The Organelle Proteome of the DT40 Lymphocyte Cell Line*

Stephanie L. Hall‡, Svenja Hester§, Julian L. Griffin‡, Kathryn S. Lilley§, and Antony P. Jackson‡¶

A major challenge in eukaryotic cell biology is to understand the roles of individual proteins and the subcellular compartments in which they reside. Here, we use the localization of organelle proteins by isotope tagging technique to complete the first proteomic analysis of the major organelles of the DT40 lymphocyte cell line. This cell line is emerging as an important research tool because of the ease with which gene knockouts can be generated. We identify 1090 proteins through the analysis of preparations enriched for integral membrane or soluble and peripherally associated proteins and localize 223 proteins to the endoplasmic reticulum, Golgi, lysosome, mitochondrion, or plasma membrane by matching their density gradient distributions to those of known organelle residents. A striking finding is that within the secretory and endocytic pathways a high proportion of proteins are not uniquely localized to a single organelle, emphasizing the dynamic steady-state nature of intracellular compartments in eukaryotic cells. Molecular & Cellular Proteomics 8:1295–1305, 2009.

The chicken pre-B cell line DT40 exhibits a remarkably high ratio of targeted to random integration for transfected DNA constructs. This property is unusual in vertebrate cell lines and enables targeted gene disruption experiments to be carried out with relative ease (1). Consequently, DT40 has become a major research tool for the molecular dissection of a wide range of cellular and biochemical mechanisms in a vertebrate context, including membrane traffic, signal transduction, and cell cycle (2).

Proteins in eukaryotic cells are organized according to their functions within a dynamic network of membranes. Localization is therefore paramount in assigning functions to uncharacterized proteins and understanding the processes occurring in subcellular compartments. An increased knowledge of the protein localization within the DT40 cell line would be of great value. Traditional localization methods such as immunofluorescence microscopy are typically low throughput and are more suitable to the study of specific proteins of interest rather than the cataloguing of large numbers of proteins. Recent developments in proteomics have made it possible to analyze the protein composition of organelles using a variety of different approaches. Several groups have utilized label-free quantitative proteomics in the high throughput assignment of proteins to subcellular compartments. In one approach, protein correlation profiling, proteins from enriched organelle fractions are quantified by peptide ion intensity measurements (3, 4). Other similar methods employ quantitation by spectral counting, recording the number of ions detected per protein (5, 6). Localization of organelle proteins by isotope tagging (LOPIT)1 is a complementary approach, which employs isotope labeling for quantitation (7–9). Rather than processing each sample separately as in label-free techniques, differentially labeled fractions are pooled early in the LOPIT protocol. This has the important advantage of reducing the points at which variation might be introduced into the data.

LOPIT begins with the partial separation of organelles by density gradient centrifugation and relies on the assumption that proteins from each organelle co-fractionate. Protein profiles along the gradient are quantified by the use of isotopically coded tags in conjunction with two-dimensional liquid chromatography of peptides and tandem mass spectrometry. Multivariate statistical techniques are then used to assign localizations to proteins by comparing their gradient profiles to those of established organelle markers in an unbiased manner. The major strength of such an approach is that it enables residents of different subcellular compartments to be resolved even if their gradient distributions overlap, and genuine organelle constituents can be readily distinguished from contaminants.

Here we use LOPIT to produce the first proteomic analysis of the major organelles of DT40. We have reproducibly iden-
tified 1090 proteins through the parallel analysis of preparations enriched for integral membrane or soluble and peripherally associated proteins. We use the distributions of 102 known organelle resident proteins as a basis to assign a further 223 proteins to five organelles: 79 to the endoplasmic reticulum (ER), 42 to the Golgi, 2 to the lysosome, 31 to the mitochondrion, and 69 to the plasma membrane (PM). We also demonstrate the resolution of components of the vesicular transport machinery. A striking finding is that a high proportion of identified proteins are not localized to a single organelle. This indicates that at steady state a substantial fraction of proteins are in transit between compartments, emphasizing the dynamic nature of intracellular organelles in eukaryotic cells. Our results represent the first application of LOPIT to a vertebrate system, provide the first organelle proteomic analysis of any lymphocyte cell line, and establish a major resource for the DT40 community.

**EXPERIMENTAL PROCEDURES**

**Cells**—The DKOR cell line was derived from the DT40 cell line by inactivation of both endogenous alleles of clathrin heavy chain and replacing them with human clathrin under the control of a tetracycline-regulatable expression system (10). Cells were maintained under clathrin-expressing conditions in RPMI 1640 with 10% fetal bovine serum (HyClone, Thermo Scientific), 1% chicken serum (Isobiotec). Debris and nuclei were pelleted by centrifugation for 10 min at 3000 × g for 1.5 h at 4 °C in a SW50.1 rotor (Beckman Coulter). Crude membranes collected from the interface were adjusted to an iodixanol cushion and centrifuged at 100,000 × g for 3 h at 4 °C in a VTi65.1 rotor (Beckman Coulter), and the supernatant was removed for separate analysis. The membrane pellet, enriched for integral proteins, was washed with 1 ml of H2O, repelleted and solubilized in 100 μl of 25 mM triethylammonium bicarbonate, 8 mM urea, 2% Triton X-100, 0.1% SDS. Protein concentrations were determined using the BCA protein assay kit (Thermo Scientific). Two independent density gradient fractionations were performed.

