Expression Clustering Reveals Detailed Co-expression Patterns of Functionally Related Proteins during B Cell Differentiation

A PROTEOMIC STUDY USING A COMBINATION OF ONE-DIMENSIONAL GEL ELECTROPHORESIS, LC-MS/MS, AND STABLE ISOTOPE LABELING BY AMINO ACIDS IN CELL CULTURE (SILAC)*⑤

Edwin P. Romijn‡§¶, Chantal Christis§‖, Marnix Wieffer**, Joost W. Gouw‡, Asier Fullaondo‡ ‡‡, Peter van der Sluijs**, Ineke Braakman‖, and Albert J. R. Heck‡§§

B cells play an essential role in the immune response. Upon activation they may differentiate into plasma cells that secrete specific antibodies against potentially pathogenic non-self antigens. To identify the cellular proteins that are important for efficient production of these antibodies we set out to study the B cell differentiation process at the proteome level. We performed an in-depth proteomic study to quantify dynamic relative protein expression patterns of several hundreds of proteins at five consecutive time points after lipopolysaccharide-induced activation of B lymphocytes. The proteome analysis was performed using a combination of stable isotope labeling using [$^{13}C_6$]leucine added to the murine B cell cultures, one-dimensional gel electrophoresis, and LC-MS/MS. In this study we identified 1,001 B cell proteins. We were able to quantify the expression levels of a quarter of all identified proteins (*i.e.* 234) at each of the five different time points. Nearly all proteins revealed changes in expression patterns. The quantitative dataset was further analyzed using an unbiased clustering method. Based on their expression profiles, we grouped the entire set of 234 quantified proteins into a limited number of 12 distinct clusters. Functionally related proteins showed a strong correlation in their temporal expression profiles. The quality of the quantitative data allowed us to even identify subclusters within functionally related classes of proteins such as in the endoplasmic reticulum proteins that are involved in antibody production.   *Molecular & Cellular Proteomics 4:1297–1310, 2005.*

B cells play a vital role in the immune response as they produce antibodies that recognize specific antigens that are foreign to the body and potentially pathogenic. In the absence of pathogens B lymphocytes do not secrete antibodies. Resting cells have a small cytoplasm with scarce endoplasmic reticulum (ER)[1] cisternae. Upon encounter with antigens, B lymphocytes start to proliferate rapidly and differentiate primarily into Ig-secreting plasma cells (1, 2). The differentiation process involves a major morphological change of the cell including a massive development of secretory organelles, most notably the endoplasmic reticulum (3). Because the changes in the differentiation process of B lymphocytes into plasma cells are so dramatic, it is logical that many proteins will be involved that individually may exhibit different temporal expression patterns. One of the core benefits of proteomics is the ability to systematically identify and quantify individual proteins expressed in a cell or tissue at a reasonably high throughput. Until recently, quantitative proteomics was performed mostly by a combination of two-dimensional gel electrophoresis, whereby protein expression patterns are related to spot densities on the 2D gel, followed by mass spectrometry for identification. We have used this approach to investigate the B cell differentiation process, which allowed us to identify clusters of co-expressing abundant proteins that step by step prepared the B cell for its task as an antibody-producing cell (4). In the present work we evaluated the B cell differentiation process with a comprehensive approach using a combination of isotope labeling, 1D gel electrophoresis, and mass spectrometry (5–7). We chose to use metabolic labeling with stable isotope-labeled amino acids added to the growth medium of B cells, more recently termed SILAC (8–14). As it is desirable to have as many peptide pairs as possible for quantification we selected leucine because of all tryptic pep-

[1] The abbreviations used are: ER, endoplasmic reticulum; 1D, one-dimensional; 2D, two-dimensional; LPS, lipopolysaccharide; SILAC, stable isotope labeling by amino acids in cell culture; PLGS 2.0, ProteinLynx Global Server 2.0; SOTA, Self Organizing Tree Algorithm; Cnx, calnexin; Crt, calreticulin; PDI, protein-disulfide isomerase; AG-II, α-glucosidase II; BiP, immunoglobulin heavy chain binding protein; TRAP, translocon-associated protein.

tides with five or more amino acids 68% contain leucine (9). We used $^{13}C_6$ instead of $^2H_3$ isotope-labeled leucine because it has been well documented that deuterium-labeled peptides do not necessarily co-elute in nano-LC with their unlabeled counterparts (15). Following SILAC labeling we used 1D gel electrophoresis, nano-reversed phase LC-MS, and nano-reversed phase LC-MS/MS to identify more than 1,000 B cell proteins. A quarter of these proteins (*i.e.* 234) could be quantified at each of the five time points. This extensive set of dynamic protein expression patterns was clustered based on their co-expression patterns. Identification of the proteins revealed that many functionally related proteins ended up in the same cluster because they displayed similar temporal protein expression profiles. Compared with the previous 2D gel-based approach the present SILAC/1D gel-based approach allowed us to quantify ~2.5 times as many proteins, now including more membrane proteins and less abundant proteins. Using the unbiased clustering approach and the accurate quantitative data we were able to even identify subclusters within functionally related classes of proteins such as in the ER proteins that are involved in antibody production.

EXPERIMENTAL PROCEDURES

Cell Culture, Media Composition, and Activation of B Cells—We used the I.29μ^+ (IgM, λ) lymphoma cell line as model B lymphocytes (1). Cells were cultured in suspension in RPMI 1640 medium-Leu (Invitrogen) supplemented with 10% dialyzed fetal calf serum, Glutamax (1 mM), penicillin (100 units/ml), streptomycin (100 μg/ml), sodium pyruvate (1 mM), β-mercaptoethanol (50 μM) (all from Invitrogen), and either [$^{12}C_6$]leucine (50 mg/liter, Sigma) or [$^{13}C_6$]leucine (50 mg/liter, Cambridge Isotope Laboratories, Inc., Andover, MA). The cells cultured in normal leucine-containing medium were used for differentiation. After 7 days of culture, differentiation was induced with 20 μg/ml LPS (Sigma). Cells were harvested 0, 1, 2, 3, and 4 days after activation.

Preparation of Protein Samples—Cells were counted and mixed in known ratios, which were corrected for the protein content per cell (4). We added a fixed number of labeled cells of day 0 (no activation) to a variable number of unlabeled cells of days 0, 1, 2, 3, and 4 after activation. Cells were washed twice in ice-cold 0.25% (w/v) sucrose in 25 mM HEPES-KOH (pH 7.0) and lysed in a small volume of the same solution containing 1% Triton X-100 and protease inhibitor mixture. Protein mixtures were resolved by SDS-PAGE gradient gel (7.5–15%) and Coomassie-stained. In total five lanes were loaded containing mixtures of labeled day 0 cells and unlabeled cells 0, 1, 2, 3, and 4 days after activation, respectively.

Liquid Chromatography-Mass Spectrometric Analysis—The five different 1D gel lanes were cut into 35 slices and subjected to in-gel reduction, alkylation, and tryptic digestion as described previously (4). A Voyager DE-STR (Applied Biosystems) MALDI-TOF mass spectrometer with α-cyanohydroxycinnamic acid as the matrix was used to monitor the incorporation of the [$^{13}C_6$]leucine label. For LC-MS/MS analysis a nano-LC system (Agilent 1100) was coupled to a Q-TOF-micro mass spectrometer (Micromass, Manchester, UK). Peptide mixtures were trapped on an Aqua™ C_{18} reversed phase column (Phenomenex; column dimensions, 2 cm × 100 μm; packed in-house) at a flow rate of 10 μl/min, and the peptide separation was done using an Aqua C_{18} reversed phase column (Phenomenex; column dimensions, 15 cm × 75 μm; packed in-house) and a gradient of 0–50% B in A (A = 0.1 M acetic acid; B = 80% acetonitrile, 0.1 M

acetic acid) in 90 min and at a constant flow rate of 200 nl/min. Each digest was analyzed twice, once in LC-MS mode for quantification and once in LC-MS/MS mode for identification. Fragmentation of the peptides was performed in data-dependent mode, and mass spectra were acquired in continuum mode.

Protein Identification and Data Analysis—Proteins were identified either by using ProteinLynx Global Server 2.0 (PLGS 2.0) (Micromass) or Mascot (Matrix Science, London, UK) with a non-redundant proteome set of Swiss-Prot and TrEMBL entries prepared for *Mus musculus* or Mouse International Protein Index (version 1.28) as a protein sequence database, respectively (16). Spectra were searched with a peptide mass tolerance of 1.2 Da, with fragment tolerance of 0.3 Da, and with strict trypsin specificity. A maximum of one missed cleavage was allowed, and carbamidomethylated cysteine and oxidized methionine were set as variable modifications. We identified more than 1,000 proteins using on average four peptide MS/MS spectra per protein whereby the individual Mascot score for each peptide MS/MS spectrum was at least 25. The FatiGO web tool was used to classify some of the proteins on bases of their cellular localization (17, 18). Proteins with no annotated gene ontology were classified manually using literature databases. Proteins were quantified either by using PLGS 2.0 or MSQuant. Because at present only quantification using ICAT is supported by PLGS 2.0 relative quantification was obtained by defining a variable modification of [$^{12}C_6$]leucine (1·10^{-4}-Da mass difference) and [$^{13}C_6$]leucine (6.0201-Da mass difference), and these variable modifications were used during both identifications and quantification analyses.

