Regular Article

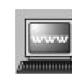
Increased quantitative proteome coverage with $^{13}C/^{12}C$-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme

Eugene C. Yi[1], Xiao-jun Li[1], Kelly Cooke[1], Hookeun Lee[1], Brian Raught[1], Andrew Page[2], Victoria Aneliunas[2], Phil Hieter[2], David R. Goodlett[1] and Ruedi Aebersold[1]

[1] Institute for Systems Biology, Seattle, WA, USA
[2] Center for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, British Columbia, Canada

Quantitative protein profiling using the isotope-coded affinity tag (ICAT™) method and tandem mass spectrometry (MS) enables the pair-wise comparison of protein expression levels in biological samples. A new version of the ICAT reagent with an acid-cleavable bond, which allows removal of the biotin moiety prior to MS and which utilizes ^{13}C substitution for ^{12}C in the heavy-ICAT reagent rather than ^{2}H (for ^{1}H) as in the original reagent (Gygi, S. P., Rist, B., Gerber, S. A., Frantisek, T., Gelb, M. H., Aebersold, R., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol*. 1999, *17,* 994–999), was investigated. We developed and validated an MS data acquisition strategy using this new reagent that results in an increased number of protein identifications per experiment, without losing the accuracy of protein quantification. This was achieved by following a single survey (precursor) ion scan and serial collision-induced dissociations (CIDs) of four different precursor ions observed in the prior survey scan. This strategy is common to many high-performance liquid chromatography-electrospray ionization (HPLC-ESI)-MS shotgun proteomic strategies, but heretofore not to ICAT experiments. This advance is possible because the new ICAT reagent uses ^{13}C as the "heavy" element rather than ^{2}H, thus, eliminating the slight delay in retention time of ICAT-labeled "light" peptides on a C18-based HPLC separation that occurs with ^{2}H and ^{1}H. Analyses using this new scheme of an ICAT-labeled trypsin-digested six protein mixture as well as a tryptic digest of a total yeast lysate, indicated that about two times more proteins were identified in a single analysis, and that there was no loss in accuracy of quantification.

Received: March 31, 2004
Revised: June 16, 2004
Accepted: July 21, 2004



Keywords:
Cation-exchange chromatography / Isotope-coded affinity tag / Quantitative analysis / Reversed-phase microcapillary chromatography

1 Introduction

During the past several years, quantitative proteome profiling has benefited from incorporating the traditional stable isotope dilution theory into global scale or discovery-based proteomic experiments that use mass spectrometers as detectors to allow the pair-wise study of differentially expres-

Correspondence: Dr. Eugene C. Yi, Institute for Systems Biology, 1441 N 34th, Seattle, WA 98103, USA
E-mail: eyi@systemsbiology.org
Fax: +1-206-732-1299

Abbreviations: ICAT, isotope-coded affinity tag; **MS^2**, tandem mass spectrometry; **SCX**, strong cation-exchange; **SIC**, single ion chromatogram; **TCEP**, tris(2-carboxyethyl)phosphine

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

sed proteins [1–6]. Two basic scenarios to incorporate stable isotopes into peptides, *i.e.*, either *in vivo* labeling during cell culture, often referred to as metabolic labeling [7], or *in vitro* chemical labeling post-isolation, have been described [8]. The *in vitro* labeling method was pioneered by Gygi *et al.* [9], by introducing isotope-coded affinity tags (ICAT™), as reagents used in these experiments (Fig. 1A). To simplify the complex peptide pool for subsequent mass spectrometric (MS) analysis, while maintaining high proteome coverage, the first-generation ICAT reagents were conjugated to peptides by selective alkylation of cysteine residues [10, 11]. Cysteine was chosen because as a rare amino acid it allows the complexity of the subsequent pool of peptides to be reduced by approximately an order of magnitude, but being a functionally essential amino acid, most proteins have at least one cysteine. These original ICAT (o-ICAT) reagents were made as either an isotopically heavy form, containing eight deuterium atoms (*i.e.*, d8-ICAT), or an isotopically light form containing no deuterium (d0-ICAT) (Fig. 1A). The new version of ICAT reagents incorporate an acid-cleavable bond, which allows removal of the biotin moiety prior to MS and uses ^{13}C substitution for ^{12}C in the heavy-ICAT reagent rather than ^{2}H (for ^{1}H) as was used in the first generation reagent (Fig. 1B). The use of ^{13}C as the heavy element ensures that ICAT-labeled peptides co-elute on a reversed-phase C18-based support, which is important for quantification [12–14]. Figure 2 provides a flow chart of the ICAT sample preparation process modified for the newly available acid cleavable ICAT (cl-ICAT); regardless of reagent status (*i.e.*, old or new) the protocol is identical, except for the final acid cleaving step [10, 11].

