The Association of Biomolecular Resource Facilities Proteomics Research Group 2006 Study

RELATIVE PROTEIN QUANTITATION*

Christoph W. Turck,[a,b] **Arnold M. Falick,**[c] **Jeffrey A. Kowalak,**[d] **William S. Lane,**[e] **Kathryn S. Lilley,**[f] **Brett S. Phinney,**[g] **Susan T. Weintraub,**[h] **H. Ewa Witkowska,**[i] **and Nathan A. Yates**[j]

The determination of differences in relative protein abundance is a critical aspect of proteomics research that is increasingly used to answer diverse biological questions. The Association of Biomolecular Resource Facilities Proteomics Research Group 2006 study was a quantitative proteomics project in which the aim was to determine the identity and the relative amounts of eight proteins in two mixtures. There are numerous methodologies available to study the relative abundance of proteins between samples, but to date, there are few examples of studies that have compared these different approaches. For the 2006 Proteomics Research Group study, there were 52 participants who used a wide variety of gel electrophoresis-, HPLC-, and mass spectrometry-based methods for relative quantitation. The quantitative data arising from this study were evaluated along with several other experimental details relevant to the methodologies used. *Molecular & Cellular Proteomics 6:1291–1298, 2007.*

An important aspect of current biological research is the design of experimental systems that can yield quantitative results. Therefore, it is not surprising that there is keen interest in the quantitative analysis of proteins because proteins are of such fundamental importance in all cellular processes. The results obtained from large scale quantitative analyses of protein expression will undoubtedly play an important role in advancing our understanding of metabolic events, cellular systems, and pathogenic mechanisms of disease.

The drive toward quantitative studies has been greatly facilitated by the development of new technologies and chemistries. A large number of methods have been developed for the determination of differences in protein expression, including methods based on gel electrophoresis, MS, immunochemistry, and others. Some of these methods have been in use for many years, whereas others have been implemented or modified more recently (1). There are two general quantitative approaches: absolute and relative. The present study deals only with the latter, mainly due to the fact that the majority of current proteomics studies are concerned with relative measurements among two or more samples and not absolute quantities. This approach is frequently referred to as comparative proteomics. Although it was not called comparative proteomics at the time, relative quantitative measurements of proteins on a large scale have been conducted ever since the development of two-dimensional (2D)[1] gel electrophoresis (2, 3). Following the establishment of sensitive and high throughput mass spectrometric methods over 10 years ago (4, 5), strategies for the relative quantitation of proteins based on stable isotope labeling techniques were implemented in the field of quantitative proteomics (6). These methods have been complemented recently by label-free mass spectrometry approaches that are based on spectral ion currents (7, 8).

Currently there is a considerable choice of experimental approaches to study the proteome in a quantitative manner, but there have been few studies in which these techniques have been compared and contrasted for a given sample set (9). One reason for this is that many of these techniques are relatively expensive to implement, and a given laboratory may

From the [a]Max Planck Institute of Psychiatry, D-80804 Munich, Germany, [c]Howard Hughes Medical Institute Mass Spectrometry Laboratory, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, [d]Laboratory of Neurotoxicology, National Institute of Mental Health, Bethesda, Maryland 20892, [e]Microchemistry and Proteomics Analysis Facility, Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, [f]Department of Biochemistry and Cambridge Centre for Proteomics, University of Cambridge, Cambridge CB2 1QW, United Kingdom, [g]Genome Center, University of California, Davis, California 95616, [h]Department of Biochemistry, University of Texas Health Science Center, San Antonio, Texas 78229, [i]UCSF Biomolecular Resource Center Mass Spectrometry Facility and Department of Cell and Tissue Biology, University of California, San Francisco, California 94143, and [j]Molecular Profiling Proteomics, Merck Research Laboratories, Merck & Company, Rahway, New Jersey 08854

---

[1] The abbreviations used are: 2D, two-dimensional; ABRF, Association of Biomolecular Resource Facilities; PRG, Proteomics Research Group; 1D, one-dimensional.

only have access to a limited number of approaches. Carrying out a comprehensive comparison of these techniques by one laboratory is, therefore, quite challenging.

