Published in final edited form as:

J Proteome Res. 2009 December ; 8(12): 5412–5422. doi:10.1021/pr900158n.

Quantitative Serum Proteomics Using Dual Stable Isotope Coding and *nano* LC-MS/MS

Hong Wang*, Chee-Hong Wong, Alice Chin, Jacob Kennedy, Qing Zhang, and Samir Hanash

Fred Hutchinson Cancer Research Center, Seattle, Washington 98109

Abstract

Stable isotope coding technique in combination with mass spectrometry has emerged as a powerful tool to accurately identify and differentially quantify proteins within complex protein mixtures. We present a novel methodology to increase the yield of quantified proteins while maintaining a high stable-isotopic labeling efficacy. With this approach, intact proteins in complex biological sample such as sera are labeled with the designated dual stable isotope coding (DSIC) systems. In brief, intact proteins are coded sequentially with acrylamide to label Cysteine residues (Cys) and with succinic anhydride to label Lysine residues (Lys). Protein samples coded with this dual stable isotope are subjected to an online 2D-HPLC fractionation. The resolved protein fractions are individually digested with trypsin and analyzed with *nano* LC-MS/MS. Our results show that the DSIC labeling efficiency is 100% for Cysteine (Cys) labeled with acrylamide and 98% for Lysine (Lys) labeled with succinic anhydride. A comparative analysis of DSIC labeling and single labeling of Cysteine residues was made. Analysis of an entire anion-exchange chromatography subfraction of sera yielded 165 identified proteins (criteria: error rate <5% and unique peptides ≥ 2), 104 of which were quantified using the single labeling method (i.e., Cysteine acrylamide labeling only). In contrast, using same criteria for identification, a total 185 proteins were identified and 174 proteins were quantified using the DSIC labeling technique.

Keywords

Dual Stable Isotope Coding (DSIC); Quantitative Serum Proteome; Biomarker

Introduction

Stable-isotope labeling method coupled with mass spectrometry (MS) analysis has emerged as a powerful technology for identification and concurrent quantification of proteins in complex biological mixtures such as serum. The basis for stable isotope coding strategy is when two different biological samples have pairs of chemically identical peptides with different stable-isotope composition (e.g., ^{13}C versus ^{12}C , ^{18}O versus ^{16}O or ^{15}N

* Correspondence should be addressed to: Hong Wang, Ph.D., Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., Seattle, WA 98109. Phone 206-667-5716. Fax 206-667-2537. hongwang@fhcrc.org..

Supporting Information Available: Information on Supplementary 1–4: proteins identified from one anion-exchange fraction of Cys-labeled immunodepleted serum samples, and of DSIC-labeled immunodepleted serum samples. This material is available free of charge via the Internet at <http://pubs.acs.org>.

versus ^{14}N), they can be differentially presented because of their mass difference in mass spectra. Thus, a computed ratio between signal intensities for such peptide pairs should be a direct and accurate measurement of the abundance ratio between the two peptides. There are currently three major applications in quantitative proteomics: (a) metabolic stable-isotope labeling,^{1,2} (b) isotope coding by chemical reaction,^{3–15} and (c) stable-isotope incorporation via enzyme reaction.^{16–19}

Metabolic stable-isotope labeling methods are applied to cells cultured in stable isotope-enriched media versus normal media to yield relative abundance of specific proteins. These particular methods are only applicable to cultured cells or small organisms and cannot be applied to other biological specimens such as cell lysates and biofluids (urine, serum or plasma). Stable isotopes can also be incorporated into peptides by proteases such as Trypsin, Lys-N or Glu-C.^{15–18} The digestion is performed in H_2^{18}O water and enzymatic oxygen exchange occurs at the carboxyl group of the generated peptides. A restriction of this approach is stable isotope incorporation accomplished at the peptide level, and one or two oxygen exchange leads to variability in peptide spacing and the mass offset of 2 Da which is not sufficient to optimally separate the isotopic envelopes for differential quantification. Given that the samples have to be processed separately until tryptic digestion, variability that may result from uneven protein loss during sample preparation may affect the derived data.

Isotope coding by chemical reaction to label specific amino acid residues is applicable to all protein samples. Several coding strategies have been developed to this effect such as the isotope-coded affinity tags (ICAT),^{3,13} a method by which proteins are modified with the Cysteine (Cys) reactive group (iodo group) and the biotin tag in the ICAT reagent allows the specific isolation of the modified Cys-containing peptides by avidin affinity chromatography. The isolated Cys-containing peptide mixtures are analyzed by LC-MS/MSMS. In addition to a Cys-directed stable isotope coding, several other coding strategies have been developed such as the isotope-coded protein labeling (ICPL)⁹ approach that targets all amino groups at the protein level using nicotinoyloxy succinimide (Nic-NHS), and an isobaric tag for relative and absolute quantification (iTRAQ)¹⁰ technique that relies on a tag delivering a specific reporter ion for quantification in MSMS spectra, but with an isobaric mass at the MS level. Disadvantages in some of those approaches for quantitative proteomics include the high cost of labeling reagents, restricted quantification to the single amino-acid coded peptides such as Cys-coded in ICAT or Lys-coded in ICPL and iTRAQ. In particular with iTRAQ, there is a reliance on a relatively large hydrophobic organic molecule, so that the stable isotope labeling reaction has to be in nonaqueous buffer system which usually causes protein precipitation and loss during the labeling process. Furthermore, large labeling organic molecules complicate the interpretation of MSMS sequence spectra due to fragmentation of the labeled tag as well during the MSMS scan. Although several label-free approaches have been reported,^{20–26} clear disadvantages of label-free methods include multiple opportunities for quantification error during parallel sample processing, thus, hindering the ability to create high accuracy and reproducibility during LC and MS runs.

In the current study, we present a novel methodology that is an alternative enhancement to the labeling approaches previously described for quantitative analysis of proteins with pairs of stable isotope-labeled and unlabeled chemical modifications to target two specific amino acid residues Cysteine (Cys) and Lysine (Lys) with respect to using smaller organic molecules acrylamide and succinic anhydride. We designate this labeling technique as the dual stable isotope coding (DSIC). We investigated the labeling efficiency of DSIC labeling with a standard protein mixture and assessed the influence of different enzymes used for protein digestion based on results of LC-MS/MS data analysis. We apply the DSIC labeling technique, together with intact-protein based online 2D-HPLC fractionation and nonsplit *nano* flow-rate dual-column parallel capillary HPLC system coupled with *nano* ESI LTQ-FT, for this quantitative serum proteomics study.

Materials and Methods

Materials

Glu-C, Trypsin and DL-dithiothreitol (DTT) were from Promega. $^{13}\text{C}_3$ -acrylamide (1,2,3- $^{13}\text{C}_3$, 99%) and $^{13}\text{C}_4$ -succinic anhydride (1,2,3,4- $^{13}\text{C}_4$, 99%) were from Cambridge Isotopic Laboratories, Inc. Sodium hydroxide (98%), guanidine hydrochloride (99%), $^{12}\text{C}_4$ -succinic anhydride (99%), transferrin (95%), α 1-Acid glycoprotein, and fetuin were from Sigma-Aldrich. $^{12}\text{C}_3$ -acrylamide (99.5%) was from Fluka. Bovine serum albumin (BSA), sodium chloride, ammonium bicarbonate, urea, acetonitrile, tris (hydroxymethyl) aminomethane, hydrochloric acid, 2-propanol and HPLC grade water were from Fisher Scientific. Formic acid (99%) was purchased from Pierce and trifluoroacetic acid (99%) was from Supelco.