Two protein mixtures, one labeled with the light-ICAT reagent and the other with the heavy-ICAT reagent, are combined and digested with trypsin (Fig. 2). The complex peptide mixture is then fractionated by strong cation-exchange chromatography (SCX), and cysteine-containing ICAT-labeled peptides are purified from unlabeled peptides by passing the mixture over a monomeric avidin-agarose column. Each SCX-avidin fraction of ICAT-labeled peptides is subjected to on-line reversed-phase microcapillary column (RP-LC) analysis using a C18-based support. Peptides eluting from

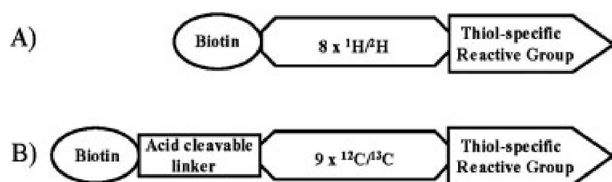


Figure 1. Block diagram of the (A) original and (B) acid-cleavable ICAT™ reagents. The original ICAT reagent consists of three elements: (i) thiol-specific reactive group that binds to cysteine residues; (ii) a nondeuterated (light reagent) or deuterated (heavy reagent, where 8 ^{2}H atoms replaced 8 ^{1}H atoms, and (iii) a biotin tag used for affinity isolation of the ICAT-labeled peptides. The acid-cleavable ICAT reagent uses nine ^{13}C atoms to replace nine ^{12}C atoms and has an acid susceptible linker that allows biotin to be removed when incubated with strong acid.

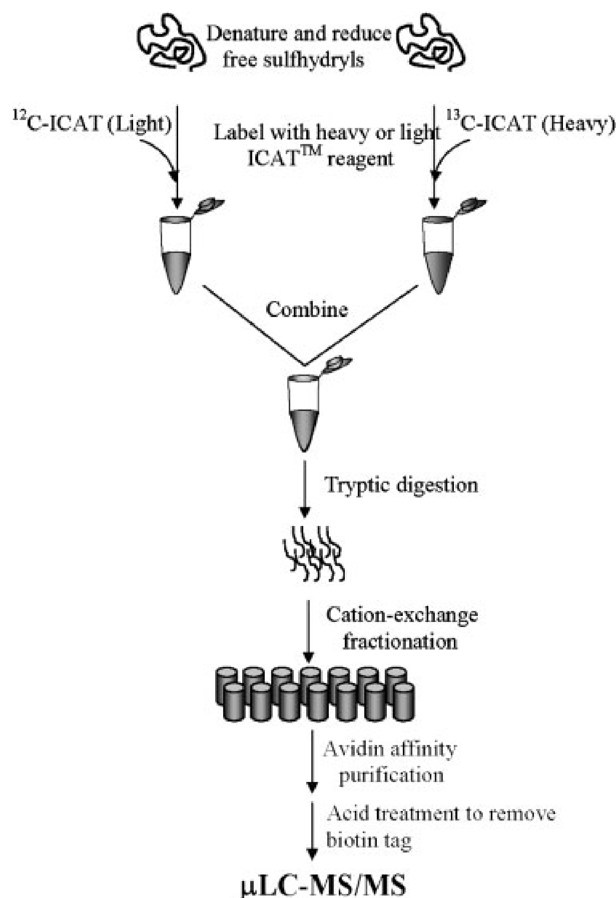


Figure 2. Flow chart of the quantitative analysis of protein profiles with cleavable ICAT™ reagent.

this column are ionized by ESI and specific ions are selected for collision induced dissociation (CID) in a data-dependent fashion that most often begins with the base peak in a survey scan. The instrument software then sequentially selects peptides for CID in descending order of signal intensity, *i.e.*, top-down, and continuously alternates between MS survey scan and tandem MS mode throughout the chromatographic separation. The information from each of these RP-µLC-MS/MS analyses is then converted into protein identifications and relative protein ratios using software tools developed at the Institute for Systems Biology [3, 15].

The initial implementation of the ICAT-based MS data acquisition scheme, where each MS scan is followed by a single tandem MS (MS^2) scan, produced good-quality ICAT ratios, but the number of protein identifications was limited because 50% of the scans during each µLC separation were survey scans. Adopting the more conventional approach, where each survey scan is followed by multiple MS^2 scans on unique parent ions would obviously increase the number of identifications. In fact, if each survey scan could be followed by four MS^2 scans of unique parent ions, this would quadruple the number of MS^2 conducted. However, when this

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

acquisition scheme was conducted on peptides labeled with the original ICAT reagents, the quality of the observed ICAT ratios decreased over those obtained using a single survey scan followed by a single MS^2 scan. This problem arose because the original ICAT-reagents used ^2H as the heavy element, and heavy ICAT-labeled peptides did not co-elute with the light ICAT-labeled peptides of identical sequence. Attempts to increase the number of MS^2 scans between survey scans resulted in ICAT ratios with greater variability because too much chromatographic time passed prior to returning to the MS scan mode from which single ion current traces were generated. Because peptides labeled with the new cl-ICAT reagents co-elute, we were able to explore the more common ion trap data acquisition scheme, where each MS scan is followed by CID of multiple unique parent ions prior to returning to a new survey scan. Here, we present the results of our study comparing the quality of data generated using the new cl-ICAT reagents and new acquisition scheme. We clearly observed twofold increase in the number of proteins identified in a single analysis, without loss of quantification data quality.