Statistical Cluster Analysis—Relative quantification values were used in the calculation considering day 0 as the internal standard reference. Protein ratios were normalized using the average ratio per day summed over all proteins present. In general, we observed that the protein content of the differentiated cells was much higher than that of the resting B lymphocytes. Therefore, nearly all proteins showed an absolute up-regulation in the plasma cells relative to the B lymphocytes. An apparent down-regulation of proteins reported here actually means that these proteins are less up-regulated relative to other proteins. These normalized data were used as input for The Institute of Genomic Research (TIGR) MultiExperiment Viewer (TM4) software (19). The Self Organizing Tree Algorithm (SOTA) was used to cluster the data using the following parameters: Pearson correlation as a distance with 10 cycles and a maximum cell diversity of 0.8. This algorithm is a neural network that grows adopting the topology of a binary tree, and the result of the algorithm is a hierarchical cluster obtained with the accuracy and robustness of a neural network (20). The SOTA tree summarizes the expression patterns of all proteins. Each branch represents the centroid expression profile of a group of proteins. The expression graphs, as depicted in Fig. 5, show the expression profile of each protein inside each SOTA branch. An additional clustering was done using Pearson correlation distance and average linkage to cluster proteins from the same expression group.

RESULTS

Fig. 1 provides a schematic overview of the experimental workflow followed in our experiments. First murine B cell lymphocytes were grown in media containing either normal leucine or [$^{13}C_6$]leucine for up to 7 days. After 7 days, unlabeled B cells were activated by adding LPS. Subsequently [$^{13}C_6$]leucine-labeled resting B lymphocytes were added (as the internal standard) to activated unlabeled B cells (0, 1, 2, 3, and 4 days after activation). Cell lysis and 1D gel electrophoresis were performed on the pooled mixtures. Each of the
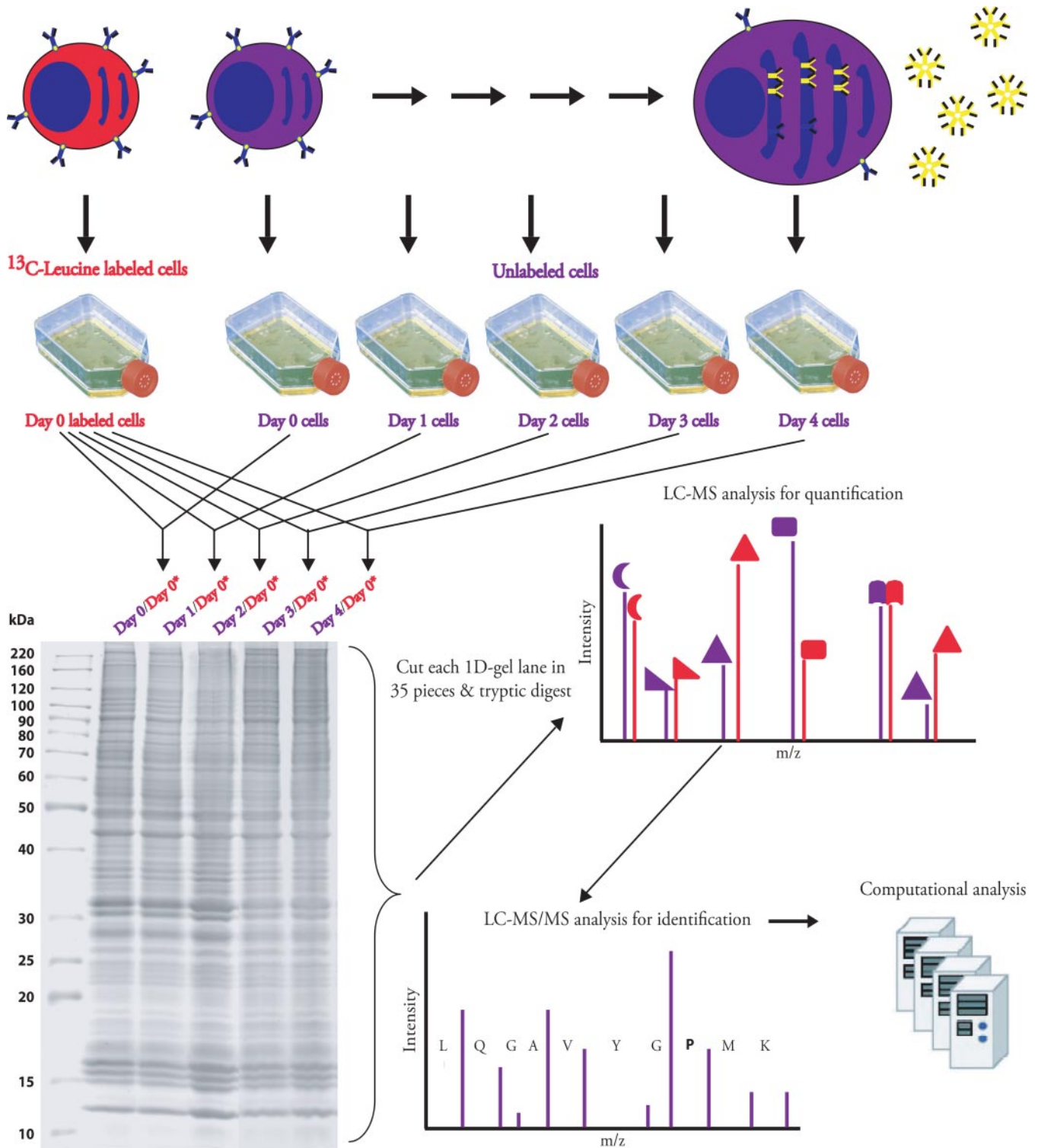
FIG. 1. **Schematic overview of the experimental procedure.** Unlabeled B cells (*purple*) were activated using LPS to differentiate into antibody-secreting plasma cells (also *purple*). Unlabeled cells from days 0, 1, 2, 3, and 4 were mixed with [$^{13}C_6$]leucine-labeled cells (*red*). These mixtures were lysed and separated on a single dimension gel. Each gel lane was cut into 35 pieces, and the proteins were in-gel digested using trypsin. The resulting complex peptide mixtures were analyzed by LC-MS for quantification and LC-MS/MS for identification. In the figure the ion signals in *red* indicate peptides originating from the labeled protein pool; the *purple* ion signals belong to the activated cells. The different labels illustrate peptides originating from different proteins.

resulting five 1D gel lanes was cut into 35 pieces, and each piece was in-gel digested using trypsin. The resulting tryptic peptides were separated and analyzed using nano-LC-MS/MS to identify and nano-LC-MS to quantify the proteins from the resulting peptide pairs. Before we performed our quantitative proteomic analysis we extensively validated our method, which we describe here first.

Validation of Complete Incorporation of Isotope Label— Murine B cells were grown in media containing either normal leucine or $^{13}C_6$ isotope-labeled leucine, *i.e.* with all carbon atoms replaced by ^{13}C. Leucine is an essential amino acid, and we therefore anticipated a 100% incorporation of the label as observed in previous similar studies (8, 9). Mass spectrometric analysis of random B cell proteins indeed showed incorporation of the $[^{13}C_6]$leucine. To illustrate the labeling efficiency we show in Fig. 2 typical peptide mass spectra, in this case of a tryptic peptide containing three leucines. As can be seen from these data, full incorporation of the ^{13}C-substituted leucine is achieved (reflected by the mass
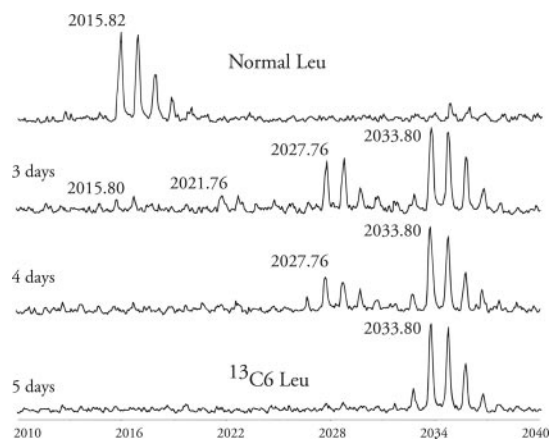

FIG. 2. **Incorporation of $[^{13}C_6]$leucine in proteins at different time points.** Medium containing "normal leucine" was replaced by medium containing $[^{13}C_6]$leucine, and the incorporation was followed as a function of time. Samples were obtained at 1, 2, 3, 4, and 5 days (days 1 and 2 not shown). The effect of partial incorporation of the label becomes clearer when looking at peptides containing more leucines. Therefore, ion signals of the peptide VILHLKEDQTEYLEER, which contains three leucines, are shown. The bottom MALDI spectrum shows that complete incorporation of $[^{13}C_6]$leucine is achieved at day 5.

shift of $3 \times 6 = 18$ Da) after 5 days following substitution of the normal leucine with $[^{13}C_6]$leucine in the growth medium. At earlier time points after the change in growth medium we still detected peptide ion signals wherein the leucine residues were only partly substituted by $[^{13}C_6]$leucines. Importantly we did not observe any residual unlabeled peptides 5 days after the change of growth media (see Fig. 2), which is critical because these would interfere with the peptides from the second unlabeled sample. We concluded that $[^{13}C_6]$leucine is a suitable amino acid to metabolically label our mammalian B cells to equilibrium. To confirm that the B cells grown in this particular $[^{13}C_6]$leucine-containing medium still could be activated and differentiate into plasma cells we validated the system as described previously (4). The secreted form of IgM was barely detectable before LPS treatment but increased thereafter continuously for 4 consecutive days, and cells increased in size and changed morphology following activation. Altogether these data (not shown) confirm that the $[^{13}C_6]$leucine-labeled I.29μ^+ (IgM, λ) lymphoma cell line, grown in this specific medium, represents an appropriate model for B cell differentiation.