**iTRAQ Labeling**—Seven fractions were selected for analysis and divided into two groups for labeling with four-plex iTRAQ reagents (Applied Biosystems). Two four-plex comparisons were performed across the integral membrane samples (comparisons A and B), alongside two four-plex comparisons of the corresponding soluble/peripheral protein samples (comparisons C and D), as illustrated in supplemental Fig. S1. One common fraction was analyzed in each comparison to enable relative quantitation across all samples from a gradient, and 100 μg protein was labeled from each fraction. Two complete biological replicate experiments were performed. Fractions were labeled with iTRAQ reagents as follows; comparisons A and C: fraction 1 - reagent 114, 4–115, 13 and 14 (pooled to give a comparable protein concentration to other fractions) - 116, 21–117, comparisons B and D: 1–114, 9 and 10 (pooled) - 115, 16–116, 18–117. Samples were reduced with 4 mM Tris(2-carboxyethyl) phosphine (20 °C, 1 h) and cysteines blocked with 8 mM methyl methanethiosulfonate (20 °C, 10 min). Samples were diluted with 50 mM triethylammonium bicarbonate to reduce the urea concentration to <1 M, and proteins were digested with trypsin overnight at 37 °C (Promega; 2.5 μg added at 0, 1, and 2 h). Peptides were lyophilized, resuspended in 100 μl of 0.25 M triethylammonium bicarbonate, 75% ethanol, added to one unit of the corresponding iTRAQ reagent and incubated for 1 h at 20 °C. Unreacted reagents were hydrolyzed with 100 μl water for 15 min at 20 °C, and labeled peptides were pooled and lyophilized.

**Cation Exchange Chromatography and LC-MS/MS Analysis**—Peptide separation was performed using a BioLC HPLC system (Dionex) and polysulfoethyl A column (PolyLC) (2.1 × 200-mm inner diameter). Each pooled iTRAQ-labeled sample was resuspended in 5 ml of 20% ACN, 10 mM KH2PO4−H2PO4, pH 2.7, loaded, and washed for 60 min (200 μl/min). Peptides were eluted in an increasing KCI concentration in a series of linear gradient steps as follows: 0–5 min, 0–30 mM; 5–70 min, 30–125 mM; 70–105 min, 125–500 mM; 105–115 min, 0.5–1 M, 115–125 min, 1 M constant. Fractions were collected at 1-min intervals for the first 60 min and 2-min intervals thereafter. Peptides were lyophilized and resuspended in 0.1% formic acid (FA). LC-MS/MS analysis was performed using an Ultimate nano-LC system (Dionex) coupled to a QSTAR XL mass spectrometer (Applied Biosystems). Samples (60 μl) were loaded onto an Acclaim PepMap C18 trap column (5 mm × 300 μm inner diameter; Dionex) and washed for 25 min with 0.1% FA (20 μl/min). Reverse-phase chromatography was performed using a PepMap100 C18 analytical column (25 cm × 75 μm inner diameter; Dionex) with a flow rate of 150 nl/min. Peptides were eluted in an increasing concentration of ACN in 0.1% FA using the following linear gradient steps: 0–165 min, 5–32%; 165–190 min, 32–50%; 190–215 min, 50–90.5%. The QSTAR XL was operated in information-dependent acquisition mode in which 1 s MS scans were performed (400–1600 m/z) followed by 3 s product ion scans (100–1580 m/z) on the two most intense doubly and triply charged peptides.

**Processing of Mass Spectral Data**—QSTAR files were processed using wiff2dta (11) to generate centroided and uncentroided peak lists containing m/z and intensity information for each product ion spectrum. MASCOT v2.0.01 (Matrix Science) was used to search centroided peak lists against the IPI chicken database (v3.20; European Bioinformatics Institute) using the following modifications: fixed, iTRAQ four-plex (Lys), iTRAQ four-plex (N-terminal), MMTS (Cys); variable, oxidation (Met), iTRAQ four-plex (Tyr). MS and MS/MS tolerances were 0.2 Da and 0.5 Da, respectively. Peak lists were searched against a reversed version of the database to determine a MASCOT peptide score resulting in a false identification rate of <1%. Peptides scoring below this threshold were deleted. Normalized iTRAQ reporter ion areas were calculated from uncentroided peak lists using i-Tracker software (12). The Genome Annotating Proteome Pipeline system (13) was used to link identification information with quantitation data in a MySQL database (MySQL 4.0; MySQL AB). Peptides were quantified if they were unique to a single protein and at least three reporter ion peaks exceeded a 15 count intensity threshold (8).

**Multivariate Data Analysis**—Analysis was performed on proteins that were identified in both independent density gradients and quantified in all seven fractions analyzed from each. Proteins were required to be identified by multiple distinct peptides in at least one of the replicates. Because fraction 1 was labeled with reagent 114 in each four-plex experiment, iTRAQ reporter ion areas from comparisons B and D were adjusted so the 114 m/z peak areas of each peptide matched those of comparisons A and C, respectively. Reporter ion
ratios were derived from the normalized reporter ion areas, yielding 21 data points for each protein. Ratios were imported into SIMCA-P+ 11 (Umetrics) and pre-processed by logarithmic transformation and univariate scaling. Datasets were analyzed by principle components analysis (PCA) and proteins represented in two-dimensional scores plots. PCA is an unsupervised clustering algorithm used in this case to co-cluster proteins with similar profiles across the fractions. Proteins of known localization were used to build partial least squares-discriminant analysis (PLS-DA) models, which were used to predict the localizations of the remaining proteins. PLS-DA is a supervised extension of PCA and was used for predicting class membership. Independent PLS-DA models were built for the membrane proteins (comprising four classes: ER, Golgi, mitochondrion, and PM) and soluble/peripheral and complete datasets (five classes: ER, Golgi, mitochondrion, PM, and lysosome). The lowest scoring training set proteins were used as thresholds for class membership.

**Immunofluorescence Microscopy**—Antibodies and dilutions used were goat anti-IgM (1:100; Bethyl Laboratories), rabbit anti-clathrin heavy chain (1:200; Abcam), rabbit anti-Rab4 (1:200; Abcam), mouse anti-rab5 (1:50; BD Biosciences). The following were a gift from Dr. A. Peden (Cambridge Institute for Medical Research, Cambridge, UK): rabbit anti-sec22 (1:50) (14); rabbit anti-syntaxin 7 (1:200) (15); rabbit anti-syntaxin8 (1:100), mouse anti-syntaxin 16 (1:200) (15). Secondary antibodies used were Cy3-conjugated goat anti-rabbit IgG (1:100; Jackson ImmunoResearch Laboratories), fluorescein isothiocyanate-conjugated donkey anti-goat IgG (1:100; Abcam), fluorescein isothiocyanate-conjugated sheep anti-rabbit IgG (1:150; Sigma Aldrich) and Cy3-conjugated goat anti-mouse IgG (1:200; Jackson ImmunoResearch Laboratories).