2 Materials and methods

2.1 ICAT labeling standard proteins with deuturated reagent

The standard proteins used were bovine serum albumin (BSA), chicken ovalbumin (OVAL), bovine catalase (CATA), bovine α-lactalbumin (LCA), horse myoglobin (MYG), and bovine serotansferrin precursor (TRFE) (Sigma, St. Louis, MO, USA). Labeling was performed as described previously [3–7, 10, 11]. 100 µg proteins in two separate tubes was dissolved in 100 µL of labeling buffer (0.05% w/v SDS, 6 M urea, 5 mM EDTA, 200 mM Tris buffer, pH 8.3). Disulfide bonds were reduced by adding tris(2-carboxyethyl) phosphine (TCEP) (Pierce, Rockford, IL, USA) to a final concentration of 5 mM. A fivefold molar excess of ICAT reagent (Applied Biosystems, Foster City, CA, USA) was added to the reduced protein samples to achieve a final concentration of 1.2 mM. The reactions were mixed well and gently shaken using a tube rocker for 90 min at room temperature in the dark. The reactions were quenched by adding approximately 10-fold molar excess of DTT and followed by incubating at room temperature for 5 min. Prior to tryptic digestion, both light and heavy ICAT-labeled protein mixtures were combined and diluted with water to a 1 M final urea concentration. To the diluted protein mixtures, trypsin of 1:50 w/w (trypsin/protein) was added. The mixture was then incubated overnight at 37°C and quenched by adding acetic acid to pH 3.0. The trypsinized sample was desalted by a cation-exchange cartridge and ICAT-labeled peptides were then isolated by passing over a monomeric avidin cartridge following the manufacturer's protocol (Applied Biosystems).

2.2 ICAT labeling standard proteins with cleavable ICAT reagent

The labeling procedure was identical as described for labeling with the deuturated ICAT reagent except that the reagent was ^{13}C-based acid-cleavable ICAT reagent (Applied Biosystems). The biotin affinity tag from ICAT-labeled peptides was removed following the procedure and the reagents provided by the manufacturer (Applied Biosystems). Samples were dried under vacuum and resuspended in 0.1 mL of a proprietary cleaving/scavenging reagent (Applied Biosystems), then incubated at 37°C for 120 min. Cleaved samples were dried under vacuum, and resuspended in 6% CH_3CN, 0.1% TFA.

2.3 Preparation and labeling yeast protein extracts with cleavable ICAT reagents

Two 1 L *Saccharomyes cerevisiae* grown in YPD media was centrifuged (1500 × *g* for 10 min at 4°C), and the cell pellets were washed twice in 1.5 mL PBS containing 50 µM PMSF. Pellets were solubilized in 1 mL 10% TCA and incubated on ice for 60 min. The resulting protein precipitates were pelleted (20 000 × *g* for 10 min at 4°C) and washed twice in 1 mL 80% ice-cold acetone. Washed protein pellets were resuspended in 1 mL ICAT labeling buffer (200 mM Tris, pH 8.3, 6 M urea, 5 mM EDTA, 0.05% w/v SDS), and the protein concentration of each sample was determined using the bicinchoninic acid assay (Pierce). An aliquot of 3 mg of each sample was adjusted to 1 mL volume with ICAT labeling buffer. Samples were reduced by the addition of TCEP to 5 mM and incubation at 37°C for 60 min. Either the isotopically heavy (^{13}C) or light (^{12}C) cleavable ICAT reagent (ABI, Foster City, CA, USA) was added to a concentration of 1.4 mM, and each reaction allowed to proceed at 37°C for 2 h. DTT was then added to 12 mM, and the samples were incubated for 5 min at room temperature to quench the labeling reaction. The two ICAT-labeled samples were combined and brought up to 15 mL with 20 mM Tris, pH 8.3, 5 mM EDTA. Sequencing-grade trypsin (160 µg, approx. 1/40 w/w; Promega, Madison, WI, USA) was added, and the combined sample was digested overnight at 37°C. The trypsinized sample was adjusted to pH 2.5 with 85% phosphoric acid. Using the Vision™ workstation (ABI), tryptic peptides were fractionated *via* SCX HPLC (2.1 × 200 mm polysulfoethyl A; PolyLC, Columbia, MD, USA). Peptides were eluted with buffer B (5 mM KH_2PO_4, pH 3, 600 mM KCl, 25% CH_3CN) using a gradient of 0 – 15% buffer B over 30 min, 15 – 60% over 20 min, and 60 – 100% over 15 min, at 0.2 mL/min. Thirty fractions of 0.4 mL were collected. Each cation-exchange fraction was then run over an avidin cartridge (Applied Biosystems), and washed sequentially with 2 × PBS, pH 7.2, 1 × PBS and 50 mM NH_4HCO_3, pH 8.3/20% CH_3OH. Biotinylated peptides were eluted with 400 µL 0.4% trifluoroacetic acid (TFA)/30% CH_3CN into 1/10th volume of neutralizing buffer (100 mM Na_2PO_4, pH 10). Samples were dried under vacuum, and resuspended in 0.1 mL of a proprietary cleaving/scavenging reagent (ABI), then incubated at 37°C for 2 h. Cleaved samples were dried under vacuum, and resuspended in 6% CH_3CN, 0.1% TFA.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

2.4 Protein identification and quantification by automated data-dependent MS^2 with modified data acquisition schemes

The ICAT-labeled peptides were analyzed by LC-ESI-MS/MS using an ion-trap mass spectrometer (LCQ-DECA; Thermofinnigan, San Jose, CA, USA). MS data were acquired by varying the number of CID attempts on different peptide ions prior to returning the instrument to survey the scan mode. For example, the instrument would continuously alternate between (i) a survey scan and a single CID event (MS $\times$ MS^2) or (ii) a single survey scan followed by up to four sequential data-dependent MS/MS scans before returning to a survey scan (MS $\times$ 4 MS^2). The ICAT labeled peptides were identified using the SEQUEST, PeptideProphet, and ProteinProphet [16–18]. Relative quantification was performed using ASAPRatio software programs [3, 15]. Peptides and proteins were considered as "identified" probability scores were $\geq$ 0.9. Detail features of the software are available at: http://www.proteomecenter.org/software.php).

