*Anal. Chem.* **2002,** *74,* 4969−4979

# Acid-Labile Isotope-Coded Extractants: A Class of Reagents for Quantitative Mass Spectrometric Analysis of Complex Protein Mixtures

Yongchang Qiu,* Eric A. Sousa, Rodney M. Hewick, and Jack H. Wang

*Proteomics/Protein Chemistry Department, Wyeth, 87 CambridgePark Drive, Cambridge, Massachusetts 02140*

**Quantitative mass spectrometry using stable isotope-labeled tagging reagents such as isotope-coded affinity tags has emerged as a powerful tool for identification and relative quantitation of proteins in current proteomic studies. Here we describe an integrated approach using both automated two-dimensional liquid chromatography/mass spectrometry (2D-LC/MS) and a novel class of chemically modified resins, termed acid-labile isotope-coded extractants (ALICE), for quantitative mass spectrometric analysis of protein mixtures. ALICE contains a thiol-reactive group that is used to capture all cysteine (Cys)-containing peptides from peptide mixtures, an acid-labile linker, and a nonbiological polymer. The acid-labile linker is synthesized in both heavy and light isotope-coded forms and therefore enables the direct relative quantitation of peptides/proteins through mass spectrometric analysis. To test the ALICE method for quantitative protein analysis, two model protein mixtures were fully reduced, alkylated, and digested in solution separately and then Cys-containing peptides covalently captured by either light or heavy ALICE. The reacted light and heavy ALICE were mixed and washed extensively under rigorous conditions and the Cys-containing peptides retrieved by mild acid-catalyzed elution. Finally, the eluted peptides were directly subjected to automated 2D-LC/MS for protein identification and LC/MS for accurate relative quantitation. Our initial study showed that quantitation of protein mixtures using ALICE was accurate. In addition, isolation of Cys-containing peptides by the ALICE method was robust and specific and thus yielded very low background in mass spectrometric studies. Overall, the use of ALICE provides improved dynamic range and sensitivity for quantitative mass spectrometric analysis of peptide or protein mixtures.**

The sequencing of the human genome and the rapid accumulation of DNA sequences of other species have paved the way for large-scale comprehensive analyses of gene products (including both RNA and proteins). An important aspect in such analyses is to quantify the changes in abundance of those gene products in a cell or tissue under different conditions. Using techniques such as gene chips,[1] differential display PCR,[2] and serial analysis of gene expression,[3] global profiling at the RNA level has generated a vast amount of useful information in terms of understanding gene expression and its regulation. However, this information cannot be used to predict the expression levels and functional states of the corresponding proteins,[4−6] the actual biological effector molecules for the genes, since the downstream events, translation and posttranslational modifications, are beyond the scope of RNA profiling. Therefore, proteomics, the analysis of protein complements, has become essential to gain accurate and complete information of complex biological processes.

Global profiling of the entire proteome of cells or tissues in different states ideally requires a combination of techniques to achieve both confident identification and accurate quantitation of complex protein mixtures in a high-throughput manner. However, the realization of this goal has turned out to be a daunting task. This is mainly due to the extremely diverse physicochemical properties and wide dynamic range of proteins expressed in cells/tissues, with the additional complication of posttranslational modifications. Traditionally, protein mixtures have been separated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). After staining, the proteins are quantified by densitometric analysis of the individual protein spots. Then the gel spots of interest are excised and subjected to in-gel enzymatic digestions to produce peptides that are subsequently analyzed by mass spectrometry. These peptides, hence their corresponding proteins, are identified by database searching, which correlates the masses of peptides or their fragment ions with protein or nucleotide sequences in the databases. This approach has been playing an instrumental role in proteomic studies.[7] However, gel spot recognition and quantitation is time-consuming and not very accurate, and 2D-PAGE as a separation tool fails to analyze hydrophobic, very large or small, extremely acidic or basic proteins. Most importantly, without prior sample fractionation or

* Corresponding author. Telephone: 617-665-8747. Fax: 617-665-8878. E-mail: ycqiu@wyeth.com.

(1) Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science* **1995**, *270*, 467−70.
(2) Liang, P.; Pardee, A. B. *Science* **1992**, *257*, 967−71.
(3) Velculescu, V. E.; Zhang, L.; Vogelstein, B.; Kinzler, K. W. *Science* **1995**, *270*, 484−7.
(4) Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. *Mol. Cell. Biol.* **1999**, *19*, 1720−30.
(5) Anderson, L.; Seilhamer, J. *Electrophoresis* **1997**, *18*, 533−7.
(6) Urquhart, B. L.; Cordwell, S. J.; Humphery-Smith, I. *Biochem. Biophys. Res. Commun.* **1998**, *253*, 70−9.
(7) Anderson, N. G.; Matheson, A.; Anderson, N. L. *Proteomics* **2001**, *1*, 3−12.

10.1021/ac0256437 CCC: $22.00 
Published on Web 09/05/2002

enrichment, 2D-PAGE can only detect highly abundant proteins due to its limited loading capacity.[8] Severe deterioration in resolution because of streaking and electrophoretic distortion is commonly observed with high loads in an attempt to visualize less abundant proteins. Recently, Gygi et al.[9] showed that the combination of 2D-PAGE and MS failed to detect proteins from yeast genes with codon bias values of <0.1 (lower abundance proteins) despite a relatively high sample load (0.5 mg) and an extended electrophoretic separation. However, fully half of all yeast genes have codon bias values less than 0.1. Therefore, this 2D-PAGE/MS approach is not able to provide a complete picture of all the protein expression level changes.

In light of these limitations of 2D-PAGE as a separation tool, Link et al. proposed a "shotgun" strategy for global protein identifications.[10−12] In their method, protein mixtures are first enzymatically digested without prior separation at the protein level, and the resultant peptides are analyzed by automated data-dependent multidimensional LC/MS. Database searching based on raw fragment ion data then yields the identities of each peptide and thus its corresponding protein. This method certainly is less discriminative since peptides are much easier to separate than proteins and without the severe losses during the separation process. Multidimensional LC/MS is particularly useful when dealing with partially purified protein complexes.[13] However, more complex peptide mixtures, such as a digest of a total cell lysate, can still easily overwhelm the separation ability of two-dimensional liquid chromatography and thus exceed the current acquisition speed limit of data-dependent MS/MS analysis. Furthermore, no quantitative information of proteins can be obtained since the MS signals of peptides are extremely variable.[14]

Recently, several reports described the use of stable isotope-labeled modifying reagents for relative protein quantitation directly based on mass spectrometric data. Such methodologies may be divided into two categories based on their isotope labeling techniques. The first type uses in vivo metabolic labeling[15−17] in which proteins are labeled when they are synthesized. As a result, the ratio of the original amount of protein in two different cell states can be strictly maintained if metabolic scrambling of the label does not occur. However, in vivo metabolic labeling is often impossible, for example, in humans, and therefore is not generally applicable. The second type involves external stable isotope labeling of a particular amino acid residue or of the N- or C-terminus of proteins. In this case, it is essential for accurate quantitation that the labeling reaction is highly specific and stoichiometric. For this very reason, the cysteine residue becomes the natural target for the introduction of a label because the free thiol group is far more nucleophilic than any other natural amino acid side-chain group. Although there were reports of modifications of other amino residues and N- or C-terminus,[18−23] they generally had issues with the lack of labeling specificity and efficiency. Labeling on basic residues such as lysine can further complicate the MS data interpretation.[24] Alternatively, external labeling methods may not involve the use of stable isotopes. For example, Cagney and Emili[25] described a method called mass-coded abundance tagging (MCAT) in which lysine residues were converted to homoarginine for one of the two peptide mixtures that were to be compared. The relative quantity of a protein between the two samples was calculated using the ratio of mass peak intensities between the peptides containing lysine residues and similar peptides containing homoarginine residues.

