Analysis of the beta-cell N-glyco surface proteome using label-free quantitative mass spectrometry, and targeted candidate biomarker validation in mouse models of type 2 diabetes

Author(s): Danzer, Carsten Friedrich

Publication Date: 2009

Permanent Link: https://doi.org/10.3929/ethz-a-005940078

Rights / License: In Copyright - Non-Commercial Use Permitted
Analysis of the beta-cell N-glyco surface proteome using label-free quantitative mass spectrometry, and targeted candidate biomarker validation in mouse models of type 2 diabetes.

Submitted to
ETH ZURICH
for the degree of
DOCTOR OF SCIENCES

by
CARSTEN FRIEDRICH DANZER
Diploma-Biologist, University of Konstanz

Date of birth 21. September 1977
Citizen of Germany

accepted on the recommendation of
Professor Dr. Wilhelm Krek
Professor Dr. Claes Wollheim
Professor Dr. Ruedi Aebersold

2009
# Table of contents

1. Summary ............................................................................................................................... 4

2. Zusammenfassung: ............................................................................................................... 6

3. Abbreviations: ....................................................................................................................... 8

4. Introduction ........................................................................................................................ 11
   4.1. The pancreatic beta-cell ................................................................................................. 11
      4.1.1. The beta-cell in the islet of Langerhans ................................................................. 11
      4.1.2. Environmental influences on insulin secretion and beta-cell metabolism ............... 13
   4.2. Diabetes mellitus ............................................................................................................ 15
      4.2.1. Classifications of diabetes mellitus .......................................................................... 15
      4.2.2. Genetic susceptibility ............................................................................................. 17
   4.3. Animal models of disease ............................................................................................... 18
      4.3.1. Current rodent models in the study of beta-cell function ........................................ 18
      4.3.2. C57BL/6J High fat diet mouse model ................................................................. 18
      4.3.3. Leptin deficient ob/ob mouse model (Lep ob/ob) .................................................. 19
      4.3.4. Leptin receptor deficient db/db mouse model (db/db mice) .................................... 20
      4.3.5. ZDF rat model (fa/fa rats) .................................................................................... 20
   4.4. Proteomics ...................................................................................................................... 21
      4.4.1. Proteome-wide expression analysis ......................................................................... 21
      4.4.2. Membrane proteomics .......................................................................................... 22
      4.4.3. Quantitative proteome analysis ................................................................................ 23
      4.4.4. Labeling techniques for quantification .................................................................... 24
      4.4.5. Label-free quantification ....................................................................................... 25
      4.4.6. Absolute quantification ......................................................................................... 26
      4.4.7. Proteomics studies with pancreatic islets and beta-cells .......................................... 26

5. Results .................................................................................................................................. 28
   5.1. MIN6 Cell Surface Capturing catalogue ....................................................................... 28
   5.2. MIN6 Whole Cell Capturing catalogue ....................................................................... 30
   5.3. Human islets Whole Cell Capturing catalogue ............................................................. 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4. Islet/beta-cell N-glyco protein catalogue</td>
<td>32</td>
</tr>
<tr>
<td>5.5. Technical and biological variation</td>
<td>34</td>
</tr>
<tr>
<td>5.6. Quantitative CSC of MIN6 cells stimulated with 25 mM D-glucose</td>
<td>34</td>
</tr>
<tr>
<td>5.7. Quantitative CSC of MIN6 cells stimulated with 25 mM D-glucose plus 20 nM GLP-1</td>
<td>38</td>
</tr>
<tr>
<td>5.8. Beta-cell biomarker studies</td>
<td>43</td>
</tr>
<tr>
<td>5.9. Diet induced obesity (DIO) mouse model blood analysis</td>
<td>44</td>
</tr>
<tr>
<td>5.10. ob/ob mouse model blood analysis</td>
<td>48</td>
</tr>
<tr>
<td>5.11. Results from shotgun approach with DIO plasma samples</td>
<td>52</td>
</tr>
<tr>
<td>5.12. Results from MRM approach in DIO and ob/ob mouse models</td>
<td>53</td>
</tr>
<tr>
<td>6. Discussion</td>
<td>62</td>
</tr>
<tr>
<td>6.1. The N-glyco proteomes of beta-cells and islets</td>
<td>62</td>
</tr>
<tr>
<td>6.2. Surface protein dynamics under glucose treatment</td>
<td>63</td>
</tr>
<tr>
<td>6.3. Surface protein dynamics under glucose in combination with GLP-1 treatment</td>
<td>70</td>
</tr>
<tr>
<td>6.4. Candidate Biomarker evaluation in mouse models of T2D disease</td>
<td>81</td>
</tr>
<tr>
<td>7. Material and Methods</td>
<td>88</td>
</tr>
<tr>
<td>7.1. Animal models</td>
<td>88</td>
</tr>
<tr>
<td>7.2. Mouse blood plasma sampling</td>
<td>88</td>
</tr>
<tr>
<td>7.3. Mouse blood plasma analysis</td>
<td>88</td>
</tr>
<tr>
<td>7.4. Intraperitoneal Glucose Tolerance Testing (IPGTT)</td>
<td>88</td>
</tr>
<tr>
<td>7.5. Mass Spectrometry</td>
<td>88</td>
</tr>
<tr>
<td>7.6. LC-MS/MS-Analysis</td>
<td>89</td>
</tr>
<tr>
<td>7.7. Database search</td>
<td>90</td>
</tr>
<tr>
<td>7.8. Cell Surface Capturing (CSC)</td>
<td>91</td>
</tr>
<tr>
<td>7.9. Whole Cell Capturing (WCC)</td>
<td>92</td>
</tr>
<tr>
<td>7.10. Directed LC-MS/MS</td>
<td>92</td>
</tr>
<tr>
<td>7.11. Label-free Quantification</td>
<td>93</td>
</tr>
</tbody>
</table>
Table of contents

7.12. N-glycosite blood sample processing ................................................................. 93
7.13. Reference Peptides ............................................................................................ 94
7.15. Off-gel electrophoresis ...................................................................................... 94
7.16. Chemicals ......................................................................................................... 94
7.17. Cell lines and culture conditions ....................................................................... 95
7.18. Human Pancreatic Islets .................................................................................. 95
7.19. Insulin secretion assay and insulin content ....................................................... 95
7.20. Q-PCR ............................................................................................................. 95
7.21. Bioinformatics analysis ...................................................................................... 96
7.22. Primer pairs ...................................................................................................... 96
7.23. Statistics .......................................................................................................... 97

8. Acknowledgement .................................................................................................. 98

9. References: .......................................................................................................... 101

10. Curriculum vitae .................................................................................................. 117
1. Summary

The pancreatic beta-cell, which is embedded in the functional unit of the islet of Langerhans, is the key player in maintaining blood glucose homeostasis. This is a systemic and complex interplay of food uptake, all inner organs, and several signaling pathways in particular inside the pancreatic beta-cell. The cell surface proteome as the origin of cell signaling is the entry point feeding the complex network of signaling events that control insulin vesicle exocytosis. Therefore we applied the Cell Surface Capturing (CSC) (Tian et al., 2007) technology to the murine beta-cell line MIN6-B1. Using this as a discovery approach, we selectively enriched N-glycosylated peptides at the exterior of the beta-cell, and generated a high quality mouse beta-cell surface N-glyco protein catalogue. The obtained results allowed relative, label-free quantification of cell surface proteins under two different stimuli: glucose alone, and in combination with GLP-1. After one hour of stimulation with 25 mM glucose plus 20 nM GLP-1 we found 24 proteins significantly regulated. Among others we found a 0.90-fold log2 ratio decrease in abundance at the surface of the GLP-1 receptor, and a 0.30-fold log2 ratio increase of the GLUT1 transporter.

Furthermore we applied the Whole Cell Capturing (WCC) technology to human pancreatic islets, and murine MIN6-B1 cells. This technology enriches for N-glycopeptides in an organelle independent manner. Combining these results with the CSC data we created a comprehensive data-set of the mouse, and human beta-cell surface N-glyco proteome. Since we compare a murine beta-cell line with primary human islets, the overlap of the proteomes can be considered as surface proteins being expressed in mouse, and human primary beta-cells. In comparison to former proteomics approaches investigating islets or beta-cell lines (D'Hertog et al., 2007; Dowling et al., 2006; Fernandez et al., 2008; Johnson et al., 2006; Ortsater and Bergsten, 2006; Ortsater et al., 2007), this approach provides in-depth quantitative insight into the beta-cell surface proteome, and its changes upon different stimuli in cell culture.

The created database of the comprehensive information on the beta-cell surface N-glyco proteome was processed using bioinformatics approaches to identify potential candidate biomarkers, which were validated in two different mouse models for type 2 diabetes (DIO and ob/ob). For this validation we used the target Multiple Reaction Monitoring proteomics approach, which enables the absolute quantification of a defined large number of peptides in a complex sample using synthetic reference peptides. This approach enabled us to test 317 potential biomarker candidates with synthetic AQUA peptides in mice blood samples,
whereas 54 of these peptides matching to 48 proteins could be quantified. Four of these proteins (MFAP4, PLTP, VASN, DSG2) showed a difference in abundance, when comparing the control with the disease group. In addition we also applied an unbiased label-free proteomics approach with the samples obtained from the DIO mouse model. The overall peptide pattern changes enabled us to distinguish the control from the disease group after 16 weeks of feeding. At this time-point we found significantly regulated proteins involved in blood coagulation, which could be due to an early onset of thrombosis and/or atherosclerosis, both known to take place in disease models involving obesity.

Finally we can say that we successfully demonstrated that the mass spectrometry based approaches we applied are feasible tools to indentify and quantify cell surface proteins. Furthermore we demonstrated that this methodology enables us to quantify potential biomarkers in blood samples derived from mouse models of Type 2 Diabetes.
2. Zusammenfassung


Des Weiteren wandten wir noch die „Whole Cell Capturing (WCC)“ Technik auf die murine MIN6 B1 beta-Zell Linie und primäre humane Langerhanssche Inseln an. Diese Methode erlaubt es ebenfalls N-glycosylierte Peptide anzureichern, sie ist aber nicht spezifisch für Oberflächenproteine. Die Kombination dieser Ergebnisse mit denen des CSC erstellten wir einen umfangreichen Datensatz des murinen und des humanen beta-Zell Oberflächen Proteoms. Für die Überschneidung der drei Proteome kann angenommen werden, dass die enthaltenen Oberflächenproteine tatsächlich an der Oberfläche von primären beta-Zellen von Maus und Mensch exprimiert sind, da wir eine murine beta-Zell Linie mit primären humanen Inseln verglichen. Im Vergleich zu früheren Proteomics Ansätzen, welche sich mit der Untersuchung von Langerhansschen Inseln oder beta-Zell Linien beschäftigen (D’Hertog et al., 2007; Dowling et al., 2006; Fernandez et al., 2008; Johnson et al., 2006; Ortsater and Bergsten, 2006; Ortsater et al., 2007), liefert dieser Ansatz einen tiefen quantitativen Einblick in das beta-Zell Proteome und seine Änderung auf Grund verschiedener Stimuli in Zellkultur.


