Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry

David K. Han[1], Jimmy Eng[2], Huilin Zhou[2], and Ruedi Aebersold[2]*

An approach to the systematic identification and quantification of the proteins contained in the microsomal fraction of cells is described. It consists of three steps: (1) preparation of microsomal fractions from cells or tissues representing different states; (2) covalent tagging of the proteins with isotope-coded affinity tag (ICAT) reagents followed by proteolysis of the combined labeled protein samples; and (3) isolation, identification, and quantification of the tagged peptides by multidimensional chromatography, automated tandem mass spectrometry, and computational analysis of the obtained data. The method was used to identify and determine the ratios of abundance of each of 491 proteins contained in the microsomal fractions of naïve and *in vitro*–differentiated human myeloid leukemia (HL-60) cells. The method and the new software tools to support it are well suited to the large-scale, quantitative analysis of membrane proteins and other classes of proteins that have been refractory to standard proteomics technology.

Proteins that are integral to or associated with lipid membranes perform a wide range of essential cellular functions. Pores, channels, pumps, and transporters facilitate the exchange of membrane-impermeable molecules between cellular compartments and between the cell and its extracellular environment. Transmembrane receptors sense changes in the cellular environment and, typically via associated proteins, initiate specific intracellular responses. Cell-adhesion proteins mediate cell-specific interactions with other cells and the extracellular matrix. Lipid membranes also provide a hydrophobic environment for biochemical reactions that is dramatically different from that of the cytoplasm and other hydrophilic cellular compartments[1].

Because of their accessibility, membrane proteins, in particular those spanning the plasma membrane, are also of considerable diagnostic and therapeutic importance. Antisera to proteins that are selectively expressed on the surface of a specific cell type (cluster of differentiation, CD) have been used extensively for the classification of cells and for their preparative isolation by fluorescence-activated cell sorting or related methods. Membrane proteins (as exemplified by Her2/neu), the abundance of which is modulated in the course of a disease, are commonly used as diagnostic indicators and, less frequently, as therapeutic targets[2]. A humanized monoclonal antibody (Herceptin; Genentech, Palo Alto, CA) that specifically recognizes Her2/neu receptors is the basis for a successful therapy of breast cancer[3], and antibodies to other cell-surface proteins are also undergoing clinical trials as anticancer agents[4]. Moreover, the majority of current effective therapeutic agents for diseases such as hypertension and heart disease are receptor antagonists that target and selectively modify the activity of specific membrane proteins[5]. It is therefore apparent that a general technique capable of systematically identifying membrane proteins and of accurately detecting quantitative changes in the membrane protein profiles of different cell populations or tissues would be of considerable importance for biology and for applied biomedical research.

A main objective of proteomics research is the systematic identification and quantification of the proteins expressed in a cell, tissue, or body fluid. The standard approach to proteomics has been the combination of high-resolution two-dimensional gel electrophoresis (2D-E) and mass spectrometry. Proteins in complex protein mixtures are separated by 2D-E, detected by staining, and quantified based on their staining intensity. Proteins are then cut out of the gel, digested, and identified by one of several mass spectrometric techniques[6].

In spite of significant incremental progress[7–11], the separation of intramembrane, membrane-associated, and other hydrophobic proteins by 2D-E has remained a significant challenge, mainly because of the difficulty of maintaining protein solubility during 2D-E. Using an alternative approach—the combination of SDS–PAGE for the one-dimensional separation of membrane proteins and microcapillary liquid chromatography–electrospray ionization–tandem mass spectrometry (μLC-ESI-MS/MS) for the analysis of the peptides generated by digesting the proteins migrating to a particular zone of the gel—Simpson and co-workers identified 284 proteins, including 92 membrane proteins[12]. Although this method is suitable for cataloguing proteins contained in membrane fractions, it is inherently not quantitative and therefore not suitable for the detection of differences in the membrane-protein profile of cells representing different states.

In this report we describe the systematic identification and the quantitative analysis of differentiation-induced changes of proteins in the microsomal fraction of HL-60 cells using ICAT reagents, automated peptide tandem mass spectrometry, and a suite of new software tools for data analysis.

[1]*University of Connecticut Health Center, 263 Farmington Ave., Farmington, CT 06030-0002.* [2]*Institute for Systems Biology, 4225 Roosevelt Way NE, Seattle, WA 98105-6099.* **Corresponding author (raebersold@systemsbiology.org).*

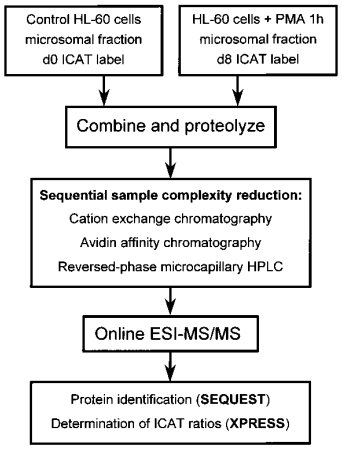


Figure 1. Schematic representation of the quantitative proteomics procedure. Proteins in the microsomal fraction of naïve or PMA-treated HL-60 cells were labeled with ICAT reagents, combined, and analyzed as described in the text.

