Quantitative Proteomic Analysis Using Isobaric Protein Tags Enables Rapid Comparison of Changes in Transcript and Protein Levels in Transformed Cells*

Richard D. Unwin‡§, Andrew Pierce‡, Rod B. Watson¶, David W. Sternberg‖, and Anthony D. Whetton‡§‡‡

Isobaric tags for relative and absolute quantitation, an approach to concurrent, relative quantification of proteins present in four cell preparations, have recently been described. To validate this approach using complex mammalian cell samples that show subtle differences in protein levels, a model stem cell-like cell line (FDCP-mix) in the presence or absence of the leukemogenic oncogene TEL/PDGFRβ has been studied. Cell lysates were proteolytically digested, and peptides within each sample were labeled with one of four isobaric, isotope-coded tags via their N-terminal and/or lysine side chains. The four labeled samples are mixed and peptides separated by two-dimensional liquid chromatography online to a mass spectrometer (LC-MS). Upon peptide fragmentation, each tag releases a distinct mass reporter ion; the ratio of the four reporters therefore gives relative abundances of the given peptide. Relative quantification of proteins is derived using summed data from a number of peptides. TEL/PDGFRβ leukemic oncogene-mediated changes in protein levels were compared with those seen in microarray analysis of control and transfected FDCP-mix cells. Changes at the protein level in most cases reflected those seen at the transcriptome level. Nonetheless, novel differences in protein expression were found that indicate potential mechanisms for effects of this oncogene. ***Molecular & Cellular Proteomics 4:924–935, 2005.***

The development of approaches for measurement of relative expression of proteins between two (or more) samples is an essential aspect of systems biology. A common technique for this type of proteomic study has been two-dimensional gel electrophoresis, where proteins are separated by isoelectric point and molecular weight, and spot patterns are compared by sophisticated computer algorithms. Proteins of interest are then identified by MS. However, gel-based approaches have drawbacks, including amount of material required, reproducibility, and limiting of sensitivity by protein loss. These are balanced against recent advances in protein staining that allow intragel comparison of protein quantity from two or three samples (1).

An alternative approach to proteomic analysis is LC-MS (2). This provides an increased sensitivity compared with gel-based approaches and can catalogue protein present in a sample. However, relative quantitation of proteins using LC-MS is challenging. Quantification by analyzing two samples in parallel and comparing their mass spectrometric profiles is not feasible. Isotopic labeling of peptides, however, does allow two samples to be analyzed in a single experiment. Isotope-coded affinity tagging using ICAT reagent technology indicates peptide source, with peak height giving relative quantity (3). Comparing the ICAT reagent approach with two-dimensional gel electrophoresis, however, demonstrates that neither offers comprehensive coverage of a proteome (4). This is true in part because many proteins (and therefore peptides) do not contain cysteine, the amino acid used for covalent attachment of the isotopomer in ICAT reagents. Thus, much information can be discarded in the form of non-labeled peptides, whereas two-dimensional gel electrophoresis excludes many large, hydrophobic, and basic proteins.

Stable isotope labeling with amino acids in cell culture uses isotopes of essential amino acids (for example deuterated leucine) to label cells in culture (5). The samples are mixed, proteolytically digested, and run in LC-MS experiments. All leucine-containing peptides appear as "heavy" and "light" peaks, giving relative protein abundance. This elegant method can only be used on cultured cells; it is unsuitable for study of primary material. For ICAT reagent and stable isotope labeling with amino acids in cell culture technologies, the labeled peptides have different masses in an MS scan; this increases

From the ‡Faculty of Medical and Human Sciences, University of Manchester, Christie Hospital, Withington, Manchester, M20 9BX, United Kingdom, §Mass Spectrometry Laboratory, Paterson Institute for Cancer Research, Christie Hospital, Withington, Manchester, M20 9BX, United Kingdom, ¶Applied Biosystems Inc., Lingley House, 120 Birchwood Boulevard, Warrington, Cheshire, WA3 7QH, United Kingdom, and ‖Mount Sinai Medical School, One Gustave L. Levy Place, New York, New York 10029

Received December 2, 2004, and in revised form, April 22, 2005
Published, MCP Papers in Press, April 22, 2005, DOI 10.1074/mcp.M400193-MCP200

© 2005 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

the complexity of the MS spectra and necessitates that MS/MS is therefore performed on the same peptide (the heavy and light labeled versions) twice, wasting analysis time.

Novel labeling reagents can overcome some of the limitations described above. Isobaric tags for relative and absolute quantitation (iTRAQ)[1] use reagents that enable up to four samples to be analyzed within the same experiment. The labels consist of a protein-reactive group that labels all free amines (*i.e.* will label at the N terminus of all peptides and also the side chain of internal lysine residues), a balance group and a reporter group (6). The labels are isobaric, with a different distribution of isotopes between the reporter and balance groups. Hence, each labeled peptide appears at the same mass in an MS scan, but upon fragmentation in the mass spectrometer, the label dissociates and releases the reporter group as a singly charged ion of masses 114.1, 115.1, 116.1, or 117.1, respectively. Relative peak area indicates the contribution of each sample to total peptide present, providing a measure of relative abundance. The balance group is also lost and the remaining peptide fragments, which all have addition of the same mass (*i.e.* the protein reactive group) provide data from which to infer the peptide sequence.

The t(5,12) translocation found in chronic myelomonocytic leukemia results in the expression of the leukemogenic tyrosine kinase TEL/PDGFRβ and activation of the PDGFRβ tyrosine kinase domain (7). This stem cell disease has been modeled by expressing the TEL/PDGFRβ in a multipotent hematopoietic stem cell line, FDCP-Mix. The effects of oncogenic expression can be subtle yet lead to profound changes in cellular development. Perhaps unlike signaling for proliferation or apoptotic suppression, the appropriate tools for immediate analysis of potential effectors of altered development are not freely available. Herein, we report the validation and use of iTRAQ reagents on the FDCP-Mix TEL/PDGFRβ system as a paradigm for rapid, systematic definition of oncogenic processes using proteomics and the value of iTRAQ in permitting direct comparison of transcriptome data.

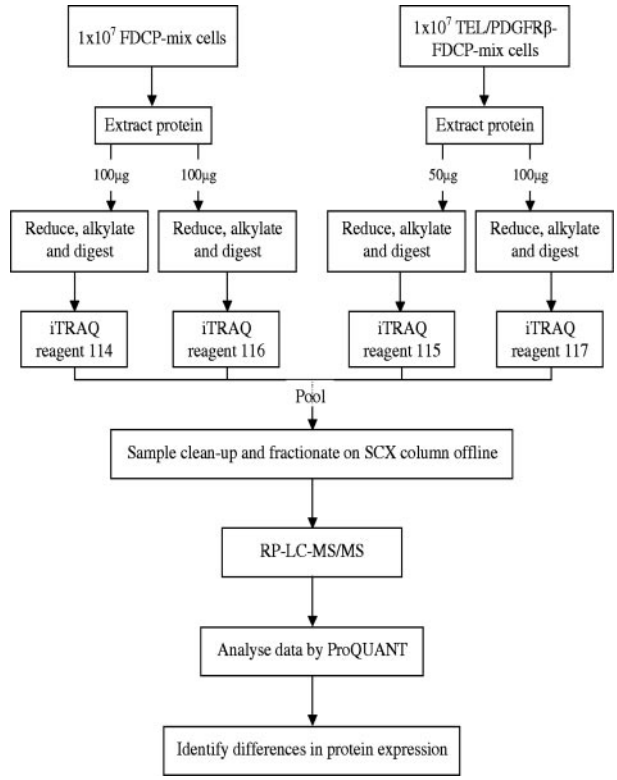


FIG. 1. **Workflow for the identification of changes in the proteome induced by TEL/PDGFRβ in the FDCP-Mix cell line, using iTRAQ reagents.** Three LC-MS/MS runs were performed to this protocol. To determine reproducibility of both the labeling reaction and the MS analysis internal controls were included. Confirmation of relative quantitation was thereby derived. The expression between FDCP-Mix and FDCP-Mix-Tel/PDGFRβ cells was compared twice in each MS analysis (114 *versus* 117 and 116 *versus* 117). The strong cation exchange step shown was employed to remove free iTRAQ reagent as well as to fractionate peptides for separate analyses by reversed-phase LC/MS/MS. PROQUANT is a program designed to integrate data from these isobaric tag experiments for relative protein quantification.