A recent development in external isotope labeling methods was the invention of the isotope-coded affinity tags (ICAT).[26] In the ICAT method, the Cys-containing peptides from two different samples are not only differentially labeled with heavy and light versions of the ICAT but also selectively isolated from non-Cys-containing peptides. Such selective retrieval of Cys-containing peptides greatly reduces the complexity of the peptide mixture and thus increases the dynamic range of the mass spectrometric analysis.[21,27] Coupled with multidimensional chromatography, the ICAT strategy has been successfully applied to quantify differentiation-induced microsomal proteins.[28] However, the ICAT method also has some limitations that stem mainly from its reliance on biotin−avidin binding to isolate the Cys-containing peptides. For example, it is difficult to completely remove nonspecifically but tightly bound contaminants from the matrix during the washing steps, and the elution of Cys-containing peptides from the avidin column can be incomplete due to irreversible binding.[29] In addition, the capacity of the biotin−avidin binding is relatively low. Finally, the relatively large size of the ICAT label along with its susceptibility to fragment under normal CID conditions can complicate the MS/MS analysis of labeled peptides and subsequent database searching.[30] Recently, an


(8) Aebersold, R.; Rist, B.; Gygi, S. P. *Ann. N. Y. Acad. Sci.* **2000**, *919*, 33−47.
(9) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9390−5.
(10) Wolters, D. A.; Washburn, M. P.; Yates, J. R., 3rd. *Anal. Chem.* **2001**, *73*, 5683−90.
(11) Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd. *Nat. Biotechnol.* **2001**, *19*, 242−7.
(12) Yates, J. R., 3rd; McCormack, A. L.; Schieltz, D.; Carmack, E.; Link, A. *J. Protein Chem.* **1997**, *16*, 495−7.
(13) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., 3rd. *Nat. Biotechnol.* **1999**, *17*, 676−82.
(14) Mann, M. *Nat. Biotechnol.* **1999**, *17*, 954−5.
(15) Veenstra, T. D.; Martinovic, S.; Anderson, G. A.; Pasa-Tolic, L.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 78−82.
(16) Martinovic, S.; Veenstra, T. D.; Anderson, G. A.; Pasa-Tolic, L.; Smith, R. D. *J. Mass Spectrom.* **2002**, *37*, 99−107.
(17) Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6591−6.
(18) Munchbach, M.; Quadroni, M.; Miotto, G.; James, P. *Anal. Chem.* **2000**, *72*, 4047−57.
(19) Goodlett, D. R.; Keller, A.; Watts, J. D.; Newitt, R.; Yi, E. C.; Purvine, S.; Eng, J. K.; von Haller, P.; Aebersold, R.; Kolker, E. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1214−21.
(20) Regnier, F. E.; Riggs, L.; Zhang, R.; Xiong, L.; Liu, P.; Chakraborty, A.; Seeley, E.; Sioma, C.; Thompson, R. A. *J. Mass Spectrom.* **2002**, *37*, 133−45.
(21) Wang, S.; Regnier, F. E. *J. Chromatogr., A* **2001**, *924*, 345−57.
(22) Stewart, I. I.; Thomson, T.; Figeys, D. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2456−65.
(23) Yao, X.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 2836−42.
(24) Ji, J.; Chakraborty, A.; Geng, M.; Zhang, X.; Amini, A.; Bina, M.; Regnier, F. *J. Chromatogr., B* **2000**, *745*, 197−210.
(25) Cagney, G.; Emili, A. *Nat. Biotechnol.* **2002**, *20*, 163−70.
(26) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994−9.
(27) Spahr, C. S.; Susin, S. A.; Bures, E. J.; Robinson, J. H.; Davis, M. T.; McGinley, M. D.; Kroemer, G.; Patterson, S. D. *Electrophoresis* **2000**, *21*, 1635−50.
(28) Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 946−51.
(29) Moseley, M. A. *Trends Biotechnol.* **2001**, *19*, S10−6.
(30) Borisov, O. V.; Goshe, M. B.; Conrads, T. P.; Panisko, E. A.; Veenstra, T. D.; Smith, R. D. *49th ASMS Annual Conference on Mass Spectrometry and Allied Topics*, Chicago, IL, 2001.

improved version of ICAT using solid-phase isotope tagging has also been described.[31]

Here, we describe a new approach using both an automated two-dimensional nanoLC/MS system and a novel class of chemically modified resins, termed acid-labile isotope-coded extractants (ALICE), for quantitative mass spectrometric analysis of protein mixtures. Briefly, following denaturation, proteins are digested in solution and then the Cys-containing peptides covalently captured by ALICE. After extensive washing, these peptides are retrieved by mild acid-catalyzed elution and directly subjected to automated two-dimensional LC/MS for identification and LC/MS for accurate quantitation. Like ICAT, ALICE labels and extracts Cys-containing peptides from complex protein digests. Furthermore, Cys-containing peptides are covalently attached to the ALICE resin and thus the nonspecifically bound species can be removed completely. The peptides are also readily eluted from the resin with high yield under mild MS-compatible solvent conditions. Finally, the ALICE label is small in size and does not undergo multiple fragmentations under CID condition.

## EXPERIMENTAL SECTION

**Materials.** 6-Aminocaproic acid-$d_{10}$ was custom-made by Medical Isotopes (Pelham, NH). Sieber amide resin was purchased from Novabiochem (San Diego, CA). All other chemicals or solvents were purchased either from Aldrich (Milwaukee, WI) or Sigma (St. Louis, MO). All of the eight model proteins were purchased from Sigma. Sequencing grade modified trypsin and Lys-C protease were obtained from Roche Diagonostics Corp. (Indianapolis, IN).

**Synthesis of the ALICE.** A solution of maleic anhydride (0.98 g, 10.0 mmol in 15 mL of acetic acid) was added to a solution of 6-aminocaproic acid (1.31 g, 10 mmol in 5 mL of acetic acid). For the synthesis of heavy ALICE, 6-aminocaproic acid-$d_{10}$ was used instead. The resulting mixture was stirred at room temperature for 2 h and then heated to reflux (oil bath temperature at about 110−120 °C) for 4.5 h. The acetic acid was removed in a vacuum, and 3.3 g of a light yellow solid was obtained. This solid was chromatographed (20% ethyl acetate in hexane and then 50% ethyl acetate in hexane) and gave 0.92 g of pure target compound **1** [6-(2,5-dioxo-2, 5-dihydropyrrol-1-yl)hexanoic acid; 43% yield].

The protected polymer, Sieber amide resin (1 g, 0.15 mmol/g), was stirred in *N*,*N*-dimethylformamide (DMF) (8 mL) and then piperidine (2 mL) was added. The reaction mixture was stirred for 10 min; the solid was filtered and washed with methylene chloride and then dried under vacuum. This dry solid was then again stirred with piperidine (2 mL) in DMF (8 mL) for another 10 min. The thin-layer chromatography (TLC) was recorded and showed no trace of the fluorenylmethyoxycarbonyl (Fmoc). The solid was then filtered, washed with methylene chloride, and dried under low pressure to give ∼1 g of the free amine polymer **2**.

The deprotected polymer **2** (1 g, 0.15 mmol/g) was stirred in DMF (10 mL). To this mixture was added sequentially the target compound **1** (0.095 g, 0.45 mmol), 1-hydroxybenzotriazole (HOBT) (0.06 g, 0.45 mmol), and *N*,*N*-dicyclohexylcarbodiimide (DCC) (0.102 g, 0.5 mmol). The reaction mixture was stirred for 3 h and the solid filtered and washed successively with ethyl acetate, ether, and methylene chloride. The solid was dried in a vacuum to give ∼1 g of the final product (ALICE, Figure 1).

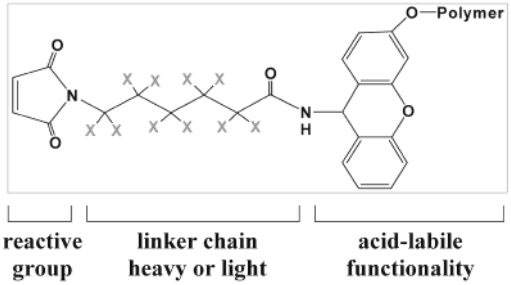


Figure 1. Chemical structure for the acid-labile isotope-coded extractants (ALICE): X = H for light ALICE (D0), mass addition 210.100 Da; X = D for heavy ALICE (D10), mass addition 220.163 Da.

**Isolation and MS Analysis of Cys-Containing Peptides from BSA Digests Using Light ALICE.** A 2-mg sample of bovine serum albumin (BSA) was solublized in 200 $\mu$L of 8 M urea, 200 mM ammonium bicarbonate, and 20 mM $CaCl_2$. A 5-$\mu$mol sample of tributylphosphine (TBP) predissolved in 20 $\mu$L of acetonitrile (ACN) was added into the solubilized protein mixture, and the resulting solution was incubated at 37 °C for 1 h. To the protein mixture was added 11 $\mu$mol of methylmethanethiosulfonate (MMTS), and the mixture was vortexed for 10 min. The protein solution was diluted 1:1 with 100 mM ammonium bicarbonate and then 40 $\mu$g of Lys-C (2% w/w) was added. This mixture was incubated at 37 °C for 5 h. The resulting solution was diluted 1:1 with water, and then proteins were further digested with trypsin (2% w/w) at 37 °C for 15 h. Half of the resulting peptide solution was saved for MS analysis, and the other half was dried and reconstituted with 50% ACN/200 mM sodium phosphate (pH 7.2). Disulfide bonds on the Cys-containing peptides were reduced with TBP (2.5 $\mu$mol) at 37 °C for 1 h. Then 25 mg of the ALICE resin (∼5.75 $\mu$mol of reactive sites) was added into the peptide solution and the solution vortexed for 1 h at room temperature. The reaction mixture was loaded onto a glass column with a Teflon stopcock and the resin washed sequentially with 5 mL each of the following solvents: 50% ACN; 30% ACN; 90% ACN; 100% ACN; 100% dichloromethane (DCM).