Results and discussion

The human cell line HL-60 is a well-characterized *in vitro* model for the study of cellular differentiation[13]. Exposure to 12-phorbol 13-myristate acetate (PMA) induces the cell line to differentiate from a nonadherent state into a monocyte-like, adherent, spreading, and morphologically altered state[13]. Anticipating that the induced morphological changes, particularly increased cell adherence, would be correlated with changes in the cell-surface protein profile, we isolated the microsomal fraction from naïve and PMA-treated HL-60 cells and subjected them to a protocol for quantitative proteome analysis. The procedure is schematically illustrated in Figure 1.

Cell lysates from 1×10^8 naïve (control) or PMA-treated HL-60 cells were fractionated by differential ultracentrifugation and the microsomal fractions were isolated following a standard procedure, as described in the Experimental Protocol[14]. The samples were redissolved in a buffer containing 0.5% SDS (wt/vol), and 100 µg total protein from each sample were reduced and labeled with the isotopically normal (d0, control) or heavy (d8, PMA-treated sample) form of the sulfhydryl-specific ICAT reagent[15]. The labeled samples were combined, digested with trypsin, and the resulting peptide mixture separated by multidimensional chromatography. First, cation-exchange chromatography was used to separate the sample into 30 fractions and remove remaining proteases, ICAT reagent, and SDS. Second, individual fractions were subjected to affinity chromatography on a small monomeric avidin column. Finally, recovered biotinylated peptides were separated and analyzed by µLC-ESI-MS/MS. The need for the extensive fractionation before MS/MS for such complex peptide mixtures is well illustrated by the data shown in Figure 2. The chromatogram of peptides eluting from the cation-exchange column, with the collected fractions indicated, is given in Figure 2A, while Figure 2B shows the chromatogram of base peaks obtained from the µLC-MS analysis of the peptides contained in ion-exchange chromatography fraction 18, indicating the presence of numerous compounds in the sample. In Figure 2D appear the base peak chromatograms for all the fractions collected from the cation-exchange column, indicating the distribution of the large number of peptides detected in the sample over the available chromatographic space. Figure 2C shows the mass spectrum of the analytes eluting from the reverse-phase column separating the peptides in fraction 18 during the 30 s period indicated in Figure 2B. ICAT-labeled peptides were apparent as paired signals with a mass differential of 8 mass units for singly charged peptides. Numerous peptide peaks were detected, most of which appeared paired with another signal. The automated precursor ion selection afforded by the mass spectrometer, together with the exclusion of peptides previously sequenced within a user-defined time window (dynamic exclusion), permitted the analysis by collision-induced dissociation (CID) of 12 of the peptides eluting in the 30 s time window, which are marked by asterisks. The CID spectra of the two peptide ions that are numbered 1 and 2 in Figure 2C, their amino acid sequences, and the ratios of abundance of their parent proteins in the naïve and PMA-treated cells are shown, respectively, in Figure 3A and B. The two peptides differed in their chromatographic retention time by 4 s, and their parent proteins were identified by sequence database searching as the transmembrane tyrosine phosphatase CD45 and the calcium pump ATC2, and the calculated ratios of d0:d8 (control:stimulated) as 1:0.72 and 1:1.2, respectively. From these data, it is clear that the analysis of a peptide mixture as complex as a tryptic digest of microsomal fractions from human cells with less prefractionation would have resulted in the omission of a large number of peptides from CID because of the inability of the mass spectrometer to select for fragmentation of all or even most of the peptide ions coeluting in any chromatographic time window.

The analysis by µLC-ESI-MS/MS of all 30 ion-exchange fractions generated a total of 25,891 CID spectra. Figure 3C represents a flowchart indicating the primary CID steps required to process the spectra to identify and quantify microsomal proteins. Because analysis of each ion-exchange fraction takes 100 min of automated sequencing time, it took a total of 50 h of analysis time to complete 30 fractions. Searching of the National Cancer Institute human sequence database (consisting at the time of 73,656 protein sequences) using the SEQUEST (refs 16, 17) analysis software identified in excess of 5,000 cysteine-containing peptides. Conservative analysis of each of these identified peptides using the stringent criteria described in the Experimental Protocol, as well as by manual inspection of each spectrum, resulted in 491 conclusively identified proteins. From the complexity of data generated by the experiment, it was immediately apparent that new software tools to quantify the proteins and to compile the results were essential. We therefore created two new tools named XPRESS and INTERACT to quantify and organize the peptides identified by the SEQUEST sequence database searching software.

First, we analyzed the reproducibility of the d0:d8 ratios for different peptides from a parent protein, which is likely a good indicator of the potential of the method to quantify peptides in such complex mixtures. As an example, all the CD45-derived peptides detected in the study are shown in Figure 4. CD45 contains 23 cysteine residues, of which 13 were detected and quantified as individual cysteine-containing peptides. Their respective positions in the CD45 sequence are shown in Figure 4A. These 13 peptides were each detected and analyzed between 1 and 10 times, representing a collective total of 65 independent measurements. They eluted in ion-exchange fractions spanning a significant part of the separation range (Fig. 4B; first experiment). The data indicated variability in quantification of <20%. Furthermore, consistent variability was found when two additional small-scale experiments were carried out in a single LC-MS/MS analysis for each experiment (Fig. 4B). This extraordinary precision of the measurements can only be realized if the samples in different experiments to be compared are highly reproducible.

