Quantitative Proteomic Comparison of Rat Mitochondria from Muscle, Heart, and Liver*[S]

Francesca Forner‡§¶, Leonard J. Foster‖, Stefano Campanaro‡, Giorgio Valle‡, and Matthias Mann§‡‡**

Mitochondria, through oxidative phosphorylation, are the primary source of energy production in all tissues under aerobic conditions. Although critical to life, energy production is not the only function of mitochondria, and the composition of this organelle is tailored to meet the specific needs of each cell type. As an organelle, the mitochondrion has been a popular subject for proteomic analysis, but quantitative proteomic methods have yet to be applied to tease apart subtle differences among mitochondria from different tissues or muscle types. Here we used mass spectrometry-based proteomics to analyze mitochondrial proteins extracted from rat skeletal muscle, heart, and liver tissues. Based on 689 proteins identified with high confidence, mitochondria from the different tissues are qualitatively quite similar. However, striking differences emerged from the quantitative comparison of protein abundance between the tissues. Furthermore we applied similar methods to analyze mitochondrial matrix and intermembrane space proteins extracted from the same mitochondrial source, providing evidence for the submitochondrial localization of a number of proteins in skeletal muscle and liver. Several proteins not previously thought to reside in mitochondria were identified, and their presence in this organelle was confirmed by protein correlation profiling. Hierarchical clustering of microarray expression data provided further evidence that some of the novel mitochondrial candidates identified in the proteomic survey might be associated with mitochondria. These data reveal several important distinctions between mitochondrial and submitochondrial proteomes from skeletal muscle, heart, and liver tissue sources. Indeed approximately one-third of the proteins identified in the soluble fractions are associated predominantly to one of the three tissues, indicating a tissue-dependent regulation of mitochondrial proteins. Furthermore a small percentage of the mitochondrial proteome is unique to each tissue. ***Molecular & Cellular Proteomics 5:608–619, 2006.***

From the ‡Centro Ricerche Interdipartimentale Biotecnologie Innovative (CRIBI), University of Padua, V.le Colombo 3, I-35121 Padua, Italy, ‖Centre for Proteomics, Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada, and §Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany

Received, September 12, 2005, and in revised form, December 22, 2005

Published, MCP Papers in Press, January 14, 2006, DOI 10.1074/mcp.M500298-MCP200

Different mammalian tissues have distinct energy needs, and mitochondria morphology can vary widely, although the structure is not exclusively linked to respiration (1). Morphology and structure of mitochondria in mammalian skeletal muscle, heart, and liver are very different (2). In skeletal muscle mitochondria are distributed between sarcomeres and tightly embedded in the microfilaments (F-actin) and microtubules. Indeed the intimate association of mitochondria with the myofilaments minimizes diffusion distance and facilitates conversion of chemical energy to mechanical work. Mammalian skeletal muscle is a highly specialized tissue composed of fibers with a diverse range of properties. The aerobic type I fibers contain many mitochondria and rely on oxidative phosphorylation for ATP. Type IIa fibers are very rich in small mitochondria and are characterized by a high capacity for oxidative ATP generation. These mitochondria can only oxidize glycolytic products unlike those in type I fibers that can also utilize fatty acids and ketones. Type IIb fibers rely upon anaerobic glycolysis, contain few mitochondria, and have high levels of glycogen (3). Given these functional differences it is reasonable to assume that mitochondrial and mitochondria-related proteins are present in different amounts depending on the specific energy requirements of each muscle. Moreover the number of mitochondria per cell appears to vary significantly from cell type to cell type with estimates ranging from a few hundred to a few thousand per cell. Interestingly in recent reports mitochondria have been defined as "cellular mitochondrion" as they resemble a highly dynamic organelle that is often organized as a continuous reticulum (4) but that can fragment depending on the cell state.

Current models of the mitochondrial inner membrane depict it as continuous, and recently published electron tomography pictures show that cristae are continuous with the inner boundary membrane through a small number of narrow tube-like connections (4–6). It has been hypothesized that the intracristal and intermembrane space might exist as two separated compartments. The number and morphology of cristae are likely to reflect the response of the mitochondria to the energy demands of the cell. Highly folded, lamellar cristae with a large surface area are typically found in muscle and neurons where the respiratory rate is the greatest. So far, little is known about the control of cristae morphology in relation to function and activity and the control of this morphology by nuclear genes during development and differentiation (1).

Given such widely varying morphologies it is reasonable to expect that the protein composition of these organelles may vary significantly from tissue to tissue. To gain a better understanding of the organization of mitochondria we focused

our attention on the intermembrane space- and matrix-enriched fractions, investigating their protein composition at a quantitative level (Supplemental Fig. 1). Although mitochondrial matrix is quite well characterized with respect to the main metabolic pathways, it is widely accepted that many other proteins remain to be discovered. In addition, very little is known about the relative abundance of matrix proteins in the different tissues. However, given the structural complexity of mitochondria, the possibility of cross-contamination between subfractions must be taken into account for a correct interpretation of the results.

Recently several mitochondria proteomic studies have been published based on yeast (7), human (8), and mouse (9) tissues. In our previous study, we elucidated the protein composition of mitochondria from several mouse tissues (heart, brain, liver, and kidney), revealing that ~50% of the mitochondrial proteins occur in all the analyzed tissues (9). Moreover a quantitative comparison of different tissues by means of cDNA microarray expression analysis highlighted functional classes of genes that seem to be regulated in a tissue-specific manner. In the present study we characterized mitochondria from three *Rattus norvegicus* tissues: heart, liver, and skeletal muscle, the latter of which had not yet been analyzed by proteomics. This analysis provides new insights into the tissue-specific expression of mitochondrial proteins and reveals a number of candidates uniquely expressed in skeletal muscle and liver mitochondria.

EXPERIMENTAL PROCEDURES

Isolation of Rat Tissue Mitochondria—Male Wistar rats weighing ~300 g were starved overnight but had access to water *ad libitum*. After killing the animal by cervical dislocation, the liver was quickly removed and washed three times in 150 ml of 250 mM sucrose, 10 mM Tris-HCl, and 0.1 mM EGTA (pH 7.4). The buffer was supplemented with 2 mM PMSF (Applichem). The liver was then minced with scissors and homogenized with an Ikawerk Potter homogenizer (Teflon pestle). After diluting the homogenate to 300 ml with the same buffer, large cellular debris and nuclei were pelleted by centrifuging for 10 min at 600 × g_{max}. Mitochondria were pelleted by centrifuging the supernatant for 10 min at 7000 × g_{max}, and after suspending the pellet in 5 ml of the buffer, the sample was centrifuged again. Skeletal muscle and heart mitochondria were isolated as described previously (10). Skeletal muscle mitochondria were isolated from all rat leg muscles, whereas heart mitochondria were obtained by homogenizing 10 rat hearts for each experiment.

Mitochondrial pellets were gently resuspended in a small volume of buffer (250 mM sucrose, 0.1 mM EGTA, and 10 mM Tris-HCl, pH 7.4) to a final protein concentration of 8–15 mg/ml and further purified on a density gradient with 50% (v/v) Percoll (Amersham Biosciences) in 10 mM Tris-MOPS and 0.1 mM EGTA, pH 7.4. The lower band corresponding to mitochondria was recovered (11). Preparations of skeletal muscle mitochondria presented a relatively wider upper band even though care was taken to remove the "fluffy" layers from the pellets during isolation. Fractions (2–4 ml) were removed from the gradient with a pipette and washed twice before use.