Cys-containing peptides were eluted from the resin with 5% TFA in dichloromethane slowly flowed through the resin over ∼3 h. The resulting peptide solution was then dried and reconstituted with 0.1 M acetic acid in water. Both the isolated Cys-containing peptides and the complete peptide mixture derived from the same amount of original protein (BSA) were subjected to automated on-line nanoLC/MS/MS analysis. The HPLC gradient was a ramp of 4−60% solvent B (solvent A, 0.1 M acetic acid/1% ACN; solvent B, 0.1 M acetic acid/90% ACN) over 45 min.

**Sample Preparation for Quantitative MS Analysis Using ALICE.** Two test protein mixtures were prepared with each containing the same eight proteins. The first protein mixture (A) contains lysozyme (10 nmol), α-lactalbumin (50 nmol), ovalbumin (25 nmol), catalase (50 nmol), $\beta$-lactoglobulin (38 nmol), bovine serum albumin (50 nmol), ribonuclease (50 nmol), and trypsinogen (50 nmol). The second protein mixture (B) contains lysozyme (50 nmol), α-lactalbumin (10 nmol), ovalbumin (50 nmol), catalase (25 nmol), $\beta$-lactoglobulin (50 nmol), bovine serum albumin (38 nmol), ribonuclease (50 nmol), and trypsinogen (50 nmol). Protein

(31) Zhou, H.; Ranish, J. A.; Watts, J. D.; Aebersold, R. *Nat. Biotechnol.* **2002**, *20*, 512−5.

mixtures A and B were each solubilized in 325 μL of a solution containing 6 M urea, 5% CHAPS, and 50 mM Tris-HCl at pH 8.0. Disulfide bonds of proteins were reduced by incubating with TBP (11.3 μmol in 6.3 μL of 2-propanol) at 37 °C for 1 h. The free thiols on cysteine residues were then methylthiolated by adding 34 μmol of MMTS predissolved in 3.5 μL of 2-propanol. The reaction was allowed to proceed for 30 min at room temperature, and the reaction mixtures were diluted four times with a solution of Tris-HCl (50 mM, pH 8.0). Proteins were then digested with trypsin (5% w/w) at 37 °C for 16 h. Forty-two percent of the total peptide mixtures was retained for future studies such as 2D-LC/MS/MS analysis, and the remaining 58% (187 nmol of total protein) was dried and then reconstituted with 1.5 mL of 60% ACN/40% 100 mM Tris-HCl (pH 7.0). Disulfide bonds on the Cys-containing peptides were reduced by TBP (18.7 μmol) at 37 °C for 1 h. Each solution was then vacuum concentrated for 10 min to remove excess TBP and ACN and then reconstituted to the previous volume using 100% ACN. To each solution was added 55 μmol of either light or heavy ALICE resins (3× TBP molar equivalent) and the solution was stirred for 1 h at room temperature. The reactions were quenched by the addition of β-mercaptoethanol (BME) to a final concentration of 1%.

The two protein mixtures were then combined and loaded onto a column (fritted glass type with Teflon stopcock), and the resin was washed with the following solvent in sequence: (i) 50 mL of a 50:50 ACN/water solution, (ii) 50 mL of pure ACN, (iii) 50 mL of a 50:50 ACN/DCM solution, and (iv) 50 mL of pure DCM. Finally, Cys-containing peptides were eluted using a solution of 5% TFA in DCM. The resulting peptide solution was dried and reconstituted with 2% ACN/0.1 M acetic acid/water. The reconstituted peptide solution was directly subjected to LC/MS and 2D-LC/MS/MS analyses.

**Automated Two-Dimensional LC/MS/MS Analysis.** Automated 2D-LC/MS/MS was accomplished using the system as shown in Figure 2 (see Scheme 1 in the figure). Specifically, an ion trap mass spectrometer (LCQDeca, Finnigan Corp., San Jose, CA) was fitted with an ABI 140C microgradient syringe pump system (Applied Biosystems, Framingham, MA), as the reversed-phase pump (RP), and an Agilent 1100 series binary pump, as the strong cation exchange (SCX) and desalting pump. The pumps were attached to a VICI 10-port microbore two-position valve with a microelectric actuator (Valco Instruments Co., Houston, TX). A 50 × 1 mm SCX column, polySULFOETHYL A (PolyLC, Columbia, MD), was attached to port 9. The two 75 μm × 10 cm IntegraFrit columns (New Objectives, Woburn, MA) packed with 10-μm YMC-Gel C18 beads (YMC, Wilmington, NC) were attached between ports 2 and 5 and ports 7 and 10, respectively. The nano-LC spraying tip is a 75-μm-i.d. PicoFrit column (New Objectives) that was packed to a length of 3 cm with C18 beads to provide optimal chromatographic separation.

Automation between the mass spectrometer, pumps, and valve was accomplished using contact closures. First, the sample was loaded onto the SCX column using a Rheodyne injection valve (Rheodyne, Rohnert Park, CA) with the port valve at position 10 as shown in Scheme 2 of Figure 2 so that any unbound peptides would bind to the RP-18 column and elute in fraction 0. With this dual C18 column design, while one C18 column (column A in Scheme 1) is on-line with the mass spectrometer for peptide separation, the second C18 column (column B in Scheme 1) is being regenerated, loaded with the peptide sample eluted from the SCX column, and desalted. After each HPLC gradient run is completed, the flow configurations of the two C18 columns were switched using the 2-position 10-port valve (Scheme 2 in Figure 2) so that the time delay for equilibrating, sample loading from SCX, and desalting was effectively eliminated. Briefly, one C18 column was equilibrated with 100% SCX-A for 20 min each at a flow rate of 1 μL/min and then peptide fractions were eluted from the SCX column onto the equilibrated C18 column using the following salt elution steps: (i) 5, (ii) 10, (iii) 15, (iv) 20, (v) 30, (vi) 40, (vii) 50, (viii) 65, (ix) 85, (x) 98, (xi) 98, (xii) 98, and (xiii) 98%, SCX-B/SCX-A, for 20 min at 1 μL/min. After each salt elution, the salt in the C18 column was rinsed with 100% SCX-A at 1 μL/min for 20 min and then 200 nL/min for the remainder of time. During the same time period, peptides were eluted from the other C18 column into the mass spectrometer using a linear gradient of 1−65% solvent B over 75 min and then 65−98% RP-B over the next 7 min at a flow rate of 400 nL/min. The column was then washed using 98% solvent B for 5 min and then 98−1% solvent B over the next 3 min. Mobile-phase buffers were as follows: solvent A, 0.1 M acetic acid/1% ACN; solvent B, 0.1 M acetic acid/90% ACN; SCX-A, 0.1 M acetic acid/1% ACN; SCX-B, 500 mM KCl. Automated data-dependent MS analysis was carried out using the dynamic exclusion feature built into the MS acquisition software. Each MS scan was followed by three MS/MS scans of the first three most intense peptide mass peaks to obtain as many CID spectra as possible.

**Quantitation of Proteins Using LC/ES/MS.** The eluted peptide mixture was mass measured using a Micromass Q-ToF$^2$ mass spectrometer (Micromass, Manchester, U.K.) equipped with an ABI 140 C microgradient syringe pump system (Applied Biosystems, Framingham, MA). Briefly, about one hundredth of the total eluted peptide mixture was injected onto a 75 μm × 10 cm PicoFrit column (New Objectives) packed with YMC−Gel 10 μM C18 beads (YMC). The PicoFrit column was directly connected to a Z-spray ion source with zero dead volume. The peptides were separated using a linear gradient of 4−65% solvent B over 75 min with a flow rate of 250 nL/min. Mobile-phase buffers were as follows: RP-A, 0.1 M acetic acid/1% ACN; RP-B, 0.1 M acetic acid/90% ACN. Only MS data were acquired throughout the chromatographic procedure to ensure the complete mass peak representation of all the ionized peptides at any given time point. The ratios of the heavy (D10) and light (H10) ALICE-labeled peptides were calculated from their relative peak heights. The relative quantitation of a protein from two different samples was determined by averaging the ratios of the heavy (D10) and light (H10) ALICE-labeled peptides that are unique to the same protein.