Protein Identification by 1D Gel LC-MS/MS—After separating the protein mixtures on a 1D gel the proteins were in-gel digested with trypsin. The resulting peptide mixtures were separated, identified, and quantified. Using a Q-TOF MS/MS instrument ~25,000 MS/MS attempts were performed per run, resulting in on average ~2,200 peptide identifications. This difference resulted from a combination of the following factors: (*a*) not all proteins were/are present in the consulted database; (*b*) we did not consider any post-translational modifications, more than two mis-cleavages, or non-tryptic activity; (*c*) several peptides were selected for MS/MS analysis more than once; and (*d*) there was poor fragmentation of singly charged peptides. Initially PLGS 2.0 was used to identify peptides, but the data were reanalyzed using Mascot to check consistency with other identification engines. Exclusive identifications were observed by either engine, indicating that both identification algorithms are not comprehensive. To identify as many peptides as possible the entire dataset was analyzed twice by both PLGS 2.0 and Mascot. Detailed results on the total numbers of peptide identifications are presented in Supplemental Table S1 with an overview of the data

TABLE I

Number of peptides and proteins identified on days 1, 2, 3, and 4 after activation and the number of peptides containing leucine that were used for quantification of the proteins

Day	Peptides identified	Proteins identified	Peptide pairs for quantification	Proteins quantified
1	1,492	468	1,011	375
2	1,916	497	1,102	402
3	3,087	731	1,382	438
4	2,416	584	1,308	416
Total identifications[a]		1,001		
Total quantifications[b]			932	234

[a] Number of total unique protein identifications.

[b] Number of total unique peptide and protein quantifications on all days.

Relative expression levels of proteins of isotope-labeled and unlabeled non-activated B cells (*i.e.* day 0 and day 0*) randomly selected from different gel pieces of the 1D gel. The last two columns show the determined relative expression ratio day 0/day 0* with the calculated standard deviation. The sixth column describes the number of peptides identified for each of these proteins with the number of leucine residue-containing peptides that could be used for quantification within parentheses. Overall these data show that the variation in expression levels between these isotope-labeled and unlabeled non-activated cells is rather small with an averaged standard deviation of just 0.07. For more detailed data on all peptides see Supplemental Table S2.

	Gel piece	Swiss-Prot accession no.	Molecular mass	Description	No. peptides (Leu)	Day 0/day 0*	S.D.
			kDa				
1	4	P17742	17.8	Peptidyl-prolyl cis-trans isomerase	5 (3)	1.13	0.06
2	10	P57759	28.8	Erp29	3 (3)	0.93	0.05
3	12	P06151	36.5	Lactate dehydrogenase 1	7 (5)	0.95	0.03
4	13	P14869	34.2	60 S ribosomal protein P0	5 (3)	1.10	0.05
5	14	P05201	46.2	Glutamate-oxaloacetate transaminase	7 (6)	0.98	0.02
6	15	P09411	44.5	Phosphoglycerate kinase 1	4 (4)	0.96	0.02
7	16	Q9D8E6	47.1	Ribosomal protein L4	6 (4)	0.98	0.03
8	17	Q61753	56.5	3-Phosphoglycerate dehydrogenase	6 (5)	1.03	0.05
9	18	Q8CAY6	60.4	T-complex protein 1, subunit B	7 (6)	1.01	0.02
10	19	Q9D6Z1	65.4	Nucleolar protein Nop56	5 (4)	0.93	0.07
Average S.D.							0.07

in Table I. The ~2,200 peptide sequences per 1D gel lane led to the identification of ~550 proteins. The identification of ~550 proteins per time point resulted in a total number of 1,001 unique unambiguous protein identifications (see Supplemental Table S1) of which 252 proteins could be identified at every time point during B cell differentiation. These data do not imply that the other proteins are uniquely expressed only at a particular time point but reveal that the analysis of our samples could still be further expanded.

Protein Quantification by 1D Gel LC-MS — For quantification of the identified proteins it is a requirement to have a leucine residue present in one of the peptides used to identify the protein. The PLGS 2.0 software was used to quantify the peptides identified in the LC-MS/MS run containing either $[^{12}C_6]$leucine or $[^{13}C_6]$leucine using the LC-MS run of the same sample. $[^{12}C_6]$Leucine and $[^{13}C_6]$leucine were defined as variable modifications to search the database, and for that reason all peptides could be identified containing either one of those two residues. For the peptides that were identified by Mascot (see above), PLGS 2.0 could not be used because it can only quantify self-identified peptides. In this case MS-Quant was used to quantify. To rule out differences in ratios between quantification programs, several randomly selected peptides were quantified both by MSQuant and PLGS 2.0. No significant variations between these determined expression ratios were observed validating the use of both algorithms (data not shown).

As a validation step we first set out to measure relative expression patterns of the non-activated isotope-labeled and unlabeled cells, *i.e.* both at day 0. In theory all these proteins should have an expression ratio of 1 whereby possible variation provides information about biological variation and/or experimental errors. Table II shows the determined ratios in expression of randomly picked proteins taken from different parts of the 1D gel. Overall the data presented in Table II show that the variation in expression levels between these labeled and unlabeled non-activated cells is rather small. From these data, combined with the results on the growth rate of these cells and activation patterns as monitored by immunofluorescence and flow cytometry (data not shown), we conclude that the stable isotope labeling as such has no significant effect on the B cell proteome and that the biological variation between our two different cell pools is small.

To evaluate the quantification results, individual protein expression ratios were inspected in detail. Ratios and standard deviations for all identified proteins were calculated when possible (see Supplemental Table S2). To illustrate quantification at the individual peptide level we have depicted in Fig. 3 a single peptide pair ((R)TGEAIVDAALSALR(Q)) from the ER-resident protein P5 (Q8BK54) during B cell differentiation. The intensity of the isotope-labeled "day 0" internal standard was set to 100% during all days of the experiment, and the intensity of the unlabeled peptide represents the quantity of the protein during B cell differentiation. Already from this single peptide it is evident that the ER-resident protein P5 shows a gradual increase in expression during B cell differentiation with a maximal expression at day 4, representing nearly a 2-fold up-regulation. To illustrate the typical variation in quantification by using several peptides of a single protein we have depicted in Fig. 4 detailed information on ratios from 10 different leucine-containing peptide pairs observed in the LC-MS/MS data for another ER-resident protein, BiP (P20029). When averaging the data of these peptides the standard deviations were all below 0.05 for all days. The data provided in Supplemental Table S2 reveal that for most proteins in our dataset a number of two-peptide pairs already
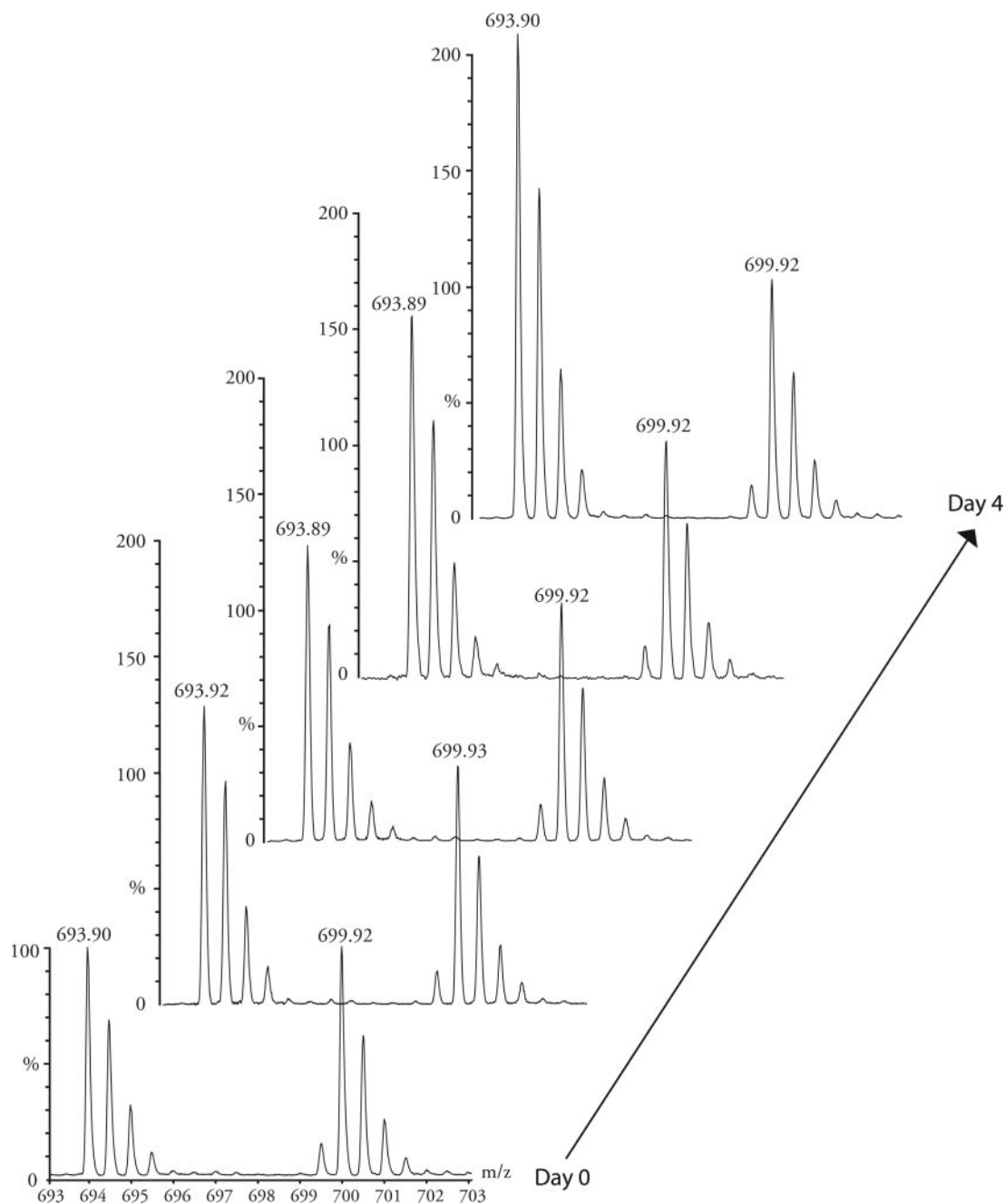
FIG. 3. **Dynamic quantification shown at the single peptide level during B cell differentiation.** To illustrate quantification at the peptide level the enlarged mass spectra of a single peptide pair ((R)TGEAIVDAALSALR(Q)) from the ER-resident protein P5 are shown as measured during the B cell differentiation process. The peptide contains two leucines resulting in a mass difference of 12 Da. Because the peptide is doubly charged the peptide pair is separated by 6 Thomson. The intensity of the unlabeled peptide (at m/z 694) represents the quantity of the intact protein during the experiment as normalized to the protein present before activation, represented by the unlabeled peptide (at m/z 700). P5 shows a maximal expression at day 4 after activation when it is ~2-fold up-regulated.

