Published in final edited form as:
J Proteome Res. 2006 December ; 5(12): 3345–3354. doi:10.1021/pr060322n.

Characterization of the Human Pancreatic Islet Proteome by Two-Dimensional LC/MS/MS

Thomas O. Metz[1], **Jon M. Jacobs**[1], **Marina A. Gritsenko**[1], **Ghislaine Fontès**[2,3], **Wei-Jun Qian**[1], **David G. Camp II**[1], **Vincent Poitout**[2,3], and **Richard D. Smith**[1,*]

[1] Biological Science Division and Environmental Molecular Sciences Laboratory Pacific Northwest National Laboratory Richland, Washington, USA

[2] Pacific Northwest Research Institute, Seattle, Washington, USA

Abstract

The pancreatic beta-cell plays a central role in the maintenance of glucose homeostasis and in the pathogenesis of both type 1 and type 2 diabetes mellitus. Elucidation of the insulin secretory defects observed in diabetes first requires a better understanding of the complex mechanisms regulating insulin secretion, which are only partly understood. While there have been reports detailing proteomic analyses of islet cell lines or isolated rodent islets, the information gained is not always applicable to humans. Therefore, definition of the human islet proteome could contribute to a better understanding of islet biology and lead to more effective treatment strategies. We have applied a two-dimensional LC-MS/MS-based analysis to the characterization of the human islet proteome, resulting in the confident identification of 29,021 different tryptic peptides covering 3,365 proteins (≥ 2 unique peptide identifications per protein). As expected, the three major islet hormones (insulin, glucagon, and somatostatin) were detected, as well as various beta-cell enriched secretory products, ion channels, and transcription factors. In addition, significant proteome coverage of metabolic enzymes and cellular pathways was observed, including the integrin signaling cascade and the MAP kinase, NF-κβ, and JAK/STAT pathways. The resulting peptide reference library provides a resource for future higher throughput and quantitative studies of islet biology.

Keywords

pancreatic islets; liquid chromatography; electrospray; mass spectrometry; proteomics; linear ion trap

Introduction

A better functional characterization of the human pancreatic beta-cell is required to improve the treatment of both type 1 and type 2 diabetes mellitus. In type 1 diabetes, islet transplantation remains a relatively unsuccessful procedure[1], and there is a need to define better predictive criteria of islet function *in vivo*. In type 2 diabetes, elucidation of the insulin secretory defects that play a pathogenic role in this disease first requires a better understanding of the complex regulatory mechanisms of insulin secretion itself.

Proteomics, or the analysis of the entire protein complement expressed by the genome of an organism, represents a potentially powerful tool for predicting islet transplantation success or

* Corresponding author: P.O. Box 999 Richland, WA 99352 Phone: (509) 376-0723 Fax: (509) 376-2303 rds@pnl.gov .
[3]Current Address: Department of Medicine University of Montréal Montréal, Québec, Canada

identifying beta-cell functional defects. While several proteomic studies describing the effects of interleukin-1β[2,3], rosiglitazone[4], or increased glucose concentration[5,6] on isolated rodent islets have been reported, the information gained can not always be extrapolated to human islets. Currently, there is a lack of original research articles involving proteomic characterization or study of human pancreatic islets. Hu *et al*.[7] have applied two-dimensional gel electrophoresis (2-DE) followed by matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI TOF-MS) in the proteomic profiling of whole human pancreas, resulting in the identification of 302 proteins. However, because the proteins identified in that investigation are a mix of both endocrine and exocrine tissue, it is difficult to use these data for the study of isolated pancreatic islets. More recently, Ahmed *et al*.[8] have also applied 2-DE and MALDI TOF-MS in the two-dimensional protein mapping of isolated human pancreatic islets. They report the detection of 744 protein spots, with 130 spots corresponding to 66 different protein entries covering enzymes, chaperones, cellular structural proteins, cellular defense proteins, signaling molecules, and transport proteins. These investigations represent a good first step towards the characterization of the human pancreatic islet proteome but are lacking the depth of proteome coverage needed to accurately characterize complex and high dynamic range mammalian tissue samples. In this study, we report the use of sensitive two-dimensional liquid chromatography (LC) separations coupled with ion-trap tandem MS analysis for the global proteomic characterization of isolated human pancreatic islets, resulting in the identification of 29,021 unique peptides corresponding to 3,365 proteins (≥ 2 unique peptide identifications per protein). Further, an additional 1,560 proteins were characterized by 1 unique peptide identification, but, in general, lower confidence is placed in proteins identified by only 1 unique peptide when employing bottom-up proteomics. By utilizing normalized peptide elution time and calculated masses, a mass and time (MT) tag database has been created for the identified peptides, which facilitates the use of accurate mass MS for quantitative identification of the peptides in subsequent high-throughput LC-MS analyses. This work represents the most extensive characterization of the human islet proteome to date and provides a peptide reference library that may be utilized in future comparative studies of human islet biology.

Materials and Methods

Reagents

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO) unless stated otherwise. Ammonium bicarbonate and methanol were purchased from Fisher Scientific (Fair Lawn, NJ) and sequencing grade trypsin was purchased from Promega (Madison, WI). Ammonium formate was obtained from Fluka (St. Louis, MO), while bicinchoninic acid (BCA) Protein Kit was purchased from Pierce (Rockford, IL). Purified, deionized water, >18 MΩ, (Nanopure Infinity ultrapure water system, Barnstead, Newton, WA) was used to make all aqueous biological and HPLC buffers.

Human pancreatic islets

Approval for the conduct of this programmatic research was obtained from the Institutional Review Boards of the participating institutions. Human pancreatic islets were isolated at the Islet Cell Resource at Puget Sound Blood Center (Seattle, WA) using the two-layer method, as described[9,10]. Isolated islets were then cultured for 72 h in CMRL 1066 culture medium supplemented with 10% Fetal Bovine Serum, also as described[9]. Aliquots of 800 islets each were hand-picked from individual donor preparations, washed, and centrifuged at 1200 rpm at 4°C for 5 min. After removal of the supernatant, the islet pellets were frozen at −80°C. Islets from 5 different donors were pooled to create a uniform sample of ~4,000 total islet equivalents for protein extraction and analysis.