Subfractionation of Mitochondria—The suspension of mitochondria (10–20 mg·ml^{-1} in 250 mM sucrose, 0.1 mM EGTA, and 10 mM Tris-HCl, pH 7.4) was diluted with 5 volumes of 10 mM Tris-HCl, pH 7.4, and gently stirred on a magnetic stirrer at 4 °C for 20 min. The hypotonically shocked mitochondria were then sedimented at 20,000 × g_{max} for 20 min, and the supernatant from this step was considered to contain intermembrane space proteins (12). Shocked mitochondria were resuspended in 10 mM Tris-HCl, pH 7.4, to a protein concentration of about 2 mg·ml^{-1} and left on ice to allow further swelling of the mitochondrial matrix space. After 5 min, "shrinking buffer" (one-third of the suspension volume) containing 1.8 M sucrose, 8 mM ATP, and 8 mM $MgCl_2$, adjusted to pH 7.4 with KOH, was added. The suspension was mixed carefully by three strokes in a loose fitting Dounce homogenizer. After 5 min, 10-ml aliquots were ultrasonicated for 3 × 5 s in an ice-ethanol bath. Total mitochondrial membranes were sedimented at 125,000 × g_{max} for 60 min. The supernatant, representing the matrix fraction, was filtered with Amicon Ultra-4 filters (molecular weight cutoff, 10,000; Millipore) by centrifugation at 6000 × g_{max} for 30 min.

Gel-enhanced LC-MS/MS and LC-MS/MS of Mitochondrial Proteins—For the in-gel separation approach, 40 μg of total mitochondrial protein (Bradford assay) were loaded onto 4–12% gradient polyacrylamide precast gels (NuPAGE Novex bis-Tris[1] gels, 1 mm, Invitrogen) and stained with colloidal Coomassie (Invitrogen). Gel lanes were cut into ~10 slices and subjected to in-gel tryptic digestion essentially as described previously (13). In brief, following complete destaining, gel slices were cut, washed with 50 mM ammonium bicarbonate, and shrunk with ethanol. Reduction/alkylation of proteins was performed with 10 mM DTT and 55 mM iodoacetamide. After two wash steps with ammonium bicarbonate/ethanol, the gel was dried with ethanol and incubated with 12.5 ng/μl trypsin (Promega) in 50 mM ammonium bicarbonate at 4 °C for 15 min. The supernatant was then discarded and replaced with 50 mM ammonium bicarbonate, and the reaction was allowed to proceed overnight at 37 °C. The reaction was stopped with 1% TFA, 0.5% acetic acid, and 3% acetonitrile, and the supernatant was recovered. Additional peptide extraction steps were performed with 30% acetonitrile and 100% acetonitrile. Supernatants were concentrated to low volume and then diluted to ~200 μl with 0.5% acetic acid, 3% acetonitrile, and 1% TFA.

For the in-solution approach, aliquots corresponding to 10 μg of total protein (Bradford assay) were diluted to 25 μl with dilution buffer (6 M urea, 2 M thiourea, 10 mM Hepes, pH 8), reduced, alkylated, and digested to peptides with trypsin and endopeptidase Lys-C as described previously (14). Resulting tryptic peptides were desalted and concentrated on reversed phase C_{18} StageTips (15). Liquid chromatography was performed on a 20-cm fused silica emitter (75-μm inner diameter) packed in-house with reversed phase ReproSil-Pur C_{18} (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). The injection volume was 2.5 μl, and the flow rate was 250 nl/min after a tee splitter. The mobile phase consisted of a step gradient of solvent A (0.5% acetic acid) and solvent B (0.5% acetic acid and 80% acetonitrile): 20-min injection time at 2% B, 2–8% B in 5 min, 8–33% B in 75 min, 33–50% B in 10 min, 50–80% B in 5 min, and 80% B for 5 min. The eluate was electrosprayed directly into a QSTAR pulsar triple quadrupole-time-of-flight mass spectrometer (ABI/MDS Sciex, Toronto, Canada). A spray voltage of 2400 V was applied. MS scans were performed for 1 s to select four intense peaks, and subsequently four MS/MS scans (1.5 s each) were performed. Raw data were processed with the software IDA (a feature of Analyst® QS, MDS Sciex) and searched against the mouse International Protein Index database (Human+Mouse+Rat, July 2004) (ftp.ebi.ac.uk/pub/databases/IPI/

[1] The abbreviations used are: bis-Tris, 2-[bis(2-hydroxyethyl)-amino]-2-(hydroxymethyl)propane-1,3-diol; GeLC, gel-enhanced LC (one-dimensional gel electrophoresis followed by liquid chromatography separation); GO, Gene Ontology; XIC, extracted ion chromatogram; PCP, protein correlation profile.

current/) using the Mascot Server, version 2.0, with the following parameters: trypsin specificity (16), one missed cleavage, cysteine carbamidomethylation as a fixed modification, methionine oxidation and protein amino-terminal acetylation as variable modifications, and ESI-Q-TOF fragmentation. Sequences of human keratins, porcine trypsin, and endopeptidase Lys-C were added to the database to reduce false positive identifications arising from fragmentation of those peptides. Protein identifications based on at least two different peptides were accepted. Only proteins identified with at least five peptides were validated by default without manual checking of the peptide fragmentation spectra. Furthermore only fragmentation spectra exhibiting a correct peptide sequence tag of at least three amino acids and consisting of at least four consecutive y-ions (17) were accepted. A total of 10,316 unique peptide sequences with scores greater than 16, more than 6 amino acids in length, and mass accuracies better than 30 ppm were sequenced in this study and led to the identification of 689 different proteins based on two or more peptides. 514 of these had a sequence coverage higher than 20% (Supplemental Table S1). MSQuant (msquant.sourceforge.net), open source software developed by our group (18), was used to parse and recalibrate peptide information from Mascot result files and then send it to a relational database.

Quantitative Analysis of Mass Spectrometry Data—MSQuant was used to extract and integrate ion chromatograms for all peptides. Where a peptide was not identified by tandem mass spectrometry, its elution time was predicted through correlation with parallel samples of other tissues to extract the correct ion current (19). Due to limitations in the linear dynamic range of the mass spectrometer, when measuring differences greater than 10-fold we report ratios greater than this limit with no standard deviation (14). The average R^2 value for the correlation of elution times between two consecutive LC-MS/MS experiments was greater than 0.99, indicating the reliability of elution time prediction. Peptide ion volumes (time $\times$ ion count $\times$ mass window (Da·s)) for peptides from each tissue were calculated by integration of the extracted ion chromatogram (XIC) (20).

Those proteins identified with only one peptide were excluded from the quantitation, and for the remaining proteins we considered the two/three highest intensity values within each set of peptides and calculated the sum of these *top 3* (or *top 2*) values (xPAI) (21). This result was a *partial sum* representing the quantitative value associated to that protein. By summing all the partial sums we obtained a *total sum* representing the global intensity of all the proteins identified in the tissue (*i.e.* corresponding to the total protein mass measured by Bradford). Each partial sum was then recalculated as percentage of the total sum to get a *relative percentage* of each protein with respect to the others in the same sample. The contribution of known contaminants (myoglobin, hemoglobin, myosin, actin, and human keratin) was subtracted from the total sum before calculation of the relative percentages. This normalization of the quantitative data was necessary to overcome errors associated with the initial Bradford assay and due to the global efficiency of proteolysis and to the loading and instrument response variability among different sample injections.

Protein Correlation Profiling—A crude skeletal muscle mitochondrial preparation (0.3 mg, Bradford) was resuspended with 250 mM sucrose, 0.1 mM EGTA, and 10 mM Tris-HCl, pH 7.4, to a final concentration of 1 mg/ml and layered onto a 14-ml continuous sucrose gradient (1.0-2.0 M) containing protease inhibitors (Roche Applied Science). Mitochondria were resolved by isopycnic centrifugation in a Sorvall Discovery 100 centrifuge (90 min at 26,500 rpm at 4 °C). Four fractions around the mitochondria band visible on the continuous gradient were collected and analyzed by gel-enhanced LC (GeLC)-MS/MS (see above). After using MSQuant to extract and integrate ion chromatograms, ion intensity profiles for each detected peptide were reconstructed across the gradient fractions, normalized, and then compiled into protein correlation profiles (PCPs) (19). PCPs of ATP synthase, fumarate hydratase, and NADH dehydrogenase were then used to establish a consensus profile for mitochondrial proteins against which all other PCPs were evaluated. In this case contaminating proteins would be expected to have a fractionation profile very different from that of the mitochondrial proteins. A χ^2 value was then calculated as a goodness of fit metric. Any protein where χ^2 was $\leq$0.2 was considered to be mitochondrial, whereas any protein where χ^2 was $>$0.2 was considered to be a contaminant.