lead to reasonably accurate quantifications with rather small standard deviations. Although the variation in our experiment between different peptides is relatively small, it is still beneficial to have as many peptide pairs as possible to enable accurate protein quantification and meaningful error analysis.

The results presented in Figs. 2 and 4 and Table II validate our approach, and thus we advanced to analyze and quantify all identified proteins present in all 1D gels. From the ~2,200 peptides identified per 1D lane, typically ~1,150 peptide pairs were found. Per 1D lane this resulted in ~550 protein identifications of which almost 400 could be quantified (see Table I and Supplemental Table S2). From the total number of 1,001
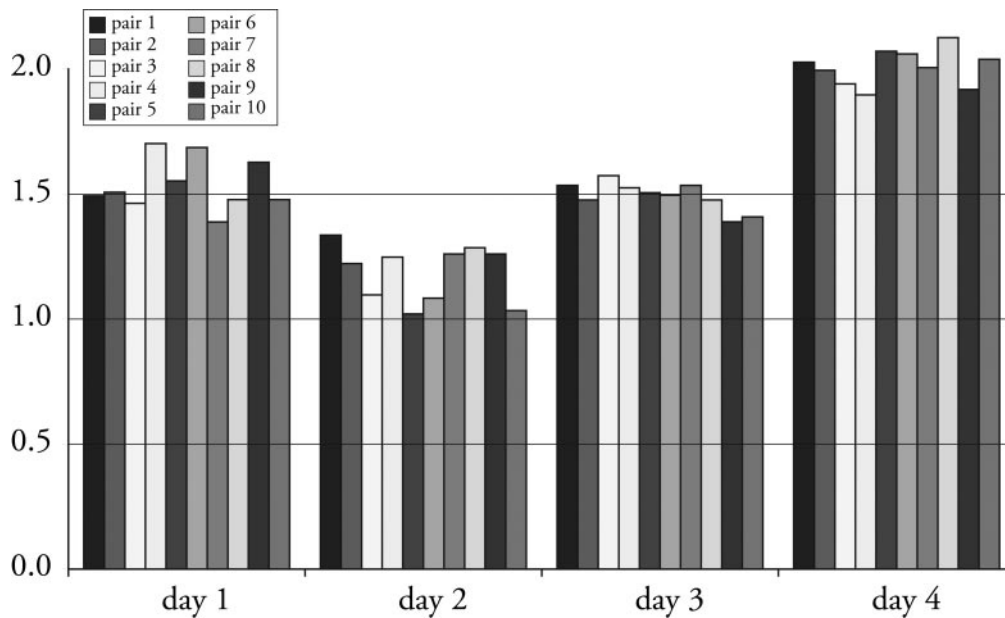
FIG. 4. **Variation in dynamic quantification for different peptides from the same protein during B cell differentiation.** For illustrative purposes detailed information on the ER-resident protein BiP are presented for which 10 peptide pairs could be used for quantification at all of the five time points. All peptide pairs show strikingly similar relative expression patterns. The standard deviations were smaller than 0.05.

uniquely identified proteins, 252 proteins were identified at every time point after activation. Only 234 of these proteins could be quantified across all time points because not all detected peptides contained leucine residues. A full overview of all detected 1,001 proteins is given in Supplemental Table S1; a full overview of the relative expression levels of the 234 proteins quantified at each time point is given in Supplemental Table S2.

Dynamic Proteome Profiling of B Cell Differentiation—In this study we identified more than 1,000 proteins from differentiating B cells, and we also quantified the relative expression levels of about a quarter of these proteins at all measured time points during B cell differentiation. Our set of quantitative proteome data is too large to discuss the relative expression of each individual protein by itself. Therefore, we analyzed our dynamic quantitative proteome data by an (in the microarray field) well accepted clustering approach. For an unbiased analysis the quantitative dataset of 234 times five entries (*i.e.* the different time points) was mathematically clustered using the clustering program TM4 (see "Experimental Procedures") whereby proteins with similar patterns in expression levels are grouped together. To prevent loss of significance of the clustering, only the 234 proteins that could be quantified on every day were used as input for the program. Using SOTA analysis we could group the total of 234 proteins into 12 distinct clusters, allowing a maximum diversity of 0.8 within a single cluster. Expression profiles of all proteins and the average expression profile per cluster are plotted in Fig. 5. The SOTA clustering process provided sufficient proteins per cluster (ranging from 10 to 37) with reasonable cell diversity numbers. The number of proteins per cluster and the diversity within the cluster are shown in Fig. 5 as well.