Human Serum Sample Preparation

The Fred Hutchinson Cancer Research Center Review Board approved the sample procurement protocol. A serum sample from a healthy male subject was prepared and filtered through 0.22 μm cellulose acetate filters prior to immunodepletion. Top-six high abundance proteins (i.e., albumin, IgG, IgA, antitrypsin, haptoglobin and transferrin) were removed from 1 mL of serum with Hu-6 high capacity multi immuno-affinity removal column according to manufacturer's instructions from Agilent Technologies (10 mm i.d. $\times$ 100 mm Long, Wilmington, DE) using the HPLC system (Shimadzu Corporation, Columbia, MD) controlled with Workstation EZStart 7.4 SP1. The flow-through fraction containing the low-abundance proteins was then collected and concentrated to 500 μL with Amicon Ultra Centrifugal Filter Devices (3-kDa M_r cutoff, Millipore Corporation, Billerica, MA) by centrifugation at 3000 rpm for 1 h.

Dual Stable Isotope Coding (DSIC) of a Standard Protein Mixture

A total of 1.0 mg of a protein mixture, consisting of four standard proteins, transferrin, fetuin, BSA, and α 1-Acid glycoprotein, was prepared in DSIC labeling buffer (6.0 M Guanidine-HCl, pH 9.0). Proteins were reduced by adding 0.70 mg of DTT (100 mg/mL) to the protein mixture and the mixed protein solution was left for 30 min at 65 $^{\circ}\text{C}$ for sufficient reduction with DTT. After reduction, the protein solution was cooled to room temperature and subjected to Cys-labeling reaction by adding 7 mg of acrylamide to the protein solution.

The mixed protein solution was incubated in the dark for 1 h at room temperature. Subsequently, 10 mg of succinic anhydride was added to the Cys-labeled protein solution to start the Lys-labeling reaction. pH was maintained at 9.0 by adding 6 M NaOH as needed. The mixed protein solution was left in the dark for 2 h at room temperature to complete the Lys-labeling reaction. The dual labeled protein mixture was desalted by a reversed-phase column (Poros R2/10, Applied Biosystems, Foster City, CA) to remove the labeling buffer and surplus labeling reagents. The desalted sample was then further lyophilized to dry powder by freeze-drying (Labconco, Kansas City, MO).

Efficiency of Cysteine-Acrylamidation and Lysine-Succinylation

Four standard proteins, transferrin, fetuin, BSA, and α 1-Acid glycoprotein, were individually dissolved in labeling buffer (6.0 M Guanidine-HCl, pH 9.0) at a concentration of 10 mg/mL. A total of 1.0 mg of standard protein pool consisting of equal amount of four standard proteins (i.e., 250 μ g of each protein) was sequentially acrylamidated and succinylated following the dual labeling method described above. Five different amounts of acrylamide and succinic anhydride were used for this experiment. The dual stable isotope labeled protein mixture was digested with Trypsin and analyzed by *nano* flow-rate RPLC LTQ-FT. Cysteine and Lysine containing peptides identified from each of the four standard proteins were used for calculating the efficiency of acrylamidation and succinylation based on their peak area constructed with Qual Browser software (version 2.0.7, Thermo Fisher Scientific, Inc.).

Quantitative Analysis of DSIC

We applied the DSIC labeling method to the quantitative analysis of the standard protein mixture (transferrin, fetuin, BSA, and α 1-Acid glycoprotein). Each of four standard proteins was individually dissolved in labeling buffer (6.0 M Guanidine-HCl, pH 9.0) at a concentration of 10 mg/mL. Three pools of the four standard proteins were prepared with different molar ratios for each protein. A total of 1 mg of pooled protein mixtures was sequentially incubated with 7 mg of acrylamide (1,2,3- 12 C3, Light) or 7 mg of acrylamide (1,2,3- 13 C3, Heavy). Then, 10 mg of succinic anhydride (1,2,3,4- 12 C4, Light) was added to the sample that was labeled with acrylamide (1,2,3- 12 C3, Light) and 10 mg of succinic anhydride (1,2,3,4- 13 C4, Heavy) was added to the sample that was labeled with acrylamide (1,2,3- 13 C3, Heavy) by following the dual labeling method described above. Afterward, a combination of the light and heavy dual coded protein mixture was mixed into three specific molar ratio pairs (10:1, 1:1, and 1:10) and compared (12 C3-acrylamide: 13 C3-acrylamide and 12 C4-succinic anhydride: 13 C4-succinic anhydride).

Quantitative Human Serum Protein Analysis Using DSIC

A total of 10 mL of DSIC labeling buffer (6.0 M Guanidine-HCl, pH 9.0) was added to 500 μ L of concentrated immunodepleted serum sample (see Human Serum Sample Preparation) to replace the Immunodepletion Buffer A (Agilent Technologies, Wilmington, DE) with the DSIC labeling buffer. The sample solution was concentrated to 500 μ L with Amicon Ultra Centrifugal Filter Devices (3-kDa M_r cutoff, Millipore Corporation, Billerica, MA) by centrifugation at 3000 rpm for 1 h. Protein concentration in the concentrated samples was

measured by Bradford using Bio-Rad Protein Assay (Bio-Rad Laboratories, Hercules, CA). Samples of equivalent protein amount (4 mg of protein in each sample, two samples were marked as Sample #1 and Sample #2) were reduced by adding 2.8 mg of DTT (100 mg/mL) to each sample, respectively. The samples were kept for 30 min at 65 °C for sufficient reduction with DTT. After reduction, samples were cooled to room temperature and subjected to DSIC labeling reaction. In brief, 28 mg of $^{12}\text{C}3$ -acrylamide was added to the Sample #1 and 28 mg of $^{13}\text{C}3$ acrylamide was added to the Sample #2 to start the Cys labeling reaction. The samples were kept in the dark for 1 h at room temperature to complete the Cys acrylamidation. Subsequently, 40 mg of $^{12}\text{C}4$ -succinic anhydride was added to Sample #1 as well as 40 mg of $^{13}\text{C}4$ -succinic anhydride was added to Sample #2 to perform the Lys labeling reaction. pH was monitored and maintained at 9.0 by adding 6 M NaOH as needed. The two samples were kept in the dark for 2 h at room temperature to complete the Lys succinylation. Then, the two samples were combined and mixed. The sample mixture was then further desalted by reversed-phase column (Poros R2/10, Applied Biosystems, Foster City, CA) and lyophilized to dry powder by a freeze-drying system (Labconco, Kansas City, MO).

Online 2D-HPLC Intact-Protein Based Fractionation for DSIC Labeled Immunodepleted Human Serum

The DSIC labeled immunodepleted human serum sample was separated with an in-house built online 2D-HPLC system.²⁷ Briefly, the sample was separated by a first-dimensional anion-exchange column (WP column, 7.5 mm i.d. $\times$ 150 mm Long, Column Technology, Inc., Fremont, CA) using an eight-step salt elution. Protein fractions from each of the eight-step salt elution were automatically transferred onto the second dimension, a reversed-phase column (Poros R2/10, 4.6 mm i.d. $\times$ 100 mm Long, Applied Biosystems, Foster City, CA) in which proteins were further separated with a 30 min RPLC gradient-elution. The first-dimension anion-exchange chromatography mobile-phase **A** is 20 mM Tris, 3% 2-propanol, 4 M Urea, pH 8.5 and mobile-phase **B** is 20 mM Tris, 3% 2-propanol, 4 M Urea, 1 M NaCl, pH 8.5. The flow-rate is 1.0 mL/min. The eight-step salt elution includes 0% **B**, 5% **B**, 10% **B**, 15% **B**, 20% **B**, 30% **B**, 50% **B** and 100% **B**. The second-dimension reversed-phase chromatography mobile-phase **A** is 95% H_2O , 5% Acetonitrile, 0.1% TFA and mobile-phase **B** is 90% Acetonitrile, 10% H_2O , 0.1% TFA. The flow-rate is 2.4 mL/min. The gradient elution is 0–5 min, 5% **B**; 5–6 min, 20% **B**; 6–28 min, 65% **B**; 28–30 min, 80% **B**; 30–32 min, 80% **B**; 32–33 min, 90% **B**; 33–35 min, 90% **B**; 35–36 min., 5% **B**. The protein fractions from 2D-HPLC separation were lyophilized to dry powder and kept at -20°C until further application.