Bioinformatics—For BLASTing (ftp.ncbi.nih.gov/blast/executables/) our sequences against the mouse Mitop database (672 sequences of the mitochondrial reference set, ihg.gsf.de/mitop2/showmousemask.do) we considered two proteins to be orthologous if they shared at least 70% amino acid sequence within the aligned segment and had BLAST e-values below 10^{-4}. 490 of 887 sequences (including those identified with just one peptide) did not have an orthologue based on these criteria.

In the GeLC-MS analysis, the protein sequences in the three lists (skeletal muscle, heart, and liver) were then BLASTed against each other to produce a list of "unique" candidate sequences. Skeletal muscle-specific and liver-specific lists were generated that contained proteins identified in that tissue with no identified orthologue in the other tissues. Sequences lacking a cellular localization in the Swiss-Prot database were predicted with Mitopred (22, 23) (mitopred.sdsc.edu), and our entire inventory was cross-referenced with the mitochondrial proteins we had reported earlier (9). Proteins found not to be unique to skeletal muscle for any reason were rejected. The same procedure was applied to obtain a list of liver-unique candidates.

To compare quantitative proteomic data with microarray data we downloaded 71 different rat microarray experiments that used the Affymetrix RG-U34A platform, each composed of 8798 different expression values. These experiments contained the expression profiles of 28 different tissues, including skeletal muscle (gastrocnemius and soleus, skeletal muscle generic), heart, and liver. As the microarray experiments were performed by different laboratories, we first normalized the data using the DNA-Chip Analyzer (dChip) program (24) (www.dchip.org/). The Affymetrix gene identifiers corresponding to the protein accession numbers obtained here were retrieved using NetAffx (www.affymetrix.com/analysis/netaffx/batch_query.affx). Of the 887 proteins identified by MS, the corresponding genes for 364 were found in the RG-U34A platform. Hierarchical clustering was performed with dChip where the distance between two genes was defined as $1 - r$ where r is the Pearson correlation coefficient.

RESULTS

Survey of Rat Mitochondrial Proteins—The principle aims of this study were to quantify tissue-specific differences in the mitochondrial matrix and to analyze the intermembrane space in different tissues. Through empirical observation, the swelling and shrinking treatment of initially highly pure mitochondria led to the most complete and effective separation of these two fractions. In total we prepared submitochondrial fractions from several different male rats and analyzed them across 15 different experiments by means of both protein in-gel and in-solution digestion followed by LC-MS/MS.

We then asked whether all these proteins had been reported in previous mitochondrial proteomic studies. Approximately 55% of the proteins identified in this study (887, including those identified with one peptide) did not have an apparent orthologue in Mitop (25) (this database includes the proteins identified by Mootha *et al.* (9)). Fifty-five of those with

no orthologue were skeletal muscle structural proteins or peroxisomal proteins and thus may be contaminants. The remaining 420 proteins were generally more abundant in skeletal muscle, a tissue where mitochondria have not been analyzed previously by proteomics.

Quantitative Analysis of Mitochondrial Proteins in Skeletal Muscle, Heart, and Liver—For the quantitative analysis of proteins we applied the in-solution digestion approach to matrix and intermembrane space proteins from heart, skeletal muscle, and liver. By correlating the elution times of sequenced peptides (see "Experimental Procedures") we were able to use the ion intensity of each peptide as a measure of relative abundance between the three tissue samples (19) (Fig. 1). Those proteins identified with only one peptide were excluded from the quantitation, and for the remaining proteins we used the three highest peptide intensity values (if available) from each as a measure of protein abundance. The contribution of known contaminants was excluded from the calculation. The reproducibility of this method of quantification was evaluated using the data for 75 protein hits measured in triplicate analyses of samples from the same animal. The ratio of the standard deviation to the arithmetic mean of intensities across all three measurements averaged 9% and ranged from 1 to 29% (Supplemental Fig. 2).

A protein was considered to be primarily associated with one tissue if it was at least 3 times more abundant there than in both of the other tissues. Table I contains those proteins associated exclusively with one tissue in at least three independent experiments or in two experiments if they also displayed a high mitochondrial targeting prediction score, and the complete list of analyzed proteins can be found in Supplemental Table S2. In Supplemental Table S2 we also included the mRNA abundance of those proteins that had a corresponding gene in the microarray and that displayed at least a 3-fold abundance preference for one tissue *versus* the others. By simply calculating every pairwise ratio (*e.g.* heart/liver and heart/skeletal muscle) for the microarray data, we estimated that on average 84% of the gene ratios changes in the same direction as the protein ratios. As expected, in the mitochondrial matrix we also identified "soluble" subunits of the electron transport chain that are well known to be exposed on the matrix face of the inner membrane and could have been partially removed from the membrane upon sonication (26, 27). This is only one of the possible explanations for the identification of small soluble polypeptides of cytochrome *c* oxidase (28). Indeed the matrix fraction could reasonably contain inner membrane vesicles as partial cross-contamination is unavoidable in mitochondrial preparation with the presently available isolation procedures. Therefore the accuracy of quantitation of the inner membrane soluble subunits is most likely affected by this problem. The b subunit of ATP synthase connects the F_1 and F_0 sectors of the enzyme and extends from the inner membrane to near the top of F_1 on the matrix side (29). Here we can again hypothesize that

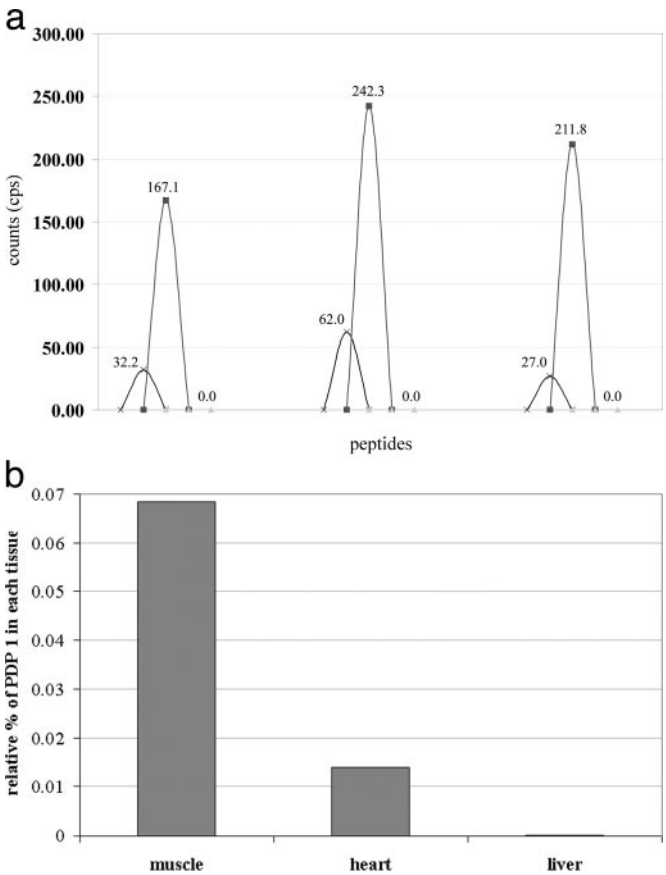

FIG. 1. **Example of relative quantitation for [pyruvate dehydrogenase (lipoamide)]-phosphatase 1 (*PDP 1*).** *a*, XIC of three of the MS-sequenced peptides (ALLPILQWHK, VTLGQMHGLLTER, and RVIESGPDQLNDNEYTK, displayed on the *abscissa* in this order) with the XICs corresponding to heart, skeletal muscle, and liver, respectively. XICs are offset from each other for visualization purposes. The protein was assigned to skeletal muscle based on the partial sums method described under "Experimental Procedures." This assignment confirms the ratios obtained from direct integration of the XICs. ×, heart; ■, skeletal muscle; ▲, liver. *b*, the sum of the XICs of the top 3 peptides in each tissue represents a partial sum that was then recalculated as percentage of the total sum (global intensity of all the proteins identified in the tissue corresponding to the total protein mass measured by Bradford) to get a relative percentage of each protein with respect to the others in the same sample (as represented by the bar chart). In general, the top 3 peptides might not always be exactly the same in different tissues. The ratios between relative percentages in different tissues provided the abundance of the protein in the tissues (see "Experimental Procedures"). As both the muscle/heart and muscle/liver ratios of [pyruvate dehydrogenase (lipoamide)]-phosphatase 1 were >3, this protein was considered associated with skeletal muscle mitochondria. *cps*, counts/s.