**Database Searching.** Peptide sequences were identified using the TurboSequest program[32] and the NCBI nonredundant protein database. Search parameters included a differential modification of +10 Da (mass difference between the heavy and light ALICE) and a static modification of +210 Da (mass addition due to modification by a light ALICE) on cysteine residues. The searches were done both with and without tryptic cleavage specificity to ensure maximum number of peptide identifications. All peptide

(32) Yates, J. R., 3rd; Eng, J. K.; McCormack, A. L. *Anal. Chem.* **1995**, *67*, 3202−10.

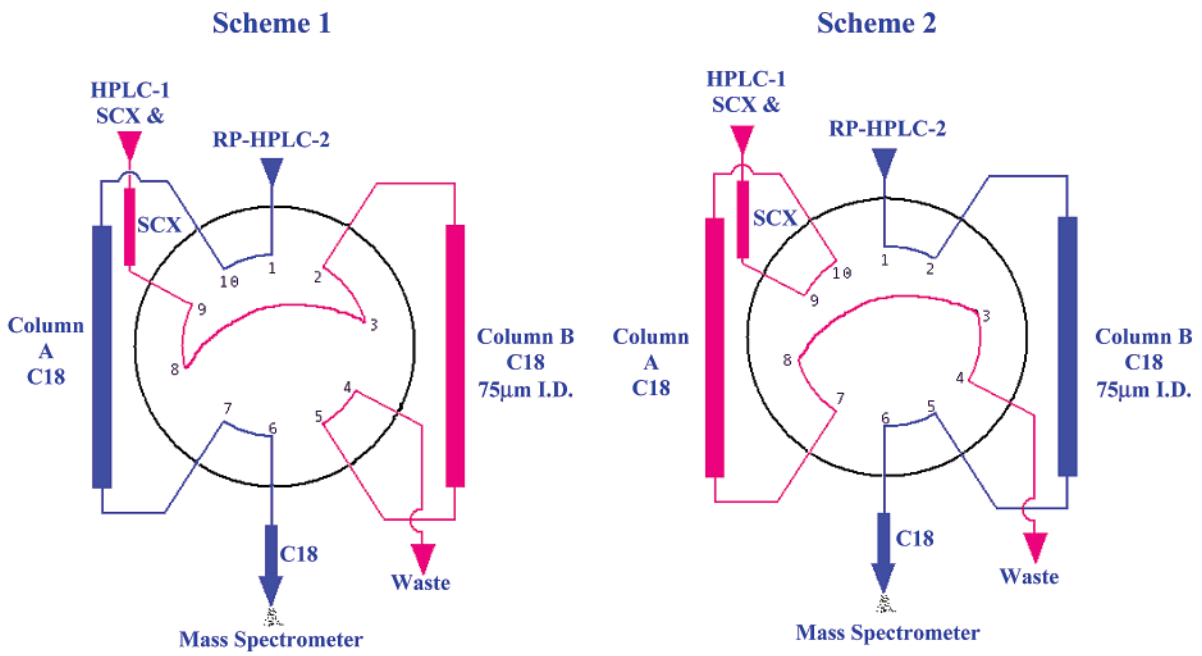


Figure 2. Schematic of the automated 2D-LC/MS system. The synchronization between a mass spectrometer, two HPLC systems, and a 10-port valve through contact closure enables parallel separation of peptides by both strong cation exchange and C18 and thus eliminates the time delay for sample loading, desalting, and column regeneration. The salt from SCX is directed to waste instead of the mass spectrometer and therefore prevents potential clogging at the spraying tip.

sequences listed in this article had Sequest scores Xcorr > 2.0. Peptides with an Xcorr between 2.0 and 3.0 were also confirmed by manual examination of their corresponding raw MS/MS spectra.

## RESULTS AND DISCUSSION

**Design and Synthesis of ALICE.** Chemically, ALICE is composed of three functional parts: a thiol-reactive group that is used to specifically capture cysteine (Cys)-containing peptides from a peptide mixture, a linker that can be isotopically labeled, and a nonbiological polymer with an acid-labile anchor group (Figure 1). A maleimido group was initially chosen as the thiol-reactive group rather than an iodoacetyl group because of concern with the stability of the latter. Furthermore, the use of a maleimido group may avoid the fragment ions that result from the potential cleavage of the amide bond since it would require two amide bond cleavages to release a fragment ion from the maleimido structure. The length of the linker was designed to allow 10 deuterium substitutions so that it is large enough to avoid the overlapping of mass peaks for the heavy and light ALICE-labeled peptides. In addition, the mass difference between heavy and light ALICE-labeled peptides cannot be mistaken for other common mass differences such as $m/z$ at 6 (3+ for water loss), 7 (2+ for methylene group), 8 (2+ for methionine oxidation), and 9 (2+ for water loss). Finally, the chemistry of the acid-labile anchor group (Sieber amide) has been well established for solid-phase peptide synthesis.[33] It was chosen because the products can be eluted from the resin with high yield using very mild MS-compatible solvents (1−5% TFA in DCM or ACN).

**Reactivity of ALICE with Cys-Containing Peptides.** To test the reactivity of ALICE, bovine serum albumin was reduced, and free cysteine residues were blocked by MMTS and then digested by trypsin to produce a mixture of peptides. Half of the digest was reacted with light ALICE after reduction with TBP. After extensive washing, the captured Cys-containing peptides were eluted using 5% TFA in ACN (Figure 3). The mole equivalent amounts of peptides both before and after the ALICE reaction were subjected to a 60-min nanoLC/MS/MS analysis during which the mass spectrometer was constantly alternating between one MS survey scan followed by three MS/MS scans. The peptides identified by subsequent database searching were summarized in Table 1. Only Cys-containing peptides were found in the sample eluted from the ALICE resin. This finding indicates that the nonspecifically bound peptides were completely removed from the resin using solutions with a high percentage of organic solvents. Peptides with more than one cysteine residue were all uniformly modified by light ALICE. As an example, a CID spectrum of a peptide containing three cysteine residues labeled by ALICE is shown in Figure 4 in which the first two modified cysteine residues are adjacent to each other. This indicates that there was little or no steric hindrance during the reaction between peptides and the reactive group attached to the ALICE resin. The total number of peptides (including nonspecific and incomplete tryptic cleavage products) identified from the crude digest was more than twice that from the ALICE sample. However, all 35 cysteine residues in BSA were identified from the sample eluted off the ALICE resin while two of the short Cys-containing peptides (C*ASIQK and LC*VLHEK) were not identified from the crude BSA digest (Table 1). The fact that some Cys-containing peptides were not detected in the crude digest emphasizes the importance of the use of ALICE in reducing sample complexity by enabling the selective recovery of Cys-containing peptides. Compared to the crude BSA digest that was also subjected to nanoLC/MS/MS analysis, the recovery

(33) Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 2107−10.

Table 1. Peptides Identified from Sample 1 (Total Digest) and Sample 2 (Peptides Eluted from ALICE Resin)