Cells were centrifuged (350 rpm, 3 min) onto coverslips pre-coated with 0.05% polylysine and fixed with 4% paraformaldehyde in phosphate-buffered saline (10 min, RT). Cells were washed six times in phosphate-buffered saline containing 0.5% saponin, 0.1% Triton X-100 (wash buffer), and blocked with 5% bovine serum albumin in wash buffer (1 h, RT). Primary antibody incubations were carried out in blocking buffer for 3 h, RT. Following further washes, cells were incubated with appropriate secondary antibodies in blocking buffer for 1 h, RT. Cells were washed, and coverslips were mounted with Vectashield (Vector) and examined using an Olympus Fluoview IX81 laser scanning confocal microscope.

**RESULTS**

**Protein Identification and Distribution**—We maximized protein resolution and coverage by separating integral membrane from soluble/peripheral proteins within gradient fractions and analyzing them in parallel. We performed two overlapping four-plex comparisons on each set of samples across a gradient, quantifying peptides in seven fractions in total. Two independent density gradient separations were carried out. Two-dimensional liquid chromatography and tandem MS analysis of iTRAQ-labeled peptides resulted in the reproducible identification of 495 proteins in the integral membrane protein-enriched sample and 862 soluble/peripherally associated membrane proteins, with 290 proteins common to both samples. TMHMM v2.0 (16) predicted 41% of the enriched sample to be integral membrane proteins with at least one transmembrane span. In contrast, only 9% of the soluble/peripheral proteins contained putative transmembrane helices, and the majority of these were in the subset of proteins overlapping the two preparations.

Since the integral membrane-enriched and soluble/peripheral protein samples originated from the same gradient fractions, a third dataset was created for each replicate by combining the data from the two. This complete dataset contained all of the proteins identified in each gradient analysis, and quantitation was carried out on all peptides from each protein irrespective of whether they originated from the analysis of membrane or soluble samples. Processing the data in this way resulted in 1090 proteins being identified and quantified in both replicates. All protein and peptide information is presented in supplemental Tables S1 and S2.

The relative iTRAQ reporter ion intensities between fractions indicate the distributions of a protein within the density gradient, and residents of different organelles exhibit distinct profiles. This is illustrated in Fig. 1 using data from proteins identified in the integral membrane protein-enriched sample. Residents of the ER, Golgi, mitochondrion, and PM can be readily distinguished despite the incomplete gradient separation of organelles. There is some variation in the distributions of Golgi residents, with GCP372 and GalNAcT1 proteins displaying slightly different profiles to the other seven marker proteins. This may reflect differing distributions within the distinct structural subcompartments of the Golgi.

**Assignment of Proteins to Organelles**—For each protein, 21 reporter ion ratios were derived from the seven normalized reporter ion areas obtained in each experiment (supplemental Table S3). PCA was then used to identify patterns in each multivariate dataset and reduce it to a smaller set of latent variables, the principal components, which represented the dominant correlated variation in the dataset, constructed to be orthogonal to one another. PCA was performed on the integral membrane protein-enriched, soluble/peripheral, and complete protein data from each replicate independently. Each set of proteins was displayed on a PCA scores plot (Fig. 2) in which the axes represent the first and second principal components, and the scores given to each protein describe its position in relation to the other proteins contributing to the model. Proteins with similar localizations and, therefore, density gradient distributions show correlated variation and are given similar principal component scores and cluster together on the plot. This is illustrated by the distinct clusters formed by known organelle markers (Fig. 2).

Clusters corresponding to the ER, Golgi, mitochondrion, and PM are clearly separated on the scatter plot of the integral membrane protein-enriched dataset. The same organelle clusters are evident within the soluble/peripheral proteins alongside an additional grouping of lysosomal residents. The unclassified proteins in the lower part of the soluble/peripheral protein plot are largely cytosolic constituents. It should be noted that each of our PCA models had six principal components explained 65% of the variation within each dataset, and clusters such as the ER and mitochondrion that appeared to overlap on some of
the two-dimensional plots of PC1 and PC2 could be resolved by the additional components (supplemental Fig. S2).

Within the scatter plots, some proteins cluster with the known organelle markers and localizations can be assigned to these proteins based on this co-variation. Defining the positions of organelle cluster boundaries requires a supervised pattern recognition approach, so PLS-DA was applied. PLS-DA was used to classify observations in our multivariate datasets based upon training sets of observations with known class membership. Independent PLS-DA models were built for the membrane protein-enriched dataset ($Q^2 = 0.798$ and 0.802 for replicates 1 and 2, respectively, where a $Q^2 > 0.40$ is considered a robust model), the soluble/peripheral proteins ($Q^2 = 0.617$ and 0.518) and the complete dataset ($Q^2 = 0.683$ and 0.612). The training sets used to build the models (supplemental Table S4) comprised of four classes for the membrane proteins (ER, Golgi, mitochondrion, and PM) and five classes, also including the lysosome, for the soluble/peripheral and complete datasets. A protein was required to be assigned to the same organelle in both experimental repeats of an analysis to ensure high confidence classification.

A total of 223 proteins were assigned subcellular localizations by PLS-DA (supplemental Table S5): 79 to the ER, 42 to the Golgi, 2 to the lysosome, 31 to the mitochondrion, and 69 to the PM. The position of these proteins in relation to the known markers is shown on the PCA scatter plot of the complete dataset in Fig. 2. Many proteins were classified in more than one dataset, and there were no cases where these classifications disagreed.

Approximately 70% of proteins were not assigned localization. This amounted to 53% of the proteins identified in the membrane-enriched sample and 79% of the soluble/peripheral proteins, with the difference because of the larger proportion of cytosolic proteins in the latter. The unclassified proteins also include residents of organelles that could not be included in PLS-DA training sets because of the absence of a sufficient number of marker proteins, such as the endosomes, and proteins found in multiple compartments within the cell. The latter exhibit density gradient distributions that are combinations of those organelles in which they reside and appear on PCA plots in an intermediate location.