either the matrix fractions contained some submitochondrial particles or that the sonication affected the anchoring of the small membrane-spanning domain. In skeletal muscle mitochondria the differentially expressed [pyruvate dehydrogenase (lipoamide)]-phosphatase 1, pyruvate dehydrogenase phosphatase regulatory subunit, and the chaperone activity of bc_1 complex protein are all involved in the regulation at dif-

TABLE I

Subset of tissue-associated proteins by quantitative correlation

These proteins were associated exclusively with one tissue in at least three independent experiments or in two experiments if they also displayed a high mitochondrial targeting prediction score. The organelle information is based on the Swiss-Prot and GO annotations (Acc. no., Swiss-Prot/NCBI accession code; M, known mitochondrial protein; M.pred/PSORT, mitochondrial prediction scores by means of Mitopred or PSORT II (psort.nibb.ac.jp); % ratio, averaged ratios of the partial sums as indicated; H, heart, SM, skeletal muscle, L, liver; No. id., number of experiments that validated the association with the specified tissue). Chr, chromosome.

Protein description	Acc. no.	Organelle	M.pred	PSORT	% ratio		No. id.
					H/SM	H/L	
			%	%			
Heart							
Acetyl-CoA synthetase 2, similar to	XP_215897	M			6.7	9.9	4
Acyl coenzyme A thioester hydrolase	Q6IMX8	M			3.4	7.8	3
ATP synthase B chain	P19511	M			5.2	3.0	3
DNA segment, Chr 10, Johns Hopkins University 81 expressed	XP_215368		100		3.5	3.0	2
Heart fatty acid-binding protein	P07483		Non-M	4.3	≥10	≥10	4
Leucine zipper-EF-hand-containing transmembrane protein	XP_223541		84.6	17.4	≥10	≥10	3
NADH dehydrogenase (ubiquinone) 1α subcomplex 2, 8 kDa	XP_214570	M			9.5	7.3	3
NADH-ubiquinone oxidoreductase,13-kDa B subunit	Q63362	M			≥10	7.4	3
Propionyl-CoA carboxylase, β chain	P07633	M			4.2	7.2	3
RIKEN cDNA 2310005E10, similar to	XP_216117		84.6		≥10	≥10	2

Protein description	Acc. no.	Organelle	M.pred	PSORT	% ratio		No. id.
					SM/H	SM/L	
			%	%			
Skeletal muscle							
[Pyruvate dehydrogenase (lipoamide)]-phosphatase 1	O88483	M			7.1	≥10	3
Chaperone activity of bc_1 complex-like	XP_341163	M			8.6	3.0	4
NAD^+-specific isocitrate dehydrogenase, a-subunit	Q99NA5	M			3.5	≥10	3
Parvalbumin	P02625		Non-M		4.2	3.0	4
Pyruvate dehydrogenase phosphatase regulatory subunit	XP_226475	M			5.8	≥10	3
Hypothetical protein (similar to TU12B1-TY)	IPI00389238		Non-M		≥10	≥10	3
Voltage-dependent anion channel 1	Q9Z2L0	M			≥10	5.3	3
Voltage-dependent anion channel 2	P81155	M			9.9	5.1	3
Voltage-dependent anion channel 3	Q9R1Z0	M			6.8	≥10	3
Williams-Beuren syndrome critical region protein 21	XP_341105		92.3		3.0	5.1	2

Protein description	Acc. no.	Organelle	M.pred	PSORT	% ratio		No. id.
					L/H	L/SM	
			%	%			
Liver							
2310005D12Rik protein	P14604	M			6.1	6.3	3
2-Amino-3-ketobutyrate coenzyme A ligase	XP_345857	M			≥10	≥10	3
Aa2-174, similar to	Q7TMB6		Non-M		≥10	≥10	3
Ba1-651	Q7TP27	M			8.5	4.3	3
Carbamoyl-phosphate synthase	P07756	M			≥10	≥10	3
D4Ertd765e hypothetical protein, similar to	XP_216479		92.3		5.4	≥10	2
Farnesyl pyrophosphate synthetase (FPP synthetase)	P05369	Cytosol			≥10	6.7	3
Glutamate dehydrogenase	P10860	M			≥10	5.6	3
Glutaryl-CoA dehydrogenase, similar to	XP_344745	M			8.0	6.0	3
Glutathione *S*-transferase	P24473	M			≥10	≥10	3
Hydroxymethylglutaryl-CoA synthase	P22791	M			≥10	≥10	3
Hypothetical protein FLJ30596, similar to	XP_345168		Non-M		≥10	≥10	3
Hypothetical protein	XP_346610		Non-M		≥10	≥10	3
Hypothetical protein FLJ21963, similar to	XP_235174		92.3		5.1	8.4	3
KIAA1630 protein, similar to	XP_341558		92.3		≥10	≥10	3

TABLE I–*continued*

Protein description	Acc. no.	Organelle	M.pred	PSORT	% ratio		No. id.
					H/SM	H/L	
			%	%			
Kidney-specific protein	O70490		99		≥10	≥10	2
Kynurenine aminotransferase I	Q64602	M			3.1	3.3	3
Lactamase, β2, similar to	XP_216316		92.3		≥10	≥10	2
Maleylacetoacetate isomerase	P57113	M			≥10	≥10	3
Medium-chain acyl-CoA synthetase	XP_341918	M			≥10	≥10	3
NADPH:adrenodoxin oxidoreductase	P56522	M			4.1	5.5	3
NIPSNAP-related protein, similar to	XP_232955			69.6	≥10	≥10	2
Nonspecific lipid transfer protein	P11915	M			≥10	≥10	3
Ornithine aminotransferase, mitochondrial precursor	P04182	M			6.6	5.2	3
Ornithine carbamoyltransferase	P00481	M			≥10	≥10	3
Putative short-chain dehydrogenase/reductase	Q9WVK3	M			≥10	≥10	3
RIKEN cDNA 0610010D20, similar to	XP_215248	M			≥10	≥10	3
Sarcosine dehydrogenase	O88499	M			≥10	≥10	3
Transthyretin (4L369), similar to	XP_215112		92.3		≥10	≥10	2
Weakly similar to multifunctional aminotransferase	Q08415		Non-M		7.0	9.9	3
Weakly similar to SA, rat hypertension-associated	Q7TMB6		84.6		≥10	≥10	2

ferent levels of energy production. In particular, the latter protein was extremely abundant in all the analyzed samples. The a subunit of NAD^+-specific isocitrate dehydrogenase, a complex that is also involved in a key step of regulation, was associated with skeletal muscle in three independent experiments. The b subunit of this complex, although more abundantly in skeletal muscle, was identified in only one of the experiments and therefore was not included in Supplemental Table S2. This might reflect limitations in the mass spectrometry analysis. By globally comparing the skeletal muscle- and heart-associated proteins (Table I and Supplemental Table S2) we observed that proteins involved in the electron transport and in the oxidative phosphorylation are more abundant in heart mitochondria (suggesting a higher energy demand), whereas enzymes with regulative functions are more abundant in skeletal muscle. Surprisingly the voltage-dependent anion channels were also identified in the skeletal muscle matrix, although these proteins are classically thought to reside in the outer mitochondrial membrane. Their occurrence in the matrix preparation may be a reflection of the extremely high abundance of these proteins in skeletal muscle mitochondria.

