Quantitative proteomic analysis of single pancreatic islets

Leonie F. Waanders, Karolina Chwalek, Mara Monetti, Chanchal Kumar, Eckhard Lammert, and Matthias Mann

Abstract

Technological developments make mass spectrometry (MS)-based proteomics a central pillar of biochemical research. MS has been very successful in cell culture systems, where sample amounts are not limiting. To extend its capabilities to extremely small, physiologically distinct cell types isolated from tissue, we developed a high sensitivity chromatographic system that measures nanogram protein mixtures for 8 h with very high resolution. This technology is based on splitting gradient effluents into a capture capillary and provides an inherent technical replicate. In a single analysis, this allowed us to characterize kidney glomeruli isolated by laser capture microdissection to a depth of more than 2,400 proteins. From pooled pancreatic islets of Langerhans, another type of “miniorgan,” we obtained an in-depth proteome of 6,873 proteins, many of them involved in diabetes. We quantitatively compared the proteome of single islets, containing 2,000–4,000 cells, treated with high or low glucose levels, and covered most of the characteristic functions of beta cells. Our ultrasensitive analysis recapitulated known hyperglycemic changes but we also find components up-regulated such as the mitochondrial stress regulator Park7. Direct proteomic analysis of functionally distinct cellular structures opens up perspectives in physiology and pathology.

Mass spectrometry (MS) is inherently an extremely sensitive analysis technique: single proteins have been analyzed at the attomole level many years ago (1, 2). Similarly, it is now possible to identify thousands of proteins in single experiments and identification and quantitation of a comprehensive proteome has been reported (3). However, the complexity of proteomic mixtures typically necessitates micrograms or milligrams of starting material—protein amounts that can readily be extracted from cell or tissue homogenates (4, 5). The analysis of small tissue substructures with specialized functions is medically important but has remained much more difficult. Ideally these structures are separated from the surrounding tissue to enrich proteins specific to these areas. Although such selection is possible by laser-capture microdissection (LCM), successful use of this technology has so far overwhelmingly been in the areas of RNA and DNA assays. In the few cases where LCM was combined with MS analysis, many thousands cells were collected (taking presumably several hours or days of selection), to obtain the protein amounts typical for MS (6). Only a few investigators have attempted to measure smaller cell numbers. Luider and coworkers reported 1,000 identified proteins from 3,000 cells and a similar number from LCM-selected breast cancer tissues, but they did not sequence peptides by mass spectrometry, relying instead on mass measurements only (7, 8).

Main obstacles in analyzing small protein amounts are the considerable sample losses during preparation and the limited “dynamic range” of liquid chromatography-tandem mass spectrometry (LC-MS/MS) systems used in proteomics. Sample preparation efficiency has improved by the advent of polished sample tubes (e.g., Eppendorf Protein LoBind) and MS-friendly detergents (9, 10), thereby allowing shorter extraction and digestion protocols and reduced unspecific surface binding.

The second important obstacle, the dynamic range, refers to the number of different ion species that can be observed simultaneously. With highly complex samples, as are typical for proteomic experiments, low intensity ions are likely to be masked by more abundant species. To also detect these low abundance proteins, extensive fractionation is commonly applied. But this is not possible with very low starting amounts as sample losses would be prohibitive.

Recently, we reported the development of a chromatographic system that allows direct reanalysis of the injected sample without losing signal intensity (11). In this LC-MS interface, termed RePlay, post-column effluent is split with one part directed to on-line LC-MS/MS analysis and the other one stored in a capture capillary for another round of analysis after the first one is completed. We envisioned that this setup should also be valuable for analyses where starting amount is limiting, and therefore developed the system for high sensitivity “single shot” analysis. We used optimized sample preparation methods and the RePlay system in combination with the sensitive LTQ-Orbitrap mass spectrometer. Due to the split, flow rates could be reduced 5-fold compared to normal operation in our laboratory while maintaining normal backpressures and a continuous stable flow and spray. Importantly, the re-analysis of the sample does not lose signal intensity because the peptide concentration is not affected by the split. Thus, we inherently obtain a technical replicate, increasing quantitation accuracy, even from the smallest proteomic samples. In addition the doubled gradient time increases peptide sequencing events per sample.