Protein extraction and enzymatic digestion

Proteins from 4,000 human pancreatic islet equivalents were extracted and digested using two different protocols to provide complementary protein coverage of the same sample: a urea/thiourea mixture containing 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and an organic mixture based on 2, 2, 2-trifluoroethanol (TFE).

For the urea/thiourea-based protocol, 2,000 islet equivalents were lysed and the proteins denatured and reduced by incubation with 7 M urea, 2 M thiourea, 1% CHAPS, and 5 mM tributylphosphine (TBP) in 100 mM ammonium bicarbonate, pH 8.4, at 60°C for 2 h in a Thermomixer R (Eppendorf, Hamburg, Germany), with constant shaking at 800 rpm. The samples were then diluted 10-fold with 100 mM ammonium bicarbonate containing 1 mM calcium chloride prior to enzymatic digestion with trypsin (1:50 w/w, trypsin:protein) for 3 h at 37°C. The digests were cleaned using strong-cation exchange (SCX) SPE tubes (Discovery DSC-SCX, Supelco, Bellefonte, PA), and bound peptides were eluted with 80% methanol containing 15% sodium hydroxide. Peptide samples were then concentrated to 50 μL by Speed-Vac (Thermo Savant, Holbrook, NY) and final concentrations determined by BCA protein assay. The samples were then stored at −80°C until time for LC-SCX fractionation.

For the organic-based protocol, 2,000 islet equivalents were initially sonicated (Branson 1510, Danbury, CT) in 50% TFE for 6 min at 30 sec intervals with chilling on ice between sonications in order to lyse cells. The samples were subsequently incubated at 60°C for 2 h with constant shaking to denature proteins, followed by reduction with 5 mM TBP for 45 min at 60°C. The samples were then diluted 10-fold with 100 mM ammonium bicarbonate containing 1 mM calcium chloride prior to enzymatic digestion with trypsin (1:50 w/w, trypsin:protein) for 3 h at 37°C. Peptide samples were then concentrated to 50 μL by Speed-Vac (Thermo Savant, Holbrook, NY) and final concentrations determined by BCA assay. The samples were then stored at −80°C, as above.

Strong-cation exchange fractionation of enzymatic digests

SCX fractionation of enzymatic digests was performed as previously described[11,12]. Briefly, human pancreatic islet peptides were diluted with 850 μL of 10 mM ammonium formate (pH 3.0) in water containing 25% acetonitrile and fractionated by SCX chromatography on a Polysulfoethyl A 200 mm × 2.1 mm column (PolyLC, Columbia, MD) that was preceded by a 10 mm × 2.1 mm guard column of the same material. The separations were performed at a flow rate of 0.2 mL/min using an Agilent 1100 series HPLC system (Agilent Technologies, Santa Clara, CA), with mobile phases consisting of 10 mM ammonium formate (pH 3.0) in water containing 25% acetonitrile (A), and 500 mM ammonium formate (pH 6.8) in water containing 25% acetonitrile (B). After loading 300 μL of sample onto the column, the gradient was maintained at 100% A for 10 min. Peptides were then separated using a gradient from 0 to 50% B over 40 min, followed by a gradient of 50-100% B over 10 min. The gradient was then held at 100% B for 10 min. A total of 30 fractions each were collected for islet digests whose proteins were extracted by either urea/CHAPs or TFE, and all fractions were dried under vacuum.

Reversed-phase capillary LC-MS/MS analyses

Dried peptide fractions were reconstituted in 30 μL 25 mM ammonium bicarbonate, pH 7.8, and analyzed using a custom-built capillary LC system coupled online to a linear ion trap mass spectrometer (LTQ; ThermoElectron, Waltham, MA) by way of an in-house manufactured electrospray ionization interface as previously described[11-13]. The reversed-phase capillary column was prepared by slurry packing 5-μm Jupiter C18 bonded particles (Phenomenex, Torrence, CA) into a 150 μm × 65 cm fused silica capillary (Polymicro Technologies, Phoenix, AZ) that utilized a 2-μm stainless steel retaining screen within a stainless steel union (Valco

Instruments Co., Houston, TX). The mobile phase consisted of 0.2% acetic acid and 0.05% TFA in water (A) and 0.1% TFA in 90% acetonitrile/10% water (B). Mobile phases were degassed on-line using a vacuum degasser (Jones Chromatography Inc., Lakewood, CO), and the HPLC system was equilibrated at 5,000 psi with 100% mobile phase A for initial starting conditions. After loading 10 μg of peptides onto the column, the mobile phase was held at 100% A for 20 min. Exponential gradient elution was performed by increasing the mobile-phase composition from 0 to 77% B over 100 min, using a stainless steel mixing chamber, followed by column washing at ~100% B for 10 minutes. To identify the eluting peptides, the LTQ was operated in a data-dependent MS/MS mode (*m/z* 400-2,000), in which a full MS scan was followed by ten MS/MS scans using a normalized collision energy of 35%. A dynamic exclusion window of 1 min was used to discriminate against previously analyzed ions. The temperature of the heated capillary and the ESI voltage were 200°C and 2.2 kV, respectively.

Additionally, two unfractionated peptide digests from both the urea and TFE protocols were analyzed in duplicate, as described above.

MS/MS data analysis and protein categorization

SEQUEST analysis software[14] was used to match the MS/MS fragmentation spectra with sequences from the April 2005 IPI human database (version 3.05), containing 49,161 entries. The criteria selected for filtering followed methods based upon a human reverse-database false-positive model which has been shown to give ≥95% confidence at the peptide level[15]. Using the reverse database approach, the false discovery rate (FDR) estimate for this dataset was determined to be 1.7% at the peptide level and 0.3% at the protein level when using the filter criteria described below.

Briefly, protein identifications were retained if their identified peptide sequence met the following criteria: 1) SEQUEST DelCn2 value (normalized Xcorr difference between top scoring peptide and second highest scoring peptide in each MS/MS spectrum) of ≥ 0.10 and 2) SEQUEST correlation score (Xcorr) ≥ 1.6 for charge state 1+ for fully tryptic peptides; Xcorr ≥ 2.4 for charge state 2+ and fully tryptic peptides and Xcorr ≥ 4.3 for charge state 2+ and partially tryptic peptides; Xcorr ≥ 3.2 for charge state 3+ and fully tryptic peptides and Xcorr ≥ 4.7 for charge state 3+ and partially tryptic peptides. In addition, only those proteins identified by ≥ 2 unique peptides were retained.