Since its creation in 2001, the Association of Biomolecular Resource Facilities (ABRF) Proteomics Research Group (PRG) has carried out research studies that have addressed various topics relevant to the field of proteomics. The study topics have included the identification of components in a protein mixture (10), determination of the sites of phosphorylation in a protein (11), differentiation of protein isoforms (12), and *de novo* peptide sequencing.[2] In 2006, the PRG organized a quantitative proteomics study that was designed to evaluate methodologies that are available for relative quantitation of proteins among experimental samples. This study was considered to be particularly timely in that an increasing number of proteomics facilities were being asked to provide large scale quantitative proteomics measurements for their user groups. It was the expectation of the PRG that proteomics facilities would use a variety of strategies to carry out the experiments needed for the study, thereby providing a more comprehensive dataset generated from more diverse experimental approaches than would be possible by a single group. In addition, it was anticipated that the study samples would aid laboratories in gaining experience with methods for carrying out such analyses.

MATERIALS AND METHODS

All proteins were purchased from Sigma-Aldrich. Deionized water (18.2 megaohms; Milli-Q Gradient A10 from Millipore, Bedford, MA) was used to prepare all solutions. For proteins supplied in small quantities (carbonic anhydrase I, glycogen phosphorylase, horseradish peroxidase, and lactoperoxidase), the requisite volume of water was added to each vial to generate the individual stock solutions. Samples of the remaining proteins (β-casein, BSA, catalase, and ribonuclease A) were weighed on an analytical balance and separately dissolved in water. The exact protein content of each individual stock solution was assessed by amino acid analysis. Three stock mixtures were then prepared: 1) a mixture containing the four proteins that were present at the same concentration in both samples, 2) the remaining proteins for sample A, and 3) the remaining proteins for sample B. Each sample, A and B, received an equal volume of stock mixture 1 and a volume of stock mixture 2 or 3 as appropriate to achieve the same total protein amount in both samples. Each vial was dried in a vacuum centrifuge and stored at room temperature for no longer than 1 week prior to mailing.

RESULTS AND DISCUSSION

The study sample consisted of two protein mixtures, each of which contained the same eight proteins. Four of the proteins were present at a 1:1 ratio in both samples, whereas the others were present at varying ratios over a range of approximately 2 orders of magnitude. The participating laboratories were asked to identify the proteins and determine their relative

---

quantities within the two samples. In addition to comparing the ability of different techniques to identify and provide relative quantitation of proteins, other goals of the 2006 PRG study included assessment of the participants' level of confidence and consistency in the quantitative data and evaluation of the ability of software to determine quantitative differences between samples. The PRG also hoped to provide participating laboratories with a means to evaluate their proficiency in identifying and quantifying proteins in a mixture. From a comparison of the results obtained by different strategies, participants would be able to gauge their own capabilities and establish realistic expectations for the approaches that were used.

Because this was the first study of its kind, the PRG wanted it to be suitable for as many different quantitative methods as possible. In preparation for the study, the PRG evaluated 12 proteins for possible inclusion in the study. A one-dimensional (1D) SDS-PAGE separation of the 12 proteins is shown in Fig. 1. All proteins were obtained from a commercial source and showed various degrees of heterogeneity as evidenced by the minor components detected in each lane in addition to the band for the expected protein. No efforts were made to characterize or remove any of the protein impurities and/or isoforms. Eight proteins that showed the least heterogeneity and that had a range of molecular weights and isoelectric points were then selected for the study: β-casein, BSA, carbonic anhydrase I, catalase, glycogen phosphorylase, lactoperoxidase, horseradish peroxidase, and ribonuclease A. Sample mixtures were prepared using the ratios listed in Table I as described under "Materials and Methods." Although four of the eight protein pairs were present in identical amounts in the two mixtures, participants were informed about the identity of only one of them (BSA) so that they could use this protein for normalization. The relative protein ratios between the two samples did not exceed 1:100 because it was thought that the technologies currently available could not accurately measure higher ratios. The most challenging case was that of glycogen phosphorylase with a ratio of 1:76 between the two samples. In sample A this protein was only present at 3 pmol. The total amount of protein in each sample was ~80 μg. Sufficient amounts of the individual proteins were provided so that participants using less sensitive methods such as gel electrophoresis in combination with Coomassie staining would be able to carry out the analysis more than once. The study was designed such that the total amount of protein in each sample was the same. Fig. 2 is an image from the analysis of the combined samples by DIGE (13). This analysis also demonstrated protein heterogeneity due to the presence of protein isoforms (in some cases) as well as unknown impurities.