3 Results and discussion

3.1 Quantification of a standard protein mix using the acid cleavable ICAT reagent and a modified data acquisition scheme

A standard six protein mixture was labeled using either the original ICAT or the new acid-cleavable ICAT reagents. Two separate samples, one using the cl-ICAT and one the o-ICAT, were prepared such that a 1:1 heavy:light ratio of peptides would be observed for each when analyzed by RP-µLC-MS analysis. These two samples were used to test three different MS data acquisition schemes, where each survey scan was followed by the selection of either 1, 2, or 4 unique precursor ion for CID before returning to a new survey scan. Data obtained from the different schemes were analyzed to test whether the accuracy of quantification was maintained as the number of MS/MS (MS^2) scans prior to returning to the MS survey scan was increased. Data were analyzed using SEUQEST to match peptide tandem mass spectra to peptide sequences in a database, and ASAPRatio [15] to calculate relative ratios of o- and cl-ICAT-labeled peptide pairs. Prior to quantification, SEQUEST database search results were verified using PeptideProphet [17], a software tool that evaluates an accurate probability of the observed tandem mass spectrum matching a given peptide sequence in the database.

Figure 3 shows the reconstructed single ion chromatogram (SIC) of peptide pairs labeled with isotopically light and heavy forms of both o-ICAT and cl-ICAT reagents. The SIC chromatograms shown in Fig. 3 were obtained from four different data acquisition schemes. Fig. 3A shows the results from use of the o-ICAT reagent, where each survey scan was followed by a single MS^2 scan. With the cl-ICAT reagent three acquisition schemes were tested. Each survey scan was followed by either one (Fig. 3B), two (Fig. 2C) or four MS^2 scans (Fig. 3D) on unique precursor ions. In Fig. 3A the heavy o-ICAT-labeled peptide in the bottom panel Fig. 3A elutes prior to the light o-ICAT-labeled peptide in the top panel of Fig. 3A. Since ion selection for CID from complex samples is random [16], typically only one of the ICAT-labeled peptide pairs is subjected to CID in the time available in a single µLC experiment. When peptides labeled with the o-ICAT reagent are analyzed using a data acquisition scheme that carries out multiple MS^2 events prior to returning to the survey scan, it

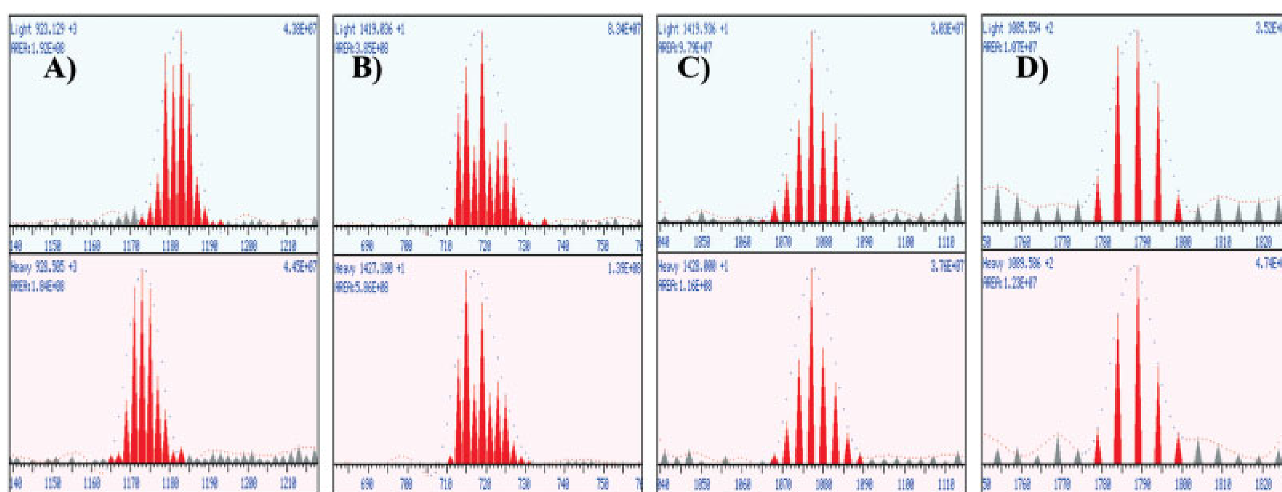


Figure 3. Reconstructed SICs for individual ICAT-labeled peptides from RP-µLC/MS analysis of peptides labeled with isotopically light and heavy forms of (A) the original ICAT™ reagents (d0/d8) using a data acquisition scheme that continuously alternates between 1 survey scan and 1 MS^2 scan; and the cleavable ICAT™ reagents (^{12}C9/^{13}C9) using a data acquisition scheme that continuously alternates between 1 survey scan; and (B) 1 MS^2 scan, (C) 2 MS^2 scans, or (D) 4 MS^2 scans. The reconstructed peptide elution profiles were obtained using the XPRESS quantification software program [3].