Abschliessend lässt sich sagen, dass die von uns angewandten massenspektrometrischen Ansätze nützliche Werkzeuge sind um Oberflächenproteine der Zelle zu identifizieren und zu quantifizieren. Ausserdem konnten wir zeigen, dass diese Methoden dazu verwendet werden können, um potentielle Biomarker im Blut von Typ 2 Diabetes Maus Modellen zu quantifizieren.
Abbreviations

3. Abbreviations

AA  amino acids
ATP  adenosine triphosphate
BCS  bovine calf serum
BMI  body mass index
bp   base pair
CDC  Centers for Disease Control and Prevention
cDNA complementary desoxyribonucleic acid
CSC  Cell Surface Capturing
CT   C-terminal
2DGE 2-dimensional gel electrophoresis
DAPI 4’6’-Diamidine-2’-Phenylindole Dihydrochloride
DIO  Diet induced obesity
DMEM Dulbecco’s modified Eagle’s Medium
dNTP desoxyribonucleoside-triphosphate
ECL enhanced chemi-luminescence
EDTA ethylenediamine tetra-acetic acid
EtOH ethanol
FCS fetal calf serum
FFA free fatty acids
FITC fluorescein
GI  gastrointestinal
GIIS glucose induced insulin secretion
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLP-1</td>
<td>glucagon-like-peptide-1</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine triphosphate</td>
</tr>
<tr>
<td>HBSS</td>
<td>Hank’s balanced salt solution</td>
</tr>
<tr>
<td>ICAT</td>
<td>isotope-coded affinity tags</td>
</tr>
<tr>
<td>IPGTT</td>
<td>intraperitoneal glucose tolerance test</td>
</tr>
<tr>
<td>IPI</td>
<td>International Protein Index</td>
</tr>
<tr>
<td>IPITT</td>
<td>intraperitoneal insulin tolerance test</td>
</tr>
<tr>
<td>iTRAQ</td>
<td>isobaric tag for relative and absolute quantitation</td>
</tr>
<tr>
<td>kb</td>
<td>kilo-base (i.e. 1000 nucleotides)</td>
</tr>
<tr>
<td>KCl</td>
<td>potassium chloride</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>m</td>
<td>milli (10^-3)</td>
</tr>
<tr>
<td>μ</td>
<td>micro (10^-6)</td>
</tr>
<tr>
<td>MgCl2</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>MODY</td>
<td>maturity-onset of diabetes of the young</td>
</tr>
<tr>
<td>MRM</td>
<td>multiple reaction monitoring</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>Na2HPO4</td>
<td>sodium hydrogen phosphate</td>
</tr>
<tr>
<td>NaAc</td>
<td>sodium acetate</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NaH2PO4</td>
<td>sodium dihydrogen orthophosphate</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>NaOH</td>
<td>sodium hydroxide</td>
</tr>
<tr>
<td>NGS</td>
<td>normal goat serum</td>
</tr>
<tr>
<td>NIDDM</td>
<td>non-insulin dependent diabetes mellitus</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase PCR</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SILAC</td>
<td>stable isotope labeling with amino acids in cell culture</td>
</tr>
<tr>
<td>T2D</td>
<td>type 2 diabetes</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N’,N’-Tetramethylethylenediamine</td>
</tr>
<tr>
<td>VEGF</td>
<td>vascular endothelial growth factor</td>
</tr>
<tr>
<td>WCC</td>
<td>Whole Cell Capturing</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WT</td>
<td>wild-type</td>
</tr>
<tr>
<td>ZDF</td>
<td>Zucker diabetic fatty (rat)</td>
</tr>
</tbody>
</table>
4. Introduction

In our affluent society the number of diseases connected to metabolic disorders is constantly increasing. These disorders are summarized as the metabolic syndrome, comprising of arteriosclerosis, coronary diseases, and type 2 diabetes (T2D). Since the prevalence rate of patients diagnosed as type 2 diabetics doubled in the United States (U.S.) between 1990 and 2005, the Centers for Disease Control and Prevention (CDC) characterized it as an epidemic. Observations of an increasing frequency in young people, and even in children being diagnosed as type 2 diabetics, correlates with increasing rates of obesity in this population. Therefore the traditional consideration of T2D as a disease of adults is not true anymore. The World Health Organization (WHO) reports at least 200 million people with diabetes, whereas this number is expected to double by 2030. They also report about 3.2 million deaths per year attributed to complications of diabetes, which cause up to 15% of total health care costs, depending on the region.

4.1. The pancreatic beta-cell

4.1.1. The beta-cell in the islet of Langerhans

The beta-cells are the major cell type involved in maintaining blood glucose homeostasis. Beta-cells are organized into clusters of cells called Islet of Langerhans, which is crucial to maintain beta-cells functionality. An islet consists of 65-80% beta-cells (which secrete insulin) in its core, and on the periphery about 15-20% alpha-cells (which secrete glucagon), 3-10% delta-cells (which secrete somatostatin), PP-cells 3-5% (which secrete pancreatic polypeptide), and <1% epsilon-cells (which secrete grehlin). Studies in rodents showed that every islet consists of 2000 to 4000 cells (Kulkarni, 2004). Every islet is surrounded by exocrine pancreatic tissue that produces, and releases digestive enzymes, which are released into the gastrointestinal (GI) system. The number of islets per pancreas varies dramatically between species. In mice approximately 1000 to 2500 islets are estimated (Bock et al., 2005), whereas for a human pancreas the number is thought to be close to 2 million (Kulkarni, 2004). The architecture of islets differs between species (figure 1), particularly between rodents and humans. Rodent islets have a predominant core of beta-cells, whereas human islets show a distribution of alpha- and beta-cells in close relationship over the entire islet cell cluster (Suckale and Solimena, 2008).
Introduction

The most important function of the beta-cell is the production of insulin and in response to elevated glucose levels the appropriate secretion of the hormone. Upon insulin secretion the liver decreases glucose production, and at the same time glucose uptake, utilization and storage in fat and muscle is increased. Glucagon is the antagonist of insulin which, after release from the alpha-cells, reverses the actions of insulin. The hormone somatostatin is derived from the delta-cells, and serves as a controller in the balance of insulin and glucagon, thereby inhibiting the release of both proteins. The role of the pancreatic polypeptide is still unclear.

As endocrine glands the islets of Langerhans release their hormones directly into the bloodstream rather than through ducts. For the release of their hormones they rely on a dense network of fenestrated capillaries to keep proper functionality (Lammert et al., 2003b). Natural islet hyperplasia occurs during development, due to a higher production and secretion of insulin. This in turn triggers the expression of vascular endothelial growth factor A (VEGFA) to increase islet vascularization (Lammert et al., 2003a).

If the natural demand for insulin changes in the long-term, beta-cells compensate by adapting functionality and mass. This is for example the case during pregnancy, where beta-cell mass increases to handle the enlarged metabolic demand, and decreases to normal levels after the end of pregnancy (Blondeau et al., 1999; Van Assche et al., 1980). This plasticity of beta-cell mass can partially explain why more than 70% of obese individuals never develop diabetes. Even though insulin resistance can be found in most of the cases of obesity, the onset of type 2 diabetes still requires failure of the beta-cells, the reasons for which are not entirely understood. Nevertheless this failure must originate from external stimuli reaching the cells via the plasma membrane either by direct cell-to-cell contact, metabolites, soluble mediators such as hormones or most probably by a combination of all these. In either way the surface proteome, including receptors, channels, enzymes and adhesion molecules is crucial to integrate the information from outside into the internal signaling pathways involved in the beta-cell metabolism and signaling.
4.1.2. Environmental influences on insulin secretion and beta-cell metabolism

The function of the pancreatic beta-cell as a glucose sensor, and the secretion of appropriate amounts of insulin to keep blood glucose levels are well established. Insulin secretion is not only influenced by glucose, but also by hormonal, neuronal, and nutritional signals. Nevertheless, glucose is the main regulator of insulin secretion, while other factors such as glucagon-like-peptide-1 (GLP1), free fatty acids (FFA), or amino acids have no effect by themselves, but serve as amplifier of glucose induced insulin secretion (GIIS). GIIS for itself is dependent on a metabolic pathway that couples glucose uptake to increased levels of ATP production (Newgard et al., 2002). The subsequent changes in the ATP/ADP ratio results in the closure of potassium channels, which induces a depolarization of the plasma membrane, and activation of voltage dependent calcium channels (Safayhi et al., 1997). This results in a rise of intracellular calcium, which provides the signal for the fusion of primed insulin-containing vesicles with the plasma membrane (Barg, 2003). The initiating event in this process is the cellular uptake of glucose via the cell surface facilitative glucose transporter.
protein GLUT2 (SLC2A2) that functions to transport glucose in proportion to circulating levels (Thorens et al., 2000). Thus, GLUT2 is considered a crucial part of the beta-cell glucose sensing mechanisms. Knockout of pancreatic beta-cell GLUT2 in mice leads to hyperglycemia, hypoinulinemia, hyperglucagonemia, and early death. The knockout can be fully rescued by re-expression of both GLUT1 ($K_m \sim 3$ mM) and GLUT2 ($K_m \sim 17$ mM), which means that normal GSIS is not dependent on transporter affinity but on the rate of uptake at stimulatory glucose concentrations.

To control the effects of insulin on glucose homeostasis, the body uses hormones like glucagon, ghrelin, and leptin to reduce the insulin secretion. Glucagon is secreted during fasting, suppresses insulin secretion, and at the same time stimulates hepatic glucose output. In the pancreatic beta-cell it stimulates cAMP formation (GLP-1 and PACAP do the same), and suppresses the insulin gene expression via activation of the inducible cAMP early repressor (ICER) (Hussain et al., 2000). Another suppressor of insulin secretion is the hormone leptin, which activates the $K_{ATP}$ channels, and thereby disables the depolarization of the plasma membrane (Kieffer et al., 1997). Grehlin is suggested to influence insulin secretion due to a tonic regulation of beta-cells, and prompting the inhibition of both insulin and somatostatin secretion (Egido et al., 2002).

The most famous enhancer of insulin secretion is the hormone GLP-1 due to its anti-diabetic action, which includes the inhibition of glucagon release, inhibition of beta-cell apoptosis, increase in insulin biosynthesis, and increase beta-cell proliferation (Ahren, 1998; Drucker, 2006). It remains controversial whether there is a direct or indirect effect on the glucagon secretion by alpha-cells, since some have reported GLP-1 receptor expression in alpha-cells (Heller et al., 1997), whereas others have reported the contrary (Tornehave et al., 2008). GLP-1 is a very powerful incretin hormone produced in the L-cells of the distal ileum and colon (Ahren, 1998). It is generated by tissue-specific posttranslational processing of proglucagon (MacDonald et al., 2002) and up-regulated by nutrients such as glucose or fatty acids. After binding, the GLP-1 receptor signals to the alpha subunit of $G_s$, with subsequent activation of adenylate cyclase (AC) and increase of cAMP levels (Doyle and Egan, 2007), thereby potentiating glucose stimulated insulin secretion. It inhibits the beta-cell $K_{ATP}$ channels mediated via cAMP/PKA, which results in plasma membrane depolarization followed by $Ca^{2+}$ influx through voltage-dependent $Ca^{2+}$ channels (VDCC), which triggers the exocytosis of insulin vesicles. With the threshold of an extracellular glucose concentration of 4 mmol/L glucose, the insulinotropic effect of GLP-1 is glucose dependent. Experiments in rodents
show that the peptide also has long-term effects on the islets, since it stimulates beta-cell proliferation and inhibits apoptosis (Drucker, 2003). GLP-1 also delays gastric emptying and reduces body weight by initiating satiety and reducing food intake (Drucker, 2006). Even though GLP-1 was thought to be a potential drug for type 2 diabetes treatment (Gutniak et al., 1992), it took a long time till the first compound came to the market. The reason for this is the fast inactivation by the enzyme DPP4, thereby limiting the half life for GLP-1 to 1-2 minutes (Holst and Deacon, 1998). Therefore GLP-1 itself cannot be used a therapeutic since it would need very frequent injections or continuous infusion. The drug developing companies overcame this difficulty with two different strategies: one is the development of a compound, in particular a peptide, which mimics GLP-1 and is resistant to DPP4 cleavage to increase the half life of GLP-1 (Bond, 2006); and the other one is to inhibit the activity of DPP4 to increase the amount of endogenous GLP-1 (Ahren, 2008). Both strategies have been successful and are already on the market (Byetta, Amylin/Eli Lilly) or in late clinical development (RR1583, Roche). The big step forward in the development of GLP-1 mimetics was the discovery of exendin 4, which was isolated from a lizard called the Gila monster. The peptide homology between GLP-1 and exendin 4 is only 53%, but it binds to the GLP-1 receptor with similar affinity as native GLP-1 and most importantly, it is resistant to DPP4 (Bond, 2006).

