Statistical Model to Analyze Quantitative Proteomics Data Obtained by $^{18}$O/$^{16}$O Labeling and Linear Ion Trap Mass Spectrometry

APPLICATION TO THE STUDY OF VASCULAR ENDOTHELIAL GROWTH FACTOR-INDUCED ANGIOGENESIS IN ENDOTHELIAL CELLS

Inmaculada Jorge‡§, Pedro Navarro‡§, Pablo Martínez-Acedo‡§¶, Estefanía Núñez‡, Horacio Serrano‡, Arántzazu Alfranca¶, Juan Miguel Redondo¶, and Jesús Vázquez‡**

Statistical models for the analysis of protein expression changes by stable isotope labeling are still poorly developed, particularly for data obtained by $^{18}$O/$^{16}$O labeling. Besides large scale test experiments to validate the null hypothesis are lacking. Although the study of mechanisms underlying biological actions promoted by vascular endothelial growth factor (VEGF) on endothelial cells is of considerable interest, quantitative proteomics studies on this subject are scarce and have been performed after exposing cells to the factor for long periods of time. In this work we present the largest quantitative proteomics study to date on the short term effects of VEGF on human umbilical vein endothelial cells by $^{18}$O/$^{16}$O labeling. Current statistical models based on normality and variance homogeneity were found unsuitable to describe the null hypothesis in a large scale test experiment performed on these cells, producing false expression changes. A random effects model was developed including four different sources of variance at the spectrum-fitting, scan, peptide, and protein levels. With the new model the number of outliers at scan and peptide levels was negligible in three large scale experiments, and only one false protein expression change was observed in the test experiment among more than 1000 proteins. The new model allowed the detection of significant protein expression changes upon VEGF stimulation for 4 and 8 h. The consistency of the changes observed at 4 h was confirmed by a replica at a smaller scale and further validated by Western blot analysis of some proteins. Most of the observed changes have not been described previously and are consistent with a pattern of protein expression that dynamically changes over time following the evolution of the angiogenic response. With this statistical model the $^{18}$O labeling approach emerges as a very promising and robust alternative to perform quantitative proteomics studies at a depth of several thousand proteins. Molecular & Cellular Proteomics 8:1130–1149, 2009.

Quantitative proteomics, which may be defined as the study of global changes in the expression level of proteins, is a field that has experienced great development in the last years (1, 2). In these studies two or more protein extracts from tissues, cells, or body fluids in changing environmental or physiological conditions are assayed to determine the presence of proteins exhibiting an alteration in their expression levels. Proteins have traditionally been separated, visualized, and subjected to relative quantification by two-dimensional gel electrophoresis and protein poststaining (3, 4) or prelabeling (5, 6). The spots showing quantitative differences were then excised and proteolyzed, and the resulting peptides were used to identify the proteins by MS (7).

Alternative strategies to the gel-based approaches have emerged in recent years; in these techniques, also called “shotgun proteomics,” protein extracts are digested without prior separation, and the complex peptide mixture is separated by multidimensional chromatography coupled to MS/MS analysis (8). The MS/MS spectra are then used for automated identification of protein components in a database. The development of stable isotope labeling protocols has allowed these techniques to produce relative quantitative information (1) and has considerably boosted the field in the last years. Stable isotopic labeling can be achieved by chemical (9, 10), metabolic (11), or enzymatic techniques (12–14). Enzymatic labeling with $^{18}$O tags is usually performed after protein digestion by incubating the resulting peptides with a proper endoprotease, usually trypsin (15). The protease catalyzes the incorporation of two $^{18}$O atoms at the carboxyl-terminal end of peptides and thus increases their mass by 4 Da. The labeled sample is then combined with the non-labeled one, and the two samples are concurrently processed and analyzed so that peptides are detected in MS as doublets, and the relative intensity of the two signals may be used to estimate the relative concentration of the corresponding proteins in the original samples (12). Because the isotopic patterns of the two peptide forms usually overlap, these analyses have routinely been done by using medium or high resolution MS analyzers. Ion traps can produce medium resolution mass spectra (or “ZoomScans”) of selected ions over...
a limited m/z range, and in our laboratory we have recently demonstrated that using these scanning modes and taking advantage of the high scanning speed of the linear ion trap it is possible to make accurate quantitative measurements without compromising the ability of this machine to perform high throughput peptide identification (16). In a recent refinement of the original technique and on the basis of a kinetic exchange model, quantification was performed by fitting the isotopic envelope composed by the two peptide forms to a theoretical curve, allowing a simultaneous calculation of the relative proportion of peptides in the original samples and of the specific labeling efficiency of each one of the peptide pairs (14). This method eliminates artifacts produced by incomplete oxygen exchange in subsets of peptides that have a low labeling efficiency and may be considered indicative of false expression changes and has been shown to be suitable for the quantification of whole proteomes (14).

Because of the large amounts of differential expression data gathered in these studies, one of the bottlenecks of current quantitative proteomics is determining whether the observed protein expression changes are correct or produced artificially. Although changes in protein expression may be considered significant when above a specified threshold, it is clearly preferable to use more rigorous statistical methods to establish statistical significance. In this regard, statistics applied to quantitative proteomics are considerably less developed in comparison with those currently applied in the microarray field. The multiple testing problem and the control of the false discovery rate (FDR)\(^1\) or the p value (17) on microarray data have been the subject of detailed discussion. A more recent controversy has centered on the validity of the assumption that p values generated by commonly used microarray statistics follow a uniform distribution (18). In clear contrast, published studies addressing statistical issues in the quantitative proteomics field are scarce. The majority of published quantitative proteomics studies have used the established two-dimensional gel electrophoresis approach and use the p value to identify protein species showing significant expression changes. The need to use robust experimental designs (19) and to control for FDR in differential electrophoresis studies has been claimed only recently (20). The majority of published DIGE studies use the Student’s t test to identify significant changes in expression (20); however, this test assumes independent sampling, normality of the data, and homogeneity of the variance, and the validity of these assumptions has not been demonstrated in a general context. The situation is not better with expression studies obtained using stable isotope labeling approaches. Although a great effort has been devoted to the development of bioinformatics tools for automated analysis of MS data and calculation of peptide ratios for each one of the several isotope labeling strategies currently available (21, 22), existing analytical methods for the statistical determination of significant expression changes are scarce (for a review, see Ref. 22). Almost all existing methods have in common that statistical significance is calculated assuming an underlying normal distribution for the null hypothesis from which Student’s t test (23, 24) or p values (25) are obtained. However, the general validity of this assumption has not been analyzed in detail. Only recently have studies begun to address the reliability and sources of errors in iTRAQ analysis (26, 27), and a profile likelihood algorithm has been used to estimate confidence intervals of protein abundance ratios determined by \(^{16}\)O/\(^{18}\)O metabolic labeling (28, 29); however, no specific statistical models have still been proposed to deal with data produced by enzymatic \(^{16}\)O/\(^{18}\)O labeling. Particularly large scale test studies aimed to establish the distribution of the null hypothesis, similar to those performed with microarray data (18), are lacking.

Angiogenesis, the formation of new blood vessels from a preexisting vascular bed, is determinant for the setup and progress of relevant pathophysiological processes: it is essential in embryo development, the menstrual cycle, wound healing, chronic inflammatory diseases, tumor growth, metastasis formation, and other processes. Angiogenesis comprises a series of finely regulated steps in which endothelial cells play a crucial role. The positive balance between pro- and antiangiogenic factors promotes the acquisition of a proangiogenic phenotype by endothelial cells, therefore triggering the formation of new vessels. Among positive regulators of angiogenesis, members of the vascular endothelial growth factor (VEGF) family are the best characterized. These factors regulate biological functions of endothelial cells involved in new vessel formation, proliferation, survival, migration, extracellular matrix remodeling, and other processes. Endothelial responses to VEGF are elicited through tyrosine kinase receptors among which VEGFR-2 (KDR; Fk-1) is considered the principal mediator of the proangiogenic activity. VEGF binding to this receptor gives rise to the activation of signal transduction routes (30), which ultimately lead to either the transcription of target genes or the posttranslational modification of presynthesized proteins as in the case of endothelial NO synthase (31). Antiangiogenic therapies that interfere directly or indirectly with VEGF production and the mechanisms responsible for its biological activity have already been successfully used in patients (32). Hence the study of mechanisms underlying biological actions promoted by VEGF in endothelial cells may be of great interest for the treatment of angiogenesis-dependent processes. However, to date only a few studies have been performed to analyze the effect of VEGF on endothelial cells by differential expression approaches, mainly by microarray techniques (33–35). Only two

\(^1\) The abbreviations used are: FDR, false discovery rate; HUVEC, human umbilical vein endothelial cell; VEGF, vascular endothelial growth factor; SCX, strong cation exchange; ANXA1, annexin A1; RCN1, reticulocalbin; TB, \(\beta\)-tubulin; iTRAQ, isobaric tags for relative and absolute quantitation; RP, reverse phase; Hsp, heat shock protein.
quantitative proteomics studies have been performed (36, 37); these studies have been done by using differential gel electrophoresis and after exposure of human umbilical vein endothelial cells (HUVECs) to VEGF for considerably long times (24 h or more). Shotgun approaches in conjunction with ICAT labeling have also been applied to study the effect of 24-h incubation of the proangiogenic factor Sokotrasterol sulfate on HUVECs (38). These results, obtained after prolonged exposure to the stimuli, might be reflecting variations in protein expression due to indirect effects of the treatments used. Therefore, studies performed at shorter times of treatment would be necessary to minimize the presence of indirect changes induced by VEGF stimulation.

In this work we performed a large scale test experiment using the $^{18}$O/$^{16}$O labeling technique on the proteome of HUVECs and developed a random effects statistical model for the null hypothesis that verifies the local assumptions of normality and variance homogeneity. We further demonstrate that with this model it is possible to detect significant and consistent protein expression changes in response to two short time (4- and 8-h) incubations with VEGF. These times were chosen with the idea of detecting effects more directly related to VEGF and to further analyze the dependence of the expression changes with the incubation time. The majority of the changes observed in this work have not been described previously, even using microarray approaches. The detected changes are consistent with a pattern of protein expression that dynamically changes over time following the evolution of the angiogenic response.

**EXPERIMENTAL PROCEDURES**

**Cell Culture and Protein Extraction**—HUVECs were isolated from human umbilical cord veins. The vein was cannulated and incubated with 0.1% collagenase (Roche Applied Science). Following removal of the collagenase, cells were pooled and established as primary cultures in medium 199 (BioWhittaker, Walkersville, MD) containing 20% FCS (Invitrogen). HUVECs were grown on dishes coated with 0.5% gelatin (Sigma) in medium 199 supplemented with 20% FCS, 50 µg/ml bovine brain extract, and 100 µg/ml heparin (Sigma). For VEGF treatment, cells were starved overnight in 0.5% FCS and then incubated for different times with 50 ng/ml VEGF (Peprotech, London, UK).

HUVECs (1.5 x 10^6 cells) were lysed at 4 °C in lysis buffer (50 mM Tris-HCl, pH 7.5, 400 mM NaCl, 1 mM EDTA, 2.5 mM EGTA, 1% Triton X-100, 1 mM DTT, 10 mM β-glycerocephosphate, 1 mM sodium vanadate, 1 mM NaF, 1 mM PMSF, 1 µg/ml aprotinin, 1 µg/ml leupeptin, 1 µg/ml pepstatin). For Western blot experiments, proteins (8 µg/lane) were resolved by SDS-PAGE and transferred to nitrocellulose membranes. The membranes were then probed with antibodies anti-annexin A1 (Santa Cruz Biotechnology, Santa Cruz, CA), anti-triose-phosphate isomerase (Abcam, Cambridge, UK), anti-reticulocalbin (Bethyl, Montgomery, TX), and anti-β-tubulin (Sigma). Immunoreactive bands were detected by an enhanced chemiluminescent substrate (ECL, GE Healthcare).