Functional annotation of all 234 proteins used in the clustering analysis was achieved using genome and protein databases (Supplemental Table S2). We chose for a classification into 12 different categories. In Fig. 6 the results of this functional annotation of all 234 proteins is represented in several pie charts whereby the largest functional groups contained about 40 proteins (*i.e.* metabolism and translation), and the smallest groups still contained at minimum four proteins (*i.e.* cell cycle and membrane transport). The pie chart of the complete set of proteins revealed that the B cell proteome is functionally broadly covered in our dataset. Next to the pie chart of all 234 proteins, pie charts of some selected clusters are also depicted in Fig. 6, in particular for those clusters that showed selective enrichment for specific functional groups of proteins. For instance, when focusing on cluster 3, which contains proteins that show a continuous gradual increase of protein expression during the B cell differentiation process, the pie chart reveals that whereas ER folding proteins represent only 6% of all quantified proteins, they represent 44% of all proteins present in cluster 3. Similarly cluster 4 is enriched for proteins involved in the immune system (from 5 to 35%). The proteins in cluster 4 show on average first a relative decrease in protein expression followed by a gradual increase during the differentiation process. Cluster 9 is enriched for cytoskeleton proteins (from 7% in the complete protein set to 35% in this cluster), and cluster 11 is enriched for proteins involved in protein synthesis (*i.e.* primarily ribosomal subunits; from 19 to 47%). These results indicate that functionally related proteins exhibit to a large extent co-expression patterns during the B cell differentiation process.
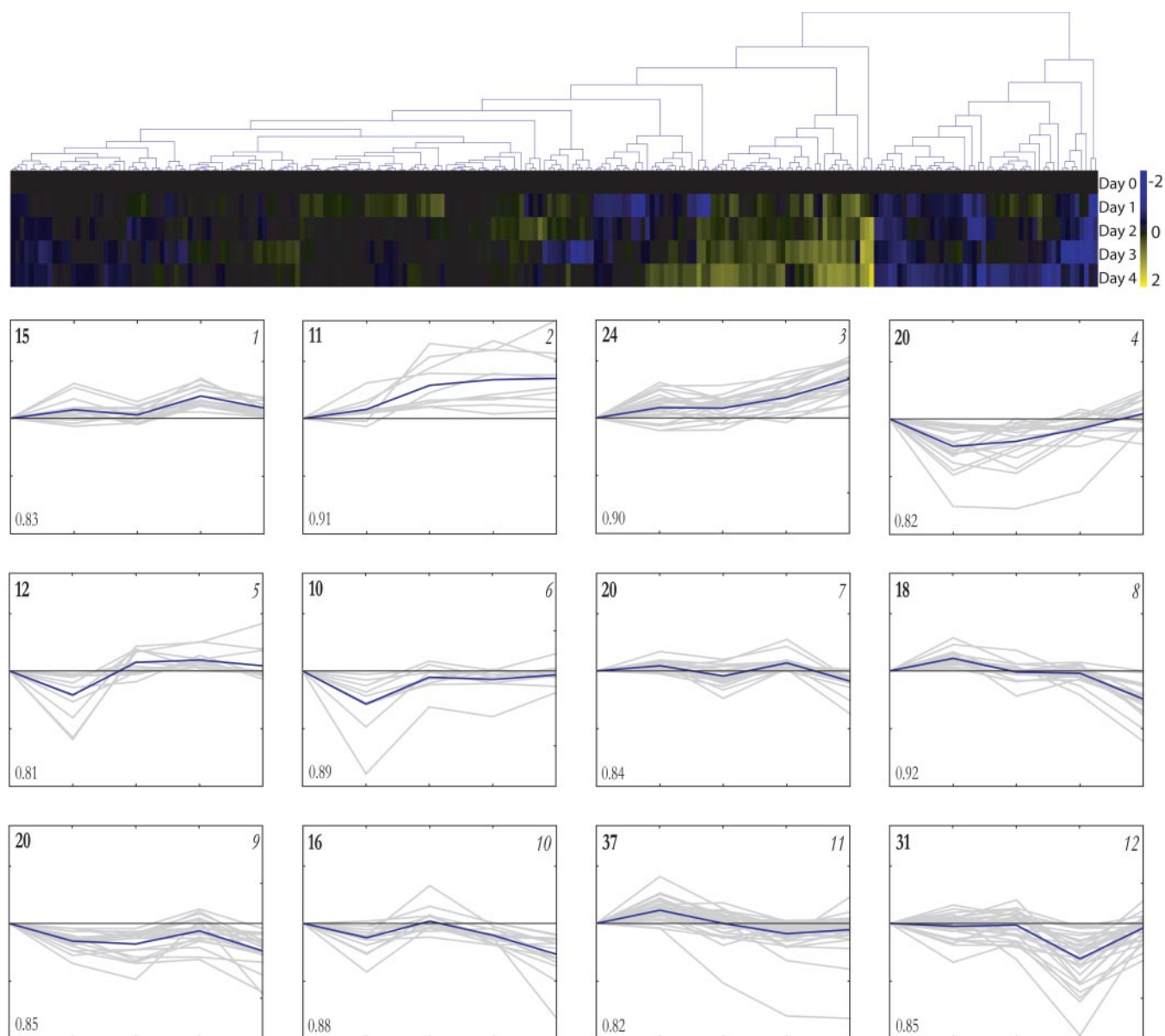
FIG. 5. **Clustering analysis of expression profiles of 234 proteins monitored during B cell differentiation.** The 234 proteins were clustered into 12 clusters based on their expression profiles. The SOTA cluster tree is shown at the *top* of the figure. Below in *blue* the mean expression profile of the proteins in each cluster is shown. The expression graphs also show in *gray* the expression profile of each individual protein inside each SOTA cluster. For each distinct cluster the cluster number is given in the *right upper corner*, the number of proteins present is given in the *left upper corner*, and the average cell diversity is given in the *left bottom corner*. Detailed information on the proteins within each individual cluster can be found in Supplemental Table S2.

Proteins Involved in Protein Folding in the Endoplasmic Reticulum—Because B cell differentiation involves a massive increase of the secretory pathway and especially the ER (4), we performed a more thorough analysis on the ER-resident proteins involved in folding. We quantified 22 ER-resident proteins of which 16 were found at each time point. Fig. 7 and Table III show their expression data and profiles. From these data it is evident that most ER proteins follow a similar general trend: a gradual up-regulation in protein expression during B cell differentiation. More thorough analysis showed that some

subtle differences and stronger deviations were present in their expression kinetics predominantly at the early time points after activation. Several proteins (*e.g.* calreticulin, GRP94, ERP57, and BiP) showed a quite rapid up-regulation at day 1 that fell back at day 2 after which protein levels became up-regulated again at days 3 and 4. Other proteins (*e.g.* GRP170) showed a somewhat opposite pattern in which at day 1 the up-regulation was below the average of that day (*e.g.* a down-regulation) after which it increased continuously. Calnexin behaved quite differently from all other well known
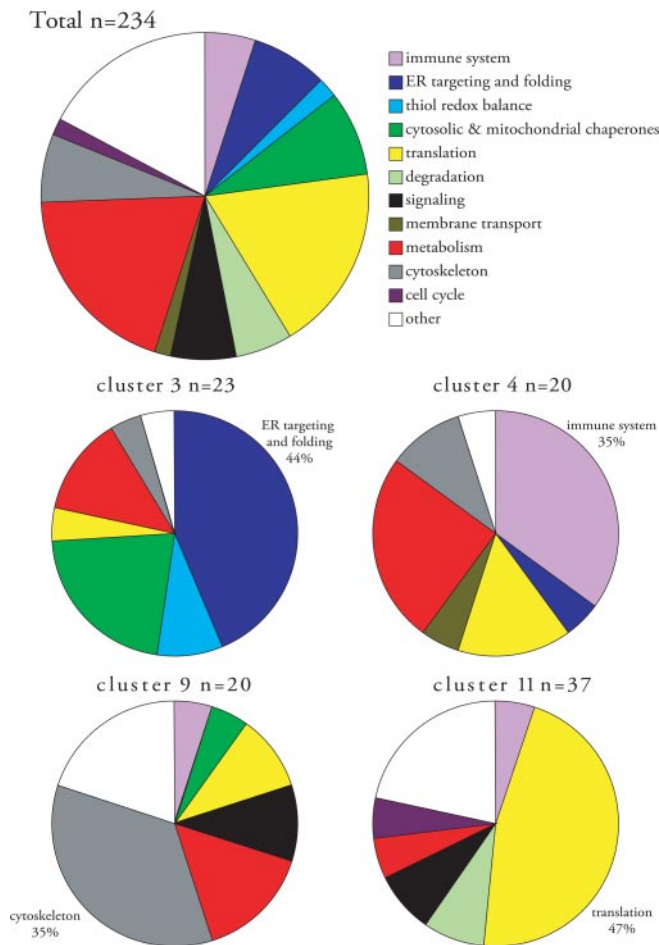
Total n=234

- ▨ immune system
- ▨ ER targeting and folding
- ▨ thiol redox balance
- ▨ cytosolic & mitochondrial chaperones
- ▨ translation
- ▨ degradation
- ▨ signaling
- ▨ membrane transport
- ▨ metabolism
- ▨ cytoskeleton
- ▨ cell cycle
- ☐ other

cluster 3 n=23   cluster 4 n=20

cluster 9 n=20   cluster 11 n=37

FIG. 6. **Functional annotation of the 234 proteins monitored during B cell differentiation and clustered by the SOTA method.** At the *top* of the figure a pie chart is shown of the complete set of 234 proteins. Next to the pie chart of all proteins monitored in our dynamic quantitative profiling of the B cell differentiation process, pie charts of selected clusters are depicted. These pie charts reveal that cluster 3 is considerably enriched for proteins involved in protein folding in the ER. Proteins functionally involved in the immune system are preferentially located in cluster 4, and most cytoskeletal proteins end up in cluster 9, whereas nearly half of the proteins in cluster 11 are proteins involved in translation, in particular ribosomal subunits. More detailed information on individual proteins can be found in Supplemental Table S2.

ER proteins involved in protein folding. In contrast to most of the other ER folding proteins the expression level ratio did not peak at day 4 but instead remained constant with a slight maximum earlier in differentiation at day 2.

DISCUSSION

In this study we aimed for an in-depth proteome analysis of B cell differentiation. We therefore followed the differentiation process of a B cell line using a combination of stable isotope labeling, SDS-PAGE, and LC-MS/MS. After activation with LPS, the B cells differentiated into antibody-secreting plasma cells. We followed this differentiation process during 5 days and performed a thorough proteome analysis at five time

points (at each day after activation) in which we identified proteins and quantified their relative expression levels. We clustered the identified proteins based on their expression patterns and observed that these clusters were enriched in functionally related proteins. This confirmed our previous hypothesis that B cell differentiation proceeds via sequential waves of changes in protein expression (4). We not only found that ER-resident proteins were linearly up-regulated as expected, but using the current accurate method of quantification we also identified subclusters of ER proteins.