Because this study is the first large-scale analysis of the membrane proteome from the HL-60 human cell line, we could not conclude that a complete microsomal fraction was identified and

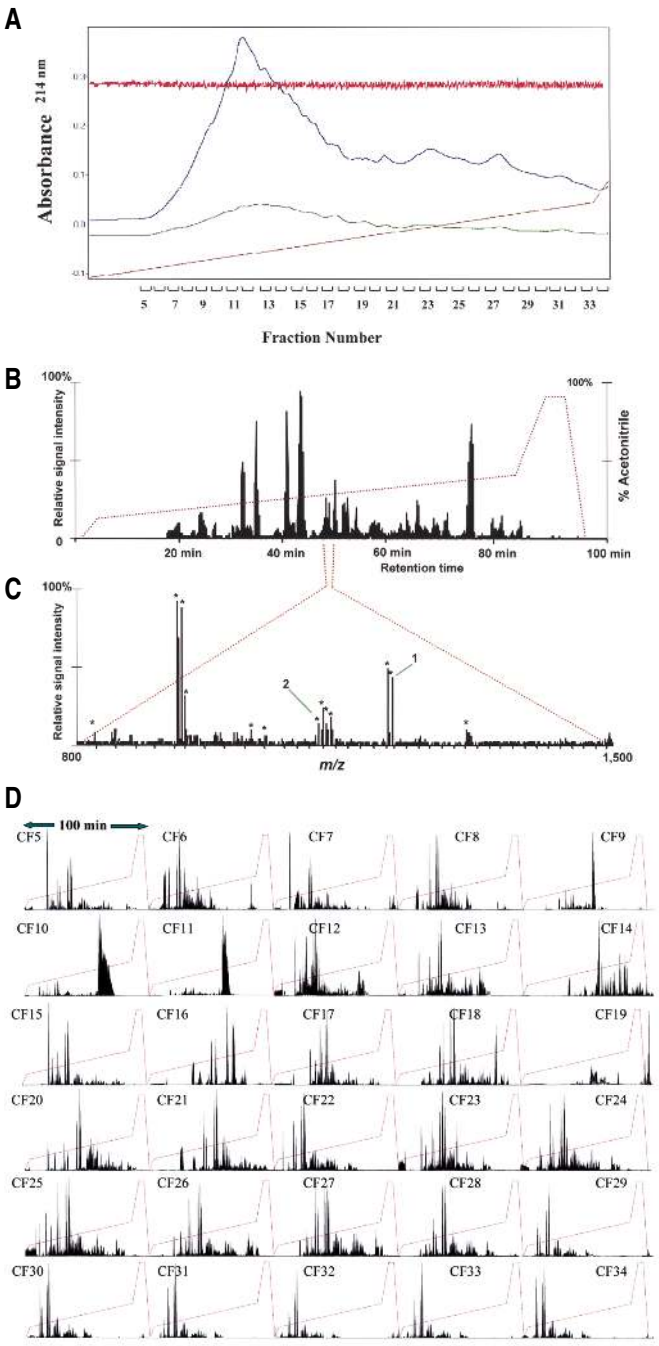


Figure 2. Multidimensional liquid chromatography tandem mass spectrometric analysis of a complex peptide mixture. (A) Distribution of peptides contained in tryptic-digested HL-60 microsomal fraction on a strong cation-exchange chromatography column. Peptides were detected by absorbance at 214 nm (blue line) and 280 nm (green line). Solvent gradient (red line), and pressure (pink line) are also indicated. Collected fraction numbers are shown on the x-axis. (B) Analysis of the biotinylated, cysteine-containing peptides contained in cation-exchange fraction 18 by μLC-ESI-MS/MS. Ion chromatogram displaying the base peak (most intense ion signal in each MS scan) as a function of retention time. Dotted line indicates the percentage acetonitrile solvent gradient used to develop the reverse-phase capillary column. (C) MS spectrum of peptides detected in the 30 s time window indicated in (B). Signals indicated with asterisk (*) were ICAT-labeled peptides that were subjected to automated CID. The data obtained from the peptides numbered as 1 and 2 are shown in Figure 3. (D) Base peak ion chromatogram of all of the cation-exchange fractions collected, indicating ICAT peptide distribution. Red line indicates solvent gradient.

quantified. Indeed, a subset of proteins that lack cysteine residues, very low abundance proteins, and very hydrophobic proteins will not be analyzed using this technique. The advantage of multidimensional ICAT analysis is that the complexity of the peptide mixture is reduced, and thus, analysis of a higher number of unique proteins and associated quantification is achieved for a given time during the automated LC-MS/MS procedure. Because the technique selects for peptides that contain free cysteine residues before labeling, it was relatively rare to find post-translationally modified peptides. However, this lack of post-translationally modified peptides could easily be overcome by simply analyzing avidin flowthrough peptides separately in a 2D-chromatography approach as first described by Yates and colleagues[17], thus allowing better coverage as well as providing quantitative information.

From our analysis, it is apparent that ICAT analysis yields maximal information for protein quantification, whereas studies that are targeted toward post-translational modifications such as protein phosphorylation would require enrichment steps, as we have demonstrated[18]. We imagine that different types of experimentally measured information would be integrated in a bioinformatics platform such as INTERACT for biologically meaningful interpretation.

Sequence database searching with the CID spectra obtained by this study resulted in ~5,000 identified, cysteine-containing peptides that matched 1,025 proteins. Of these, the 491 proteins listed in Supplementary Table 1 and annotated with their identity, database access code, and d0:d8 ratio in naïve and PMA-treated cells were considered conclusively identified and quantified because, for each of these proteins, at least two different peptides were identified with highly significant database matching scores and consistent quantification (see Supplementary Tables 1 and 2 in the Web Extras page of *Nature Biotechnology* Online). Grouping of the conclusively identified and quantified proteins by their presumed functions (Fig. 5) confirmed the expected prevalence of transmembrane and membrane-associated proteins in the microsomal fractions analyzed. Furthermore, the study identified a number of proteins presumed to be of moderate to low abundance, including regulatory proteins (protein kinases, phosphatases, GTPases), membrane receptors, and channels (see Supplementary Tables 1 and 2).