EXPERIMENTAL PROCEDURES

Cell Line Preparation and Culture—FDCP-Mix cells were transduced with TEL/PDGFRβ using a murine stem cell retroviral vector as described previously (8). Cells were routinely cultured in Fishers medium with 20% (v/v) horse serum supplemented with 10 ng/ml IL-3 (R&D Systems, Minneapolis, MN). Conditions for inducing myeloid differentiation FDCP-Mix cells were induced to differentiate by two methods as described previously (9). Briefly cells were cultured in Iscove's modified Dulbecco's medium supplemented with pre-selected fetal calf serum (20% (v/v)) and a combination of recombinant murine granulocyte macrophage-colony-stimulating factor (50 units/ml; Biogen IDEC, Zug, Switzerland), recombinant human macrophage-colony-stimulating factor (5 ng/ml; Amgen Biologicals, Thousand Oaks, CA) plus recombinant murine IL-3 (0.1 ng/ml; Calbiochem, Nottingham, UK). Cells were prepared with a Cytospin centrifuge and stained with May Grunwald-Giemsa, and differential morphology was scored for greater than 100 cells per slide.

Western Blotting—Western blotting was carried out with standard protocols using a monoclonal antibody to the kinase domain of PDGFRβ or anti-phosphotyrosine antibodies (BD PharMingen, Oxford, UK). The phosphoprotein content (serine, threonine, and tyrosine) was measured by separation of total cell lysates on 10%T SDS-polyacrylamide gels and staining with Pro-Q diamond stain (Molecular Probes, Leiden, The Netherlands) per the manufacturer's instructions.

iTRAQ Reagent Labeling—An overview of the workflow is shown in Fig. 1. 1–2 $\times$ 10^7 cells were lysed in 250 μl of 0.5 M triethylammonium bicarbonate (Sigma-Aldrich, St. Louis, MO) + 0.05% (w/v) SDS on ice for 20 min with regular vortexing. Protein was centrifuged at $>$10,000 $\times$ *g* for 20 min at 4 °C, supernatant was removed, and protein quantified using the modified Bradford protein assay (Bio-Rad Laboratories). Protein (50 or 100 μg) in 20 μl of 0.5 M triethylammonium bicarbonate/0.05% SDS was reduced by addition of 2 μl of 50 mM tris-(2-carboxyethyl)phosphine and incubation at 60 °C for 1 h. Reduced cysteine residues were then blocked by addition of 1 μl of

[1] The abbreviations used are: iTRAQ, isobaric tags for relative and absolute quantitation; PDGFRβ, platelet-derived growth factor receptor subunit β; C/EBP, CCAAT/enhancer-binding protein.

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

200 mM methylmethanethiosulfate in isopropanol and incubation at room temperature for a further 10 min. Protein was then digested by addition of 10 μl of trypsin at 0.5 μg/μl and incubated at 37 °C overnight.

To label the peptides with iTRAQ reagent (Applied Biosystems, Warrington, UK), one unit of label (defined as the amount required to label 100 μg of protein) was thawed and reconstituted in 70 μl of ethanol, with vortexing for 1 min. The reagent solution was added to the digest and incubated at room temperature for 1 h. Labeling reactions were then pooled before analysis.

Peptide Fractionation and Mass Spectrometry—To remove excess, unbound iTRAQ reagent and to simplify the peptide mixture before reversed-phase LC-MS/MS, peptides were washed and fractionated off line using a strong cation exchange column (Applied Biosystems). In brief, the peptide mixture was diluted 10-fold in loading buffer (10 mM potassium phosphate in 25% (v/v) acetonitrile, pH 3.0), and the pH was checked to ensure it remained between 2.5 and 3.3. Sample mixture was slowly injected onto the strong cation exchange cartridge and was washed with a further 1 ml of loading buffer to remove salts, tris-(2-carboxyethyl)phosphine, and unincorporated iTRAQ reagent. Samples were eluted from the column using 500-μl volumes of elution buffer (10 mM potassium phosphate in 25% (v/v) acetonitrile) containing increasing concentration of KCl. Salt concentrations used were 50, 100, 150, 200, 250, 300, 350, and 500 mM. Each salt fraction was then concentrated and dried in a SpeedVac (Thermo Electron, Waltham, MA).

Dried peptide fractions were resuspended in 250 μl of 2% (v/v) acetonitrile/0.1% (v/v) formic acid. For each analysis, 60 μl of the peptide sample was loaded onto a 15-cm reversed phase C18 column (75 μm i.d.) using an UltiMate pump (LC Packings, Amsterdam, The Netherlands) and separated over a 120-min solvent gradient from 5.9% (v/v) acetonitrile/0.1% (v/v) formic acid to 41% (v/v) acetonitrile/0.1% (v/v) formic acid on-line to a QSTAR XL mass spectrometer (Applied Biosystems). Data was acquired using an independent data acquisition protocol in which, for each cycle, the two most abundant multiply charged peptides (2^+ to 4^+) above a 10 count threshold in the MS scan with *m/z* between 480 and 2000 were selected for MS/MS. Each peptide was selected twice and then dynamically excluded (±50 milli-mass units) for 40 s.

Data Analysis—Data were searched against a mouse KBMS3.0 protein database from the Celera Discovery System (Applied Biosystems). The database allowed for iTRAQ reagent labels at N-terminal residues, internal K and Y residues, and the methylmethanethiosulfate-labeled cysteine as fixed modification, plus one missed cleavage. Search parameters within ProQUANT were set with an MS tolerance of 0.15 Da, an MS/MS tolerance of 0.1 Da, and a minimum confidence score of 20. ProQUANT pooled data from all LC-MS runs. Assessment of these parameters for peptide and protein identification is described in Supplemental Table I.

Transcriptome Analysis—RNA from FDCP-Mix and TEL/PDGFRβ FDCP-Mix samples was prepared using TRIzol (Invitrogen) in triplicate samples and then cleaned using MinElute RNeasy Clean up kit (Qiagen, Valencia, CA) per the manufacturer's instructions. Transcriptome analysis was undertaken using murine MOE430A Affymetrix chips by the CR-UK Affymetrix microarray facility, Paterson Institute (Manchester, UK). -Fold changes were calculated from the scaled data, where appropriate *t*-test analysis was applied.

RESULTS

Expression of TEL/PDGFRβ in the Multipotent Hematopoietic Cell Line FDCP-Mix Inhibits Differentiation but Has No Effect on Growth Factor Dependence—FDCP-Mix cells were transduced to express TEL/PDGFRβ (Fig. 2*a*). Previous experience with Ba/F3 cells transfected with BCR/ABL and TEL/PDGFRβ have shown that TEL/PDGFRβ has comparatively small effects on tyrosine phosphorylation (data not shown). However, TEL/PDGFRβ has small but significant effects on protein tyrosine phosphorylation and also significantly affects total phosphoprotein level (Fig. 2, *b* and *c*).

Differentiation-blocked cell lines can become growth factor-independent when expressing TEL/PDGFRβ (10). This did not occur in TEL/PDGFRβ-FDCP-Mix cells. FDCP-Mix cells differentiate to form mature cells when cultured in the appropriate cytokines (11). Expression of TEL/PDGFRβ inhibited this development (Fig. 2*d*). Culture conditions that induce myeloid development to 100% postmitotic cells (macrophages, neutrophils) in control FDCP-mix cells gave no such effect in TEL/PDGFRβ-transfected cells. After 20 days in culture, colony-forming clonogenic cells were still present that had primitive myeloid cell or blast cell morphology. This effect on differentiation led us to systematically analyze potential differences in transfected and non-transfected cells using transcriptomic and proteomic methods. To do this, the proteomic method required validation.