Here we demonstrate that our workflow enables the analysis of small “miniorgans”: we examined mouse glomeruli selected by LCM and single islets of Langerhans, handpicked after purification. In both cases the starting material was less than 400 ng of proteins, —but still resulted in more than 2,000 protein identifications. Besides high confidence identifications, we demonstrate quantitation of protein expression levels in single islets of Langerhans. These islets were stimulated with high or low glucose before LC-MS/MS and we identified many responders.

Author contributions: L.F.W., K.C., E.L., and M. Mann designed research; L.F.W. and K.C. performed research; L.F.W., M. Monetti, and C.K. analyzed data; and L.F.W. and M. Mann wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

1Present address: Philips Research, Molecular Diagnostics, High Tech Campus 12, 5656 AE Eindhoven, The Netherlands.

2Present address: Lilly Singapore Centre for Drug Discovery, 8A Biomedical Grove #02–05, Immunos, Biopolis, 138648, Singapore.

3To whom correspondence should be addressed. E-mail: mmann@biochem.mpg.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0908351106/DCSupplemental.
Like anti-oxidants and proteins involved in glucose catabolism or vesicle secretion processes.

**Results and Discussion**

There are two challenges in the measurement of very small in vivo cell populations by liquid chromatography-tandem mass spectrometry (LC-MS/MS): efficient recovery of the proteins and high analytical sensitivity. To meet the first challenge, we implemented a workflow with a minimal number of handling steps starting from LCM (Fig. 1A) or miniorgan purification (Fig. 1B) and resulting in a trypsin digested proteome (Materials and Methods). LC-MS/MS sensitivity was improved in three ways (Fig. 1C): (i) by splitting the flow after chromatographic peptide separation, thereby providing a technical replicate from a single sample loading without loss of signal [RePlay (11)]; (ii) by reducing the effective flow rate to 50 nL/min and employing smaller diameter columns; and (iii) by separating the peptide mixture with long gradients (4 h) to increase peptide sequencing events by the high-resolution LTQ-Orbitrap mass spectrometer.

First, we tested this high-sensitivity workflow on mouse glomeruli, about 50 of which we isolated by LCM from histological slices (Materials and Methods). Glomeruli are the filtering units in the kidney and consist of only a few hundred cells in mice. Automated analysis proved robust and the chromatographic setup retained excellent average peak width (< 22s) in direct and replicate ("replay") runs (Fig. 1D). Overall signal was the same in both runs and individual peptide intensities, which ranged over more than four orders of magnitude, correlated highly between them \( R = 0.92 \) (Fig. S1A). Data processing with MaxQuant (12) resulted in the identification of 2406 distinct proteins from a single injection. Virtually all known glomeruli proteins are contained in this dataset (Table S1). Repeat measurements resulted in 2,402 and 2,377 proteins identified and the three datasets shared 86% of protein identifications (Fig. S1B). For comparison, our recent proteomic analysis of 20 fractions of mouse liver tissues using 1,000-fold more material resulted in a similar number of identifications (13).

Islets of Langerhans are another type of miniorgan essential for human health. In mice, 2,000 islets are dispersed throughout the pancreas, of which the insulin producing β cell is the most predominant (14). To obtain an in-depth islet proteome we pooled purified islets and analyzed them after peptide fractionation (Materials and Methods). Triplicate analysis employing RePlay and material from six mice resulted in the identification of 7,014 proteins with at least two peptides and a False Discovery Rate (FDR) of 1% (Table S2). The largest previous islet proteome is less than half the size and almost completely contained in our proteome (94%) (15). This list consists of (i) islet specific proteins, (ii) proteins shared between islets and the rest of the pancreas, and (iii) proteins from the rest of the pancreas that contaminate the islet preparation. To distinguish these classes we measured an in-depth proteome of complete pancreas homogenate, of which the islets constitute less than 2%, and quantified proteins against the islet proteome by label-free quantification in MaxQuant. Elastase, trypsin, and other digestive enzymes produced by the exocrine pancreas were more than 30-fold more abundant in the total pancreas proteome compared to purified islets. After filtering out proteins with a similar behavior, we retain a high quality islet proteome of 6,873 proteins. Our data indicate that islets, which are highly specialized tissue structures, use at least a third of the genes in the genome. Moreover, several proteins were identified as the pancreas specific isoform of the gene, and our data provides isoform specific tissue expression information for many more.