For most proteins identified in the microsomal fraction, the abundance in naïve and PMA-treated cells did not change significantly. Unchanged abundance was particularly prevalent for ribosomal proteins, cytoskeletal proteins, metabolic enzymes, and the majority of cell-surface receptors and channel proteins (see Supplementary Tables 1 and 2 in the Web Extras page of *Nature Biotechnology* Online). In contrast, some of the membrane-associated signal transduction proteins and enzymes showed significant changes in their d0:d8 ratios, pointing toward regulatory mechanisms that are directly or indirectly consequential to the PMA treatment. For example, the PMA-induced, coordinated downregulation of farnesyl-diphosphate farnesyl-transferase (squalene synthetase, SQS, FTFD_HUMAN) (refs 19–24) (20-fold reduction) and of a number of proteins associated with the plasma membrane via lipid anchors (see Supplementary Tables 1 and 2) is compatible with a model in which localization and recruitment of proteins to and from the membrane via protein prenylation is controlled by farnesyl pyrophosphate (FPP) utilization. SQS represents a branchpoint in the mevalonate pathway[21] that plays a key role in proportioning the available FPP pool either for cholesterol/sterol biosynthesis or for the synthesis of nonsteroidal products such as the precursors of protein prenylation, a post-translational modification that typically results in the localization of the modified protein to membranes[25]. Activation of SQS, which has been shown to

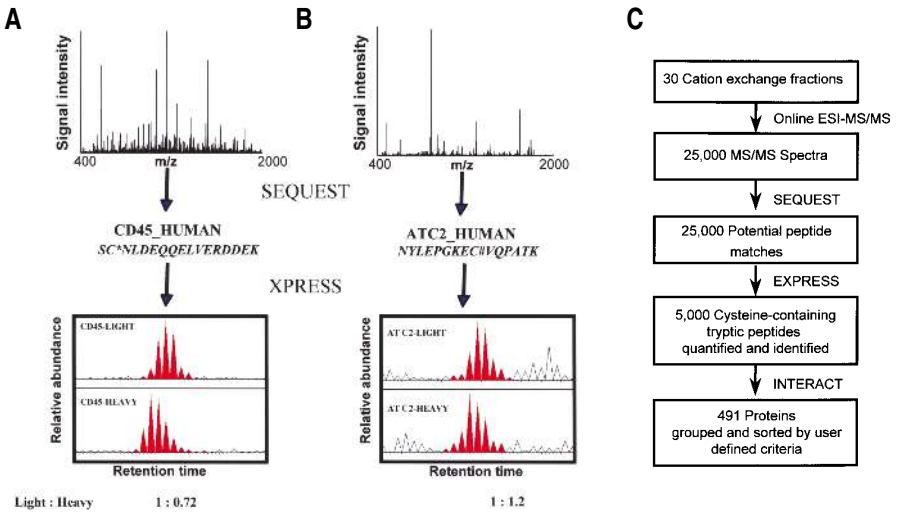


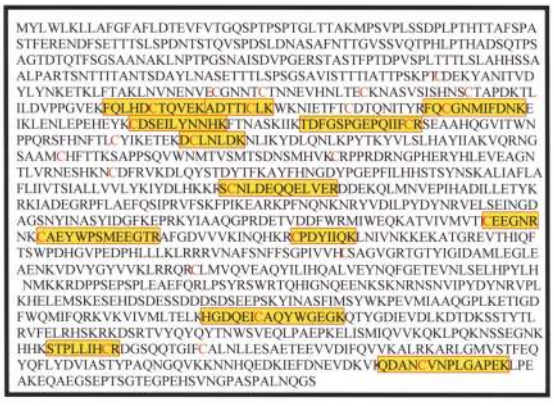
Figure 3. Post-MS processing of data from selected peptides. Two peptides numbered as 1 (A) and 2 (B) in Figure 2C were subjected to alternating MS and MS/MS scans. For each peptide we show the CID spectrum, the identity of the parent protein determined by sequence database searching using SEQUEST[16] (C* designates cysteine labeled with the heavy form of ICAT reagent, while C# designates cysteine labeled with the light form of ICAT reagent), the data indicating the relative abundance, and the calculated d0:d8 ratio obtained using XPRESS software. (C) Flowchart indicating the steps applied to process the initial CID spectra to identify and quantify microsomal proteins.

coincide with proteolysis-induced release from the membrane[22,26], consumes the available FPP for the generation of squalene, a precursor of cholesterols and sterols[21]. Thus, activation of this enzyme could be associated with a depletion of FPP available for protein prenylation, thus providing a possible explanation for the observed PMA-induced downregulation of prenylated proteins such as RhoG and Rac1 in the microsomal fraction, coincidental with the disappearance of microsomal SQS. The central role of SQS in prenyl metabolism is supported by observations from the inflammatory response. Downregulation of SQS in response to cytokines and endotoxin and the coordinated increase in HMG-CoA reductase activity have been suggested to result in an increased pool of isoprenoid intermediates for acute inflammatory response[27].