This same quantitative approach was applied to analyze subfractions from liver and skeletal muscle mitochondria. Matrix- and intermembrane space-enriched fractions were prepared from both tissues and analyzed as described above. The association of a protein with one subfraction was based on the percentage ratio obtained from the quantitation, and proteins were again considered primarily associated with one subfraction if they were 3-fold more abundant in that subfraction. This study was performed on liver and skeletal muscle mitochondrial samples (Supplemental Tables S3 and S4). Reducing the sample complexity resulted in a sensitivity enhancement.

Despite the reduced sample complexity offered by mitochondrial subfractionation, some low abundance proteins could still be missed by in-solution digestion and direct LC-MS/MS analysis. Therefore, we used GeLC-MS (30) to further resolve matrix proteins from skeletal muscle and liver prior to LC-MS/MS analysis. Potential redundancy in the results from this analysis was reduced using BLAST (see "Experimental Procedures") to arrive at a list of unique candidate sequences. Table II contains the list of 21 proteins found only in skeletal muscle. Similarly a list of 238 proteins unique to liver was obtained from the BLAST analysis of 814 protein sequences obtained with GeLC-MS/MS (not shown).

Estimation of Mitochondrial Specificity—Purification protocols for mitochondria have been refined by several groups over many years. Despite this, the exquisite sensitivity of modern mass spectrometers has revealed that it is very difficult, if not impossible, to purify these or any other organelles to homogeneity, so potential contaminants must be dealt with in an unbiased way. One method to estimate the level of contamination is to express the sum total of known contaminants as a fraction of the whole protein mass. By calculating the partial sum values for myoglobin, hemoglobin, myosin, actin, and peroxisomal and endoplasmic reticulum proteins as a fraction of total sums we found that mitochondrial preparations from heart contained the highest levels of contaminants (5–14%) followed by muscle at 2–4% and by liver at 1–3% (Fig. 2).

A more accurate and unbiased method for assessing contamination is PCP (19), a method we have developed to overcome the specificity limitations inherent in biochemical isolation of subcellular fractions. With PCP it was possible to determine a characteristic common profile for mitochondrial and mitochondria-associated proteins based on the peptide quantitation across a number of sucrose gradient fractions. In

TABLE II

Proteins identified uniquely in skeletal muscle

Matrix protein samples were resolved and identified by GeLC-MS/MS. Upon BLASTing of the muscle sequences against the sequences identified in heart and liver matrix samples, we obtained a list of 21 proteins found only in skeletal muscle (Acc. no., Swiss-Prot or National Center for Biotechnology Information (NCBI) accession code; No. peptides, number of red or bold red peptides used for the identification by Mascot (All top ranking peptides are shown in red in a Mascot report. The first time a peptide match appears in the report it is shown as bold red (Matrix Science).)). If the protein was also identified in an intermembrane space fraction, then "ims" is indicated. The last two columns report the mitochondrial targeting score with Mitopred (only when >60%) and the subcellular localization of the protein if it had an annotation in the Swiss-Prot database (M, mitochondrial; SR, sarcoplasmic reticulum; ER, endoplasmic reticulum; IEA, inferred from electronic annotation). Proteins in bold either have a high prediction score or represent known mitochondrial proteins.

Name	Acc. no.	No. peptides	+Ims	Mitopred	Organelle
				%	
47-kDa heat shock protein	P29457	2			ER
Elongation factor G 2	**XP_226707**	**4**			M
AMP deaminase 1, similar to	P10759	19			
ATP-binding cassette, family F	**XP_138739**	**2**		**92.30**	
Calmodulin-like 3	P62161	2			
Calsequestrin, skeletal muscle isoform	P19633	2	ims		SR
Decorin precursor	Q01129	2			Secreted
DJ842G6.1.1	**XP_215857**	**2**		**92.30**	
Enigma protein, isoform 1	Q9Z1Z9	3	ims		M (IEA)
Hepatitis C virus core-binding protein 6	**XP_341574**	**2**		**92.30**	
Lactotransferrin		2			Secreted
MGC4767 protein	**XP_222227**	**2**		**100**	
Mitsugumin 23, similar to	**XP_342026**	**2**		**84.60**	
Hypothetical protein	Q5XIG1	5			
Protein CGI-51, similar to	XP_217005	11			
RIKEN cDNA 1810015H18	XP_215286	2			
Sarcoplasmic/ER calcium ATPase 1	Q64578	24	ims		
Seryl-aminoacyl-tRNA synthetase 2, similar to	**XP_214884**	**10**			M
Splice isoform DPII of desmoplakin	XP_214884	5			Desmosomes
TOM22, similar to	**Q75Q41**	**2**			M
Hypothetical protein (similar to TU12B1-TY)	IPI00389238	13			

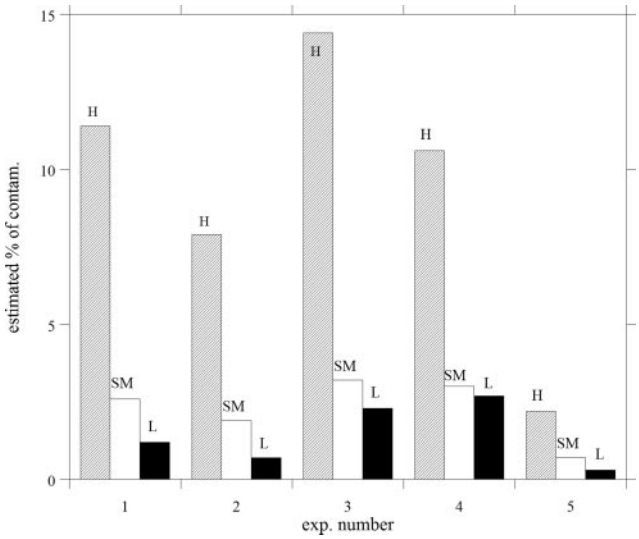

FIG. 2. **Mitochondria contamination estimate in heart, skeletal muscle, and liver mitochondria through five independent experiments.** Percentage of contamination is calculated based on partial sum values as a fraction of total sums for a number of non-mitochondrial proteins. *H*, heart; *SM*, skeletal muscle; *L*, liver. Heart mitochondria contain the highest percentage of non-mitochondrial proteins (mainly structural proteins such as myosin and actin) followed by skeletal muscle and liver. *contam.*, contamination.

organelle preparations there is always some degree of other cellular components co-fractionation, but PCP is extremely effective in discriminating organelle-associated proteins through their profile across the examined fractions. Indeed proteins not associated with mitochondria showed a completely different profile and could thus be easily distinguished. Here we applied PCP to confirm the mitochondrial location of a number of candidate proteins with uncertain or previously undefined subcellular localization. We used the PCPs of three accepted mitochondrial enzymes, fumarate hydratase, NADH dehydrogenase, and ATP synthase, as a standard against which to compare all other measured PCPs. Where the χ^2 value indicated a closely matching profile (see "Experimental Procedures") proteins were considered to be mitochondria-localized (Table III).

Hierarchical Cluster Analysis of mRNA Expression Data—Hierarchical clustering of expression data can provide valuable insight into co-regulated sets of genes, revealing previously undiscovered components of a pathway or an organelle (9). To this end we used hierarchical clustering of 8798 rat mRNA expression data from 28 tissues to identify novel putative mitochondrial proteins. Publicly available datasets collected on the RG-U34A chip set were clustered using dChip (see "Experimental Procedures") to identify gene clusters en-

TABLE III
Protein correlation profiling

The first group in the table represents a number of known mitochondrial proteins with $0.05 < \chi^2 < 0.2$. The second group includes a number of high confidence proteins not previously experimentally identified in mitochondria but with $0.05 < \chi^2 < 0.2$. Actin might perhaps be explained in terms of a tight association with the mitochondrial membrane. The last group shows the χ^2 value of some contaminants of mitochondria. VDAC, voltage-dependent anion channel.