**In-solution Digestion**—Protein concentration in the samples was measured using the Bradford protein assay (39). Samples were desalted and buffer-exchanged on a prepacked PD-10 column (GE Healthcare). After denaturation of proteins with 9 M urea, disulfide bonds from cysteinyl residues were reduced with 10 mM DTT for 1 h at 37 °C, and then thiol groups were alkylated with 50 mM iodoacetamide for 1 h at room temperature in darkness. Samples were diluted to reduce urea concentration below 1.4 M and digested using sequencing grade trypsin (Promega, Madison, WI) overnight at 37 °C using a 1:20 (w/v) trypsin/protein ratio (40). Digestion was stopped by the addition of 1% TFA and by trypsin denaturation by reduction/alkylation as described above.

**Trypsin-catalyzed $^{18}$O Labeling**—Samples containing trypptic peptides were desalted using Oasis HLB (hydrophilic-lipophilic balance) Extraction cartridges (Waters, Milford, MA), dried down, and subjected to labeling with either H$_2^{16}$O or H$_2^{18}$O (95%; Isotec, Miamisburg, OH) in 100 mM ammonium acetate, pH 6.0, 20% CH$_3$CN in the presence of immobilized trypsin (Pierce) (41) at a 1:200 (v/w) trypsin/protein ratio. Peptides from the control samples were labeled with $^{16}$O, and peptides from the VEGF-treated samples were labeled with $^{18}$O. Labeling was stopped by addition of 5 mM ammonium formate, pH 3.0. Labeled samples were reduced and alkylated as indicated above, and tryptic beads were removed using a physical filter. The two labeled samples were then mixed and dried down.

**Peptide Fractionation by Strong Cation Exchange Liquid Chromatography**—Peptides were desalted onto a Vydc C$_{18}$ RP column (4.6 x 250 mm) using a Beckman Gold HPLC system by one-step elution with 40% CH$_3$CN containing 5 mM ammonium formate, pH 3.0. Clean peptide pools were fractionated by strong cation exchange (SCX) liquid chromatography using a 2.1 x 200-mm BioBasic SCX column (Thermo Electron, Cheshire, UK) operating at 200 µl/min. Peptides were eluted using the following gradient: 0% mobile phase B for 4 min, 0–10% B in 30 min, 10–40% B in 40 min, 40–100% B in 2 min, and 100% B for 8 min (solvent A: 5 mM ammonium formate, pH 3.0, 25% CH$_3$CN; solvent B: 500 mM ammonium formate, pH 3.0, 25% CH$_3$CN) (42). Peptides were collected in 1-min fractions.

**LC-MS/MS Analysis and Peptide Identification**—Each SCX fraction was analyzed by LC-MS/MS using a Surveyor LC system coupled to a linear ion trap mass spectrometer (model LTQ, Thermo Finnigan, San Jose, CA) as described previously with minor modifications (43, 44). Peptides were concentrated and desalted onto an RP precolumn (0.32 x 30 mm, BioBasic C$_{18}$, Thermo Electron) and eluted on line onto an analytical RP column (0.18 x 150 mm, BioBasic C$_{18}$, Thermo Electron) operating at 2 µl/min and using the following gradient: 5% B for 15 min, 5–14% B in 15 min, 14–30% B in 155 min, 30–95% B in 7 min, and 95% B for 3 min (solvent A: 0.1% formic acid; solvent B: 0.1% formic acid, 80% CH$_3$CN). The LTQ was operated in a data-dependent ZoomScan and MS/MS switching mode (45, 46). Zoom target parameters, number of microscans, normalized collision energy, and dynamic exclusion parameters were as described in Lopez-Ferrer et al. (16).

Protein identification was carried out as described previously (43). The MS/MS raw files acquired on the mass spectrometer were searched against the human Swiss-Prot database using the SEQUEST algorithm (Biorworks 3.2 package, Thermo Finnigan). The same collection of MS/MS spectra were searched against a pseudo-inverted database constructed from the same database. SEQUEST searches were performed allowing both optional (methionine oxidation and lysine and arginine modification of +4 Da) and fixed modifications (cysteine carboxamidomethylation). Determination of FDR as a measure of statistical significance of peptide identification was performed by using the probability ratio method (47).

**Peptide Quantification and Statistics Analysis**—Peptide quantification from ZoomScan spectra was performed as described previously (14, 16) using QuixToT, a program written in C# in our laboratory, that automatically opens the raw files and peptide identification results and sorts out the ZoomScan spectra corresponding to identified peptides (48). The quantification in QuixToT is performed using an
algorithm described previously (14). Only ZoomScan spectra corresponding to peptide matches with a false discovery rate lower than 5% were used for quantification. ZoomScan spectra were fitted to a theoretical curve so that the best fit parameters allow a simultaneous determination of the peptide concentration in the two samples ($A$ and $B$, respectively, expressed in units of area) and of the labeling efficiency $f$ (14). The proportion of non-labeled and mono- and dilabeled peptide in the labeled sample ($\beta_0$, $\beta_1$, and $\beta_2$, respectively) can be calculated from these parameters as described previously (14). The theoretical curve used to fit the spectra is a mixed Gaussian/double exponential distribution (14), and their parameters were not constrained during the fitting process. The best fit values of the parameters $\sigma$ (scale parameter, which determines the half-width of the peak) and $\beta$ (relative proportion of double exponential component) and of the labeling efficiency $f$ were used to eliminate unreliable quantifications. The typical value of $\sigma$ for the LTQ under the scanning conditions used in this work (16) was around 0.085 Da; an increase in this value indicated that the curve was not consistently fitting an isotopic peak. Similarly values of $\beta$ or $f$ above one were indicative either that the curve cannot be fitted by the theoretical curve or that the relative isotopic peak proportion could not be explained by the superimposition of three peptide components (unlabeled, mono-, and dilabeled), respectively. In the initial filtering step (see Table I) ZoomScans were automatically considered as unreliable and were eliminated when the best fit parameters verified any of the following conditions: $\sigma > 0.12$, $\beta > 1.1$, and $f > 1.2$.

**Calculation of Fitting Weights**—As described above, ZoomScan spectra were fitted to a theoretical function that depends on peptide concentration in the two samples. Expressing relative peptide concentrations in a base 2 logarithmic scale, if $A$ and $B$ are the peptide amounts expressed in units of area in the control and treated samples, respectively, the result of a quantification made in scan $s$ coming from peptide $p$ derived from protein $q$ is expressed as $x_{qps} = \log_2(A/B)$. The fitting weight of the scan is calculated using the following formula,

$$v_{qps} = \frac{T^2}{\text{MSDc} + \text{MSDi}}$$  \hspace{1cm} (Eq. 1)

where $T$ is a measure of peptide concentration in units of area, MSDc is the mean squared deviation between the experimental and theoretical ZoomScan spectra around a central window spanning the cluster of isotopic peaks, and MSDi is the mean squared deviation around a lateral window at the left or the right of the cluster. If $A + B_0 > B_1 + B_2$ then the lateral window is taken at the left, and $T = A$; otherwise the lateral window is taken at the right, and $T = B$. This procedure ensures that the fitting weight is proportional to the squared value of the peptide concentration and inversely proportional to the variance of the theoretical function; besides, it ensures that the weight does not suffer an abrupt decrease when there are large expression changes and also takes into account the presence of adjacent signals when they may artifactually interfere with the quantification. The variance in the log$_2$ ratio values is calculated from the fitting weight using the following formula: $\sigma^2_{qps} = k/v_{qps}$ where $k$ is a constant that for the LTQ and the scanning conditions used in this work has the value 0.17 (see supplemental information).

**Calculation of Averaged Means and Variances at the Scan, Peptide, and Protein Levels**—When the peptide amounts in the two samples are identical, $x_{qps} = 0$. Experimental deviations from this value are assumed to come from a systematic error, $\mu$, in the ratio in which the two samples are mixed; from deviations of protein concentration due to biological variability and errors committed during the process of preparation of the protein extracts, $\rho_q^p$; from deviations in peptide concentration due to peptide preparation from the protein extracts, $\beta_{qps}$; and from errors committed during the quantification of the peptide pair from the scan, $\xi_{qps}$, giving the following equation.

$$x_{qps} = \mu + \rho_q^p + \beta_{qps} + \xi_{qps}$$  \hspace{1cm} (Eq. 2)

Let us assume that $\beta_{qps}$ and $\rho_q^p$ are normally distributed, i.e., $\beta_{qps} \sim N(0, \sigma_{\beta}^2)$ and $\rho_q^p \sim N(0, \sigma_\rho^2)$, and that the peptide and protein variances, $\sigma_{\beta}^2$ and $\sigma_\rho^2$, respectively, are constant. Let us also assume that $\xi_{qps}$ is normally distributed according to $\xi_{qps} \sim N(0, \sigma_\xi^2 + k/v_{qps})$ where $\sigma_\xi^2$ is the scan variance, $v_{qps}$ is the fitting weight, and $k$ is the constant described in the previous section. The log$_2$ ratio value associated to each peptide is calculated as a weighted average of the scans used to quantify the peptide, and the value associated to each protein is similarly the weighted average of its peptides. Besides a grand mean is calculated as a weighted average of the protein values. The statistical weight associated to each scan, peptide, and protein is the inverse of their local variances, whereas the inverse of the variances of the averaged values is the sum of the inverses of variances of the values used to calculate the average. Hence the statistical weights are given by the following.

$$w_{qps} = \frac{1}{\frac{k}{v_{qps}} + \sigma_\xi^2}$$  \hspace{1cm} (Eq. 3)

$$w_{qp} = \frac{1}{\frac{1}{\sum w_{qps} x_{qps}} + \sigma_\rho^2}$$  \hspace{1cm} (Eq. 4)

$$w_q = \frac{1}{\frac{1}{\sum w_{qp} x_{pq}} + \sigma_\beta^2}$$  \hspace{1cm} (Eq. 5)

And the peptide and protein averages and the grand mean, which is an estimate of $\mu$, are given, respectively, by the following.

$$x_{qp} = \frac{\sum w_{qps} x_{qps}}{\sum w_{qps}}$$  \hspace{1cm} (Eq. 6)

$$x_q = \frac{\sum w_{qp} x_{pq}}{\sum w_{qp}}$$  \hspace{1cm} (Eq. 7)

$$x = \frac{\sum w_{q} x_{q}}{\sum w_{q}}$$  \hspace{1cm} (Eq. 8)

To take into account the systematic error $\mu$ of the experiment, protein ratios are normalized by subtracting the grand mean. The algorithm used to estimate the values of $\sigma_\beta^2$, $\sigma_\rho^2$, and $\sigma_\xi^2$ as well as their associated confidence intervals is described in the supplemental information.

