

# Top-Down Quantitation and Characterization of SILAC-Labeled Proteins

Leonie F. Waanders, Stefan Hanke, and Matthias Mann

Proteomics and Signal Transduction, Max-Planck Institute for Biochemistry, Martinsried, Germany

Stable isotope labeling by amino acids in cell culture (SILAC) has become a popular labeling strategy for peptide quantitation in proteomics experiments. If the SILAC technology could be extended to intact proteins, it would enable direct quantitation of their relative expression levels and of the degree of modification between different samples. Here we show through modeling and experiments that SILAC is suitable for intact protein quantitation and top-down characterization. When SILAC-labeling lysine and/or arginine, peaks of light and heavy SILAC-doublets do not interfere with peaks of different charge states at least between 10 and 200 kDa. Unlike chemical methods, SILAC ensures complete incorporation—all amino acids are labeled. The isotopic enrichment of commercially available SILAC amino acids of nominally 95% to 98% shifts the mass difference between light and heavy state but does not lead to appreciably broadened peaks. We expressed labeled and unlabeled Grb2, a 28 kDa signaling protein, and showed that the two forms can be quantified with an average standard deviation of 6%. We performed on-line top-down sequencing of both forms in a hybrid linear ion trap orbitrap instrument. The quantized mass offset between fragments provided information about the number of labeled residues in the fragments, thereby simplifying protein identification and characterization. (J Am Soc Mass Spectrom 2007, 18, 2058–2064) © 2007 American Society for Mass Spectrometry

The use of mass spectrometry to analyze biological samples has evolved tremendously over the last decade, and mass spectrometry-based proteomics, in particular, has become an indispensable part of modern life sciences [1]. In recent years, the need for quantitative as opposed to qualitative proteomics experiments has become apparent, and a large number of different approaches have been developed, mainly based on incorporation of a “light” and “heavy” stable isotope tag [2]. In chemical approaches, such as in the original ICAT and in the iTRAQ method, the label is reacted with a functional group of an amino acid. In metabolic approaches, the label is instead incorporated by living cells through protein turnover. Our laboratory has previously described a metabolic labeling method called stable isotope labeling by amino acids in cell culture (SILAC, [3]). In SILAC, an essential amino acid has been substituted by its stable isotope counterpart in the medium in which the cells grow, and this “heavy” amino acid is hence incorporated into all expressed proteins. SILAC allows very accurate peptide quantitation in an automated, high throughput experimental setup.

In contrast to this “bottom-up” approach, “top-down” proteomics seeks to characterize intact proteins. Although not widely used in biological research yet, top-down proteomics has unique potential because it

can characterize the complete primary structure of the proteins, including modifications that may be missed in the bottom-up approach. Until now, most top-down studies have focused on mass measurement of the protein, its identification in databases using MS or MS/MS data, or on measuring protein modifications. Hardware and software developments, such as improved resolution and sensitivity, better fragmentation techniques, and increasingly automated software have recently made top-down identification and characterization much faster [4–7]. These developments should make quantitative rather than qualitative top-down proteomics more feasible. So far, however, very few studies have combined the top-down approach with protein quantitation. Gordon et al. demonstrated that relative molecular ion intensities can be used for intact protein quantitation [8]. Kelleher and coworkers investigated the use of <sup>15</sup>N labeling of yeast proteins for intact protein quantitation and determined 50 protein ratios [9]. In the same paper, the authors also chemically labeled yeast proteins with acrylamide and iodoacetamide. They concluded that the use of stable isotopes is preferred, since it prevents chromatographic shifts during LC-MS separation and make quantitation easier and more accurate. Furthermore, a general problem of chemical labeling strategies is the fact that it is not 100% complete in terms of incorporation. Due to steric hindrance in the intact proteins, not all amino acids react evenly well with the isotope labeled reagent, and therefore the degree of labeling is difficult to control.

Address reprint requests to Dr. Matthias Mann, Department of Proteomics and Signal Transduction, Max-Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany. E-mail: [mmann@biochem.mpg.de](mailto:mmann@biochem.mpg.de)

Although quantitative approaches are rare in the top-down proteomics field, stable isotopes have been used for intact protein analysis for a number of years. In 2000, Smith and coworkers used deuterated leucine to improve the identification of *E. coli* proteins as the mass offset in the full scan indicated the number of leucines present in the protein [10]. They also substituted other essential amino acids such as Ile, Phe, Arg, His, and Lys to determine their number in the protein [11]. “Heavy” and “light” labeled proteins were well separated as long as the labeled amino acid was present at least three times per protein.

We have recently investigated the LTQ-Orbitrap, a hybrid linear ion trap orbitrap mass spectrometer, for top-down proteomics [12]. We found that the good sensitivity and excellent mass accuracy and resolution are sufficient for fast and reliable protein identification and characterization. We now extend this work by investigating whether the SILAC technology is also applicable to intact protein quantitation and characterization. We show feasibility by modeling SILAC-labeling for proteins up to 220 kDa, and by quantifying a 28 kDa signaling protein that was expressed in medium with normal and with heavy arginine and lysine residues.