Analysis of a Standard Mix of Proteins Using iTRAQ Reagent Labels—A defined six-protein mix that had been enzymatically digested with trypsin was used to confirm the accuracy of ratiometric quantitation of the iTRAQ reagents. The tryptic digest was halved, and each half was labeled with either reagent 116 or 117. These differentially labeled digests were mixed at various ratios (1:1, 2:1, and 1:3) and analyzed by LC-MS/MS. A representative spectrum is shown in Fig. 3*a*. Relative quantitation of proteins by iTRAQ reagent technology was both accurate and reproducible for five proteins (Fig. 3*b*). The sixth protein in the standard mixture was not detected with sufficient peptides to allow quantitation. Overall 117:116 ratios of 0.9699, 0.5885, and 3.1748 were obtained. We confirmed that no peptides remained "unlabeled" by analysis of parent ion masses derived for the MS analysis and comparison with theoretical tryptic digests of the six proteins. All isobaric forms of the iTRAQ reagent tag labeled equally efficiently.

Identification of TEL/PDGFRβ Induced Alterations in the Proteome—The potential value of the system was further examined on lysates from mammalian FDCP-Mix cells and FDCP-Mix cells transduced with the TEL/PDGFRβ leukemic oncogene. In these experiments, microarray data was available for comparison from the FDCP-Mix cells described above. The experimental design is outlined in Fig. 1. All four labels were used, allowing the use of internal controls and replicates within the same LC-MS experiment for this paradigm study. This experimental design allowed the accuracy of quantitation to be verified by including a 1:1 ratio of control sample (114:116) and a 2:1 ratio of the TEL/PDGFRβ FDCP-Mix cell (117:115), as shown in Fig. 1. Furthermore, an internal replicate for the comparison of TEL/PDGFRβ-expressing and non-expressing cells is thus provided. The experiment was

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

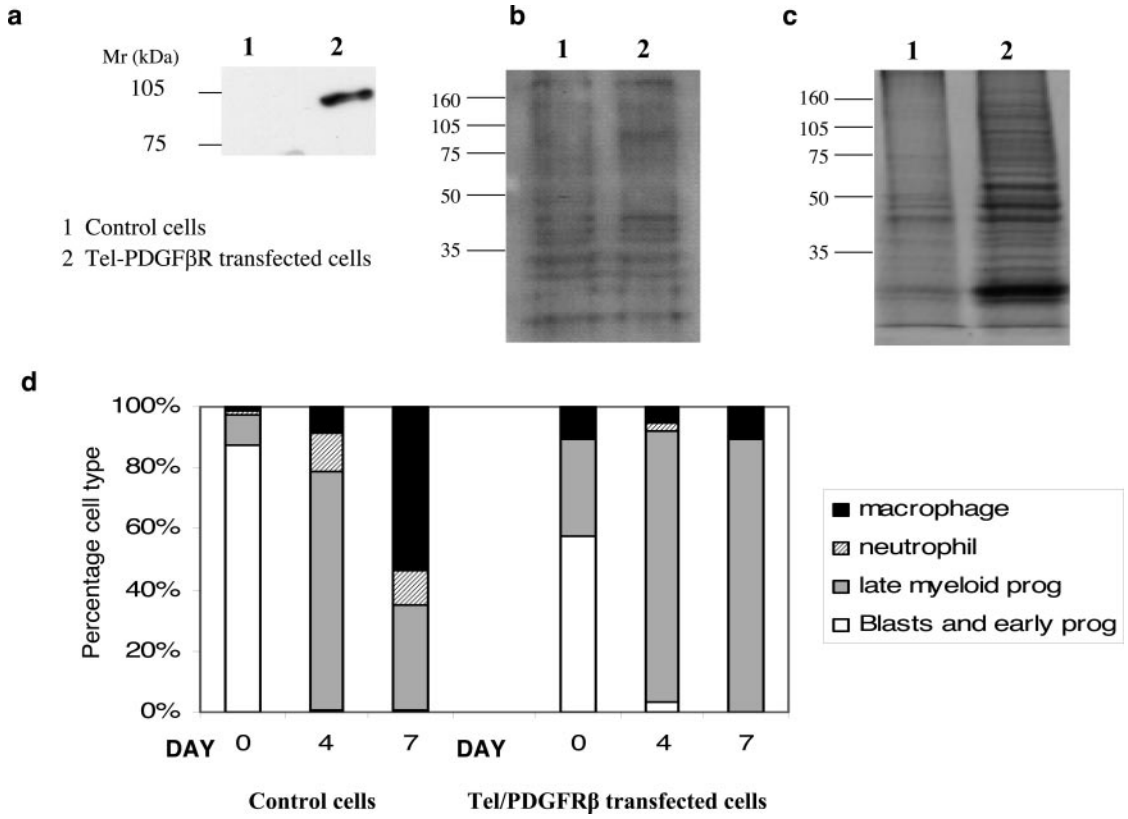


FIG. 2. **Consequences of expression of TEL/PDGFβR in FDCP-Mix cells.** Expression and consequence of TEL/PDGFβR in transfected cells as analyzed by Western blot analysis of total cell lysates using anti-PDGFβR (*a*) and anti-phosphotyrosine (*b*) antibodies. *c*, level of total protein phosphorylation as assessed by staining whole-cell lysate with Pro-Q Diamond. In each of the three gel images, *lane 1* indicates control cells and *lane 2* indicates TEL/PDGFRβ-transfected cells. *d*, effect of TEL/PDGFβR on differentiation as assessed by culturing control and TEL/PDGFβR-transfected FDCP-Mix cells in 0.1 ng/ml IL-3 and 50 units/ml granulocyte macrophage-colony-stimulating factor plus 5 ng/ml macrophage-colony-stimulating factor. Cells were seeded at 1×10^5/ml, and morphology was determined on days 0, 4, and 7. Results shown are the mean of three experiments; S.E. was <10% in all cases. Late progenitor cells indicate metamyelocytes, myelocytes, and promonocytes.

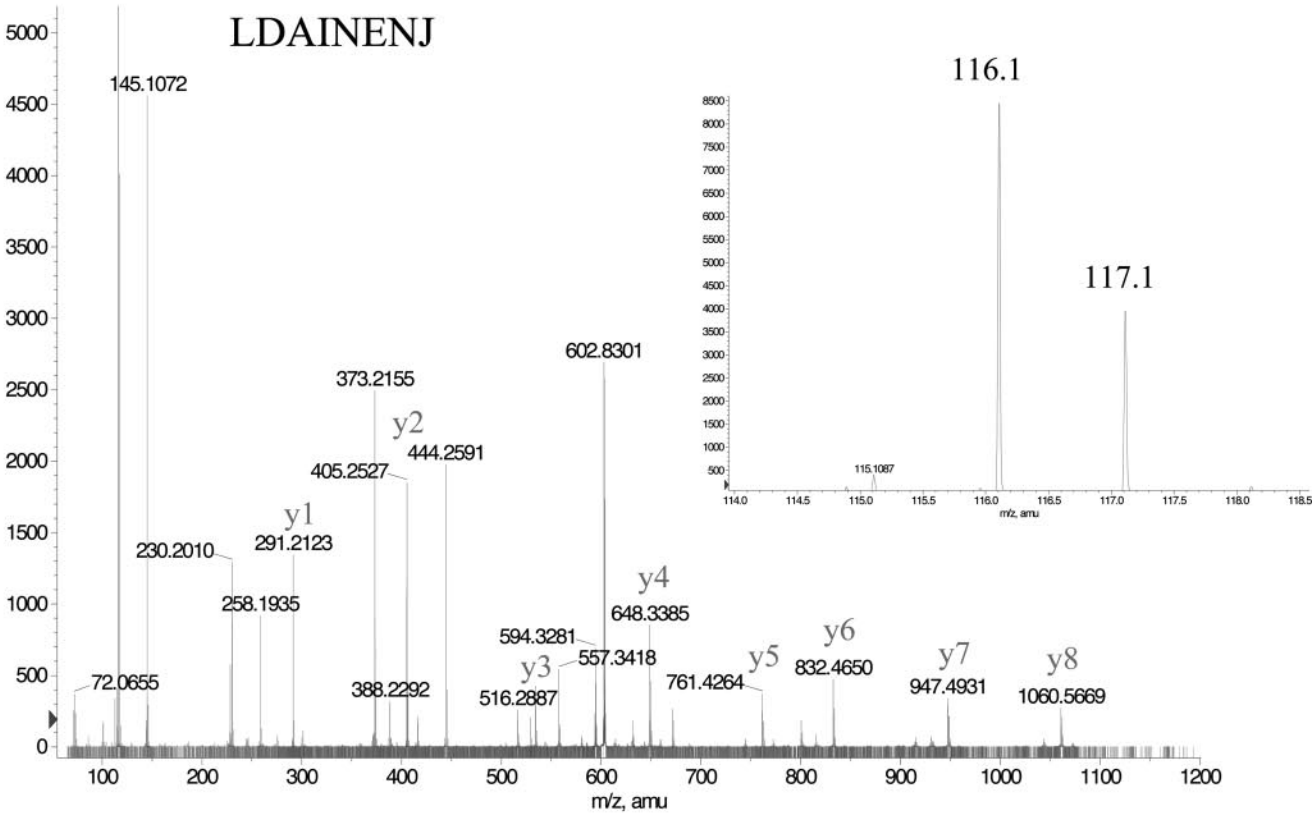
run in triplicate. Labeling efficiency with iTRAQ reagents was high (>99%), and both intra- and interexperiment quantitation was highly reproducible, with overall ratios of 1.064, 0.967, and 1.093 when samples were mixed 1:1 and 2.127, 2.128, and 2.217 when samples were mixed 2:1. It is noteworthy that ratios detected were normally distributed; therefore, statistical tests such as Student *t* test can be applied. In addition, replication within the experiment achieved an acceptable standard, with an average 116:114 ratio of 1.072 32 ± 0.07 (S.D.; $n = 3$) (Fig. 4*a*), and the replicate FDCP-Mix *versus* FDCP-Mix-TEL/PDGFRβ analyses (117:114 *versus* 117:116) had a correlation coefficient of 0.93, confirming the reproducibility of the technique. Data for individual peptides and proteins was reproducible across replicate experiments (Fig. 4*b*). In a typical experiment, 1120 peptides from 347 proteins were found. Analysis of the ratios of identified proteins showed significant and reproducible differences between samples in a total of 13 proteins (Fig. 4*c* and Table I). It is noteworthy that such a plot shows an extremely tight clustering around the 1:1 ratio value, indicative of few changes and a robust, reproducible technique. Examples of the spectra produced by iTRAQ reagent-labeled peptides are shown in Fig. 5, *a–c*, where peptides showing no change between samples, down-regulation or up-regulation by TEL/PDGFRβ expression are illustrated.