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

can be difficult to locate the partner ICAT-labeled peptide because the time it takes to return to survey scan mode may be longer than the duration of the chromatographic peak of the ICAT-peptide pair. Consequently, it was observed when using the o-ICAT reagents that the accuracy of relative quantification decreased as the number of MS^2 events following each survey scan is increased beyond one (data not shown). This forced the data to be acquired by continuously switching between MS and MS^2 modes, rather than spending more time in MS^2 mode between survey scans. However, by virtue of ^{13}C incorporation in the cl-ICAT reagent, the heavy and light cl-ICAT labeled peptides co-elute during C18-based RP-μLC, making location of the begin/end of the SIC trace for the peptide that was not selected for CID much easier. Panels B, C, and D of Fig. 3 contain the results from use of the three different data acquisition schemes with peptides labeled with the cl-ICAT reagent. In each case the peptide pair labeled with the light cl-ICAT reagent is shown in the top panel and the peptide labeled with the heavy cl-ICAT reagent in the bottom. Here one can easily observe the co-elution of cl-ICAT-labeled peptide pairs and location of the begin/end of each SIC is easier than with the o-ICAT reagents.

Table 1 shows results from experiments where the quantification software ASAPRatio was used to measure relative ratios of select identified peptides obtained from data acquired as described in Fig. 3. Because of the manner in which the standard solutions were prepared, the calculated and observed relative ICAT ratios are expected to be 1:1. Any deviation between calculated and observed relative ratios should be a result of the experimental conditions. Table 1 shows that the error assignments between observed and expected vary from 1.1% for chicken ovalbumin to 38.4% for bovine catalase. The ICAT ratios of individual proteins were calculated by taking the average heavy-to-light (H:L) ratio of uniquely identified ICAT-labeled peptides, weighed by their errors [15]. While there is a slight difference in the ratios from the theoretical ratio of 1:1 between peptides labeled with the o- and cl-ICAT reagents, the different MS data acquisition schemes (survey scan followed by 1, 2, or 4 unique ions selected for CID) gave fairly consistent results for each protein. Results from comparison of the three different data acquisition schemes using the heavy and light cl-ICAT-labeled peptides provided similar ratios to those shown in column two for the o-ICAT reagent. The average error from all four experiments in Table 1 increased from ~ 10 to 20% as the number of MS^2 events increased from 1 to 4, and average precision decreased slightly. Note that when five or more unique ions were selected for CID prior to returning to a survey scan the accuracy in quantification was greater than 20% (data not shown). This was primarily because there were not enough MS survey data points to accurately reconstruct peptide SIC elution profiles, *i.e.*, too much time elapsed prior to returning to each survey scan. It is expected that this problem will be exaggerated for peptides with a low signal-to-noise ratio. Thus, while it could lead potentially to more protein identifications, it was decided not to pursue use of more than four MS^2 events between each new survey scan because quantitative error was outside acceptable limits.

3.2 Application to protein expression profiling of yeast

We further evaluated the use of the different data acquisition schemes to determine if there was the expected increase in the number of protein identifications from addition of extra MS^2 events per survey scan. As a test, a sample of a yeast lysate where proteins were labeled with the cl-ICAT-reagents was used. Data was collected using four different MS data acquisition schemes (selection of 1, 2, 3, or 4 unique ions selected for CID before returning to a survey scan) for a few of SCX HPLC fractions from a larger experiment using cl-ICAT labeled yeast lysate proteins. We then analyzed and compared the efficiency of the different MS data acquisition schemes in terms of protein identification and quantification.

Table 1. Quantification results on a six protein (1:1) mixture

Protein[a]	o-ICAT 1 MS^2[b]	cl-ICAT 1 MS^2[b]	cl-ICAT 2 MS^2[b]	cl-ICAT 4 MS^2[b]	Rel. diff (%)[c]
SW:ACTA_HUMAN	1.15 (0.20)	0.98 (0.12)	1.44 (0.06)	1.11 (0.21)	13.2
SW:ALBU_BOVIN	1.03 (0.06)	0.89 (0.03)	0.74 (0.05)	0.77 (0.05)	11.5
SW:CATA_BOVIN	1.13 (0.20)	1.12 (0.16)	1.13 (0.24)	0.69 (0.07)	38.4
SW:LCA_BOVIN	0.96 (0.06)	0.91 (0.07)	0.96 (0.06)	0.88 (0.09)	3.3
SW:OVAL_CHICK	1.06 (0.06)	0.95 (0.08)	1.03 (0.07)	0.94 (0.05)	1.1
SW:TRFE_BOVIN	0.98 (0.03)	1.24 (0.05)	1.17 (0.06)	1.13 (0.05)	8.9
Average of proteins	1.05 (0.08)	1.02 (0.14)	1.08 (0.23)	0.92 (0.18)	12.7
Average rel. error (%)	9.3	8.3	8.4	9.3	

a) Peptides were filtered by probability ≥ 0.9
b) Results were reported as ratio (± error) (light:heavy)
c) Between cl-ICAT 1 MS^2 and cl-ICAT 4 MS^2