Generally, we observed a PMA-induced reduction of the Src family kinases, PKC isoforms, small GTPases, and phosphatases present in the microsomal fractions of PMA-treated HL-60 cells (see Supplementary Tables 1 and 2 in the Web Extras page of *Nature Biotechnology* Online). However, specific members of these enzyme families showed a PMA-induced increase in microsomal abundance. Specifically, we detected increases in the microsomal abundance of PKC-theta, Yes, and the phosphatases PTN-1 and PTN-3. Conversely, PKC-α and -β1, Lck, Src itself, and the phosphatases PP1A and PTN-2 showed decreased microsomal concentrations after PMA treatment. These results indicate the presence of previously unidentified, isoform-specific mechanisms that control *de novo* synthesis and/or intracellular distribution of regulatory proteins during cellular attachment and differentiation in HL-60 cells[28]. While it is well established that PMA-induced activation of PKC and the consequential initiation of signal transduction events result in changes in cell-surface properties, cellular attachment, and the subsequent differentiation of HL-60 cells[29–31], the specific enzyme isoforms that orchestrate these complex sequence of events are not known. The divergent patterns of the PKC-theta compared with the PKC-α and -β1, together with analogous observations made for members of the Src protein tyrosine kinase family and protein phosphatases, suggest new directions for probing the mechanism of PMA-induced differentiation of HL-60 cells.

As with any discovery science method[32], the data obtained by this study are not by themselves sufficient to explain the mechanisms of PMA-induced changes in HL-60 cells, such as cell differentiation and adherence. The data do, however, implicate new proteins in these processes and suggest hypotheses that can subsequently be tested by hypothesis-driven research methods or by further discovery science experiments. The data obtained by our method also do not, by themselves, explain the mechanisms by which protein concentrations in the microsomal fractions change. Changes in the rates of synthesis or degradation, the intracellular redistribution of a constant protein pool, or a combination of

A



B

First experiment

CD 45 Peptide identified	Cation fraction number	Number of times identified	Quantification (d0: d8)	Standard deviation
FQLHDCTQVEK	21, 22	4	1 : 0.86	0.04
ADTTICLK	6, 7	4	1 : 0.83	0.07
FQCGNMIFDNK	9, 10	7	1 : 0.78	0.03
CDSEILYNNHK	20, 21	8	1 : 0.75	0.05
TDFGSPGEPQIIFCR	7, 8	2	1 : 0.74	0.05
ETEKDCLNLDK	20	2	1 : 0.83	0.04
SCNLDEQQELVER	18	4	1 : 0.84	0.14
CEEGNR	12	1	1 : 0.75	0.0
CAEYWPSMEEGTR	8, 9	8	1 : 0.75	0.05
CPDYIIQK	8, 9	5	1 : 0.73	0.04
HGDQEICAQYWGEGK	23, 24	10	1 : 0.72	0.07
STPLLIHCR	21, 22	7	1 : 0.71	0.09
QDANCVNPLGAPEK	17	3	1 : 0.80	0.14
		mean ratio	1:0.77	0.05

Second experiment

ADTTICLK	N/A	2	1: 0.81	N/A
FQCGNMIFDNK	N/A	1	1:0.79	N/A
CDSEILYNNHK	N/A	1	1:0.88	N/A
TDFGSPGEPQIIFCR	N/A	1	1:0.78	N/A
SCNLDEQQELVERDDEK	N/A	2	1:0.86	N/A
CAEYWPSMEEGTR	N/A	1	1:0.77	N/A
HGDQEICAQYWGEGK	N/A	1	1:0.83	N/A

Third experiment

ADTTICLK	N/A	2	1:0.82	N/A
CDSEILYNNHK	N/A	2	1:0.81	N/A
SCNLDEQQELVERDDEK	N/A	1	1:0.79	N/A
TDFGSPGEPQIIFCR	N/A	1	1:0.84	N/A
HGDQEICAQYWGEGK	N/A	2	1:0.78	N/A

Figure 4. Consistency of redundant protein quantification. A total of 13 peptides from the CD45 transmembrane protein tyrosine phosphatase CD45 were identified and quantified. (A) CD45 protein sequence with the identified, ICAT reagent-labeled peptides indicated in boxes with yellow background. (B) Redundancy and accuracy of quantitation: the amino acid sequence of the identified peptides, the cation-exchange fraction(s) in which each peptide was detected, the number of times each peptide was identified and quantified, and the abundance ratio and its standard deviations are indicated. Results are shown from three independent experiments: one large scale (first experiment) and two smaller scale (second and third experiments).

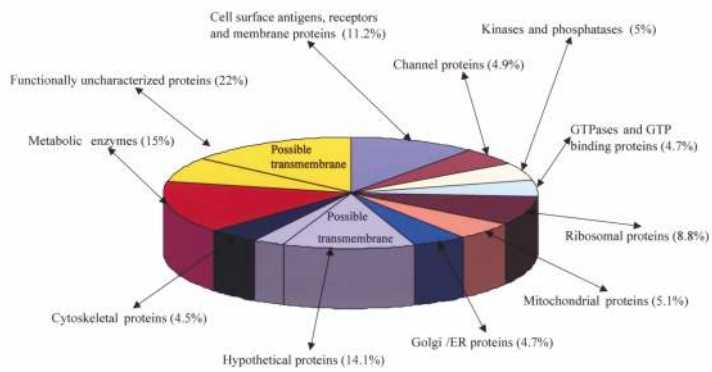


Figure 5. Categories of proteins identified in this study. The 491 proteins identified and quantified in this study were classified by broad functional criteria. The numbers in parentheses indicate the percentage fraction of identified proteins represented by each category. Some proteins are represented in more than one category (e.g., CD45 was counted as both a transmembrane protein and a protein phosphatase). Protein sequences from hypothetical proteins (81% containing possible transmembrane α-helices) and functionally uncharacterized proteins (78% containing possible transmembrane α-helices) were analyzed for the possible presence of transmembrane helices using the Tmpred tool available on EXPASY website (http://www.ch.embnet.org/software/TMPRED_form.html).