## Experimental

### Modeling

The theoretical applicability of SILAC for protein quantitation was tested by modeling the isotopic and charge state clusters of proteins of different molecular weight. Based on the amino acid frequencies as determined by McCaldon and Argos [13], we calculated the mass of a hypothetical sequence with 100 residues (~11 kDa) and extrapolated this to 55 and 220 kDa. For the modeling we used the Isotopica web application, developed by the groups of Fernandez-de-Cossio and Takao [14]. We first simulated SILAC-labeling with  $^{13}\text{C}_6^{15}\text{N}_4$ -arginine and  $^{13}\text{C}_6^{15}\text{N}_2$ -lysine, inducing a mass increment of 10.008 and 8.014 Da, respectively, assuming that 100% of the substituted carbon and nitrogen atoms were “heavy.”

The influence of incomplete labeling and isotope enrichment were tested by calculating the frequency of different protein forms in an arginine-labeled 55 kDa protein with approximately average amino acid distribution. To test the effects of imperfect isotope enrichment of the SILAC amino acids, we varied the relative abundance of the substituted  $^{13}\text{C}$  and  $^{15}\text{N}$  in the labeled amino acid from 100% to a lower value such as 98% or 95%.

To simulate incomplete amino acid incorporation, we calculated the relative abundance of protein forms when the chance of incorporation of light labels was  $P = 0.02$  or  $P = 0.05$ , assuming a binomial distribution. For every protein form, the isotope distribution of the 34th charge state was predicted with Isotopica (resolution 60,000) and protein forms were weighted according to their probabilistic frequencies.

### *In Vitro* Expression and Purification of Grb2

The gene for growth factor receptor-bound protein 2 (Grb2) was purchased from RZPD German Resource Center for Genome Research, Berlin, Germany (clone RZPD0834A0934D). It was transferred into pDEST17 via Gateway-cloning for T7-RNA polymerase dependent expression as an N-terminally His<sub>6</sub>-tagged protein.

All chemicals and enzymes were purchased from Sigma (Taufkirchen, Germany) or Roche Applied Science (Mannheim, Germany) at the highest purity.

Recombinant N-terminally His<sub>6</sub>-tagged Grb2 was expressed using a cell-free system prepared as described in reference [15] with slight modifications. *E. coli* S30 lysate was prepared from BL21(DE3)RIL cells (Novagen, Darmstadt, Germany). T7 RNA polymerase was also expressed in BL21(DE3)RIL containing the vector pAR1219 described in ref [16], but the enzyme was not purified. Instead, we prepared another lysate from this IPTG-induced culture and added 60  $\mu\text{L}$  for this lysate to 400  $\mu\text{L}$  of the standard lysate to 1 mL of reaction volume. The concentration of the reaction components were adjusted to 57 mM HEPES-KOH buffer (pH 8.2), 2 mM DTT, 1.2 mM ATP, 0.85 mM each of CTP, GTP and UTP, 100 mM creatine phosphate, 130  $\mu\text{g}/\text{mL}$  creatine kinase, 2.0% PEG 8000, 0.64 mM 3',5'-cyclic AMP, 34  $\mu\text{M}$  L(-)-5-formyl-5,6,7,8-tetrahydrofolic acid, 175  $\mu\text{g}/\text{mL}$  *E. coli* total tRNA, 90 mM potassium glutamate, 80 mM ammonium acetate, 12 mM magnesium acetate, 2.0 mM each of the 20 amino acids, and 6.7  $\mu\text{g}/\text{mL}$  of plasmid DNA. The reaction mixture was incubated for 2 h while shaking at 600 rpm at 30 °C.

For the heavy SILAC labeled Grb2, the normal arginine and lysine were replaced with their heavy-isotope counterparts  $^{13}\text{C}_6^{15}\text{N}_4$ -Arginine and  $^{13}\text{C}_6^{15}\text{N}_2$ -Lysine, also at a concentration of 2 mM. The reaction mixture contained the target protein mostly as precipitate, which was solubilized with 6 M guanidinium chloride and subsequently purified employing  $\text{Ni}^{2+}$ -affinity chromatography according to the manufacturer's protocol for purification of denatured protein using Ni-NTA sepharose (Qiagen, Hilden, Germany) and spin columns (Mo-BiTec, Goettingen, Germany). The purified protein was dialyzed against distilled water to remove imidazole and guanidinium, which again led to precipitation of the protein.

### Sample Preparation and Mass Spectrometry

Heavy and light labeled forms of Grb2 were aliquoted separately and stored at  $-80^\circ\text{C}$  before use. Before mass spectrometric analysis, the proteins were washed on RP-C<sub>18</sub> StageTips [17], eluted, and mixed in the desired ratio.