**ER**—We assigned 79 proteins to the ER. Several of these are annotated as “hypothetical” and have no significant ho-
mology to known proteins in other organisms. SignalP analysis of the amino acid sequences of our ER-classified proteins revealed 29 to contain predicted N-terminal signal peptides, which target proteins to the ER during translation (17). Ten proteins contain a C-terminal di-lysine motif and two an N-terminal di-arginine motif, both characteristic ER retrieval signals found in the cytoplasmic domains of transmembrane proteins (18, 19). Five proteins contain KDEL or a closely related sequence at their C termini, which is a retention signal for soluble proteins of the ER lumen (20). These proteins are listed in supplemental Table S6.

Several proteins assigned to the ER are likely to be involved in protein folding and modification. These include two chaperones, annotated as endoplasmin and hypoxia up-regulated 1, five protein disulfide isomerases and chicken homologs of dolichol-phosphate-mannose synthase, and an oligosaccharyl transferase subunit. Another functional class of ER residents is the enzymes involved in lipid metabolism, including lanosterol synthase and an acyl-CoA synthetase. The analysis also resulted in the localization of the small GTPase Rab18 to the ER. Rab proteins are members of the Ras-like GTPase superfamily and are distributed to distinct compartments of the endomembrane system, where they regulate transport between organelles (21). Rab18 has been previously associated with the ER and lipid droplets in other cell types (22, 23).

Seven ER-classified proteins probably function in the nucleus, including a nucleoporin (transmembrane protein 48) and importin 7, a nuclear import receptor (24). Although nuclei were removed from the preparations prior to density gradient separation, the presence of some such proteins is unsurprising because the ER forms a continuous membrane network with the nuclear envelope (25). In fact, the amino acid sequence of importin 7 contains a C-terminal di-lysine motif, potentially facilitating its retrieval to the ER membrane from where it can diffuse to the nuclear envelope.

Golgi—Our study assigned 42 proteins to the Golgi, three of which are likely to be false positive classifications of cytosolic residents: a proteasome subunit, tubulin-specific chaperone A, and microtubule-associated protein 4. The remainder of the classified proteins includes three that are annotated as hypothetical with no significant homology to other known proteins,
and a number of transmembrane superfamily members with poorly characterized function that can represent targets for future investigation.

Our Golgi residents include integral membrane transporters, protein modification enzymes, and calcium-binding proteins. Several proteins are likely to be involved in membrane fusion and vesicular transport. These include SNAP29 and three SNAREs (membrin, syntaxin 12, and VT12), which mediate the specific targeting of transport vesicles. Two sorting nexin (SNX) homologs were also assigned to the Golgi, SNX2, and SNX6. In mammalian cells these are components of the retromer complex that retrieve lysosomal enzyme receptors from endosomes and return them to the trans-Golgi network (TGN) (26, 27).

A striking finding is the relative paucity of Golgi proteins involved in glycosylation compared with previous LOPIT analyses (7, 8). Our gradient separation conditions were optimized to give the maximum overall resolution of all major organelles, and this approach inevitably limits a fully comprehensive analysis of each organelle proteome. The previous studies were carried out in Arabidopsis where the Golgi is heavily involved in the biosynthesis of cell wall polysaccharides, and it is probable that glycosylation enzymes are significantly less abundant in DT40.

Lysoosome—Only two further steady-state residents of the lysosome were determined using our stringent criteria for organelle classification. The first, gamma-glutamyl hydrolase, is an enzyme involved in folate and anti-folate metabolism that has been identified in a variety of other organisms (28) and localized to lysosomes in human cells (29). The second is homologous to the Xenopus protein designated LOC548400, and Pfam analysis of the amino acid sequence (30) reveals it to be a member of the papain family of cysteine peptidases.

Of the nine lysosomal marker proteins in our complete dataset only one, LIMP II, was an integral membrane protein. This illustrates a prominent feature of our analysis; a lysosomal protein cluster was only apparent in the soluble/peripheral protein sample. Rather than this being a problem with protein detection, we suggest that this reflects the steady-state protein distribution within the cell, and that there are actually relatively few membrane proteins uniquely localized to the lysosome. Although other proteomic studies focusing on lysosome membranes have reported a wide range of constituents (31, 32), these have been based on the analysis of enriched lysosomal preparations and represent catalogues of all the proteins associated with the organelle. Included are residents of multiple compartments or proteins only transiently associated with the lysosome. In contrast, LOPIT analyses the steady-state partitioning of proteins to establish their primary site of residency. A number of proteins previously reported as lysosomal (31, 32) were identified in our analysis but did not cluster with the established lysosomal markers, indicating that they were not unique to the lysosome or predominantly localized there in DT40. Examples include nicas-trin, a component of the gamma-secretase complex, which is located close to our PM cluster on PCA plots, and Niemann-Pick C1, found in an intermediate location close to the lysosome and Golgi clusters. Niemann-Pick C1 is involved in cholesterol trafficking and has been characterized as a late endosome resident that transiently associates with lysosomes and the TGN (33), consistent with our data.

Mitochondrion—We assigned 31 residents to the mitochondrion using a large training set of 40 markers. Only one, a proteasome subunit, appears to be a misassignment. Five proteins are designated hypothetical with no putative function or significant homology to other known proteins. We identified many proteins involved in classical mitochondrial functions such as the tricarboxylic acid cycle, electron transport, and ATP synthesis. Our mitochondrial residents also include metabolite transporters and prohibitin 2, part of an inner mitochondrial membrane protein complex that acts as a chaperone in the assembly of respiratory chain complex subunits (34, 35). Although the functions of most of our assigned proteins were consistent with inner membrane or matrix localizations, two likely residents of the outer membrane were detected, one a putative porin (MGC81430 protein) and the other a sorting and assembly machinery component 50 homolog. The latter is an essential subunit of the sorting and assembly machinery complex that integrates beta-barrel proteins into the outer membrane (36, 37).