As commented above, the local variances of the scan determinations and of the peptide and protein averages are the inverse of the scan weights in Equations 3–5, respectively. These variances may be used to determine whether a particular value significantly deviates from average more than expected according to the assumed normal distribution. Defining $P(\mu, \sigma^2, \lambda_x)$ as the two-tailed probability that the
we performed three large scale experiments using 1 mg of the proangiogenic factor VEGF in HUVECs. For this purpose,

| TABLE I |
|-------------------|
| Identification and quantification statistics of the three large scale experiments |
| Test | 4 h | VEGF | 8 h |
| SCX fractions<sup>a</sup> | 71 | 64 | 92 |
| MS/MS spectra obtained | 142,585 | 134,590 | 149,713 |
| MS/MS spectra yielding a peptide match<sup>b</sup> | 13,701 | 9,171 | 12,196 |
| Number of unique peptides identified | 4,878 | 4,806 | 3,876 |
| Number of unique proteins identified | 2,461 | 1,982 | 2,085 |
| Scans remaining after initial filtering<sup>c</sup> | 9,786 | 7,315 | 8,145 |
| Scans used for the analysis<sup>d</sup> | 7,640 | 5,097 | 4,391 |
| Number of unique peptides quantified | 2,271 | 2,556 | 1,218 |
| Number of unique proteins quantified | 1,246 | 1,278 | 890 |
| Scan variance (δ<sub>1</sub>) (95% C.I.) | 0.018 (0.016–0.020) | 0.024 (0.022–0.026) | 0.031 (0.028–0.034) |
| Peptide variance (δ<sub>2</sub>) (95% C.I.) | 0.021 (0.014–0.028) | 0.019 (0.014–0.028) | 0.014 (0.006–0.023) |
| Protein variance (δ<sub>3</sub>) (95% C.I.) | 0.0007 (0–0.009) | 0.004 (0–0.012) | 0.003 (0–0.011) |
| Significant expression changes (FDR<sub>1</sub> ≤ 5%) | 1 | 26 | 37 |
| Significant expression changes (FDR<sub>2</sub> ≤ 10%) | 1 | 32 | 46 |

<sup>a</sup> Fractions where peptides are detected.

<sup>b</sup> With an FDR of identification equal or lower than 5%, calculated using the probability ratio method (see “Experimental Procedures”).

<sup>c</sup> “Scans” refer to ZoomScan spectra. Initial filtering was made using the following criteria: σ < 0.12, β < 1.1, t < 1.2, and v<sub>avg</sub> > 3.1.

<sup>d</sup> Remaining after second filtering: elimination of scans from C-terminal peptides and peptides containing methionine residues and missed cleavage sites, and subpeptides from the latter.

<sup>e</sup> C.I., confidence interval.

value x<sub>t</sub> taken from sample t deviates from a normal distribution with mean μ and variance σ<sup>2</sup>, the probabilities that a given scan, peptide, or protein measurement is an outlier with respect to the corresponding average may be estimated, respectively, by the following.

\[ P_{qps} = P(x_{qp}, w_{qp}^{-1}, x_{qps}) \]  
(Eq. 9)

\[ P_{qp} = P(x_{qp}, w_{qp}^{-1}, x_{qp}) \]  
(Eq. 10)

\[ P_{q} = P(x_{q}, w_{q}^{-1}, x_{q}) \]  
(Eq. 11)

Detection of the presence of outliers is made by using multiple hypothesis testing and controlling for the FDR (49–52), defined as the proportion of values expected to deviate by chance alone within the population of observed outliers. Because the number of expected random changes is the product of the probability, as defined in Equations 9–11, multiplied by the total number of determinations at each level, the FDR at each one of the three levels is calculated by

\[ FDR_{qps} = \frac{P_{qps} \cdot NS}{O(p_{qps})} \]  
(Eq. 12)

\[ FDR_{qp} = \frac{P_{qp} \cdot NP}{O(p_{qp})} \]  
(Eq. 13)

\[ FDR_{q} = \frac{P_{q} \cdot NQ}{O(p_{q})} \]  
(Eq. 14)

where NS, NP, and NQ are the total number of scans, peptides, and proteins, respectively, and O(p) is the observed number of values detected in each population with a probability equal to or less than p.

RESULTS

Strategy for Quantitative Proteomics by 16O/18O Labeling and Analysis by Linear Ion Trap Mass Spectrometry—In this work, 18O-based quantitative proteomics analyses were performed to determine protein expression changes induced by the proangiogenic factor VEGF in HUVECs. For this purpose, we performed three large scale experiments using 1 mg of total protein extracts from untreated and VEGF-treated cells. In the two first experiments the cells were incubated in the presence of VEGF for 4 and 8 h, respectively. As a positive control of VEGF stimulation, cells were routinely checked for VEGF-induced increase in COX-2 expression and ERK-1,2 activation by Western blot analysis of these cell extracts (not shown) (33). In the third experiment two identical aliquots from a total HUVEC protein extract containing 1 mg of protein each were compared; this large scale control experiment was used to test for the null hypothesis and to study the robustness of the statistical model used to interpret the quantitative data. Finally a replica of the 4-h VEGF stimulation experiment was also performed at a smaller scale using 100 μg of each HUVEC protein extract.

The protein extracts were trypsin-digested in solution, and the resulting peptides were labeled postdigestion with either 18O or 16O in the presence of trypsin. The control and stimulated samples were then mixed and separated by SCX chromatography. Each fraction was analyzed by RP HPLC on line with an LTQ linear ion trap mass spectrometer using large gradients of 170 min; hence each one of the experiments took several weeks to complete. As described in a previous work (16), the LTQ was programmed to perform a ZoomScan spectrum and then an MS/MS spectrum over the most intense ions detected; the first scan was used to quantify the 16O/18O-labeled peptide pair, and the second scan was used to identify the peptide in a human database. About 10,000 MS/MS spectra were assigned to peptide sequences at an FDR of 5% using the probability ratio method (47), corresponding to around 4000 unique peptides and 2000 unique proteins in each one of the three large scale experiments (Table I). A typical distribution of identified peptides and proteins along the SCX fractions is illustrated in Fig. 1A. Proteins identified
in these experiments showed a considerable but not complete overlap (not shown), and combining the four experiments a total of 3878 unique proteins could be identified in this study.

Quantification was performed on the ZoomScan spectra corresponding to identified peptides only using a method previously developed in our laboratory that allowed control of labeling efficiency of each one of the peptides (14). In this method ZoomScan spectra were fitted to a theoretical curve, and the peptide concentration (in units of area) in the two samples and the labeling efficiency were determined simultaneously (14). As explained under “Experimental Procedures,” the best fit parameters were used as an initial criterion to automatically filter out unreliable quantifications. As indicated in Table I, in this initial filtering step about 20% of the scans corresponding to identified peptides were considered as un-
reliable. Scans corresponding to peptides not containing a basic residue in the carboxyl-terminal position, corresponding to peptides located at the carboxyl-terminal end of proteins, were also eliminated at this initial filtering step; these scans were present in a very low proportion with respect to the total number of spectra.

Monitoring of labeling efficiency at the individual peptide level was critical to achieve trustable quantification results in these large scale experiments. In initial attempts using previously published protocols (14, 16) we observed a slow but significant $^{18}$O/$^{16}$O back-exchange over the time needed to analyze all SCX fractions. This phenomenon was not originally observed when performing shorter experiments. It was necessary to improve labeling conditions until peptides remained stably labeled even after being stored for several months (not shown). In our hands it was critical to achieve a complete inactivation of residual trypsin activity, which was in part responsible for the slow back-exchange as observed by other authors (54). This was efficiently accomplished by using trypsin covalently bound to cross-linked agarose beads, which are eliminated by filtration, and by subjecting the filtered material to reduction and alkylation. Besides acidic hydrolysis of labeled peptides (54) was carefully avoided by maintaining pH at 3.0. Under these conditions all the peptides remained effectively labeled, having a labeling efficiency higher than 0.8 in all cases as illustrated in Fig. 1, B–E.

Analysis of the Null Hypothesis—To analyze the null hypothesis we studied the distribution of log$_2$ ratios obtained in the large scale test experiment where a proteome was compared against itself. We observed that peptides containing missed cleavage sites or oxidized methionine residues tended to deviate frequently from the expected 1:1 ratio, suggesting that these peptides did not reliably reflect protein concentration. To avoid these potential artifacts from altering the distribution of data under the null hypothesis, a second filtering step was performed, eliminating peptides that were partially digested or peptides containing oxidized methionines and also subpeptides belonging to the former and peptides containing non-oxidized methionine residues (Table I); the behavior of these peptides was analyzed later in more detail (see below). Partially digested peptides were assumed to be those containing at least one missed cleavage site that did not contain a proline residue after the arginine or lysine residues. The distribution of log$_2$ ratio values of filtered data is presented as a frequency histogram in Fig. 2A. As shown, the data were symmetrically distributed and had a Gaussian-like shape. The presence of outliers was inspected by fitting the histogram to a normal distribution and counting how many peptide pairs had a ratio that deviated from the mean value at a $p < 0.05$ confidence level. Using this commonly used criterion, more than 200 peptides were detected as having significant expression changes, corresponding to 72 proteins.

The number of significantly altered proteins decreased when the FDR was used as a criterion; thus, more than 40 peptides, corresponding to 20 proteins, were detected as outliers at a 5% FDR.

As these results were unacceptable because application of these criteria to the real experiments would have resulted in a large number of false expression changes, we analyzed in more detail whether the normal distribution was a good model for this kind of data. For this end, a cumulative normalized frequency plot was constructed and fitted to the theoretical curve corresponding to a Gaussian function; as shown, clear deviations were evident between the theoretical curve and the experimental data points (Fig. 2B). Besides the data were analyzed using a normal probability plot (Fig. 2C); in this kind of plot the ordered observations are represented against the inverse of the standard normal cumulative and
**Statistical Model for Quantitative Proteomics by \(^{18}\text{O}\) Labeling**

**Fig. 3.** Classification of data into normal populations according to the fitting weight. *A*, distribution of fitting weights as a function of \(\log_2\) ratios. *Inset*, results of the D'Agostino normality test (56) applied in sliding windows of 200 points ranked according to fitting weight; shown are values taken by the \(Y\) parameter (56) as a function of the median fitting weight in each window. 99% confidence levels for \(Y\) values for a normal distribution (2.9 to 5.6) are indicated in the *inset*. For comparison, the arrow in the \(Y\) axis of the *inset* indicates the \(Y\) value (-330) obtained by applying the test to the entire population. The minimum fitting weight threshold used in this study (3.1) is indicated by *discontinuous lines* in the main figure and in the *inset*. *B* and *C* are cumulative frequency and normal probability plots, respectively, as in *Fig. 2*, corresponding to the window indicated in *A*.

are very useful to detect sources of deviation of normality (55). As shown, the data considerably deviated from the expected values (*Fig. 2C, straight line*) and showed the characteristic shape indicative of leptokurtosis (55). Finally the data failed to pass a normality test (56) (*Fig. 3A, inset, arrow*). These findings demonstrated that the Gaussian distribution, which is the statistical model most commonly used to analyze this kind of data, is not adequate to describe the behavior of the null hypothesis.

**A Random Effects Statistical Model for the Null Hypothesis:**

Calculation of Variance Components and of Peptide and Protein Averages—We hypothesized that the failure to follow a normal distribution was due to the presence of peptide pairs that were quantified with different accuracy so that the variance in the entire population of data was not homogeneous. With this idea in mind, we analyzed whether the data could be classified into discrete categories having homogeneous variance. For this purpose, we considered that quantifications performed with peaks having a larger intensity, and hence a larger peak area, should be more accurate than those performed with low intensity peaks; hence a plot of the squared intensity against the \(\log_2\) ratio values was constructed. As shown in *Fig. 2D*, the dispersion of ratio values around the mean ratio clearly diminished when peak intensity increased, suggesting that the accuracy of quantification and hence the variance was intensity-dependent. We also observed the presence of numerous peptide pairs having a relatively high intensity that showed significant deviations from the expected value, indicating that the squared intensity alone could not be used as a criterion to classify data into homogeneous categories.