All experiments were done on a LTQ-Orbitrap (Thermo Fisher, Bremen, Germany), coupled to a nanoLC system (Agilent, Waldbronn, Germany). Online protein separation was performed by use of 75  $\mu\text{m} \times 150$  mm IntegraFrit columns (New Objective, Berlin, Germany) packed in-house with 5  $\mu\text{m}$  RP-C18 beads (Reprosil-Pur Aq, 200 Å pore size, Dr. Maisch, Ammerbuch-Entringen,

Germany). The column was connected to a short nano-spray needle and spraying voltage was kept low (<2 kV) to prevent oxidation [18]. During loading and washing, the flow was set to 500 nL/min; whereas during the actual gradient the flow rate was 250 nL/min. Buffers compositions were 0.5% acetic acid in mQ distilled water (Buffer A) and 0.5% acetic acid in acetonitrile (Buffer B). The proteins eluted in a 30 min gradient from 40% to 90% of Buffer B. CID fragmentation was performed in the LTQ, but MS/MS ions were detected in the orbitrap mass analyzer.

### Data Analysis

The MS scans were deconvoluted with Xtract software (Thermo Fisher) and matched with the expected sequence. The light form of Grb2 was used to determine the sequence of the linker between the His-tag and the protein.

CID fragments observed in light or in both light and heavy MS/MS spectra were searched in ProSight PTM (developed by The Kelleher group, University of Illinois, Urbana-Champaign, IL), but without success. We therefore suspected they were internal fragments and searched them in Mascot (Matrix Science Ltd., London, UK) after increasing light fragment masses by one water molecule (18.0152 Da). A peptide mass fingerprint search was performed against the human IPI database (version 3.19) and a small database with the predicted His-tagged Grb2 sequence with 10 ppm maximum mass deviation and without enzyme specificity.

## Results and Discussion

### Simulation of SILAC Quantitation for Intact Proteins

We first wanted to investigate whether SILAC would be applicable for proteins of all sizes and whether any overlap between charge states and labeled states could occur. To this end, we used Isotopica, a software package freely available on the Internet, and developed by Fernandez-de-Cossio and Takao [14], to model the effect of SILAC labeling. Isotopica can predict the isotopic cluster of specific charge states and can also calculate the mono-isotopic and average mass of all charge states (up to 50) for both heavy and light labeled proteins. As shown in Figure 1a–c, SILAC labeling with heavy arginine and lysines results in mass offsets that are clearly distinguishable in a spectrum. The thickness of the lines is approximately equal to the full width of the isotopic cluster. Even for a 220 kDa protein (Figure 1c) the two lines are separate, indicating that at least theoretically the isotopic clusters of the light and heavy protein forms do not overlap with consecutive charge states.

Because the mass offset induced by SILAC is in the range of a few hundred Da, when both arginine and lysine are labeled, post-translational modifications with

a mass less than this would not overlap with either the heavy or the light form of the protein. For example, several phosphorylation sites on a 50 kDa protein would not overlap with the heavy version, which is ~500 Da higher in mass.

The simulation also allowed us to visualize the effect of incomplete isotope enrichment. In principle, SILAC amino acids should be labeled with  $^{13}\text{C}$  or  $^{15}\text{N}$  to 100% at the substituted sites but in practice commercial sources guarantee isotope enrichment between 95% and 98%. As described in the Experimental section, we modeled this by varying isotope abundances in the heavy form of an average 55 kDa arginine-labeled protein. Interestingly, incomplete isotope enrichment does not broaden the peaks but it does shift the mass of the heavy form to a lower value (Figure 1d). It follows from this that accurate mass measurement to determine the molecular weight of the intact protein should be performed on the light and not on the heavy form of the SILAC protein pair.