4.2. Diabetes mellitus

4.2.1. Classifications of diabetes mellitus

About 10% of diabetes mellitus cases are diagnosed as type 1 diabetes which is an immune-mediated deficiency and characterized by loss of the insulin-producing beta-cells of the islets of Langerhans in the pancreas. No preventive measure is known which can be taken to protect from type 1 diabetes. Most of the affected people are generally healthy and of a healthy weight when onset occurs. Especially in the early stages of disease patients remain normal in their sensitivity and responsiveness to insulin. The majority of type 1 diabetes patients are children, but it can also affect adults. The patients have to monitor their blood glucose levels very carefully and inject artificial insulin to keep their blood glucose homeostasis. Thanks to the ‘Edmonton protocol’ it is possible to successfully transplant pancreatic islets to compensate for type I diabetes. In combination with immuno suppressants they remain functional for about 5 years then they lose their functionality (Shapiro et al., 2000).
Introduction

Figure 2. Schematic overview of the cellular processes that are involved in stimulating insulin secretion from pancreatic beta-cells (Ahren, 2009).

The most common form of diabetes mellitus is type 2 diabetes which is characterized by peripheral insulin resistance or sensitivity in combination with reduction of insulin secretion due to failure or finally loss of pancreatic beta-cells. In early stages increased blood insulin levels are found, which can be explained by the loss of insulin sensitivity of peripheral tissues. At these early stages the treatment for type 2 diabetes is often planning a diet in combination with exercising that the patients lose weight, and can keep their blood glucose levels better under control. In most of the cases these measures are not enough to bring blood glucose levels down to a near normal range. The next step is treatment with oral drugs to lower the blood glucose levels. The classical one used in most cases is metformin, which is a biguanide drug. It lowers blood glucose levels by decreasing the amount of glucose which is released by the liver. In addition it increases the insulin sensitivity, which leads to a higher glucose uptake by the peripheral tissues and a higher efficiency of insulin. Other types of drugs are the sulphonylurea drugs: glibenclamide, gliclazide, glimepiride, gliclizide, tolbutamide, and glipizide. All of them work by increasing the insulin production of the beta-cells. The thiazolidinediones, which are commonly called glitazones exist as two types called
Introduction

Pioglitazone and rosiglitazone. They increase the sensitivity of insulin responsive cells and therefore lower the blood glucose levels. Usually they are used in combination with metformin or one of the sulphonylureas. Nateglinide and repaglinide are not commonly used, as they have a similar action as the sulphonylureas. They have a very fast effect on insulin release, which does not last very long, so each dose has to be taken shortly before the meals.

A completely different type of drug is acarbose, which works by delaying the absorption of carbohydrates from the gut. Therefore, it is able to reduce the peaks of blood glucose which occur after meals. It is not used very often since many people develop gut-related side-effects such as bloating, wind, and diarrhea when taking acarbose.

If people continue to have high blood glucose levels despite lifestyle changes in combination with oral drugs insulin must be injected to lower blood glucose. A study of Hedley et al. showed that more than 85% of the patients which were diagnosed as type 2 diabetics were either overweight or obese (Hedley et al., 2004). This suggests a connection of the development of T2D to the body weight. The world health organization (WHO) defines overweight or obesity via the body mass index (BMI) which is body weight divided by the square of the person’s size \([\text{kg/m}^2]\). According to this a BMI >25 defines the person as overweight and >30 is the classification for obese.

4.2.2. Genetic susceptibility

Certain gene defects that lead to the development of diabetes were classified in a subgroup called maturity-onset of diabetes of the young (MODY) (Tattersall, 1998). The term MODY originates from the time, when diabetes mellitus was considered to have two main forms: juvenile-onset and maturity-onset. The modern understanding of the pathophysiology of diabetes has changed, therefore the concept and usage of MODY has been refined. Today it does not include any form of type 2 diabetes and is a synonym for dominantly inherited, monogenic defects of insulin secretion occurring at any age. The different forms of MODY are all due to ineffective insulin production or release by pancreatic beta-cells. These defects are caused by different transcription factors as well as important regulatory proteins such as hepatocyte nuclear factor 4 alpha (HNF4alpha, MODY1) (Yamagata et al., 1996a), glucokinase (GK, MODY2) (Froguel et al., 1993), hepatocyte nuclear factor 1 alpha (HNF1alpha, MODY3) (Yamagata et al., 1996b), insulin promoter factor 1 (IPF1/PDX1, MODY4) (Stoffers et al., 1997), hepatocyte nuclear factor 1 beta (TCF2, MODY5) (Horikawa et al., 1997), and neurogenic differentiation 1 (NEUROD1, MODY6) (Malecki et al., 1999).
4.3. Animal models of disease

4.3.1. Current rodent models in the study of beta-cell function

The study of animal models to investigate and understand the molecular mechanisms of diseases is of great importance, especially in a complex setting as in the metabolic syndrome in particular diabetes, in which various organs and tissues are affected. Several tissue specific knockout as well as transgenic rodent models exist that help to decipher the molecular aberrations underlying the development of type 2 diabetes. Thereby, animal models allow not only studying the influence of genetic variations but also allow following up pre-diabetic stages as well as the onset and the development of type 2 diabetes.

In contrast to humans, rodent models have a rather defined genetic background (inbred strains) and phenotyping as well as experiments can be performed in a defined environment. These factors reduce the noise from other genetic and environmental impacts. Further advantages are a high reproduction rate and the possibility for tissue sampling.

However, it is clear that rodent models do not phenocopy perfectly all metabolic characteristics of humans. Whereas basal glucose levels of 7.8 mmol/L and higher are normal for most inbred mouse strains, they would be considered as diabetic for humans. Therefore there are suggestions that only mice with basal glucose levels equal or higher than 13.9 mmol/L should be considered as diabetic (Clee and Attie, 2007). However, there is no absolute characterized standard yet due to these discrepancies. Further differences are found in islet structure as well as in the kinetics of the response on glucose or other insulin secretagogues. In addition, some rodent strains are completely protected from reaching a diabetic stage. The following subchapters give a short overview over some important rodent models that were either used or were of particular interest for the presented research.

4.3.2. C57BL/6J High fat diet mouse model

The C57BL/6J strain is often referred to as the “classical” laboratory mouse strain and it is the most commonly used reference strain in studies of diabetes (Clee and Attie, 2007) as several studies showed diabetes susceptibility (Rebuffe-Scrive et al., 1993; Surwit et al., 1988). Unlike the diabetes resistant C57BL/6A strain, the C57BL/6J develops clear diabetes with fasting blood glucose levels greater than 13.3 mmol/L when on a high fat diet. However, refeeding these mice using normal chow leads ultimately to a complete reversal of the diet-induced obesity and diabetes (Parekh et al., 1998). There is only one particular C57BL/6J strain known, which originates from the Jackson laboratories showing a change in the glucose
tolerance and having a general metabolic imbalance (Freeman et al., 2006a). This metabolic disorder is caused by a mutation of the nicotinamide nucleotide transhydrogenase (Nnt). Nnt is linked to insulin secretion, glucose metabolism and oxidative stress (Freeman et al., 2006a) and was shown to cause glucose intolerance in C57BL/6J mice (Freeman et al., 2006b).

4.3.3. Leptin deficient ob/ob mouse model (Lep ob/ob)

This mouse model has a nonsense point mutation in the leptin gene (ob mutation) leading to an inactive protein. The leptin hormone is produced and secreted by adiposites and signals satiety to the brain. Due the complete absence of the hormone in homozygous animals, the signal cannot be transmitted and as a consequence these mice feel permanently hungry. The obese mice (ob/ob) are phenotypically distinguished from their lean littermates (ob/+ and +/+ ) at about 26 days of age on basis of body weight. The obese mice gain rapidly in weight and have a marked obesity at 5 weeks of age. They are further characterized by hyperphagia, transient hyperglycemia and markedly elevated plasma insulin levels, which could be associated with an increase in number and size of the beta-cells of the islets of Langerhans (Baetens et al., 1978). Serum insulin level is increased by day 17-21 and hypoglycemia develops. However, because of the rapid increase in serum insulin level, hyperglycemia becomes a prominent feature of the obese mouse (19-22 mmol/L). Marked insulin resistance associated with loss of insulin receptors in several tissues follows hyperinsulinemia (Lord and Atkins, 1985). Serum insulin levels reach a peak level by 6-8 month (25-30 ng/mL), and the blood glucose level seems to normalize (Herberg and Coleman, 1977; Westman, 1968). The rate of lipogenesis in the liver and the adipose tissue is more than doubled and both intraperitoneal and subcutaneous deposits of fat are increased (Bray and York, 1979; Herberg and Coleman, 1977). The fat deposit continues despite of food restriction and exercise in young ob/ob mice (Dubuc et al., 1984), indicating a more efficient absorption from the intestine and a higher metabolic utilization of nutrients (Bihler and Freund, 1975; Bray and York, 1979). Gluconeogenesis is enhanced despite the hyperinsulinemic state in obese mice (Bray and York, 1979; Herberg and Coleman, 1977). The adrenal glands of obese mice are enlarged compared to control and the enlargement is related to the cortical part, where an increased synthesis of corticosterone was found. The secretion of glucagon is also elevated and it has been suggested that the hyper function of the alpha-cells in the pancreas of obese mice are involved in the pathogenesis of the obese-hyperglycemic syndrome (Herberg and Coleman, 1977). The morphological appearance of the pancreas in the obese mice is characterized by a marked hypertrophy and hyperplasia of the islet of Langerhans. Both beta- and alpha-cells are increased in number and size, however up to 90 % of the cells in the islets
are beta-cells. The size of the islets is related to the bodyweight, and the islet will decrease in size following caloric deprivation (Hellman, 1965; Hellman et al., 1961). Depending on the genetic background the associated phenotypes can differ (see C57BL/6J and A) (Coleman and Hummel, 1973). For example, the C57BL/KsJ strain (derived from an accidental genetic contamination of C57BL/6) is far more susceptible to diabetes than the original strain (Clee and Attie, 2007).

