.Z. O'Farrell, H.M. Goodman, P. O'Farrell, Cell 12 (1977) 1133.
[3] J. Klose, Humangenetik 26 (1975) 231.
[4] A. Shevchenko, O.N. Jensen, A.V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, H. Boucherie, M. Mann, Proc. Natl. Acad. Sci. USA 93 (1996) 14440.
[5] S.D. Patterson, D. Thomas, R.A. Bradshaw, Electrophoresis 17 (1996) 877.
[6] M.P. Rout, J.D. Atchison, A. Suprapto, K. Hjertaas, Y. Zhao, B.T. Chait, J. Cell Biol. 148 (2000) 635.
[7] O.N. Jensen, T. Houthaeve, A. Shevchenko, S. Cudmore, T. Ashford, M. Mann, G. Griffiths, J. Krijnse Locker, J. Virol. 70 (1996) 7485.
[8] A. Shevchenko, M. Wilm, O. Vorm, O.N. Jensen, A.V. Podtelejnikov, G. Neubauer, P. Mortensen, M. Mann, Biochem. Soc. Trans. 24 (1996) 893.
[9] M. Fountoulakis, H. Lagen, S. Evers, C. Gray, B. Takacs, Electrophoresis 18 (1997) 1193.
[10] T. Nystrom, N.H. Albertson, K. Flardh, S. Kjellegerg, FEMS Microbiol. Ecol. 74 (1990) 129.
[11] F.C. Neidhardt, R.A. VanBogelen, Biochem. Biophys. Res. Commun. 100 (1981) 894.
[12] W. Weinmann, C.E. Parker, L.J. Deterding, D.I. Papac, J. Hoyes, M. Przybylski, K.B. Tomer, J. Chromatogr. A 680 (1994) 353.
[13] D. Figeys, A. Ducret, R. Aebersold, J. Chromatogr. A 763 (1997) 295.
[14] G.A. Valaskovic, N.L. Kelleher, F.W. McLafferty, Science 273 (1996) 199.
[15] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[16] M.P. Washburn, D. Wolters, J. Yates, Nat. Biotechnol. 19 (3) (2001) 242.
[17] A.J. Link, J. Eng, D.M. Schieltz, E. Carmack, G.J. Mize, D.R. Morris, B.M. Garvik, J. Yates, Nat. Biotechnol. 17 (1999) 676.
[18] K.K. Unger, K. Racaityte, K. Wagner, T. Miliotis, L.E. Edholm, R. Bischoff, G. Marko-Varga, J. High Resolut. Chromatogr. 23 (2000) 259.
[19] J. Ji, A. Chakraborty, M. Geng, X. Zhang, A. Amini, M. Bina, F. Regnier, J. Chromatogr. B 745 (2000) 197.
[20] Y. Nakagawa, T. Yip, M. Belew, J. Porath, Anal. Biochem. 75 (1981) 168.
[21] S.D. Patterson, C.S. Spahr, E. Daugas, S.A. Susin, T. Irinopoulou, C. Koehler, G. Kroemer, Cell Death Differ. 7 (2000) 137.
[22] C.S. Spahr, S.A. Susin, E.J. Bures, J.H. Robinson, M.T. Davis, M.D. McGinley, G. Kroemer, S.D. Patterson, Electrophoresis 21 (2000) 1635.
[23] M. Geng, J. Ji, F. Regnier, J. Chromatogr. A 870 (2000) 295.
[24] M. Geng, X. Zhang, M. Bina, F. Regnier, J. Chromatogr. B 752 (2001) 293.
[25] M.C. Posewitz, P. Tempst, Anal. Chem. 71 (1999) 2883.
[26] A.M. Crestfield, S. Moore, W.H. Stein, J. Biol. Chem. 238 (1963) 622.
[27] F. Regnier, A. Amini, A. Chakraborty, M. Geng, J. Ji, L. Riggs, C. Sioma, S. Wang, X. Zhang, LC–GC 19 (2001) 200.
[28] P. Hensen, G. Lindeberg, L. Anderson, J. Chromatogr. 627 (1992) 125.
[29] S. Wang, F. Regnier, J. Chromatogr. A 924 (2001) 345.
[30] R. Zhang, C. Sioma, S. Wang, F. Regnier, Anal. Chem. 73 (2001) 5142.
[31] N.F. Tabachnik, P. Blackburn, C.M. Peterson, A. Cerami, J. Pharmacol. Exp. Ther. 220 (1982) 243.
[32] K. Brocklehurst, J. Carlsson, M.P.J. Kierstan, E.M. Crook, Methods Enzymol. 34 (1974) 531.
[33] F. Fex, C.B. Laurell, E. Thulin, Eur. J. Biochem. 75 (1977) 181.
[34] B.S. Baines, K. Brocklehurst, P.R. Carey, M. Jarvis, E. Salih, A.C. Storer, Biochem. J. 233 (1986) 119.