The PRG provided the following information to the participants.

1. The sample set consists of two mixtures, A and B.
2. Both mixtures contain eight major proteins, each present in various amounts.

FIG. 1. **One-dimensional SDS-poly-acrylamide gel electrophoresis of 12 commercial protein preparations.** Individual protein preparations and a mixture of all 12 proteins were separated in a 12% gel and then stained with SYPRO Ruby. The proteins indicated in *red* were selected for the PRG2006 study.
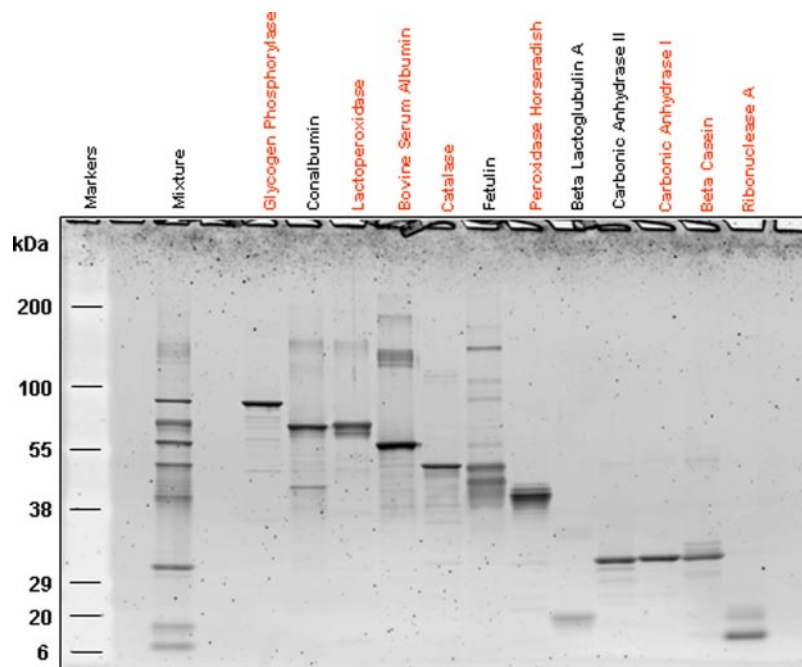
TABLE I

Proteins in PRG2006 study samples

Protein[a]	Molecular mass	Quantity		Ratio (A/B)
		A	B	
	kDa	*pmol*		
β-Casein	24.0	598	150	4:1
Bovine serum albumin	66.6	195	195	1:1
Carbonic anhydrase I	28.9	28	90	1:3
Catalase	57.5	299	59	5:1
Glycogen phosphorylase	97.4	3	235	1:76
Lactoperoxidase	77.5	300	300	1:1
Horseradish peroxidase	43.3	298	298	1:1
Ribonuclease A	13.7	296	296	1:1

[a] Proteins were purchased from Sigma-Aldrich: β-casein, catalog number C-6905; bovine serum albumin, A-0281; carbonic anhydrase I, C-6653; catalase, C-40; glycogen phosphorylase, P-1261; lactoperoxidase, L-8257; horseradish peroxidase, P-6782; ribonuclease, R-4875.

3. In some cases, isoforms and/or contaminant(s) are present (as typically encountered for real life samples).

4. The total amount of protein in each tube is ~80 μg. The individual proteins are present on average at 300 pmol, ranging from ~3 to 600 pmol.