4.3.4. Leptin receptor deficient db/db mouse model (db/db mice)
Different to the ob/ob mouse model, these mice carry a mutation (db mutation) in the leptin receptor (LEPR) leading to the consequence that leptin cannot exert its function. Therefore, the phenotypical effects are identical to the Lep ob/ob and the severity of diabetes depends again on the genetic background.

4.3.5. ZDF rat model (fa/fa rats)
Due to a mutation in the Leptin receptor gene the Zucker obese (fa/.fa) rat is the equivalent rat model to the db/db mouse (Takaya et al., 1996). In addition to obesity the animals have hyperinsulinemia, insulin resistance glucose intolerance, and hyperlipidemia (Shafrir, 1992). On the other hand the obese Zucker diabetic fatty (ZDF) male rat is derived from inbreed hyperglycemic Zucker obese rats, and shows a hyperglycemia going along with insulin resistance, hyperlipidemia, and obesity, which makes the animal useful as a non-insulin dependent diabetes mellitus (NIDDM) model (Peterson et al., 1990). The development of hyperglycemia starts between 7 and 9 weeks of age accompanied with obesity and the onset of insulin resistance. Initially the plasma insulin levels remain normal, then elevate, and finally decrease to nearly undetectable levels by the age of 20 weeks. This course of disease is similar to the progression seen in NIDDM human patients (Johnson et al., 1990) and therefore the ZDF rat is a very good model to study the progression of disease. Since only male rats become hyperglycemic under a normal diet, the female obese rats are mostly used as controls. Nevertheless hyperglycemia can also be induced in female rats using a high fat diet containing 25.9 kcal% fat and higher (Corsetti et al., 2000).
4.4. Proteomics

4.4.1. Proteome-wide expression analysis

Differences in cellular states, such as the secretion of insulin upon glucose stimulation in the pancreatic beta-cell are reflected in changes in gene expression, in particular the levels of messenger ribonucleic acid (mRNA) and the levels of proteins as the final product. Measuring mRNA levels of cells kept under different conditions using microarray technology is a well established methodology. Even though microarrays are in principle able to cover the whole transcriptome their ubiquity and tremendous usefulness also have certain limitations. Some of the problems are technology related, as there are issues concerning reproducibility across platforms and laboratories (Canales et al., 2006). More importantly, microarray analysis does not reflect the actual abundance or the activity of the proteins. Since protein activity depends not only on the levels of the corresponding mRNA but also on post-translational modifications, degradation and localization, microarray technology does not reveal anything about the biological effectiveness of a mature protein (Gygi et al., 1999b).

Even though there has been remarkable progress in developing new mass spectrometric technology to investigate biological questions at a protein level, they still have one big drawback: proteins cannot be amplified as oligonucleotides can be. Therefore the sensitivity of the machines used and the methodology to process and enrich the proteins becomes essential. In order to make the proteins accessible for MS sequencing, protein samples get digested into peptides, which generates several tens of thousands of proteolytic peptides in a complex sample mixture. Due to technical limitations not all of these peptides can be fragmented and can be finally identified during one analysis run. Furthermore it is estimated that less than 10% of all peptides present in one run contribute to an useful peptide identification (Kuster et al., 2005). These and other limitations have made leaders in the field doubt the capability of standard MS-based “shotgun” proteomics in achieving a comprehensive proteome of a complex biological sample e.g. total cell lysate (Domon and Aebersold, 2006; Nielsen et al., 2006). Nevertheless mass spectrometry based proteomics is the method of choice to identify and accurately quantify proteins contained in complex sample mixtures (Aebersold and Mann, 2003).
4.4.2. Membrane proteomics

Recent progress in improved instrumentation, more advanced software tools for the analysis of proteomics data sets (de Godoy et al., 2006) and in depth sequencing strategies (Schmidt et al., 2008) enables the identification of hundreds to thousands of proteins in a single proteomics experiment of a complex protein mixture. However resolving complete proteomes of highly complex samples can still not be achieved. Therefore it is necessary to reduce sample complexity, wherefore scientists use various approaches to enrich low abundant proteins of interest. Fractionation is one method based on for example affinity procedures, which enables the isolation of distinct groups of proteins with similar features (Zhang et al., 2005). Application of such a concept include the selective isolation of cysteine-containing peptides (Gygi et al., 1999a), phosphorylated peptides (Bodenmiller et al., 2007), N-glycosylated peptides (Zhang et al., 2003), set of N-terminal peptides (McDonald et al., 2005), and specific sub-cellular fractions and organelles (Yates et al., 2005). Such procedures simplify the complexity of crude cell or tissue extracts, and thereby increase the probability to detect low abundant proteins as well as functional units of proteins. The sub-compartmental organization of eukaryotic cells provides a unique opportunity to link proteomic data with functional units. It was always important for biologists to identify and validate the presence of proteins in organelles to further understand their molecular function and mechanisms involved. Cellular membranes form an essential barrier in living cells to their external environments, and they serve to compartmentalize intracellular organelles within the eukaryotic cell. Depending on their localizations in the cell or even in the whole organism, these membranes have very different and specialized functions. Although nearly one third of all existing proteins are predicted to be embedded in biological membranes, membrane proteins are traditionally understudied due to difficulties with solubilization, separation, and identification. These difficulties have limited studies of surface proteins including adhesion molecules, channel transporter proteins, cell surface receptors, and enzymes, which are crucial for the sensing, induction and the catalysis in response to changes in the environment. The importance of plasma membrane proteins in sensing the cell’s environment, their major roles in diseases, and their potential as drug targets makes it obligatory to further investigate this class of proteins. With the recent advancement in technology to quantify N-glycosylated plasma membrane proteins (Schiess et al., 2009), new surface proteomes can be discovered and quantified upon different stimuli.
4.4.3. Quantitative proteome analysis

Since living cells are highly dynamic systems one cannot expect a static situation in which there will be an on/off situation of a particular protein. Therefore it is necessary to quantify the results, when studying biomarkers by comparing either different cell types or comparing different perturbations of the cellular system. For example one could compare the pattern of normal versus cancer tissue or investigate changes in protein patterns of a particular cell type upon different perturbations. In any setting it is not sufficient to look for on/off situations of protein abundance, but it is rather necessary to quantify the acquired data and finally create protein ratios for the different situations. To achieve this, antibodies have been traditionally employed, which limits the number of proteins being observed to a limited collection of molecules for which antibodies are available. For certain situations, this approach is still a powerful tool as in the fields of hematology, immunology, and pathology for research, diagnosis, and therapy, where antibodies are raised against cluster of differentiation (CD) proteins to investigate expression changes (Zola et al., 2005). However, not many cell types, nor diseases can be identified with the currently available set of CD antibodies. Due to this it is of great interest to increase the number of quantifiable proteins. The most promising approach to observe a “non limited” set of proteins is quantitative proteomics, which can be applied with different methodologies for quantification to cells and tissues.

Although proteomic technologies and methodologies are very advanced, the quantification of changes between two or more states of a complex cellular system is one of the most challenging tasks in proteomics. Due to the ionization differences and/or detectability of the intensity, a peak in a mass spectrum is not a very good indicator for the amount of peptide in the sample, due to the ionization differences and/or detectability of the intensity. Nevertheless, intensity differences in the peaks of one and the same peptide reflect differences in its abundance within the sample. Therefore, it is valid to use mass spectrometry based proteomics for relative quantification of proteins between two or more samples.
4.4.4. Labeling techniques for quantification

Most of the approaches in achieving proteomics quantification use labeling techniques. This is done using stable isotopes, either by metabolic labeling (SILAC) (Ong et al., 2002), chemical labeling (ICAT) (Gygi et al., 1999a) or labeling by covalent binding (iTRAQ) (Zieske, 2006). In either case the samples are labeled separately, combined and analyzed together. The major difference is the step, when the labeling of the sample takes place. In the case of ICAT and iTRAQ the labeling of the proteins happens after the cells are disrupted, whereas with the SILAC approach the proteins are labeled from the beginning of the experiment. This enables a combination of the samples from the very beginning of the processing, especially in cases where fractions are created (figure 3). This avoids bias, which might occur from differences in processing the single samples. Therefore SILAC is today the most popular method applied in quantitative proteomics. In this approach one batch of cells is grown on medium containing normal amino acids, whereas the other batch is grown in medium containing heavy stable isotopes labeled amino acids (non-radioactive). To reach a labeling efficiency of about 95% the cells have to be passaged 6 to 8 times before performing...
the experiment (Ong and Mann, 2006). The major advantage of this approach is that both cell populations can be combined directly after perturbation, proteins can be isolated and analyzed together by mass spectrometry. The identical peptide pairs with the different isotope compositions can be differentiated with the MS due to their mass difference. Finally the relative abundance of the two peptides can be calculated from the ratio of peak intensities in the mass spectrum, which finally reflects the corresponding protein abundance in the samples. The clear strength of all labeling approaches is its high accuracy. There are also limitations: Labeling techniques allow only the comparison of up to 3 different samples, the dynamic range is limited to 2 logs and for example SILAC can be normally only applied for cell culture, even though recently a SILAC mouse was used in a quantitative proteomics experiment (Kruger et al., 2008).

4.4.5. Label-free quantification

In contrast to the labeling based techniques, label-free approaches have the clear advantage of simple biochemical workflows, high dynamic range, multiple samples comparison, and wide proteome coverage. For label-free methods there are two ways to quantify two or more samples, of which first is the direct comparison of the mass spectrometric signal intensity derived from the peptide, and second the assignment of the number of acquired spectra for the peptide as an indicator for the respective abundance in the given sample. An alternative way of quantitative proteomics is the measurement of absolute peptide abundance by using synthetic reference peptides (discussed later).

When comparing proteomes using a label-free approach the processing, and the analyses of the different samples is done sequentially. Therefore it is important that this is executed in exactly the same way each time. These accurate requirements to maintain reproducibility can be achieved using standard operating procedures and automated systems. After MS analysis the intensities of the same peptide measured in two separate runs are compared to determine their relative abundance. Due to technical limitations not all peptides are selected for fragmentation within a complex mixture (Kuster et al., 2005). Therefore, it is critical to find and quantify the same peptide in different runs, even if it has been sequenced in only one run. One way to solve this problem is using high mass accuracy and high resolution mass spectrometers in combination with deep sequencing to generate a reference map with peptides of interest using the SuperHirn software (Mueller et al., 2007). This enables to quantify peptides in a focused way and saves machine running time. The clear disadvantage of a label-free approach is the multiple possible errors during sample processing. The overall
normalization of the single runs and spiking in reference peptides in the samples are ways to minimize these limitations (Meng et al., 2007). For peptide signals close to the noise, the quantitative accuracy is determined by the variation in the background and as such cannot be quantified.

4.4.6. Absolute quantification

Using synthetic heavy labeled peptides is the way to perform absolute quantification of peptide amounts in the sample. It starts becoming broadly used and is known as AQUA (absolute quantification of proteins) (Gerber et al., 2003). The simplest setting for such a measurement is to supplement a known quantity of a stable isotope labeled peptide to a protein digest and subsequently compare the signal with the endogenous peptide in the sample. This approach is attractive for studies addressing, for example the analysis, and validation of potential biomarkers in a large number of clinical samples (Pan et al., 2005).