The mitochondrion plays a critical role in cell death signaling pathways, and one of our classified proteins is known to function in this capacity. Peptidyl-tRNA hydrolase 2, also known as Bcl2-inhibitor of transcription 1 (Bit1), acts as a pro-apoptotic factor when released into the cytosol. Bit1 forms a complex with the transcriptional regulator AES and induces caspase-independent apoptosis upon loss of cell attachment (38).

Plasma Membrane—Consistent with the PM being the site of interaction between the cell and its environment, 17 (24.6%) of the 69 proteins localized here are transporters. Also identified were exocyst component 8, Sec8, and Sec15-like 1, three components of the exocyst complex involved in the docking of exocytic vesicles with the PM (39). Proteins involved in signal transduction, such as protein tyrosine phosphatases and receptor kinases, were significantly represented. Nine small GTPases were among the classified proteins, including two members of the Rab family, Rab8b and Rab35. Both have been localized to the PM in other cell types, and whereas the precise function of Rab8b remains unclear, Rab35 is also found in early endosomal compartments and is believed to regulate an endocytic recycling pathway (40–43).

Among the PM assignments is myristoylated alanine-rich protein kinase C substrate (MARCKS), a protein kinase C substrate that is membrane-bound in its unphosphorylated form (44). MARCKS proteins mediate cross-talk between signaling processes involving protein kinase C and calmodulin and are implicated in the regulation of the actin cytoskeleton.
Several more of our PM-classified proteins are also likely to be involved in membrane-cytoskeletal interactions. Although typically cytosolic, the association of these proteins with PM components evidently results in their co-fractionation. Four septin GTPase homologs were identified; septins form oligomeric complexes and are believed to coordinate changes in cytoskeletal and membrane organization (45). We also classified a Wiskott-Aldrich Syndrome Protein (WASP) family member and a Wiskott-Aldrich Syndrome-interacting protein. Wiskott-Aldrich Syndrome proteins function in the transmission of actin-remodeling signals from receptors and GTPases to the cytoskeleton (46). Similar roles are played by Ena/VASP (47) and ezrin, which cross-links actin filaments with PM components such as CD44 (48, 49).

Coated Vesicle Components—Among the proteins unasigned to the major organelles are components of the vesicular transport systems that mediate protein trafficking between intracellular compartments. There are three major classes of coated transport vesicle: COPI- and COPII-coated and clathrin-coated vesicles (CCVs). We identified coat components from all three, and they localized to distinct areas of the PCA plot (Fig. 3). All coat proteins are cytosolic until recruited onto a membrane; therefore their localizations represent average positions between their assembled and unbound forms.

Clathrin light chain was detected in our analysis and is expected to be a Wiskott-Aldrich Syndrome-Interacting Protein (WASP) family member and a Wiskott-Aldrich Syndrome-interacting protein. Wiskott-Aldrich Syndrome proteins function in the transmission of actin-remodeling signals from receptors and GTPases to the cytoskeleton (46). Similar roles are played by Ena/VASP (47) and ezrin, which cross-links actin filaments with PM components such as CD44 (48, 49).

Coated Vesicle Components—Among the proteins unasigned to the major organelles are components of the vesicular transport systems that mediate protein trafficking between intracellular compartments. There are three major classes of coated transport vesicle: COPI- and COPII-coated and clathrin-coated vesicles (CCVs). We identified coat components from all three, and they localized to distinct areas of the PCA plot (Fig. 3). All coat proteins are cytosolic until recruited onto a membrane; therefore their localizations represent average positions between their assembled and unbound forms.

Clathrin light chain was detected in our analysis and is found in an intermediate location on the PCA plot that reflects its distribution between CCVs, clathrin lattices found on endosomal compartments and unassembled, cytosolic protein. The outer CCV coat formed by clathrin is linked to membrane cargo by heterotetrameric adaptor protein (AP) complexes, and different APs are involved in distinct trafficking pathways. We detected three of the four classical APs, identifying two subunits each of AP1 and AP3 and one of AP2. AP1 and AP3 proteins are found close together on the PCA plot reflecting their similar intracellular distributions, with both having been localized to the TGN and endosomal membranes (50). AP1 is associated with protein sorting between the TGN and endosomes, and the more poorly characterized AP3 is believed to mediate trafficking to endosomes and lysosomes. A different localization is exhibited by AP2, which is found closer to the known PM residents in line with its function in receptor-mediated endocytosis.

COPI-coated vesicles traffic primarily from the Golgi to the ER and between Golgi cisternae (51). Seven COPI subunits are divided into two subcomplexes that are functional equivalents of clathrin and the APs. We identified six COPI proteins and these cluster together (Fig. 3). It should be noted that although this cluster overlaps that of the ER on the two-dimensional plot, PLS-DA only classified the β1-subunit as an ER resident, indicating that the two clusters are largely distinguishable by the additional dimensions of the dataset.

The COPII complex is responsible for ER to Golgi transport (52). The three COPII-coated vesicle components we identified do not form a tight cluster. The small GTPase Sar1a and its activating protein Sec23 are found close together on PCA plots, some distance away from the cargo-binding Sec24 subunit. Sec23 and Sec24 are known to form a cytosolic complex (53), therefore this separation is somewhat surprising. Four Sec24 isoforms, designated Sec24A–D, are expressed in mammalian systems (54) and proteins with close sequence matches to all four human Sec24 proteins are present in the chicken database. Human Sec24 isoforms show different but overlapping cargo binding preferences, and Sec24A is most important for ER export (55). It is possible that a functional difference between Sec24 proteins exists in DT40, and Sec24 is more greatly associated with another Sec24 isoform rather than the Sec24C homolog we identified.