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

We first evaluated the data from the two data acquisition schemes (selection of one and four unique ions selected for CID) using an in-house developed software visualization tool called Pep3D (19). As shown in Fig. 4, the software is used to represent LC-MS data by plotting *m/z versus* chromatographic time *versus* relative ion intensity. Importantly, it also allows the user to annotate all ions that (i) were selected for CID during an LC-MS/MS analysis and (ii) that subset of ions that were selected for CID and resulted in a sequence match in the database. This software provides a visual display of the efficiency of the LC-MS/MS process. Results from the yeast lysate analysis are shown in Fig. 4, where ions selected for CID are shown in blue and ions selected for CID that matched a peptide sequence in a database with a PeptideProphet probability score ≥0.9 are shown in hues ranging from green to red. All remaining *m/z* values that were not selected for CID are displayed in the gray scale. Fig. 4 shows results where each survey scan was followed by one MS^2 (A and C) or four MS^2 (B and D). Not surprisingly, the percentage of MS^2 scans leading to positive identification during each LC-MS/MS analysis is roughly the same, at 11% (244/2174) and 12% (495/4029). However, the total number of peptide matches exceeding the 0.9 probability score cutoff is about two-times higher when each survey scan is followed by four MS^2 scans. Figures 4C and D show an expansion of the region from 800–1000 *m/z* and 50 to 113 min regions from Figs. 4A and B, respectively. In these expanded plots it is obvious that more ions were selected for CID in panel D than C, however, the percentage of those acquired tandem mass spectra matched to a peptide sequence was roughly equivalent, at 28% (95 out of 338) for experiments shown in A and C and 26% (167 out of 652) for those shown in B and D.

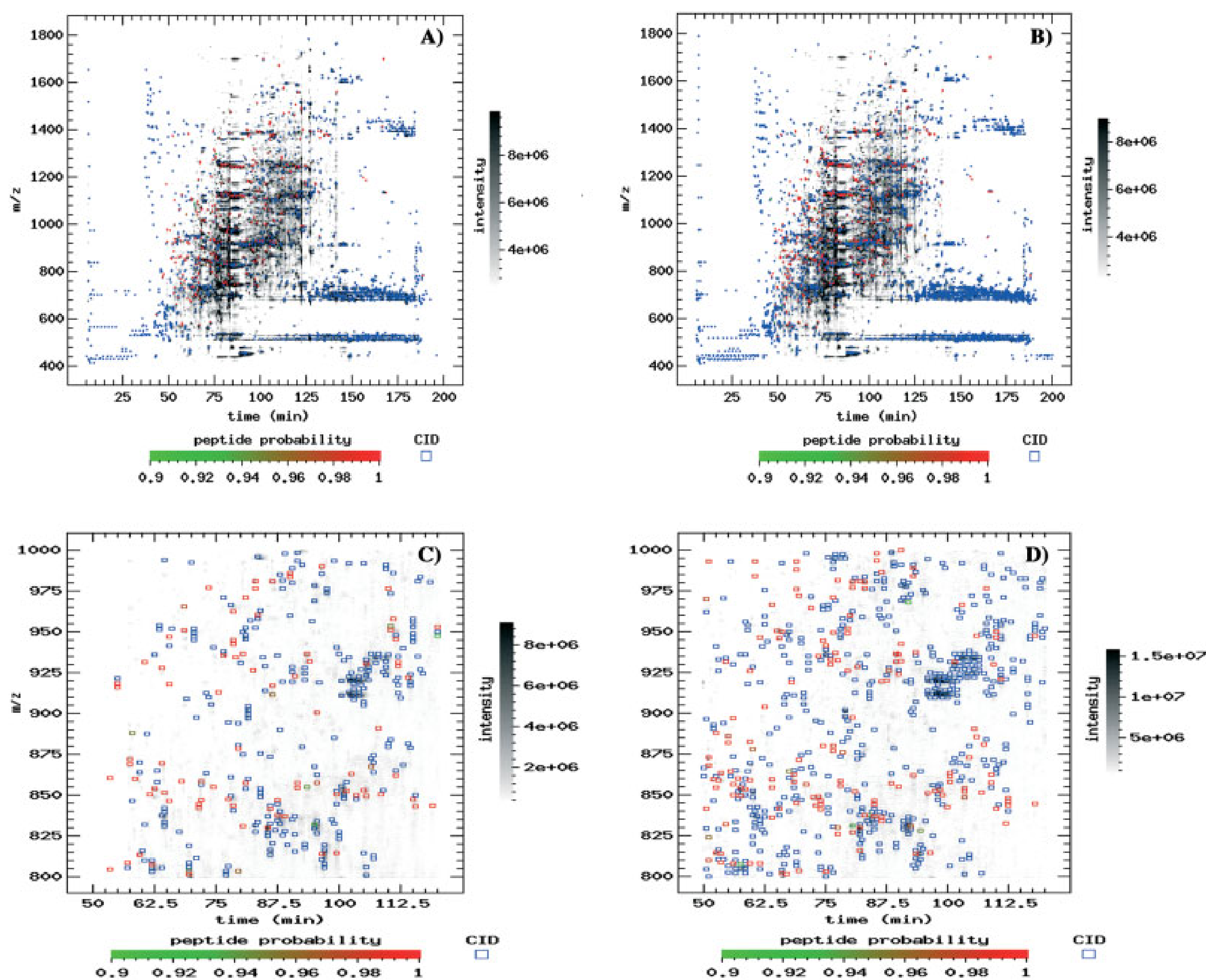


Figure 4. Pep3D display of MS scans, MS/MS scans, and identified peptides with ≥0.9 probability score. The ratio of identified peptides to MS/MS attempted from one survey scan followed by (A) a single MS/MS and (B) four MS/MS data acquisition schemes are 244/2174 (11%) and 495/4029 (12%), respectively. Pep3D display of selected mass range and retention time of A (C) 95/338 (28%) and B (D) 167/652 (26%).