One practical drawback of this approach is that one has to ‘guess’ how much of the labeled standard should be added to a sample. This amount may be different for all proteins of interest as their expression levels (used here in the sense of protein abundance rather than protein synthesis) may differ greatly within a sample. Another limitation is the specificity of the spiked standard as there are likely multiple isobaric peptides present in the mixture. Both of these issues can be addressed using the multiple reaction monitoring method (MRM), which uses a triple quadrupole MS to monitor both the intact peptide mass and one or more specific fragment ions during the LC-MS experiment. This method eliminates the uncertainty of the peptide assignment by using the combination of retention time, peptide mass, and fragment mass, and improves the quantification range to 4-5 orders of magnitude (Wolf-Yadlin et al., 2007). However, the critical step in an MRM experiment is the selection of the target proteins, and the corresponding synthetic peptides. The selection of the protein set can be based on previously performed experiments or publicly available information.

4.4.7. Proteomics studies with pancreatic islets and beta-cells

Many protein profiling efforts have been undertaken during the last decade using 2-dimensional gel electrophoresis (2DGE) to create protein maps which should reflect the protein expression pattern of a particular cell type or organ in a specific state. The first mouse islet proteome catalogue was published by Sanchez (Sanchez et al., 2001) and contained 44 unique proteins. Later others tried to elucidate more of the islet proteome with same approaches and similar coverage rates (Ahmed and Bergsten, 2005; Nicolls et al., 2003). This low coverage can be explained by the fact that the 2DGE technique preferentially detects
proteins of high abundance (Gygi et al., 2000). The most comprehensive islet proteome study till today was published by Metz (Metz et al., 2006) where he identified 3’665 proteins by analyzing human islets with a 2D LC/MS/MS approach. To overcome the complexity of the sample and get a higher coverage of the proteome one way is the sub-cellular fractionation of samples prior to mass spectrometric analysis. One approach applying this was the isolation of insulin granules from INS1-E cells followed by proteomics analysis (Brunner et al., 2007; Hickey et al., 2009). With this method Brunner et al unraveled 130 organelle specific proteins, which is a higher number compared to the first 2DGE approaches. Most of the identified proteins could be classified as vesicular or membrane bound, which definitely speaks for a more focused approach to gain deeper insights into the beta-cell proteome.

In the presented study, we explore the N-glyco surface of the pancreatic beta-cell line MIN6, the whole cell N-glyco-proteome of the same cell line, and of primary human islets. We quantify surface proteome changes in the MIN6 cells upon stimulation with glucose and in combination with GLP-1. The created surface catalogues in combination with public available data were undertaken bioinformatics analysis to predict potential biomarker for the progression of type 2 diabetes (T2D). These candidate biomarkers were tested using an MRM approach in the plasma of two different mouse models (DIO and ob/ob) for T2D over the progression of disease. In addition a shotgun proteomics approach was used to investigate global, non beta-cell specific changes in the same samples.
5. Results

The aim of this study was to perform a quantitative analysis of the beta-cell’s surface proteome under different stimuli. We combined the selective enrichment of cell surface N-glycopeptides with a label-free, quantitative proteomics method that provides the sensitivity and the sample throughput needed to analyze multiple samples. To demonstrate the sensitivity and specificity of the applied technology, we generated a murine beta-cell (MIN6) N-glycoprotein catalogue applying Cell Surface Capturing (CSC). This catalogue was complemented by Whole Cell Capturing (WCC) of the murine beta-cells and human pancreatic islets (figure 4).

For quantification experiments we first tested the reproducibility of the technology and the biological variation. Changes in the cell surface N-glyco-proteome induced by specific perturbations were monitored in a quantitative way relative to the corresponding controls.

As proof for a successful stimulation and functional reaction of the cells, the increase of insulin secretion upon stimulation with either glucose and GLP-1 or glucose only was measured. To validate our approach, the well established internalization of the GLP-1R upon the stimulation of beta-cells with glucose and the GLP-1 peptide was monitored.

5.1. MIN6 Cell Surface Capturing catalogue

In a first discovery approach we identified the N-glycosite surface proteome from the murine beta-cell line MIN6 to generate a reference map for further comparative analyses by label-free quantitative MS.

The N-glycosites were identified by LC-MS/MS after their isolation via the CSC method that is based on selective affinity labeling and solid-phase capturing of N-glycosylated cell surface peptides (Wollscheid and al., 2009). The data were stored in the “Meta database of beta-cells in diabetes” and are represented in a LC-MS (retention time (RT) vs. m/z) feature map (biodata.ethz.ch) in which the identified N-glycosites were annotated with their amino acid sequence.
Results

Figure 4. Schematic drawing of selective enrichment for N-glyco proteins from murine MIN6 cells and human islets. The Cell Surface Capturing (CSC) was only applied to the MIN6 cells, whereas the Whole Cell Capturing (WCC) was used for both MIN6 cells and human pancreatic islets. The N-glycosites were enzymatically released and analyzed with a mass spectrometer. The spectra were searched against the database mouse/human IPI protein database using SEQUEST (Keller et al., 2002b). Combined with deep sequencing using SuperHirn (Mueller et al., 2007) we were able to generate three catalogues. To annotate a protein as high confident, it requires identification with at least two N-glyco peptides and an overall ProteinProphet (Nesvizhskii et al., 2003) probability score of higher than 0.9.

To generate the MIN6 CSC catalogue we combined the results of several experiments in which we enriched cell surface N-glycopeptides from MIN6 cells and performed LC-MS/MS analysis. The fragment ion spectra acquired from all performed LC-MS/MS runs were searched against the mouse IPI protein database (v3.26, 66'440 entries). A total of 2’063 MS/MS spectra were assigned at a PeptideProphet (Keller et al., 2002a) probability threshold of ≥ 0.9 to peptide sequences of which 85% (1’757) matched to peptides containing the N-X-S/T motif, indicating the presence of a glycan at that site in the intact protein. All these peptides could be positively assigned as true N-glyco peptides due to the mass shift caused by the enzymatic deamidation at the site of the glycan attachment (mass difference: 0.98604 Da). These results demonstrate the high specificity of the method for N-glycosites. Overall, the
assigned spectra represented 1'729 unique N-glycosites (PeptideProphet probability score ≥ 0.9, FDR 1%) matching to 259 unique N-glyco proteins. According to SOSUI (Hirokawa et al., 1998) 80.5% (206 proteins) of the identified cell surface N-glycoproteins contained at least one or more transmembrane domain (TMD). 3.9% (10 proteins) of all proteins in the MIN6 Cell Surface Catalogue were predicted to be attached to the membrane via a GPI anchor (Fankhauser and Maser, 2005). Proteins holding neither a TMD nor a GPI anchor are annotated as either secreted (6.6%; 17 proteins) or mitochondrial (1.6%; 4 proteins).

Our analysis show that 88.5% (230 proteins) of the MIN6 CSC catalogue proteins match the GO15 (gene ontology) category “membrane” (GO:0016020), 52.3% (136 proteins) match the category “plasma membrane” (GO:0005886). This shows a high specificity of the enrichment method. Only 2 proteins are annotated as “cytosolic” (GO:0005829). However, these two intracellular proteins were found in several independent experiments and cannot be excluded as being contaminants.

Receptor activity (GO:0004872) was annotated for 71 proteins of the catalogue, whereas 34 were found to have a transmembrane receptor activity (GO:0004888). Within this category 7 G-protein coupled receptors could be mapped, which are generally known to be of low abundance (Sarramegna et al., 2003). One of these receptors is the GLP-1R, which further allowed us to relatively quantify the receptor abundance at the cell surface after GLP-1 stimulation.

5.2. MIN6 Whole Cell Capturing catalogue

To determine the specificity for the enrichment for exterior surface peptides, we lysed MIN6 cells according to the WCC procedure. The follow up analysis were performed as described prior in the CSC catalogue section. The data were stored in the “Meta database of beta-cells in diabetes” and are represented in a LC-MS (retention time (RT) vs. m/z) feature map in which the identified N-glycosites were annotated with their amino acid sequence. To generate the MIN6 WCC catalogue we combined the results of all performed experiments in which we enriched N-glycopeptides from MIN6 cell lysates and performed LC-MS/MS analysis.
Results

Figure 5. The catalogue comparison shows an overlap between all three proteomes of 104 proteins, whereas MIN6 CSC and MIN6 WCC only have 83 proteins in common. Comparing the human islet N-glyco proteome to the MIN6 CSC 16 are in common, whereas 91 proteins overlap in comparison with the MIN6 WCC. (orange: MIN6 CSC; red: MIN6 WCC; blue: human islets WCC).

Compared to the MIN6 CSC catalogue (259) we overall identified more proteins in the WCC catalogue (489) (figure 5). The total number of TMD containing proteins is higher in the CSC approach compared to the WCC (82.0% vs. 71.6%), which speaks for the surface specificity of the CSC method. The number of proteins annotated as being secreted is 14.0% instead of 6.0% when compared with the CSC approach, whereas the identified GPI-anchored proteins remain nearly the same in both cases (figure 6).

5.3. Human pancreatic islet Whole Cell Capturing catalogue

Due to the low amount of available material, the human pancreatic islets from 2 different patients were processed according to the WCC protocol. The follow up analysis was performed as described in the WCC catalogue section for MIN6 cells. The data are stored in the “Meta database of beta-cells in diabetes” and are represented in a LC-MS (retention time
Results

(RT) vs. m/z) feature map, where the identified N-glycosites are annotated with their amino acid sequence. To generate the Hs ISLET WCC catalogue we combined the results derived from two different patient samples in which we enriched N-glycopeptides from human islets lysates and performed LC-MS/MS analysis.

Overall, compared to the MIN6 CSC, we identified more proteins in the Hs ISLET WCC catalogue, though the total number of TMD containing proteins is higher by the CSC approach. The absolute number of proteins annotated as being secreted is more than triple compared to the CSC approach, while the identified GPI-anchored proteins remain the same in number.

5.4. Islet/beta-cell N-glyco protein catalogue

Combination of the three catalogues results in a total of 743 unique N-glyco surface proteins. Gene ontology (GO) analysis revealed amongst others 22 proteins with an annotation for kinase activity, 14 proteins with a G-protein coupled receptor activity (GPCR) and 59 proteins with a peptidase activity, whereas 29 had an annotation for metallo peptidase activity (Rawlings et al., 2008). The epidermal growth factor receptor (EGFR) was the only protein found in all generated catalogues exhibiting kinase activity. In the GPCR category the G protein-coupled receptor 56 (GPR56) is the only protein appearing in all catalogues. In the metallo-peptidase category 6 proteins are annotated in the different catalogues, namely ADAM10, CPD, CPD, CPM, ENPEP, LNPEP.