B Cell Proteome Coverage—We analyzed the B cell proteome at five time points after LPS addition to follow the differentiation of resting cells to antibody-secreting plasma cells. This analysis led to the identification of more than 1,000 proteins, which covered a broad variety of cellular functions. In Fig. 8 we constructed a "virtual" two-dimensional gel of all 1,001 identified proteins for which we have used the theoretical pI and molecular weight of the proteins as listed for the Swiss-Prot and TrEMBL entries in the *M. Musculus* database. The comprehensive nature of our current approach is underscored by the fact that we detected proteins in a wide range of pI values and molecular masses: from 3.6 to 12.2 and from 5 to 200 kDa, respectively. When we consider the practical range of commonly used 2D gels to be approximately pI 3–9 and molecular mass 18–120 kDa, many identified proteins (217 and 168, respectively) fall outside this range. Although the proteome coverage is far more comprehensive in our current approach than in the earlier 2D gel-based approach (4) (10 times more identifications), it is still far from complete. Experimental undersampling using the current LC-MS setup is one of the reasons for this; the undersampling is due in part to the fact that in the experimental setup not all peptides present in the extremely complex mixtures are selected for MS/MS analysis. However, being aware of these limitations, we have created a substantial quantitative dataset in which proteins from many different functional groups are represented. This allows us to provide a good representation of what is happening in the B cells at proteome level during differentiation.

Protein Expression Quantification Using SILAC—Using the SILAC approach with [^{13}C]leucine as isotope-labeled amino acid we were able to quantify the relative expression levels of about a quarter (*i.e.* 234) of the 1,001 identified proteins at all measured time points during the B cell differentiation process. The number of peptide pairs used for quantification ranged from 1 to ~30 per protein with ~75% of the quantified proteins by at least two peptide pairs (see Supplemental Table S2). The standard deviation was shown to be quite small when a peptide could be quantified by at least two peptide pairs. Only a rather small fraction (less than 10%) could not be quantified because the detected peptides used for identification lacked leucine amino acids, indicating that leucine is a good choice in SILAC. We conclude that the combination of 1D gel electrophoresis, LC-MS/MS, and SILAC (using [^{13}C]leucine) is an excellent method in quantitative proteom-
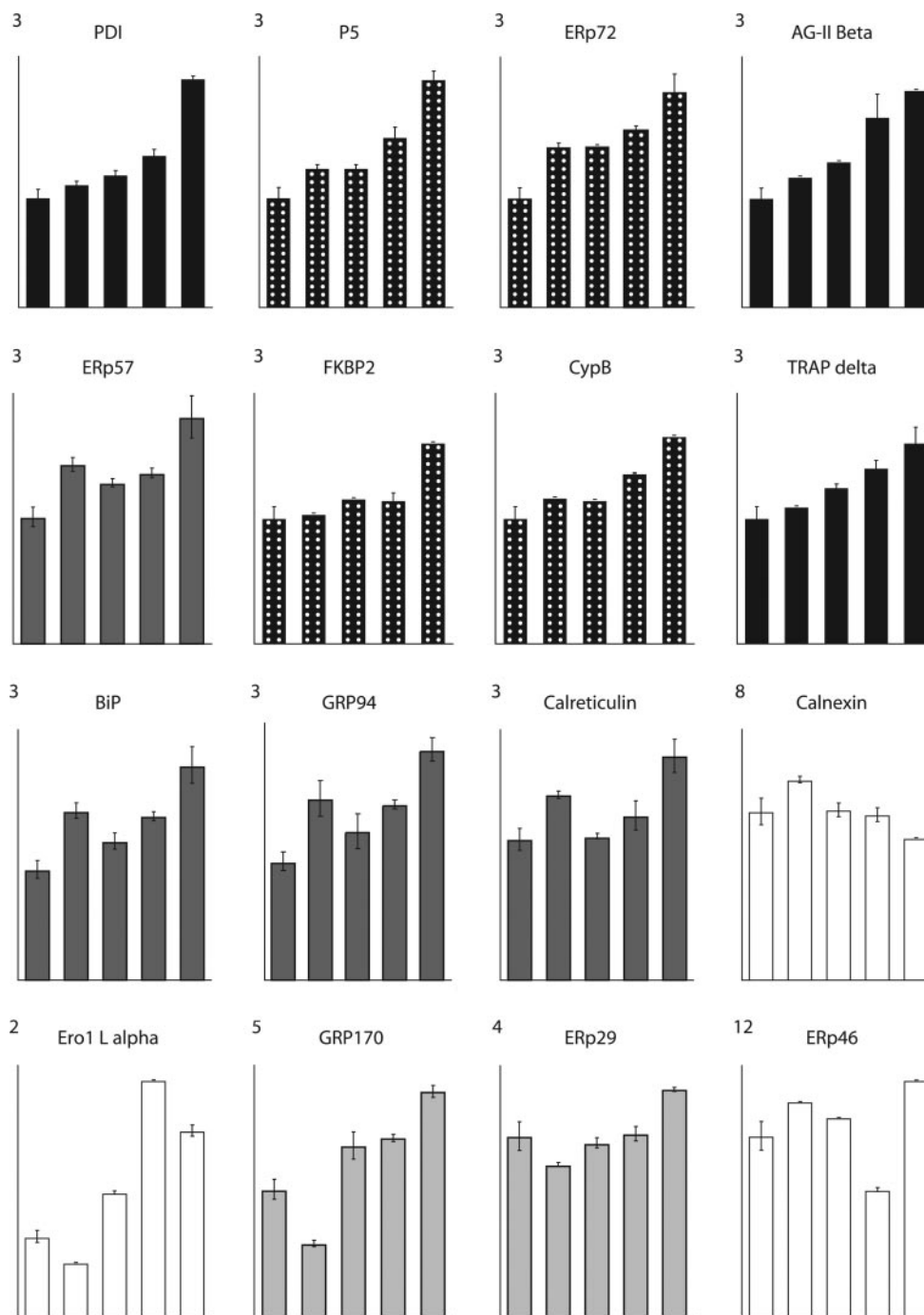
FIG. 7. **Detailed expression kinetic profiles of ER-resident proteins.** The histograms show the relative expression patterns of quantified ER-resident proteins. The cluster number is indicated in the *upper left corner* of each histogram. Eleven of these proteins fall in cluster 3. The other ER-resident proteins are up-regulated during differentiation as well with Cnx as the notable exception. Subclusters can also be observed. PDI, AG-IIβ, and TRAPδ are continuously up-regulated as well as P5, ERp72, FKBP2, and CypB. These last four proteins have somewhat slower kinetics. ERp57, BiP, GRP94, and Crt show an initial up-regulation at day 1. GRP170 and ERp29 in contrast show an initial relative down-regulation at day 1. Note that PDI, P5, ERp72, ERp57, ERp29, and ERp46 belong to the oxidoreductase family; FKBP2 and CypB are peptidyl-prolyl cis-trans isomerases; BiP, GRP94, and GRP170 are chaperones; and Crt and Cnx are lectin chaperones.

ics, allowing the analysis of detailed co-expression patterns of functionally related proteins during B cell differentiation.

In quantitative proteomics based on stable isotope labeling there are alternative approaches to introduce the chemically identical but mass-differentiated stable isotope tags in proteins (7, 21), broadly classified into two groups based on the moment that the tag is incorporated: (*a*) biological incorporation where labeling of the protein/peptide is achieved by

TABLE III
Identification and relative quantification of ER proteins during B cell differentiation

Values given at days 1, 2, 3, and 4 are relative quantities to day 0 (no activation). No Leu means that the protein was identified on a particular day but not quantified. No ID means that the protein was not identified on that particular day. New cluster numbers correspond to clusters in Fig. 6, and old cluster numbers correspond to clusters presented previously (4). SER, sarco(endo)plasmic reticulum; DDOST, dolichyl-diphosphooligosaccharide-protein glycosyltransferase.

	Day 0	Day 1	Day 2	Day 3	Day 4	Old cluster	New cluster
BiP	1	1.54	1.26	1.49	1.96	V	3
ERp57	1	1.42	1.27	1.35	1.80	V	3
PDI	1	1.12	1.21	1.39	2.09	V	3
P5	1	1.27	1.27	1.55	2.08	V	3
ERp72	1	1.47	1.48	1.63	1.98	V	3
AG-IIβ	1	1.19	1.33	1.74	1.99	V	3
GRP170	1	0.57	1.35	1.57	1.78	V	5
Calreticulin	1	1.32	1.02	1.17	1.60	V	3
Rotamase	1	1.16	1.14	1.35	1.65		3
GRP94	1	1.40	1.07	1.33	1.68	V	3
ERp46	1	1.19	1.10	0.70	1.31	V	12
TRAPδ	1	1.09	1.26	1.40	1.60		3
Calnexin	1	1.19	1.01	0.98	0.84	III	8
FKBP2	1	1.03	1.15	1.14	1.60		3
AG-IIα	1	0.97	No ID	1.49	No Leu		
ERp29	1	0.84	0.96	1.01	1.26	V	4
DDOST 48 kDa	1	No ID	No ID	No Leu	1.56		
SER ATPase 3	1	No ID	No ID	0.90	No ID		
Ero1α	1	0.67	1.56	2.99	2.35		2
TRAPα	1	No ID	No ID	No ID	1.020		
ERp44	1	No Leu	No Leu	No ID	No ID	V	
SER ATPase	1	No ID	No Leu	No Leu	No ID		