| peptides identified only from sample 1 | peptides identified from both samples 1 and 2 | peptides identified only from sample 2 |
|---|---|---|
| (F)LYEYSR | (F)GDELC*K | (R)C*ASIQK |
| (K)SLHTLF | (K)GAC*LLPK | (R)LC*VLHEK |
| (C*)TLPDTEK | (R)NEC*FLSHK | (F)SQYLQQC*PFDEHVK |
| (K)YLYEIAR | **(K)C*C*AADDK** | |
| (K)ATEEQLK | (K)SHC*IAEVEK | |
| (K)DLGEEHFK | **(R)C*C*TKPESER** | |
| (K)AEFVEVTK | **(K)C*C*TESLVNR** | |
| (K)QTALVELLK | **(K)EC*C*DKPLLEK** | |
| (K)LVNELTEFAK | (K)YIC*DNQDTISSK | |
| (R)FKDLGEEHFK | (K)SLHTLFGDELC*K | |
| (K)LVVSTQTALA | (K)QNC*DQFEK | |
| (K)KQTALVELLK | (K)QNC*DQFEKLGEY | |
| (K)HLVDEPQNLIK | (K)LKPDPNTLC*DEFK | |
| (Y)AVSVLLR | **(K)LKEC*C*DKPLLEK** | |
| (R)HPEYAVSVLLR | **(K)EYEATLEEC*C*AK** | |
| (R)RHPEYAVSVLLR | **(K)DDPHAC*YSTVFDK** | |
| (Y)FYAPELLYYANK | **(R)MPC*TEDYLSLILNR** | |
| (K)LGEYGFQNALIVR | **(R)M*PC*TEDYLSLILNR** | |
| (R)RHPYFYAPELLYY | **(R)ETYGDMADC*C*EK** | |
| (K)TVMENFVAFVDK | (K)LFTFHADIC* | |
| (K)TVMet-oxENFVAFVDK | (F)HADIC*TLPDTEK | |
| (Y)FYAPELL | (K)LFTFHADIC*TLPDTEK | |
| (Y)APELLYYANK | **(K)TC*VADESHAGC*EK** | |
| (Y)FYAPELLYYANK | (R)RPC*FSALTPDETYVPK | |
| (R)HPYFYAPELLYYANK | (R)NEC*FLSHKDDSPDLPK | |
| (F)LGSFLYEYSR | **(K)YNGVFQEC*C*QAEDK** | |
| (K)DAFLGSFLYEY | **(K)EC*C*HGDLLEC*ADDR** | |
| (K)DAFLGSFLYEYSR | (K)GLVLIAFSQYLQQC*PFDEHVK | |
| (K)NYQEAKDAFLGSF | (K)DAIPENLPPLTADFAEDKDVC*K | |
| (K)VPQVSTPTLVEVSR | | |
| (R)KVPQVSTPTLVEVSR | | |
| (Y)TRKVPQVSTPTLVEVSR | | |
| (K)DAIPENLPPLTADFAEDK | | |

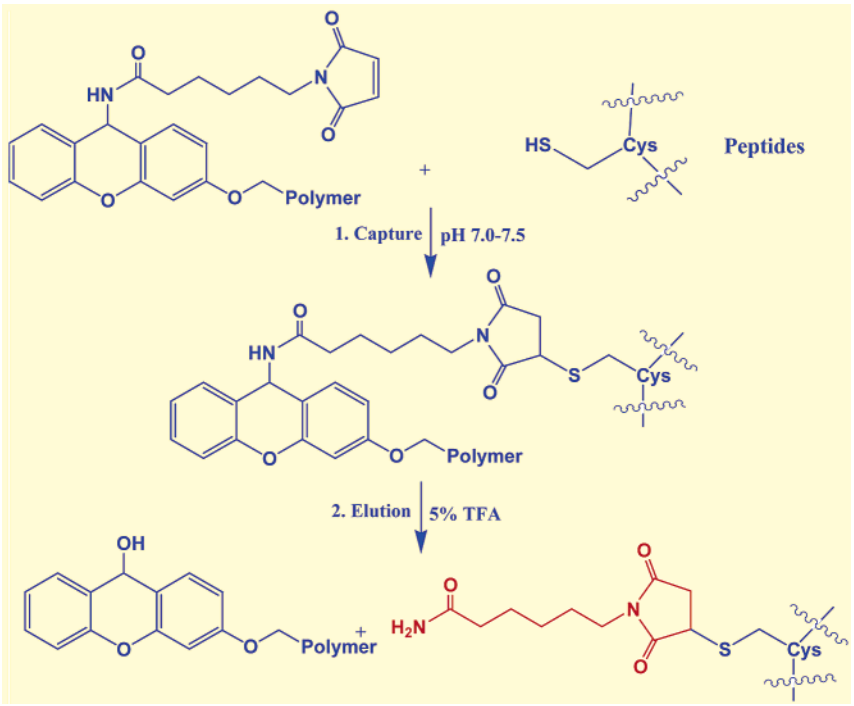


Figure 3. Chemical reactions involved in the specific isolation of cysteine-containing peptides using ALICE.

of the Cys-containing peptides using ALICE was estimated at ~90% based on the ratio of the TIC peak intensities between the same peptides before and after ALICE reaction. When the ALICE reaction was carried out at higher pH (pH >8.3), the side product derived from the opening of the maleimide ring became noticeable (data not shown).

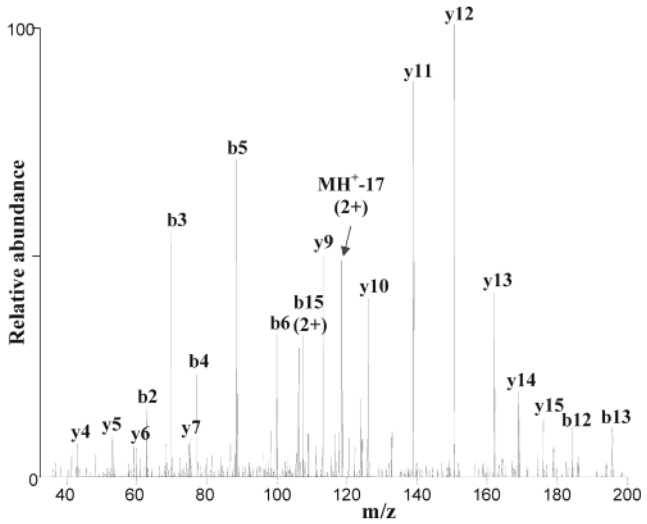


Figure 4. Fragment ion spectrum of a peptide from BSA with three cysteine residues that were uniformly labeled by light ALICE (peptide sequence: C*C*AADDKEAC*FAVEGPK; C*, ALICE-labeled cysteine residue).

**CID of the ALICE-Labeled Peptides.** One required property for an ideal isotopic labeling agent is that the labeled amino acid residue does not fragment more easily than an average peptide amide bond under normal CID conditions. Otherwise, the additional sequence ions resulting from the fragmentation of the label itself will complicate the subsequent data analyses such as database searching for protein identification.[30] To further explore the CID properties of the ALICE-labeled peptides, both the crude BSA digest and the isolated ALICE-labeled peptides were subjected to an automatic switching experiment on a Q-TOF mass spectrometer. During this experiment, the MS instrument was constantly alternating between one MS survey scan and two subsequent MS/MS scans if there were new doubly or triply charged ions above the preset threshold. The collision energy used was charge-dependent and optimized to produce abundant higher $m/z$ sequence ions. Under these conditions, no sequence ions with partial ALICE labels were found among all CID spectra acquired. As an example, in Figure 5, a complete y ion series (doubly and triply charged ions were also counted) was obtained for an ALICE-labeled peptide (GLVLIAFSQYLQQC*PFDEHVK; C*, ALICE-labeled cysteine residue). The ALICE modification was clearly indicated by an increase of 210.10 in $m/z$ value for the y ion series starting from y8 in comparison to the theoretical y ion series without any modifications. Compared to the CID spectrum of the corresponding MMTS-labeled peptide (spectrum not shown), the only noticeably different ion is the immonium ion of the light ALICE-labeled cysteine residue at $m/z$ 286.13 (marked as C* in Figure 5). All other fragment ions were identical without or with the corresponding mass addition (plus 210.10 Da for ALICE, plus 45.99 Da for MMTS).

**Quantitative Analysis of a Defined Protein Mixture Using ALICE.** To test the feasibility of the ALICE strategy for quantitation of protein mixtures, two protein mixtures that contained the same eight proteins but in different amounts were prepared (see Experimental Section). These two mixtures were compared and analyzed using the experimental protocol illustrated in Figure 6 . Briefly, the two protein mixtures were subjected to in-solution trypsin digestion after reduction of the disulfides and subsequent protection of the resultant free thiols. Then the peptides containing the protected cysteine residues were reduced again and reacted with either heavy or light ALICE. After the ALICE reaction, samples were combined and washed rigorously to remove all the noncovalently bound species using various organic solvents under neutral pH conditions. Finally, the captured Cys-containing peptides were eluted using mild solvents that are compatible with mass spectrometric analysis. The eluted samples were dried and then subjected to 2D-LC/MS/MS analysis and subsequent database searching for protein identification and LC/MS for relative quantitation. This analysis showed good correlation of the experimental data with the theoretical ratios of the same proteins in the two different mixtures (Table 2).

**Automated 2D-LC/MS/MS.** In light of the need for better spatial and temporal resolution of peptides, an automated two-dimensional LC/MS/MS system was constructed as illustrated in Figure 2. The first dimension was a strong cation exchange column that was separated from the C18 column in order to increase the initial sample loading capacity. More importantly, the dual C18 column configuration allowed simultaneous separation of peptides by both SCX and C18 columns and thus cut the overall analysis time in half compared to a single-column configuration.[34] Furthermore, the salt eluted from the SCX column was directed to waste instead of the mass spectrometer and therefore prevented the potential clogging at the spraying tip. One noticeable limitation of this new configuration was that the resolution of peptides using the C18 column was somewhat impaired by the dead volume in the switching valve (TIC peak broadening, data not shown). However, the use of a small C18 PicoFrit column as the spraying tip effectively eliminated this resolution loss.