To remove redundantly identified proteins, the program ProteinProphet was utilized[16]. All peptides which passed these filter criteria were given the identical score of one, and entered into ProteinProphet for redundancy analysis only. This condensed the number of proteins detected from an initial 4,015 to a combined total of 3,365 proteins reported as identified. A list of confidently identified (≥ 2 unique peptides) proteins is provided as Supporting Information. In addition, a separate list of 1,560 proteins identified by only 1 unique protein is also provided as Supporting Information.

Ingenuity Pathways Analysis

Canonical pathway mapping was performed using the Ingenuity Pathways Analysis (IPA) application (www.ingenuity.com). A dataset containing Gene IDs for each identified protein as obtained from the NCBI Entrez Gene database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene) was uploaded as an Excel spreadsheet using the template provided in the IPA application, and each gene identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base.

Results

Proteomic coverage of human pancreatic islets

The initial goal of this investigation was to obtain the most comprehensive proteomic characterization of human islet tissue as possible to be used as a baseline for further studies. To accomplish this goal, we utilized a bottom-up proteomics approach, which first entails detecting and identifying peptide sequences via tandem mass spectrometry and subsequently linking those peptide sequences to their respective proteins during downstream data analysis. In this study, a total of 509,599 and 478,508 tandem mass spectra were collected for islets denatured/digested using the urea and TFE protocols, respectively. Figure 1 shows a schematic of the overall approach and results. Two main protein denaturation/digestion protocols were followed to obtain a more diverse sampling of the islet proteome and to determine the optimum protocol, if any, to be applied in future studies. A total of 29,021 peptides were detected and identified in the study, corresponding to 3,365 proteins identified after redundancy reduction as described under Materials and Methods. Of the total number of peptides reported, 98.8% were found to be fully tryptic peptides. This value was preserved when considering those peptides identified independently in each denaturation/digestion condition, *i.e*. 98.8% of the peptides identified from either the urea or TFE protocol were also fully tryptic. In addition, 1.9% of those proteins identified by only 2 unique peptides were identified by 1 partially tryptic peptide; no protein was identified by only 2 partially tryptic peptides. Importantly, 60% of the reported peptides lacked a missed cleavage event, *i.e*. either without internal Lys or Arg present or with an internal Lys or Arg present but immediately followed by a Pro, Glu, Asp, Lys, or Arg residue. Those peptides containing 1, 2, or ≥ 3 missed cleavage events comprised 32%, 7%, and 1% of all reported peptides, respectively. In general, 92% of the reported peptides contained one or no missed cleavage events. Some of the missed cleavages observed may be due to the relatively short digestion time (3 h). However, the identity and number of peptide and protein identifications did not change significantly in a separate evaluation of 3 h and 6 h digestion time in a time-course study of tryptic digestion, using *Shewanella oneidensis* as model sample (unpublished data). Indeed, the number of peptide and protein identifications declined dramatically for >6 h digestion time, possibly due to non-specific tryptic cleavage of proteins. Therefore, the missed cleavages identified in this study are likely due to individual protein characteristics such as degree of denaturation, accessibility of sequences to trypsin, etc.

To estimate the peptide and protein FDR in this study, the MS/MS data were processed using the SEQUEST algorithm in conjunction with a reverse protein sequence database, as described above. The FDR was calculated as the # of unique false peptide identifications divided by the total # of unique peptide identifications (from both forward and reverse protein sequence database searches). Using identical peptide filtering criteria as the forward protein sequence database search, the reverse protein sequence database search yielded a total of 519 unique peptides for a peptide FDR of 1.7%. These peptides mapped to 10 proteins (≥ 2 unique peptides per protein), corresponding to a protein level FDR of 0.3%.

A detailed list of proteins identified in this study, together with their molecular weight, sequence coverage, and number of unique peptides are provided as supplemental material. The percentage of overlapping peptides detected under both digestion conditions was relatively low (28%) but with a much higher overlap at the protein level (61%). This may be explained, in part, by the fact that different digestion conditions and environments (*i.e*. aqueous versus organic) alter the solubility and structure of individual proteins and, therefore, the availability of peptide residues for tryptic digestion. Thus, peptides identified as unique to either of the two denaturation/digestion protocols are more likely to converge into the same protein, yielding a higher protein overlap compared to the peptide overlap. This demonstrates that there may be an advantage in pursuing different sample preparation protocols to increase proteome coverage at the peptide level, but it becomes somewhat dilutive at the protein level. Alternatively, the

low overlap at the peptide level could also be a result of undersampling of low abundance peptides in samples from either digestion protocol; indeed, ~50% of filter passing peptides were identified from a single tandem mass spectrum.

Population of the mass and time tag database

We have developed the accurate mass and time (AMT) tag approach[17-20] in order to increase the throughput of LC-MS-based proteomics experiments. This approach involves first storing peptides identified from relatively low-throughput LC-MS/MS experiments as mass and time (MT) tags in a SQL Server database. Features from subsequent high-throughput LC-FTICR experiments of similar samples are then matched against the MT tag database to identify peptides, exploiting the high resolution and mass measurement accuracy of the FTICR instrument. Those peptides in LC-FTICR experiments that match entries in the MT tag database are then considered AMT tags.