Comparison of iTRAQ™ Reagent Data to cDNA Microarray—Relatively few changes are seen in cDNA microarrays from TEL/PDGFRβ-transfected and control FDCP-Mix cells, reflecting the relatively small effect of this oncogene on protein tyrosine phosphorylation. This paucity of change in the transcriptome led us to confirm changes using proteomics. Analysis of microarray data revealed that, with a 1.5-fold change cut-off, a total of 105 transcripts show a significant increase with TEL/PDGFRβ expression, and 159 transcripts show decrease (see Supplemental Table I).

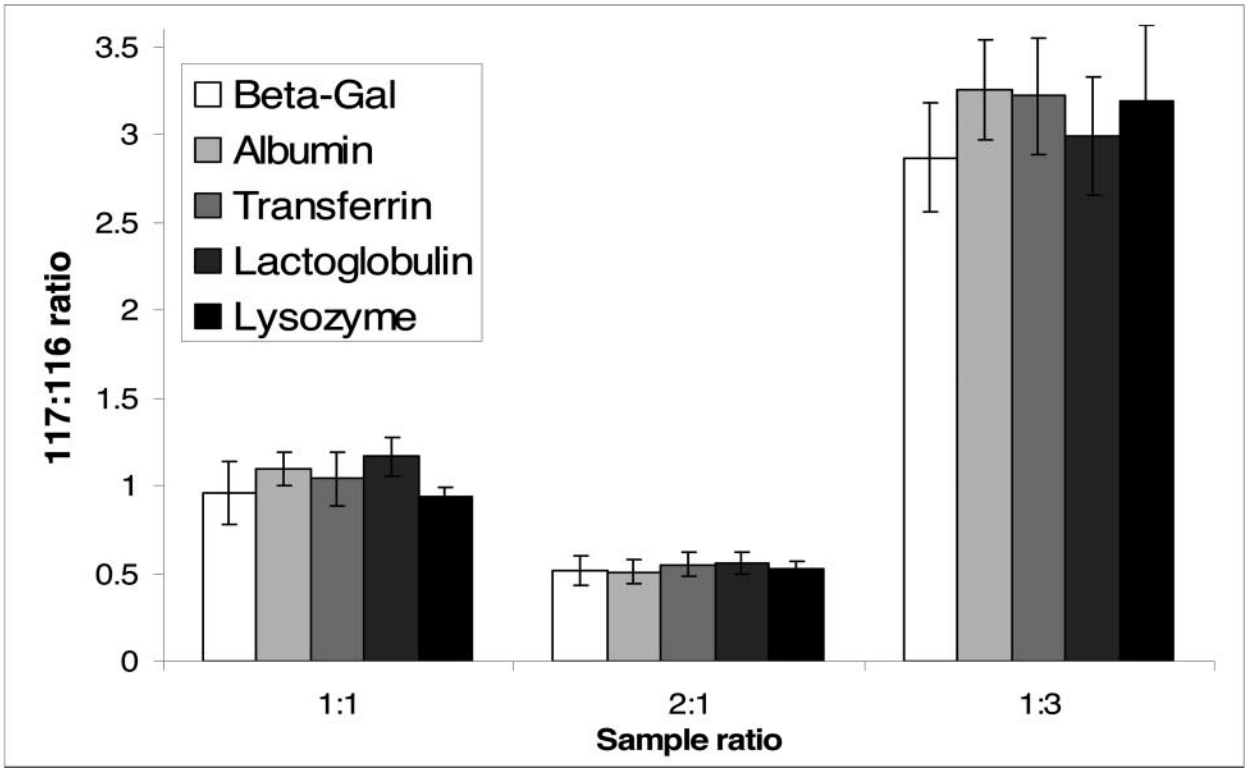
The data derived permitted comparison of individual transcript/protein levels. All of the differences identified by iTRAQ reagents (Table I), along with 87 unchanged proteins, were analyzed, giving a set of 100 proteins/transcripts in total as examples of the data recorded. The average ratio of the