5. The ratios of proteins between mixtures A and B vary up to 1:100.

6. Four of the eight proteins are present at a 1:1 ratio with BSA being one of them.

7. The proteins were mixed from aqueous solutions that also contained small amounts of salt and then dried.

The protein identifications and quantitative results generated by the participants were collected using a web-based survey system (SurveyMonkey) that has proven in previous studies to be a quick and easy way for web-based data input.
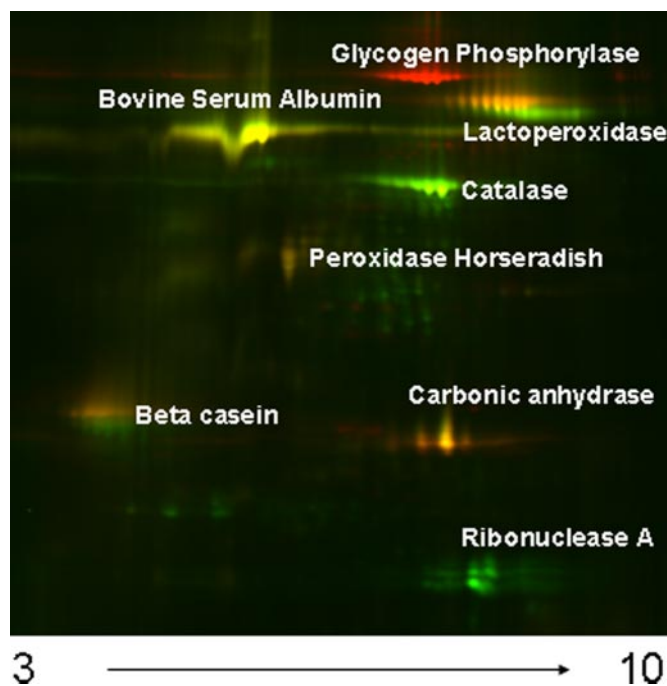

FIG. 2. **Differential in-gel electrophoresis of the two mixtures prepared for the PRG2006 study sample.** Sample A was labeled with Cy3 dye, and sample B was labeled with Cy5 dye. Proteins from both samples were then combined and separated in a 3–10 pH non-linear gradient IPG strip in the first dimension and a 12% SDS gel in the second dimension.

Each participant was also asked to provide information about the methods they used for sample preparation and data analysis as well as any other details relevant to the study. The use of this web-based questionnaire greatly facilitated interpretation and consolidation of the submitted data by the PRG. As

TABLE II
Methods used by participants in PRG2006 study

Method[a]	Number
Gel-based (37.5%)	
1D Coomassie	4
2D Coomassie	4
2D silver-stain	2
1D fluorescence	4
2D fluorescence	1
2D DIGE	3
2D radioactivity	1
MS/isotope (41.7%)	
ICPL	4
iTRAQ	12
$^{16}O/^{18}O$	5
O-Methylisourea	1
MS/non-isotope (20.8%)	
Ion current	10
Spectral count	1

[a] 2D radioactivity, $^{125}I/^{131}I$; iTRAQ, isobaric tags for relative and absolute quantitation; ICPL, isotope-coded protein label.

in the past, all steps of data collection, evaluation, and presentation were conducted in an anonymous manner.

Of the 92 laboratories requesting study samples, 52 returned analysis results. This represents a very good response rate compared with previous studies and reflects the great interest in quantitative proteomics in the community. Data were submitted by both ABRF members and non-members from academia, government, and industry (including instrument manufacturers and biotechnology and pharmaceutical companies).

For determining the relative quantities of the proteins in the two samples, a little over one-third of the participants used gel electrophoresis-based methods, whereas approximately two-thirds of the laboratories used MS-based techniques (Table II). In the case of 1D gel electrophoresis, protein detection included Coomassie Blue (Coomassie), silver stain (silver), and fluorescent stains such as SYPRO Ruby (fluorescent). For 2D gel electrophoresis, in addition to Coomassie, silver, and fluorescent protein detection, DIGE-based approaches and, in one case, radioactive labels (14) were used. Two-thirds of the participants who used MS-based techniques used stable isotope labeling for the analysis. These included isotope-coded protein label (isotope ICPL) (15), isobaric tags for relative and absolute quantitation (isotope iTRAQ) (16), *O*-methylisourea (isotope *O*-methylisourea) (17), and ^{18}O labeling (isotope $^{16}O^{18}O$) (18). The remaining third of the participants who used MS-based techniques utilized a label-free approach based on either ion current or spectral counting. To evaluate the data, the percent error of the expected and observed ratios for each of the eight proteins was assessed using the following formula: percent error of ratio = [|(expected ratio − observed ratio)|/expected ratio] × 100.

Evaluation of the results submitted by the study participants is shown in Figs. 3 and 4. Individual values for percent error of

ratio for seven proteins (all study proteins excluding BSA, which was reported to study participants as being present at a 1:1 ratio) (Fig. 3*A*) and average values (shown both for seven proteins or six proteins excluding glycogen phosphorylase, the most challenging protein to be quantitated in this study) (Fig. 3*B*) are shown for all 13 methods that were used by the participants. Fig. 4 summarizes an assessment of the success of each method in quantifying the relative abundances of the seven proteins present at unknown ratios in the two mixtures.

It is important to emphasize that the PRG evaluation of data received from this study was never intended to promote or support any particular analytical method or type of instrumentation. The number of submitted responses was insufficient to yield any statistically significant measure of the ability of any method to get "the correct answer." Another essential point is that the results of this study are not only dependent on the absolute capabilities of the methods used but also on the experience levels of the scientists who performed the analyses because some of the participating laboratories were conducting these analyses for the first time. The study was undertaken with the goal of helping laboratories to both improve and expand the range of their own capabilities and to provide them with a means of testing the techniques that they use. It is only with these essential caveats in mind that any trends can be deduced from the 52 datasets that were submitted.

Not surprisingly, most of the protein identifications were conducted by MS. Only one laboratory used Edman degradation in combination with MS. Due to the relatively large quantities of material provided, most respondents were able to identify all eight proteins (data not shown). This demonstrates that identification by MS of proteins that are present in ample amounts has become routine.

In contrast to the clear success in protein identification, there was substantial variability in the accuracy of the reported quantitation results. The differences between the expected and observed protein ratios were found to be method- and protein-dependent. For the majority of methods, the greatest percent error of ratio was seen for glycogen phosphorylase (expected ratio of 1:76; Figs. 3 and 4). MS using ion current ($n = 10$) or spectral counting ($n = 1$) and 2D gel electrophoresis using radioactivity ($n = 1$) yielded ratios closest to the expected for glycogen phosphorylase.

Electrophoresis-based methods showed greater variability of percent error of ratio than MS-based approaches for this study (Fig. 3). Interestingly the results obtained using electrophoresis were closer to the expected values for proteins present at a 1:1 ratio than for proteins present at other ratios (Fig. 4, *A* and *B*). For electrophoresis, the lowest percent error of ratio was seen for 2D DIGE ($n = 3$) and 2D radioactivity ($n = 1$), whereas 2D Coomassie ($n = 4$) and 2D fluorescence ($n = 1$) exhibited relatively high values for percent error of ratio (Fig. 4*B*).

Overall electrophoresis showed greater variability of percent error of ratio than MS-based methods. Ratios obtained
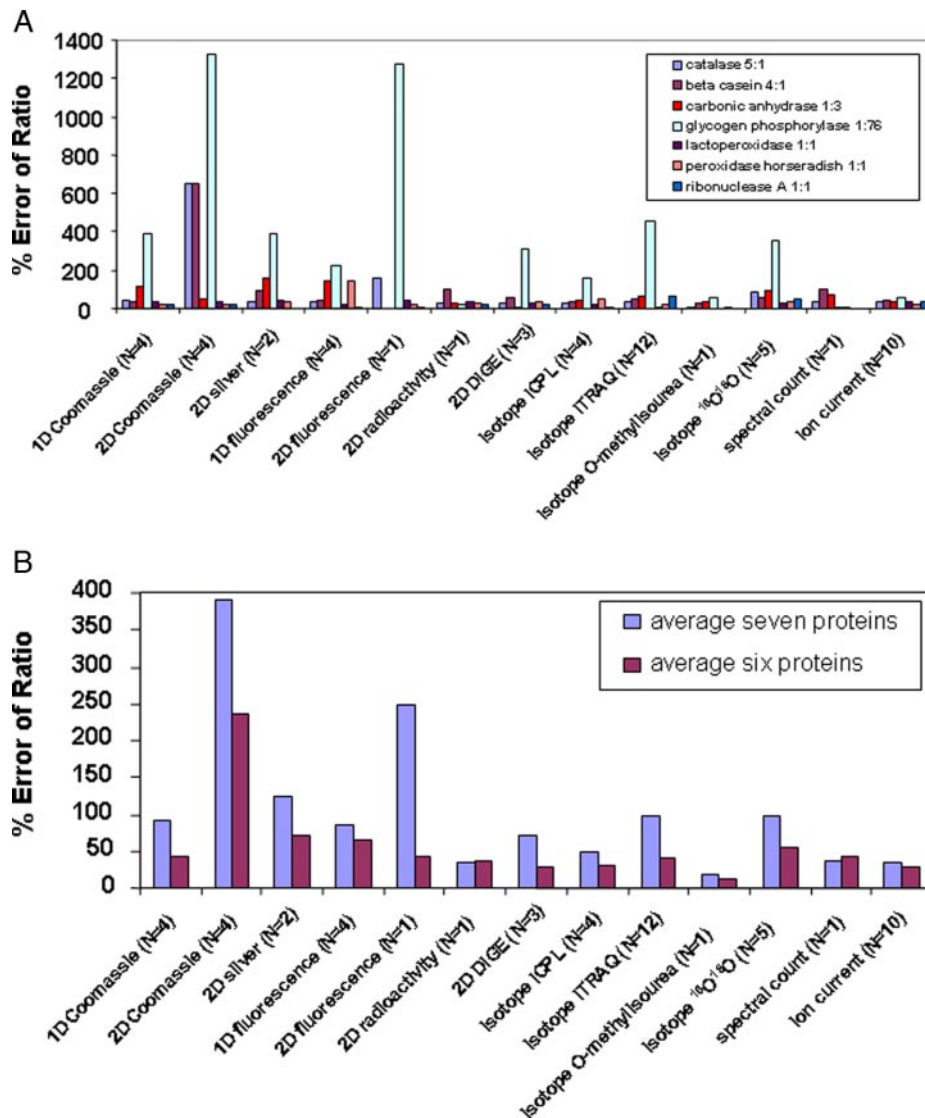
FIG. 3. **Percent error of ratio of relative quantification of proteins in the sample mixtures.** *A*, results for individual proteins for each method utilized by participants. *B*, average results for seven proteins for each method with and without inclusion of results for glycogen phosphorylase. The number of responses for each method are shown in *parentheses*. Note that results for BSA are not included because the ratio for this protein (1:1) was provided to participants. *iTRAQ*, isobaric tags for relative and absolute quantitation; *ICPL*, isotope-coded protein label.

by MS ion current ($n = 10$) or spectral counting ($n = 1$) were as close to the expected values as those obtained by stable isotope labeling ($n = 23$) (Figs. 3 and 4C). When MS with stable isotope labeling was used, the percent error of ratio was evenly distributed for all proteins except glycogen phosphorylase (Figs. 3 and 4C), whereas MS using ion current or spectral counting showed an even distribution of the percent error of ratio for all proteins including glycogen phosphorylase (Fig. 4). Quantitation by MS was not affected by prior separation of intact proteins by electrophoresis (data not shown). This result might have been different if a more complex protein mixture had been present.

Replicate analysis of the samples yielded results that were closer to the expected values as compared with analyses performed only once (Fig. 5). This was particularly true for electrophoresis-based methods. However, for the MS-based methods, additional analyses (triplicate and quadruplicate) did not yield further improvement of the data. The majority of participants stated that they did not rely on software alone for data analysis but performed additional manual validation of their results (data not shown). This is in line with reports of participants in previous PRG studies dealing with other aspects of proteomics (12).[2] It appears as though proteomics software that was available at the time of the 2006 study was not capable of fully automatic data processing for quantitative proteomics. As such, the study participants relied heavily on additional, manual evaluation and validation of the results by experienced core facility personnel.
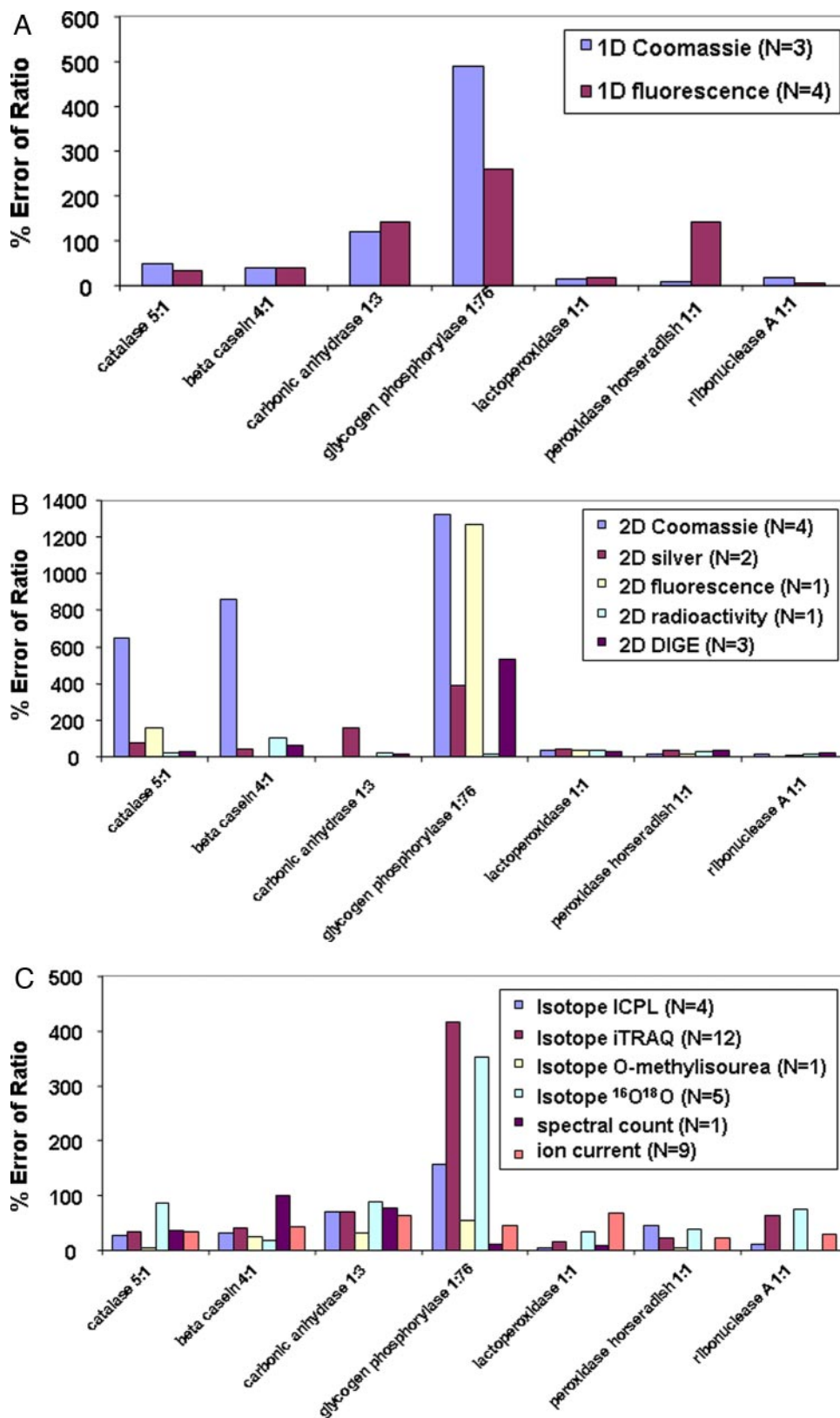
FIG. 4. **Average percent error of ratio results for relative quantitation of the study proteins (excluding BSA) grouped by analysis method: *A*, 1D PAGE; *B*, 2D PAGE; *C*, MS-based methods.** The number of responses for each method is shown in *parentheses*. *iTRAQ*, isobaric tags for relative and absolute quantitation; *ICPL*, isotope-coded protein label.
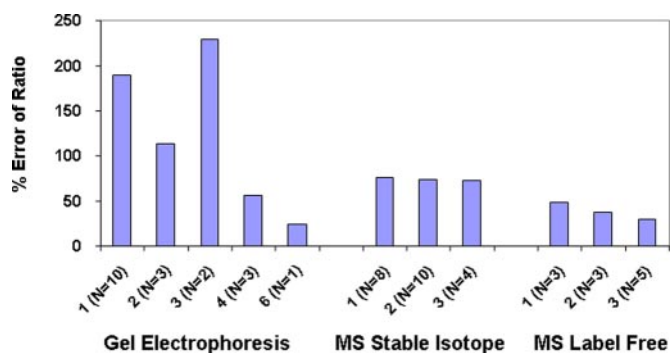
FIG. 5. **Influence of the number of times an analysis was repeated on the percent error of ratio.**

According to the results of the report survey, quantitative proteomics was being offered as a service in about half of the participants' facilities. The great majority of respondents felt that there is definitely a demand for relative protein quantitation for projects submitted to their laboratories. Judging by the participants' responses about the difficulty of the study and their confidence in their quantitation data, the level of experience with quantitative proteins was of major significance in determining the success of the analysis.