Fig. 8. **Two-dimensional map of all 1,001 identified B cell proteins.** To construct this virtual 2D map the theoretical pI and molecular weight of the identified proteins were taken as listed for the Swiss-Prot and TrEMBL entries in the *M. musculus* database. The practical boundaries of the standard 2D gel proteomics approach are indicated by the *rectangular box*. Nearly 40% of the identified proteins fall outside this box.
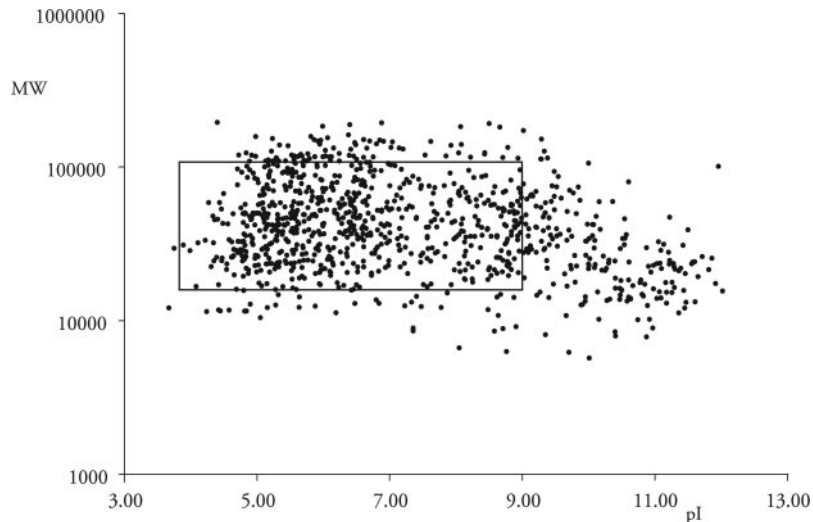
growing cells in media enriched in stable isotope-containing compounds as done in the present SILAC approach (8) and (*b*) chemical incorporation, which relies on the use of a derivatization agent for chemical modification of proteins in a site-specific manner after harvesting of the proteins, such as in the ICAT and ITRAQ (isobaric tagging reagents) approaches (22–24). ICAT offers the advantage of not requiring gel separation and thus may seem a less laborious alternative compared with SILAC. Additionally ICAT can be applied to any cells or even body fluids and tissue. However, the SILAC approach chosen here offers the earliest time point to introduce isotope labels without the need of any *in vitro* derivatization steps, which may induce experimental variation. In contrast to chemical incorporation methods, *e.g.* the ICAT approach (22), no significant reduction in sample complexity is obtained in SILAC. Therefore, using metabolic labeling potentially higher protein coverage can be achieved, and the number of peptide pairs used for quantification per protein will generally be higher.

Functionally Related Proteins Have Similar Expression Patterns—During differentiation from resting B lymphocyte to plasma cell, the architecture of the cell is completely rearranged because a significant increase of secretory machinery is needed to sustain the mass production of antibody molecules. This change involves dramatic rearrangements in the proteome of the cell. Most notably, ER-resident proteins are up-regulated gradually and continuously during differentiation (Fig. 7).

In line with previous findings we can distill from our data that B cells carefully prepare for their role as antibody factory. Functionally related proteins are regulated in groups: proteins needed for energy supply and protein production are up-regulated early after activation; ER-resident proteins are continuously up-regulated; and finally, when the secretory capacity has increased, IgM production starts (4). The clustering of the proteins identified in the present study reveals that functionally related proteins largely follow similar expression kinetics during B cell differentiation.

As an example, the proteins in cluster 4 were first decreased in relative expression followed by a gradual increase during the differentiation process. This cluster is enriched in proteins involved in the immune system such as the Ig light chain λ. Ig heavy chain μ followed a somewhat similar expression pattern, although it fell in cluster 5. These clusters share similar global kinetics (see Fig. 5), showing an initial decrease in relative expression followed by an increase in relative expression of proteins at later time points. These results are in line with our previous conclusions, summarized above, that the B cells first have to prepare themselves before the large scale production of antibodies can take off.

Proteins involved in translation, in particular ribosomal subunits, were strongly enriched in cluster 11 (from 19% in the total protein pool to 47%). The general expression pattern of this cluster (Fig. 5) shows a strong initial up-regulation followed by down-regulation relative to the total protein increase during 2 days after which the protein level increased again at the final day of differentiation. After the B cell received the signal for differentiation, additional ribosomal subunits were produced. This extra increase in production was not maintained during the differentiation process, however. Electron microscopy studies have shown that resting B cells are rich in cytosolic ribosomes (data not shown), which might be relocated to the ER membrane to facilitate the production of ER-resident proteins and IgM.

Cluster 9, which contained proteins formally down-regulated over the course of differentiation, was enriched in cytoskeletal proteins. The down-regulation of the cytoskeletal proteins is only a relative down-regulation caused by our normalizing procedure whereby we corrected for the overall increase in protein content of the differentiated cells (see "Experimental Procedures"). Of the 16 cytoskeletal proteins we identified, seven ended up in cluster 9. Most other cytoskeletal proteins we identified were relatively down-regulated over time as well but with different kinetics.

Subclusters of Related Proteins within a Cluster—Although most functionally related proteins showed similar expression patterns, not all of them did. The cytosolic and mitochondrial chaperones for example were not clearly enriched in any cluster. Instead they were found to be spread over clusters 2, 3, 5, 7–9, 11, and 12 whereby the majority fell in clusters 3 and 7. Almost all these chaperones increased at the first day after activation, which allowed them to immediately assist in the need for elevated energy and protein production that underlies the making of a secretory plasma cell.

Diffuse clustering appeared for the subunits of the hetero-oligomeric molecular chaperone complex chaperonin-containing t-complex polypeptide 1 (we were able to quantify seven of the eight known subunits), which assists protein folding in the eukaryotic cytosol. Three subunits of this complex ended up in cluster 7, two were in cluster 3, one was in cluster 4, and one was in cluster 11 (see Supplemental Table S2).

Another example of diffuse clustering is observed for the ribosomal proteins: although 66% of proteins present in the ribosome resided in only two clusters, *i.e.* clusters 11 (40%) and 12 (26%), not all subunits showed identical expression patterns (see Supplemental Table S2). These subclusters may represent the changing function of a protein complex, which is accomplished by exchanging subunits or changing their stoichiometry. Indeed other quantitative proteome studies showed nonhomogeneous expression patterns of different subunits of protein complexes as a response to a change in the environment (13, 25, 26). The proteasome, another example of a large protein complex, responds to interferon-γ by forming a so-called immunoproteasome. Following exposure to interferon-γ, expression of specific proteasome subunits is induced that alters the catalytic function and specificity of the proteasome, enhancing production of certain major histocompatibility complex class I epitopes (27). Perhaps this is only one of various possible changes in proteasome composition as we observed proteasomal subunits in seven different clusters.

Another example of protein complexes in which subunits and cofactors are not regulated together are the Hsp70 and Hsp90 chaperones. These chaperones not only assist in folding of newly synthesized proteins but also control the conformational regulation of a variety of client proteins involved in signal transduction, cell proliferation, and apoptosis (see Hohfeld *et al.* (28)). To be functional, these chaperones associate with several co-chaperones, which stimulate and regulate their function. Interaction with Hop for example stimulates the cooperation of Hsp70 and Hsp90, which is important in signal transduction. When the chaperones interact with C-terminal Hsc70-interacting protein on the other hand, degradation of substrate proteins is promoted (29). Therefore, although functionally related proteins in general will follow similar expres-

sion patterns, fine tuning of the precise function is accomplished by changing subunit composition or partnerships, resulting in the appearance of expression subclusters. We found this to be particularly true for the expression patterns of ER-resident proteins.

Regulation of ER-resident Proteins—Because the major difference between a B lymphocyte and a plasma cell is the size of the secretory organelles and in particular the ER, regulation of ER-resident proteins is of special interest. Upon clustering, 11 of the 16 fully quantified ER-resident proteins ended up in cluster 3, which consists of the proteins that were continuously and gradually up-regulated. The other quantified ER-resident proteins also showed an increase in expression level but followed different kinetics (Fig. 7). Calnexin (Cnx) was a major exception because, after an initial up-regulation at the 1st day after activation, the level of this protein gradually decreased relative to the other proteins. In our previous study Cnx also was the only ER-resident protein that deviated from all other ER-resident proteins. Cnx and its soluble homolog calreticulin (Crt) are lectin chaperones with affinity for monoglucosylated proteins. These two proteins have largely redundant functions, but specific substrates display a preference for Cnx (30). The up-regulation of Crt but not Cnx may indicate that secretory IgM subunits preferentially interact with Crt.

Disulfide bond formation, a redox reaction, is very prominent in IgM production because over 100 disulfide bonds need to be formed in each assembled IgM molecule. To allow rapid introduction of disulfide bonds into IgM by oxidoreductases, the ER needs to have a large oxidative capacity. Ero1α is very important for maintaining ER-resident oxidoreductases such as PDI and ERp57 in an active state. We found a remarkably strong up-regulation of Ero1α upon differentiation, which clearly reflects this need for oxidizing power. Whereas conditions in the ER allow formation of disulfide bonds, other cellular compartments have a relatively reducing milieu. The peroxiredoxin protein family is important for maintaining proper redox conditions in these compartments. The expression patterns of the peroxiredoxins we identified show a continuous up-regulation synchronized with the ER proteins. When the ER expands an excess of electrons is produced, as a consequence of disulfide bond formation, which induces pressure on the redox conditions of the cell. Additional buffer capacity in the cytosol and mitochondria might be needed to maintain the proper redox conditions in these compartments.