Database searching with raw MS/MS data from the analysis of the eluted peptide mixture using SEQUEST identified numerous peptides. Again, only ALICE-modified Cys-containing peptides were found (data not shown). A total of 80 out of the 87 cysteine residues contained in the sequences of the eight proteins were covered in these identified peptide sequences. Many of these cysteine residues were seen multiple times since about half of the peptides identified were derived from nonspecific or incomplete tryptic cleavages. These nonspecific or incomplete enzymatic cleavages increased the overall complexity of the sample mixture. On the other hand, such cleavages can facilitate the identification of cysteine residues that are located in peptides with very low or high molecular weights that tend to be intractable to CID analysis. For example, a peptide resulting from the incomplete tryptic cleavage of trypsinogen with the sequence (K)C*LKAPILSDSSC*K-(S) was identified together with (K)APILSDSSC*K(S). As a result, the first cysteine residue of this longer peptide was identified whereas complete trypsin digestion would have produced a peptide, (K)C*LK(A), that would be too small to be detected by a mass spectrometer for a CID analysis. Another observation was that the same peptides were often found in nanoLC/MS runs that corresponded to adjacent SCX elution fractions and this is probably due to the relatively poor resolution of peptides by SCX column chromatography compared to that achieved on a C18 column. Since the overall number of different peptides that can be

(34) Yates, J. R., 3rd; Link, A. J.; Schieltz, D. *Methods Mol. Biol.* **2000**, *146*, 17−26.

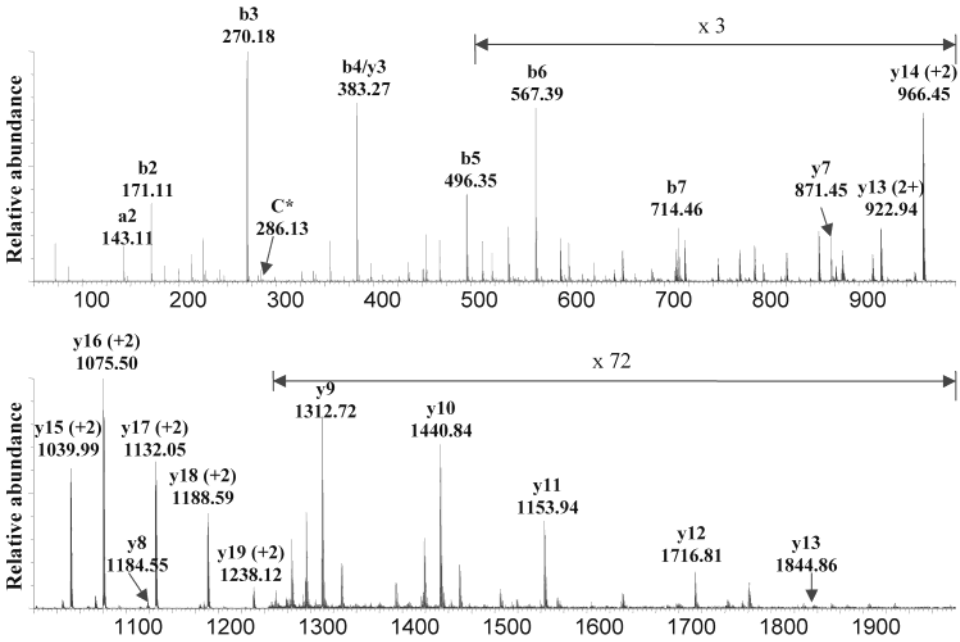


Figure 5. Low-energy CID spectrum of a light ALICE-labeled peptide generated from a Q-TOF mass spectrometer (peptide sequence: GLVLIAFSQYLQQC*PFDEHVK). The peak labeled as C* is the immonium ion of the light ALICE-labeled cysteine residue (*m*/*z* 286.13).

potentially sequenced in a given time period is reduced due to the redundancy in MS acquisition, this issue of resolution may become more evident when increasingly complex protein mixtures such as whole cell lysates are analyzed.

**Chromatographic Properties of Light and Heavy ALICE-Labeled Peptides.** Peptides having the same amino acid sequence but modified by either light or heavy ALICE were readily resolved by nanoLC (reversed-phase C18). The peptides modified by heavy ALICE eluted a few seconds earlier than the same peptides modified by light ALICE (Figure 7). The retention time differences between the two isotopic forms of the same peptide varied with the amino acid sequence of each individual peptide and generally became larger as the number of modified cysteine residues in their sequence increases. This is certainly consistent with several earlier reports.[35−37] Zhang et al.[35] also reported that resolution of isotopic forms of a peptide could cause substantial quantitation errors in quantitative analysis of proteins if not considered properly.

**Quantitation of Protein Mixtures Using LC/MS.** To obtain an accurate quantitation of the components of a protein mixture, the light and heavy ALICE-labeled peptide mixture was analyzed by LC/MS using the orthogonal quadrupole-TOF instrument that was programmed to acquire an MS spectrum constantly throughout the entire nanoLC gradient run. All the spectra containing both mass peaks of heavy and light ALICE-labeled peptides were combined to produce a composite MS spectrum. Typically 10−30 spectra were combined since an average peak duration for a peptide was about 20−60 s and each individual spectrum was acquired within 2 s with an interscan time of 0.1 s. The intensity ratio of the monoisotopic peaks of the heavy and light ALICE-labeled peptides in the composite MS spectra were calculated individually for each charge state (Figure 8). Peaks that overlapped with other mass peaks or had abnormal peak shapes due to poor data statistics (typically because of the weakness of a signal) were not included in the calculation (see Table 2). The overall quantitation of a protein from two different samples was determined by averaging all the peak intensity ratios of heavy and light peptide pairs that were derived from the same protein. To determine the identities of the ALICE-labeled peptide detected by orthogonal quadrupole-TOF, the MS spectra were recalibrated using internal standards (known peptide peaks from BSA). Given the earlier identification of peptides by 2D-LC/MS/MS analysis, the peptide sequences were easily assigned to each of the peak pairs based on their *m*/*z* value, charge state, presence of heavy and light pairs, and elution profile (percentage of solvent B when the peptide pairs were eluted off the C18 column). Since the average chromatographic peaks for each heavy and light peptide pair were about 30−60 s in width, the chance for either one of these heavy and light mass pairs to partially overlap with other mass peaks were still very high in this 90-min nanoLC run. As a result, many mass pairs were not useful for quantitation purposes and thus not used for calculation of the final average intensity ratios or simply not listed in Table 2. The final quantitation results were summarized in Table 2. The percentage errors for quantitation were generally less than 5% except for ovalbumin (16%). Since the intensity ratios for all individual peptide pairs from ovalbumin were consistently higher than the corresponding theoretical ratios and the overall standard deviation was relatively low (0.04), this larger deviation for ovalbumin was likely a result of a systematic error, such as a pipetting error.

(35) Zhang, R.; Sioma, C. S.; Wang, S.; Regnier, F. E. *Anal. Chem.* **2001**, *73*, 5142−9.
(36) Parker, K. C.; Griffin, T.; Gygi, S.; Aebersold, R. *48th ASMS Conference on Mass Spectrometry and Allied Topics*, Long Beach, CA, 2000.
(37) Filer, C. N. *J. Labeled Compd. Radiopharm.* **1999**, *42*, 169−97.