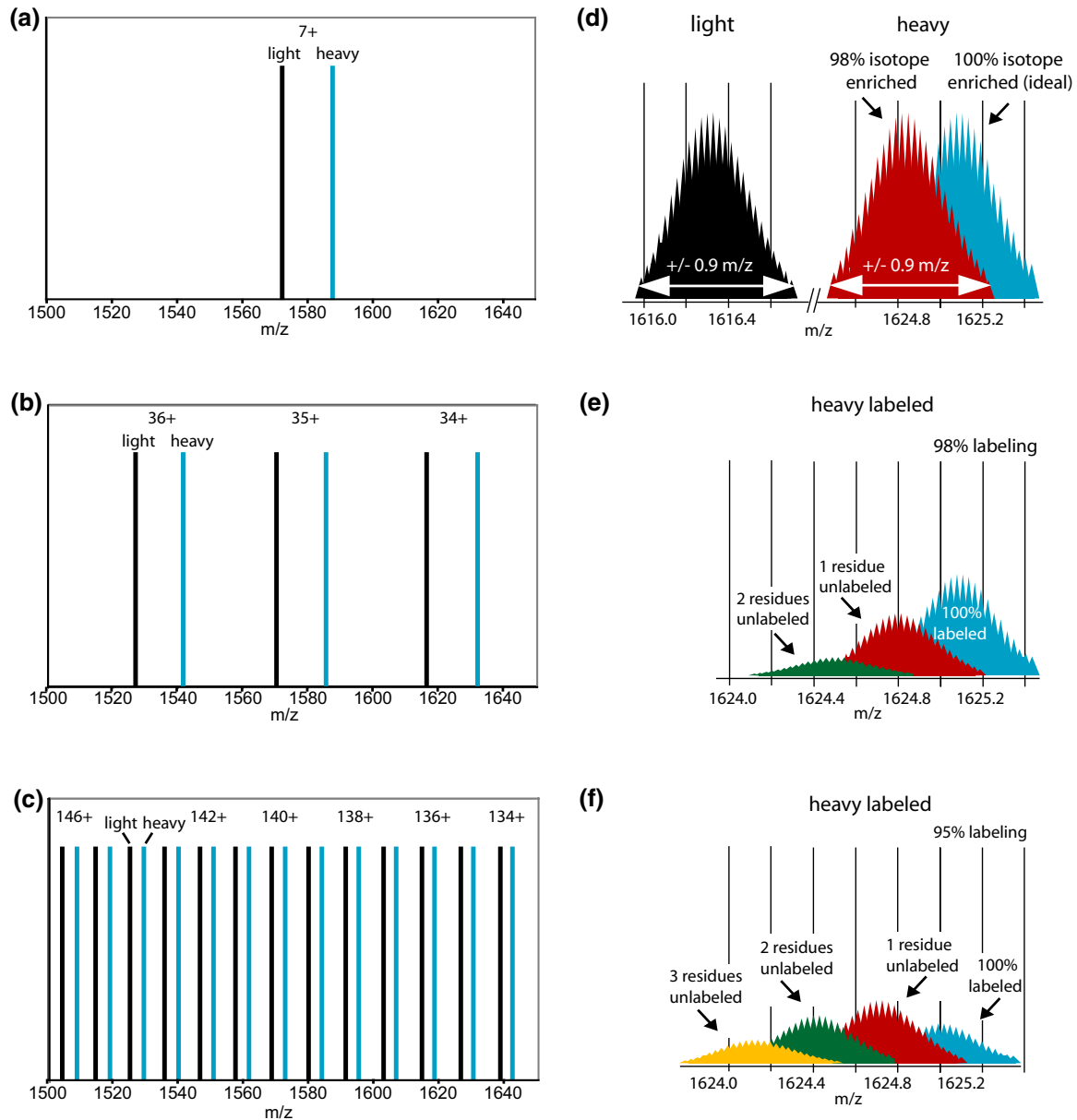
Next, we modeled the effect of incomplete labeling, that is, we modeled the case that not every light amino acid was replaced by its heavy counterpart. This could happen during SILAC cell adaptation, while light amino acids are still present. As Figure 1e shows, even incomplete labeling at the two percent level (no more than one amino acid in our example), causes the signal to split into at least three states and this reduces the overall intensity and signal to noise (S/N) dramatically. At the 5% level, severe splitting into at least four states occurs, and the heavy peak is broadened at least 3-fold. Therefore, complete labeling is a precondition for successful quantitative top-down analysis. Fortunately, it is not difficult in SILAC experiments to achieve complete labeling; the only requirement is to grow the cells for a sufficient number of cell doublings.

Note, however, that this broadening is very likely to occur in chemical labeling strategies because it is very difficult to achieve close to 100% labeling on the desired amino acids while preventing any labeling of untargeted sites.

### Expression and Mass Spectrometry of SILAC-Labeled Grb2

To experimentally test the feasibility and accuracy of quantitation by SILAC for an intact protein, we expressed HIS-tagged Grb2, a signaling protein with a calculated monoisotopic mass of 27,789.758 Da in vitro in normal media and in media with heavy arginine and lysine. Light and heavy protein forms were measured separately and mixed in ~1:1 or 2:1 ratio. The mixture was analyzed by online HPLC MS on the LTQ-Orbitrap (see Figure 2).

The resolving power of the orbitrap was sufficient to observe the isotopic clusters. By measuring the heavy labeled protein only, we obtained a single population and conclude that we achieved 100% label incorporation (data not shown).