Localization of the B Cell Receptor—DT40 is a pre-B cell line and thus expresses the B cell receptor, immunoglobulin (Ig)M, on the cell surface (56). We identified IgM chain and a protein with 94% sequence identity to Igλ, and expected the Ig chains to cluster with the plasma membrane markers. Instead both were found in an intermediate location on our PCA plot (Fig. 4, panel A) and were unclassified by PLS-DA. This implies that at steady state a significant proportion of IgM is contained within intracellular compartments. The proximity of IgM to clathrin on the PCA plot prompted us to examine the colocalization of these proteins within the cell. Immunofluorescence microscopy confirmed the IgM distribution to significantly overlap that of clathrin, with both displaying punctate intracellular and plasma membrane staining (Fig. 4, B–D). IgM also displays partial colocalization with the early/recycling endosome marker Rab4 (57) (Fig. 4, E–G).
Intracellular Distributions of Unassigned Proteins—Many proteins were found close to organelle clusters on PCA plots but remained unclassified by PLS-DA. To validate our experimental approach and confirm that these proteins do not have single steady-state localizations, we examined the intracellular distributions of some additional proteins by immunofluorescence microscopy. We focused on three unclassified SNAREs, sec22, syntaxin 7, and syntaxin 8, which all exhibited intermediate positions on our PCA plots (Fig. 5A), and compared their distributions to that of syntaxin 16, one of the Golgi training set proteins used for PLS-DA. Sec22 is reportedly involved in ER to Golgi transport (58) and in DT40 exhibited a punctate staining pattern throughout the cell, consistent with a localization within transport vesicles. Sec22 displayed a partial colocalization with syntaxin 16 (Fig. 5B). Syntaxin 8 was found close to the lysosome and Golgi clusters on our PCA plot (Fig. 5B), and it has been reported to be present in the TGN, endosomes, and lysosomes in other cell types (59). In DT40, syntaxin 8 exhibited a significant colocalization with syntaxin 16 but displayed a punctate distribution that extended beyond the Golgi (Fig. 5B). Syntaxin 7 was located close to the Golgi cluster on our PCA plot (Fig. 5A). Immunofluorescence revealed some overlap in distribution with syntaxin 16, but with considerable additional punctate perinuclear staining (Fig. 5B). Since syntaxin 7 has been localized to endosomal compartments in other cell types (59, 60), we also compared its distribution to that of the early endosome marker Rab5 and found there to be a partial overlap (Fig. 5B).

Reproducibility—To assess the reproducibility of the LOPIT approach we performed two independent density gradient fractionations. The principal component scores given to each protein in the two replicates were strongly correlated, with correlation coefficients of 0.89 obtained for the integral membrane protein-enriched sample, 0.80 for the soluble/peripheral proteins and 0.85 for the combined dataset when PC1 scores from replicates 1 and 2 were compared. The relationship between PC1 scores is illustrated in Fig. 6 using the combined dataset. PCA plots of the datasets from each replicate experiment are shown in supplemental Fig. S3. Proteins clustered...
Within the same organelles in different experiments, although the relative positions of organelle clusters showed some variability.

**DISCUSSION**

We have employed the LOPIT technique (7–9) to identify proteins and examine their subcellular distributions in the DT40 cell line. For this initial study, we designed a gradient protocol to resolve the major organelles. However, it should be emphasized that density gradient separation conditions can be readily modified to increase the resolution of particular organelles of interest. We identified 1090 different proteins with 99% confidence in two experiments, performing separate analyses of integral membrane protein-enriched and soluble/peripheral protein samples. Significantly more proteins were identified from each gradient separation using this approach compared with previous LOPIT analyses in other organisms (7, 8).

We performed two independent density gradient fractionations, allowing us to confirm the reproducibility of the LOPIT approach. There was a strong correlation between the principal components scores assigned to each protein in successive experiments. A greater degree of experimental variation was seen in the soluble/peripheral protein samples compared with the integral membrane protein-enriched samples. This was a probable consequence of the higher proportion of cytosolic proteins within the former. Since cytosolic constituents are not contained within a limiting membrane throughout the fractionation procedure, their gradient distributions are more likely to vary between experiments than those of organelle residents. This observation is important for the design of future LOPIT experiments, particularly where changes in protein localization under different conditions are to be compared. For studies concentrating on integral membrane protein distribution, enriching for these proteins by removing the soluble/peripheral protein components of the samples would ensure maximal reproducibility. Alternatively, combining the existing LOPIT methodology with a technique such as stable isotope labeling by amino acids in cell culture (SILAC) (61) would enable samples to be pooled prior to gradient fractionation, hence reducing experimental variability.

Emerging from our study is a picture of a highly dynamic system in which a large number of proteins are in transit between intracellular compartments, with 70% of our identified proteins not localizing to a single organelle. These included cytosolic constituents and steady-state residents of more than one compartment in addition to trafficking proteins, and all exhibited density gradient profiles that did not match those of the known organelle residents. The intermediate positions of these proteins on PCA plots represent their average localizations and provide valuable information about their distributions between the compartments in which they reside.

One limitation of techniques such as LOPIT is their reliance on previously established organelle marker proteins to define protein clusters within a PCA plot. This can limit the ability of the technique to localize novel proteins to poorly characterized or less abundant organelles. An example of this in our study was the late endosome, of which we identified constituents such as Rab7 but could not determine sufficient other proven residents to confidently define a cluster. LOPIT also relies on previously reported marker proteins being steady-state residents of a single organelle, and in a system where a high proportion of proteins are trafficking between compartments this cannot necessarily be assumed. This was illustrated by the distribution of the B cell receptor IgM, which was found some distance away from the PM cluster on PCA plots, and was confirmed by immunofluorescence microscopy to have a significant intracellular presence. It is possible that our cells are being stimulated by a chicken serum component in culture and are routinely internalizing IgM in CCVs and returning it to the plasma membrane via endosomal compartments.

Having demonstrated that the LOPIT method can be applied successfully to DT40 and established protein distributions within normal cells, we have generated a valuable resource for the DT40 community and laid the foundation for extensions of the study to examine changes in protein localization under perturbation of conditions. The variant of the DT40 cell line that we used here was derived by inactivation of both endogenous alleles of chicken clathrin heavy chain gene and the expression of human clathrin cDNA under the control of a tetracycline-regulatable expression system (10). By comparing LOPIT PCA plots of cells expressing or deficient in clathrin, we aim to highlight those proteins whose localization changes in the absence of clathrin. Utilizing such a powerful proteomic technique within the functional analyses made possible by DT40 will yield major insights into the dynamic processes occurring within eukaryotic cells.