In all catalogues derived either from MIN6 cells or human islets there is a clear enrichment for proteins predicted to have one or two transmembrane domains (TMD) using the SOSUI (Hirokawa et al., 1998). About 10% of all proteins have a prediction for “0” TMDs, this number is smaller in the WCC of both human and mouse material compared to the CSC. The number of proteins predicted to be secreted is also increased in the WCC samples compared to the CSC. Whereas the number of predicted GPI anchored proteins is roughly the same in all three catalogues. The GO analysis showed a high number of typical surface proteins types like receptors, transporter, peptidases and calcium ion binding, but also low abundant protein classes like GPCRs and ion channels (figure 6).
Results

Figure 6. Analysis of the generated surface proteomes for TMD domains, Gene ontology (GO) annotation (Harris et al., 2004) and KEGG (Kanehisa et al., 2008) pathway enrichment show the surface specificity of the applied approach. SOSUI (Hirokawa et al., 1998) and GPI analysis (Fankhauser and Maser, 2005) of the MIN6 CSC, the MIN6 WCC and the human islets WCC. Predictions could be made for 256 (out of 259) proteins annotated in the MIN6 CSC, whereas 216 (84.4%) contain one or more than one TMD or a GPI-anchor. For 486 (out of 489) proteins in the MIN6 WCC a prediction was made, whereas for 347 (71.4%) of these either one or more than one TMD or a GPI-anchor was found. In the human islet WCC catalogue 391 (out of 393) were predicted, whereas 260 (66.5%) of these predicted proteins have one or more than one TMD or a GPI-anchor. GO cellular component analysis of the identified proteins show that 88.5% belonged to membrane (GO:0016020), 52.3% to plasma membrane (GO:0005886), 0.8% to cytosol (GO:0005829), 6.9% to cytoplasmic vesicle (GO:0031410), 6.2% to endoplasmic reticulum (GO:0005783), and 4.6% to golgi apparatus (GO:0005794). The GO analysis for the MIN6 WCC gave 73.0% annotated to membrane (GO:0016020), 31.3% to plasma membrane (GO:0005886), 1.0% to cytosol (GO:0005829), 5.9% to cytoplasmic vesicle (GO:0031410), 16.0% to endoplasmic reticulum (GO:0005783), and 8.6% to golgi apparatus (GO:0005794). The human islet WCC contained 62.7% GO annotation, 15% for membrane (GO:0016020), 29.9% for plasma membrane (GO:0005886), 0.8% for cytosol (GO:0005829), 7.1% for cytoplasmic vesicle (GO:0031410), 12.9% for endoplasmic reticulum (GO:0005783), and 4.3% for golgi apparatus (GO:0005794). We further checked the catalogues for the abundance of receptor activity (GO:0004872), calcium ion binding (GO:0005509), peptidase activity (GO:0008233), transporter activity (GO:0005215), TMD receptor activity (GO:0004888), kinase activity (GO:0016301), ion channel activity (GO:0005216), and GPCR activity (GO:0004930). In detail the kinase activity, the G-protein coupled receptors, and the metallo-peptidases are listed according to their abundance in the catalogues.
5.5. Technical and biological variation

The technical variation of the experimental setting was tested in MIN6 cells cultured in normal culturing medium containing 25 mM glucose. The cell surface labeling and N-glyco peptide enrichment was performed as described in the material and methods section. Three different runs of the same sample were analyzed with the MS and the outcomes were plotted against each other. The $R^2$ showed variations between 0.9741 (sample 1 vs. 2) and 0.9909 (sample 2 vs. 3) (figure 7A).

The biological variation was determined in three independent experiments (3 different passages), where MIN6 cells were cultured in normal culturing medium containing 25 mM glucose. The cell surface labeling and the N-glyco peptide enrichment was performed as described in the material and methods section. Three different runs of a single experiment were analyzed with FT-LC-MS/MS, and from three independent biological experiments the averages were plotted against each other. The $R^2$ showed variations between 0.8713 (experiment 2 vs. 3) and 0.9641 (sample 1 vs. 2) (figure 7B).

5.6. Quantitative CSC of MIN6 cells stimulated with 25 mM D-glucose

The MIN6 cells were grown to 80% confluency, starved for 4 hours in DMEM medium containing 2.5 mM D-glucose and stimulated with DMEM full medium containing 25mM D-glucose for one hour.

The following processing was analogous to the one described for the MIN6 Cell Surface Capturing catalogue. The label free quantification was performed according to the description in the methods section.

After one hour of stimulation with 25 mM D-glucose 15 surface proteins were found to be significantly regulated ($\leq$0.20 log2 ratio, p>0.05). The measured regulations are in the range of 0.30-fold up-regulation and 0.40-fold down-regulation.
Results

Figure 7. Technical and biological variation: (a) The technical variation was tested with three independent mass spectrometer runs measuring the same N-glyco peptide preparation. The single runs are plotted against each other and resulted in the following values: (■) 1 vs. 2: \( R^2 = 0.9741 \); (●) 1 vs. 3: \( R^2 = 0.9805 \); (▲) 2 vs. 3: \( R^2 = 0.9909 \). (d) The biological variance was defined with three independent N-glyco peptide preparations of three different passages of MIN6 cells under normal culturing conditions without treatment. The single experiments are plotted against each other and resulted in the following values: (red) 1 vs. 2: \( R^2 = 0.9641 \); (green) 1 vs. 3: \( R^2 = 0.9391 \); (blue) 2 vs. 3: \( R^2 = 0.8713 \).

None of the proteins found being regulated are known to be internalized or more abundant at the cell surface upon the applied stimulus. Due to methodological limitations it was not possible to detect GLUT2 (SLC2A2) at the beta-cell surface, which is known to be up-regulated upon glucose stimulation (Bonny et al., 1995; Waeber et al., 1996). To ensure successful stimulation we only processed samples, which showed a more than 0.20-fold increase of insulin secretion after 1 hour of stimulation.
Results

Following are the regulated proteins, described according to their regulation beginning with the most up-regulated one:

- The integrin alpha chain family protein ITGA2 (+0.31, p=0.007) was found to be highly regulated in the stimulated state compared to the unstimulated state. This class of proteins is expressed in a wide variety of cell types. Being an integrin, ITGA2 is involved in cell adhesion, but also in cell-surface mediated signaling events, still there is nothing known about its regulation in beta-cells.

- The low density lipoprotein receptor LDLR (+0.30, p=0.007) is known to be responsible for cholesterol up-take into the cell, but not in connection with glucose stimulation. LDL is pro apoptotic, whereas HDL is protective (Cnop et al., 2002; Roehrich et al., 2003).

- Nectin-1 or poliovirus receptor-related 1 PVRL1 (+0.29, p=0.049) and its family members are mostly studied as cell-cell adhesion protein, but recently thought to be involved in proliferation (Takai et al., 2008).

- The Taurine transporter SLC6A6 (+0.25, p=0.014) is not known to be regulated by glucose stimulation, but systemic application of Taurine itself improves glucose homeostasis and islet function in mice (Carneiro et al., 2008).

- The transferrin receptor TRFC (+0.25, p=0.025) is a Wnt (Rohrs et al., 2009) target gene and expressed as an immediate response (below 1 hour) after stimulation with WNT-3a.

- The IGF2R (+0.24; p=0.007) binds both insulin-like growth factor 2 (IGF2) and mannose 6-phosphate. The receptor itself functions in intracellular trafficking of lysosomal enzymes, the activation of transforming growth factor beta, and the degradation of IGF2.

- The protein tyrosine phosphatase PTPRN (+0.23, p=0.025) is a well known autoantigen in type 1 diabetes (Notkins et al., 1998). Synergistic effects of glucose and growth hormone signaling in islet cells is thought to be mediated by PTPRN and STAT5 (Mziaut et al., 2006).

- The granule marker and tetraspanin family member CD63 (+0.22, p=0.010) is regulated in the same way as with GLP-1 plus D-glucose stimulation.

- The transmembrane protein TMEM16f (+0.22, p=0.005) is a Ca^{2+} regulated Cl⁻ secretory channel protein (Rock et al., 2009), which is not further characterized.
Results

- ENPP4 (+0.21, p=0.042) has a putative function of being an ectonucleotide pyrophosphatase/phosphodiesterase, but this is not further investigated.

- Interleukin-1 is cytotoxic to pancreatic beta-cell (Bergholdt et al., 2000) and we found its signal transducing receptor ILR1R (-0.23, p=0.031) internalized upon glucose stimulation. After binding of interleukin-1, the receptor is internalized mediated by presenilin-1 (Elzinga et al., 2009a). Though it is thought that glucose and endoplasmic reticulum calcium channels regulate HIF-1 beta via presenilin in pancreatic beta-cells (Dror et al., 2008).

- The mouse specific MHC I antigen was found less abundant at the cell surface (-0.24, p=0.042).

- The aminopeptidase A (APA) ectopeptidase ENPEP (-0.28, p=0.025) is an integral membrane-bound zinc metalloprotease that cleaves aspartic and glutamic acidic residues from the N-terminus of a number of protein substrates (Tonna et al., 2008).

- The carcinoma associated antigen TACSTD1 (-0.30, p=0.025) is decreased in surface abundance after 1 hour of stimulation with D-glucose.

- The tetraspanin family protein TSPAN15 (-0.41, p=0.010), most of these families members are cell-surface proteins that are characterized by the presence of four hydrophobic domains. The proteins mediate signal transduction events that play a role in the regulation of cell development, activation, growth and motility.
Results

Figure 8. After 1 hour stimulation of MIN6 cells with 25 mM D-glucose 15 proteins showed a change in surface abundance after applying the filtering rules. The relative change ranged from +0.31 to -0.41 log2 ratio, whereas 10 were up-regulated and 5 down-regulated. Red (control), blue (stimulation with 25 mM D-glucose). All values indicated show a deregulation of <0.20 log2 ratio, and have a p>0.05.

5.7. Quantitative CSC of MIN6 cells stimulated with 25 mM D-glucose plus 20 nM GLP-1

The MIN6 cells were grown to 80% confluency, starved for 4 hours in DMEM medium containing 2.5 mM D-glucose and stimulated with DMEM full medium containing 25 mM D-glucose supplemented with 20 nM GLP-1 for one hour. The following processing was analogue to the one described for the MIN6 Cell Surface Capturing catalogue. The label free quantification was performed according to the description in the methods section. After one hour of stimulation with 25 mM D-glucose plus 20 nM GLP-1, 24 surface proteins were found to be significantly regulated (<0.20 log2 ratio, p>0.05). The regulation is in the range of 0.32-fold log2 ratio up-regulation and 0.58-fold log2 ratio down-regulation. None of the proteins found being regulated is known to be internalized or more abundant at the cell surface upon the applied stimulus. Besides this automated monitoring of surface protein abundance changes, we found the GLP-1R receptor internalized, seen with two N-glyco
Results

peptides reducing their abundance of 0.90-fold log2 ratio in one of our experiments (figure 10).

The following proteins were changed upon glucose plus GLP-1 stimulation:

- The Zn\(^{2+}\) dependent amino peptidase LNPEP was found being +0.32 fold higher in its surface abundance after stimulation. The protein is known to reside in intracellular vesicles with the insulin-responsive glucose transporter GLUT4 and is responsive to insulin whereon it redistributes to the cell surface (Keller, 2004). GLUT4 itself is known to be expressed in pancreatic islets of humans and rats (Kobayashi et al., 2004), but could not be found with our approach.

- The abundance of the protein tyrosine phosphatase PTPRN2 was (+0.30, p=0.040) is known to reduce glucose stimulated insulin secretion when over-expressed in MIN6 cells (Doi et al., 2006). A higher abundance could also be observed after one hour of stimulation with 25 mM D-glucose only, even though not significantly (+0.20, p=0.12), but the raw data show a clear trend to be up-regulated.