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 2 shows the number of peptides, the number of unique peptides, and the number of proteins identified from 2 of the 40 SCX fractions from a larger ICAT experiment, analyzed using the four different data acquisition schemes. Peptides and proteins were considered as "identified" (*i.e.*, tandem mass spectrum matched a sequence in the database) when their PeptideProhpet and ProteinProphet probability score were ≥ 0.9. Since the same peptide sequence may be identified multiple times, each identification counts as a peptide while each identified peptide sequence counts as a unique peptide. About 1/3 unique peptides were identified only once. Clearly both the number of tandem mass spectra that match a peptide sequence and the number of protein identifications increase as the number of MS^2 scans increase. When four MS^2 scans (1×4 in Table 2) instead of a single MS^2 scan follow each survey scan there was a 67% increase in the number of matches of peptide sequence to tandem mass spectra, which in turn led to a 41% increase in protein identifications. An increase in the number of proteins identified from a single peptide hit was in the same order of magnitude. Note that the number of peptide sequence matches with $P \geq 0.9$ does not grow linearly with the increase in number of MS^2 scans. This has more to do with the fact that a certain amount of MS^2 scans are wasted when a peptide ion is selected at suboptimal times during the LC introduction. For example, consider that a peptide ion selected early in its elution profile may not result in the production of as high a quality tandem mass spectrum as one selected at the apex of the elution profile, where the effective concentration is highest. Nevertheless, these results demonstrate, as expected, that one survey scan followed by four MS data acquisition schemes allows more CID spectra to be acquired and more peptide tandem mass spectra to be matched with sequences in a database.

We then examined the question whether or not determination of the ICAT ratio would change by increasing the length of time between survey scans which alone are used for quantitative determination. In order to find the accuracy and consistency in measuring protein abundance ratios from data acquired by the different MS acquisition schemes, the software program ASAPRatio [15] was used. For each protein, ASAPRatio evaluated its abundance ratio and the corresponding standard deviation. The relative error, which is defined as the standard deviation of a protein abundance ratio divided by the ratio, describes the accuracy of an experimental measurement. The relative error in the different data acquisition schemes was compared as shown in Fig. 5A. The accumulated distribution of protein relative error, which describes the fraction of proteins whose relative error is less than or equal to a given value, in the different data acquisition schemes remained close to each other over the whole range. These results indicate that one survey scan followed by four MS scans is comparable to other data acquisition schemes in the accuracy of the protein ratios. The relative difference is defined as $2|r1-r2|/(r1+r2)$ where $r1$ and $r2$ are ratios of the same protein but measured in two different data acquisition schemes. The relative difference describes the consistency between the two experimental measurements. The accumulated distribution of the relative difference, which describes the fraction of proteins whose relative difference is less than or equal to a given value, between protein abundance ratios measured in the 4 MS^2 data acquisition scheme, and the corresponding ratios measured in other schemes is plotted in Fig. 5B. It shows that the same measurement of the relative ICAT ratios is consistent from one MS^2 scan out to four MS^2 scans.

Table 2. Numbers of peptides and proteins identified by the four different data acquisition schemes

Setup[a)]	Peptides[b)]	Unique[c)]	Proteins	Avg. peptides/ protein	Singlets[d)]
1 × 1	463	340	160	2.89	79
1 × 2	682	463	190	3.59	79
1 × 3	833	570	220	3.79	79
1 × 4	883	567	226	3.91	88

a) 1 × 1 = 1 MS followed by 1 MS^2, *etc.*
b) All peptides were filtered with protein probability ≥ 0.9, as determined by ProteinProphet
c) Unique peptides.
d) Proteins identified with a single hit