these factors could lead to the same apparent result. More extensive experiments using ICAT protein profiling that include additional subcellular fractions and additional time-resolved observations would be expected to differentiate the factors underlying the observed changes. In addition and in contrast to membrane-protein profiling methods based on the quantitative analysis of messenger RNAs (mRNAs) associated with the rough endoplasmatic reticulum[33] or the prediction of membrane association by computational analysis, direct proteomic analyses of microsomal fractions are able to identify proteins that are membrane-associated by post-translational mechanisms such as acylation or prenylation, or by protein–protein interactions, and to quantify changes in their abundance.

Conclusions. In this study, we have unambiguously identified 491 proteins and determined their relative abundance in the microsomal fractions of naïve and *in vitro*–differentiated human HL-60 cells. This represents the largest number of membrane or membrane-associated proteins identified to date and the first study in which the relative abundance of such proteins was directly and systematically determined. The data were generated by an approach consisting of post-isolation protein labeling with ICAT reagents, multidimensional chromatography of complex peptide mixtures, automated tandem mass spectrometry, and a suite of software tools for data analysis and interpretation. The successful application of this strategy to a class of proteins that to date have proved particularly difficult to analyze suggests that quantitative protein profiling of microsomal fractions is now technically feasible. Future comparison of identical membrane protein samples with 2D-E and 3D-ICAT will point out the advantages of each method and will be very useful for the field of proteomics[34].

Experimental protocol

Preparation of microsomal fractions, ICAT labeling, and peptide separation. HL-60 cells were grown in RPMI medium supplemented with 10% FBS to a density of 1×10^6 cells/ml, and microsomal fractions were prepared according to Walter and Blobel[14]. Equal numbers of cells (1×10^8) were either mock-treated or treated with 10 nM PMA for 1 h at 37°C. Cells were then washed with ice-cold PBS, swelled for 20 min, and lysed in a hypotonic buffer (20 mM Tris-HCl, pH 7.4, 10 mM $MgCl_2$, 10 mM $CaCl_2$) supplemented with protease inhibitor cocktail (Boehringer Mannheim, Indianapolis, IN). Cells were then spun for 15 min at 3,000 *g* to remove nuclei and cytoskeletal components. The supernatant was further spun for 2 h at 100,000 *g* to pellet the microsomal fraction. The microsomal pellet contained plasma membrane, golgi, endoplasmic reticulum, mitochondria, lysosomes, and all other membrane-bound vesicles separated from soluble cytosol[14]. Western analysis of known membrane, mitochondria-associated proteins, and known cytosolic proteins in identically prepared subcellular fractions from human lymphocytes indicated at least 20-fold enrichment of membrane proteins in the microsomal fraction compared to total cell lysates (Han *et al.*, unpublished results). Resulting pellets were dissolved by boiling in labeling buffer (50 mM Tris-HCl pH 8.3, 5 mM EDTA, 0.5% SDS), and the protein concentration was measured.

Equal amounts of protein from the control or PMA-treated cells (100 μg each) were then reduced with 5 mM tributylphosphine for 30 min at 37°C. Proteins from the control cells and PMA-treated sample were labeled with isotopically light- or heavy-ICAT reagent (200 nM each for 5- to 10-fold molar excess) for 90 min at room temperature with gentle stirring. Reagents and detailed protocol for ICAT labeling and analysis can be obtained from Applied Biosystems as a Commercial Kit (ABI, Framingham, MA). Labeled samples were mixed, diluted 10-fold, and digested with trypsin. The resulting peptide mixture was fractionated by three-step chromatography: (1) cation-exchange chromatography using a 2.1 mm × 20 cm Polysulfoethyl A column (Poly LC Inc., Columbia, MD) at a flow rate of 200 μl/min, buffer A, 20 mM KH_2PO_4, 25% acetonitrile, pH 3.0; buffer B, 20 mM KH_2PO_4, 350 mM KCl, 25% acetonitrile, pH 3.0; (2) affinity chromatography using a self-packed ultralink-monomeric avidin column (Pierce, IL) using 400 μl of packed beads in a glass Pasteur pipette; and (3) reverse-phase capillary chromatography using a 75 μm × 10 cm self-packed C18 (Monitor, Column Engineering, Ontario, CA) column at a flow rate of 250 nl/min. Peptide identification by CID was carried out in an automated fashion using the dynamic-exclusion option on a Finnigan LCQ ion trap mass spectrometer followed by automated data processing for protein identification (SEQUEST)[16] and quantitation (XPRESS). A human protein database (73,656 entries) was used to search the uninterpreted MS/MS spectra.