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

a



b



Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

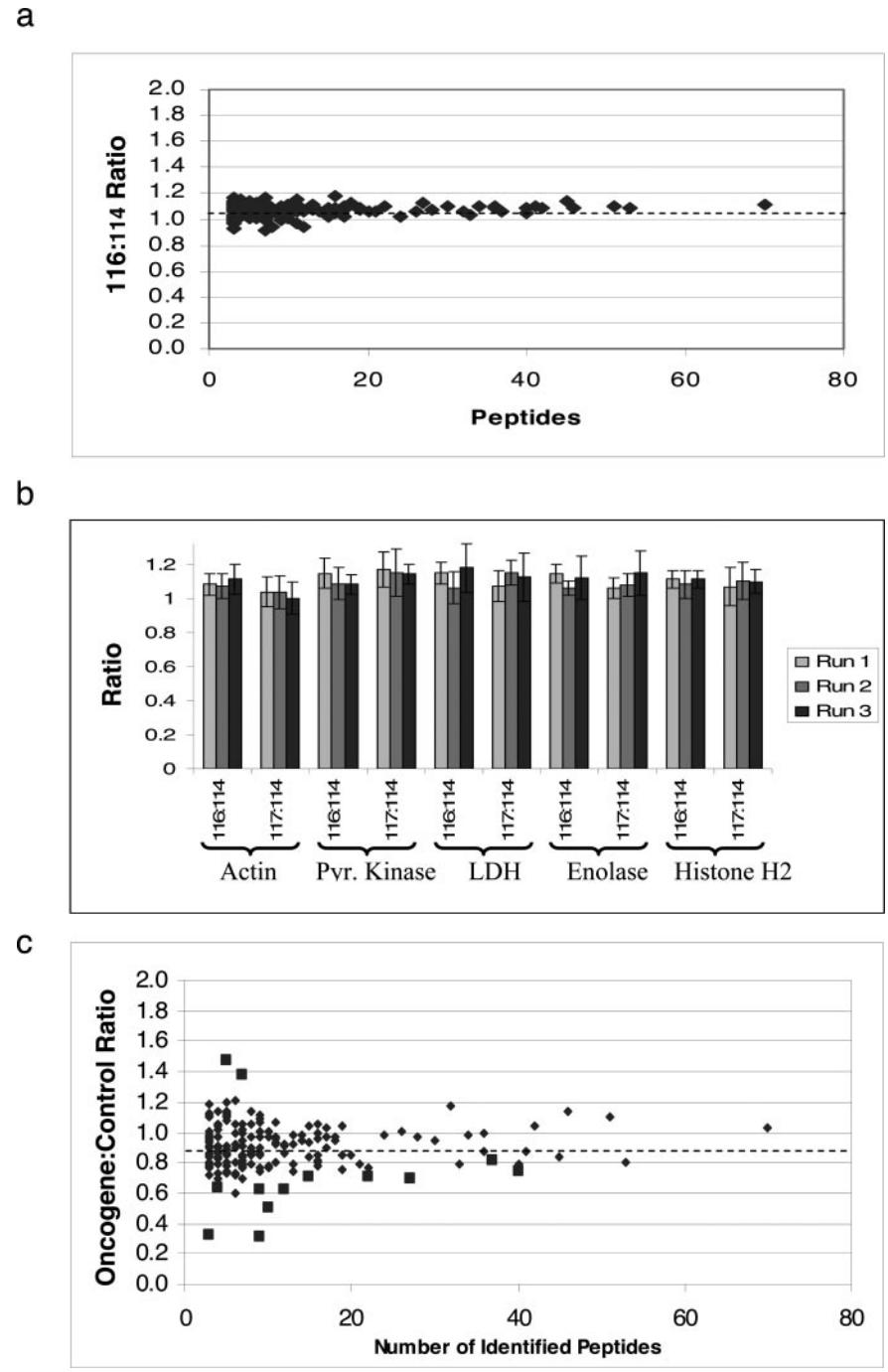


FIG. 4. **Ratiometric analysis of proteins from FDCP-mix and FDCP-mix-TEL/PDGFRβ cell lines using iTRAQ reagents.** *a*, plot to show the reproducibility of ratio determination in the internal control experiment comparing 100 μg of control protein labeled with either the 114 or 116 iTRAQ reagent labels. The *x*-axis indicates the number of peptides identified (using MS/MS) from each protein. Each *spot* represents one protein identified. *Dotted line* shows the average 116:114 ratio across the complete data set. *b*, reproducibility of the iTRAQ reagent-relative quantitation: the ratios for five representative proteins over three experiments using the FDCP-Mix and FDCP-Mix-TEL/PDGFRβ cell lines are shown. Error bars are defined as in Fig. 2*b*. *c*, plot showing the distribution of the oncogene/control (117:114) ratios of proteins identified in the iTRAQ reagent experiment, plotted against the number of peptides identified and therefore quantified from that protein. Unchanged proteins are shown by ♦, and differentially expressed proteins are shown by ■. *Dotted line* represents the mean FDCP-Mix TEL/PDGFRβ:FDCP-Mix ratio from the complete data set.

mRNA expression between FDCP-Mix and FDCP-Mix-TEL/PDGFRβ for each specific transcript was calculated and compared with the 117:114, 117:116, and 116:114 ratios (TEL/PDGFRβ/control 1, TEL/PDGFRβ/control 2, and control 1/control 2, respectively) from iTRAQ reagent experiments (Table I). The cDNA array data from these 100 proteins con-

FIG. 3. **Ratiometric analysis of five standard proteins labeled with iTRAQ reagents 116 or 117 and mixed at 1:1, 1:2, and 3:1 ratios.** *a*, a representative spectrum of peptide LDAINENJ (where *J* represents an internal lysine labeled with iTRAQ reagent via the ϵ-amino acid in its side chain) from β-lactoglobulin derived for a 2:1 mixture of 116- and 117-labeled samples. The peptide sequence is derived from the fragment ion pattern; *inset*, the relative abundances were determined by comparison of the areas under the 116.1 and 117.1 peaks. *b*, summary of the ratios for the five standard proteins analyzed. Error bars are S.D. of *n* observations, where *n* is equal to the number of peptides identified in that protein.

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

TABLE I

The relative quantification of protein levels and mRNA levels in FDCP-Mix cells and TEL/PDGFRβ-FDCP-Mix cells

For comparison, 100 proteins (randomly selected proteins plus those in which changes between the two cell lines were seen using the isobaric tag quantification approach) were selected. Random selection was made via search databases (generating the accession codes shown: emb, EMBL; gb, GenBank™; pir, Protein Information Resource; rf, RefSeq (NCBI); spt, Swiss-Prot; trm, TrEMBL.). The total peptides columns refers to the number of peptides from that protein fragmented within three replicate experiments. The relative ratios for each protein were calculated as a weighted value combining averages from three replicate experiments. Ratios shown are: average control/control, an internal control, in which the same amount of the same lysate has been compared (ratio of 116:114, see Fig. 2*b*); average (1) oncogene/control, the ratio of protein in the TEL/PDGFRβ-transfected cells *vs.* control FDCP-Mix cells containing data from comparison of 117:114 iTRAQ label relative quantities; average (2) oncogene/control, containing data from comparison of 117:116 ratios. Also shown is the fold change in the mRNA level (-fold cDNA) for that transcript/protein (oncogene/control) according to microarray analysis. Sequences whose expression is significantly altered ($p < 0.05$) at the mRNA level are shown in italics.

Accession	Protein	Total peptides	Average control/control[a]	Average (1) oncogene/control[a]	Average (2) oncogene/control[a]	-Fold cDNA
pir | S31975	14-3-3 protein ϵ	7	1.115	1.061	0.981	1.034
rf | NP_853613.1	14-3-3 protein γ	8	1.093	1.072	1.006	1.042
spt | P63101	14-3-3 protein ζ	8	1.111	1.043	0.968	1.106
spt | P47911	60S ribosomal protein L6	10	1.059	0.785	0.745	0.995
rf | NP_031501.1	Acidic ribosomal phosphoprotein P0	7	1.065	0.789	0.759	1.028
spt | P51881	ADP, ATP carrier protein, fibroblast isoform	18	1.108	0.999	0.919	0.940
spt | Q60604	Adseverin	36	1.066	1.000	0.974	1.040
spt | P47738	Aldehyde dehydrogenase, mitochondrial	5	1.059	1.074	1.034	0.926
spt | P56480	ATP synthase β chain, mitochondrial	10	1.123	1.238	1.127	1.058
Emb | CAA27396.1	β-Actin	64	1.123	1.061	0.956	0.988
spt | Q9CWJ9	Bifunctional purine biosynthesis protein	7	1.079	1.691	1.661	
Emb | CAA05361.1	Bip	13	1.138	0.976	0.885	0.835
gb | AAA62450.1	Calnexin	15	1.012	1.010	1.021	0.984
Spt | P14211	Calreticulin (CRP55)	41	1.109	0.939	0.867	0.963
Spt | P97742	Carnitine *O*-palmitoyltransferase	4	1.167	1.024	0.915	1.199
pir | HHMS60	Chaperonin groEL	13	0.892	1.538	1.757	0.913
rf | NP_031662.1	Chaperonin subunit 2 (β)	9	1.032	0.746	0.749	1.078
spt | P18760	Cofilin, non-muscle isoform	5	1.081	1.077	1.015	1.008
prf | 1513495A	Cu/Zn superoxide dismutase	10	1.110	1.056	0.973	0.996
rf | NP_031834.1	Cytochrome *c*, somatic	9	1.088	0.887	0.829	0.885
spt | P49717	DNA replication licensing factor MCM4	5	1.075	0.878	0.827	1.234
Trm | Q8BMR3	Dolichyl-diphosphooligosaccharide	6	1.126	0.917	0.821	1.085
sp | O70251	Elongation factor 1-β	4	1.038	0.952	0.935	0.957
spt | P10126	Elongation factor 1-α1	64	1.079	0.797	0.754	1.080
spt | P08113	Endoplasmin precursor (GRP94)	17	0.999	1.106	1.176	0.876
rf | NP_075608.1	Enolase 1, α non-neuron	28	1.093	1.159	1.073	1.036
rf | NP_031933.1	Eukaryotic translation elongation factor 2	27	1.117	0.771	0.705	0.977
spt | P26040	Ezrin	12	1.081	0.969	0.917	0.976
Trm | Q9EQR0	Fatty acid synthase (Fasn protein)	6	1.022	0.821	0.824	1.003
spt | P35550	Fibrillarin (Nucleolar protein 1)	8	1.050	0.933	0.900	1.030
rf | XP_127565.4	Filamin B, β	22	1.052	0.862	0.836	
rf | XP_207130.3	Filamin, α	10	1.002	0.908	0.953	
spt | P05064	Fructose-bisphosphate aldolase A	28	1.054	1.260	1.224	1.041
rf | NP_032181.1	Glucose phosphate isomerase 1	13	1.052	1.037	1.018	1.033
spt | P26443	Glutamate dehydrogenase, mitochondrial	7	1.114	1.055	0.988	0.986
rf | NP_032110.1	Glyceraldehyde-3-phosphate dehydrogenase	29	1.060	1.126	1.087	0.987
spt | P04187	Granzyme B(G,H)	26	1.111	1.194	1.118	0.987
rf | NP_034566.1	H2A histone family, member X	27	1.089	0.817	0.767	1.189
pir | JQ1983	H3.3 like histone MH921	8	1.066	0.799	0.758	0.997
rf | NP_034610.1	Heat-shock protein 1α	36	1.051	0.835	0.810	0.948
Emb | CAA56631.1	High mobility group protein	16	1.054	1.146	1.107	1.040
rf | NP_085112.1	Histone 1, H1a	28	1.029	1.243	1.267	0.880
rf | NP_056601.1	Histone 1, H1c	19	1.112	0.949	0.875	0.827
rf | NP_835503.1	Histone 1, h2bg	32	1.062	1.261	1.217	1.009
pir | S45110	Histone H2A	45	1.142	0.846	0.762	1.138
gb | AAH58529.1	Histone H4	30	1.111	0.928	0.856	1.095
spt | Q99020	hnRNP A/B	8	1.049	1.123	1.097	0.985
gb | AAH06694.1	hnRNP K protein	11	1.179	1.077	0.927	1.129