Because quantification of peptide pairs was made by fitting each ZoomScan spectrum to a theoretical function (14), deviations from the computed curve are indicative of errors committed in the calculation of the ratios from the fitted parameters. As explained under “Experimental Procedures,” it is possible to estimate the variance of the \(\log_2\) ratio values from the mean squared deviation of the fitted curve. The influence of contaminating peaks adjacent to the peptide pair peaks was also estimated from the mean squared deviations computed in narrow windows at the left and right of the main isotopic peaks. These mean squared deviations together define a statistical parameter, \(v_{qps}\), defined as the fitting weight of scan \(p\) coming from peptide \(q\) derived from protein \(s\) (Equation 1). The fitting weight, in a \(\log_2\) ratio scale, is expected to be inversely proportional to the variance produced by deviations between the theoretical and the experimental spectrum profiles. Because this parameter contains the squared peak intensity, spectra having higher intensity peaks have higher fitting weights than those displaying lower signal to noise ratios. In addition, the fitting weights tend to acquire lower values when the theoretical isotopic envelope fails to describe the experimental peaks or when co-eluting peaks are superimposed or adjacent to the main isotopic peaks. In *Fig. 3A* the fitting weights are plotted versus \(\log_2\) ratio values; as shown, an excellent correlation between \(\log_2\) ratio dispersion and fitting weight was obtained. The data were ranked according to decreasing fitting weights, and the normality test was then applied in sliding windows containing 200 ZoomScan spectra each. As shown in *Fig. 3A, inset*, the normality test was satisfactorily passed for all scan weights above a minimum threshold value. Besides the cumulative frequency distributions within each window could be satisfactorily fitted by Gaussian curves (*Fig. 3B*), and the shape of the curves in normal probability plots did not indicate apparent deviations.
from normality (Fig. 3C). This result strongly suggested that, above the minimum threshold, the fitting weight may be used to classify the data set into discrete categories having homogeneous variance; in other words, that scans having similar fitting weights have similar log2 ratio variances. Besides these variances may be directly calculated from the fitting weights (see “Experimental Procedures”). In addition, this criterion allows using the minimum threshold value as an efficient cutoff parameter to eliminate low quality quantifications whose variance cannot be statistically controlled (Fig. 3A, horizontal dashed line, and Fig. 3A, inset, vertical dashed line).

During the analysis of the relation between fitting weight and scan variance, we observed that the variance tended to be a constant, non-zero value when fitting weight increased (supplemental Fig. S1). This finding indicated the presence of a constant residual scan variance component, termed $\sigma^2_{R}$, that was independent of the curve fitting; in other words, that a “perfect” quantification (in terms of the goodness of fitting and signal to noise ratio of ZoomScan spectra) still retains a constant error source that is independent of scan quality and hence of signal intensity. This is due to variations in the signal produced by the MS detector; this variance component is approximately constant because of the logarithmic nature of the scale used to analyze peptide ratios. To further analyze the behavior of this error source, we selected a subpopulation of data containing high quality quantifications only by choosing an arbitrary fitting weight threshold (supplemental information) above which the effect of curve fitting errors was expected to be negligible. We observed that this population passed the normality test, suggesting that the error source produced by the MS detector is homogeneous and can be reasonably modeled by a Gaussian distribution (not shown).

In addition, errors are also expected at the peptide level because not all the peptides derived from the same protein produce exactly the same quantitative result. Similarly a further error source has to be considered at the protein level because not all the proteins are produced at exactly the same concentration due to the manipulation of the sample. These three error sources were analyzed by assuming a random effects, hierarchical statistical model. Because the population of high quality quantifications was found to display a normal behavior, we thought it reasonable to assume that in this population these three error sources have a constant variance and are normally distributed. Finally because the fitting weight is a property that only depends on the spectrum profile and is not related to the peptide and protein nature, we also found it reasonable to assume that peptide and protein variances are constant for all the scans in the population. The mathematical details of the statistical model are described under “Experimental Procedures.”

Because as explained above quantifications at the scan level have a variance component due to curve fitting whose magnitude depends on the fitting weight, not all the quantifications are made with the same accuracy. Therefore when different quantifications from the same peptide are considered together they have to be weighted to estimate the averaged peptide ratio; this is done by using as statistical weight the inverse of the scan variance (Equations 3 and 6). In turn, averaging of different scans increases the accuracy of the peptide mean so that, in the absence of other error sources, the inverse of the peptide variance may be calculated by adding up the inverses of scan variances; this is equivalent to adding up scan weights. Note that this is the general procedure to calculate the variance of a weighted mean and that if there are N scans and all of them have the same variance then the variance of the mean takes the well known form, i.e. the scan variance divided by N. The final expression for the statistical weight associated to each peptide that measures accuracy of quantification at the peptide level is finally calculated by taking into account the contribution of errors at the scan level and the error produced by peptide preparation, which are measured by the peptide variance $\sigma^2_p$ (Equations 4 and 7). Similar considerations explain the formulas used to calculate the weighted average and the statistical weight at the protein level (Equations 5 and 8).

To apply the model to the practice it is not only necessary to estimate the variance components at the scan, peptide, and protein levels but also to determine how reliable these estimations are. This was accomplished in each one of the experiments, as explained in the supplemental information, by using the subset of high quality quantifications described in the previous section and applying an algorithm to estimate the variance components that best explain the structure of the data. The algorithm was tested by constructing data sets of similar size and structure, introducing normally distributed random errors having preset variance components, and checking that the estimated variances coincide with the original ones (supplemental Fig. S3). These simulations were particularly critical to ensure that the variance estimates were non-biased due to the particular structure of these data where most of the peptides and proteins were quantified by one or two scans and peptides, respectively (supplemental Fig. S3). These simulations were also used to calculate the 95% confidence intervals of the estimated variances.

Application of the algorithm to the three large scale experiments gave the results presented in Table I. As shown, an excellent agreement was found in the scan and peptide variances among the different experiments. The protein variances were not significantly different from zero in any of the cases, suggesting that the experimental design of the comparative study was adequate to determine the presence of significant expression changes at the protein level. In general, these results indicate that the three variance components can be efficiently calculated from the data and hence that the random effects model is in practice suitable for these kind of studies.

**Analysis of Outliers at the Scan and Peptide Levels—**As explained before, the scan populations having the same fitting weight (Fig. 3A) behaved as normal distributions. This as-
assumption was further checked by analyzing the presence of outliers. For this end, the variances estimated in the previous section were used to calculate peptide and protein averages and the grand mean using Equations 3–8. In this step the population of scans remaining after the second filtering was used, and hence scans belonging to peptides containing missed cleavage sites or subpeptides derived from these or peptides containing methionine residues were not taken into account. The presence of outliers at the scan level was analyzed by checking for the presence of scans having a deviation from the peptide average higher than that expected by chance alone. The probability that a scan deviates from the peptide average was calculated using Equation 9 and the corresponding error rate (FDR$_{qp}$) using Equation 12. Determinations having an FDR$_{qp}$ lower than 5% were visually inspected. In this step some scan outliers appeared whose deviation from the peptide average was explained by the presence of other isobaric peptides or contaminants, elevated background noise, or poor curve fitting. After eliminating these artifacts, the number of true scan outliers was extremely low (Fig. 4, *left panels*); only a maximum of four outliers per experiment was detected among several thousand scans. A similar analysis was performed to detect peptide quantifications that significantly deviated from the corresponding protein averages; this was done using as a criterion the parameter FDR$_{qps}$ (Equations 10 and 13). The number of true peptide outliers was again almost negligible (Fig. 4, *right panels*).

As commented above, an additional source of variability at the peptide level was due to methionine oxidation and partial protein digestion; because these factors were found with some frequency to introduce considerable deviations in the ratios, they were analyzed in more detail. Oxidation of methionines occurs spontaneously, and this process may take place at different extents in the two samples so that the relative proportion of oxidized and non-oxidized peptides may deviate from the true peptide ratio. To check this phenomenon, the dispersion from protein averages in the log$_2$ ratio values corresponding to peptides containing non-oxidized methionines was compared with those of methionine-oxidized peptides (Fig. 5, *A–C*). A visual inspection of this figure immediately suggested that with some frequency both oxidized and non-oxidized forms of methionine-containing peptides tended in general to deviate from their corresponding protein averages. Partial protein digestion was also an impor-

![Fig. 4. Analysis of outliers at the scan (A–C) and peptide (D–F) levels. Dark points indicate outliers at the scan (FDR$_{qps}$ < 5%) or peptide levels (FDR$_{qp}$ < 5%) in the test experiment (A and D) and the 4-h (B and E) and 8-h VEGF incubation experiments (C and F).](image-url)
tant source of deviations at the peptide level because partially digested peptides may not accurately reflect the protein concentration in the two samples. A similar approach was used to analyze this effect. The log2 ratios corresponding to peptides containing a missed cleavage were superimposed with those of completely digested peptides and also with subpeptides whose sequence is located inside a larger quantified peptide of the same protein. As shown in Fig. 5, D–F, the ratios of partially digested peptides were displaced from the expected values, whereas those of subpeptides tended to be displaced to opposite values. In general, these kind of artifacts due to oxidation and digestion are easily detected as outliers by using the parameter FDR_{qp} as in Fig. 4; in this work we have judged these data as unreliable for protein quantification and have therefore eliminated them from the analysis.

As shown in Table I, once poor quality determinations and partially digested and methionine-containing peptides were discarded, the number of scans used for the quantitative analysis was considerably lower than the number of MS/MS spectra yielding a positive peptide identification. It is important to note here that this is not due to the 18O labeling technique by itself because, as shown in Fig. 1, the labeling efficiency is well controlled in these experiments. Similarly it is not due to the statistical model used in this work because the number of outliers at the scan and peptide levels is negligible. Most of the poor quality scans eliminated in the first filtering are due to the limited resolving power of the ZoomScan mode in the linear ion trap, whereas the methionine oxidation and partial digestion effects are inherent to peptide-centric quantification approaches based on postdigestion labeling.

Analysis of Outliers at the Protein Level in the Test Experiment, and Symmetry and Consistency of Detected Expression Changes—From a purely statistical viewpoint, outliers at the protein level are significant expression changes and are detected by using the same strategy as that used at the scan and peptide levels; protein outliers present in the test experiment should, however, be considered as true artifacts. The outliers were analyzed by inspecting for the presence of protein values with an FDR_{qp} lower than 10%. Only one outlier was identified among more than 1200 proteins in the test experi-

Fig. 5. Effect of methionine oxidation and partial digestion on relative peptide quantification. In A–C, scans corresponding to peptides containing oxidized (Mox) and non-oxidized methionines (M) are indicated by black points and white circles, respectively. In D–F, scans corresponding to peptides containing missed cleavage (cleav.) sites, except those with a proline after the basic residue, are indicated by a black point; subpeptides derived from a partially digested peptide identified in the same experiment are indicated by a white circle. Results correspond to the test experiment (A and D) and the 4-h (B and E) and 8-h VEGF incubation experiments (C and F).
ment. A further validation was done by studying the protein log2 ratio distribution according to the estimated variance of each one of the proteins. For this purpose log2 ratios at the protein level ($x_q$) were converted to standard normal values ($z_q$) by subtracting the grand mean and dividing by the standard deviation, i.e. $z_q = (x_q - \bar{x}) / \sqrt{w_i}$. As shown in Fig. 6A, left inset, the normalized cumulative frequency distribution of the standardized variable behaved according to the expected standard normal distribution $z_q \sim N(0,1)$, and in the corresponding normal probability plot (Fig. 6A, right inset) no deviations from normality became apparent. These findings provided a final and conclusive support of the validity of the statistical model.