In this study, a human pancreatic islet peptide MT tag database was populated with those peptides identified by the SEQUEST algorithm in LC-MS/MS datasets. The peptide sequences were used to calculate the monoisotopic masses of identified peptides, which were subsequently added to the MT tag database. Similarly, elution time information for each peptide was utilized as an additional identifier in the MT tag database. Because a peptide is often observed in several different analyses with similar, but slightly different elution profiles, one must normalize the observed elution time data for the separate analyses to generate normalized elution time (NET) values, allowing for effective comparison of multiple analyses. Each LC-MS/MS analysis is normalized independently, regressing the observed peptide elution times in a given analysis against the predicted NET for each identified peptide[21]. The predicted NET for a given peptide is determined by employing a neural network-based model, developed using training data from 20 species and over 140,000 unique filtered peptide identifications[21]. Predicted NET values generally range from 0 to 1, and are based on the composition and order of the first 8 and last 8 residues in a peptide (all residues when the peptide is 16 residues long or less). The model considers all 20 amino acid residues, in addition to alkylated cysteine. Once the observed NET values for all peptides identified in an LC-MS/MS or LC-MS analysis have been determined, one can compare the observed NET values with the predicted NET values to exclude peptide identifications that exhibit drastically different observed and predicted values. The observed NETs for the remaining peptides are then incorporated into the MT tag database.

Islet proteomic MT tag data will be made publicly available on the Pacific Northwest National Laboratory Proteomic National Center for Research Resources (NCRR) website (http://ncrr.pnl.gov).

Pathway analysis of the human islet proteome

The sequence coverage of proteins involved in the glycolytic/gluconeogenic and citrate cycle pathways is shown in Tables 1 and 2, respectively. All major enzymes were identified along with minor (in terms of sequence coverage) isoforms. A median value of 63% sequence coverage (as calculated using the Protein Coverage Summarizer available at http://ncrr.pnl.gov/software/) is reported for the glycolytic/gluconeogenic enzymes (major isoforms only, n = 13), although two of these were found at relatively low sequence coverage: pyruvate carboxylase and splice isoform 1 of hexokinase at 19% and 12%, respectively. Similarly, a median value of 43% sequence coverage is reported for the citrate cycle enzymes (major isoforms only but including various subunits as appropriate, n = 15). Only two enzyme components exhibited relatively low sequence coverage: 2-oxoglutarate dehydrogenase E1 component and isocitrate dehydrogenase (NAD) subunit gamma at 21% and 13%, respectively. However, the majority of these enzymes are over 100 kDa such that the coverage reported still

provides a fairly significant number of peptides (7 to13) available for future comparative quantitative studies.

Figure 2 is a canonical pathway view of the JAK/Stat signaling pathway, as generated by the IPA application. Our method provides approximately 54% coverage of this pathway, based on unique protein identification (≥ 2 unique peptides per protein). This indicates a high probability for assessing the functional aspect of human islet isolates by a global proteomics-based approach, using coverage of the JAK/Stat signaling pathway as a metric. Figures 3 and 4 show canonical pathway views of the integrin signaling and NF-κB signaling pathways, respectively, as generated by the IPA application. The approach taken in this investigation provides roughly 76% and 30% coverage of the integrin and NF-κB pathways, respectively, based on unique protein identifications. Several isoforms of the same gene product are typically mapped by the IPA application; however, only the general protein is plotted in the canonical pathways view. To determine which isoforms are actually present, the user must place the mouser cursor overtop the shaded protein. For example, placing the cursor overtop the Ras gene product in the NF-κB signaling pathway reveals that of the three possible isoforms (HRas, KRas, or NRas), only HRas was identified in our analysis.

In general, of the 3,365 proteins identified by ≥ 2 unique peptides in this study, 3,232 (96%) were found to have Gene IDs as obtained from the NCBI Entrez Gene database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene). Of these, 2,269 (70%) were eligible for generating networks; the remainder were structural proteins or those which could not be assigned to a pathway.

It is important to note that the pathway figures obtained from the IPA application represent the theoretical representation of the given pathway as supported by data in the current scientific literature. The mapping of proteins detected in this study to those proteins in the shown pathways only indicates that, in general, the approach applied here will likely be successful in assessing islet viability and functionality through detection of specific proteins involved in common biochemical pathways.

Discussion

A proteomics-based model for determining islet viability and functionality will depend highly on the extent of islet proteome coverage. In this work, we have constructed a peptide MT tag database containing ~30,000 peptides which pass stringent SEQUEST Xcorr and DelCn2 filtering criteria described under Research Design and Methods; the criteria selected for filtering follows methods based upon a human reverse-database false-positive model which has been shown to give ≥95% confidence for the entire protein dataset[15]. To our knowledge, this work represents the most extensive proteomic coverage of isolated human islets to date. It is difficult to compare the extent of islet proteome coverage in our study with that of Hu *et al*.[7], who utilized 2-DE and MALDI TOF-MS in the proteomic characterization of whole human pancreas. However, of the 302 proteins identified in that study, we have identified 267 (88%); 17 of the remaining 35 proteins appear to be plasma contaminates. This suggests that there may be inherent differences in the proteome of whole pancreas versus isolated islets, or that the whole pancreata in the former study may not have been completely washed free of high abundance plasma proteins. Of the top 20 high abundance plasma proteins[22], we have identified 8 by two or more unique peptides: albumin (48 peptides), complement C3 (15 peptides), haptoglobin (5 peptides), alpha-1-antitrypsin (26 peptides), alpha-1-antichymotrypsin (13 peptides), vitamin d-binding protein (9 peptides), ceruloplasmin (3 peptides), and complement C4 (10 peptides). The relatively low number of unique peptides identified for these proteins (insulin was identified by 21 unique peptides) indicates that contamination due to plasma was minimized during the preparation and isolation of the

pancreatic islets used in this study. Indeed, peptides corresponding to fibrinogen, transferrin, alpha-2-macroglobulin, alpha-1-acid glycoprotein, alpha-2-hs-glycoprotein, and apolipoproteins (A-I, A-II, and B-100) were not detected. Interestingly, alpha-1-antitrypsin has been shown[23] to be produced by an islet cell type exclusive of the hormone-secreting cells; therefore both plasma contamination and the islets themselves may be contributing to the total concentration of this protein in the preparation. A more direct comparison of the current work may be made to that of Ahmed *et al*.[8], who also utilized 2-DE and MALDI TOF-MS in the proteomic characterization of isolated human islets. Of the 66 proteins identified in that study, we have identified 61. In addition, we have identified these proteins with an approximate 100% increase (median) in sequence coverage. This illustrates the ability of two-dimensional LC-MS/MS to provide high quality front-end separations of complex mixtures, which ultimately reduces the number of ions delivered to the MS detector and increases the dynamic range of the measurement. Although by nature human islet preparations are always contaminated with some exocrine tissue, the islets used for proteomic analysis were hand-picked prior to extraction, giving confidence that most of the proteins identified are indeed from islet cells. Clearly, the stress associated with the isolation procedure is likely to induce the expression of a number of proteins. Therefore, the proteomic profiles obtained in this study do not represent that of the native islet *in situ* but of the isolated islet *in vitro*. Despite this limitation, we believe that this approach will prove useful in identifying predictive criteria of islet function after transplantation.