**Acknowledgments**—We thank Pawel Sadowski and Dr. Gavin Jarvis for help with data processing; Dr. Andrew Peden for providing antibodies; and Dr. Nick Pugh for help with confocal microscopy.
The Organellar Proteome of DT40

1. Buerstedde, J. M., and Takeda, S. (1991) Increased ratio of targeted to random integration after transfection of chicken B cell lines. Cell 67, 179–188

2. Buerstedde, J. M., and Takeda, S. (eds) (2006) Reviews and Protocols in DT40 Research. Springer, New York, NY

3. Foster, L. J., de Hoog, C. L., Zhang, Y., Xie, X., Mootha, V. K., and Mann, M. (2006) A mammalian organelle map by protein correlation profiling. Cell 125, 187–199

4. Wiese, S., Gronemeyer, T., Ofman, R., Kunze, M., Grou, C. P., Almeida, J. A., Eisenacher, M., Stephan, C., Hainen, Y., Schollenberger, L., Korosec, T., Waterham, H. R., Schliwers, W., Erdmann, R., Berger, J., Meyer, H. E., Just, W., Azevedo, J. E., Wanders, R. J., and Warscheid, B. (2007) Proteomics characterization of mouse kidney peroxisomes by tandem mass spectrometry and protein correlation profiling. Mol. Cell. Proteomics 6, 2045–2057

5. Gilchrist, A., Au, C. E., Hiding, J., Bell, A. W., Fernandez-Rodriguez, J., Lesimple, S., Nagaya, H., Roy, L., Gosline, S. J., Hallett, M., Paomet, J., Kearney, R. E., Nilsson, T., and Bergeron, J. J. (2006) Quantitative proteomics analysis of the secretory pathway. Cell 127, 1265–1281

6. Kisinger, C., Cox, B., Kannan, A., Chung, C., Hu, P., Ignatchenko, A., Scott, M. S., Gramolini, A. O., Morris, Q., Hallett, M. T., Rossant, J., Hughes, T. R., Frey, B., and Emili, A. (2006) Global survey of organ and organellar protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 125, 173–186

7. Dunkley, T. P., Watson, R., Griffin, J. L., Dupree, P., and Lilley, K. S. (2004) Localization of organelle proteins by isotope tagging (LOPIT). Mol. Cell. Proteomics 3, 1128–1134

8. Dunkley, T. P., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, T. R., Frey, B., and Emili, A. (2006) Double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. EMBO J. 13, 1696–1705

9. Pelham, H. R. (1990) The retention signal for soluble proteins of the endoplasmic reticulum. Trends Biochem. Sci. 15, 483–486

10. Grosshans, B. L., Ortiz, D., and Novick, P. (2006) Rab5 and their effectors: achieving specificity in membrane traffic. Proc. Natl. Acad. Sci. U. S. A. 103, 11821–11827

11. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795

12. Jackson, M. R., Nilsson, T., and Peterson, P. A. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. EMBO J. 9, 3153–3162

13. Shutez, M. P., Peterson, P. A., and Jackson, M. R. (1994) An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. EMBO J. 13, 1696–1705

14. Chao, D. S., Kuo, C. S., and Scheller, R. H. (1997) Protein interactions regulating vesicle transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells. Cell 89, 477–487

15. Voeltz, G. K., Rolls, M. M., and Rapoport, T. A. (2002) Structural organization of the endoplasmic reticulum. EMBO Rep. 3, 84–93

16. Martin, S., Driessen, K., Nixon, S. J., Zerial, M., and Parton, R. G. (2005) Rab35 regulates localization of Rab18 to lipid droplets and induces their close apposition to the endoplasmic reticulum-derived membrane. J. Cell. Sci. 118, 2601–2611

17. Martin, S., Driessen, K., Nixon, S. J., Zerial, M., and Parton, R. G. (2005) Rab35 regulates localization of Rab18 to lipid droplets: effects of lipolytic stimulation and inhibition of lipid droplet catabolism. J. Biol. Chem. 280, 42325–42335

18. Shadforth, I. P., Dunkley, T. P., Hooper, H. B., Schram, C. H., and Bessant, C. (2008) Automated software for the confident identification of human peptides from tandem mass spectra. J. Proteome Res. 7, 189–197

19. Elsasser, H. P., Mann, M., and Hasilik, A. (2007) Integral and associated mitochondrial proteins. Cell 129, 145–150

20. Schieffer, C., Lemstrom, R., Petraglia, F., Mann, M., and Buggenhout, V. (1997) Peptide atlas of the human proteome. Science 275, 954–956

21. Elsasser, H. P., Mann, M., and Hasilik, A. (2007) Integral and associated lysosomal membrane proteins. Traffic 8, 1676–1686

22. Schweizer, M., and Göttlicher, M. (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Mol. Biol. Cell 173, 1521–1525

23. Takeda, S., Mochida, A., and Takeda, J. (2004) A mammalian organelle map by protein correlation profiling. Cell 125, 173–186

24. Kozjak, V., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H. E., Gruen, S., Gribanov, O., and Krammer, F. H. (2004) Interchangeable but essential functional elements of SNX1 and SNX2 in the association of retromer with endosomes and the trafficking of mannose 6-phosphate receptors. Mol. Cell. Biol. 24, 343–353

25. Schroder, B., Wrocklage, C., Pan, C., Jager, R., Kosters, B., Schafer, H., Schmitt, J., Elsasser, H. P., Mann, M., and Hasilik, A. (2007) Integral and associated lysosomal membrane proteins. Traffic 8, 145–150

26. Bagshaw, R. D., Mahuran, D. J., and Callahan, J. W. (2005) A proteomic analysis of lysosomal integral membrane proteins reveals the diverse composition of the organelle. Mol. Cell. Proteomics 4, 133–143

27. Higgins, M. E., Davies, J. P., Chen, F. W., and Ioannou, Y. A. (1999) Niemann-Pick C1 is a late endosome-resident protein that transiently associates with lysosomes and the trans-Golgi network. Mol. Genet. Metab. 68, 1–13