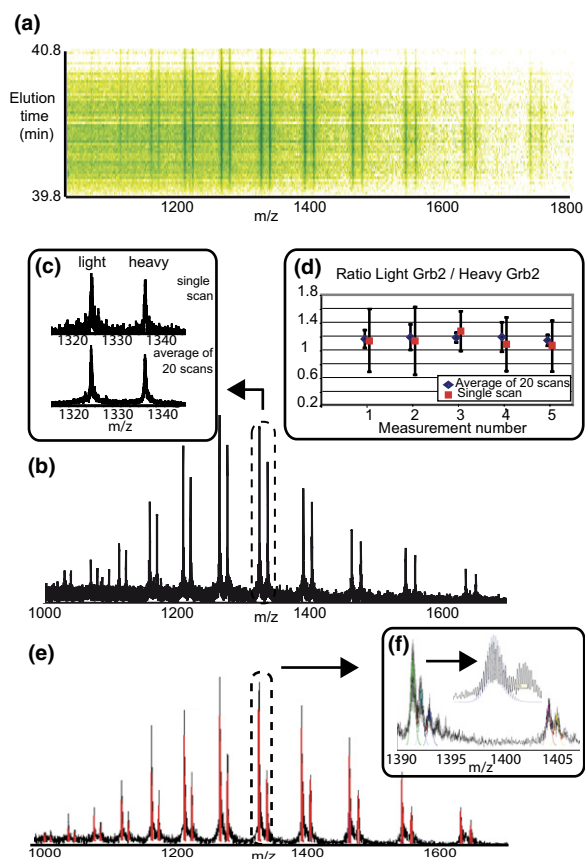


**Figure 1.** Simulation of heavy and light SILAC labeled proteins of various molecular weights with an average amino acid distribution. (a)–(c) SILAC-labeling with  $^{13}\text{C}_6^{15}\text{N}_4$ -Arg and  $^{13}\text{C}_6^{15}\text{N}_2$ -Lys increases the protein mass by approximately 1%, resulting in well separable isotope clusters, independent of protein size. The isotopic clusters of light and heavy labeled proteins of 11, 55, and 200 kDa in the  $m/z$  range 1500 to 1650 are depicted in panels (a)–(c), respectively. The thickness of the lines indicates the full width of the isotope cluster, which for all size proteins remains below 1 Thomson. (d)–(f) The effects of incomplete isotope enrichment and incomplete mass labeling are modeled, based on the 34+ peak of a theoretical protein of 55 kDa with average amino acid composition, labeled with heavy arginine and measured with 60,000 resolving power. (d) A 98% isotope enrichment of  $^{13}\text{C}$  and  $^{15}\text{N}$  in the heavy amino acids results in a shift of the total heavy isotopic cluster from 100% enrichment. The width and the intensity of the isotopic cluster remain unchanged. (e) In contrast, if only 98% of the arginines and lysines are labeled with heavy amino acids, the result is a significant spread of the signal. (f) With 95% labeling the signal is even more spread and the total intensity is reduced to 33% of the original signal.

### *Influence of Isotope Enrichment in Heavy Label*

Above, we modeled the influence of incomplete isotope enrichment of the SILAC amino acid. The SILAC amino acids used in our experiment were specified at 98% for  $^{13}\text{C}$ . For  $^{15}\text{N}$  they were specified as 95% for Lys and 98% for Arg. Thus about 2% of the labeled carbon and

nitrogen atoms in their stable isotope labeled amino acids should be light ( $^{12}\text{C}$  and  $^{14}\text{N}$ ). However, when we used this value to predict the mass offset between the heavy and light labeled Grb2 protein forms, we could not explain an additional mass difference of 3 Da for the heavy labeled Grb2 form in comparison with the calcu-



**Figure 2.** Experimental spectra of mixed Grb2 forms, showing the experimental feasibility of SILAC. (a) Contour plot of the eluting SILAC labeled protein pair. The graph is color coded with light green corresponding to low and dark green to high intensity. (b) Single spectrum of mixed light and heavy Grb2 forms, showing multiple charge states of both forms with isotopic clusters that do not overlap. The ratio between heavy and light Grb2 is constant over the different charge states and allows accurate quantitation based on relative intensities. (c) Averaging of spectra leads to significant noise reduction and improves quantitation accuracy. (d) Intensity ratios between light and heavy labeled Grb2, based on the eight most intense charge states of single or averaged spectra (20 scans) for 5 different LC-MS-runs were determined with their 95% confidence interval error bars. (e) An experimentally obtained spectrum was overlaid with the Isotopica prediction (red), calculated using 99% heavy isotope enrichment. (f) The inserts show that the simulation is accurate down to the isotopic resolution, indicating higher isotope enrichment of Arg and Lys than specified by manufacturers.

lated one of 264.3 Da. Since the sequence of the protein is known and confirmed by the light labeled protein, this mass difference had to be caused by the heavy amino acids.