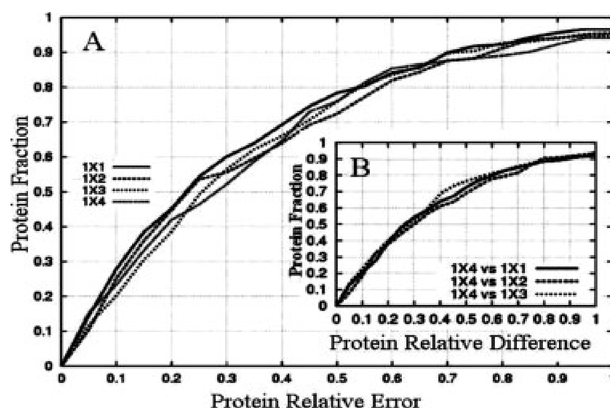


Figure 5. (**A**) Accumulated distribution of relative error of yeast protein abundance ratios in different μLC-MS^2 data acquisition schemes. Relative error is defined as the standard deviation of a protein ratio divided by the ratio. Protein fraction represents the corresponding fraction of proteins whose relative error is less than or equal to a given value. Different μLC-MS^2 data acquisition schemes are indicated by different line types in the figure. (B) Accumulated distribution of protein relative difference of yeast protein abundance ratios obtained in 4 MS^2 scheme and in other schemes. The relative difference is defined as $2|r1-r2|/(r1+r2)$ where $r1$ and $r2$ are ratios of the same protein but measured in two different data acquisition schemes. The protein fraction gives the corresponding fraction of proteins whose relative difference is less than or equal to a given value. Comparisons between different μLC-MS^2 data acquisition schemes are indicated by different line types in the figure, where each survey scan is followed by 1 MS^2 (1X1), 2 MS^2 (1X2), 3 MS^2 (1X3), and 4 MS^2 (1X4) scans prior to returning to a new survey scan.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

4 Concluding remarks

Substitution of ^{12}C in the new cl-ICAT reagent with ^{13}C to create the heavy reagent instead of ^{2}H for ^{1}H as was used in the o-ICAT reagent, eliminates a slight delay in elution time between the heavy and light ICAT-labeled peptides during HPLC separation on a C18-based support. We demonstrated that this simple change allowed a more efficient type of data acquisition scheme to be used with the ICAT strategy. Instead of single-peptide tandem mass spectra are acquired between survey scans, multiple MS^2 scans were performed. The optimal number of MS^2 scan following on MS scan was 4. The total acquisition time of survey MS scan followed by four MS^2 scans was 7.2 s, while the same number of MS^2 scans required 19.2 s if survey MS scan was followed by only a single MS^2 scan. It is apparent that the multiple MS^2 scan approach offers 2.5 times more CID attempts than the alternate MS and MS^2 scan approach in the time elapsed. This greatly increased the number of protein identifications from any single LC-MS analysis, without compromising the quantitative results, because the use of a $^{13}C/^{12}C$ combination for heavy/light ICAT reagents eliminates any chromatographic delay from a C18 support. Furthermore, we acquired up to four tandem mass spectra of peptides between survey scans without any loss in quantitative information. Finally, when we analyzed a complex tryptic digest of all soluble yeast proteins, we increased the number of protein identifications by 33% by acquiring four instead of one tandem mass spectrum between each survey scan.

This project has been funded in part with federal funds from the National Heart, Lung, and Blood Institute, National Institute of Health, under contract No. NO1-HV-28179 and from National Institutes of Health grant No. R33 CA93302. The ISB is the recipient of a generous gift from Merck and Co.

5 References

[1] Aebersold, R., Goodlett, D. R., *Chem. Rev.* 2001, *101*, 269–295.

[2] Goodlett, D. R., Yi, E. C., *Funct. Integr. Genomics* 2002, 138–153.

[3] Han, D. K., Eng, J., Zhou, H., Aebersold, R., *Nat. Biotechnol.* 2001, *19*, 946–951.

[4] Baliga, N. S., Pan, M., Goo, Y. A., Yi, E. C. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *23*, 14913–14918.

[5] Shiio, Y., Donohoe, S., Yi, E. C., Goodlett, D. R., Aebersold, R., Eisenman, R. N., *EMBO J.* 2002, *21*, 5088–5096.

[6] Guina, T., Purvine, S.O., Yi, E. C., Eng, J. *et al.*, *PNAS* 2003, *100*, 2771–2776.

[7] Ong, S. E., Kratchmarova, I., Mann, M., *J. Proteome Res.* 2003, *2*, 173–181.

[8] Aebersold, R., Mann, M., *Nature* 2003, *422*, 198–207.

[9] Gygi, S. P., Rist, B., Gerber, S. A., Frantisek, T. *et al.*, *Nat. Biotechnol.* 1999, *17*, 994–999.

[10] *Cleavable ICAT™ Reagent Kit for Protein Labeling, Applied Biosystems, Foster City, CA, USA*.

[11] Yi, E. C., Goodlett, D. R., (Eds.), *Current Protocols in Protein Science*, Wiley, New York, 2004.

[12] Zhang, R., Sioma, C. S., Wang, S., Regnier, F. E., *Anal. Chem.* 2001, *73*, 5142–5149.

[13] Oda, Y., Owa, T., Sato, T., Boucher, B. *et al.*, *Anal. Chem.* 2003, *75*, 2129–2165.

[14] Hansen, K. C., Schmitt-Ulms, G., Chalkley, R. J., Hirsch, J. *et al.*, *Mol. Cell Proteomics* 2003, *2*, 299–314.

[15] Li, X.-J., Zhang, H., Ranish, J. A., Aebersold, R., *Anal. Chem.* 2003, *75*, 6648–6657.

[16] Eng, J. K., McCormack, A. L., Yates III, J. R., *J. Am. Soc. Mass Spectrom.* 1994, *5*, 976–989.

[17] Keller, A., Nesvizhskii, A. I., Kolker, E., Aebersold, R., *Anal. Chem.* 2002, *74*, 5383–5392.

[18] Nesvizhskii, A. I., Keller, A., Kolker, E., Aebersold, R., *Anal. Chem.* 2003, *75*, 4646–4658.

[19] Li, X.-J., Pedrioli, P., Eng, J., Martin, D. *et al.*, *Anal. Chem.* 2004, *76*, 3856–3860.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim