

Shotgun proteomics using the iTRAQ isobaric tags

Kunal Aggarwal, Leila H. Choe and Kelvin H. Lee

Advance Access publication date 10 May 2006

Abstract

Shotgun proteomic methods involving isobaric tagging of peptides enable high-throughput proteomic analysis. iTRAQ reagents allow simultaneous identification and quantitation of proteins in four different samples using tandem mass spectrometry (MS). In this article, we provide a brief description of proteome analysis using iTRAQ reagents and review the current applications of these reagents in proteomic studies. We also compare different aspects of protein identification including protein sequence coverage and proteome coverage obtained using iTRAQ reagents with those using other shotgun proteomic techniques. We briefly discuss the issue of isotope purity correction in measured peak areas during protein quantitation using iTRAQ reagents. Finally, we conclude with some of the current challenges in MS-based proteomic analysis that are limiting protein identifications obtained by different shotgun proteomic methods.

Keywords: isobaric tag; iTRAQ; shotgun proteomics; tandem mass spectrometry

INTRODUCTION

Catalysed by recent advances in quantitative mass spectrometry (MS), shotgun proteomics can provide a relatively high-throughput assessment of changes in protein expression. This approach to proteome analysis usually involves differential isotope labelling of proteins and peptides either metabolically (*in vivo*), enzymatically or chemically using external reagents/tags [1–4]. The labelled peptides are separated by multidimensional liquid chromatography (LC), and the resolved peptides are analysed using MS. The peak areas or intensities observed in the MS spectra of eluted peptides are used to quantify relative peptide (protein) abundance. Shotgun proteomic methods involving isotopic labelling address some of the limitations faced in traditional gel-based proteomic approaches, namely, an inability to analyse highly basic/hydrophobic proteins. However, these approaches still suffer from some limitations such as an inability to multiplex (analyse multiple samples in parallel) and to quantify zero protein expression level (because of the need to observe mass shifts by MS).

In contrast to isotopic labelling strategies, shotgun proteomic methods involving isobaric tagging of peptides enable simultaneous identification and quantification of peptides using tandem MS [5] and permit parallel proteome analysis of more than two samples. One such method commercialized by Applied Biosystems is called iTRAQ and uses four amine specific isobaric reagents to label the primary amines of peptides from four different biological samples [6]. The labelled peptides from each sample are mixed, separated using two-dimensional LC and analysed using MS and tandem mass spectrometry (MS/MS). Because of the isobaric nature of these reagents, the same peptide from each sample appears as a single peak in the MS spectrum, thus reducing the complexity in the MS spectrum when compared with an isotopic labelling technique such as ICAT (isotope coded affinity tag) [3]. Upon collision-induced dissociation, the iTRAQ-tagged peptides fragment to release reporter ions (at 114.1, 115.1, 116.1 and 117.1 m/z) and b - and y -ion series among other fragments. The peak area of the reporter ions

Corresponding author. Kelvin H. Lee, School of Chemical and Biomolecular Engineering, Cornell University, 120 Olin Hall, Ithaca, NY 14853-5201, USA. E-mail: KHL9@cornell.edu

Kunal Aggarwal is a graduate student in the School of Chemical and Biomolecular Engineering, Cornell University, NY, USA.

Leila H. Choe is a Research Support Specialist in the School of Chemical and Biomolecular Engineering, Cornell University, NY, USA.

Kelvin H. Lee is Associate Professor in the School of Chemical and Biomolecular Engineering, Cornell University and is the Director of the Cornell Institute for Biotechnology and Life Science Technologies, NY, USA.

are used to assess relative abundance of peptides and consequently the proteins from which they are derived [6, 7].

In this article, we review some of the recent studies employing isobaric reagents for proteomic analysis. We also compare different aspects of protein identification and quantitation obtained in isobaric tagging experiments with those observed in other shotgun proteomic experiments.

RECENT WORK USING ISOBARIC TAGS

The iTRAQ approach has been successfully applied to a variety of prokaryotic and eukaryotic samples including *Escherichia coli*, yeast, human saliva, human fibroblasts and mammary epithelial cells to identify and quantify the proteins in these samples. Hardt *et al.* [8] used iTRAQ reagents to study the diurnal effects on the proteomic composition of human parotid saliva collected at four different time points during a day. Cong *et al.* [9] compared proteomes of human fibroblasts in four different biological states: replicatively senescent (under permanent growth arrest), stress-induced prematurely senescent, quiescent and young replicating, to identify the signature proteins of each biological state. Zhang *et al.* [10] combined the iTRAQ labelling approach with immunoprecipitation to quantify tyrosine phosphorylated peptides in four different populations of epidermal growth factor-stimulated mammary epithelial cells. Tyrosine phosphorylated peptides were enriched using immunoprecipitation after labelling with iTRAQ reagents. This work suggests that protein quantitation using iTRAQ labels is not limited to cell-culture-derived samples. DeSouza *et al.* [11] used iTRAQ labels to identify five potential protein markers for endometrial cancer. They also identified a different set of four proteins using cleavable ICAT (cICAT) as potential markers for endometrial cancer. They identified different proportions of proteins in various functional categories using these techniques. A greater percentage of proteins identified using cICAT as compared with iTRAQ were involved in signalling. iTRAQ analysis was able to identify a greater number of ribosomal proteins and transcription factors [11]. We were also able to identify some low-abundance proteins and transcription factors in *E. coli* using iTRAQ labels [7]. iTRAQ reagents have been observed to result in an improved signal-to-noise ratio and an enhanced

fragmentation with increased signal intensities in matrix-assisted laser desorption ionization (MALDI) tandem time-of-flight (TOF) MS [8]. This may result in the identification of more peptides.

Like other shotgun proteomic approaches, the isobaric tagging strategy provides multiple independent measures of the relative abundance of a protein. Protein quantitation results obtained using iTRAQ have been verified by analysing standard mixtures of proteins of known proportions [6, 12]. Differential expression of selected proteins in various samples as detected using iTRAQ analysis has also been confirmed qualitatively using western blot analysis [9]. We have employed different statistical outlier exclusion methods on iTRAQ data to obtain statistically relevant estimates of relative protein abundance [7]. We have also compared protein quantitation of *E. coli* samples obtained using iTRAQ reagents with that using two-dimensional electrophoresis (2-DE). We observed that the two protein quantitation approaches provide similar quantitation results. Most of the protein quantitation data from the two techniques was consistent within a 2-fold change. However, iTRAQ was observed to provide a more consistent protein quantitation as compared with 2-DE. About 95% of the proteins were quantified with a coefficient of variation <0.53 using isobaric tags and <0.81 using 2-DE. Further, iTRAQ was observed to provide a more consistent quantitation in comparison with 2-DE for proteins stained with lower intensity on 2-DE gels [13].

In the rest of this article, we compare iTRAQ technology with other shotgun proteomic approaches under different aspects of proteome analysis. Specifically, we use the iTRAQ data collected during proteomic analysis of *E. coli* for comparison. The details of protein extraction, labelling and quantitation procedures used in this work have been published previously [7]. Briefly, tryptic peptides from *E. coli* cells expressing *rhsA* elements at four different levels were differentially labelled with iTRAQ reagents. The complex mixture of labelled peptides was separated using strong cation-exchange (SCX) fractionation followed by reversed-phase high pressure liquid chromatography (HPLC). The resolved peptides were studied using MALDI TOF/TOF analysis on a 4700 Proteomics Analyzer running v2.0 software. The spectral data was searched against a database of the translations of all genome coding sequences from the *E. coli* K-12 genome (4289 ORFs) with semi-trypsin specificity to identify

the peptide sequence and the source protein [14]. The search parameters used were as follows: 50 ppm precursor mass tolerance, 0.3 Da MS/MS peptide mass tolerance, two maximum missed cleavages, variable methionine oxidation and two fixed and three variable modifications associated with the labelling chemistry. The peak areas of the signature ions, corresponding to the isobaric reagents, in the MS/MS spectra were corrected for overlapping isotopic contributions as described previously [15]. The corrected peak areas were then used to calculate the relative abundance of each detected peptide and the corresponding protein from which it was derived. In this study, 23 139 MS/MS spectra were collected. 5063 peptides were matched to 780 unique proteins with high MASCOT ion scores ($P < 0.05$).

PROTEIN SEQUENCE COVERAGE

The total protein sequence coverage obtained while performing protein identification depends on the number of unique peptide matches per protein and the length of the matched peptide.

Peptide matches per protein

Because iTRAQ reagents target all peptides for labelling, multiple peptides from the same protein can be detected during MS. Ross *et al.* [6] detected 4.5

peptides per protein on average during yeast proteome analysis using iTRAQ labels. Cong *et al.* [9] identified 45% of the proteins (108 out of 240) using at least two significant peptides in the proteome analysis of human fibroblasts. In our study, more than 65% of the proteins (527 out of 780) were identified with at least two peptide matches ($P < 0.05$) per protein [7]. On average, 6 peptides were matched per protein for all the identified proteins.

Peptide length

A normal distribution profile of peptide length is generally observed for high-scoring peptides detected using MS, irrespective of the type of MS ionization or the MS instrumentation used [16, 17]. A similar peptide length distribution profile is observed for peptides originating from isobarically labelled samples in our study (Figure 1). Fewer peptides shorter than seven amino acids are matched with a high confidence. This bias is observed in MALDI MS experiments because peptides shorter than six or seven amino acids typically appear in the low-mass region in a MALDI MS spectrum, which is otherwise dominated by matrix peaks; hence, these peaks are not used in the database search. An average peptide length of 12–23 amino acid residues has been reported in the literature for peptides detected using electrospray (ESI) MS (using QSTAR and LTQ mass spectrometers) and MALDI MS (using 4700

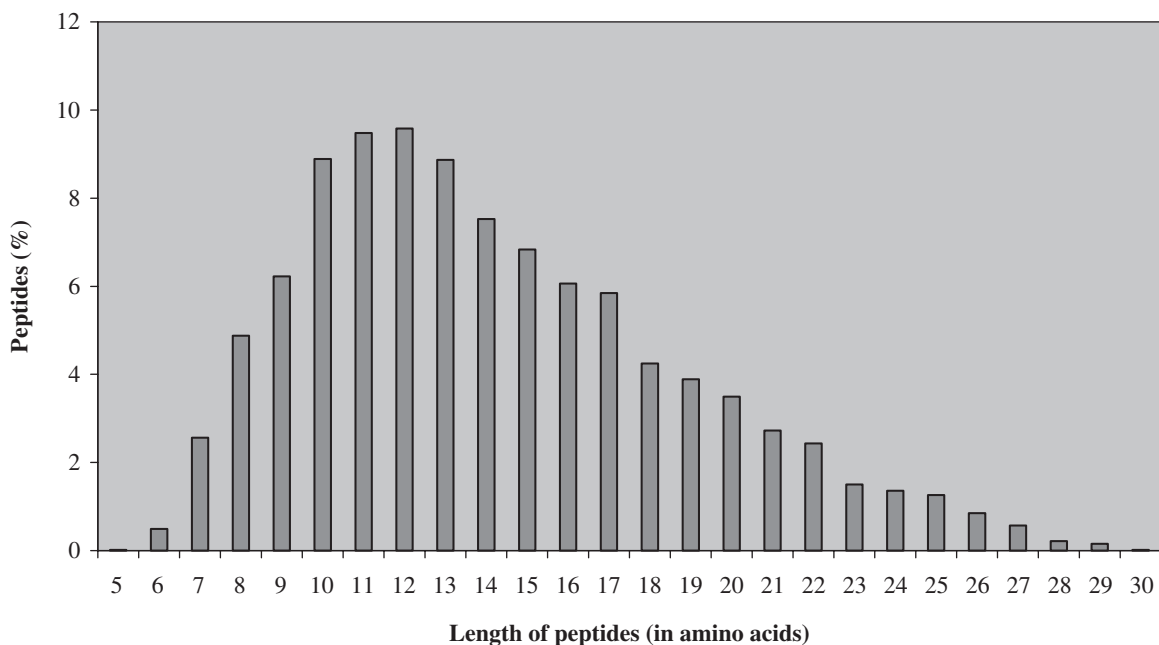


Figure 1: Length distribution of peptides matched with a high score ($P < 0.05$).

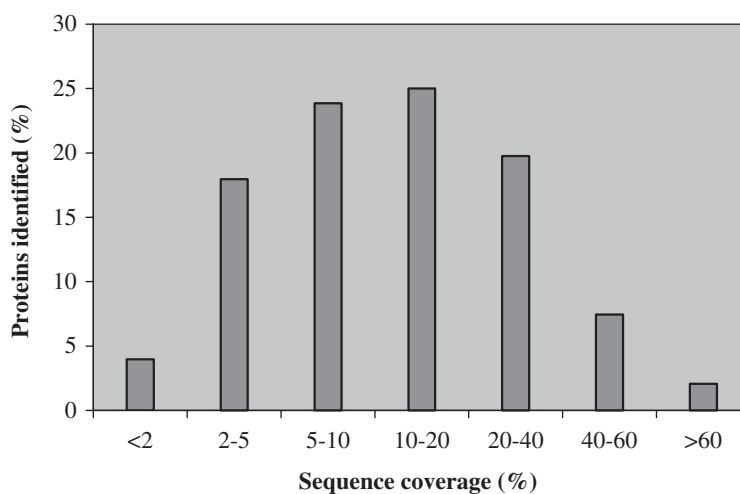


Figure 2: Sequence coverage distribution of proteins identified. The percentage values on the y-axis are based on a total of 780 proteins identified in this study.

Proteomics Analyzer) [17]. Consistent with the values reported in literature, an average peptide length of 14 amino acids was observed for high-scoring peptides in this study. Sixty-five percent of these peptides were at least 12 amino acids long. We also observed a correlation between the peptide length and MASCOT ion score or significance level. For peptides shorter than 12 amino acids, the average significance level of detection decreased with the peptide length.

Sequence coverage

Protein sequence coverage can be calculated using sequence information of non-redundant peptides matched with a protein. A sequence coverage of 2–77% has been obtained in *E. coli* proteome analysis using a 2D-LC MALDI TOF/TOF approach [18]. Using iTRAQ labels to study the proteome of rat pancreatic zymogen granule membranes, Chen *et al.* [19] obtained protein sequence coverage of 0.1–27%. Here, we observe a sequence coverage of 0.8–83.5% was observed for all the proteins identified. Fifty percent of the proteins were identified with sequence coverage of at least 10% (Figure 2). However, only 10% of the proteins were identified with sequence coverage >40%.

AMINO ACID CONTENT OF PEPTIDES

Zhen *et al.* [18] have reported the frequency of appearance for *E. coli* amino acids in high-scoring

peptides analysed using 2D-LC MALDI TOF/TOF. They observed a higher percentage of negatively charged residues in high-scoring peptides compared with their natural abundance. The presence of acidic residues in a peptide may result in more intense and comprehensive fragment ions thus enabling a higher confidence identification [18, 20]. Thus, a greater number of high-scoring peptides can be expected to be rich in negatively charged residues. Similar to Zhen *et al.* [18], a higher percentage of negatively charged residues in high-scoring peptides are also observed here with respect to all *E. coli* proteins (Table 1). On average, 16.8% of the residues in the high-scoring peptides are negatively charged. A lower abundance of positively charged residues in the high-scoring peptides (7%) is observed with respect to Zhen *et al.* [18] (10.8%) and compared with all *E. coli* proteins (12.5%). This may be because 33% of the high-scoring peptides in this study have neither lysine nor arginine at their ends due to the inclusion of semitryptic peptides. For tryptic peptides only (i.e. 67% of the high-scoring peptides), the observed frequency of positively charged residues is 10%, which is similar to that observed in all *E. coli* proteins and reported in the work of Zhen *et al.* [18].

PROTEOME COVERAGE

Proteins belonging to all functional categories have been detected using iTRAQ-based shotgun proteomic methods in different organisms [6, 7, 11]. Because of their relative abundance, a majority of the

Table 1: Frequency of different types of amino acid residues

Type of residues	Frequency of residues observed (%)		
	in peptides using isobaric tags	in Zhen et al. [18]	in all <i>E. coli</i> proteins
Non-polar (aliphatic, G,A,P,V,L,I,M)	48.5	42.8	47.8
Non-polar (aromatic, F,Y,W)	9.0	7.2	8.2
Polar (uncharged, S,T,C,N,Q)	18.8	23.2	20.3
Polar (positive, K,H,R)	7.0	10.8	12.5
Polar (negative, D,E)	16.8	16.0	11.2

Table 2: Molecular weight distribution of proteins identified

MW (in kDa)	Number of proteins in class	Number of proteins in class identified using isobaric tags	Total proteins in class identified using isobaric tags (%)
0–20	1044	163	15.6
20–40	1726	272	15.8
40–60	902	202	22.4
60–80	255	74	29.0
80–100	128	44	34.4
100–120	42	14	33.3
> 120	38	11	28.9

proteins identified in these studies are involved in housekeeping functions including biosynthesis, cell processes and metabolism. However, there does not appear to be a bias against any particular functional class of proteins during identification. Proteins from multiple cellular locations were identified in human fibroblasts using iTRAQ labelling strategy [9]. Chen *et al.* [19] have used 2D-gel electrophoresis and 2D-LC with MS to identify proteins from rat pancreatic zymogen granule membranes. Using iTRAQ reagents to quantify the enrichment of intrinsic membrane proteins in a purification process, they were able to distinguish intrinsic membrane proteins from soluble and peripheral membrane proteins. The cellular location of the proteins identified in the current study was determined using EcoCyc database [21]. About 9.4% of the total proteins identified in our study (73 out of 780) are known to be membrane proteins. This set of identified proteins constitutes 12.9% of all the known *E. coli* membrane proteins.

The extent of proteome coverage obtained using a proteomic analysis strategy can also be assessed in terms of the percentage of proteins identified across different ranges of molecular weight (MW) and isoelectric point (pI). Traditional gel-based approaches to protein identification are limited in

their ability to detect proteins with extremes in MW and pI. Previous studies have reported the use of shotgun proteomic methods to detect proteins that lie on the extreme ends of MW and pI [22].

MW dependence of proteins identified

Proteins ranging across a wide spectrum of MW were identified using isobaric tags. Table 2 lists the total number of *E. coli* proteins within a certain theoretical MW range and the number of proteins belonging to that range identified in this work. For example, 202 proteins identified in this work have MW between 40 and 60 kDa. This group is 22.4% of the total *E. coli* proteins with a MW of 40–60 kDa (902 proteins). It is interesting to note that about 21% of the total proteins identified (163 out of 780) have MW < 20 kDa, and this group represents only 15.6% of the total *E. coli* proteins with MW < 20 kDa. Generally, the number of trypsin cleavage sites decreases with the protein size, and thus the theoretical number of peptides available for detection decreases. This observation was also made with peptide matches per protein increasing with the size of the protein identified (Figure 3a). Interestingly, uniform average sequence coverage of 10–15% was obtained across all MW ranges of proteins identified, except for proteins with

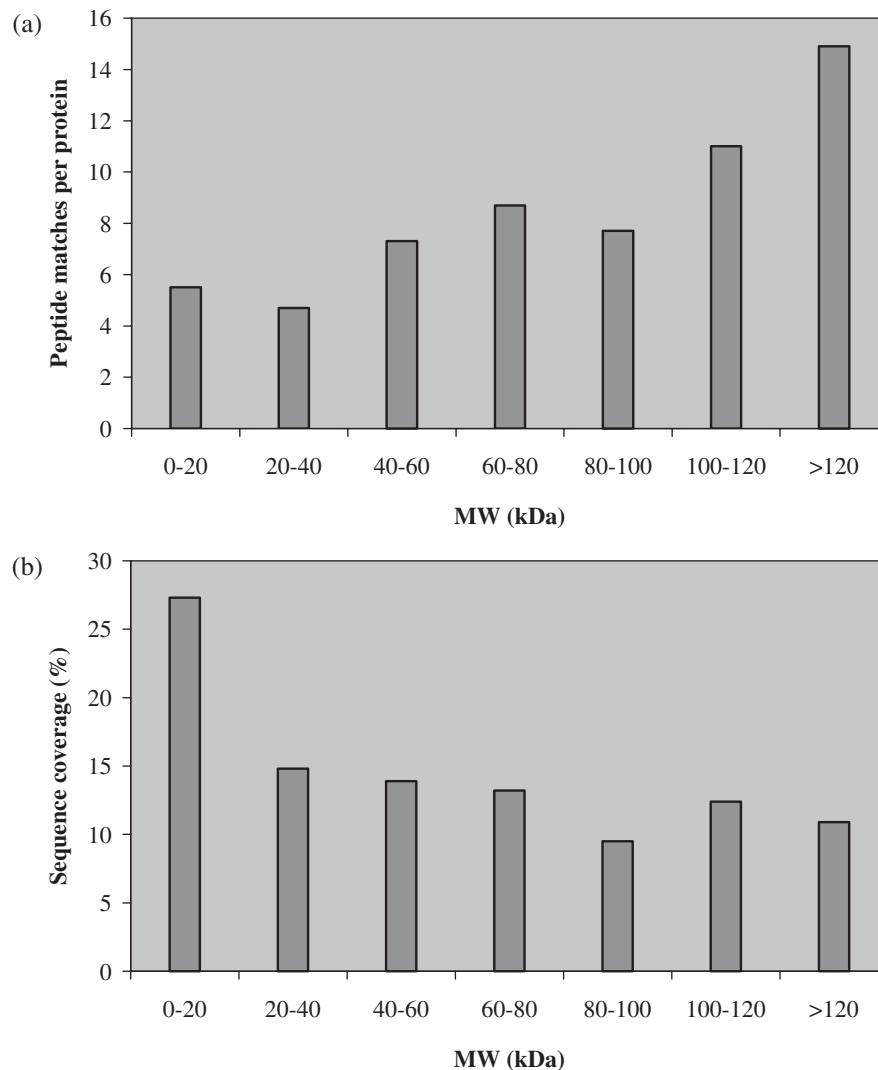


Figure 3: Average values of (a) peptides matched per protein and (b) sequence coverage for a total of 780 proteins identified in this work. The identified proteins have been grouped into different MW bins according to their theoretical size.

MW < 20 kDa, where an average sequence coverage of 27% was observed (Figure 3b).

pI dependence of proteins identified

Proteins within a wide range of pI were also detected in this study. Table 3 lists the total number of *E. coli* proteins predicted to be within a certain pI range and the number of proteins belonging to that range identified in this work. About 26.3% of the total *E. coli* proteins predicted to have pI 5–6 were identified in this study (387 out of 1470). The largest fraction of the total proteins identified in this work (387 corresponding to 49.6% of 780) has a pI in the 5–6 range. This observation is expected because this group is also the largest based on predicted pI. Only

one protein with a predicted pI < 4 was identified in this work. Proteins with pI < 4 have fewer arginine/lysine residues for undergoing tryptic digestion, thus reducing the number of available peptides for detection. Interestingly, 34.5% of all the proteins with a predicted pI > 11 were identified in this study.

Detection of highly expressed proteins

As an indirect measure of the extent of proteome coverage obtained using iTRAQ-based shotgun proteomics, we estimated the number of proteins corresponding to the highly expressed genes that were detected in this study. DNA microarray analysis was performed on the same samples that were used for proteomic analysis. Briefly, RNA was stabilized

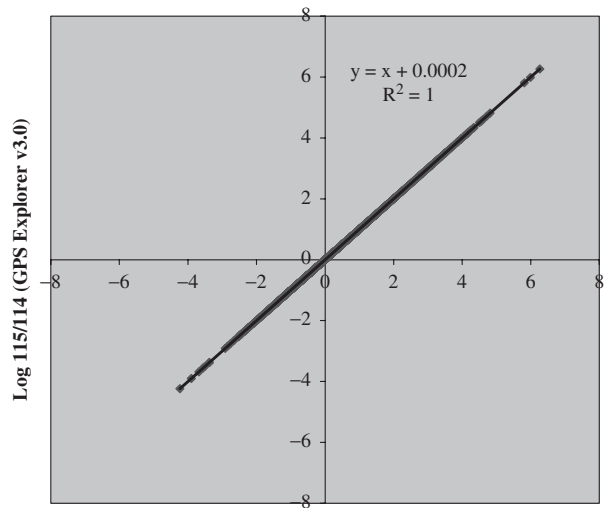
Table 3: Isoelectric point (pI) distribution of proteins identified

pI	Number of proteins in class	Number of proteins in class identified using isobaric tags	Total proteins in class identified using isobaric tags (%)
<4	11	1	9.1
4–5	507	109	21.5
5–6	1470	387	26.3
6–7	732	125	17.1
7–8	203	25	12.3
8–9	421	44	10.5
9–10	631	57	9.0
10–11	130	22	16.9
>11	29	10	34.5

Table 4: mRNA abundance of genes used for proteomic analysis

Class	Average mRNA signal
All <i>E.coli</i> proteins	4568.4
Proteins identified using isobaric tags	12065.6
Proteins not identified	2950.8

by suspending cells in RNA Later (Ambion, Austin, TX, USA) and extracted using MasterPure RNA purification kits (Epicentre, Madison, WI, USA). cDNA fragments were created by two rounds of reverse transcriptase polymerase chain reaction using random hexamer primers [23]. The resulting cDNA was then fluorescently labelled and hybridized to *E. coli* antisense GeneChip probe arrays (Affymetrix, Santa Clara, CA, USA). Intensity data from scanned chip images was analysed using GC-RMA procedure [24] available in the Genetraffic software (Iobion, La Jolla, CA, USA) to calculate gene expression values. An average mRNA signal of 12,065.6 was observed for all the proteins identified in this study (Table 4). This value is more than four times the average mRNA signal of the proteins not detected in this study (2950.8 corresponding to 3547 genes). Of the proteins not detected, 5.2% have an average mRNA signal >12,065.6. This observation is consistent with our identification of the majority of the highly expressed proteins in this study. The detected proteins were not limited to those highly expressed; 21 different transcription factors and other low-abundance proteins were also identified as well as quantified in this study [7].



Log 115/114 (Purity correction algorithm similar to i-Tracker)

Figure 4: Comparison of the purity correction algorithm provided by GPS Explorer v3.0 to our algorithm that is mathematically similar to the one implemented in i-Tracker. Log peak area ratios (115/114) calculated after correction using either algorithm have been plotted.

PURITY CORRECTION FOR QUANTITATION

The isobaric tags used for proteome analysis differ in the isotopic compositions of nitrogen, carbon and oxygen but have identical masses [6]. Due to isotopic contamination in tags, the peak area for each reporter ion will have some contribution from other reporter ions. Shadforth *et al.* [15] have presented a detailed procedure to calculate true peak areas that account for overlapping isotopic contributions using the reagent purity values provided by the manufacturer when using electrospray ionization. This procedure has been implemented in i-Tracker software. Although, the authors of the software observed a high correlation between the relative peak area values (measure of relative peptide abundance) obtained using i-Tracker and ProQuant (Applied Biosystems), there were several peptides in which the relative peak area values computed by the two software programs did not match [15]. We separately derived a similar algorithm for purity correction in Microsoft Excel and applied it to the raw peak areas obtained from 4700 Proteomics Analyser version 2.0 software (Applied Biosystems). A very high correlation was observed between the relative peak areas obtained using GPS Explorer v3.0 (Applied Biosystems) and those calculated using our purity correction algorithm that is mathematically similar to the one implemented in the i-Tracker (Figure 4).

A maximum difference of 3% was observed in the relative peak area values obtained with our method versus GPS Explorer v3.0. It is possible that the difference between the values obtained from ProQuant and i-Tracker, as seen in the work of Shadforth *et al.* [15] is due to some other systematic errors. For example, ProQuant and i-Tracker may interpret mass spectra peak areas in different ways (i-Tracker uses non-centroided tandem mass peak lists). In this study, we used raw peak areas for our method.

CONCLUSIONS

Shotgun proteomics experiments involving iTRAQ reagents have been successful in identifying and quantifying proteins across a variety of prokaryotic and eukaryotic samples. One of the major advantages of iTRAQ technology is that parallel proteomic analysis of four different samples can be achieved, reducing the total analysis time. Several proteomic studies reported in the literature have benefited from this multiplexing ability of iTRAQ technology.

Proteome analysis using iTRAQ reagents compares favourably with the other shotgun proteomic approaches. The peptide length distribution profile and amino acid content of the isobarically labelled peptides detected during MS analysis are similar to those obtained using other MS-based approaches. This suggests that iTRAQ reagents do not negatively interfere with peptide fragmentation. An improved signal-to-noise ratio with increased signal intensity in MALDI TOF/TOF of isobarically tagged peptides can not only result in detection of a greater number of peptides per protein with high confidence but also in detection of some low-abundance proteins (as observed in this work and reported previously [8]). iTRAQ reagents can be used to identify and quantify proteins across diverse MW and pI ranges, functional categories and cellular locations. iTRAQ reagents also allow multiple, independent measures of protein abundance in the same experiment, enabling statistical estimates of protein quantitation. Protein identification based on multiple peptide matches (versus one peptide) will improve confidence in the identification as well as in the quantitation. However, more than one-third of the proteins in all the iTRAQ studies conducted till date have been identified using only one peptide. Here, only 22% of the MS/MS collected were used to identify proteins with high confidence. Many high-quality spectra

were not matched to any protein. New search algorithms and databases for improved open reading frame assignments may allow allocation of unassigned MS/MS data to sequences. An enhanced spectra assignment will increase peptide matches per protein, thus reducing false positive identifications and enabling statistical quantitation for a greater number of identified proteins. Further, it may enhance the total proteome coverage obtained in a shotgun proteomic experiment.

Key Points

- iTRAQ reagents allow parallel proteomic analysis of four different samples, reducing the analysis time.
- Protein sequence coverage obtained using iTRAQ reagents is similar to that obtained using other shotgun proteomic approaches.
- iTRAQ reagents can be used to identify and quantify proteins across diverse MW and pI ranges, functional categories, cellular locations and abundances.
- Further developments in search algorithms and databases are needed for enhanced MS/MS spectra assignment.

Acknowledgements

We thank Applied Biosystems, Framingham, MA, for access to reagents and important discussions. This work was funded in part by the USDA Agricultural Research Service Specific Cooperative Agreement 58-1907-1-146, the National Science Foundation BES-0120315 and New York State Office of Science, Technology and Academic Research.

References

1. Cagney G, Emili A. *De novo* peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat Biotechnol* 2002; **20**:163–70.
2. Conrads TP, Alving K, Veenstra TD, *et al.* Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and N-15-Metabolic labeling. *Anal Chem* 2001; **73**:2132–9.
3. Gygi SP, Rist B, Gerber SA, *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 1999; **17**:994–9.
4. Yao XD, Freas A, Ramirez J, *et al.* Proteolytic O-18 labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Anal Chem* 2001; **73**:2836–42.
5. Thompson A, Schafer J, Kuhn K, *et al.* Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* 2003; **75**:1895–904.
6. Ross PL, Huang YLN, Marchese JN, *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004; **3**:1154–69.

7. Aggarwal K, Choe LH, Lee KH. Quantitative analysis of protein expression using amine-specific isobaric tags in *Escherichia coli* cells expressing *rhsA* elements. *Proteomics* 2005;**5**:2297–308.
8. Hardt M, Witkowska HE, Webb S, et al. Assessing the effects of diurnal variation on the composition of human parotid saliva: Quantitative analysis of native peptides using iTRAQ reagents. *Anal Chem* 2005;**77**:4947–54.
9. Cong YS, Fan E, Wang E. Simultaneous proteomic profiling of four different growth states of human fibroblasts, using amine-reactive isobaric tagging reagents and mass spectrometry. *Mech Ageing Dev* 2006;**127**:332–43.
10. Zhang Y, Wolf-Yadlin A, Ross PL, et al. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol Cell Proteomics* 2005;**4**:1240–50.
11. DeSouza L, Diehl G, Rodrigues MJ, et al. Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. *J Proteome Res* 2005;**4**:377–86.
12. Unwin RD, Pierce A, Watson RB, et al. Quantitative proteomic analysis using isobaric protein tags enables rapid comparison of changes in transcript and protein levels in transformed cells. *Mol Cell Proteomics* 2005;**4**:924–35.
13. Choe LH, Aggarwal K, Franck Z, et al. A comparison of the consistency of proteome quantitation using two-dimensional electrophoresis and shotgun isobaric tagging in *Escherichia coli* cells. *Electrophoresis* 2005;**26**:2437–49.
14. NCBI FTP site. <ftp://ftp.ncbi.nih.gov/> (12 February 2006, date last accessed).
15. Shadforth IP, Dunkley TPJ, Lilley KS, et al. i-Tracker: For quantitative proteomics using iTRAQTM. *BMC Genomics* 2005;**6**.
16. ASMS 2002 Scientific Presentations. ESI and MALDI LC/MS-MS approaches for larger scale protein identification and quantification: Are they equivalent? http://www.biotech.iastate.edu/facilities/protein/seminars/ABI_ASMS_2002/presentations_index.htm (12 February 2006, date last accessed).
17. Elias JE, Haas W, Faherty BK, et al. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat. Methods* 2005;**2**:667–75.
18. Zhen YJ, Xu NF, Richardson B, et al. Development of an LC-MALDI method for the analysis of protein complexes. *J Am Soc Mass Spectrom* 2004;**15**:803–22.
19. Chen X, Walker AK, Strahler JR, et al. Organellar proteomics: Analysis of pancreatic zymogen granule membranes. *Mol Cell Proteomics* 2006;**5**:306–12.
20. Qin J, Chait BT. Preferential fragmentation of protonated gas-phase peptide ions adjacent to acidic amino-acid-residues. *J Am Chem Soc* 1995;**117**:5411–2.
21. Encyclopedia of *Escherichia coli* K12 Genes and Metabolism. <http://ecocyc.org/> (12 February 2006, date last accessed).
22. Washburn MP, Wolters D, Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol* 2001;**19**:242–7.
23. Rosenow C, Saxena RM, Durst M, et al. Prokaryotic RNA preparation methods useful for high density array analysis: comparison of two approaches. *Nucleic Acids Res* 2001;**29**: art. no. e112.
24. Wu Z, Irizarry RA, Gentleman R, et al. A model based background adjustment for oligonucleotide expression arrays. *Johns Hopkins University, Biostatistics Working Papers*. 2004: Working Paper 1.