Accession no.	Description	χ^2
Known mitochondrial proteins		
P15999	ATP synthase, α chain	0.14
P10719	ATP synthase, β chain	0.16
P14408	Fumarate hydratase	0.10
XP_215176	NADH dehydrogenase (ubiquinone) flavoprotein 1	0.05
XP_214977	$NADP^+$-specific isocitrate dehydrogenase	0.17
Q60587	Trifunctional enzyme, β subunit	0.12
XP_217267	Ubiquinol-cytochrome *c* reductase complex core 1	0.01
P32551	Ubiquinol-cytochrome *c* reductase complex core 2	0.11
Q9Z2L0	VDAC 1	0.13
Mitochondrial residents validated by PCP		
Q7TQ85	Ac1164, similar to (*in Swiss-Prot "mitochondrial" only by annotation*)	0.08
P03996	Actin, aortic smooth muscle	0.09
Q9JLG5	Brain protein 44-like protein	0.01
XP_216748	Coq6 protein, similar to	0.08
XP_227080	Elongation factor TU	0.20
XP_341250	NipSnap1 protein, similar to	0.06
XP_341644	Polymerase (RNA) II polypeptide C, similar to (*likely contaminant?*)	0.20
XP_217005	Protein CGI-51, similar to	0.11
P24142	Repressor of estrogen receptor activity (prohibitin)	0.08
Q91V64	Tumor-related protein	0.20
Likely contaminants		
XP_225259	Desmoplakin	0.86
Q6IFV1	Keratin complex 1, gene 14	1.23
Q63280	Keratin type II, cytoskeletal 5	0.83

riched in mitochondrial and putative mitochondrial genes that corresponded with proteins identified by mass spectrometry. We edited dChip to assign a novel GO term representative of the list of our 364 putative mitochondrial genes that had a corresponding protein identified by mass spectrometry (we named it "Mito_MS" term) and ran dChip to cluster all the genes in the platform. This resulted in the identification of three main gene clusters enriched in our mitochondrial genes to different extents. In particular, one of the clusters contained 187 Affymetrix oligonucleotide probes (corresponding to 163 different genes, Supplemental Table S5), and the corresponding proteins for 85 of these genes were also identified with mass spectrometry, representing 52% of the cluster members. The idea was that due to the high enrichment in mitochondrial genes of this cluster, it was likely that other members of the cluster may also be components of the mitochondria. We then used the mitochondrial prediction algorithm Mitopred to evaluate the 130 proteins for which we could not retrieve a mitochondrial GO annotation and obtained scores greater than 60% for 65 of these proteins. Of these, 35 had been identified by mass spectrometry, representing a class with extremely high confidence of mitochondrial localization because of independent evidence (occurrence in the cluster, MS-based organellar proteomics, and bioinformatic prediction). This group of proteins is marked in Supplemental Table S5 as "class 2," whereas "class 1" represents genes with mitochondrial annotation. Genes for which the corresponding proteins had high prediction scores with Mitopred but were not identified by mass spectrometry are marked as "class 3." Similarly we grouped the genes that were not predicted as mitochondrial as "class 4" (if identified by MS) and "class 5" (not identified by MS). Because the class 2 subset of proteins fulfills independent experimental criteria, it is very likely that it represents novel mitochondrial proteins (a selection of the most interesting of these proteins is reported in Table IV). This result demonstrates that the integration of different technologies is very effective in gaining new insights into the composition and organization of cells.

Functional Classification of Proteins and RNA—As an orthogonal approach to discovering tissue-specific functional classes of mitochondrial proteins we used the GoMiner program (31) to group proteins based on their Gene Ontology annotation. In heart, "ion transporter and ATPase activity" was one of the major terms, representing the oxidative phosphorylation and energy production machinery. In skeletal muscle, the three major terms were "response to stimulus," "protein binding," and "metal ion binding." In particular, the first term included various heat shock proteins such as 10-kDa heat shock protein (P26772), heat shock protein HSP90-β (P34058), α crystallin B chain (P23928), and 47-kDa

TABLE IV

Selection of interesting non-mitochondrial annotated proteins (based on GO) that were associated with the mitochondrial gene-enriched cluster and that were identified by mass spectrometry (class 2 and class 4 in Supplemental Table S5)

The cluster was enriched in the MS-identified proteins of the "Mito_MS" term (Supplemental Table S5); therefore proteins with unknown localization associated with this cluster may be considered *putative mitochondrial* or associated with mitochondria. In this table we extracted some of the proteins from Supplemental Table S5 that were identified by mass spectrometry and that were particularly interesting to us (Accession no., Swiss-Prot accession number of the protein corresponding to the gene; Description, gene name; H, heart; S.m., skeletal muscle; columns filled with an asterisk represent differential expression of the gene based on the microarray data; M.pred, Mitopred prediction score of the protein; —, predicted as non-mitochondrial).

Accession #	Description	H	S.m.	M.pred
				%
Q7TP78	Similar to Aa2-258	*	*	100
Q7TQ85	Similar to Ac1164 (*in Swiss-Prot "mitochondrial" only by electronic annotation*)	*		100
P05065	Aldolase A	*	*	61.50
Q9JLG5	Brain protein 44-like	*	*	69.20
P23928	Crystallin, α B chain	*	*	99
Q01129	Decorin	*	*	—
P02692	Fatty acid-binding protein 3	*	*	—
P04797	Glyceraldehyde-3-phosphate dehydrogenase	*	*	—
P19804	Nucleoside-diphosphate kinase	*	*	—
P61016	Phospholamban	*		69.20
XP_238570	QIL1	*		99
XP_231617	RIKEN cDNA D130059P03	*	*	84.6
XP_341781	RIKEN cDNA 0610042E07	*		—
XP_214074	Similar to myosin light chain 2a	*		99
P08733	Similar to myosin regulatory light chain 2, ventricular/cardiac muscle isoform	*		61.6
XP_215069	Elongation factor TU, EF-Tu (similar to RIKEN cDNA 2300002G02)	*		92.3

heat shock protein (P29457) together with several antioxidant proteins including peroxiredoxin 3 (Q9Z0V6), peroxiredoxin 5 (Q9R063), peroxiredoxin 6 (O35244), NAD(P)H dehydrogenase (quinone) 1 (P05982), and catalase (P04762). Finally in liver we found that proteins related to lipid and amino acid metabolism categories were the more represented. Major protein categories in the three tissues were compared with the GO terms obtained for the major heart, skeletal muscle, and liver gene clusters (corresponding to the tissue-expressed genes) obtained from hierarchical cluster analysis of the 364 genes that had a correspondent protein identified by mass spectrometry. These clusters are indicated as 1, 2, 3, and 7 in Supplemental Fig. 3. In liver, the highest score terms were basically identical for both proteins and genes. In heart, the highest score GO term for the gene cluster 1, "lipid metabolism," was not assigned with statistical significance in the heart protein classification (except for the related "lipid binding"), although by joining the gene clusters 1, 2, and 3 into a single cluster (corresponding to heart and skeletal muscle differentially expressed genes), we obtained in addition "ion transport" and "oxidative phosphorylation" as major terms in close similarity with those obtained for proteins.

DISCUSSION

Mitochondria are morphologically well defined cytoplasmic organelles, and a great deal of information is already known about the metabolism that occurs within them and the proteins that populate them. Nevertheless it is generally agreed that there remain many more mitochondrial proteins to be discovered. This is especially true for mitochondria from different tissues as much of the proteomic work on mitochondria has been done on organelles isolated from heart and liver (8, 32, 33). Previous proteomic studies of tissue-specific differences in mitochondria have been qualitative or have used cumulative statistics (9). Our analysis of whole mitochondria and submitochondrial fractions from rat liver, skeletal muscle, and heart resulted in the high confidence identification of 689 proteins. By applying several different methods to evaluate tissue-specific differences we have developed a clearer picture of how the different metabolic requirements of the tissues examined result in mitochondria of different compositions.