Several proteins known to be specifically expressed in beta- (insulin, GAD65) or non-beta- (glucagon, somatostatin) islet cells were identified. The ultimate goal of this project is to correlate changes in the level of expression of proteins with functional performance of the islets, as assessed both *in vitro* and *in vivo*. Therefore, our approach is essentially discovery-based and will not be biased towards known proteins, because the power of the analysis will likely enable us to discover new markers of islet function. However, as a preliminary interpretation of our results, we have selected several examples of metabolic and signaling pathways known to be expressed and active in beta-cells, to illustrate the fact that several major regulators of beta-cell function and survival were identified using this approach. This analysis is not comprehensive and is limited to selected enzymes and signalling proteins as recognized by the Ingenuity Systems software. It is important to note, however, that our proteome coverage also includes a number of proteins from other functional classes, such as structural proteins, transcription factors, and proteins involved in energy homeostasis.

Anaplerotic enzymes

The functional importance of anaplerotic enzymes in beta-cell physiology is well known, and our analysis identified the expression of pyruvate carboxylase (PC), phosphoenolpyruvate carboxykinase, and malic enzyme in human islets. Flux through PC has been shown to play a major role in glucose-induced insulin secretion[24], and one might anticipate that the level of PC expression could represent a predictive criterion of islet function.

Fatty-acid metabolism

Intracellular metabolism of fatty acids is thought to generate signals that contribute to the control of insulin secretion[25]. Intracellular levels of malonyl-CoA increase in beta-cells in response to glucose, resulting in inhibition of fatty-acid oxidation and generation of intracytosolic lipid-derived signaling molecules[26]. Interestingly, our analysis confirmed that a large number of enzymes involved in fatty acid metabolism are expressed in human islets, consistent with the notion that signals derived from lipid metabolism serve as metabolic coupling factors for insulin secretion[26].

The insulin signaling pathway

Components of the insulin signaling cascade, such as Akt/protein kinase B are thought to be major regulators of beta-cell survival[27]. Although the main mechanism of activation of the pathway is debated (e.g. activation of the insulin receptor, of the insulin-like growth-factor-1 (IGF-1) receptor, or direct modulation of IRS2 by intracellular signals), it is likely that a number of signals converge towards Akt activation and promotion of beta-cell survival[28].

The MAP kinase pathway

Mitogen-activated and stress-activated kinases have been implicated in many aspects of pancreatic beta-cell function and survival. For instance, they mediate some of the pro-apoptotic effects of chronically elevated glucose and cytokines on beta-cells[29], and the Erk pathway has been recently shown to be involved in the regulation of insulin gene transcription by glucose[30].

The JAK/STAT pathway

The Janus kinase/signal transducers and activators of transcription pathway mediate, at least in part, cytokine-induced cell death and are particularly important to consider in the context of islet transplantation[31].

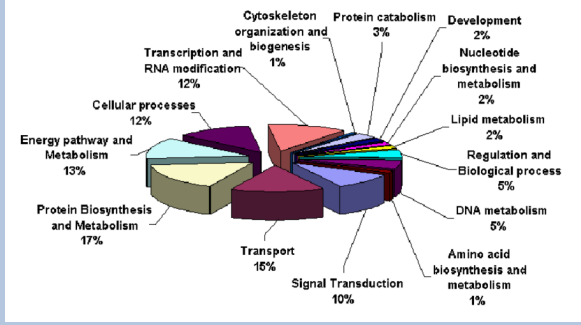
Integrins

Little is known on the function of integrins and integrin receptors in beta-cell biology; however, recent evidence suggests that these proteins may be important not only for pancreas development but also to maintain the sophisticated architecture of the mature islet of Langerhans[32]. A recent publication highlighted the role of cell-extracellular matrix (ECM) contacts in the maintenance of glucose-regulated insulin secretion[33]. Since the isolation process is likely to disrupt the bonds between islet cells and the ECM, it is tempting to speculate that the integrity of the integrin receptor pathway might represent a valuable marker of the preservation of the islet structure and function prior to transplantation.

In conclusion, this analysis represents, to our knowledge, the most extensive proteome coverage of isolated human islets to date. The initial identification of important metabolic enzymes and proteins involved in insulin signaling can provide the basis for future quantitative analyses, the results of which can then be correlated to islet function both *in vitro* and *in vivo*. Meanwhile, the mass and time tag database represents a useful reference tool for studies in human islet biology.

Synopsis

Two-dimensional LC-MS/MS-based analysis of human islets has resulted in the confident identification of 29,021 different tryptic peptides covering 3,365 proteins. As expected, the three major islet hormones (insulin, glucagon, and somatostatin) were detected, as well as various beta-cell enriched secretory products, ion channels, and transcription factors. Significant proteome coverage of metabolic enzymes and cellular pathways was observed, including the integrin signaling cascade and the MAP kinase, NF-κβ, and JAK/STAT pathways.



Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH/NIDDK grant R21 DK070146. We thank Dr. JoAnna Reems (Puget Sound Blood Center, Seattle, WA) and Dr. Paul Robertson (Pacific Northwest Research Institute, Seattle, WA) for providing isolated human islets. We also thank Dr. Matthew Monroe for software and informatics support. Portions of this research were supported by the NIH National Center for Research Resources (RR18522). Work was performed in the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy (DOE) national scientific user facility located on the campus of Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is a multiprogram national laboratory operated by Battelle Memorial Institute for the DOE under contract DE-AC05-76RLO 1830.