Once the statistical model was validated, a set of experiments was performed to check the robustness of the new labeling protocol and the accuracy and symmetry of calculated protein ratios. One experiment addressed the consistency of results when the labeling protocol inverted the sample that was subjected to $^{18}$O labeling. In this experiment no bias was found toward the sample being labeled with $^{18}$O, and the data showed a good symmetry (supplemental Fig. S4A). In another experiment the theoretical and experimental protein ratios were compared in two aliquots from the same proteome doped with a set of proteins at different ratios; an excellent agreement was found between the expected and calculated values (supplemental Fig. S4B).

Finally the consistency of the method was analyzed by comparing the expression changes induced by VEGF after 4-h incubation in the large scale experiment and its technical replica at a smaller scale. As expected, not all the proteins quantified in the large experiment could be detected in the small experiment; however, among the proteins quantified in the two experiments, all the proteins displaying significant expression changes in the large experiment were observed to maintain an expression change of the same sign and similar extent in the small experiment even when the significance threshold was relaxed to a 35% FDR (Fig. 7). Hence the method showed an excellent consistency among two technical replicas.

Analysis of Protein Expression Changes Induced by VEGF Treatment in Endothelial Cells—In clear contrast with the test experiment, significant expression changes could be detected when cells were subjected to VEGF incubation for 4 and 8 h (Fig. 6, B and C, and Tables II and III). Some of the observed changes were moderate, although they were statistically significant. The protein values having the highest protein weights in the plots of Fig. 6 corresponded mainly to proteins quantified by several peptides. For instance, the von Willebrand factor showed only a 1.33-fold decrease after 4-h incubation; however, this protein was quantified by 12 peptides so that the expression change was associated with an FDR lower than 0.01%. Other proteins detected as statistically significant were quantified by a lower number of peptides but showed a larger expression change.

In this study, 32 and 46 proteins were found to be differentially expressed after VEGF treatment for 4 and 8 h, respectively. Most of the differentially expressed proteins in cells treated with VEGF for 4 h were down-regulated (84%) rather than up-regulated (16%) at 4 h (Fig. 6). The difference between the proportions of down- and up-regulated proteins was less remarkable (65 versus 35%, respectively) in cells treated with VEGF for 8 h, suggesting a biphasic expression pattern. This finding was consistent with the fact that the majority of proteins showing expression changes were different at 4 and 8 h. According to the Gene Ontology annotations, most of the differentially expressed proteins at 4 h are cytoplasmic and seem to be implicated in protein folding, whereas at 8 h the majority of changing proteins are distributed not only in the cytoplasm but also in the intracellular membrane-
bound organelle part, being primarily implicated in the cellular response to diverse stimulus.

Among the three experiments the relative proportion of a total of 1319 unique proteins was quantified in response to VEGF stimulus for 4 or 8 h. Among these proteins we found the majority of HUVEC proteins previously reported to show significant expression changes in response to VEGF (36, 37) or the angiogenic factor sokotrasterol sulfate (38), but very few of the proteins previously reported as differentially expressed were found to change their expression levels in our study. Among the proteins detected as differential expression or the angiogenic factor sokotrasterol sulfate (38), but very significant expression changes in response to VEGF (36, 37) and the algorithm we have developed to control and correct for the labeling efficiency has also been published previously (14). Controlling labeling efficiency for each one of the quantified peptides was particularly critical to develop an appropriate labeling protocol for these long lasting high throughput studies and to achieve consistent quantitative results. It should be noted that isotope labeling performance has also been claimed as potentially problematic in other labeling approaches. For instance, efficient using of stable isotope label-

DISCUSSION

In this work we describe the largest proteomics study performed to date in HUVECs. More than 3800 unique proteins were identified among four different experiments; from these, we were able to quantify more than 1300 unique proteins using stable 18O labeling and linear ion trap mass spectrometry. These results conclusively demonstrate the suitability of the ZoomScanning technique in the linear ion trap for relative peptide quantification using this labeling approach. The scanning conditions necessary to perform alternate ZoomScan and MS/MS spectra were described in a previous work (16), and the algorithm we have developed to control and correct for the labeling efficiency has also been published previously (14). Controlling labeling efficiency for each one of the quantified peptides was particularly critical to develop an appropriate labeling protocol for these long lasting high throughput studies and to achieve consistent quantitative results. It should be noted that isotope labeling performance has also been claimed as potentially problematic in other labeling approaches. For instance, efficient using of stable isotope label-
Detection of significant expression changes in endothelial cells in response to VEGF was possible by the development of a new statistical model for the analysis of peptide and protein expression changes. The model is characterized by three main features. First, the model considers that not all relative quantifications of peptide pairs are performed with the same accuracy. Therefore, the variance cannot be assumed to be homogeneous in the entire population, and it is necessary to classify the data into subsets where measurements have the same variance. This is efficiently done using the fitting weight, a parameter that measures the dispersion in quantifications of peptide pairs are performed with the same sequence, it is possible to assign a variance due to quantification from the isotopic envelope to each one of the scan measurements. Other works using $^{14}$N/$^{15}$N metabolic label-

### Table II

| Protein name                                                                 | Accession number | $(x_{\alpha} - x_{\beta}) \pm$ S.D. ($n$) | -Fold change | $p_{\alpha}$ | FDR$_{\alpha}$ |
|------------------------------------------------------------------------------|------------------|------------------------------------------|--------------|-------------|---------------|
| Calfreticulin (precursor)                                                     | P27797           | 1.03 ± 0.14 (2)                          | 2.04 down    | 3.3e−14     | 0.00          |
| Cofilin-2                                                                    | Q9Y281           | 1.86 ± 0.26 (1)                          | 3.62 down    | 2.1e−12     | 0.00          |
| Tyrosyl-tRNA synthetase, cytoplasmic                                          | P54577           | 1.53 ± 0.22 (1)                          | 2.88 down    | 8.9e−12     | 0.00          |
| Triose-phosphate isomerase                                                    | P60174           | 1.03 ± 0.16 (2)                          | 2.04 down    | 3.3e−11     | 0.00          |
| Calumenin (precursor)                                                         | O43852           | 0.96 ± 0.15 (2)                          | 1.94 down    | 4.3e−10     | 0.00          |
| von Willebrand factor (precursor)                                             | P04275           | 0.41 ± 0.08 (12)                         | 1.33 down    | 4.3e−07     | 0.01          |
| UDP-glucose:glucosylprotein glucosyltransferase 1 (precursor)                | Q9NYU2           | 1.13 ± 0.23 (1)                          | 2.20 down    | 4.8e−07     | 0.01          |
| Multifunctional protein ADE2                                                  | P22234           | 1.18 ± 0.23 (1)                          | 2.27 down    | 4.0e−07     | 0.01          |
| Rho GTPase-activating protein 29                                              | Q52LW3           | 1.22 ± 0.26 (1)                          | 2.33 down    | 2.1e−06     | 0.04          |
| Annexin A1                                                                   | P04083           | 0.44 ± 0.10 (5)                          | 1.36 down    | 8.3e−06     | 0.1           |
| Eukaryotic translation initiation factor 5B                                   | O60841           | 0.61 ± 0.14 (3)                          | 1.53 down    | 7.7e−06     | 0.1           |
| Reticulocalbin-1 (precursor)                                                  | Q15293           | 0.55 ± 0.13 (3)                          | 1.46 down    | 2.5e−05     | 0.3           |
| Heat shock protein Hsp90−α                                                    | P07900           | 0.36 ± 0.09 (9)                          | 1.28 down    | 3.0e−05     | 0.3           |
| γ-Interferon-inducible protein Iif-16                                         | Q16666           | 0.77 ± 0.20 (1)                          | 1.71 down    | 7.6e−05     | 0.6           |
| YTH domain family protein 2                                                   | Q9Y5A9           | 0.97 ± 0.25 (1)                          | 1.97 down    | 8.7e−05     | 0.7           |
| Flotillin-1                                                                   | O75955           | 0.92 ± 0.24 (1)                          | 1.89 down    | 9.7e−05     | 0.7           |
| Lupus La protein                                                              | P05456           | 0.75 ± 0.20 (1)                          | 1.68 down    | 0.0002      | 1.3           |
| RR1-like protein B                                                            | Q14684           | 0.91 ± 0.25 (1)                          | 1.88 down    | 0.0002      | 1.3           |
| T-complex protein 1 subunit $\theta$                                          | P50990           | 0.47 ± 0.13 (3)                          | 1.38 down    | 0.0002      | 1.5           |
| Obg-like ATPase 1                                                             | Q9NTK5           | 0.56 ± 0.16 (2)                          | 1.48 down    | 0.0004      | 2.4           |
| Coatomer subunit 8                                                           | P35606           | 0.59 ± 0.17 (2)                          | 1.50 down    | 0.0006      | 3.4           |
| Phosphoglycerate mutase 1                                                     | P18669           | 0.63 ± 0.19 (1)                          | 1.55 down    | 0.0009      | 4.5           |
| Cation-dependent mannose 6-phosphate receptor (precursor)                    | P20645           | 0.77 ± 0.23 (1)                          | 1.70 down    | 0.0011      | 5.1           |
| Cytochrome b-c$_1$ complex subunit 1, mitochondrial (precursor)               | P31930           | 0.54 ± 0.17 (2)                          | 1.45 down    | 0.0014      | 5.9           |
| tRNA (cytosine-5)-methyltransferase NSUN2                                    | Q08J23           | 0.80 ± 0.25 (1)                          | 1.74 down    | 0.0014      | 6.1           |
| Carboxyl reductase (NADPH)                                                   | P16152           | 0.76 ± 0.25 (1)                          | 1.69 down    | 0.0021      | 8.3           |
| Staphylococcal nuclease domain-containing protein 1                           | Q7KZF4           | 0.41 ± 0.13 (3)                          | 1.33 down    | 0.0022      | 8.7           |
| Proteasome subunit $\alpha$ type-3                                             | P25788           | −0.76 ± 0.19 (1)                         | 1.70 up      | 7.6e−05     | 0.6           |
| Replication protein A 14-kDa subunit                                          | P35244           | −0.81 ± 0.24 (1)                         | 1.75 up      | 0.0006      | 3.1           |
| $\alpha$-Endosulfine                                                         | O43768           | −0.80 ± 0.24 (1)                         | 1.75 up      | 0.0009      | 4.3           |
| Squamous cell carcinoma antigen recognized by T-cells 3                      | Q15020           | −0.76 ± 0.23 (1)                         | 1.69 up      | 0.0010      | 4.6           |
| Methionine adenosyltransferase 2 subunit $\beta$                              | Q9NWL9           | −0.73 ± 0.23 (1)                         | 1.65 up      | 0.0017      | 6.7           |
ing have also described that peptide quantification variance is not homogeneous but correlated with signal to noise ratio (28, 29); that observation is in agreement with the results presented in this work because scan variance decreases as a function of the fitting weight, which in turn is proportional to the square of peptide peak area (Equation 1), a parameter that reflects the effect of peak intensity in a manner similar to the signal to noise ratio.