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

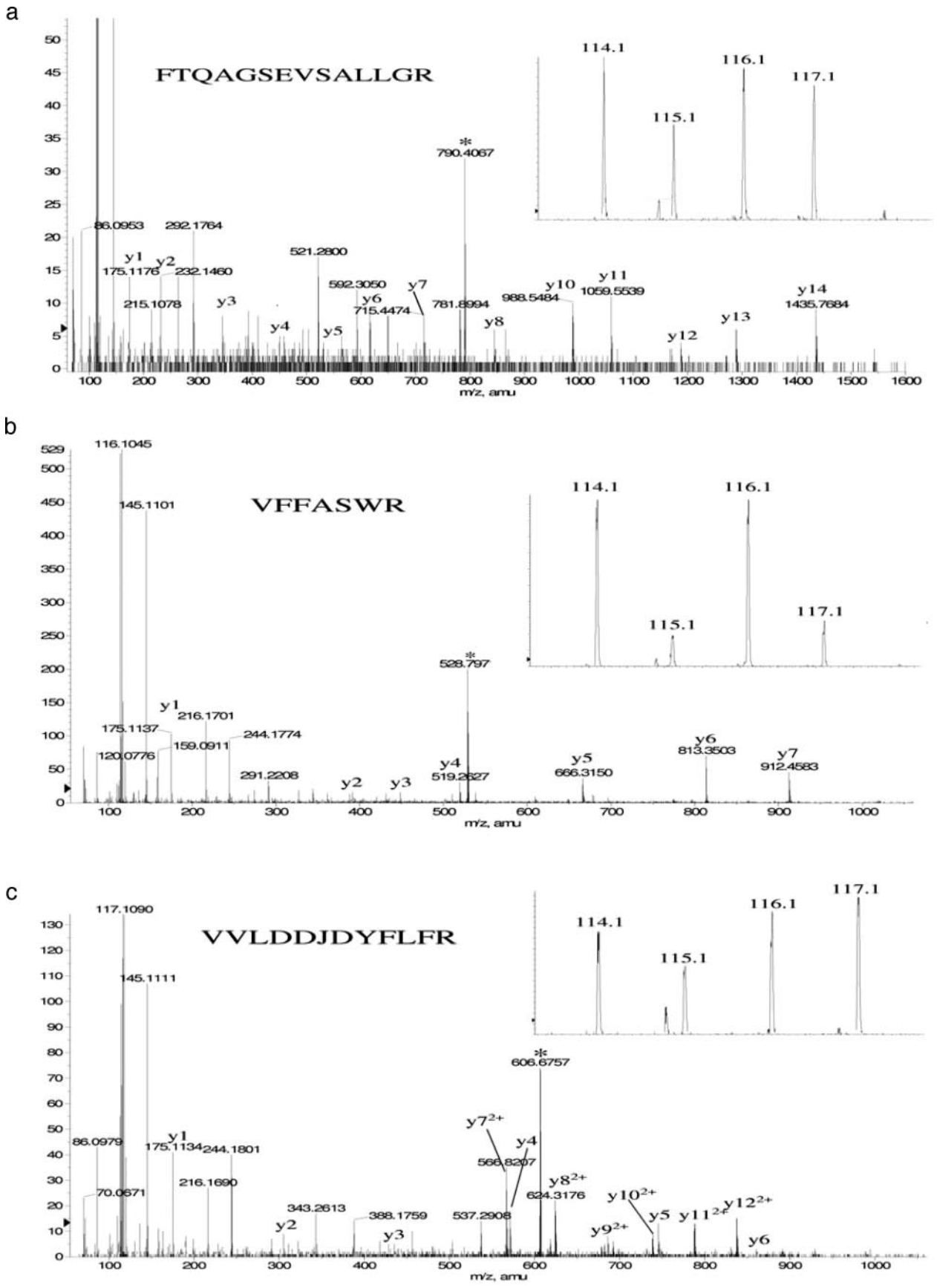
TABLE I—*continued*

Accession	Protein	Total peptides	Average control/ control	Average (1) oncogene/ control	Average (2) oncogene/ control	-Fold cDNA
rf | NP_034829.1	Lactate dehydrogenase 1, A chain	36	1.124	1.040	0.941	1.021
pir | JN0066	Leukotriene-A4 hydrolase	8	1.073	0.868	0.827	0.878
emb | CAA65761.1	M2-type pyruvate kinase	51	1.104	1.137	1.055	1.071
gb | AAA39509.1	Malate dehydrogenase	9	1.036	1.000	0.978	0.945
rf | NP_620084.1	Methylenetetrahydrofolate dehydrogenase 1	5	1.045	0.954	0.932	1.022
rf | NP_034963.1	Moesin	14	1.080	0.975	0.923	0.809
trm | Q9EQK1	NADP+-specific isocitrate dehydrogenase	11	1.136	0.930	0.851	0.838
trm | Q8VDD5	Nonmuscle heavy chain myosin II-A	24	1.063	1.004	0.957	1.069
sp | P09405	Nucleolin (protein C23)	17	1.093	0.892	0.829	0.968
trm | Q8VDW0	Nuclear RNA helicase, DECD variant of DEAD box family	13	1.088	1.085	1.021	0.918
rf | NP_077155.1	Nucleolar protein 5A	4	1.053	1.248	1.222	0.952
spt | Q61937	Nucleophosmin	13	1.099	0.799	0.805	0.905
pir | A46557	Nucleoside-diphosphate kinase	9	1.103	0.962	0.895	0.963
spt | Q01768	Nucleoside diphosphate kinase B	7	1.129	0.906	0.830	0.920
spt | P24369	Peptidyl-prolyl cis-trans isomerase B	6	1.110	1.330	1.199	1.015
rf | NP_032933.1	Peptidylprolyl isomerase A	10	1.119	1.114	1.002	1.051
spt | Q99K85	Phosphoserine aminotransferase	5	1.051	1.117	1.113	1.103
rf | NP_035202.1	Profilin 1	9	1.055	1.046	1.004	1.114
spt | P27773	Protein disulfide isomerase A3	11	0.984	0.954	0.986	0.931
rf | NP_033417.1	RAN, member RAS oncogene family	14	1.067	1.216	1.180	0.983
spt | Q61599	Rho GDP-dissociation inhibitor 2	4	1.110	1.026	0.941	0.879
rf | NP_038749.1	Ribosomal protein L7a	8	1.005	0.937	0.925	1.072
rf | NP_079862.1	Ribosomal protein L15	8	0.951	0.862	0.921	1.083
rf | NP_075029.1	Ribosomal protein L23	18	1.119	0.928	0.859	0.998
rf | NP_035430.1	Ribosomal protein S7	6	1.105	0.929	0.849	1.052
rf | NP_084043.1	Ribosomal protein S9-like	7	1.123	0.929	0.849	1.003
rf | NP_080809.1	Ribosomal protein S13	9	1.120	0.903	0.828	0.987
gb | AAA16796.1	Ribosomal protein S18	7	1.070	0.801	0.763	1.092
trm | Q7TNQ6	Ribosomal protein, large P2	6	1.071	0.709	0.684	1.015
spt | Q08943	Structure-specific recognition protein 1	6	1.110	1.104	0.921	1.086
rf | NP_063932.1	Succinate-CoA ligase, GDP-forming, α	7	1.071	1.117	1.066	0.901
gb | AAA40338.1	T complex polypeptide 1	7	0.972	0.905	0.970	1.219
spt | P26039	Talin 1	12	0.976	1.076	0.986	1.029
spt | Q9WVA4	Transgelin 2	10	1.047	1.006	0.982	1.089
spt | P40142	Transketolase	16	1.074	0.835	0.790	0.922
pir | JL0149	Transplantation antigen P198	5	1.043	0.839	0.816	1.089
pir | ISMST	Triosephosphate isomerase	11	1.055	1.310	1.284	1.160
rf | NP_666228.1	Tubulin, β, 2	54	1.109	0.642	0.602	1.081
gb | AAH43053.1	Valosin containing protein	15	1.068	1.019	0.981	0.985
Differentially expressed proteins as determined by iTRAQ						
spt | Q64433	10-kDa heat shock protein	5	1.079	1.513***	1.402***	0.989
spt | P45634	60S ribosomal protein L10	12	1.072	0.624*	0.593*	1.013
gb | AAH21937.1	Aldo-keto reductase family 1, C13	22	1.105	0.770*	0.798*	1.072
spt | P28293	*Cathepsin G*	*9*	*1.059*	*0.617****	*0.621****	*0.532*
pir | A45935	DnaK-type molecular chaperone hsc70	37	1.057	0.813**	0.794***	0.892
rf | NP_853613.1	Eukaryotic translation initiation factor 5A	12	1.137	0.498***	0.448***	0.961
gb | AAA37866.1	84-kDa heat-shock protein	34	1.087	0.712***	0.668***	0.955
rf | NP_031542.1	Heterogeneous nuclear ribonucleoprotein D	7	1.068	1.382**	1.337*	0.996
spt | Q61233	*L-plastin*	*40*	*1.081*	*0.748**	*0.702***	*0.745*
spt | P43430	*Mast cell protease 8*	*3*	*1.062*	*0.318**	*0.312***	*0.352*
Trm | Q7TMS4	*Myeloperoxidase protein*	*10*	*1.035*	*0.307****	*0.302****	*0.219*
Spt | P09103	*Protein disulfide isomerase*	*27*	*1.122*	*0.696****	*0.631****	*0.763*
spt | Q60931	Voltage-dependent anion-selective channel protein 3	5	1.058	0.638**	0.629**	1.146

[a] Statistical significance of the change as determined by Student *t* test is depicted (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

tained five differences. These were cathepsin G, L-plastin, Mast cell protease 8, myeloperoxidase, and protein disulfide isomerase. Regarding the iTRAQ experiments, all five transcriptomic differences were mirrored in the protein changes seen. It is noteworthy that, with respect to developing studies on mechanisms of transformation, a set of proteins showed

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019



Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