References

1. Close NC, Hering BJ, Eggerman TL. Results from the inaugural year of the Collaborative Islet Transplant Registry. Transplant Proc 2005;37(2):1305–1308. [PubMed: 15848704]
2. Larsen PM, Fey SJ, Larsen MR, Nawrocki A, Andersen HU, Kahler H, Heilmann C, Voss MC, Roepstorff P, Pociot F, Karlsen AE, Nerup J. Proteome analysis of interleukin-1beta-induced changes in protein expression in rat islets of Langerhans. Diabetes 2001;50(5):1056–1063. [PubMed: 11334408]
3. Sparre T, Christensen UB, Larsen P. Mose, Fey SJ, Wrzesinski K, Roepstorff P, Mandrup-Poulsen T, Pociot F, Karlsen AE, Nerup J. IL-1beta induced protein changes in diabetes prone BB rat islets of Langerhans identified by proteome analysis. Diabetologia 2002;45(11):1550–1561. [PubMed: 12436339]
4. Sanchez JC, Converset V, Nolan A, Schmid G, Wang S, Heller M, Sennitt MV, Hochstrasser DF, Cawthorne MA. Effect of rosiglitazone on the differential expression of diabetes-associated proteins in pancreatic islets of C57Bl/6 lep/lep mice. Mol. Cell. Proteomics 2002;1(7):509–516. [PubMed: 12239279]
5. Collins HW, Buettger C, Matschinsky F. High-resolution two-dimensional polyacrylamide gel electrophoresis reveals a glucose-response protein of 65 kDa in pancreatic islet cells. Proc. Natl. Acad. Sci. USA 1990;87(14):5494–5498. [PubMed: 2196569]
6. Collins H, Najafi H, Buettger C, Rombeau J, Settle RG, Matschinsky FM. Identification of glucose response proteins in biological models of beta-cell adaptation to chronic high glucose exposure. J. Biol. Chem 1992;267(2):1357–1366. [PubMed: 1730656]
7. Hu L, Evers S, Lu Z, Shen Y, Chen J. Two-dimensional protein database of human pancreas. Electrophoresis 2004;25(3):512–518. [PubMed: 14760645]
8. Ahmed M, Forsberg J, Bergsten P. Protein profiling of human pancreatic islets by two-dimensional gel electrophoresis and mass spectrometry. J. Proteome Res 2005;4(3):931–40. [PubMed: 15952740]
9. Matsumoto S, Qualley SA, Goel S, Hagman DK, Sweet IR, Poitout V, Strong DM, Robertson RP, Reems JA. Effect of the two-layer (University of Wisconsin solution-perfluorochemical plus O_2)

method of pancreas preservation on human islet isolation, as assessed by the Edmonton Isolation Protocol. Transplantation 2002;74(10):1414–1419. [PubMed: 12451242]

10. Matsumoto S, Kandaswamy R, Sutherland DE, Hassoun AA, Hiraoka K, Sageshima J, Shibata S, Tanioka Y, Kuroda Y. Clinical application of the two-layer (University of Wisconsin solution/perfluorochemical plus O2) method of pancreas preservation before transplantation. Transplantation 2000;70(5):771–774. [PubMed: 11003355]
11. Liu T, Qian WJ, Gritsenko MA, Camp DG 2nd, Monroe ME, Moore RJ, Smith RD. Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. J. Proteome Res 2005;4(6):2070–2080. [PubMed: 16335952]
12. Wang H, Qian WJ, Chin MH, Petyuk VA, Barry RC, Liu T, Gritsenko MA, Mottaz HM, Moore RJ, Camp DG 2nd, Khan AH, Smith DJ, Smith RD. Characterization of the mouse brain proteome using global proteomic analysis complemented with cysteinyl-peptide enrichment. J. Proteome Res 2006;5(2):361–369. [PubMed: 16457602]
13. Shen Y, Zhao R, Belov ME, Conrads TP, Anderson GA, Tang K, Pasa-Tolic L, Veenstra TD, Lipton MS, Udseth HR, Smith RD. Packed capillary reversed-phase liquid chromatography with high-performance electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for proteomics. Anal. Chem 2001;73(8):1766–1775. [PubMed: 11338590]
14. Eng JK, McCormack AL, Yates JR 3rd. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. J. Am. Soc. Mass Spectrom 1994;5(11):976–989.
15. Qian WJ, Liu T, Monroe ME, Strittmatter EF, Jacobs JM, Kangas LJ, Petritis K, Camp DG 2nd, Smith RD. Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. J. Proteome. Res 2005;4(1):53–62. [PubMed: 15707357]
16. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem 2003;75(17):4646–4658. [PubMed: 14632076]
17. Qian WJ, Monroe ME, Liu T, Jacobs JM, Anderson GA, Shen Y, Moore RJ, Anderson DJ, Zhang R, Calvano SE, Lowry SF, Xiao W, Moldawer LL, Davis RW, Tompkins RG, Camp DG 2nd, Smith RD. Inflammation and the Host Response to Injury Large Scale Collaborative Research Program: Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using 16O/18O labeling and the accurate mass and time tag approach. Mol. Cell. Proteomics 2005;4(5):700–709. [PubMed: 15753121]
18. Adkins JN, Monroe ME, Auberry KJ, Shen Y, Jacobs JM, Camp DG 2nd, Vitzthum F, Rodland KD, Zangar RC, Smith RD, Pounds JG. A proteomic study of the HUPO Plasma Proteome Project's pilot samples using an accurate mass and time tag strategy. Proteomics 2005;5(13):3454–3466. [PubMed: 16052625]
19. Shen Y, Strittmatter EF, Zhang R, Metz TO, Moore RJ, Li F, Udseth HR, Smith RD, Unger KK, Kumar D, Lubda D. Making broad proteome protein measurements in 1-5 min using high-speed RPLC separations and high-accuracy mass measurements. Anal. Chem 2005;77(23):7763–7773. [PubMed: 16316187]
20. Zimmer JS, Monroe ME, Qian WJ, Smith RD. Advances in proteomics data analysis and display using an accurate mass and time tag approach. Mass Spectrom. Rev 2006;25(3):450–482. [PubMed: 16429408]
21. Petritis K, Kangas LJ, Ferguson PL, Anderson GA, Pasa-Tolic L, Lipton MS, Auberry KJ, Strittmatter E, Shen Y, Zhao R, Smith RD. Use of artificial neural networks for the prediction of peptide liquid chromatography elution times in proteome analyses. Anal. Chem 2003;75(5):1039–1048. [PubMed: 12641221]
22. Qian WJ, Jacobs JM, Camp DG 2nd, Monroe ME, Moore RJ, Gritsenko MA, Calvano SE, Lowry SF, Xiao W, Moldawer LL, Davis RW, Tompkins RG, Smith RD. Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry. Proteomics 2005;5(2):572–584. [PubMed: 15627965]
23. Ray MB, Zumwalt RE. Identification of alpha-1-proteinase inhibitor-containing cells in pancreatic islets. Cell Tissue Res 1986;243(3):677–680. [PubMed: 3513960]