Second, the model takes into account three additional sources of error at each one of the quantitative levels: scan, peptide, and protein. This is done by applying a hierarchical, random effects model and calculating the peptide and protein averages as well as their deviations taking into account the error contributions from each one of the three sources. Peptide and protein averages are calculated as weighted means using as statistical weights the inverse of variances of the

### Table III

| Protein name                                      | Accession number | \((x_q - \bar{x}) \pm \text{S.D. (} n \)) | Fold change | \(p_q\) | FDR\(_q\) |
|---------------------------------------------------|------------------|------------------------------------------|-------------|--------|----------|
| Tryptophanyl-tRNA synthetase, cytoplasmic         | P23381           | 1.22 ± 0.14 (1)                          | 2.34 down   | 0.00   | 0.00     |
| Purine-nucleoside phosphorylase                   | P00491           | 0.97 ± 0.14 (1)                          | 1.95 down   | 2.7e-12 | 0.00     |
| Echinoderm microtube-associated protein-like 3    | Q32P44           | 1.16 ± 0.19 (1)                          | 2.23 down   | 0.00   | 0.00     |
| Ubiquitin carboxyl-terminal hydrolase 5           | P45974           | 1.40 ± 0.24 (1)                          | 2.63 down   | 4.3e-09 | 0.00     |
| High mobility group protein B2                    | P26583           | 0.99 ± 0.16 (1)                          | 1.99 down   | 1.8e-09 | 0.00     |
| Lamin-B1                                          | P20700           | 1.01 ± 0.17 (1)                          | 2.01 down   | 7.8e-09 | 0.00     |
| Mitochondrial inner membrane protein              | Q16891           | 0.74 ± 0.16 (1)                          | 1.68 down   | 1.7e-06 | 0.02     |
| Probable palmitoyltransferase ZDHHC13             | Q8IU4H4          | 1.29 ± 0.27 (1)                          | 2.45 down   | 2.0e-06 | 0.03     |
| Nucleolin                                         | P19338           | 0.94 ± 0.20 (1)                          | 1.92 down   | 2.9e-06 | 0.03     |
| Interferon-induced transmembrane protein 3        | Q01628           | 0.88 ± 0.19 (1)                          | 1.84 down   | 2.5e-06 | 0.03     |
| Myotrophin                                        | P58546           | 1.00 ± 0.23 (1)                          | 1.99 down   | 1.4e-05 | 0.1      |
| Coflin-2                                          | Q9Y281           | 0.81 ± 0.19 (1)                          | 1.75 down   | 2.4e-05 | 0.2      |
| Nuclear receptor-binding factor 2                 | Q96F24           | 0.94 ± 0.23 (1)                          | 1.92 down   | 5.0e-05 | 0.3      |
| Casein kinase II subunit β                        | P67870           | 0.70 ± 0.17 (1)                          | 1.62 down   | 4.9e-05 | 0.4      |
| CDGS-2 Fe(3+) domain-containing protein 2         | Q9S251           | 0.86 ± 0.22 (1)                          | 1.81 down   | 8.6e-05 | 0.5      |
| Zinc transporter 1                                 | Q9VM35           | 0.88 ± 0.23 (1)                          | 1.84 down   | 0.0001 | 0.7      |
| Pirin                                             | O00626           | 0.84 ± 0.23 (1)                          | 1.79 down   | 0.0002 | 1.3      |
| Glutaredoxin-3                                    | Q76003           | 0.61 ± 0.18 (1)                          | 1.53 down   | 0.0005 | 2.3      |
| Transcription intermediary factor 1-β             | Q13126           | 0.65 ± 0.19 (1)                          | 1.57 down   | 0.0005 | 2.4      |
| Serpin H1 (precursor)                             | P50454           | 0.48 ± 0.14 (1)                          | 1.40 down   | 0.0007 | 3.0      |
| Superoxide dismutase (Cu-Zn)                     | P00441           | 0.98 ± 0.29 (1)                          | 1.97 down   | 0.0009 | 3.1      |
| UBX domain-containing protein 8                   | Q96CS3           | 0.57 ± 0.17 (1)                          | 1.48 down   | 0.0009 | 3.2      |
| Ran GTPase-activating protein 1                   | P46060           | 0.49 ± 0.15 (1)                          | 1.41 down   | 0.0009 | 3.3      |
| Interleukin enhancer-binding factor 2             | Q12095           | 0.54 ± 0.17 (1)                          | 1.45 down   | 0.0012 | 3.5      |
| 40 S ribosomal protein S14                        | P62263           | 0.48 ± 0.15 (1)                          | 1.39 down   | 0.0011 | 3.5      |
| Ras-associating death domain protein              | Q13158           | 0.74 ± 0.23 (1)                          | 1.67 down   | 0.0017 | 4.5      |
| ATP-dependent RNA helicase DDX39                  | Q00148           | 0.40 ± 0.13 (2)                          | 1.32 down   | 0.0019 | 4.8      |
| Enoyl-CoA hydratase, mitochondrial (precursor)    | P30084           | 0.47 ± 0.15 (1)                          | 1.39 down   | 0.0023 | 5.2      |
| Protein-β isoaspartate (β-aspartate) O-methyltransferase | P22061           | 0.76 ± 0.25 (1)                          | 1.70 down   | 0.0022 | 5.2      |
| Translation initiation factor elf-2B subunit δ    | Q9U10            | 0.76 ± 0.26 (1)                          | 1.69 down   | 0.0045 | 8.9      |
| Leucine-rich repeat flightless-interacting protein 1 | Q32M24          | -0.66 ± 0.18 (1)                         | 1.58 up     | 0.0003 | 1.3      |
| Eukaryotic translation initiation factor 2A       | Q9BY44           | -0.45 ± 0.13 (3)                         | 1.37 up     | 0.0003 | 1.4      |
| Patatin-like phospholipase domain-containing protein 2 | Q96AD5           | -1.06 ± 0.31 (1)                         | 2.09 up     | 0.0005 | 2.3      |
| Succinate dehydrogenase (ubiquinone) flavoprotein subunit, mitochondrial (precursor) | P31040           | -0.57 ± 0.17 (2)                         | 1.48 up     | 0.0008 | 3.0      |

\[ \Delta^1-Pyrroline-5-carboxylate synthetase \]

\[ \text{P54886} \]

\[ -0.59 \pm 0.18 \ (2) \]

\[ 1.51 \ up \]

\[ 0.0010 \]

\[ 3.1 \]

\[ \text{P60228} \]

\[ -0.60 \pm 0.18 \ (1) \]

\[ 1.51 \ up \]

\[ 0.0010 \]

\[ 3.1 \]

\[ \text{Q9NP74} \]

\[ -0.99 \pm 0.31 \ (1) \]

\[ 1.99 \ up \]

\[ 0.0012 \]

\[ 3.5 \]

\[ \text{Q13610} \]

\[ -0.60 \pm 0.19 \ (1) \]

\[ 1.52 \ up \]

\[ 0.0015 \]

\[ 4.2 \]

\[ \text{Q95140} \]

\[ -0.49 \pm 0.16 \ (1) \]

\[ 1.40 \ up \]

\[ 0.0027 \]

\[ 6.0 \]

\[ \text{P13667} \]

\[ -0.39 \pm 0.13 \ (2) \]

\[ 1.31 \ up \]

\[ 0.0034 \]

\[ 7.2 \]

\[ \text{Q94832} \]

\[ -0.39 \pm 0.13 \ (2) \]

\[ 1.31 \ up \]

\[ 0.0038 \]

\[ 8.1 \]

\[ \text{P08107} \]

\[ -0.66 \pm 0.23 \ (1) \]

\[ 1.58 \ up \]

\[ 0.0040 \]

\[ 8.5 \]

\[ \text{Q99536} \]

\[ -0.39 \pm 0.14 \ (2) \]

\[ 1.31 \ up \]

\[ 0.0044 \]

\[ 8.7 \]
values from which the average is calculated, and their vari-
ances are similarly calculated by integrating the variance at
each level with the variances of the data used to calculate the
average. This procedure is not only sound from the statistical
viewpoint but also integrates the results in a coherent and
intuitive manner. Thus, the averaged values are skewed to-
ward the determinations that are more accurate \( i.e. \)
that have lower variance), and lower quality determinations have little
effect on the average and hence do not need to be eliminated.
Besides the model takes naturally into account the fact that
quantifications at any of the higher levels (peptide or protein)
are more reliable when they are calculated from two or more
quantifications from the lower levels (scans or peptides, re-
spectively). Moreover peptide and protein measurements
are assumed to have their own source of error, and this error
remains despite that the error transmitted from lower levels
is decreased by extensive scan and peptide averaging,
respectively. If these higher level errors are not taken into
account, the peptides measured with a very large number of
scans would have an almost negligible error, and protein
averages would be completely skewed toward these pep-
tides. In clear contrast, our hierarchical model assigns a
similar weight to two peptides when they are measured by
at least one good scan with little influence of the actual
number of scans with which they are measured. Quality
weights that are the inverse of variances have recently been
used in the microarray field; by assigning lower weights to
less reproducible arrays in different experiments the power
to detect differential expression was reported to increase
\( \text{(67).} \)

The model also address two of the most controversial
points in quantitative proteomics: what is the minimum num-
er of peptides required to consider protein quantification
reliable, and what is the lowest expression change that may
be considered statistically significant? In our model the accu-
racy of quantification is measured by its associated variance,
which depends not only on the number of peptides with which
it is measured but also on the accuracy of each one of the
peptide measurements. Hence a protein quantified by two
peptide measurements with low accuracy may have a higher
variance than a protein quantified by only one peptide with
higher accuracy. The statistical analysis will then produce a
population of proteins with large heterogeneity with regard to
their number of peptides and their associated variance,
and the general rule will then be that a protein expression change
may be called as significant when its ratio significantly devi-
ates from the global distribution above the expected limits
imposed by its own accuracy. It can be argued that a protein
expression change is less prone to artifacts when it is de-
tected by more than one peptide; however, in the experiments
presented here all the peptides analyzed, except for a remark-
ably negligible number of exceptions, show a ratio dispersion
that lies within the limits expected according to their variance.
In such a context of absence of outliers it is reasonable to call
as significant an expression change for proteins quantified by
only one peptide.

Two main sources of peptide outliers are methionine ox-
idation and variations in protein digestion. Oxidation of methi-
onines may take place at a different extent in the two samples,
and this source of error is technically very difficult if not

---

**Fig. 8.** Western blot analysis of expression changes of some of the proteins quantified in this study. A, protein extracts from HUVECs incubated in the absence or presence of VEGF for various times as indicated were separated by SDS-PAGE, blotted, and analyzed using specific antibodies against ANXA1 (37), RCN1 (36), triose-phosphate isomerase (TPIS), and TB. The position of molecular mass standards are indicated at the left. In B the stained bands were quantified by densitometry, and the resulting data were plotted as the ratio of band intensities obtained in the presence and absence of VEGF. Statistical analysis by conventional regression methods indicated the existence of a significant correlation between protein ratio and incubation time \( p < 0.02 \) for ANXA1, RCN1, and triose-phosphate isomerase; in the case of TB the correlation was not statistically significant \( p > 0.3 \).
impossible to control. For the same reason, peptides containing missed cleavage sites indicate incomplete trypsin digestion at these sites, the extent of which may again be different in the two samples, making protein quantification unreliable. Although these two potential sources of artifacts are obvious, they have not been analyzed before and were apparently not taken into account in the studies devoted to protein quantification by stable isotope labeling.

A very useful graphical interpretation of the quantitative results is obtained by plotting protein weights (i.e. the inverse of protein variances) versus log$_2$ ratio values (Fig. 6). As shown in this figure and in Tables II and III, there are proteins displaying very moderate expression changes that are highly significant due to their higher protein weight, whereas others, displaying higher expression changes, cannot be considered significant because they have lower protein weights. The protein weight is expected to increase with protein concentration due to an increase in the number of quantified peptides; therefore these peptide-centric methods will naturally tend to produce better results with higher abundance proteins. Not surprising, in this work we were able to detect moderate protein expression changes that were not detected as significant by microarray approaches.