- The abundance of the glucose transporter SLC2A1 (GLUT1) increased upon stimulation (+0.27, p=0.002). GLUT1 is expressed in most cell types (an exception being normal beta-cells) and represents an established HIF-target gene product. It is up-regulated in many cancer cells to facilitate glucose uptake to meet the increased energy demand associated with high rates of growth and proliferation (Shaw, 2006). In pancreatic beta-cells GLUT2 is thought to be the major glucose transporter, which we could not measure due to methodological issues.

Islets lacking GLUT2 lose their first phase of glucose stimulated insulin secretion, but not the second phase. This is thought to be a compensatory effect of GLUT1/3, whereas its overall protein abundance does not change (Guillam et al., 2000). We could show GLUT2 mRNA expression in the MIN6 cell lines we used, but could not detect GLUT2 neither in western blot nor in immunofluorescence (data not shown).

The tetraspanin (TSPAN) family member CD63 (+0.26, p=0.003) is thought to be involved in neutrophils in endoplasmic reticulum or golgi export, cellular retention, and granule targeting of proNE before storage as mature NE (Kallquist et al., 2008). TSPAN family members, like TSPAN15 (+0.22 fold, p=0.010) and TSPAN3 (+0.21 fold, p=0.010) were also found to have a higher surface abundance after stimulation.
• The serine protease ST14 (+0.26, p=0.010) is thought to regulate Na⁺ channels (Planes and Caughey, 2007) and acts as an epithelial membrane activator for other proteases and latent growth factors.

• The ecto-nucleotide pyrophosphatase/phosphodiesterase ENPP1 (+0.20, p= 0.005) is known to be expressed in pancreatic islets (Goding et al., 2003), inhibits insulin-receptor (IR) signaling and, when over-expressed, induces insulin resistance in vitro and in vivo (Marucci et al., 2009).

• LTBR the lymphotoxin beta receptor (TNFR superfamily, member 3) was found to be up-regulated +0.23 (p=0.034) in our experiments. The down-regulation of LTBR controlled migration factors within the pancreatic islets prevents insulin-dependent diabetes mellitus (Lee et al., 2006).

• The netrin-1 family receptor UNC5A (+0.23, p= 0.010) is thought to be involved in apoptosis, though the function of protein up-regulation is unclear, therefore it could be either pro- or anti-apoptotic (Arakawa, 2004; Paradisi et al., 2008; Williams et al., 2006).

• The sidekick homologue SDK1 is reduced in surface abundance after stimulation (-0.21, p=0.009). There are not many functions known for SDK1, aside from being involved in cell adhesion and directing lamina-specific connections at synapsis (Yamagata and Sanes, 2008).

• The receptor protein NEO1 (-0.21,p=0.022) was shown to have a low mRNA expression in mouse pancreas (Rodriguez et al., 2007). Relative to other gene expression levels we tested for, there was little of NEO1 mRNA being expressed in the MIN6 cell line used (data not shown).

• The serine protease family gene TMPRSS4 (-0.25, p=0.010) was identified as a gene being over-expressed in pancreatic carcinoma and its up-regulation drives tumorogenesis (Dawelbait et al., 2007).

• The Zn²⁺ Transporter SLC39A14 (-0.26, p=0.013) drives the Zn²⁺ influx (Taylor et al., 2005) and is activated in early stages of adipogenesis (Tominaga et al., 2005).

• SLC1A4 (-0.33; p=0.017) is a glutamate/neutral amino acid transporter and mediates co-transport of small neutral amino acids such as Ala, Ser, Cys and Thr (Kanai and Hediger, 2004; Pinilla et al., 2001).

• Another solute carrier family protein is SLC5A3 (-0.37, p=0.010), which is a sodium/myo-inositol co-transporter, when being reduced by TNF alpha at the
transcription level, the cell accumulates myo-inositol, but reduces its uptake (Yorek et al., 1999).

- The type I integral transmembrane protein CD164 (-0.27, p=0.010) functions as an adhesion receptor (Forde et al., 2007; Watt et al., 1998), and is involved in the biosynthetic trafficking of apical proteins in the endocytic compartments (Cresawn et al., 2007).

- The cadherin superfamiy member FAT1 (-0.41, p=0.010) contains 34 tandem cadherin-type repeats, five epidermal growth factor (EGF)-like repeats and one laminin A-G domain. Its classical function is to regulate cell proliferation and planar cell polarity, but also actin dynamics and control of cell-cell interactions (Tanoue and Takeichi, 2005).

- The serine protease inhibitor SPINT2 (-0.42, p=0.017) is a Na\(^{2+}\) driven peptidase and is thought to inhibit cell growth through the apoptotic pathway (Nakamura et al., 2009).

- Cell adhesion molecule 1 CADM1 (-0.42, p>0.001) is a novel pancreatic-islet cell adhesion molecule mediating nerve-islet cell interactions. CADM1 seems to be involved in hormone secretion from islet cell tumors.

- CD24a (-0.49, p>0.001) is expressed in the developing pancreas, but not in the mature beta-cells, which led to the hypothesis of a marker for precursor endocrine cells (Cram et al., 1999). It’s expression is high in pancreatic adenocarcinoma (Sagiv et al., 2006).

- The antagonism of the G-protein coupled receptor ADRA2A (-0.58, p=0.005) increases insulin secretion and synergistically augments the insulinotropic effect of glibenclamide in mice (Fagerholm et al., 2008). The ADRA2A knockout mouse shows reduced blood glucose levels, increased insulin levels and improved glucose tolerance (Savontaus et al., 2008).
After the stimulation of MIN6 cells with 20 nM GLP-1 plus 25 mM D-glucose for one hour, 24 proteins were found to be regulated after applying the filtering rules. The relative change ranged from +0.32 to -0.58 log2 ratio, whereas 11 were up-regulated and 13 down-regulated. Red (control), blue (stimulation with 20 nM GLP-1 plus 25 mM D-glucose). All values indicated show a deregulation of <0.20 log2 ratio, and have a p>0.05.

The GLP-1 receptor (GLP1R) was found to be reduced in abundance only in a single experiment, in all other experiments with glucose plus GLP-1 stimulation we were not able to detect the peptides identifying the receptor. Therefore the regulation of the GLP1R does not show up in the list of regulated proteins, since the significance of change is below the threshold (p≤0.05). In this single experiment we found two distinct peptides which identify the receptor. Both of these peptides are regulated the same way and show a 0.90-fold log2 ratio down-regulation of the receptor at the surface. This down-regulation could not be...
observed on mRNA level, since we did not measure any expression changes of the GLP1R by Q-PCR. In addition some of the surface proteins known to be involved in GLP-1 signaling in beta-cells (Holz and Chepurny, 2005) could be quantified in this experiment as well. Most of them were found not to be regulated (DPP-4, ADAM10, EGFR, IGFR, INSR) in their abundance at the surface. The only one found to be up-regulated was an adenylate cyclase (ADCY9).

**Figure 10.** After 1 hour stimulation of MIN6 cells with 25 mM D-glucose plus 20 nM GLP-1 only the GLP-1R was found to be down-regulated 0.90-fold log2 ratio (A). This could be shown with two independent peptides matching to the GLP-1R (B). No regulation of GLP-1R was found on mRNA levels after 1 hour of stimulation (C). Green arrows indicate an up-regulation, orange arrows indicate no regulation, and red arrows indicate a down-regulation measured by label-free quantification.

### 5.8. Beta-cell biomarker studies

The N-glyco protein catalogues created by the experiments described before were used to define a set of candidate biomarkers. This list of candidate biomarkers was screened for known functions with a strong focus on beta-cells, and gene expression. Proteins which were found to be expressed ubiquitously on a high level or in large organs like the liver were excluded, since the chance to measure a change in plasma abundance of such proteins caused by beta-cells is highly unlikely. In addition we used the Peptide Atlas (Deutsch et al., 2008) as
an additional information source to find peptides which are in our candidate biomarker list and were already measured in plasma before. At the end we came up with a list of 317 synthetic peptides for potential biomarkers. These were spiked into the processed plasma samples derived from the diet induced obesity mouse models and the ob/ob mice (see below). The defined amounts of synthetic peptides in the samples allowed us to perform MS-based absolute quantification of the abundance of these particular proteins in the samples.

**Figure 11.** Workflow how the potential biomarkers were defined using the N-glyco catalogue as a kickoff, combining it with information from public available data sources like Peptide Atlas (Deutsch et al., 2008), gene expression data and known functions of the protein candidates. This defined the list of AQUA peptides, which were used for biomarker validation in plasma samples of mouse models of type 2 diabetes.

### 5.9. Diet induced obesity (DIO) mouse model blood analysis

One of the standard mouse models for type 2 diabetes (T2D) is the diet induced obesity mouse model (Clee and Attie, 2007), which uses the effect of high caloric food to induce a progression of T2D disease over the course of feeding. In a first phase, we aged the mice on normal rodent chow till they reached 5 weeks of age. Then we split the animals in two groups: one was fed with normal chow and the other with a high fat diet. During the course of the experiment, we sacrificed 5 animals per group after 2, 4, 6, 12 and 16 weeks after the start of feeding (figure 11). At these time points the animals were starved overnight, then weighed (figure 12 A), and blood plasma was sampled according to the SOP defined in the materials.
and methods section. The pancreas was removed, processed and embedded in paraffin blocks for preservation. Liver, kidney, spleen, brain, brown and white fat were preceded the same. For each of these organs a part was removed and cryopreserved. This organ database was build to have the chance to further validate data derived from the plasma biomarker studies.

**Figure 12.** Experimental layout of the DIO mouse plasma experiment showing the two phases of the experiment, and the two different groups of C57BL/6J animals fed with the according diet. At each time-point 5 animals of each group were sampled.

The processed plasma samples were stored in 55 µL aliquots at -80°C till the end of the experiment. Two aliquots of each individual were handed over to F. Hoffmann-La Roche to measure physiological plasma parameters.

The bodyweight of the animals was already significantly different in comparison to the controls after 2 weeks of high fat diet (HFD) feeding (figure 12A). However, the plasma glucose levels were similar until the 16 weeks time-point, at which the plasma glucose levels were double in the high fat diet group (figure 12B). Insulin levels increased in this group during the course of the experiment but never significantly, probably due to large variations in the samples (figure 12C). Total cholesterol was elevated in the high fat diet group over the whole course of the experiment and only slightly increased over time (figure 12D). The triglyceride levels dropped from the 2 weeks time point to 4 weeks same in both groups (figure 12E). They remained similar in the two groups, even though they were significantly higher (p ≤0.05) for the controls at 6 and 16 weeks. In the HFD diet group this parameter was higher at 4 weeks of feeding. The fructosamine levels in the HFD group were significantly elevated compared to the controls. Frustosamine levels reflect the long term blood glucose levels over a period of about 6 weeks, similar as the better known HbA1c. Approximately
each change of 3.3 mmol in average blood glucose levels will give rise to changes of 2% HbA1c or 75 µmol fructosamine values (Nathan et al., 1984).