24. Lu D, Mulder H, Zhao P, Burgess SC, Jensen MV, Kamzolova S, Newgard CB, Sherry AD. ^{13}C NMR isotopomer analysis reveals a connection between pyruvate cycling and glucose-stimulated insulin secretion (GSIS). Proc. Natl. Acad. Sci. USA 2002;99(5):2708–2713. [PubMed: 11880625]
25. Roduit R, Nolan C, Alarcon C, Moore P, Barbeau A, Delghingaro-Augusto V, Przybykowski E, Morin J, Masse F, Massie B, Ruderman N, Rhodes C, Poitout V, Prentki M. A role for the malonyl-CoA/long-chain acyl-CoA pathway of lipid signaling in the regulation of insulin secretion in response to both fuel and nonfuel stimuli. Diabetes 2004;53(4):1007–1019. [PubMed: 15047616]
26. Prentki M, Vischer S, Glennon MC, Regazzi R, Deeney JT, Corkey BE. Malonyl-CoA and long chain acyl-CoA esters as metabolic coupling factors in nutrient-induced insulin secretion. J. Biol. Chem 1992;267(9):5802–5810. [PubMed: 1556096]
27. Rhodes CJ. Type 2 diabetes – a matter of beta-cell life and death? Science 2005;307(5708):380–384. [PubMed: 15662003]
28. Tuttle RL, Gill NS, Pugh W, Lee JP, Koeberlein B, Furth EE, Polonsky KS, Naji A, Bimbaum MJ. Regulation of pancreatic beta-cell growth and survival by the serine/threonine protein kinase Akt1/PKBalpha. Nat. Med 2001;7(10):1133–1137. [PubMed: 11590437]
29. Maedler K, Storling J, Sturis J, Zuellig RA, Spinas GA, Arkhammar POG, Mandrup-Poulsen T, Donath MY. Glucose- and interleukin-1{beta}-induced {beta}-cell apoptosis requires Ca2+ influx and extracellular signal-regulated kinase (ERK) 1/2 activation and is prevented by a sulfonylurea receptor 1/inwardly rectifying K+ channel 6.2 (SUR/Kir6.2) selective potassium channel opener in human islets. Diabetes 2004;53(7):1706–1713. [PubMed: 15220194]
30. Khoo S, Griffen SC, Xia Y, Baer RJ, German MS, Cobb MH. Regulation of insulin gene transcription by ERK1 and ERK2 in pancreatic beta cells. J. Biol. Chem 2003;278(35):32969–32977. [PubMed: 12810726]
31. Flodstrom-Tullberg M, Yadav D, Hagerkvist R, Tsai D, Secrest P, Stotland A, Sarvetnick N. Target cell expression of suppressor of cytokine signaling-1 prevents diabetes in the NOD mouse. Diabetes 2003;52(11):2696–2700. [PubMed: 14578288]
32. Yashpal NK, Li J, Wheeler MB, Wang R. Expression of {beta}1 integrin receptors during rat pancreas development – sites and dynamics. Endocrinology 2005;146(4):1798–1807. [PubMed: 15618357]
33. Hammar EB, Irminger JC, Rickenbach K, Parnaud G, Ribaux P, Bosco D, Rouiller DG, Halban PA. Activation of NF-kappaB by extracellular matrix is involved in spreading and glucose-stimulated insulin secretion of pancreatic beta cells. J. Biol. Chem 2005;280(34):30630–30637. [PubMed: 15994334]

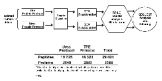


Figure 1.
Schematic of proteomic approach and overall results.

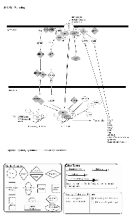


Figure 2. Proteomic coverage of the JAK/Stat signaling pathway as determined by Ingenuity Pathways Analysis

Those proteins identified in this work by two ore more peptides are shaded in grey.

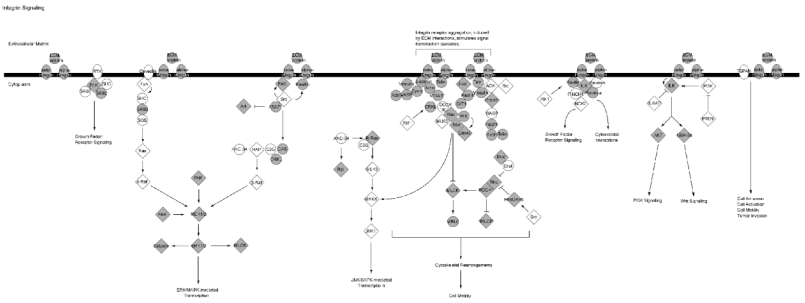


Figure 3. Proteomic coverage of the integrin receptor signaling pathway as determined by Ingenuity Pathways Analysis

Those proteins identified in this work by two or more peptides are shaded in grey.

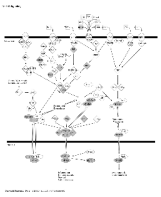


Figure 4. Proteomic coverage of the NF-κβ signaling pathway as determined by Ingenuity Pathways Analysis

Those proteins identified in this work by two ore more peptides are shaded in grey.

Table 1
Sequence coverage of the glycolytic and gluconeogenic enzymes.

The total sequence coverage and total number of unique peptides (urea and TFE protocol combined) are reported for the traditional enzymes of the glycolytic and gluconeogenic pathways. Major isoforms are also reported, and only those proteins identified by 2 or more unique peptides are shown.