In-Solution Digestion of DSIC Labeled Protein Sample

Lyophilized protein samples from the standard protein mixture were dissolved in 50 mM NH_4HCO_3 (pH 8.5) followed by in-solution digestion with Trypsin or Glu-C (Promega, Madison, WI) overnight in an incubator set at 37°C . The digestion was quenched by adding 5 μL of 1.0% formic acid solution prior to *nano* flow-rate RPLC LTQ-FT analysis. For the 2D-HPLC fractionated human serum sample, we pooled the digested RPLC fractions from the second salt-elution of anion-exchange chromatography. After pooling, a total of 12 RPLC fractions were analyzed by *nano* flow-rate RPLC ESI LTQ-FT.

Preparation of Packed Capillary Columns

A packed capillary column (75 μm i.d. $\times$ 250 mm Long) was prepared with MagicC18 packing particles (5 μm , 100 Å pore size) using a high pressure packing device (~1500 psi). The packed column was washed with 95% of mobile-phase **B** (0.1% formic acid in acetonitrile) for 30 min at 600 nL/min, and washed with 95% of mobile-phase **A** (0.1% formic acid in water) for 30 min at 600 nL/min. The packed column was conditioned by loading 500 fmol of enolase digest and developing a 1-h gradient elution, followed by blank wash-run to regenerate and equilibrate the column prior to sample analysis by *nano* flow-rate RPLC LTQ-FT.

nano Flow-Rate RPLC LTQ-FT Analysis

Digested protein samples were analyzed using our in-house built nonsplit *nano* flow-rate dual-column parallel capillary HPLC system²⁷ coupled online with the *nano* ESI LTQ-FT (with 7-T) mass spectrometer (Thermo Fisher Scientific, Inc., Waltham, MA). Mass spectrometer parameters were 2.5 kV of spray voltage (capillary voltage), 200 °C of capillary temperature, 10^5 of FT resolution, 8×10^5 of FT target value, 10^4 of ion trap target value, 1 μs microscan for FT with 850 ms of injection time, and 1 microscan for ion-trap with 100 ms of injection time. Mass spectra were collected in a data-dependent mode with the m/z range of 400–2000. The full mass spectrum (MS scan) was acquired by FTICR and tandem mass spectrum (MSMS scan) was acquired by LTQ with a 35% NCE (normalized collision energy). Acquisition of each full mass spectrum was followed by the acquisition of MS/MS spectra for the five most intense +2 or +3 ions within a one second duty cycle. The minimum signal threshold (counts) for a precursor occurring during a MS scan was set at 500 for triggering a MSMS scan. Acquired data was automatically processed through the Computational Proteomics Analysis Systems (<https://proteomics.fhcrc.org/CPAS/project/home/begin.view>).²⁸ In brief, LC-MS/MSMS data were first converted to mzXML format using ReAdW software (version 1.2) to generate the peak list for the protein database search using the X!Tandem search engine (version 2007.07.01.2). Afterward, the protein database used in the study was the International Protein Index (IPI) Human protein knowledgebase (version 3.53) and the number of currently existing protein entries in the database was 73 748. A minimum criteria for peptide matching was set at a Peptide Prophet Score 0.2. Peptides that met the criteria were then grouped by protein sequences using the Protein Prophet algorithm at an error rate of 5% or less.²⁹ In addition to the mentioned database searching criteria, the following variable modifications were also considered: Cys labeled with $^{12}\text{C}_3$ -acrylamide (71.03657@C) and Lys labeled with $^{12}\text{C}_4$ -succinic anhydride (100.01604@K), Cys labeled with $^{13}\text{C}_3$ -acrylamide (74.04663@C) and Lys labeled with $^{13}\text{C}_4$ -succinic anhydride (104.02946@K), Methionine (Met) oxidation was also considered as variable modification (15.9949@M). Quantization algorithm (Q3) developed in-house was implemented to obtain the relative quantities for each pair of peptides identified by MSMS that contained either Cys or Lys, or both Cys and Lys residues. The Q3 algorithm reconstructs the ion mass chromatograms for both light and heavy forms for each identified peptide containing either Cys or Lys, or both Cys and Lys residues and computes the intensities for each form. According to the MSMS peptide identification list, peptides containing either Cys or Lys, or both Cys and Lys with Peptide Prophet score equal or

greater than 0.75 were selected. The monoisotope mass difference between light and heavy labeled peptides was 3.01006, 4.01342, and 7.02348 Da for Cys acrylamidation alone, for Lys succinylation alone, and for Cys acrylamidation plus Lys succinylation, respectively. For peptides containing only either one Cys or one Lysine, the analysis was restricted to either the monoisotopic or the first two- ^{13}C isotopic peaks. For peptides containing more than one Cys or Lys or both Cys and Lys, the analysis was restricted to either the monoisotopic or the first five- ^{13}C isotopic peaks. The peak intensity in each MS scan was computed by centering the MS scan to the location of its maximum, a location closest to the theoretical m/z for each peak within a 20 ppm window. Each light peak was paired with its corresponding heavy peak, and if both peaks were present, this pair of peaks was considered a match. For calculating the peak area, we used the primary scan (i.e., the one that immediately preceded the MSMS scan) as well as a certain number of MS scans immediately before and after the primary scan, up to a limit of 10 scans. In the case of low intensity peaks, if no more than one pair of ions was found in the primary scan, then all of the isotopic peaks found were used, whether matched or not (i.e., if the monoisotopic peak of both heavy and light form matched, but the heavy form also presented the first ^{13}C isotope, then we considered the sum of both peaks from the heavy and only the monoisotopic peak from the light form in the final quantization). The resulting ratio was adjusted slightly for all the peptides with mass greater than 1800 Da by subtracting the heavy form from the derived quantity, the fraction that can be expected from overlapping isotopes is the light form. Proteins containing more than one pair of peptides that yielded quantization had all the ratios averaged.

Results and Discussion

Intact Protein Based Dual Stable Isotope Labeling Strategy. The Reason for Developing Dual Isotope-Labeling Strategy

Acrylamide specifically reacts with sulfhydryl (-SH)-groups of Cysteine residues in a reduced protein sample. Chait and Sechi utilized unlabeled acrylamide and deuterium-labeled acrylamide as the coding reagents to identify Cysteine-containing peptides which they inferred from the isotopic distributions.³⁰ Their results showed that protein identification was improved and relative quantification could be obtained by MALDI-TOF MS. Our group has applied $^{12}\text{C}_3$ -acrylamide and $^{13}\text{C}_3$ -acrylamide to quantitative analysis of human serum using *nano* flow-rate RPLC ESI LTQ-FT.³¹ However, the mass offset of 3 Da makes it difficult to distinguish the isotopic envelopes for single Cysteine containing peptide mass spectra acquired by ESI based mass spectrometers, even if they are equipped with a high-resolution FT analyzer. Given the fact that peptides ionized by ESI source usually have dominating charge states of +2 and +3, another concern using the Cys-labeled approach is the lower percentage of quantified proteins in quantitative proteomics studies. We have observed an average of 35% quantified proteins among ~1500 proteins identified in a serum proteomics analysis. The reduced percentage of quantified proteins using Cys-acrylamide labeling is due to the low occurrence of Cysteine in proteins. Theoretically, there is 2.2% of Cysteine in IPI-Human protein knowledgebase (version 3.56) and 1.41% of Cysteine in UniProtKB/Swiss-Prot protein knowledgebase (version 56.6). The rationale for choosing other amino acids for isotopic labeling is to target amino acids that have higher chemical