Whereas Cnx and Ero1α differed most from the overall regulational pattern of the other ER-resident proteins, we detected various more subtle differences, which allowed us to define expression subclusters of ER-resident proteins. These differences in expression patterns were most prominent at the first days after activation. PDI, TRAPδ, and the β subunit of α-glucosidase II were up-regulated at a constant rate over the days, whereas levels of P5, ERp72, FKBP2, and CypB increased faster at later days than immediately after activation. ERp57, BiP, GRP94, and Crt showed a transient peak in

expression at the 1st day after activation, whereas GRP170 and ERp29 showed a relative dip in their expression level at day 1. Multiprotein complexes in the ER were reported in several cell types, but the exact composition varies from report to report and may be determined by the specific needs of particular client proteins (31–34). The subclusters in ER-resident proteins we found therefore may reflect a rearrangement of protein complexes present in the ER to optimize them for the efficient production of antibody molecules.

Concluding Remarks—Because in recent years nano-LC-MS/MS approaches have proven to lead to a significant increase in protein identifications and thus proteome coverage when compared with 2D gel approaches (7, 35, 36), we used such an approach to achieve a more extensive proteome analysis of the B cell differentiation process. We identified 1,001 proteins, a roughly 10-fold increase as compared with our earlier 2D gel-based study. Of these 1,001 proteins we could quantify a quarter at each of the five time points addressed providing a wide ranging view of the dynamic changes of the B cell proteome during differentiation. The 234 proteins for which we obtained a complete expression pattern were clustered in an unbiased way, organized via co-expression patterns. Analysis of the proteins present in the different clusters revealed that many functionally related proteins displayed similar temporal protein expression profiles, validating our previous proposal (4) that B cells anticipate their secretory role in a multistep process. Whereas many functionally related proteins exhibited co-expression patterns, subtle differences in regulation of related proteins, in particular in the ER, were also revealed. The use of a quantitative and dynamic proteomic approach via isotope labeling and mass spectrometry thus gave us both a global and detailed view of the changes that occur at the proteome level during B cell differentiation.

REFERENCES

1. Alberini, C., Biassoni, R., DeAmbrosis, S., Vismara, D., and Sitia, R. (1987) Differentiation in the murine B cell lymphoma I.29: individual $\mu+$ clones may be induced by lipopolysaccharide to both IgM secretion and isotype switching. *Eur. J. Immunol.* **17,** 555–562

2. Shaffer, A. L., Shapiro-Shelef, M., Iwakoshi, N. N., Lee, A. H., Qian, S. B., Zhao, H., Yu, X., Yang, L., Tan, B. K., Rosenwald, A., Hurt, E. M., Petroulakis, E., Sonenberg, N., Yewdell, J. W., Calame, K., Glimcher, L. H., and Staudt, L. M. (2004) XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **21,** 81–93

3. Wiest, D. L., Burkhardt, J. K., Hester, S., Hortsch, M., Meyer, D. I., and Argon, Y. (1990) Membrane biogenesis during B cell differentiation: most endoplasmic reticulum proteins are expressed coordinately. *J. Cell Biol.* **110,** 1501–1511

4. van Anken, E., Romijn, E. P., Maggioni, C., Mezghrani, A., Sitia, R., Braakman, I., and Heck, A. J. (2003) Sequential waves of functionally related proteins are expressed when B cells prepare for antibody secretion. *Immunity* **18,** 243–253

5. Goshe, M. B., and Smith, R. D. (2003) Stable isotope-coded proteomic mass spectrometry. *Curr. Opin. Biotechnol.* **14,** 101–109

6. Krijgsveld, J., Ketting, R. F., Mahmoudi, T., Johansen, J., Artal-Sanz, M., Verrijzer, C. P., Plasterk, R. H., and Heck, A. J. (2003) Metabolic labeling of C. elegans and D. melanogaster for quantitative proteomics. *Nat. Biotechnol.* **21,** 927–931

7. Romijn, E. P., Krijgsveld, J., and Heck, A. J. R. (2003) Recent liquid chromatographic-(tandem) mass spectrometric applications in proteomics. *J. Chromatogr. A* **1000,** 589–608

8. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386

9. Ong, S. E., Kratchmarova, I., and Mann, M. (2003) Properties of ^{13}C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2,** 173–181

10. Ong, S. E., Foster, L. J., and Mann, M. (2003) Mass spectrometric-based approaches in quantitative proteomics. *Methods* **29,** 124–130

11. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21,** 315–318

12. Pratt, J. M., Robertson, D. H., Gaskell, S. J., Riba-Garcia, I., Hubbard, S. J., Sidhu, K., Oliver, S. G., Butler, P., Hayes, A., Petty, J., and Beynon, R. J. (2002) Stable isotope labelling *in vivo* as an aid to protein identification in peptide mass fingerprinting. *Proteomics* **2,** 157–163

13. Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005) Nucleolar proteome dynamics. *Nature* **433,** 77–83

14. Everley, P. A., Krijgsveld, J., Zetter, B. R., and Gygi, S. P. (2004) Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol. Cell. Proteomics* **3,** 729–735

15. Zhang, R., Sioma, C. S., Wang, S., and Regnier, F. E. (2001) Fractionation of isotopically labeled peptides in quantitative proteomics. *Anal. Chem.* **73,** 5142–5149

16. Pruess, M., Fleischmann, W., Kanapin, A., Karavidopoulou, Y., Kersey, P., Kriventseva, E., Mittard, V., Mulder, N., Phan, I., Servant, F., and Apweiler, R. (2003) The Proteome Analysis database: a tool for the *in silico* analysis of whole proteomes. *Nucleic Acids Res.* **31,** 414–417

17. Al-Shahrour, F., Diaz-Uriarte, R., and Dopazo, J. (2004) FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* **20,** 578–580

18. Pavlidis, P., Lewis, D. P., and Noble, W. S. (2002) Exploring gene expression data with class scores. *Pac. Symp. Biocomput.* 474–485

19. Saeed, A. I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis. *BioTechniques* **34,** 374–378

20. Herrero, J., Valencia, A., and Dopazo, J. (2001) A hierarchical unsupervised growing neural network for clustering gene expression patterns. *Bioinformatics* **17,** 126–136

21. Julka, S., and Regnier, F. (2004) Quantification in proteomics through stable isotope coding: a review. *J. Proteome Res.* **3,** 350–363

22. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17,** 994–999

23. Ross, P. L., Huang, Y. N., Marchese, J., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3,** 1154–1169

24. DeSouza, L., Diehl, G., Rodrigues, M. J., Guo, J., Romaschin, A. D., Colgan, T. J., and Siu, K. W. (2005) Search for cancer markers from endometrial tissues using differentially labeled tags iTRAQ and cICAT with multidimensional liquid chromatography and tandem mass spectrometry. *J. Proteome Res.* **4,** 377–386

25. Ranish, J. A., Yi, E. C., Leslie, D. M., Purvine, S. O., Goodlett, D. R., Eng, J., and Aebersold, R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33,** 349–355

26. de Lichtenberg, U., Jensen, L. J., Brunak, S., and Bork, P. (2005) Dynamic complex formation during the yeast cell cycle. *Science* **307,** 724–727

27. Barton, L. F., Cruz, M., Rangwala, R., Deepe, G. S., Jr., and Monaco, J. J. (2002) Regulation of immunoproteasome subunit expression in vivo following pathogenic fungal infection. *J. Immunol.* **169,** 3046–3052

28. Hohfeld, J., Cyr, D. M., and Patterson, C. (2001) From the cradle to the grave: molecular chaperones that may choose between folding and degradation. *EMBO Rep.* **2,** 885–890

29. Connell, P., Ballinger, C. A., Jiang, J., Wu, Y., Thompson, L. J., Hohfeld, J., and Patterson, C. (2001) The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. *Nat. Cell Biol.* **3,** 93–96

30. Molinari, M., Eriksson, K. K., Calanca, V., Galli, C., Cresswell, P., Michalak, M., and Helenius, A. (2004) Contrasting functions of calreticulin and calnexin in glycoprotein folding and ER quality control. *Mol. Cell* **13,** 125–135

31. Meunier, L., Usherwood, Y. K., Chung, K. T., and Hendershot, L. M. (2002) A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Mol. Biol. Cell* **13,** 4456–4469

32. Kuznetsov, G., Chen, L. B., and Nigam, S. K. (1997) Multiple molecular chaperones complex with misfolded large oligomeric glycoproteins in the endoplasmic reticulum. *J. Biol. Chem.* **272,** 3057–3063

33. Zhang, J., and Herscovitz, H. (2003) Nascent lipidated apolipoprotein B is transported to the Golgi as an incompletely folded intermediate as probed by its association with network of endoplasmic reticulum molecular chaperones, GRP94, ERp72, BiP, calreticulin, and cyclophilin B. *J. Biol. Chem.* **278,** 7459–7468

34. Tatu, U., and Helenius, A. (1997) Interactions between newly synthesized glycoproteins, calnexin and a network of resident chaperones in the endoplasmic reticulum. *J. Cell Biol.* **136,** 555–565

35. Washburn, M. P., Wolters, D., and Yates, J. R., III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19,** 242–247

36. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2,** 43–50