Table 2. Sequence Identification and Quantitation of the Components of a Protein Mixture Using ALICE

| protein name | peptide mass | charge state | peptide sequence identified[a] | observed ratio | expected ratio | mean ± SD | % error |
|---|---|---|---|---|---|---|---|
| α-lact-albumin | 432.20 | 2 | (K)C*EVFR(E) | 4.97 | 5 | 4.94 ± 0.20 | 1.2 |
| | 863.40 | 1 | (K)C*EVFR(E) | 5.20 | | | |
| | 432.71 | 2 | (K)ALC*SEK(L) | 4.83 | | | |
| | 860.42 | 1 | (K)ALC*SEK(L) | 4.75 | | | |
| | *622.80* | *2* | *(K)LDQWLC*EK(L)* | *weak* | | | |
| | *934.92* | *2* | *(K)FLDDDLTDDIM#C*VK(K)* | *overlap* | | | |
| | *666.28* | *2* | *(K)FLDDDLTDDIM#C*VKK(I)* | *weak* | | | |
| β-lacto-globulin | 1107.84 | 3 | (K)YLLFC*M#ENSAEPEQSLVC*QC*LVR(T) | 0.76 | 0.76 | 0.79 ± 0.04 | 3.5 |
| | 934.94 | 2 | (R)LSFNPTQLEEQC*HI(−) | 0.77 | | | |
| | 623.62 | 3 | (R)LSFNPTQLEEQC*HI(−) | 0.83 | | | |
| | *666.78* | *2* | *(K)WENDEC*AQK(K)* | *weak* | | | |
| catalase | 654.34 | 2 | (R)LC*ENIAGHLK(D) | 2.1 | 2 | 2.02 ± 0.09 | 1 |
| | 436.56 | 3 | (R)LC*ENIAGHLK(D) | 1.93 | | | |
| | 979.00 | 2 | (R)LGPNYLQIPVNC*PYR(A) | 2.01 | | | |
| | *746.36* | *2* | *(K)LVNADGEAVYC*K(L)* | *overlap* | | | |
| lysozyme | 1062.49 | 1 | (R)C*ELAAAM#K(R) | 0.2 | 0.2 | 0.21 ± 0.01 | 2.5 |
| | 989.81 | 3 | (R)NLC*NIPC*SALLSSDITASVNC*AK(K) | 0.21 | | | |
| | *739.86* | *2* | *(R)GYSLGNWVC*AAK(F)* | *overlap* | | | |
| | *531.75* | *2* | *(R)C*ELAAAM#K(R)* | *overlap* | | | |
| ovalbumin | 739.8 | 2 | (A)SM#EFC*FDVFK(E) | 0.61 | 0.5 | 0.58 ± 0.04 | 16 |
| | 700.85 | 2 | (R)ADHPFLFC*IK(H) | 0.6 | | | |
| | 467.57 | 3 | (R)ADHPFLFC*IK(H) | 0.52 | | | |
| | 838.44 | 2 | (R)YPILPEYLQC*VK(E) | 0.59 | | | |
| ribonuclease | 1189.08 | 2 | (K)HIIVAC*EGNPYVPVHFDASV(−) | 1.08 | 1 | 1.01 ± 0.11 | 1.0 |
| | 793.06 | 3 | (K)HIIVAC*EGNPYVPVHFDASV(−) | 1.16 | | | |
| | 595.04 | 4 | (K)HIIVAC*EGNPYVPVHFDASV(−) | 1.17 | | | |
| ribonuclease | 706.60 | 4 | (R)C*KPVNTFVHESLADVQAVC*SQK(N) | 0.89 | | | |
| | 922.40 | 2 | (A)C*EGNPYVPVHFDASV(−) | 1.03 | | | |
| | 608.63 | 3 | (F)VHESLADVQAVC*SQK(N) | 0.96 | | | |
| | 865.5 | 1 | (K)HIIVAC*(E) | 1.03 | | | |
| | 433.25 | 2 | (K)HIIVAC*(E) | 0.9 | | | |
| | 1239.5 | 1 | (Y)STM#SITDC*R(E) | 0.9 | | | |
| | 620.25 | 2 | (Y)STM#SITDC*R(E) | 0.84 | | | |
| | 661.33 | 2 | (K)YPNC*AYK(T) | 1.01 | | | |
| | 744.37 | 1 | (K)NVAC*K(N) | 1.15 | | | |
| trypsinogen | 580.3 | 2 | (A)PILSDSSC*K(S) | 0.87 | 1 | 1.04 ± 0.11 | 4 |
| | 1230.61 | 1 | (K)APILSDSSC*K(S) | 1.01 | | | |
| | 615.80 | 2 | (K)APILSDSSC*K(S) | 1.18 | | | |
| | 661.33 | 2 | (K)VC*NYVSWIK(Q) | 1.18 | | | |
| | 892.95 | 2 | (K)C*LKAPILSDSSC*K(S) | 1.02 | | | |
| | 595.63 | 3 | (K)C*LKAPILSDSSC*K(S) | 1.04 | | | |
| | 958.41 | 2 | (K)DSC*QGDSGGPVVC*SGK(L) | 0.98 | | | |
| BSA | 1141.6 | 1 | (C)C*TESLVNR(R) | 1.5 | 1.32 | 1.34 ± 0.10 | 2.5 |
| | 566.25 | 2 | (C)C*TESLVNR(R) | 1.28 | | | |
| | 623.35 | 2 | (H)TLFGDELC*K(V) | 1.21 | | | |
| | 1194.02 | 2 | (K)C*C*AADDKEAC*FAVEGPK(L) | 1.24 | | | |
| | 796.35 | 3 | (K)C*C*AADDKEAC*FAVEGPK(L) | 1.23 | | | |
| | 722.83 | 2 | (K)C*C*TESLVNR(R) | 1.34 | | | |
| | 650.30 | 3 | (K)DDPHAC*YSTVFDKLK(H) | 1.35 | | | |
| | 630.80 | 2 | (K)EAC*FAVEGPK(L) | 1.3 | | | |
| | 533.25 | 3 | (K)EC*C*DKPLLEK(S) | 1.41 | | | |
| BSA | 911.50 | 1 | (K)GAC*LLPK(I) | 1.48 | | | |
| | 638.80 | 2 | (K)LFTFHADIC*(T) | 1.35 | | | |
| | 638.80 | 2 | (K)LFTFHADIC*(T) | 1.51 | | | |
| | 613.65 | 3 | (K)LKEC*C*DKPLLEK(S) | 1.51 | | | |
| | 577.28 | 3 | (K)LKPDPNTLC*DEFK(A) | 1.21 | | | |
| | 786.89 | 2 | (K)SLHTLFGDELC*K(V) | 1.35 | | | |
| | 524.92 | 3 | (K)SLHTLFGDELC*K(V) | 1.35 | | | |
| | 885.37 | 2 | (K)TC*VADESHAGC*EK(S) | 1.52 | | | |
| | 590.58 | 3 | (K)TC*VADESHAGC*EK(S) | 1.52 | | | |
| | 591.62 | 3 | (K)VTKC*C*TESLVNR(R) | 1.19 | | | |
| | 798.86 | 2 | (K)YIC*DNQDTISSK(L) | 1.36 | | | |
| | 1027.43 | 2 | (K)YNGVFQEC*C*QAEDK(G) | 1.2 | | | |
| | 859.43 | 1 | (R)C*ASIQK(F) | 1.46 | | | |
| | 430.21 | 2 | (R)C*ASIQK(F) | 1.3 | | | |
| | 1051.56 | 1 | (R)LC*VLHEK(T) | 1.35 | | | |
| | 526.284 | 2 | (R)LC*VLHEK(T) | 1.27 | | | |

Table 2 (Continued)

| protein name | peptide mass | charge state | peptide sequence identified[a] | observed ratio | expected ratio | mean ± SD | % error |
|---|---|---|---|---|---|---|---|
| BSA | 947.45 | 2 | (R)M#PC*TEDYLSLILNR(L) | 1.36 | | | |
| | 631.97 | 3 | (R)M#PC*TEDYLSLILNR(L) | 1.25 | | | |
| | 1027.97 | 2 | (R)NEC*FLSHKDDSPDLPK(L) | 1.27 | | | |
| | 1017.50 | 2 | (R)RPC*FSALTPDETYVPK(A) | 1.41 | | | |
| | 678.672 | 3 | (R)RPC*FSALTPDETYVPK(A) | 1.39 | | | |
| | 882.45 | 3 | (K)GLVLIAFSQYLQQC*PFDEHVK(L) | 1.30 | | | |
| | 736.97 | 3 | (K)EC*C*HGDLLEC*ADDR(A) | 1.26 | | | |
| | *700.85* | *2* | *(F)GDELC*KVASLR(E)* | *overlap* | | | |
| | *687.67* | *3* | *(K)LFTFHADIC*TLPDTEK(Q)* | *overlap* | | | |
| | *700.37* | *3* | *(K)SLHTLFGDELC*KVASLR(E)* | *overlap* | | | |
| | *613.30* | *2* | *(K)SHC*IAEVEK(D)* | *weak* | | | |
| | *736.83* | *2* | *(R)C*C*TKPESER(M)* | *overlap* | | | |

[a] Abbreviations: M#, oxidized methionine residue; C*, heavy and light ALICE-labeled cysteine residue.