reactivity on their side-chain residues and naturally higher occurrence in proteins. On the basis of the criteria mentioned, we chose Lysine as the labeling amino acid since Lysine has high chemical reactivity on its side chain and natural high occurrence in proteins. There is 5.86% of Lysine in IPI-Human protein knowledgebase (version 3.53) and 5.88% of Lysine in UniProtKB/Swiss-Prot protein knowledgebase (version 56.6). We have surveyed *in silico* the percentage of quantifiable peptides among three different labeling strategies (Table 1) from IPI-Human protein knowledgebase (version 3.53) using Trypsin as the enzyme: Cys-labeling, Lys-labeling, and Cys-labeling plus Lys-labeling–dual stable isotope coding (DSIC). A total of 28.1% of peptides are quantifiable using Cys-labeling method. However, the percentage of quantifiable peptides was increased to 57.8% with Lys-labeling approach. Impressively, the percentage of quantifiable peptides ascended to 70.1% with DSIC labeling strategy. Interestingly, we also found that the total number of peptides generated from the *in silico* Trypsin digestion was 1 474 170 with Cys-labeling method, and 1 002 278 with DSIC labeling approach. The complexity of tryptic peptide-pool generated from DSIC-labeled proteins was strikingly reduced since there were 471 892 less peptides generated from the DSIC labeling method compared to the Cys-labeling approach. However, the number of quantifiable peptides increased from 413 915 with Cys-labeling to 702 127 with DSIC labeling. Thus, given our experience with serum/plasma proteomics using Cys-labeling and *nano* RPLC LTQ-FT, we predicted that more proteins would be quantified with DSIC labeling strategy due to the reduced complexity of the peptide mixture (from 1 474 170 peptides to 1 002 278 peptides) and the increased percentage of quantifiable peptides (from 28.1% to 70.1%). From our limited knowledge, this is a novel concept which is presented for the first time to propose a dual stable isotope labeling strategy for a quantitative proteomic study. We chose labeling reagents based on the following criteria: (1) Only $^{12}\text{C}/^{13}\text{C}$ coding is considered because $^1\text{H}/^2\text{H}$ coding usually results in a primary isotope effect that causes differential elution between the ^1H - and ^2H -labeled peptides during reversed-phase liquid chromatography (RPLC) separation and compromises measurement accuracy during LC-MS analysis;^{33,34} (2) The labeling reagents should be small molecules with strong intermolecule chemical bonding to avoid fragmentation in collision-induced dissociation (CID) which would complicate the interpretation of MS/MS spectra; (3) The labeling reagents should have relative hydrophilic properties so that the labeling process can be performed in an aqueous buffer system such that proteins will maintain good solubility during the labeling reaction; (4) The isotopic labeling reagents should be affordable (Table 2).

Considering the criteria above and comparing with currently established intact-protein isotopic labeling systems—*isotope coded affinity tag* (ICAT) and *isotope-coded protein label* (ICPL) (Table 2), our design to use $^{12}\text{C}_3/^{13}\text{C}_3$ -acrylamide as the first coding system for Cys labeling and $^{12}\text{C}_4/^{13}\text{C}_4$ -succinic anhydride as the second coding system for Lys labeling offers these enhancements. The work-flow of this proposed dual stable isotope coding (DSIC) is illustrated in Figure 1.

Efficiency of Cysteine-Acrylamidation and Lysine-Succinylation

Protein acrylamidation results from specific acrylamide reaction with sulfhydryl groups of Cys residues in reduced proteins and protein succinylation results from the reaction of ϵ -

amino group of Lys residues with succinic anhydride. Hamada et al. examined the relationship between addition of varied amounts of succinic anhydride and the succinylation level of individual proteins.³⁵ Three standard proteins were used (casein, BSA and egg white albumin), and two fish proteins (myofibrillar protein and sarcoplasmic protein) extracted from rudder-fish *Girella punctata*. No linear relationship between succinic anhydride/protein (w/w) and the succinylation level was observed except for sarcoplasmic protein which was achieved at a ratio of 1:1 (w/w). For sarcoplasmic protein, only 10% succinylation level was obtained at ratio of 1:1 (w/w). However, when the ratio of succinic anhydride/protein was 10:1 (w/w), 95% succinylation level of sarcoplasmic protein was also achieved. Thus, according to Hamada's study and in order to determine the optimized amount of acrylamide and succinic anhydride to be added in the labeling reaction, we performed a similar study. In our study, we used a mixture of four standard proteins, transferrin (M_r 77 050 Da, not including 2 N-link glycans and 1 O-link glycan), fetuin (M_r 38 419 Da, not including 3 N-link glycans and 4 O-link glycans), BSA (69 323 Da), and α 1-Acid glycoprotein (M_r 23 512 Da, not including 5 N-link glycans) as the test sample. The reason we used a protein mixture rather than individual proteins (exemplified in Hamada's study) as the test sample is because the level of acrylamidation and succinylation is dependent on the spatial structure of the microenvironment around the sulfhydryl groups of Cys and the ϵ -amino group of Lys by different proteins. Five varied amounts of acrylamide and succinic anhydride were added to 1.0 mg of protein mixture for two sequential labeling reactions as described in an earlier section, Materials and Methods. The identified peptides containing either Cys or Lys, or both Cys and Lys from each of the four standard proteins were used for calculating the efficiency of acrylamidation and succinylation. We applied two methods for calculating the labeling efficiency. One was based on peak area of identified peptides and the other was based on the number of MSMS events. Table 3 shows that 100% acrylamidation at a ratio of 5:1 (acrylamide/protein, w/w) was achieved. Table 4 shows that 98% succinylation level with a ratio of 7:1 (succinic anhydride/protein, w/w) was obtained. The labeling efficiency calculated either by peak area or by MSMS scan event was very close. Taking into consideration the complexity of real biological samples such as serum or plasma, we used the weighted ratio of 7:1 (w/w) for acrylamide/protein, and 10:1 (w/w) for succinic anhydride/protein in this study.

Enzyme Selection

Since the ϵ -amino group of Lys is blocked after succinylation, Trypsin will be restricted to only cleave peptide chains at the carboxyl side of the Arginine (Arg). Thus, it was necessary to confirm the influence different enzymes had on the extent with identifications of Lys-labeled protein samples. We chose Trypsin and Glu-C as the enzymes for in-solution digestion of the DSIC labeled protein mixture (Table 5). We also compared the results between Trypsin digestion and Glu-C digestion for the nonfractionated DSIC labeled immunodepleted serum sample (Table 6). Results from the standard protein mixture digestion indicated that Glu-C digestion does not generate more sequenced peptides or quantifiable peptides compared to results from Trypsin digestion. Although Glu-C cleaves peptide bonds C-terminally at Glutamic acid (Glu), there is a much lower cleavage rate at Aspartic acid (Asp). Our results showed that Glu-C only cleaves at the C-terminus of Glutamic acid (Glu) even if one or more Asparartic acid (Asp) coexisted in cleaved peptides

(Table 5). However, for Trypsin digestion with respect to transferrin, there were 16 sequenced peptides assigned. Among those 16 peptides, 15 peptides were used for quantification based on Cys-coding or Lys-coding or by DSIC. In contrast, only nine were sequenced and quantified peptides which were assigned to transferrin with Glu-C digestion.

We also used Trypsin and Glu-C to digest the nonfractionated DSIC labeled immunodepleted serum sample (Table 6). A total of 104 proteins were identified and 72 proteins were quantified from Trypsin digestion. In contrast, 70 proteins were identified and 45 proteins were quantified from Glu-C digestion. Interestingly, the overlap of identified proteins between Trypsin digestion and Glu-C digestion was 84% and the overlap of quantified proteins Trypsin digestion and Glu-C digestion was 73%. Thus, we chose Trypsin as the preferred enzyme for the serum proteome study using the DSIC labeling technique.