Protein Description	Molecular Weight (Da)	IPI Number	Sequence coverage (%)	No. of Unique Peptides
Hexokinase				
Splice isoform 1 of hexokinase, type I	102437	IPI00018246.3	12	7
Hexokinase domain containing 1	102478	IPI00414612.2	2	2
Glucose 6-phosphate isomerase				
Glucose-6-phosphate isomerase	62976	IPI00027497.4	59	33
Phosphofructokinase				
Splice isoform 2 of 6-phosphofructo-2-kinase	54372	IPI00220808.1	40	18
6-phosphofructo-2-kinase	60304	IPI00004511.2	10	5
Triose phosphate isomerase				
Triosephosphate isomerase 1	26653	IPI00465028.4	24	6
Similar to triosephosphate isomerase	26926	IPI00383071.1	49	16
Aldolase				
Fructose-bisphosphate aldolase A	39264	IPI00465439.4	73	36
Fructose-bisphosphate aldolase B	39317	IPI00218407.5	40	13
Fructose-bisphosphate aldolase C	39300	IPI00418262.3	44	17
Glyceraldehyde-3-phosphate dehydrogenase				
Glyceraldehyde-3-phosphate dehydrogenase, liver	35899	IPI00219018.4	92	40
Glyceraldehyde-3-phosphate dehydrogenase, muscle	35853	IPI00383758.1	22	7
Phosphoglycerate kinase				
Phosphoglycerate kinase 1	44455	IPI00169383.2	83	45
Phosphoglyceratemutase				
Phosphoglycerate mutase 1	28655	IPI00385244.4	33	7
Phosphoglycerate mutase 2	28617	IPI00218570.5	30	8
Enolase				
Alpha enolase	47008	IPI00465248.4	64	30
Beta enolase	46826	IPI00218474.4	13	3
Gamma enolase	47108	IPI00216171.2	86	37
Pyruvate kinase				
Pyruvate kinase 3 isoform 2	58025	IPI00220644.6	72	54
Pyruvate kinase M2	57744	IPI00383237.3	9	5
Fructose-1,6-bisphosphatase				
Fructose-1,6-bisphosphatase	36660	IPI00073772.4	34	10
Pyruvate carboxylase				

Protein Description	Molecular Weight (Da)	IPI Number	Sequence coverage (%)	No. of Unique Peptides
Pyruvate carboxylase	129551	IPI00299402.1	19	13
Phosphoenolpyruvate carboxykinase				
Phosphoenolpyruvate carboxykinase	70592	IPI00294380.3	40	21

Table 2
Sequence coverage of the citrate cycle enzymes.

The total sequence coverage and total number of unique peptides (urea and TFE protocols combined) are reported for the traditional enzymes of the citrate cycle. Major isoforms are also reported, and only those proteins identified by 2 or more unique peptides are shown.

ProteinDescription	Molecular Weight (Da)	IPI Number	Sequence Coverage (%)	No. of Unique Peptides
Pyruvate dehydrogenase				
Pyruvate dehydrogenase E1 component alpha subunit, somatic form, mitochondrial precursor	43268	IPI00306301.1	36	9
Splice isoform 1 of pyruvate dehydrogenase E1 component beta subunit, mitochondrial precursor	39194	IPI00003925.5	40	12
Pyruvate dehydrogenase protein X component, mitochondrial precursor	54089	IPI00298423.3	9	5
Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial precursor	65739	IPI00021338.1	16	8
Pyruvate dehydrogenase lipoamide-phosphatase 1, mitochondrial precursor	61181	IPI00218971.2	6	2
Pyruvate dehydrogenase lipoamide kinase isozyme 3, mitochondrial precursor	46909	IPI00014849.1	15	4
Citrate synthase				
Citrate synthase, mitochondrial precursor	51679	IPI00025366.4	44	12
ATP-citrate synthase	120748	IPI00021290.4	59	57
Aconitase				
Aconitate hydratase, mitochondrial precursor	85372	IPI00017855.1	51	39
Isocitrate dehydrogenase				
Isocitrate dehydrogenase (NADP), cytoplasmic	46629	IPI00027223.2	66	25
Isocitrate dehydrogenase (NADP), mitochondrial precursor	50877	IPI00011107.2	42	22
Isocitrate dehydrogenase (NAD) subunit alpha, mitochondrial precursor	39566	IPI00030702.1	43	13
Isocitrate dehydrogenase (NAD) subunit gamma, mitochondrial precursor	42767	IPI00220150.4	13	3
Splice isoform B of isocitrate dehydrogenase (NAD) subunit beta, mitochondrial precursor	42185	IPI00304417.5	32	8
Alpha-ketoglutarate dehydrogenase				
2-oxoglutarate dehydrogenase E1 component, mitochondrial precursor	113403	IPI00098902.3	24	14
Oxoglutarate dehydrogenase-like	114423	IPI00412522.3	13	8
Succinyl-CoA ligase				
Succinyl-CoA ligase (GDP-forming) alpha-chain, mitochondrial precursor	35025	IPI00295625.1	39	11
Succinyl-CoA ligase (GDP-forming) beta-chain, mitochondrial precursor	46481	IPI00096066.2	49	19
Splice isoform 2 of succinyl-CoA ligase (ADP-forming) beta-chain, mitochondrial precursor	48023	IPI00217232.1	43	13
Succinate dehydrogenase				

ProteinDescription	Molecular Weight (Da)	IPI Number	Sequence Coverage (%)	No. of Unique Peptides
Succinate dehydrogenase (ubiquinone) iron-sulfur protein, mitochondrial precursor	31609	IPI00294911.1	35	8
Succinate dehydrogenase (ubiquinone) flavoprotein subunit, mitochondrial precursor	72645	IPI00305166.1	8	3
Fumarate hydratase				
Fumarate hydratase, mitochondrial precursor	54602	IPI00296053.3	51	22
Malate dehydrogenase				
Malate dehydrogenase, cytoplasmic	36272	IPI00291005.7	55	18
Malate dehydrogenase, mitochondrial precursor	35509	IPI00291006.1	69	30