28. Nijtmans, L. G., de Jong, L., Arta Sanz, M., Coates, P. J., Berden, J. A., Back, J. W., Mulijers, A. O., van der Spek, H., and Grivel, L. A. (2000) Prohibitin acts as a membrane-bound chaperone for the stabilization of mitochondrial proteins. EMBO J. 19, 2444–2451

29. Nijtmans, L. G., Arta, S. M., Grivel, L. A., and Coates, P. J. (2002) The mitochondrial PHB complex: roles in mitochondrial respiratory complex assembly, ageing and degenerative disease. Cell. Mol. Life Sci. 59, 143–155

30. Kojda, J., Wiedemann, N., Milenkovic, D., Lohaus, C., Meyer, H. E., Guiard, B., Meisinger, C., and Pfanner, N. (2003) An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. J. Biol. Chem. 278, 48520–48523

31. Pfanner, N., Wiedemann, N., Meisinger, C., and Likhog, T. (2004) Assembling the mitochondrial outer membrane. Nat. Struct. Mol. Biol. 11, 1044–1048

32. Jan, Y., Matter, M., Pai, J. T., Chen, Y. L., Pilch, J., Komatsu, M., Ong, E., Fukuda, M., and Rusolati, E. (2004) A mitochondrial protein, Shh, mediates apoptosis regulated by integrins and Groucho/TLE corepressors. Cell 116, 751–762

33. Wang, S., and Hsu, S. C. (2006) The molecular mechanisms of the mammalian exocyst complex in exocytosis. Biochem. Soc. Trans. 34, 687–690

34. Kouranti, I., Sachse, M., Arrouche, N., Goub, D., and Echard, A. (2006) Rab35 regulates an endocytic recycling pathway essential for the termi-
nal steps of cytokinesis. *Curr. Biol.* **16**, 1719–1725

41. Patino-Lopez, G., Dong, X., Ben-Aissa, K., Bernot, K. M., Itoh, T., Fukuda, M., Kruhlak, M. J., Samelson, L. E., and Shaw, S. (2008) RAB35 and its GAP EPI64C in T-cells regulate receptor recycling and immunological synapse formation. *J. Biol. Chem.* **283**, 18323–18330

42. Sato, M., Sato, K., Liou, W., Pant, S., Harada, A., and Grant, B. D. (2008) Regulation of endocytic recycling by C. elegans Rab35 and its regulator RME-4, a coated-pit protein. *EMBO J.* **27**, 1183–1196

43. Armstrong, J., Thompson, N., Squire, J. H., Hayes, B., and Solari, R. (1996) Identification of a novel member of the Rab8 family from the rat basophilic leukaemia cell line, RBL2H3. *J. Cell Sci.* **109**, 1265–1274

44. Arbuzova, A., Schmitz, A. A., and Vergeres, G. (2002) Cross-talk unfolded: MARCKS proteins. *Biochem.* **362**, 1–12

45. Weirich, C. S., Erzberger, J. P., and Barral, Y. (2008) The septin family of GTPases: architecture and dynamics. *Nat. Rev. Mol. Cell Biol.* **9**, 478–489

46. Sechi, A. S., and Wehland, J. (2000) The actin cytoskeleton and plasma membrane connection: PtdIns(4,5)P(2) influences cytoskeletal protein activity at the plasma membrane. *J. Cell Sci.* **113**, 3685–3695

47. Krause, M., Dent, E. W., Bear, J. E., Loureiro, J. J., and Gertler, F. B. (2003) Ena/VASP proteins: regulators of the actin cytoskeleton and cell migration. *Annu. Rev. Cell Dev. Biol.* **19**, 541–564

48. Bretscher, A., Reczek, D., and Berryman, M. (1997) Ezrin: a protein requiring conformational activation to link microfilaments to the plasma membrane in the assembly of cell surface structures. *J. Cell Sci.* **110**, 3011–3018

49. Legg, J. W., and Isacke, C. M. (1998) Identification and functional analysis of the ezrin-binding site in the hyaluronan receptor, CD44. *Curr. Biol.* **8**, 705–708

50. Robinson, M. S. (2004) Adaptable adaptors for coated vesicles. *Trends Cell Biol.* **14**, 167–174

51. Bethune, J., Wieland, F., and Moelelen, J. (2006) COPI-mediated transport. *J. Membr. Biol.* **211**, 65–79

52. Tang, B. L., Wang, Y., Ong, Y. S., and Hong, W. (2005) COPII and exit from the endoplasmic reticulum. *Biochim. Biophys. Acta* **1744**, 293–303

53. Hicke, L., Yoshihisa, T., and Schekman, R. (1992) Sec23p and a novel 105-kDa protein function as a multimeric complex to promote vesicle budding and protein transport from the endoplasmic reticulum. *Mol. Biol. Cell* **3**, 667–676

54. Tang, B. L., Kausalya, J., Low, D. Y., Lock, M. L., and Hong, W. (1999) A family of mammalian proteins homologous to yeast Sec24p. *Biochem. Biophys. Res. Commun.* **258**, 679–684

55. Stoddart, A., Jackson, A. P., and Brodsky, F. M. (2005) Plasticity of B cell receptor internalization upon conditional depletion of clathrin. *Mol. Cell Biol.* **16**, 2339–2348

56. Van Der Sluijs, P., Hull, M., Zahraoui, A., Tavitian, A., Goud, B., and Mellman, I. (1991) The small GTP-binding protein rab4 is associated with early endosomes. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 6313–6317

57. Sacher, M., Stone, S., and Ferro-Novick, S. (1997) The synaptobrevin-related domains of Bos1p and Sec22p bind to the syntaxin-like region of Sed5p. *J. Biol. Chem.* **272**, 17134–17138

58. Prekeris, R., Yang, B., Oorschot, V., Klumperman, J., and Scheller, R. H. (1999) Differential roles of syntaxin 7 and syntaxin 8 in endosomal trafficking. *Mol. Biol. Cell* **10**, 3891–3908

59. Nakamura, N., Yamamoto, A., Wada, Y., and Futai, M. (2000) Syntaxin 7 mediates endocytic trafficking to late endosomes. *J. Biol. Chem.* **275**, 6523–6529

60. 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