The PRG2006 study is the first of its kind and has elicited substantial interest not only by members of the ABRF community but also by other societies and scientific journals (19–21). As stated earlier, it is important to remember that there were not enough submitted responses to draw statistically significant conclusions about the capabilities of various methods for relative quantitation. In addition, it can be anticipated that the results obtained by the participants might have been quite different if a more complex sample or smaller protein quantities had been provided. Moreover the eight easily solubilized proteins that were provided in the samples were chosen to be appropriate for all of the different approaches used in this study. It is likely that some of the techniques used by participants would only be successful with this type of sample. Another aspect that must be considered is that the success of a particular quantitation method depends greatly on both the analytes and the design of the experiment, particularly the number of replicates. The study results reported here demonstrate the range of capabilities of a variety of different methods for quantitative proteomics. It is clear, however, that advances in a number of areas are still needed. In addition, there is no consensus in the proteomics community on how to report quantitative data with regard to statistical confidence. Clearly there is a need for additional comparative assessments of methods for quantitative proteomics. This study and others that will follow will undoubtedly provide the proteomics community with valuable information on methods and approaches that can be successfully used for relative protein quantitation.

REFERENCES

1. Glish, G. L., and Vachet, R. W. (2003) The basics of mass spectrometry in the twenty-first century. *Nat. Rev. Drug Discov.* **2,** 140–150
2. Klose, J. (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* **26,** 231–243
3. O'Farrell, P. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250,** 4007–4021
4. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246,** 64–71
5. Karas, M., and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60,** 2299–2301
6. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **1,** 252–262
7. Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E. M. (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* **25,** 117–124
8. Mallick, P., Schirle, M., Chen, S. S., Flory, M. R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B., and Aebersold, R. (2007) Computational prediction of proteotypic peptides for quantitative proteomics *Nat. Biotechnol.* **25,** 125–131
9. Kolkman, A., Dirksen, E. H. C., Slijper, M., and Heck, A. J. R. (2005) Double standards in quantitative proteomics. Direct comparative assessment of difference in Gel electrophoresis and metabolic stable isotope labeling. *Mol. Cell. Proteomics* **4,** 255–266
10. Arnott, D. P., Gawinowicz, M. A., Grant, R. A., Lane, W. S., Packman, L. C., Speicher, K., and Stone, K. (2002) Proteomics in mixtures: study results of ABRF-PRG02. *J. Biomol. Tech.* **13,** 179–186
11. Arnott, D., Gawinowicz, M. A., Grant, R. A., Neubert, T. A., Packman, L. C., Speicher, K. D., Stone, K., and Turck, C. W. (2003) ABRF-PRG03: phosphorylation site determination. *J. Biomol. Tech.* **14,** 205–215
12. Arnott, D., Gawinowicz, M. A., Kowalak, J. A., Lane, W. S., Speicher, K. D., Turck, C. W., K.A., W., and Neubert, T. A. (2007) ABRF-PRG04: differentiation of protein isoforms. *J. Biomol. Tech.* **18,** 124–134
13. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. *Electrophoresis* **18,** 2071–2077
14. Poznanovic, S., Wozny, W., Schwall, G. P., Sastri, C., Hunzinger, C., Stegmann, W., Schrattenholz, A., Buchner, A., Gangnus, R., Burgemeister, R., and Cahill, M. A. (2005) Differential radioactive proteomic analysis of microdissected renal cell carcinoma tissue by 54 cm isoelectric focusing in serial immobilized pH gradient gels. *J. Proteome Res.* **4,** 2117–2125
15. Schmidt, A., Kellermann, J., and Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5,** 4–15
16. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., K., P., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P.,

Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3,** 1154–1169

17. Beardsley, R. L., Karty, J. A., and Reilly, J. P. (2000) Enhancing the intensities of lysine-terminated tryptic peptide ions in matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **14,** 2147–2153

18. Yao, X., Freas, A., Ramirez, J., Demirev, P. A., and Fenselau, C. (2001) Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **73,** 2836–2842

19. Cottingham, K. (2006) ABRF proteomics studies: quantitation and standards. *J. Proteome Res.* **5,** 740

20. Gilchrist, A., Au, C. E., Hiding, J., Bell, A. W., Fernandez-Rodriguez, J., Lesimple, S., Nagaya, H., Roy, L., Gosline, S. J. C., Hallett, M., Paiement, J., Kearney, R. E., Nilsson, T., and Bergeron, J. J. (2006) Quantitative proteomics analysis of the secretory pathway. *Cell* **127,** 1265–1281

21. Lee, T.-S. (2006) Label-free mass spec, multiplexed gels perform well in PRG quantitation study. *ProteoMonitor* **6,** 2–3