The known endocrine hormones produced by this miniorgan, such as insulin 1 and 2, glucagon, and secretogranins are among the most highly expressed proteins in the proteome judged by added peptide intensity (3) (Table S2). To contrast islet specific expression to expression in whole pancreas we categorized islet proteins by their enrichment factors (Fig. 24). More than half of the islet proteome is within a 4-fold abundance range compared...
**Fig. 2.** The proteome of pooled islets of Langerhans. (A) Classification of purified islet proteome versus pancreatic proteome, based on protein ratios between the two samples. After subtracting pancreas contaminants (141), a high confidence list of 6,873 islet proteins was obtained. Note that for proteins with less than three peptides and close to the limit of detection, accurate ratios and hence classification were not assigned (“not classified”). (B) Proteins enriched more than 4-fold (total 1,133) were categorized by GO cellular compartmentalization and (C) GO biological process. Numbers indicate proteins per category in the enriched dataset.

to whole pancreas, and these proteins are predominantly associated with house-hold functions. The pancreatic islet is an intricately assembled structure connecting the endocrine cells to the circulation and neural control. Concordantly, proteins connected to cell communication and cell adhesion make up much of the specific islet proteome. In the 1,133 proteins enriched more than 4-fold in islets, membrane and extracellular proteins comprise a large part (Fig. 2B and Table S3). This class of proteins contains various G protein coupled receptors, cAMP regulators, receptor tyrosine kinases and ion channels—potential and established targets for diabetes drugs, as well as autoantigens in type 1 diabetes (Fig. 2B and C and Fig. S1C). Interestingly, 86 of 396 identified kinases are highly enriched in the islets. Our dataset also contains many unannotated proteins, especially among those that were close to the limit of detection and therefore not quantified and categorized (Fig. 2A).

High fat and high glucose conditions are reported to negatively affect β cell function, a process that is central to the development of type 2 diabetes (16). Dynamics of β-cells are frequently studied in cell culture; however, it is physiologically more relevant to investigate these cells in their native cellular context (17). We wondered if our high sensitivity workflow was capable of analyzing single pancreatic islets. To this end, we determined single, isolated islet proteomes after 24 h at 16.7 and 5.6 mM glucose, representing hyperglycemic and basal conditions, respectively. One day of elevated glucose concentration significantly affects metabolic functions at the mRNA level (18), but does not yet evoke irreversible changes (19).

First, we assessed the proteome coverage in single islet experiments. Similar to the case of glomeruli we achieved very good correlation on the peptide level between direct and replicate runs (Fig. 3A). Protein quantitation integrates several peptide measurements and therefore the quantitative reproducibility is even higher on the protein level ($R = 0.98$) (Fig. 3B). We sequenced 15,193 different peptides originating from 2,013 unique proteins (2.5% FDR). Per biological replicate we identified on average 84% of the total number of proteins. Ninety-seven percent of the single islet proteins (1,949) were also present in the large islet proteome determined above (Fig. S1D).

The single islet proteome contained all of the islet specific hormones as well as known β cell proteins like glucose transporter 2, glucokinase, and prohormone convertase 1 and 2. Furthermore, we covered the majority of proteins involved in classical β cell functions (Fig. 3C and Fig. S1 E–G); for example, between 40 and 85% of the annotated proteins (irrespective of tissue) in glycolysis, TCA cycle, and oxidative phosphorylation, ribosomes, and secretory granules (Fig. 4).

Next we determined changes of the proteome after 24-h glucose stimulation. Quantitative comparison was based on proteins identified with greater than or equal to two unique peptides and detected in at least four out of eight runs (1,487 quantifiable proteins). With these stringent criteria and ANOVA $P < 0.01$, there were 77 up- and 65 down-regulated proteins. Some of the most significantly affected proteins are highlighted in the volcano plot in Fig. 3D. Many regulated proteins have previously been associated with hyperglycemia in animal models for diabetes or in patients (Table S4). Concordant with a recently published microarray study (18) we find a general up-regulation of glycolysis, the TCA cycle, and ATP translocation (Fig. 4). In contrast to elevated insulin transcription, we observed reduced insulin protein levels, presumably due to depletion of insulin from internal stores as a consequence of continuous secretion.