This quantitative study was based on comparing peptide ion intensities between LC/MS analyses, a method we have shown to give reliable estimates of relative protein abundances between complex samples (19). The analysis described here led to the assignment of 75 proteins associated preferentially with liver, 23 proteins associated with heart, and 22 proteins associated with skeletal muscle mitochondria (see Supplemental Table S2). In particular, 27 of the 75 liver-associated proteins were "liver-unique" according to the GeLC-MS/MS-based analysis. Our observation that the number of proteins unique to either the heart- or the skeletal muscle-associated groups was considerably less than for liver is likely due to the similarity of heart muscle and skeletal

muscle tissues and, by extension, their mitochondria. If we consider that ~400 different proteins were confidently identified on average in each single LC-MS experiment it can be estimated that about 30% of these mitochondrial proteins have a tissue-specific expression according to the criteria we have adopted ("Experimental Procedures"). Some parallels exist between these results and our previous qualitative study (9) where we had proposed a model in which half of the mitochondrial components are present in all tissues and the other half are tissue-specific. Indeed we expect that by increasing the number of analyzed tissues the number of proteins found to be tissue-specifically expressed through the quantitation would also increase.

As anticipated, the quantitative correlation also led to the identification of different proteins with high expression levels in either the matrix or the intermembrane space of the examined tissues. We focused mainly on the matrix compartment, where we observed that several of the tissue-specific proteins identified through the proteomic analysis could be correlated with the specific metabolic characteristics of the tissue. To assess the reliability of our mass spectrometry data, in the first instance we focused on the proteins that had already been extensively studied by means of biochemical methods and whose higher abundance in specific tissues is well established. As an example, we found that electron transport chain-related enzymes and subunits were highly abundant in heart mitochondrial matrix, and this is in agreement with the well known tissue-specific expression of complex IV (34, 35) (Supplemental Table S2). Similarly in agreement with what is already known, we found a high abundance of the urea cycle enzymes in the liver. These observations were an important starting point for our data analysis.

Enzymes related to the regulation of the skeletal muscle energy metabolism are highly expressed in skeletal muscle mitochondria, such as different phosphatases and kinases. At least two proteins with chaperone functions were also observed. We identified two interesting skeletal muscle-expressed proteins for which no previous experimental evidences of localization to mitochondria have been reported, namely the hypothetical protein similar to TU12B1-TY, which is reported to have 5′-nucleotidase activity, and Williams-Beuren syndrome critical region protein 21, which is involved in several biological processes, *e.g.* in xenobiotic metabolism.

Even though several metabolic pathways have been ascribed to the intermembrane space (36), the protein composition of this subcompartment is perhaps the least well known. The important role of the intermembrane space in mitochondrial protein import came to light with the discovery of Tim10p and Tim12p in yeast, two proteins that belong to a specific translocase of the inner membrane system specialized in the transport of hydrophobic proteins into the inner membrane (37). So far, efforts to find conventional chaperones in the intermembrane space have produced no results (38), and our data generally support these findings. The one exception to this was heat shock 71-kDa protein that was detected in liver mitochondria. Bioinformatic analysis of intermembrane space- or matrix-enriched proteins annotated in Gene Ontology or Swiss-Prot as non-mitochondrial provided additional support for the localization of these proteins to their respective mitochondrial substructures. H protein was identified in liver and skeletal muscle but was approximately 3 times more abundant in liver; the ratio between intermembrane fraction and matrix was also conserved in the two tissues (~3:1). HIRA-interacting protein (HIRA has been shown to directly interact with core histones) had a Mitopred score of 100% and showed a more than 3-fold enrichment in the skeletal muscle intermembrane space. For the protein "similar to Aa2-258" (100% score) evidence of its association with mitochondria emerged from hierarchical cluster analysis (Table IV), whereas NipSnap1, which was associated with the skeletal muscle intermembrane space, was putative mitochondrial in the correlation profiling analysis. Finally it is notable that some sequences had a very high mitochondrial prediction score with Mitopred, although they are associated with other organelles according to the Swiss-Prot and GO annotations. Specifically 3-ketoacyl-CoA thiolase A and acetyl-CoA acyltransferase 1 are only annotated as peroxisomal proteins, whereas aldehyde dehydrogenase 3 is associated with the endoplasmic reticulum.

As expected, the matrix-associated proteins according to this correlation were almost entirely known mitochondrial proteins. Some proteins with unknown or doubtful subcellular localization had 90–100% mitochondrial prediction score with Mitopred or PSORT II (39), such as acid phosphatase 6 (with orthophosphoric monoesters dephosphorylation activity), kidney-specific protein (AMP binding), and ES1 homolog, which is annotated as potentially mitochondrial in the Swiss-Prot database. Interestingly the gene for "kidney-specific protein" that had been described previously to exhibit a unique kidney-specific expression (40) was also associated with the liver matrix fraction in our study. Due to the high sensitivity of the mass spectrometry, it is likely that trace amounts of proteins were identified and quantitatively relevant in a relatively low protein density sample such as the intermembrane space. With few exceptions, the matrix and intermembrane space proteins appeared to be expressed in a tissue-specific manner. Indeed only programmed cell death protein 8, H protein, protein-disulfide isomerase, serum albumin, and tissue-specific isoforms of long-chain fatty acid-CoA ligase were identified in the intermembrane fraction of both skeletal muscle and liver. In mitochondria of both tissues just ~30 proteins were associated with the intermembrane space, and ~20 were associated with the matrix (based on the criteria that one protein was associated with one compartment only if it was at least 3 times more abundant there, see "Experimental Procedures"). These apparently low numbers indicate a similar composition of the soluble mitochondrial compartments, at least in the experimental conditions we adopted, even though

the correlation of the matrix and intermembrane space fractions was based on quite restrictive criteria (the 3:1 quantitative ratio for the association with one compartment).

One of the key points in organelle proteomics is the ability to identify novel proteins with high confidence. By reporting the mitochondria targeting score for many proteins, we have tried to integrate different and independent methods to provide support to the experimental identifications of novel proteins inside mitochondria. PCP provides an orthogonal, unbiased method for determining the specificity of organelle localization that avoids the potential pitfalls of sequence-based predictions or of literature annotation. For instance, the occurrence in mitochondrial preparations of proteins annotated in Gene Ontology or Swiss-Prot as cytosolic might be tissue-specific contamination but could also be either cytosolic proteins associating with the cytosolic face of mitochondria or proteins with multiple localizations. PCP analysis confirmed the mitochondrial localization of putative components identified here with mitochondria targeting prediction scores: Ac1164, 100%; 2810403L02Rik, 99%; repressor of estrogen receptor activity, 92.3%; and similar to Coq6, 92.3%. Moreover the hierarchical cluster analysis provided further support for the mitochondrial expression of Ac1164 and brain protein 44-like; indeed these genes showed expression patterns very similar to the cluster enriched in mitochondrial genes. In particular, brain protein 44-like was more abundant in skeletal muscle according to our quantitative results. The decorin gene was particularly interesting for us; indeed according to our analyses the corresponding protein was unique in the skeletal muscle mitochondria, and this tissue was investigated here for the first time.

In summary, according to the tissues we examined, the mitochondria composition showed a high degree of tissue-dependent variability. Indeed more than 200 proteins were identified uniquely in liver and not in muscle mitochondria with the highly sensitive GeLC-MS/MS. In addition, our quantitative correlation approach, by reducing the number of false negatives, showed that approximately one-third of the proteins identified in the soluble fractions can be confidently tissue-associated. Taken together, these results suggest that ~40% of the 689 proteins we identified with high confidence can be associated with one tissue either because of a tissue-unique expression or because of a differential expression.

Here we have increased to 689 the number of mitochondrial and putative mitochondrial previously identified proteins. This is a consequence of higher protein sequence coverage and of this being the first analysis of the skeletal muscle mitochondrial proteome. Moreover the exclusive use of a quantitative approach enabled a deeper insight into the organelle internal composition and the acquisition of novel information on the characteristics of the soluble compartments. In particular, we have shown that the combination of independent approaches (*i.e.* mass spectrometry-based organelle analysis, microarray cluster analysis, and bioinformatic prediction) can provide some evidence of targeting to mitochondria for novel candidates.