Quantitative Analysis DSIC Labeled Proteins

The standard protein mixture (transferrin and BSA) was labeled using DSIC labeling with three different molar ratios as described in Materials and Methods. The digested protein mixture was analyzed by *nano* flow-rate RPLC ESI LTQ-FT. Figure 2 shows the plots of all quantification data points obtained from BSA (left panel) and transferrin (right panel) with significant correlation coefficients ($R^2 = 0.91$). The coefficient of variation (CV) calculated for all sequenced Cys and/or Lys containing peptides with three different molar ratios (Light/Heavy, $^{12}\text{C}:^{13}\text{C}$) 1:10, 1:1, and 10:1 were 23.0%, 10.4%, and 18% for BSA, and 22.0%, 11.3%, and 15.0% for transferrin. The quantitative results with DSIC labeling method were more accurate than what was achieved by Cys-acrylamide labeling alone (i.e., single labeling method). In our previous study,³¹ the variation coefficient of BSA quantification was 13.2% for the molar ratio 1:1 and 35–50% for molar ratio 1:10 and 10:1. The reduced accuracy of quantification with Cys-acrylamide labeling was interference of the isotopic peaks representing 3 or 4 ^{13}C . For a peptide containing a single Cys, the mass difference between its light and heavy acrylamide-labeled forms was only 3 Da, giving rise to the overlapping of their isotope envelopes because the mass-to-charge values (m/z) were 1.5 and 1.0 for the 2+ and 3+ ion species of the differentially acrylamide-labeled peptides, respectively. However, with DSIC labeling, sequenced peptides tend to have more than one quantifiable amino acid. This is due to the higher percentage of Lys compared to Cys residues in proteins (Table 1), and a common co-occurrence of Cys and Lys in proteins given that Lys is blocked with DSIC labeling, so trypsin will only cleave the C-terminus of Arginine (Arg). Figure 3 illustrates the feature of DSIC labeled peptides identified from transferrin. Figure 3A,B shows the overlap between ^{12}C -light labeled peak and ^{13}C -heavy labeled peak in the case of doubly charged peptides containing only one isotope-coded amino acid, either Cys or Lys. Even with high resolution FTICR MS, there is ~15% overlap between the Cys- $^{12}\text{C}_3$ -acrylamide-labeled peptide and Cys- $^{13}\text{C}_3$ -acrylamide-labeled peptide for the Cys containing peptide (Figure 3A) due to the overlapping of their isotope envelopes. In contrast, there is about 2% overlap between the Lys- $^{12}\text{C}_4$ -succinic anhydride-labeled peptide and Lys- $^{13}\text{C}_4$ -succinic anhydride-labeled peptide because the difference of their mass-to-charge values (m/z) is increased to 2.0 for doubly charged Lys containing peptide (Figure 3B). Furthermore, Figure 3C,D illustrates the advantage of DSIC labeling peptides containing one of each amino acid, Cys and Lys. In ESI MS, it is common that the Trypsin-

digested peptides have different charge status such as +2 and +3 in the same spectrum. Figure 3C,D demonstrates a lack of isotope-envelop overlap given that the difference of mass-to-charge values (m/z) of the paired monoisotopic peaks is 3.5 for the doubly charged peptide (Figure 3C) and the difference of mass-to-charge values (m/z) is 2.33 for the triply charged peptides (Figure 3D). Because there is much less overlap between the paired monoisotopic peaks using DSIC labeling method, compared to Cys-acrylamide labeling alone, the quantization accuracy was significantly improved with the DSIC labeling approach as predicted. This is of substantial importance for the analysis of complex biological samples such as serum which usually yields extremely high numbers of peptides after Trypsin digestion.

Quantitative Human Serum Proteomics Using DSIC Technique

For comparison between the DSIC-labeling and single-labeling, the same amount of serum sample was also labeled by single-labeling method (Cys-acrylamidation) and processed similarly as the DSIC labeling. Results of protein identification and quantification using the two different labeling methods are summarized in Table 7. The same criteria were applied to process and analyze the LC-MS/MS data through the CPAS pipeline (<https://proteomics.fhcrc.org/CPAS/project/home/begin.view>).²⁸ With the criteria of error rate <5% and number of unique peptide ≥ 2 , a total 165 proteins were identified with the Cys-acrylamide labeling method (Supplementary Table 1), and 185 proteins were identified with DSIC (Supplementary Table 2). With the criteria of error rate <5%, unique peptide ≥ 2 and peptide prophet ≥ 0.75 , 104 proteins were quantified with the Cys-acrylamide labeling method (Supplementary Table 3) and 174 proteins were quantified with the DSIC labeling method (Supplementary Table 4). The percentage of quantified protein was 94% with DSIC and 63% with single labeling. Although an additional 3654 MSMS events were obtained with the single labeling (Cys-coding) method compared to the DSIC labeling method (Table 7), more proteins were identified with DSIC labeling. From Figures 4 and 5 and Table 8, we observed that more MSMS spectra were assigned to a few high-abundance proteins (0.1–1.0 mg/mL) in the single labeling compared to the DSIC labeling. In the single labeling, 71.41% of MSMS events were used for sequencing six high-abundance proteins including apolipoprotein A-I, apolipoprotein A-II, isoform 1 of fibrinogen alpha chain, retinol-binding protein 4, transthyretin and beta-2-glycoprotein 1 (Supplementary Table 1), and 64.55% of MSMS events were used for quantifying four high-abundance proteins including apolipoprotein A-II, isoform 1 of fibrinogen alpha chain, retinol-binding protein 4 and beta-2-glycoprotein 1 (Supplementary Table 3). However, in the DSIC labeling, 32.67% of MSMS events were assigned for sequencing four abundance proteins including complement C3 (fragment), isoform 1 of fibrinogen alpha chain, fibrinogen beta chain and hemopexin (Supplementary Table 2), and 32.89% of MSMS events were used for quantifying those four high-abundance proteins (Supplementary Table 4). Furthermore, 1.29% of MSMS events were directed to the low-abundance protein identification (~ 500 pg/mL, 2–5 MSMS events) in the DSIC labeling (Table 8 and Supplementary Table 2) and 0.78% of MSMS events were directed to the low-abundance protein identification (~ 500 pg/mL, 2–5 MSMS events) in the single labeling (Table 8 and Supplementary Table 1). In the case of protein quantification, 1.20% of MSMS events directed to the low-abundance proteins (~ 500 pg/mL, 2–5 MSMS events) in the DSIC labeling (Table 8 and Supplementary Table 4) and