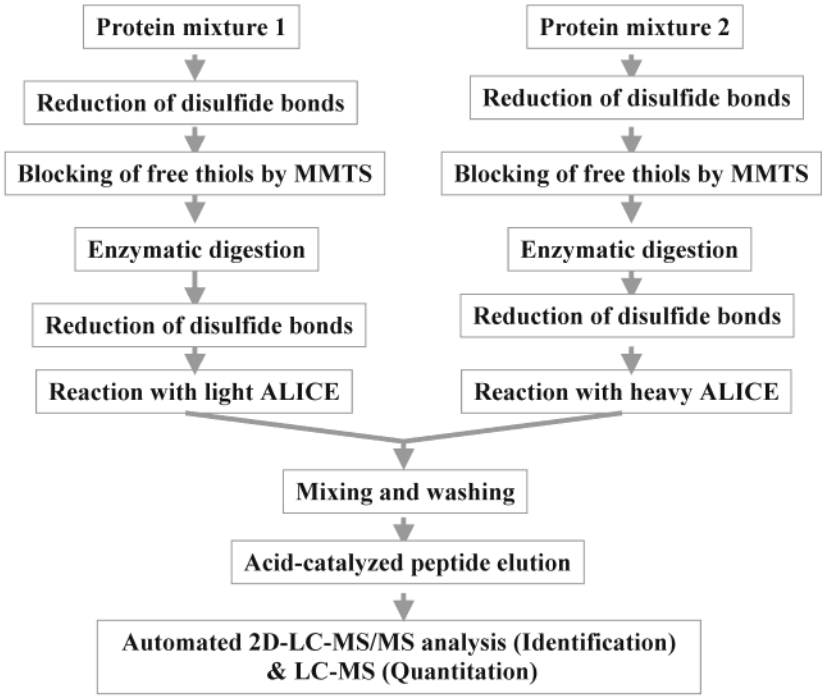


Figure 6. Typical procedure used for quantitative analysis of protein mixtures using ALICE.

A small number of the peptides identified in the previous 2D-LC/MS/MS analysis and database searching were not found in the LC/MS analysis (data not shown). Many of those peptides were partial or incomplete tryptic cleavage products with relatively weak mass intensity. This was probably due to the fact that the amount of sample loaded onto the nanoLC C18 column was less than one-tenth of that loaded onto the SCX column for 2D-LC/MS/MS analysis. This is because the loading capacity for a C18 column is much smaller than that for a SCX column of the same size.

Compared to the ICAT, one potential limitation of the ALICE strategy is that ALICE labels the peptides after rather than before enzymatic digestion since complete modification of all cysteine residues on proteins using ALICE is presumably very difficult to achieve due to the steric hindrance for the reaction between proteins and ALICE resin. However, quantitation accuracy was not affected in our study even for the quantitation results calculated from nonspecific or incomplete tryptic cleavage products. It may become an issue when the concentration of a protein in the first sample is much lower than that of the same protein in the second sample to be compared because of the reduced protease activities at low substrate concentrations in the first sample.[38] Further studies will be needed to determine the concentration limit below which the quantitation accuracy using the ALICE strategy will be affected.

## CONCLUSIONS

We have introduced a new strategy for quantitative analysis of complex protein mixtures based on a novel reagent termed acid-labile isotope-coded extractant. Like ICAT, ALICE enables accurate relative quantitation of proteins from two different samples. It is also capable of isolating Cys-containing peptides from complex peptide mixtures with high yield and thus reduces the sample complexity. However, ALICE relies on an acid-labile

(38) Quadroni, M.; James, P. *Electrophoresis* **1999**, *20*, 664−77.

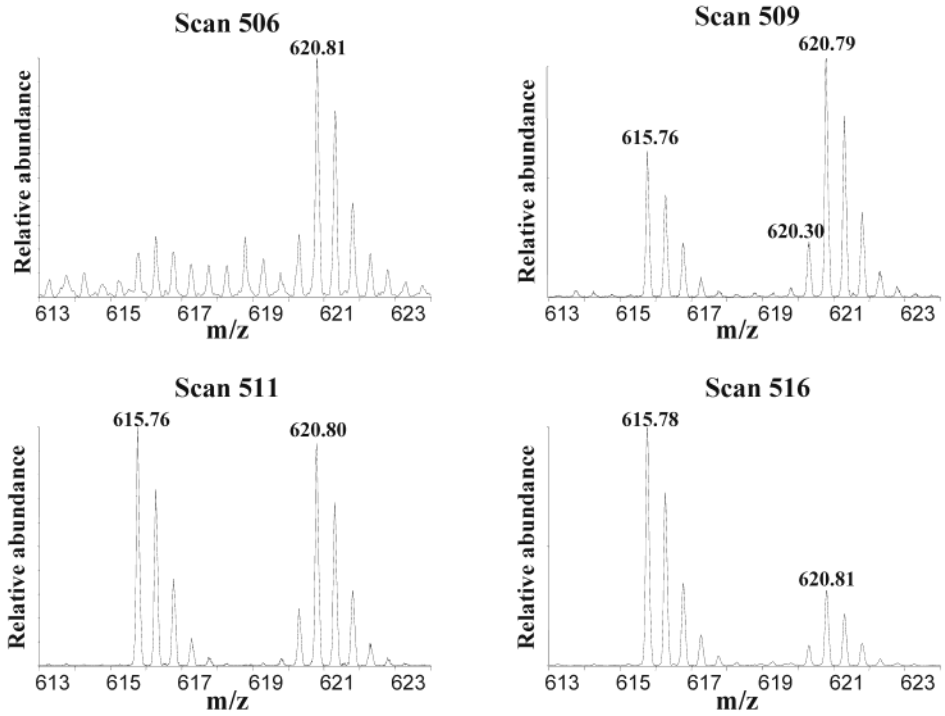


Figure 7. HPLC elution profile of the peptides labeled by heavy and light ALICE (peptide sequence from trypsinogen: APILSDSSC*K). Each MS scan consumed 2 s with an interscan time of 0.1 s. The peptide modified by heavy ALICE was eluted several seconds earlier than the same peptide modified by light ALICE during nanoLC.

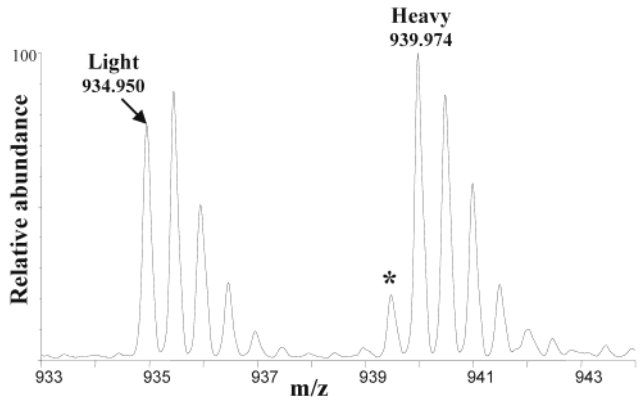


Figure 8. Composite MS spectrum (21 MS spectra combined) for the heavy and light ALICE-labeled C-terminal peptides from $\beta$-lactoglobulin: (R)LSFNPTQLEEQC*HI(−); charge state, +2. The intensity ratio of monoisotopic peak of the light ALICE-labeled peptide to that of the heavy one was calculated as 0.77 (see Table 1). *, a mass peak resulted from isotopic impurity.

covalent link to capture the Cys-containing peptides and thus allows rigorous washing to completely remove nonspecifically bound species without losing peptides of interest. This feature can potentially increase the sensitivity of mass spectrometric analysis since it minimizes the background chemical noise from upstream sample preparation. The robustness of the ALICE reaction permits the use of high percentage organic solvent throughout the procedure and thus minimizes adsorptive loss, particularly because this procedure avoids a final vacuum concentration of peptide solutions in the presence of water.[39,40] The chemical moiety attached to the Cys-containing peptides is small in size and does not fragment under normal low-energy CID conditions. Finally, ALICE has a very high capacity for capturing Cys-containing peptides and the captured peptides are easily eluted off the ALICE resin with high yield using MS-compatible acidic solvents. Combined with an automated two-dimensional nanoLC/MS system, the ALICE strategy should prove to be a general robust method for quantitative proteome analysis.

## ACKNOWLEDGMENT

We gratefully acknowledge Dr. Paul Blundell from Organix, Inc. for his help with the synthesis of heavy and light ALICE resins.

Received for review March 22, 2002. Revised manuscript received July 13, 2002. Accepted July 31, 2002.

AC0256437

(39) Speicher, K. D.; Kolbas, O.; Harper, S.; Speicher, D. W. *J. Biomol. Tech.* **2000**, *11*, 74−86.
(40) Staudenmann, W.; Hatt, P. D.; Hoving, S.; Lehmann, A.; Kertesz, M.; James, P. *Electrophoresis* **1998**, *19*, 901−8.