The third characteristic feature of our model is the assumption that the variances associated to the three error sources at the different levels are constant and common for the entire population of scans, peptides, and proteins in the same experiment. This viewpoint comes from a global approach in contrast with the local Student’s $t$ test approach commonly used in microarray or two-dimensional gel electrophoresis experiments where even being measured by the same number of determinations each protein is assumed to have a different variance. In the microarray field several groups have suggested that estimates of variance from individual genes are unreliable (18, 68, 69), and it has been pointed out that when the variance from each gene is truly unknown it makes sense to consider all of the genes on the array as arising from a single, normal distribution (18). In our case variance homogeneity at the three levels is convincingly supported by several lines of evidence. First, the population of scans having a large fitting weight (i.e. that have negligible errors due to curve fitting) pass the normality test; because the fitting weight is completely independent of the scan, peptide, and protein variances, such a result would not have been obtained if one of the error sources is not normally distributed. Second, the variances obtained at the three levels in several different experiments take essentially the same values, and their dispersion, calculated by computer simulation, indicates that they can be estimated with good accuracy. Third, the numbers of outliers at the scan and peptide levels are negligible; this result indicates that all the scans belonging to one peptide and all the peptides belonging to one protein display the same expression change. This result would have not been obtained if the error sources were not homogeneous. It should be noted that it is a common practice to eliminate scan or peptide outliers using criteria such as Dixon’s (23, 25, 70).

The mathematical details of the model are sound. A final result that supports the validity of the method is given by the practical absence of false protein expression changes in the test experiment. Although these kinds of tests have been reported in the microarray field (see for instance Fodor et al. (18)), the number of tests experiments published supporting the validity of statistical models using stable isotope labeling data are almost non-existent in the literature. A further point that remains to be studied is whether the global variance approach, and in general the statistical model presented in this work, could be applied to quantitative experiments performed using other isotope labeling methods such as iTRAQ. In this regard, it should be noted that only the error sources at the fitting and scan levels seems to be specific for the $^{18}$O labeling method and the linear ion trap analyzer used in this study; whereas the error sources at the peptide and protein levels are expected to depend only on sample preparation and protein digestion, provided that labeling efficiency effects are properly controlled. This suggests that by introducing an appropriate model for scan variance under the null hypothesis the random effects statistical model proposed here could be used to analyze ITRAQ data and also other isotope labeling methods.

Application of the statistical model to the analysis of the VEGF effect on endothelial cells allowed the detection of significant expression changes at 4- and 8-h incubations. Because of the large number of proteins quantified, the expression changes were detected as significant by controlling for the FDR. Interestingly the majority of detected proteins were down-regulated in response to VEGF after 4-h incubation. This behavior has also been observed in other studies using the same cells; thus, when HUVECs were stimulated with cancer cell-conditioned medium, a model expected to resemble tumor angiogenesis better than VEGF alone, most differentially expressed proteins (88%) were also down-regulated (37). We also observed that the expression changes were more equilibrated at 8 h, the detected proteins were different from those observed at 4 h, and the classification of proteins according to Gene Ontology categories followed a different pattern; all these results are consistent with a biphasic process. The fact that almost all the proteins showing significant expression changes are different from those observed in other works (36–38) where cells were incubated for longer times with the proangiogenic factor and the results obtained by Western blotting with a selected set of proteins using different incubation times are together consistent with a pattern of protein expression that dynamically changes over
time following the evolution of the angiogenic response. The exposure to proangiogenic stimuli could give rise either to posttranslational modifications of preexisting proteins or to de novo synthesis of angiogenesis regulators. These mediators would then act in an autocrine fashion, thus amplifying the angiogenic response at later times. We cannot, however, discard the possibility that the differences between this work and previously published works may in part be due to the much more rigorous statistical criterion used in this work to establish significant expression changes, including not only a validated null hypothesis model but also the use of FDR and not p values or -fold changes alone.

It is interesting to note that the majority of significant expression changes detected in this study have not been described before even by microarray approaches and hence give novel information about the molecular processes involved in the response to VEGF at earlier stages. Some of these proteins have been directly or indirectly associated with angiogenesis in endothelial cells. In addition to Hsp70, a chaperone reported to display down-regulation in HUVECs exposed to VEGF (36, 37), we also found a down-regulation of Hsp90, and some studies show that inhibitors of this protein have antiangiogenic activity in vivo (58, 59). Down-regulation of adipose triglyceride lipase is also consistent with a proangiogenic event because this protein acts as a receptor for pigment epithelium-derived factor (57), a strong endogenous inhibitor of angiogenesis (71). Similarly Ran GTPase-activating protein regulates the nuclear localization of PTEN, a tumor suppressor gene product that is also closely associated with angiogenesis (72), suggesting that down-regulation of Ran GTPase-activating protein by VEGF might suppress the nuclear translocation of PTEN. Our study also revealed a down-regulation of cofilin-2 both at 4 and 8 h of incubation, and it is interesting to note that VEGF has been reported to stimulate cofilin phosphorylation in HUVECs (62). Of the two aminoacyl-tRNA synthetases displaying expression changes in this study, tyrosyl-tRNA synthetase, which was detected after 4-h incubation with VEGF, has been described as a proangiogenic factor on HUVECs (73), whereas tryptophanyl-tRNA synthetase, which was detected after 8-h incubation with VEGF, has been shown to inhibit HUVEC migration and angiogenesis (64). Finally previous studies reporting that localization of nucleolin at the cell surface is modulated by VEGF in the process of angiogenesis (60) and the finding that casein kinase 2 regulates cellular processes, such as organization of nucleolar chromatin where nucleolin plays a fundamental role (74), are also in agreement with the implication of these two proteins in angiogenesis.

In conclusion, the results presented in this work suggest that the $^{18}$O labeling approach and analysis by linear ion trap mass spectrometry in conjunction with a computational algorithm that allow a precise control of labeling efficiency and a robust and validated statistical model provide a promising and semiautomated alternative to perform large scale studies of differential expression of proteins by stable isotope labeling. With this approach we were able to identify novel protein expression changes in HUVECs in response to incubations with the proangiogenic factor VEGF for short periods of time.

Acknowledgment—We thank Aurora García-Dorado for helpful discussion on analysis of variance.

* This work was supported in part by Grants BIO2006-10085 and SAF 2006-8348 from Spain’s Ministerio de Educación y Ciencia, by Grant CAM BIO/0194/2006 from Comunidad de Madrid, by Grant LSHTM-CT-2004-005033-EICOSANOX from the European Union, by the Fondo de Investigaciones Sanitarias (Ministerio de Sanidad y Consumo, Instituto Salud Carlos III, RECAVA Grant RD06/0014), and by an institutional grant from Fundación Ramón Areces (to Centro de Biología Molecular Severo Ochoa).

The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

REFERENCES

1. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. Nature 422, 196–207
2. Hanash, S. (2003) Disease proteomics. Nature 422, 226–232
3. Fievet, J., Dillmann, C., Lagniel, G., Davanture, M., Negroni, L., Labarre, J., and de Vienne, D. (2004) Assessing factors for reliable quantitative proteomics based on two-dimensional gel electrophoresis. Proteomics 4, 1939–1949
4. Smejkal, G. B., Robinson, M. H., and Lazarev, A. (2004) Comparison of fluorescent stains: relative photostability and differential staining of proteins in two-dimensional gels. Electrophoresis 25, 2511–2519
5. Yan, J. X., Devenish, A. T., Wait, R., Stone, T., Lewis, S., and Fowler, S. (2002) Fluorescence two-dimensional difference gel electrophoresis and mass spectrometry based proteomic analysis of Escherichia coli. Proteomics 2, 1682–1698
6. Hu, Y., Wang, G., Chen, G. Y., Fu, X., and Yao, S. Q. (2003) Proteome analysis of Saccharomyces cerevisiae under metal stress by two-dimen- sional differential gel electrophoresis. Electrophoresis 24, 1458–1470
7. Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., and Mann, M. (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. U. S. A. 93, 14440–14445
8. Yates, J., R., III, Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal. Chem. 67, 1426–1436
9. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using iso- tope-coded affinity tags. Nat. Biotechnol. 17, 994–999
10. Ross, P. L., Huang, Y. N., Marchesa, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Pulka, A., Purkayastha, S., Juhász, 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
11. Ong, S. E., Blagov, 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
12. Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., and Roepstorff, P. (2000) Quantitation of peptides and proteins by ma- trix-assisted laser desorption/ionization mass spectrometry using $^{18}$O-
Statistical Model for Quantitative Proteomics by 18O Labeling

labeled internal standards. Rapid Commun. Mass Spectrom. 14, 1226–1232

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

Ramos-Fernandez, A., Lopez-Ferrer, D., and Vazquez, J. (2007) Improved method for differential expression proteomics using trypsin-catalyzed 18O labeling with a correction for labeling efficiency. Mol. Cell. Proteomics 6, 1274–1286

Yao, X., Afonso, C., and Fenselau, C. (2003) Dissection of proteolytic 18O labeling: endoprotease-catalyzed 18O-to-16O exchange of truncated peptide substrates. J. Proteome Res. 2, 147–152

Lopez-Ferrer, D., Ramos-Fernandez, A., Martinez-Bartolome, S., Garcia-Ruiz, J. P., and Vazquez, J. (2008) Quantitative proteomics using 18O/16O labeling and linear ion trap mass spectrometry. Proteomics 6, Suppl. 1, S4–S11

Storey, J. D., and Tibshirani, R. (2003) Statistical significance for genome-wide studies. Proc. Natl. Acad. Sci. U. S. A. 100, 9440–9445

Karp, N. A., and Lilley, K. S. (2007) Design and analysis issues in quantitative proteomics studies. Proteomics 7, Suppl. 1, 42–50

Karp, N. A., McCormick, P. S., Russell, M. R., and Lilley, K. S. (2007) Experimental and statistical considerations to avoid false conclusions in proteomics studies using differential in-gel electrophoresis. Mol. Cell. Proteomics 6, 1354–1364

Mueller, L. N., Brusniak, M. Y., Mani, D. R., and Aebersold, R. (2008) An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. J. Proteome Res. 7, 51–61

Lau, K. W., Jones, A. R., Swainston, N., Siepen, J. A., and Hubbard, S. J. (2007) Capture and analysis of quantitative proteomic data. Proteomics 7, 2787–2799

MacCoss, M. J., Wu, C. C., Liu, H., Sadygov, R., and Yates, J. R. (2003) A correlation algorithm for the automated quantitative analysis of shot-noise and software variations in isobaric tags for relative and absolute quantitation (ITRAQ). J. Proteome Res. 2, 821–827

Boehm, A. M., Putz, S., Altrohfer, D., Sickmann, A., and Falk, M. (2007) Precise protein quantification based on peptide quantification using iTRAQ and AGC_MaximumQuant. J. Proteome Res. 6, 7214–7221

Pan, C., Kora, G., McDonald, W. H., Tabb, D. L., Ver Berkmoes, N. C., Hurst, G. B., Pelletier, D. A., Samatova, N. F., and Hettich, R. L. (2006) ProRata: a quantitative proteomics program for accurate protein abundance ratio estimation with confidence interval evaluation. Anal. Chem. 78, 7121–7131

Pan, C., Kora, G., Tabb, D. L., Pelletier, D. A., McDonald, W. H., Hurst, G. B., Hettich, R. L., and Samatova, N. F. (2006) Robust estimation of protein abundance ratios and rigorous scoring of their variability and bias in quantitative shotgun proteomics. Anal. Chem. 78, 7110–7120