0.72% of MSMS events directed to the low-abundance proteins (~500 pg/mL, 2–5 MSMS events) in the single labeling (Table 8 and Supplementary Table 3). Since the total MSMS events used in protein identification and quantification are 22 216 and 12 453 in single labeling and 18 562 and 18 440 in DSIC labeling, the difference of MSMS events assigned to the low-abundance proteins between the single labeling and DSIC labeling was quite striking. Our results indicated that the single-labeling (Cys-coding) method generated more mass spectrometer detectable peptides than the DSIC-labeling (Cys- and Lys-coding) method. Since Trypsin only cleaves at the C-terminus of Arginine when Lysine is blocked after DSIC labeling, DSIC approach may greatly simplify the peptide population of the digested protein mixtures and increase the efficiency of MS detectable peptides compared to non-Lys labeling, for example, the Cysteine-acrylamide labeling (i.e., the single labeling approach). Given the limitation of RPLC column capacity and separation resolution, it is beneficial to simplify sample complexity for LC-MS/MSMS analysis taking into account the duty-cycle for data acquisition in data-dependent acquisition (DDA) mode. The identified proteins from the single-labeling (Supplementary Table 1) and DSIC-labeling (Supplementary Table 2) were correlated with the serum protein concentration reported by Rules-Based Medicine (http://www.rulesbasedmedicine.com/case-studies/figures/serum_or_plasma-table1.html). From the reported serum protein concentration, we chose the following proteins with varying concentration as reference markers. Complement 3 (1.51 mg/mL), Fibrinogen (0.01 mg/mL) and Apolipoprotein A1 (0.38 mg/mL) were used as the milligram per milliliter (mg/mL) reference. β -2 Microglobulin (2.15 μ g/mL) and Serum Amyloid P (34.15 μ g/mL) were used as the microgram per milliliter (μ g/mL) reference. CEA (1.08 ng/mL) and Factor VII (456.1 ng/mL) were used as the nanogram per milliliter (ng/mL) reference, and VEGF (354.8 pg/mL) was used as the picogram per milliliter (pg/mL) reference. By considering the relationship among MSMS events, the percentage of identified proteins from different labeling methods and protein concentration from the referenced proteins (Figure 4), there were 3.6% of identified proteins with MSMS events over 1000 in the single-labeling, whereas there were 2.2% of identified proteins with MSMS event over 1000 in DSIC labeling. In the low-abundance protein range, there were 36% of identified proteins with 2–5 MSMS events in DSIC labeling, whereas there were 26% of identified proteins with 2–5 MSMS events in single labeling. From the discussion above, a large number of MSMS events were assigned to a few high-abundance proteins in single labeling compared to DSIC labeling. Thus, we were able to identify lower abundance proteins with DSIC labeling such as CD14 (IPI00029260), IGFBP2 (IPI00297284), IGFBP3 (IPI00018305), IGFBP4 (IPI00305380), IGFBP5 (IPI00029236), IGF1 (IPI00001610), RAB21 (IPI00007755), RARRES2 (IPI00019176), and SERPINF1 (IPI00006114). We also analyzed protein quantification with single labeling (Supplementary Table 3) and with DSIC labeling (Supplementary Table 4). Figure 5 shows the relationship between MSMS events, number of quantified proteins and protein concentration. In the low-abundance protein range, there were 32% of quantified proteins with 2–5 MSMS events in DSIC labeling, whereas there were 17% of quantified proteins with 2–5 MSMS events in single labeling. There were 4.2% of quantified proteins with MSMS events over 1000 in single labeling, whereas there were 2.7% of quantified proteins with MSMS events over 1000 in DSIC labeling. DSIC labeling yielded more quantified proteins compared to the single-labeling, particularly for proteins with a relatively low

concentration in serum, such as CD14 (IPI00029260), CEACAM19 (IPI00376238), IGF1 (IPI00001610), IGFBP2 (IPI00297284), IGFBP3 (IPI00018305), IGFBP4 (IPI00305380), IGFBP5 (IPI00029236), PCDH24 (IPI00307446), RARRES2 (IPI00019176), PI16 (IPI00301143), and PEDF (IPI00006114).

Conclusion

A novel dual stable isotope-coding (DSIC) technique was developed. This labeling approach is simple and fairly easy to implement since protein samples are reduced and labeled in the same aqueous buffer system without performing any necessary buffer exchange for the second label. DSIC has yielded more accurate measurements and a higher percentage of quantified proteins compared to the single labeling with Cys-alone. The labeling efficiency achieved is greater than 98% for both Cys-acrylamidation and Lys-succinylation. Compared with other current established isotope-labeling approaches such as ICAT and ICPL, DSIC is more affordable and applicable to complex biological sample such as serum.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics*. 2002; 1:376–386. [PubMed: 12118079]
2. Ishihama Y, Sato T, Tabata T, Miyamoto N, Sagane K, Nagasu T, Oda Y. Quantitative mouse brain proteomics using culture-derived isotope tags as internal standards. *Nat. Biotechnol*. 2005; 23:617–621. [PubMed: 15834404]
3. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol*. 1999; 17:994–999. [PubMed: 10504701]
4. Ji J, Chakraborty A, Geng M, Zhang X, Amini A, Bina M, Regnier F. Strategy for qualitative and quantitative analysis in proteomics based on signature peptides. *J. Chromatogr., B: Biomed. Sci. Appl*. 2000; 745:197–210. [PubMed: 10997715]
5. Wang S, Regnier FE. Proteomics based on selecting and quantifying cysteine containing peptides by covalent chromatography. *J. Chromatogr., A*. 2001; 924:345–357. [PubMed: 11521884]
6. Che FY, Fricker LD. Quantitation of neuropeptides in Cpe (fat)/Cpe(fat) mice using differential isotopic tags and mass spectrometry. *Anal. Chem*. 2002; 74:3190–3198. [PubMed: 12141682]
7. Brancia FL, Openshaw ME, Kumashiro S. Investigation of the electrospray response of lysine-, arginine-, and homoarginine-terminal peptide mixtures by liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom*. 2002; 16:2255–2259. [PubMed: 12478568]
8. Brancia FL, Montgomery H, Tanaka K, Kumashiro S. Guanidino labeling derivatization strategy for global characterization of peptide mixtures by liquid chromatography matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem*. 2004; 76:2748–2755. [PubMed: 15144184]
9. Schmidt A, Kellermann J, Lottspeich F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics*. 2005; 5:4–15. [PubMed: 15602776]
10. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlett-Jones M, He F, Jacobson A, Pappin DJ. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics*. 2004; 3:1154–1169. [PubMed: 15385600]

11. Panchaud A, Guillaume E, Affolter M, Robert F, Moreillon P, Kussmann M. Combining protein identification and quantification: C-terminal isotope-coded tagging using sulfanilic acid. *Rapid Commun. Mass Spectrom.* 2006; 20:1585–1594. [PubMed: 16628568]
12. Guillaume E, Panchaud A, Affolter M, Desvergnés V, Kussmann M. Differentially isotope-coded N-terminal protein sulpho-nation: combining protein identification and quantification. *Proteomics.* 2006; 6:2338–2349. [PubMed: 16526082]
13. Gartner CA, Elias JE, Bakalarski CE, Gygi SP. Catch-and release reagents for broad scale quantitative proteomics analyses. *J. Proteome Res.* 2007; 6:1482–1491. [PubMed: 17311443]
14. Panchaud A, Hansson J, Affolter M, Bel Rhlid R, Piu S, Moreillon P, Kussmann M. ANIBAL, stable isotope-based quantitative proteomics by aniline and benzoic acid labeling of amino and carboxylic groups. *Mol. Cell. Proteomics.* 2008; 7:800–814. [PubMed: 18083701]
15. Han DK, Eng J, Zhou H, Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* 2001; 19:946–951. [PubMed: 11581660]
16. Mirgorodskaya OA, Kozmin YP, Titov MI, Korner R, Sonksen CP, Roepstorff P. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards. *Rapid Commun. Mass Spectrom.* 2000; 14:1226–1232. [PubMed: 10918372]
17. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. Proteolytic ¹⁸O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* 2001; 73:2836–2842. [PubMed: 11467524]
18. Liu T, Qian WJ, Strittmatter EF, Camp DG II, Anderson GA, Thrall BD, Smith RD. High-throughput comparative proteome analysis using a quantitative cysteinyl-peptide enrichment technology. *Anal. Chem.* 2004; 76:5345–5353. [PubMed: 15362891]
19. Miyagi M, Rao KC. Proteolytic ¹⁸O-labeling strategies for quantitative proteomics. *Mass Spectrom. Rev.* 2007; 26:121–136. [PubMed: 17086517]
20. Rappsilber J, Ryder U, Lamond AI, Mann M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* 2002; 12:1231–1245. [PubMed: 12176931]
21. Lasonder E, Ishihama Y, Andersen JS, Vermunt AM, Pain A, Sauerwein RW, Eling WM, Hall N, Waters AP, Stunnenberg HG, Mann M. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature.* 2002; 419:537–542. [PubMed: 12368870]
22. Liu H, Sadygov RG, Yates JR. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* 2004; 76:4193–4201. [PubMed: 15253663]
23. Colinge J, Chiappe D, Lagache S, Moniatte M, Bougueleret L. Differential proteomics via probabilistic peptide identification scores. *Anal. Chem.* 2005; 77:596–606. [PubMed: 15649059]
24. Ono M, Shitashige M, Honda K, Isobe T, Kuwabara H, Matsuzuki H, Hirohashi S, Yamada T. Label-free quantitative proteomics using large peptide data sets generated by nano-flow liquid chromatography and mass spectrometry. *Mol. Cell. Proteomics.* 2006; 5:1338–1347. [PubMed: 16552026]
25. Levin Y, Schwarz E, Wang L, Leweke FM, Bahn S. Label-free LC-MS/MS quantitative proteomics for large-scale biomarker discovery in complex samples. *J. Sep. Sci.* 2007; 30:2198–2203. [PubMed: 17668910]
26. Wang F, Ye M, Dong J, Tian R, Hu L, Han G, Jiang X, Wu R, Zou H. Improvement of performance in label-free quantitative proteome analysis with monolithic electrospray ionization emitter. *J. Sep. Sci.* 2008; 31:2589–2597. [PubMed: 18693305]
27. Wang H, Hanash SM. Increased throughput and reduced carryover of mass spectrometry-based proteomics using a high-efficiency nonsplit nanoflow parallel dual-column capillary HPLC system. *J. Proteome Res.* 2008; 7:2743–2755. [PubMed: 18512973]
28. Rauch A, Bellew M, Eng J, Fitzgibbon M, Holzman T, Hussey P, Igra M, Maclean B, Lin CW, Detter A, Fang R, Faca V, Gafken P, Zhang H, Whiteaker J, States D, Hanash S, Paulovich A, McIntosh MW. Computational Proteomics Analysis System (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments. *J. Proteome Res.* 2006; 5:112–121. [PubMed: 16396501]

29. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 2003; 75:4646–4658. [PubMed: 14632076]
30. Sechi S, Chait BT. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* 1998; 70:5150–5158. [PubMed: 9868912]
31. Faca V, Coram M, Phanstiel D, Glukhova V, Zhang Q, Fitzgibbon M, McIntosh M, Hanash S. Quantitative analysis of acrylamide labeled serum proteins by LC-MS/MS. *J. Proteome Res.* 2006; 5:2009–2018. [PubMed: 16889424]
32. Grant-Greene MC, Friedberg F. Studies on succinylated and maleylated ATP: creatine phosphotransferase. *Int. J. Protein Res.* 1970; 2:235–240. [PubMed: 5538391]
33. Zheng R, Regnier FE. Minimizing resolution of isotopically coded peptides in comparative proteomics. *J. Proteome Res.* 2002; 1:139–147. [PubMed: 12643534]
34. Zheng R, Soma CS, Thompson RA, Xiong L, Regnier FE. Controlling detuterium isotope effects in comparative proteomics. *Anal. Chem.* 2002; 74:3662–3669. [PubMed: 12175151]
35. Hamada M, Nagai T, Kai N, Tanoue Y. Influence of added amount of succinic anhydride on the succinylation level for some proteins. *J. Natl. Fish. Univ.* 1997; 46:96–99.

Figure 1.
Scheme of dual stable isotope coding (DSIC) strategy.

Figure 2.

Quantitation of different molar ratios of BSA and Transferrin. The graphs show the linear distribution of three different molar ratios ($^{12}\text{C}:$ $^{13}\text{C} = 1:10$, $1:1$, and $10:1$) Sample was analyzed by *nano* flow-rate RPLC LTQ-FT with a 60 min elution window. Mobile-phase A, 0.1% Formic acid in water; mobile-phase B, 0.1% Formic acid in acetonitrile. Flow rate: 300 nL/min.

Figure 3.

Feature of DSIC labeling method for quantifiable peptide analysis. Large mass difference between the light-label and heavy-label was presented in DSIC approach. Sample was analyzed by *nano* flow-rate RPLC LTQ-FT with a 60 min elution window. Mobile-phase A, 0.1% Formic acid in water; mobile-phase B, 0.1% Formic acid in acetonitrile. Flow rate: 300 nL/min.

Figure 4.

Comparison of identified protein distribution in different MSMS event bins between single-labeling (Cys) and DSIC-labeling (Cys + Lys). Protein concentration range is from the referenced proteins reported from Rules-Based Medicine (http://www.rulesbasedmedicine.com/case-studies/figures/serum_or_plasma-table1.html). The fractionated DSIC labeled immunodepleted serum samples were by *nano* flow-rate RPLC LTQ-FT with a 130 min elution window. Mobile-phase A, 0.1% Formic acid in water; mobile-phase B, 0.1% Formic acid in acetonitrile. Flow rate: 300 nL/min. Proteins were analyzed using Computational Proteomics Analysis Systems through CPAS pipeline (<https://proteomics.fhcrc.org/CPAS/project>) with criteria error arte number of unique peptide 2.

Figure 5.

Comparison of quantified protein distribution in different MSMS event bins between single-labeling (Cys) and DSIC-labeling (Cys + Lys). Protein concentration range is from the referenced proteins reported from Rules-Based Medicine (http://www.rulesbased-medicine.com/case-studies/figures/serum_or_plasma-table1.html). The fractionated DSIC labeled immunodepleted serum samples were by *nano* flow-rate RPLC LTQ-FT with a 130 min elution window. Mobile-phase A, 0.1% Formic acid in water; mobile-phase B, 0.1% Formic acid in acetonitrile. Flow rate: 300 nL/min. Proteins were analyzed using Computational Proteomics Analysis Systems through CPAS pipeline (<https://proteomics.fhcrc.org/CPAS/project>) with criteria error rate number of unique peptide 2.

Table 1Significant Increase in Quantifiable Peptides with DSIC Labeling^a

[M + H] ⁺ 0.8–5.4 (kDa)	Cys labeling			Lys labeling			Cys + Lys (DSIC) labeling		
	total peptides	quantifiable		total peptides	quantifiable		total peptides	quantifiable	
		count	(%)		count	(%)		count	(%)
Human IPI (whole proteome) ^b	1 474 170	413 915	28.1	1 000 756	578 383	57.8	1 002 278	702 127	70.1

^aTrypsin is used as the enzyme for in-silico digestion.^bHuman IPI.v3.53 (Released on Dec 17, 2008).

Table 2Comparison of ICAT, ICPL and DSIC^a

	coded amino acid	form	composition	monoisotopic mass	cost per mg protein	intact protein separation
ICAT	Cysteine (C)	Light	$^{12}\text{C}_{10}\text{H}_{17}\text{O}_3\text{N}_3$	227.1267 (Da)	\$1 000.00	SDS gel
		Heavy	$^{12}\text{C}^{13}\text{C}_9\text{H}_{17}\text{O}_3\text{N}_3$	236.1572 (Da)		
ICPL	Lysine (K)	Light	$^{12}\text{C}_6\text{H}_3\text{ON}$	105.0215 (Da)	\$2 200.00	SDS gel
		Heavy	$^{13}\text{C}_6\text{H}_3\text{ON}$	111.0416 (Da)		
DSIC	Cysteine (C)	Light	$^{12}\text{C}_3\text{H}_5\text{ON}$	71.03711 (Da)	\$20.00	SDS gel or 1D HPLC or 2D-HPLC
		Heavy	$^{13}\text{C}_3\text{H}_5\text{ON}$	74.0466 (Da)		
	Lysine (K)	Light	$^{12}\text{C}_4\text{H}_4\text{O}_3$	100.0160 (Da)	\$40.00	
		Heavy	$^{13}\text{C}_4\text{H}_4\text{O}_3$	104.0295 (Da)		

^aICAT, isotope coded affinity tag; ICPL, isotope-coded protein label; DSIC, dual stable isotope coding.