This assumption is indirectly supported by the finding of decreased levels of two hormone-processing inhibitors, prodynorphin (Pdyn) and ProSAAS (Pcsk1n) (20).

Increased glucose catabolism induces oxidative stress, a process detrimental to β-cells. We find molecular evidence for this in the form of antioxidant proteins, supporting and expanding previous observations (e.g., ERO1-α and Prdx3) (Figs. 3 and 4) (21). Prdx3, up-regulated 2.4-fold, protects β-cells from mitochondrial hydrogen peroxide stress and mice overexpressing Prdx3 are more glucose tolerant upon high-fat diet feeding and better protected against hyperglycemia (22). Interestingly, DJ-1 (Park7) and Mn$^{2+}$-superoxide dismutase (Sod2) were up-regulated 2- and 1.7-fold ($P < 0.05$), respectively. DJ-1 is a mitochondrial oxidantive stress and apoptotic regulator implicated in Parkinson disease (23, 24) but has not yet been associated with β cell responses in hyperglycemia. It induces mitochondrial antioxidant proteins like Sod2 (25). To explore DJ-1’s connection to β cell dysfunction, we quantified its expression in mouse islets after high-fat diet and found the protein up-regulated 1.7-fold (Fig. S1F) similar to the short term hyperglycemic stimulation, suggesting that it is involved in β cell regulation under multiple stress conditions.

In contrast to the increased protein levels in glucose catabolism and induced stress mechanisms described above, proteins associated with vesicle secretion are reduced (Fig. 4). In the case of vesicle-associated membrane protein (VAMP)-2, synaptotagmin-like protein 4 (Syt4), and Rab3b, who interact upon Ca$^{2+}$ increase, this has been observed before (26, 27). Decreased
expression of Sytl4 reduces the number of docked vesicles to the plasma membrane, which is suggested to effect insulin secretion (19, 28). Decreased mRNA and protein expression of VAMP2 and various other secretory proteins have also been associated with impaired protein secretion in diabetic patients (29) and animal models for diabetes (30). Furthermore, we observed depletion of all three members of the sarcoplasmic/endoplasmic reticulum calcium ATPase (SERCA) complex, which has also been correlated with insulin secretion deficiencies in diabetic animal models (31). In contrast, mRNA levels of SERCA increase after 18 h of glucose stimulation (18), possibly pointing to a compensatory mechanism for acute protein depletion.

 Vesicle access to the plasma membrane is thought to be hindered by dense cortical F-actin and depolymerization of F-actin upon glucose stimulation potentiates insulin secretion (36). Thus, our single islet analysis between its expression level and the rate of insulin secretion demonstrated a beneficial aspect of this directed membrane (35). Consistently, there is a direct relation to a compensatory mechanism for acute protein depletion.

In conclusion, by optimizing the proteomic workflow, using reduced flow rates, direct replicate measurements and extremely high glucose conditions. Consistently, there is a direct relation between its expression level and the rate of insulin secretion (36). Thus, our single islet analysis directly supports an important role of cytoskeletal remodeling in high glucose conditions.

In conclusion, by optimizing the proteomic workflow, using reduced flow rates, direct replicate measurements and extremely high glucose conditions. Consistently, there is a direct relation between its expression level and the rate of insulin secretion (36). Thus, our single islet analysis directly supports a compensatory mechanism for acute protein depletion.
sodium bicarbonate, and 0.001% beta-mercaptoethanol. Subsequently, the cells were incubated for 24 h with either 16.7 mM (high) or 5.6 mM (low) glucose in Krebs-Ringer-Hepes buffer (25 mM Hepes, pH 7.4, 115 mM NaCl, 24 mM NaHCO3, 5 mM KCl, 2.5 mM CaCl2, 1 mM MgCl2, and 0.1% BSA). Single cell measurements we used a split ratio of 1:1 (MS:capture capillary), an islet measurements we used a split ratio of 1:50. After acidification by trifluoroacetic acid, peptide mixtures were separated into 12 fractions by peptide isoelectric focusing as described (40).