SEQUEST scoring criteria. Uninterpreted MS/MS spectra were searched against the National Center for Biotechnology Information (NCBI) human protein database; peptides that showed SEQUEST scores >2, and delta correlation scores of ≥0.1 were further analyzed by detailed spectral analysis. Because we were selecting for ICAT labeling, we expected cysteine-containing peptides to be identified in the sequence database search. If SEQUEST identified the peptides as either heavy- (d8) or light- (d0) labeled, we first manually inspected raw MS/MS spectra by looking for characteristic ICAT fragments (284 mass units for light- and 288 mass units for heavy-ICAT reagent). We then manually inspected the match of the major product ions with theoretically predicted product ions from the database-matched peptides. Furthermore, if the above criteria showed confirmation of either heavy- or light-ICAT peptides, we examined the chromatography profiles of heavy- and light-ICAT peptides, because heavy peptides elute 1–2 s earlier from the HPLC column. If all above criteria were fulfilled (SEQUEST score, delta correlation score, presence of cysteine in peptide sequence, ICAT fragments matching, spectral matching, and elution profile matching), we scored the peptide as conclusively identified. SEQUEST search parameters were set as follows: static modification was set to +442.2244 for ICAT-labeled cysteine, differential modification of +8 for heavy ICAT-labeled cysteine residues, +16 for oxidized methionine, +80.00 for phosphorylated serine, threonine, and tyrosine residues. We found a number of peptides with oxidized methionine residues and one phosphopeptide in our analysis.

EXPRESS and INTERACT software tools. The XPRESS software automates protein expression calculations by accurately quantifying the relative abundance of ICAT-labeled peptides from their chromatographic coelution profiles. Starting with the peptide identification, XPRESS isolates the d0 and d8 peptide elution profiles, determines the area of each peptide peak, and calculates the abundance ratio based on these areas. Furthermore, the program provides a graphical, interactive interface to the results.

The task of analyzing and reassembling the tens of thousands of data points generated in the course of this study was labor-intensive. We therefore developed a software tool termed INTERACT that greatly facilitates this process. INTERACT allows multiple data sets to be analyzed using a form-

based HTML interface interacting with a web-server common gate interface (CGI) program. INTERACT is a program written to allow the interactive analysis of SEQUEST MS/MS database search results. INTERACT takes in a single or multiple SEQUEST Summary HTML files as input and allows the user to quickly filter and sort the search results using various user-controlled criteria. The program is composed of two parts: a stand-alone executable program that reads in the Summary HTML file inputs and builds the initial data set, and a web-based CGI program that performs the interactive analysis with the user. Both parts of the program are written in the C programming language and are completely independent of either SEQUEST or the Summary program (besides using the Summary program output). The purpose of the INTERACT program is to allow a researcher to weed through the various search results present in a single or multiple SEQUEST searches.

A SEQUEST search result is typically analyzed through the Summary web page, where each search result is listed as a single line entry on the page. However, the Summary web page is a static page that does not allow the user flexibility in terms of filtering, sorting, and deleting any of the entries of the search results. INTERACT enables interactive analysis by parsing the Summary file(s), stripping out HTML tags as necessary, and treating each line in each Summary file as a data point. When the program is run, it will generate a web-based form that the user views in a web browser. The generated web page is composed of two sections or frames: the control form in the top frame and the data in the bottom frame. The control form has entry fields that allow the user to filter the data on the basis of scores, amino acid composition, protein descriptions, and enzyme digest specificity. The control form also allows individual search results to be removed or deleted from the analysis and specific residues to be highlighted. Additionally, a single last-step undo control and a global undo control are available. The current state of the filtered data and the undo steps are available when the web page is closed and reopened at a later time. The HTML links in each Summary line entry are always maintained; this allows the user to use the existing Summary CGI tools, such as the spectrum viewer, on the data set at any time. With the ability to interactively work with the search results in this matter, users are able to quickly filter the data set into more meaningful results. With INTERACT, a user can query, filter, and sort data rapidly. For the present study, application of INTERACT resulted in the three-dimensional information of the dynamically changing microsomal proteome, including (1) identity of proteins (2) relative quantity of proteins, and (3) localization of proteins during cellular attachment and differentiation of HL-60 cells. The XPRESS and INTERACT software tools will be described in more detail elsewhere.

Note: Supplementary information can be found on the Nature Biotechnology *website in Web Extras (http://biotech.nature.com/web_extras).*

Acknowledgments

We thank John Glomset, Julian Watts, Steve Gygi, and Beate Rist for helpful discussion and Julian Watts for critical reading of the manuscript. The work was supported by a grant from the University of Washington Royalty Research Fund, a grant from the Merck Genome Research Institute, grant no. 1R33CA84698 from the National Cancer Institute, and grant no. HL67569 from the National Institutes of Health.