Table 3

Added Amounts of Acrylamide and the Efficiency of Cys-Acrylamidation

acrylamide	1 (mg)		3 (mg)		5 (mg)		7 (mg)		10 (mg)	
	peak area ^a	MSMS event ^b	peak area	MSMS event	peak area	MSMS event	peak area	MSMS event	peak area	MSMS event
Transferrin	100	100	100	100	100	100	100	100	100	100
BSA	99.7	99.0%	96.4	98.2	100	100	100	100	100	100
Fetuin	100	100	100	100	100	100	100	100	100	100
α 1-Acid glycoprotein	98.8	95.2%	100	100	100	100	100	100	100	100

^aEfficiency (%) = Total peak area of labeled peptides (containing Cys or Lys or both of Cys and Lys)/[Total peak area of labeled peptides (containing Cys or Lys or both of Cys and Lys) + Total peak area of unlabeled peptides (containing Cys or Lys or both of Cys and Lys)].

^bEfficiency (%) = Total number of MSMS event of labeled peptides (containing Cys or Lys or both of Cys and Lys)/[Total number of MSMS event of labeled peptides (containing Cys or Lys or both of Cys and Lys) + Total number of MSMS event of unlabeled peptides (containing Cys or Lys or both of Cys and Lys)].

Table 4

Added Amounts of Succinic Anhydride and the Efficiency of Lys-Succinylation

succinic anhydride	1 (mg)		3 (mg)		5 (mg)		7 (mg)		10 (mg)	
	peak area ^a	MSMS event ^b	peak area	MSMS event	peak area	MSMS event	peak area	MSMS event	peak area	MSMS event
Transferrin	96.9	96.8	99.4	97.7	99.4	98.1	99.7	99.3	99.2	98.1
BSA	98.0	95.0	98.1	97.2	98.1	97.2	99.5	98.0	100	100
Fetuin	100	100	100	100	100	100	100	100	100	100
α 1-Acid glycoprotein	100	100	100	100	100	100	100	100	100	100

^aEfficiency (%) = Total peak area of labeled peptides (containing Cys or Lys or both of Cys and Lys)/[Total peak area of labeled peptides (containing Cys or Lys or both of Cys and Lys) + Total peak area of unlabeled peptides (containing Cys or Lys or both of Cys and Lys)].

^bEfficiency (%) = Total number of MSMS event of labeled peptides (containing Cys or Lys or both of Cys and Lys)/[Total number of MSMS event of labeled peptides (containing Cys or Lys or both of Cys and Lys) + Total number of MSMS event of unlabeled peptides (containing Cys or Lys or both of Cys and Lys)].

Table 5

Influence of Different Enzyme on DSIC-Labeled Standard Protein Digestion

	Trypsin	start	end	Glu-C	start	end
α 1-acid glycoprotein	EHVAHLLFLR	114	123	HVAHLLFLRD	115	124
	SDVMYTDWKKDKCEPLEKQHEKER	171	194	CLCIPRSDVMYTDWKKDKCEPLEKQHE	165	191
Transferrin	WCAVSEHEATKCQSFR	27	42	AYLAPNNLKPVVAE	89	102
	DHMKSVIPSDGPSVACVKKASYLDCIR	43	69	VAFVKHSTIFE	221	231
	KDSGFQMNQLR	122	132	LLCLDNTRKPVDE	224	256
	GKKSCHTGLGR	133	143	FQLFSSPHGKD	301	311
	SAGWNIPIGLLYCDLPEPR	144	162	GTCPEAPTDECKPVKWCALSHHE	348	370
	DQYELLCLDNTR	240	251	TTEDCIAKIMNGE	392	404
	KPVDEYKDCHLAQVPSHTVVAR	252	273	LLCLDGTRKPVVEE	580	592
	EGTCPEAPTDECKPVKWCALSHHER	347	371	YANCHLARAPNHAVVTRKDKE	593	613
	EGYYGYTGAFR	531	541	TKDLLFRDD	645	653
	KPVEEYANCHLAR	588	600			
	KDKEACVHKILR	610	621			
	SETKDLLFR	643	651			
	SETKDLLFRDDTVCLAKLHDR	643	663			
	DDTVCLAKLHDR	652	663			
	NTYEKYLGEYVKA VG NLR	664	682			
	KCSTSSLLEACTFR	683	696			
Fetuin	GYKHTLNQIDSVK VWR	55	71	QAALAAVDYINKHLPRGYKHTLNQID	39	64
	QYGFC KG SVIQKALGGEDVR	226	245	ISRAQFVPLPVSVSVE	185	200
	TPIVGQPSIPGGPVR	334	348	FAVAATDCIAKE	201	212
BSA	ETYGDMADCCEKQEPER	106	122	IAHRFKDLGEE	31	41
	FWGKYLYEIAR	157	167	FAKTCVADE	73	81
	GACLLPKIETMR	198	209	KSLHTLFGDE	88	97
	CASIQKFGER	223	232	LCKVASLRE	98	106
	ALKAWSVAR	233	241	TYGDMADCCE	107	116
	VHKECCHGDLLECADDR	264	280	KKFWGKYLYE	155	164
	RHPEYAVSVLLR	360	371	RALKAWSVARLSQKFPKAEFVE	232	253
	HPEYAVSVLLR	361	371	CADDRADLAKYICDNQDTISSKLKE	276	300
	KVPQVSTPTLVEVSR	437	451	NLPPLTADFAEDKDVCKNYQE	324	344
	SLGKVGTRCCTKPESER	452	468	AKDAFLGSFLYE	345	356
	MPCTEDYLSLILNR	469	482	YAVSVLLRLAKE	364	375
	LCVLHEKTPVSEKVTCCCTESLVNR	483	507	DYLSLILNRLCVLHE	474	488
				SLVNRRCFSALTPDETYVPKAFDE	503	527
				KLFTFHADICTLPDTE	528	543
				NFVAFVDKCCAADDKE	573	588

Table 6Influence of Different Enzyme on DSIC-Labeled Immunodepleted Serum Sample^a

	Trypsin	Glu-C
Number of identified proteins	104	70
Number of quantified proteins	72	45
Percentage of quantified proteins	69%	64%
Overlap of identified protein		84%
Overlap of quantified protein		73%

^a 1. The DSIC-labeled immunodepleted serum sample digested with Trypsin without any prefractionation for this experiment. 2. Trypsin digestion performed at 37 °C overnight and Glu-C digestion performed at room temperature overnight.

Table 7

Results of Single-Labeling (Cys-Coding) and Dual-Labeling (DSIC, Cys- and Lys-Coding) from One of Anion-Exchange Fractionated Serum Sample

	MSMS events ^a	identified proteins ^b	quantified proteins ^a	% of quantified protein
Single-labeling (Cys-coding)	22 216	165	104	63%
Double-labeling (Cys- and Lys-coding)	18 562	185	174	94%

^a Number of MSMS spectra generating sequenced peptides with Peptide Prophet 0.2.

^b Number of identified proteins with the criteria: 2 unique peptides and 5% error rate.

^c Number of quantified proteins with the criteria: 2 unique peptides, 5% error rate and peptide prophet 0.75.

Table 8

Distribution of MS/MS Events *versus* Protein Concentration Levels between Single Labeling and DSIC Labeling

protein concentration		~500 (pg/mL)	10–500 (ng/mL)	1–100 (μg/mL)	0.1–1.0 (mg/mL)	total MSMS events	
No. MSMS events		2–5	6–25	26–100	101–1000	>1001	
Identification	Single labeling (%)	0.8	3.7	5.9	18.2	71.4	22 216
	DSIC labeling (%)	1.3	3.4	6.2	56.8	32.7	18 562
Quantification	Single labeling (%)	0.7	4.9	7.2	22.6	64.6	12 453
	DSIC labeling (%)	1.2	3.3	5.8	57.0	32.9	18 440