When overlaying the experimentally obtained spectra with the prediction made by Isotopica (Figure 2b), we discovered that the matching became very accurate when isotope enrichment was 99%. Thus, in this case, the isotope enrichment of the SILAC amino acids was almost complete and significantly higher than specified. Furthermore, our measurements confirmed the simulation in Figure 1d. The rare presence of  $^{12}\text{C}$  and  $^{14}\text{N}$  instead of  $^{13}\text{C}$  and  $^{15}\text{N}$  in the heavy amino acids did not

cause problems for the analysis or quantitation, since the signal intensity and the width of the isotopic distribution remained the same. Note that the percentage of light atoms in the heavy label needs to be determined only once as it will be the same for all proteins that are labeled with that specific batch of heavy amino acids.

### Quantitation at the Protein Level Using SILAC

A contour plot of the SILAC labeled Grb2 forms showed that they eluted as a 1 min peak with no discernable retention time shift between them (Figure 2a). Complete isotopic separation and coelution should allow for very accurate quantitation. The spreading of the protein signal into multiple charge states reduces the S/N but improves quantitation, since all charge states should show the same ratio between heavy and light labeled forms. To determine the quantitation precision, the ratio between Grb2-light and Grb2-heavy was calculated for a single scan and for an averaged scan, composed of 20 separate scans over the elution profile. In both cases the eight most intense charge states were used for the calculation. We then repeated this measurement in five separate HPLC runs to demonstrate the reproducibility of the measurement. In the insert of Figure 2, the average ratio is displayed with 95% confidence interval error bars. When averaging 20 scans, the mean standard deviation was 6%, whereas ratios determined from single scans had a standard deviation of typically 18%.

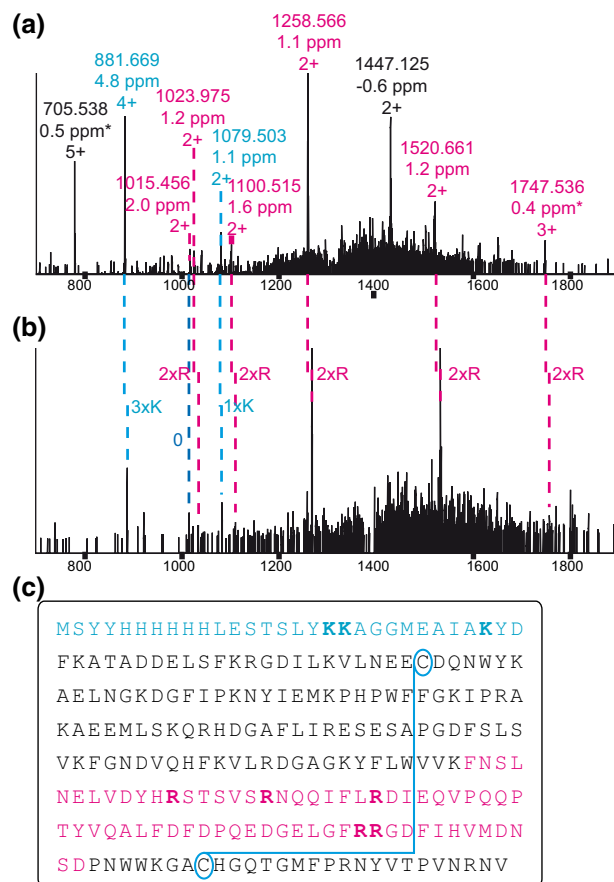
The ratio calculation can be further improved by using the complete elution profile of the proteins (see Figure 2a). Furthermore, as Ong et al. [19] and Du et al. [9] have noted, results become more accurate if they are corrected for noise, especially in case of low S/N peaks.

### Improved Assignment of Top-Down Fragments by SILAC

Besides the accurate quantitation, SILAC can also be used for identification and characterization purposes. The information in fragmentation spectra is valuable in assigning modifications and in improving the reliability of protein identification. Modifications and truncations are more easily detected and more likely correctly assigned. Due to the speed and sensitivity of the orbitrap, such MS/MS spectra are readily obtainable on small proteins [12].

To demonstrate this in a SILAC experiment, we performed data-dependent CID fragmentation in the online format, collecting MS and MS/MS scans in the orbitrap, without microscanning to keep the duty cycle short. The MS/MS scans were of relatively low intensity (Figure 3), but the resolution and mass accuracy of the fragments was sufficiently high that even low abundant ions could be clearly distinguished from the noise.

When comparing the MS/MS scans of heavy and light labeled Grb2, we observed very similar fragmen-



**Figure 3.** Identification of Grb2 is supported by direct comparison of the heavy and light CID fragmentation spectra. Mass differences between corresponding fragments observed in MS/MS spectra of the (a) light and (b) heavy Grb2 protein indicate the number of lysines and arginines per peptide and thereby simplify and confirm the peptide and protein identification. Assigned fragments have an average absolute mass accuracy of  $3.1 \pm 2.3$  ppm. The mass deviation of the ion indicated with an asterisk has been derived from the  $^{13}\text{C}$  isotope. N-terminal fragments are shown in blue and internal fragments in red. (c) In total the fragments covered 40% of the total protein sequence.