Carmeliet, P. (2000) Mechanisms of angiogenesis and arteriogenesis. Nat. Med. 6, 389–395

Papapetropoulos, A., Garcia-Cardena, G., Madr, J. A., and Sessa, W. C. (1997) Nitric oxide production contributes to the angiogenic properties of vascular endothelial growth factor in human endothelial cells. J. Clin. Invest. 100, 3131–3139

Shih, T., and Lindley, C. (2006) Bevacizumab: an angiogenesis inhibitor for the treatment of solid malignancies. Clin. Ther. 28, 1779–1802

Abe, M., and Sato, Y. (2001) cDNA microarray analysis of the gene expression profile of VEGF-activated human umbilical vein endothelial cells. Angiogenesis 4, 289–298

Minami, T., Horuchi, K., Miura, M., Abid, M. R., Takabe, W., Noguchi, N., Kohro, T., Ge, X., Abrutani, H., Hamakubo, T., Kodama, T., and Aird, W. C. (2004) Vascular endothelial growth factor- and thrombin-induced termination factor, Down syndrome critical region-1, attenuates endothelial cell proliferation and angiogenesis. J. Biol. Chem. 279, 50537–50544

Yang, S., Toy, K., Ingle, G., Zlot, C., Williams, P. M., Fuh, G., Li, B., de Vos, A. A., and Gerritsen, M. E. (2002) Vascular endothelial growth factor-induced genes in human umbilical vein endothelial cells: relative roles of KDR and Flt-1 receptors. Arterioscler. Thromb. Vasc. Biol. 22, 1797–1803

Pawlowska, Z., Baranska, P., Jerczynska, H., Kozikiolkiewicz, W., and Cierneewski, C. S. (2005) Heat shock proteins and other components of cellular machinery for protein synthesis are up-regulated in vascular endothelial cell growth factor-activated human endothelial cells. Proteomics 5, 1217–1227

Katanasaka, Y., Asai, T., Naitou, H., Onashi, N., and Oku, N. (2007) Proteomic characterization of angiogenic endothelial cells stimulated with cancer cell-conditioned medium. Biol. Pharm. Bull. 30, 2300–2307

Karsan, A., Pellet, I., Yu, L. R., Chan, K. C., Conrads, T. P., Lucas, D. A., Andersen, R., and Veenstra, T. (2005) Quantitative proteomic analysis of sotokastreol sulfate-stimulated primary human endothelial cells. Mol. Cell. Proteomics 4, 191–204

Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254

Villar, M., Ortega-Perez, I., Were, F., Cano, E., Redondo, J. M., and Vazquez, J. (2006) Systematic characterization of phosphorylation sites in NFATc2 by linear ion trap mass spectrometry. Proteomics 6, Suppl. 1, S16–S27

Sevinsky, J. R., Brown, K. J., Cargile, B. J., Bundy, J. L., and Stephenson, J. L., Jr. (2007) Minimizing back exchange in 180/160 quantitative proteomics experiments by incorporation of immobilized trypsin into the initial digestion step. Anal. Chem. 79, 2158–2162

Liu, T., Qian, W. J., Strittmatter, E. F., Camp, D. G., II, Anderson, G. A., Thrall, B. D., and Smith, R. D. (2004) High-throughput comparative peptide gel analysis using a quantitative cysteinyl-peptide enrichment technology. Anal. Chem. 76, 5345–5353

Lopez-Ferrer, D., Martinez-Bartolome, S., Villar, M., Campillos, M., Martin-Maroto, F., and Vazquez, J. (2004) Statistical model for large-scale peptide identification in datasets from tandem mass spectra using SEQUEST. Anal. Chem. 76, 6883–6880

Ortega-Perez, I., Cano, E., Were, F., Villar, M., Vazquez, J., and Redondo, J. M. (2005) c-Jun N-terminal kinase (JNK) positively regulates NFATc2 transactivation through phosphorylation within the N-terminal regulatory domain. J. Biol. Chem. 280, 20867–20878

Serrano, H., Jorge, I., Martinez-Acedo, P., Navarro, P. J., P´erez-Hern´andez, D., ´Miro Cases, E., Garcia-Dorado, D., and Vazquez, J. (2007) Quantitative proteomics of mitochondrial membrane proteins by sodium dodecyl sulphate polyacrylamide gel electrophoresis, 160/180 stable isotope labeling and linear ion trap mass spectrometry. Proteomics 7, 24–34

Jorge, I., Casas, E. M., Villar, M., Ortega-Perez, I., Lopez-Ferrer, D., Martinez-Ruiz, A., Carrera, M., Marina, A., Martinez, P., Serrano, H., Canas, B., Were, F., Gallardo, J. M., Lamas, S., Redondo, J. M., Garcia-Dorado, D., and Vazquez, J. (2007) High-sensitivity analysis of specific peptides in complex samples by selected MS/MS ion monitoring and linear ion trap mass spectrometry: application to biological studies. J. Mass Spectrom. 42, 1391–1404

Martinez-Bartolome, S., Navarro, P., Martin-Maroto, F., Lopez-Ferrer, D., Ramos-Fernandez, A., Villar, M., Garcia-Ruiz, J. P., and Vazquez, J. (2008) Properties of average score distributions of SEQUEST: the probability ratio method. Mol. Cell. Proteomics 7, 1135–1145

Navarro, P. J., Martinez-Acedo, P., Serrano, H., Jorge, I., and Vazquez, J. (2007) A fully automated and integrated bioinformatic toolset for large-scale peptide identification and quantification by 18O/16O stable isotope labeling in the 2nd Congress of the SEProt and 1st Meeting of the European Proteomics Association, Valencia, March 13, 2007, p. 119, Sociedad Española de Proteómica (SEProt), Valencia, Spain

Choi, H., and Nesiˇzvˇaski, A. I. (2008) False discovery rates and related statistical concepts in mass spectrometry-based proteomics. J. Proteome Res. 7, 47–50

Kall, I., Storey, J. D., MacCoss, M. J., and Noble, W. S. (2008) Assigning
significance to peptides identified by tandem mass spectrometry using decoy databases. J. Proteome Res. 7, 29–34
51. Tabb, D. L. (2008) What’s driving false discovery rates? J. Proteome Res. 7, 45–46
52. Fitzgibbon, M., Li, Q., and McIntosh, M. (2008) Modes of inference for evaluating the confidence of peptide identifications. J. Proteome Res. 7, 35–39
53. Hernandez, G. L., Volpert, O. V., Iniguez, M. A., Lorenzo, E., Martinez-Martinez, S., Grau, R., Fresno, M., and Redondo, J. M. (2001) Selective inhibition of vascular endothelial growth factor-mediated angiogenesis by cyclosporin A: roles of the nuclear factor of activated T cells and cyclooxygenase 2. J. Exp. Med. 193, 607–620
54. Staes, A., Demol, H., Van Damme, J., Martens, L., Vandekerckhove, J., and Gevaert, K. (2004) Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. J. Proteome Res. 3, 786–791
55. D’Agostino, R. B., Belanger, A., and D’Agostino, R. B. J. (1990) A suggestion for using powerful and informative tests of normality. Am. Stat. 44, 316–321
56. D’Agostino, R. B. (1971) An omnibus test of normality for moderate and large size samples. Biometrika 58, 341–348
57. Notari, L., Baladron, V., Aroca-Aguilar, J. D., Balko, N., Heredia, R., Meyer, C., Notario, P. M., Saravanamuthu, S., Nueda, M. L., Sanchez-Sanchez, F., Escribano, J., Laborda, J., and Becerra, S. P. (2006) Identification of a lipase-linked cell membrane receptor for pigment epithelium-derived factor. J. Biol. Chem. 281, 38022–38037
58. Kaur, G., Belotti, D., Burger, A. M., Fisher-Nielsen, K., Borsotti, P., Riccardi, E., Thillainathan, J., Hollingshead, M., Sausville, E. A., and Giavazzi, R. (2004) Antiangiogenic properties of 17-(dimethylaminoethylamino)-17-demethoxygeldanamycin: an orally bioavailable heat shock protein 90 modulator. Clin. Cancer Res. 10, 4813–4821
59. Chaudhury, S., Welch, T. R., and Blagg, B. S. (2006) Hsp90 as a target for drug development. Chem. Med. Chem. 1, 1331–1340
60. Huang, Y., Shi, H., Zhou, H., Song, X., Yuan, S., and Luo, Y. (2006) The angiogenic function of nucleolin is mediated by vascular endothelial growth factor and nonmuscle myosin. Blood 107, 3564–3571
61. Keeler, S. M., Ivie, S. E., Krutzsch, H. C., Tandie, A., Libutti, S. K., and Roberts, D. D. (2003) Angiogenesis inhibitors target the endothelial cell cytoskeleton through altered regulation of heat shock protein 27 and cofilin. Cancer Res. 63, 6405–6412
62. Gong, C., Stoletov, K. V., and Terman, B. I. (2004) VEGF treatment induces signaling pathways that regulate both actin polymerization and depolymerization. Angiogenesis 7, 313–321
63. Tzima, E., and Schimmel, P. (2006) Inhibition of tumor angiogenesis by a natural fragment of a tRNA synthetase. Trends Biochem. Sci. 31, 7–10
64. Wakasugi, K., Slike, B. M., Hood, J., Otani, A., Ewalt, K. L., Friedlander, M., Chesh, D. A., and Schimmel, P. (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. Proc. Natl. Acad. Sci. U. S. A. 99, 173–177
65. Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Scholer, H., and Mann, M. (2008) Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. Mol. Cell. Proteomics 7, 672–683
66. D’Ascenzo, M., Choe, L., and Lee, K. H. (2008) iTRAQPak: an R based analysis and visualization package for 8-plex isobaric protein expression data. Brief. Funct. Genomics Proteomics 7, 127–135
67. Ritchie, M. E., Diyagama, D., Neilson, J., van Laar, R., Dobrovic, A., Holoway, A., and Smyth, G. K. (2006) Empirical array quality weights in the analysis of microarray data. BMC Bioinformatics 7, 261
68. Baldi, P., and Long, A. D. (2001) A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. Bioinformatics 17, 509–519
69. Allison, D. B., Cui, X., Page, G. P., and Sabripour, M. (2006) Microarray data analysis: from disarray to consolidation and consensus. Nat. Rev. Genet. 7, 95–105
70. Park, S. K., Venable, J. D., Xu, T., and Yates, J. R., III (2008) A quantitative analysis software tool for mass spectrometry-based proteomics. Nat. Methods 5, 319–322
71. Dawson, D. W., Volpert, O. V., Gillis, P., Crawford, S. E., Xu, H., Benedict, W., and Bouck, N. P. (1999) Pigment epithelium-derived factor: a potent inhibitor of angiogenesis. Science 285, 245–248
72. Ueda, S., Basaki, Y., Yoshie, M., Ogawa, K., Sakisaka, S., Kuwano, M., and Ono, M. (2006) PTEN/Akt signaling through epidermal growth factor receptor is prerequisite for angiogenesis by hepatocellular carcinoma cells that is susceptible to inhibition by gefitinib. Cancer Res. 66, 5346–5353
73. Yang, X. L., Schimmel, P., and Ewalt, K. L. (2004) Relationship of two human tRNA synthetases used in cell signaling. Trends Biochem. Sci. 29, 250–256
74. Srivastava, M., and Pollard, H. B. (1999) Molecular dissection of nucleolin’s role in growth and cell proliferation: new insights. FASEB J. 13, 1911–1922