Sample Digestion. LCM and single islet samples were dissolved in 0.1% PPS Silent Surfactant (Protein Discovery) in 50 mM ammonium bicarbonate (ABC) with 1 mM DTT, heated to 95 °C for 3 min, and sonicated for 2 min. Subsequently proteins were alkylated in 1 mM chloroacetamide in ABC for 30 min and digested overnight with porcine trypsin (Promega) in an enzyme:protein ratio of 1:50. After acidification by trifluoroacetic acid, peptide mixtures were concentrated and desalted on StageTips (39) and eluted directly before LC-MS measurement. To prevent protein losses during sample preparation LoBind tubes (Eppendorf) and Maxymum Recovery tips (Axygen) were used and sample contact was minimized.

For pooled islet and pancreas proteome the peptide samples were additionally separated into 12 fractions by peptide isoelectric focusing as described (40).

High Sensitivity LC-MS Interface. The RePlay system was based on the one described in (11), but developed for high sensitivity purposes: the capture capillary was 19 m long × 30 μm inner diameter (ID); the focusing column, a pulled capillary of 22 cm × 40 μm ID was packed with 3 μm ReproSil-Pur C18-AQ resin (Dr. Maisch GmbH), as described in (41). For the LCM and single islet measurements we used a split ratio of 1:1 (MS:capture capillary), an effective flow rate of 50 nL to the MS in both direct and replicate (or replay) run, and 240 min gradients. For the pooled islet and pancreas proteome gradients of 2 h were used with 150 nL/min effective flow rates. The flow sensor was placed in the waste line to prevent potential dilution effects.

The gradient generated by an EasyLC (Proxeon) was coupled, via the RePlay system and an electrospray interface, to an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific). MS and MS/MS were acquired data-dependently of the ten most intense ions per survey spectrum, using MS/MS target values of 5,000, with a minimal intensity of 1,000 ions and a maximum fill time of 150 ms. For survey spectra one million ions were accumulated with maximum fill times of 500 ms.

Data Analysis. MS/MS data were analyzed by in-house developed MaxQuant software (v 1.0.12.27), essentially as described (12). The data were searched using Mascot (v 2.1.04, Matrix Science Ltd) against the mouse IPI database (v 3.37) supplemented with frequently observed contaminants and concatenated with reversed copies of all sequences (2 × 51,467 entries). Enzyme specificity was set to trypsin, allowing for cleavage N-terminal to proline and between aspartic acid and proline and a maximum of two missed cleavages. Carbamidomethylation was set as fixed and N- acetylation and methionine oxidation were set as variable modifications. The initial maximum allowed mass deviation was set to 7 ppm for monoisotopic precursor ions and 0.5 Da for MS/MS peaks. The required minimum peptide length was six amino acids. If the identified peptide sequence set of one protein was equal to or contained another protein’s peptide set, these two proteins were grouped together and the proteins were not counted as independent hits. Proteins were considered identified when at least two peptides were identified of which at least one was uniquely assignable to the respective sequence. The false discovery rate (FDR) at the peptide level and protein level were set to 1% for the glomeruli, pooled islet, and pancreas proteome. Based on the additional identification confidence provided by the pooled islet dataset, we accepted single unique quantitation proteins were additionally filtered, and were required to be detected in four out of eight most intense LC-MS/MS runs with at least two unique peptides.

Label-free quantitation was performed in MaxQuant. Briefly, total peptide signals were determined in the mass to charge, elution time and intensity space. For every peptide, corresponding total signals from multiple runs were compared to determine peptide ratios. Pair-wise peptide ratios were only determined when the corresponding peak is detected in both LC-MS runs. Median values of all peptide ratios of one protein then represent a robust estimate of the protein ratio. MaxQuant determines these ratios after first
ACKNOWLEDGMENTS. We thank Tami Geier, Ana Velic, and Gabriele Friedemann for assisting with glomerular collection and data analysis; Jürgen Cox for help with the data analysis; Irena Constantinova for providing the single islet picture in Fig. 1; members of the Proteomics and Signal Transduction group for help and discussions; and Tobias Waithner, Maximilian Hilger, Nina Hubner, and Tami Geier for critical reading of the manuscript. This project was supported by European Commission’s 7th Framework Program PRoteomics SPECification in Time and Space (PROSPECTS, HEALTH-F4-2008-021648), and German Research Council Grant DFG LA1216 (to K.C. and E.L.).

38. Schwartz JC, Syka JEP, Remes PM, Quarnby ST (2008) in Proc 56th ASMS Conf Mass Spectrometry and Allied Topics, WPAA039. For additional information, contact the corresponding author of this article.