**Table 1.** Fragments of Grb2 identified by Mascot

Light <i>m/z</i>	Heavy <i>m/z</i>	Charge	$\Delta\text{mass}$	No. of heavy residues	Mass deviation (ppm)	Sequence
954.775		3			5.49	MSYYHHHHHHLESTSLYKKAGGME
978.788		3			3.90	MSYYHHHHHHLESTSLYKKAGGMEA
1039.831	1045.504	3	16.03	2 Lys	(1.88) <sup>a</sup>	MSYYHHHHHHLESTSLYKKAGGMEIA
881.669	887.681	4	24.04	3 Lys	4.79	MSYYHHHHHHLESTSLYKKAGGMEIAIKYD
705.538		5			(0.50) <sup>a</sup>	MSYYHHHHHHLESTSLYKKAGGMEIAIKYD
1175.223		3			4.81	MSYYHHHHHHLESTSLYKKAGGMEIAIKYD
1015.456	1015.455	2	0.01	0	2.00	SYHHHHHHLESTSLY
1079.503	1083.511	2	8.02	1 Lys	1.10	SYHHHHHHLESTSLYK
1577.204		2			1.67	LDFDPQEDGELGFRRGDFIHVMDNSD
1520.661	1530.668	2	20.02	2 Arg	1.20	DFDPQEDGELGFRRGDFIHVMDNSD
1447.125		2			-0.57	DFDPQEDGELGFRRGDFIHVMDNSD
1100.515	1110.520	2	20.04	2 Arg	1.59	PQEDGELGFRRGDFIHVMD
1258.566	1268.573	2	20.03	2 Arg	1.07	QEDGELGFRRGDFIHVMDNSDP
1023.975	1033.983	2	20.03	2 Arg	1.18	GELGFRRGDFIHVMDNS
1747.536	1757.540	3	30.01	3 Arg	(0.42) <sup>a</sup>	FNSLNELVDYHRSTSVSRNQIIFLRDIE-QVPOOPTVYQALDFD

<sup>a</sup> $^{13}\text{C}$ -peaks were used for mass accuracy determination.

tation that appeared very useful for assigning the fragments. The mass offset between the heavy and light fragments should exactly represent the number of lysines and arginines of every fragment. Table 1 lists the fragments and their mass deviations from the calculated values. Most of the fragments were internal and likely observed because the disulphide bridge was still present in the protein.

The absolute mass deviation of the fragments was very low, on average 3.2 ppm, and the number of lysines and arginines determined by the mass increment of the heavy labeled fragments matched exactly with the proposed identification. One potential false positive hit could be eliminated both because of its mass deviation of 14 ppm and also because of the fact that its counterpart in the spectrum of the heavy form of Grb2 did not indicate the correct number of lysines. In total the single scan CID fragmentation depicted in the figure resulted in fragments covering 40% of the protein sequence. Clearly the SILAC information at the fragment level would have been very valuable to assign any protein modification, had it been present.

## Conclusions and Perspectives

Here we have investigated the applicability of the SILAC technology to top-down proteomics. Through theoretical modeling, we found that heavy and light SILAC-doublets do not interfere with each other even for very large proteins. The incomplete isotope enrichment of commercial SILAC amino acids does not cause peak broadening but does shift the mass to lower values. This effect can be modeled very precisely. Incomplete labeling, on the other hand, would lead to distribution of the signal into several peaks, substantially broadening them and decreasing the signal to noise. Fortunately, incomplete labeling can easily be avoided in SILAC-metabolic labeling experiments.

We have shown that a 28 kDa signaling protein, Grb2, can be readily quantified in the labeled versus the unlabeled form. The quantitation in a one to one mixture had a typical standard deviation of six percent and is mainly limited by the lower signal to noise in protein measurements compared with peptide measurements. Grb2 was fragmented and analyzed in an LTQ-Orbitrap instrument using single scan data from online experiments. The high mass accuracy combined with the quantized mass offsets significantly improves fragment identification when comparing tandem mass spectra of light and heavy SILAC labeled protein. The unambiguous information about the number of labeled residues per fragment or protein is a clear advantage of the SILAC technology over  $^{15}\text{N}$  labeling.

In our experiment we chose to use two amino acids, arginine and lysine that were isotopically labeled, leading to a 1% mass offset between light and heavy form. Often, it may be more convenient to label with a single amino acid, for example, lysine. This would allow direct “counting” of the number of lysines in the protein and between fragments in the “heavy” and “light” tandem spectrum. Double amino acids may be valuable when analyzing proteins with potentially a high number of modifications.

Extension of the SILAC technology to intact protein analysis should allow direct quantitation of endogenous and unprocessed proteins. However, the most interesting application could be in the direct quantitation of multiple and combinatorial modifications of regulatory proteins as a function of cellular state. This goal will require SILAC quantitation of fragments isolating these modifications. It may also require nonreductive fragmentation techniques such as ECD or ETD as well as several stages of fragmentation.