Received 20 March 2001; accepted 14 August 2001

1. Oseroff, A.R., Robbins, P.W. & Burger, M.M. Cell surface membrane: biochemical aspects and biophysical probes. *Annual Rev. Biochem.* **42**, 647–682 (1973).
2. Liu, E. *et al.* The HER2 (c-erbB-2) oncogene is frequently amplified in *in situ* carcinomas of the breast. *Oncogene* **7**, 1027–1032 (1992).
3. Shak, S. Overview of the trastuzumab (Herceptin) anti-HER2 monoclonal antibody clinical program in HER2-overexpressing metastatic breast cancer. Herceptin Multinational Investigator Study Group. *Semin. Oncol.* **26**, 71–77 (1999).
4. Eccles, S.A. Monoclonal antibodies targeting cancer: "magic bullets" or just the trigger? *Breast Cancer Res.* **3**, 86–90 (2001).
5. Drews, J. Research and development. Basic science and pharmaceutical innovation. *Nat. Biotechnol.* **17**, 406–408 (1999).
6. Aebersold, R. & Goodlett, D.R. Mass spectrometry in proteomics. *Chem. Rev.* **101**, 269–295 (2001).
7. Molloy, M.P. *et al.* Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. *Electrophoresis* **19**, 837–844 (1998).
8. Molloy, M.P. *et al.* Proteomic analysis of the *Escherichia coli* outer membrane. *Eur. J. Biochem.* **267**, 2871–2881 (2000).
9. Santoni, V., Molloy, M. & Rabilloud, T. Membrane proteins and proteomics: un amour impossible? *Electrophoresis* **21**, 1054–1070 (2000).
10. Rabilloud, T. *et al.* Analysis of membrane proteins by two-dimensional electrophoresis: comparison of the proteins extracted from normal or *Plasmodium falciparum*-infected erythrocyte ghosts. *Electrophoresis* **20**, 3603–3610 (1999).
11. Santoni, V. *et al.* Large scale characterization of plant plasma membrane proteins. *Biochimie* **81**, 655–661 (1999).
12. Simpson, R.J. *et al.* Proteomic analysis of the human colon carcinoma cell line (LIM 1215): development of a membrane protein database. *Electrophoresis* **21**, 1707–1732 (2000).
13. Birnie, G.D. The HL60 cell line: a model system for studying human myeloid cell differentiation. *Br. J. Cancer Suppl.* **9**, 41–45 (1988).
14. Walter, P. & Blobel, G. Preparation of microsomal membranes for cotranslational protein translocation. *Methods Enzymol.* **96**, 84–93 (1983).
15. Gygi, S.P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
16. Eng, J., McCormack, A.L. & Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass. Spectrom.* **5**, 976–989 (1994).
17. Link, A.J. *et al.* Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**, 676–682 (1999).
18. Zhou, H., Watts, J.D. & Aebersold, R. A systematic approach to the analysis of protein phosphorylation. *Nat. Biotechnol.* **19**, 375–378 (2001).
19. Stamellos, K.D. *et al.* Subcellular localization of squalene synthase in rat hepatic cells. Biochemical and immunochemical evidence. *J. Biol. Chem.* **268**, 12825–12836 (1993).
20. Szkopinska, A., Swiezewska, E. & Karst, F. The regulation of activity of main mevalonic acid pathway enzymes: farnesyl diphosphate synthase, 3-hydroxy-3-methylglutaryl-CoA reductase, and squalene synthase in yeast *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **267**, 473–477 (2000).
21. Goldstein, J.L. & Brown, M.S. Regulation of the mevalonate pathway. *Nature* **343**, 425–430 (1990).
22. Tansey, T.R. & Shechter, I. Squalene synthase: structure and regulation. *Prog. Nucleic Acid Res. Mol. Biol.* **65**, 157–195 (2000).
23. Gu, P., Ishii, Y., Spencer, T.A. & Shechter, I. Function–structure studies and identification of three enzyme domains involved in the catalytic activity in rat hepatic squalene synthase. *J. Biol. Chem.* **273**, 12515–12525 (1998).
24. Chen, H.T., Mehan, R.S., Gupta, S.D., Goldberg, I. & Shechter, I. Involvement of farnesyl protein transferase (FPTase) in Fc epsilonRI-induced activation of RBL-2H3 mast cells. *Arch. Biochem. Biophys.* **364**, 203–208 (1999).
25. Yokoyama, K., Goodwin, G.W., Ghomashchi, F., Glomset, J. & Gelb, M.H. Protein prenyltransferases. *Biochem. Soc. Trans.* **20**, 489–494 (1992).
26. Shechter, I. *et al.* Solubilization, purification, and characterization of a truncated form of rat hepatic squalene synthetase. *J. Biol. Chem.* **267**, 8628–8635 (1992).
27. Memon R.A. *et al.* Endotoxin, tumor necrosis factor, and interleukin-1 decrease hepatic squalene synthase activity, protein, and mRNA levels in Syrian hamsters. *J. Lipid Res.* **38**, 1620–1691 (1997).
28. Kiss, Z., Deli, E. & Kuo, J.F. Temporal changes in intracellular distribution of protein kinase C during differentiation of human leukemia KL60 cells induced by phorbol ester. *FEBS Lett.* **231**, 41–46 (1988).
29. Seibenhener, M.L. & Wooten, M.W. Heterogeneity of protein kinase C isoform expression in chemically induced HL60 cells. *Exp. Cell. Res.* **207**, 183–188 (1993).
30. Wooten, M.W., Seibenhener, M.L. & Soh, Y. Expression of protein kinase C isoforms in HL60 and phorbol ester resistant HL525 cells. *Cytobios.* **76**, 19–29 (1993).
31. Ohguchi, K., Banno, Y., Nakashima, S. & Nozawa, Y. Activation of membrane-bound phospholipase D by protein kinase C in HL60 cells: synergistic action of a small GTP-binding protein RhoA. *Biochem. Biophys. Res. Commun.* **211**, 306–311 (1995).
32. Aebersold, R., Hood, L.E. & Watts, J.D. Equipping scientists for the new biology. *Nat. Biotechnol.* **8**, 359 (2000).
33. Diehn, M., Eisen, M.B., Botstein, D. & Brown, P.O. Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. *Nat. Genet.* **25**, 58–62 (2000).
34. Gygi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y. & Aebersold, R. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. USA* **97**, 9390–9395 (2000).