changes (subtle, albeit statistically significant) in the proteome but not in the transcriptome. These included a 60S ribosomal protein subunit, aldo-keto reductase family member C13, and eukaryotic translation initiation factor 5A. This initial scan of the value of the iTRAQ approach therefore reveals that it will offer advantages over transcriptome analysis. Furthermore, specific proteins identified using iTRAQ LC-MS/MS were not assayed within the microarray. These included Filamins A and B, and bifunctional purine biosynthesis protein PURH. This assessment was confirmed after searching using alternative protein names, gene names, and accession numbers. iTRAQ therefore offers objective, non-selective sample analysis.

In the five changes showing consistency at transcript and protein level (cathepsin G, mast cell protease 8, myeloperoxidase precursor, protein disulfide isomerase, and L-plastin) the -fold changes in this set were remarkably similar to the -fold changes detected in their mRNA levels. The iTRAQ reagent and cDNA microarray data sets also agree in that most of the changes identified are subtle (2-fold or less).

DISCUSSION

iTRAQ reagent technology is a newly developed method for relative quantification of proteins from up to four samples. It has immense potential to improve the sensitivity and quality of mass spectrometric analysis of the proteome. Validation of the approach is reported here using a defined protein mixture and cell lysates from a disease model. iTRAQ reagent labels all peptides at their N terminus, along with free amines in lysine side chains, hence all of the peptide population is labeled, allowing more peptides to be quantified from each protein and increasing the quality of the data obtained by this approach. We have demonstrated that it is possible, with both standard proteins and whole-cell lysates, to label four peptide mixtures to completion and, using LC-MS/MS, to identify the relative quantities of the peptide emanating from four samples. The labeling protocol for iTRAQ reagents is simple, with few steps, so there are few opportunities for sample loss or contamination. Another advantage of iTRAQ reagent approach is that the peptides from all samples appear as one peak in MS increasing the total ion current for that peptide, simplifying the spectra, and requiring only one MS/MS experiment per peptide.

Up to 347 proteins were identified in a typical first-pass analysis of a whole-cell lysate. 400 μg of protein in a DIGE experiment yielded approximately 100 confirmed protein identities using the same mass spectrometer and LC parameters for peptide sequencing.[2] However, the number of identified proteins in our experiments is lower than that reported for other LC-MS/MS approaches. Overcoming this relative lack of sensitivity is straightforward. Standard protein identification protocols are designed to select the precursor ion with the greatest number of counts in each MS scan for MS/MS. A deeper penetration into the proteome, gaining quantification on lower abundance proteins is achievable by a combination of pre-enrichment strategies (for examples of organelles, see Dreger (12)) and/or the generation of "exclude lists," which instruct the mass spectrometer to ignore specific "high abundance" ions (defined by both mass and chromatographic retention time). On the other hand, the MS/MS experiment can be designed such that the ion(s) selected for MS/MS are the lowest two (above a predetermined threshold) in the MS spectra, rather than the highest. Preliminary investigation of this "bottom-up" selection protocol on one of the FDCP-Mix samples increased the number of proteins identified by a further 50% (data not shown). Because only a fraction of the sample is used in a single LC-MS/MS run, using several thresholds/sample loads should maximize the number of peptide/protein identifications from each sample. Therefore, global proteomic analysis or more focused procedures can be achieved using iTRAQ reagents and multiple subcellular fractions, multidimensional chromatography, plus ion exclusion lists in iTRAQ reagent-based experiments.

Several proteins are differentially expressed in the presence of TEL/PDGFRβ. Heterogeneous nuclear ribonucleoprotein D, shown to be increased by TEL/PDGFRβ expression, is an mRNA binding protein that has been implicated in tumorigenesis (13) and that is a target of the leukemogenic oncogene BCR/ABL (14). In our study TEL/PDGFRβ expression decreased Hsc70 levels, whereas previous studies have shown that increased expression of Hsc70 inhibits transformation (15). Myeloperoxidase protein, decreased by TEL/PDGFRβ, is expressed in early myeloid progenitors (16); therefore, its decreased levels in the differentiation-blocked TEL/PDGFRβ-expressing cells may provide clue as to the mechanism for this block. Myeloperoxidase expression is regulated by transcription factors such as Pu.1 and the C/EBP family (17). These proteins regulate myelopoiesis and loss of either Pu.1 or C/EBPα leads to compromised ability to produce mature

[2] A. D. Whetton and C. A. Evans, unpublished observations.

FIG. 5. **Identification of differentially expressed proteins.** *a*, representative MS/MS spectrum showing a peptide, FTQAGSEVSALLGR, from ATP synthase β that does not show an expression change upon TEL/PDGFRβ expression and the iTRAQ reagent peaks used for relative quantitation (*inset*). *b*, representative MS/MS spectrum showing a peptide, VFFASWR, from myeloperoxidase, a protein that shows a decrease in expression with TEL/PDGFRβ and the iTRAQ reagent peaks used for relative quantitation showing the decreased levels of the 115 and 117 labels compared with control protein labeled with 114 and 116 (*inset*). *c*, representative MS/MS spectrum showing a peptide VVLDDJDYFLFR (where *J* represents an internal lysine labeled with iTRAQ reagent via the ϵ-amino acid in its side chain) from 10-kDa heat-shock protein that shows an increase in expression with TEL/PDGFRβ and the iTRAQ reagent peaks used for relative quantitation showing the increased levels of the 115 and 117 labels compared with control protein-labeled with 114 and 116 (*inset*).

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

cells (18, 19). C/EBPα transcript levels are decreased 1.5-fold ($p = 0.017$) in TEL/PDGFRβ-transfected cells, although Pu.1 and other C/EBP family members are unchanged at the transcript level. Thus, the data we have derived allow further experiments on the mechanism of differentiation blockade in these cells. These can include pre-enrichment and selective searching for ions from C/EBPα in our iTRAQ reagent experiments to allow relative quantitation of transcription factor levels. In addition, Cathepsin G, which is also reduced in expression, is a serine protease highly expressed in promyelocytes (20, 21); it has a role in hematopoietic stem cell mobilization and differentiation and so may also play a role in the TEL-PDGFRβ-mediated differentiation block. Mutations in such proteases have been implicated in neutropenia (22, 23). Perhaps even more relevant is that cathepsin L has been shown to locate to the nucleus and regulate transcription via a proteolytic mechanism (24). Cathepsin G may have a similar function.

Comparison of this data set with a cDNA microarray data set from the same cell line provides relatively high levels of agreement between transcripts and protein level. All five of the changes from the 100-transcript sample set were also detected by iTRAQ; of these, all showed a similar level of change. However, the iTRAQ reagent approach identified changes in proteins that are caused by post-transcriptional effects, in that no change is seen in the levels of mRNA. A comparison of this type was previously almost impossible, because stable isotope or gel-based approaches tend to focus on identifying proteins whose expression changes, rather than the relative abundances of all proteins in a sample.

In conclusion, we have shown that iTRAQ protein labeling reagents can be employed to successfully identify proteins in which expression is potentially modified. This has the advantage of using multiple samples in a single LC-MS/MS run. The iTRAQ reagent produced high quality, reproducible data regarding relative expression levels in up to four samples. Comparison of the iTRAQ reagent data with cDNA microarray data suggests a high degree of similarity; all changes in a subset of the cDNA microarray are replicated by iTRAQ reagent analysis. However, iTRAQ reagent experiments also defined several other changes not detected by the cDNA array. iTRAQ reagent technology has great value as a new method for relative quantification of proteins in enriched complexes, organelles, and whole cell lysates.

* This work was supported by the Leukemia Research Fund and Biotechnology and Biological Sciences Research Council UK. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

S The on-line version of this article (available at http://www.jbc.org) contains Supplemental Table I.

‡‡ To whom correspondence should be addressed. Tel.: 0161-306-4182; Fax: 0161-236-0409; E-mail: awhetton@picr.man.ac.uk.

REFERENCES

1. Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., and Davison, M. (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* **1,** 377–396
2. Washburn, M. P., Wolters, D., and Yates, J. R., 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19,** 242–247
3. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999
4. Patton, W. F., Schulenberg, B., and Steinberg, T. H. (2002) Two-dimensional gel electrophoresis; better than a poke in the ICAT? *Curr. Opin. Biotechnol.* **13,** 321–328
5. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics* **1,** 376–386
6. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., 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
7. Golub, T. R., Barker, G. F., Lovett, M., and Gilliland, D. G. (1994) Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell* **77,** 307–316
8. Pierce, A., Owen-Lynch, P. J., Spooncer, E., Dexter, T. M., and Whetton, A. D. (1998) p210 Bcr-Abl expression in a primitive multipotent haematopoietic cell line models the development of chronic myeloid leukaemia. *Oncogene* **17,** 667–672
9. Heyworth, C. M., Dexter, T. M., Kan, O., and Whetton, A. D. (1990) The role of hemopoietic growth factors in self-renewal and differentiation of IL-3-dependent multipotential stem cells. *Growth Factors* **2,** 197–211
10. Carroll, M., Tomasson, M. H., Barker, G. F., Golub, T. R., and Gilliland, D. G. (1996) The TEL/platelet-derived growth factor beta receptor (PDGF beta R) fusion in chronic myelomonocytic leukemia is a transforming protein that self-associates and activates PDGF beta R kinase-dependent signaling pathways. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 14845–14850
11. Spooncer, E., Heyworth, C. M., Dunn, A., and Dexter, T. M. (1986) Self-renewal and differentiation of interleukin-3-dependent multipotent stem cells are modulated by stromal cells and serum factors. *Differentiation* **31,** 111–118
12. Dreger, M. (2003) Subcellular proteomics. *Mass Spectrom. Rev.* **22,** 27–56
13. Gouble, A., Grazide, S., Meggetto, F., Mercier, P., Delsol, G., and Morello, D. (2002) A new player in oncogenesis: AUF1/hnRNPD overexpression leads to tumorigenesis in transgenic mice. *Cancer Res.* **62,** 1489–1495
14. Perrotti, D., and Calabretta, B. (2004) Translational regulation by the p210 BCR/ABL oncoprotein. *Oncogene* **23,** 3222–3229
15. Yehiely, F., and Oren, M. (1992) The gene for the rat heat-shock cognate, hsc70, can suppress oncogene-mediated transformation. *Cell Growth Differ.* **3,** 803–809
16. Strobl, H., Takimoto, M., Majdic, O., Fritsch, G., Scheinecker, C., Hocker, P., and Knapp, W. (1993) Myeloperoxidase expression in CD34+ normal human hematopoietic cells. *Blood* **82,** 2069–2078
17. Ford, A. M., Bennett, C. A., Healy, L. E., Towatari, M., Greaves, M. F., and Enver, T. (1996) Regulation of the myeloperoxidase enhancer binding proteins Pu1, C-EBP alpha, -beta, and -delta during granulocyte-lineage specification. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 10838–10843
18. Friedman, A. D. (2002) Transcriptional regulation of granulocyte and monocyte development. *Oncogene* **21,** 3377–3390
19. Friedman, A. D. (2002) Runx1, c-Myb, and C/EBPalpha couple differentiation to proliferation or growth arrest during hematopoiesis. *J. Cell. Biochem.* **86,** 624–629
20. Heusel, J. W., Scarpati, E. M., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., Shapiro, S. D., and Ley, T. J. (1993) Molecular cloning, chromosomal location, and tissue-specific expression of the murine cathepsin G gene. *Blood* **81,** 1614–1623
21. MacIvor, D. M., Shapiro, S. D., Pham, C. T., Belaaouaj, A., Abraham, S. N., and Ley, T. J. (1999) Normal neutrophil function in cathepsin G-deficient mice. *Blood* **94,** 4282–4293

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019

22. Benson, K. F., Li, F. Q., Person, R. E., Albani, D., Duan, Z., Wechsler, J., Meade-White, K., Williams, K., Acland, G. M., Niemeyer, G., Lothrop, C. D., and Horwitz, M. (2003) Mutations associated with neutropenia in dogs and humans disrupt intracellular transport of neutrophil elastase. *Nat. Genet.* **35,** 90–96
23. Levesque, J. P., Hendy, J., Winkler, I. G., Takamatsu, Y., and Simmons, P. J. (2003) Granulocyte colony-stimulating factor induces the release in the bone marrow of proteases that cleave c-KIT receptor (CD117) from the surface of hematopoietic progenitor cells. *Exp. Hematol.* **31,** 109–117
24. Goulet, B., Baruch, A., Moon, N. S., Poirier, M., Sansregret, L. L., Erickson, A., Bogyo, M., and Nepveu, A. (2004) A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. *Mol. Cell* **14,** 207–219

Downloaded from http://www.mcponline.org/ by guest on September 17, 2019