## Acknowledgments

The authors thank Dr. Takao and Dr. Fernandez-de-Cossio for their assistance with the Isotopica software. They thank Dr. Jesper Olsen and Dr. Johannes Graumann for advice and discussion. This work was partly funded by Interaction Proteome, a 6th Framework grant by the European Union Research Directorate and by the Max-Planck Society for the Advancement of Science.

## References

- Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. *Nature* **2003**, *422*, 198–207.
- Ong, S. E.; Mann, M. Mass Spectrometry-Based Proteomics Turns Quantitative. *Nat. Chem. Biol.* **2005**, *1*, 252–262.
- Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; 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. Proteom.* **2002**, *1*, 376–386.
- Kelleher, N. L. Top-Down Proteomics. *Anal. Chem.* **2004**, *76*, 197A–203A.
- Bogdanov, B.; Smith, R. D. Proteomics by FTICR Mass Spectrometry: Top-Down and Bottom-Up. *Mass Spectrom. Rev.* **2005**, *24*, 168–200.
- Whitelegge, J.; Halgand, F.; Souda, P.; Zabrouskov, V. Top-Down Mass Spectrometry of Integral Membrane Proteins. *Exp. Rev. Proteom.* **2006**, *3*, 585–596.
- Zamdborg, L.; Leduc, R. D.; Glowacz, K. J.; Kim, Y. B.; Viswanathan, V.; Spaulding, I. T.; Early, B. P.; Bluhm, E. J.; Babai, S.; Kelleher, N. L. ProSight PTM 2.0: Improved Protein Identification and Characterization for Top-Down Mass Spectrometry. *Nucleic Acids Res.* **2007**.
- Gordon, E. F.; Mansoori, B. A.; Carroll, C. F.; Muddiman, D. C. Hydrophobic Influences on the Quantification of Equine Heart Cytochrome *c* Using Relative Ion Abundance Measurements by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Mass Spectrom.* **1999**, *34*, 1055–1062.
- Du, Y.; Parks, B. A.; Sohn, S.; Kwast, K. E.; Kelleher, N. L. Top-Down Approaches for Measuring Expression Ratios of Intact Yeast Proteins Using Fourier Transform Mass Spectrometry. *Anal. Chem.* **2006**, *78*, 686–694.
- Veenstra, T. D.; Martinovic, S.; Anderson, G. A.; Pasa-Tolic, L.; Smith, R. D. Proteome Analysis Using Selective Incorporation of Isotopically Labeled Amino Acids. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 78–82.
- Martinovic, S.; Veenstra, T. D.; Anderson, G. A.; Pasa-Tolic, L.; Smith, R. D. Selective Incorporation of Isotopically Labeled Amino Acids for Identification of Intact Proteins on a Proteome-Wide Level. *J. Mass Spectrom.* **2002**, *37*, 99–107.
- Macek, B.; Waanders, L. F.; Olsen, J. V.; Mann, M. Top-Down Protein Sequencing and MS3 on a Hybrid Linear Quadrupole Ion Trap-Orbitrap Mass Spectrometer. *Mol. Cell. Proteom.* **2006**, *5*, 949–958.
- McCaldon, P.; Argos, P. Oligopeptide Biases in Protein Sequences and Their Use in Predicting Protein Coding Regions in Nucleotide Sequences. *Proteins* **1988**, *4*, 99–122.
- Fernandez-de-Cossio, J.; Gonzalez, L. J.; Satomi, Y.; Betancourt, L.; Ramos, Y.; Huerta, V.; Besada, V.; Padron, G.; Minamino, N.; Takao, T. Automated Interpretation of Mass Spectra of Complex Mixtures by Matching of Isotope Peak Distributions. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2465–2472.
- Kigawa, T.; Yabuki, T.; Matsuda, N.; Matsuda, T.; Nakajima, R.; Tanaka, A.; Yokoyama, S. Preparation of *Escherichia coli* Cell Extract for Highly Productive Cell-Free Protein Expression. *J. Struct. Funct. Genomics* **2004**, *5*, 63–68.
- Davanloo, P.; Rosenberg, A. H.; Dunn, J. J.; Studier, F. W. Cloning and Expression of the Gene for Bacteriophage T7 RNA Polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 2035–2039.
- Rappsilber, J.; Ishihama, Y.; Mann, M. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Anal. Chem.* **2003**, *75*, 663–670.
- Morand, K.; Talbo, G.; Mann, M. Oxidation of Peptides During Electrospray Ionization. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 738–743.
- Ong, S. E.; Kratchmarova, I.; Mann, M. Properties of  $^{13}\text{C}$ -Substituted Arginine in Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *J. Proteome Res.* **2003**, *2*, 173–181.