Acknowledgments—F. F. is particularly grateful to Professor Paolo Bernardi from the University of Padua for expert advice concerning mitochondria isolation. We thank Peter Mortensen for programming assistance and members of the Center for Experimental Bioinformatics (which is supported by the Danish National Research Foundation) for fruitful discussions.

* This work was supported by Italian Telethon Foundation Grant GSP042894B. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

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

¶ Supported by a Marie Curie training site fellowship.

** To whom correspondence may be addressed. E-mail: giorgio.valle@unipd.it.

‡‡ To whom correspondence may be addressed. E-mail: mmann@biochem.mpg.de.

REFERENCES

1. Scheffler, I. E. (1999) *Mitochondria*, Wiley-Liss, New York
2. Fawcett, D. W. (1981) *The Cell*, pp. 410–485, W. B. Saunders, Philadelphia
3. Engel, A. G., and Franzini-Armstrong, C. (2004) *Myology*, pp. 87–103, McGraw-Hill, New York
4. Capaldi, R. A. (2000) The changing face of mitochondrial research. *Trends Biochem. Sci.* **25,** 212–214
5. Mannella, C. A. (2000) Introduction: our changing views of mitochondria. *J. Bioenerg. Biomembr.* **32,** 1–4
6. Frey, T. G., and Mannella, C. A. (2000) The internal structure of mitochondria. *Trends Biochem. Sci* **25,** 319–324
7. Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H. E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B., Rehling, P., Pfanner, N., and Meisinger, C. (2003) The proteome of Saccharomyces cerevisiae mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 13207–13212
8. Taylor, S. W., Fahy, E., Zhang, B., Glenn, G. M., Warnock, D. E., Wiley, S., Murphy, A. N., Gaucher, S. P., Capaldi, R. A., Gibson, B. W., and Ghosh, S. S. (2003) Characterization of the human heart mitochondrial proteome. *Nat. Biotechnol.* **21,** 281–286
9. Mootha, V. K., Bunkenborg, J., Olsen, J. V., Hjerrild, M., Wisniewski, J. R., Stahl, E., Bolouri, M. S., Ray, H. N., Sihag, S., Kamal, M., Patterson, N., Lander, E. S., and Mann, M. (2003) Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell* **115,** 629–640
10. Fontaine, E., Eriksson, O., Ichas, F., and Bernardi, P. (1998) Regulation of the permeability transition pore in skeletal muscle mitochondria. Modulation by electron flow through the respiratory chain complex I. *J. Biol. Chem.* **273,** 12662–12668
11. Mickelson, J. R., Greaser, M. L., and Marsh, B. B. (1980) Purification of skeletal-muscle mitochondria by density-gradient centrifugation with Percoll. *Anal. Biochem.* **109,** 255–260
12. Sottocasa, G. L., Kuylenstierna, B., Ernster, L., and Bergstrand, A. (1967) An electron-transport system associated with the outer membrane of liver mitochondria. A biochemical and morphological study. *J. Cell Biol.* **32,** 415–438
13. Schevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68,** 850–858
14. Foster, L. J., De Hoog, C. L., and Mann, M. (2003) Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 5813–5818
15. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670

16. Olsen, J. V., Ong, S. E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **3,** 608–614
17. Roepstorff, P., and Fohlman, J. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* **11,** 601
18. Schulze, W. X., and Mann, M. (2004) A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* **279,** 10756–10764
19. Andersen, J. S., Wilkinson, C. J., Mayor, T., Mortensen, P., Nigg, E. A., and Mann, M. (2003) Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426,** 570–574
20. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **1,** 252–262
21. Rappsilber, J., Ishihama, Y., Foster, L. J., Mittler, G., and Mann, M. (2003) Approximate relative abundance of proteins within a mixture determined from LC-MS data, in *Abstracts of the 51st American Society for Mass Spectrometry Conference in Mass Spectrometry and Allied Topics, Montreal, Canada, June 8–12, 2003*, American Society for Mass Spectrometry, Santa Fe, NM
22. Guda, C., Fahy, E., and Subramaniam, S. (2004) MITOPRED: a genome-scale method for prediction of nucleus-encoded mitochondrial proteins. *Bioinformatics* **20,** 1785–1794
23. Guda, C., Guda, P., Fahy, E., and Subramaniam, S. (2004) MITOPRED: a web server for the prediction of mitochondrial proteins. *Nucleic Acids Res.* **32,** W372–W374
24. Schadt, E. E., Li, C., Ellis, B., and Wong, W. H. (2001) Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. *J. Cell Biochem. Suppl.* Suppl. 37, 120–125
25. Andreoli, C., Prokisch, H., Hortnagel, K., Mueller, J. C., Munsterkotter, M., Scharfe, C., and Meitinger, T. (2004) MitoP2, an integrated database on mitochondrial proteins in yeast and man. *Nucleic Acids Res.* **32,** D459–D462
26. Darley-Usmar, V. M., Rickwood, D., and Wilson, M. T. (1987) *Mitochondria, a Practical Approach*, pp. 108–111, IRL Press, Oxford
27. Knowles, A. F., and Penefsky, H. S. (1972) The subunit structure of beef heart mitochondrial adenosine triphosphatase. Isolation procedures. *J. Biol. Chem.* **247,** 6617–6623
28. Patel, S. D., Cleeter, M. W., and Ragan, C. I. (1988) Transmembrane organization of mitochondrial NADH dehydrogenase as revealed by radiochemical labelling and cross-linking. *Biochem. J.* **256,** 529–535
29. Dunn, S. D., Revington, M., Cipriano, D. J., and Shilton, B. H. (2000) The b subunit of Escherichia coli ATP synthase. *J. Bioenerg. Biomembr.* **32,** 347–355
30. 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
31. Zeeberg, B. R., Feng, W., Wang, G., Wang, M. D., Fojo, A. T., Sunshine, M., Narasimhan, S., Kane, D. W., Reinhold, W. C., Lababidi, S., Bussey, K. J., Riss, J., Barrett, J. C., and Weinstein, J. N. (2003) GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol.* **4,** R28
32. Jiang, X. S., Dai, J., Sheng, Q. H., Zhang, L., Xia, Q. C., Wu, J. R., and Zeng, R. (2005) A comparative proteomic strategy for subcellular proteome research. ICAT approach coupled with bioinformatics prediction to ascertain rat liver mitochondrial proteins and indication of mitochondrial localization for catalase. *Mol. Cell. Proteomics* **4,** 12–34
33. Arnold, R. J., Hrncirova, P., Annaiah, K., and Novotny, M. V. (2004) Fast proteolytic digestion coupled with organelle enrichment for proteomic analysis of rat liver. *J. Proteome Res.* **3,** 653–657
34. Capaldi, R. A., Marusich, M. F., and Taanman, J. W. (1995) Mammalian cytochrome-c oxidase: characterization of enzyme and immunological detection of subunits in tissue extracts and whole cells. *Methods Enzymol.* **260,** 117–132
35. Linder, D., Freund, R., and Kadenbach, B. (1995) Species-specific expression of cytochrome c oxidase isozymes. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **112,** 461–469
36. Stuart, R. A., Ono, H., Langer, T., and Neupert, W. (1996) Mechanisms of protein import into mitochondria. *Cell Struct. Funct.* **21,** 403–406
37. Jarosch, E., Tuller, G., Daum, G., Waldherr, M., Voskova, A., and Schweyen, R. J. (1996) Mrs5p, an essential protein of the mitochondrial intermembrane space, affects protein import into yeast mitochondria. *J. Biol. Chem.* **271,** 17219–17225
38. Koehler, C. M., Merchant, S., and Schatz, G. (1999) How membrane proteins travel across the mitochondrial intermembrane space. *Trends Biochem. Sci.* **24,** 428–432
39. Nakai, K., and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* **24,** 34–36
40. Hilgers, K. F., Nagaraj, S. K., Karginova, E. A., Kazakova, I. G., Chevalier, R. L., Carey, R. M., Pentz, E. S., and Gomez, R. A. (1998) Molecular cloning of KS, a novel rat gene expressed exclusively in the kidney. *Kidney Int.* **54,** 1444–1454