To additionally prove the induction of type 2 diabetes, we performed an IPGTT after 12 weeks with the animals, which were sampled at the 16 weeks time-point (figure 13). This test showed a significant delay in glucose clearance by the HFD group compared to the controls. At the start value (t=-60 min) the blood glucose levels were significantly higher in the HFD, and became non significant at (t=0 min). 30 minutes after intra-peritoneal glucose injection the glucose levels of the HFD group were significantly higher. This difference becomes even more obvious when calculating the area under curve (figure 13). From these results we assume a state of T2D disease for the animals sampled at the 16 weeks time point. For the HFD group this includes a pathological increased amount of islets, which are not fully functional and a mild peripheral insulin resistance as shown by others (Rebuffe-Scrive et al., 1993; Surwit et al., 1988).
Figure 13. Physiological parameters determined of the overnight starved animals during the course of the experiment. (A) The bodyweight increased dramatically in the high fat diet (HFD) group. (B) The plasma glucose levels stayed very similar until the 16 weeks time-point, where a significant difference could be found. (C) The insulin levels increased over the course of the experiment slowly and showed a large deviation after 16 weeks. (D) The total cholesterol was statistically higher from the very beginning in the HFD group and remained basically the same. (E) The triglyceride levels dropped at 4 weeks and stayed very similar in both groups. (F) The fructosamine increased at 4 weeks in both groups but showed a significant difference between the two groups at later time-points. Normal Chow Group (drawn line); HFD Group (dotted line); * $p \leq 0.001$, ** $p \leq 0.05$
Results

Figure 14. IPGTT after 12 weeks of feeding with the 16 weeks time point group, which was sampled 4 weeks later. The curves and the calculated area under curve (AUC) show significant differences between the control (drawn line) and the HFD group (dotted line), also in the area under curve. * p ≤ 0.001; ** p ≤ 0.05

5.10. ob/ob mouse model blood analysis

As a second mouse model to validate our biomarker candidates, we have chosen the ob/ob model. These animals were bred in our facility and aged according to the course of the experiment (figure 14). At the age of 4, 8, 12, 16, 20 and 24 weeks 5 control animals (wt/wt or ob/wt) as well as 5 ob/ob animals were sacrificed, blood plasma was prepared according to the SOP, and stored in 55 µL aliquots at -80°C. Prior to the blood sampling the animals underwent food deprivation and weight was documented as displayed in figure 15A and 16A. The bodyweight of the ob/ob animals increased with age and was statistically significant higher compared to the wt controls beginning with 8 weeks of age in both genders (15A and 16A). The organs were harvested (as described for the DIO model in 3.8.) and either stored at -80°C or embedded in paraffin blocks.
Results

Figure 15. Experimental layout of the ob/ob mouse plasma experiment showing the course of the experiment, and the two different groups of ob/ob and control animals, both fed with standard diet. At each time point 5 animals of each group were sampled.

Since male C57BL/6 mice are more susceptible to high fat diet, we only used male mice in the DIO study, whereas with the ob/ob model we used both genders to also monitor potential biomarker changes, which might be caused due to sex-specific differences. The plasma glucose levels were elevated in the ob/ob compared to the controls in both gender starting from time point 8 weeks (figure 15B and 16B). The plasma insulin levels of ob/ob animals increased significantly with age in the ob/ob animals (both genders) compared to controls and dropped at the age of 24 weeks in male animals, more than in females (figure 15C and 16C). Total cholesterol was significantly doubled in the ob/ob females beginning with the age of 12 weeks, and remains like this till 24 weeks (figure 15D). For the male ob/ob animals this is more pronounced, in which the total amounts of cholesterols was increased significantly depending on the age and was more than double at the age of 24 weeks (figure 16D). The ob/ob female animals showed significantly higher triglyceride levels only at the age of 8 weeks, but at other ages they remained very similar (figure 15E). The ob/ob males had a significant increase in triglyceride levels at 12, 16, 20 weeks, which became non significant at 24 weeks of age (figure 16E). The fructosamines were significantly higher in ob/ob groups in both genders, beginning at 8 weeks of age for ob/ob females and 12 weeks for ob/ob males (figure 15F and 16F).

With the parameters we measured, we conclude that our animals show a T2D phenotype as has been described before for this ob/ob strain (Lord and Atkins, 1985).
Results

Figure 16: Physiological parameters of the overnight starved female animals determined during the course of the experiment. (A) The bodyweight increased dramatically in the female ob/ob group compared to controls. (B) The plasma glucose levels stayed significantly higher in the female ob/ob group compared to controls from the 8 weeks time point on. (C) The insulin levels increased significant over the course of the experiment in the female ob/ob group. (D) The total cholesterol stayed the same in the control group, but was significantly elevated over time in the female ob/ob group compared to controls. (E) The triglyceride levels were slightly higher in the ob/ob group, and did only statistically significant differ from the controls at 8 weeks. (F) The fructosamine levels are increased after 8 weeks of age in the ob/ob group, and showed significant difference at the age of 24 weeks. Control group (drawn line), ob/ob group (dotted line). * p ≤ 0.001; ** p ≤ 0.02
Results

**Figure 17.** Physiological parameters of the overnight starved male animals determined during the course of the experiment. (A) The bodyweight increased significantly in the male ob/ob group compared to controls. (B) The plasma glucose levels stayed significantly higher in the male ob/ob group compared to controls from the 8 weeks onwards. (C) The insulin levels increased significantly over the course of the experiment in the male ob/ob group and dropped slightly at the age of 24 weeks. (D) The total cholesterol was significantly increased from the beginning in the ob/ob group compared to the control group and further increased over time. (E) The triglyceride levels were slightly higher in the ob/ob group, which became significantly different after 12 weeks. (F) The fructosamine levels are significantly increased after 12 weeks of age in the ob/ob group and remain higher compared to the controls. Control Group (drawn line), ob/ob Group (dotted line). * $p \leq 0.001$; ** $p \leq 0.04$
5.11. Results from shotgun approach with DIO plasma samples

In addition to the highly sensitive and focused approach (MRM), we analyzed biomarker changes in the plasma of mouse model for T2D; we decided to undertake a complete unbiased shotgun approach with the DIO samples. Here we took the processed samples of all time points and applied shotgun proteomics (Nielsen et al., 2006) in a label-free quantitative way. This enabled us to quantify protein changes in plasma during progression of disease, which is not necessarily related to the mass of beta-cells or their functionality. When analyzing the obtained results from this unbiased approach, we first tested if it is possible to distinguish the two different groups by looking at overall peptide pattern changes. The dendrograms derived from this analysis showed, that we were not able to distinguish the two different groups at any time point. Further data analysis uncovered that the noise in the dataset is too high and has to be filtered. After applying a stringent cut-off, where we excluding all non N-glyco peptides, took only N-glyco peptides with a probability >0.99 and reduced the noise to a minimal cut-off of 20%, we were able to separate the two groups at the 16th week time point (figure 18B), but not at others (figure 18A). Complementary information concerning all heatmaps and dendrograms from every time-point of the DIO experiment can be found in the beta-cell database (biodata.ethz.ch).

As a representative example figure 18A shows the dendrogram combined with the heatmap at week 12 of feeding. The heatmap and the dendrogram clearly show no clustering between the control, and the HFD group.

At the 16 week time point we processed 2 aliquots of the same animal sample to monitor differences between sample pairs caused by either processing or analysis variation. As the heatmap/dendrogram for this time point shows (figure 19B), all duplicate pairs of the HFD group (HF1+HF2, HF3+HF4, HF5+HF6, HF7+HF8) and the control group (LF1+LF2, LF3+LF4, LF5+LF6, LF7+LF8, LF9+LF10) cluster together. The only sample duplicate pair which is separate from each other is HF9+HF10. Sample HF9 separates from all other groups, whereas its aliquot duplicate HF10 stays in the HFD group. When looking at sample HF9 the pattern in the heatmap looks very different from the others, which results from a lower number of identified peptides. Therefore it appears as a separate group in the dendrogram.

When analyzing the 16 weeks time point, 13 proteins were found to be regulated in the plasma of the DIO model (data filtered with a p-value cut-off $p \leq 0.05$ and the earlier defined threshold for regulation of at least 0.20-fold log2 ratio. The one found to be most up-regulated (+0.46 fold) was MFAP4 (microfibrillar-associated protein 4). It was found in both models to
be significantly regulated during the course of the experiment (figure 20). Five proteins being regulated fall in the GO category, “response to stress” (GO:0006950), for example various coagulation factors: F10 (+0.40 fold up), F9 (+0.21 fold up), and F13B (-0.28 fold down). Also in this category was complement component 1, subcomponent (CS1) (+0.22 fold up), which is a serine protease associated with two other complement components C1R and C1Q and together forming the first component of the serum complement system. The most abundant compound of blood clots is fibrin and one of its members is FGB, which was found to be up-regulated +0.27 fold and also clustering in the “response to stress” GO category.

Proteins of other categories which were deregulated upon high fat diet feeding were the C-type lectin domain family 4 member F (CLEC4F) (+0.41 fold up), prestin has the symbol PRES or SLC26A5 (+0.34 fold up), the cholinergic receptor, nicotinic, beta 1 (CHRN1B1) (-0.28 fold up). SLIT1, the human homologue of Drosophila slit, is a secreted glycoprotein that plays important roles in the development of the nervous system and other organs. Human Slits are also involved in a number of pathological situations, such as cancer and inflammation (Hohenester, 2008). It was found to be up-regulated by +0.28 fold. The phospholipid transfer protein (PLTP) was monitored in both models and was found to be +0.21 fold up at the 16 weeks time-point. The speedy homolog C (SPDYC), originally found in Xenopus is a CDKs activating class of cell cycle regulators (Cheng and Solomon, 2008) found to be down-regulated (-0.20 fold). The matrix G-protein coupled receptor GPR115, containing a GPS domain in the N-terminus and a long Ser/Thr-rich region forming mucin-like stalk, was found to be down-regulated -0.28 fold.

5.12. Results from MRM approach in DIO and ob/ob mouse models

To provide the basis for the targeted proteomics strategy we used the already generated N-glyco peptide catalogue derived from MIN6 cells and human pancreatic islets. Using bioinformatics approaches we defined the 317 potential peptide biomarkers, which were synthesized and used for targeted MRM quantification with two different mouse models. Out of these we were able to quantify 54 peptides matching to 48 proteins due to their abundance in the samples. Four out of them were regulated with significant difference at various time-points of the experiments in the HFD or ob/ob group compared to its corresponding controls.

In contrast to a typical LC-MS/MS shotgun experiments, in MRM measurements no complete fragment ion spectrum is recorded. Instead the signal intensity for the selected predefined
Results

MRM transition is acquired over time. Each transition is defined by a combination of a Q1 and a Q3 m/z value resulting in filtering at the peptide and fragment ion level. The selective targeting in MRM measurements yields high duty cycles for each transition resulting in high detection sensitivity; and high selectivity is achieved because of the two-stage signal filtering. However, the optimal choice of transitions is critical for the sensitivity and selectivity of an MRM experiment. Furthermore the targeted peptides should ionize well under ESI conditions and uniquely identify the protein. These proteotypic peptides (PTP) as well as the validation, and the optimization of specific transitions are therefore critical steps in targeted proteomics. Since only a subset of all peptides of a tryptic digest is observed by mass spectrometry, the selection of these PTPs uniquely identifying the proteins of interest is of critical importance for a targeted proteomics approach.

The major goal of the targeted MRM approach was the identification of a biomarker, which would provide information about the amount of functional beta-cells during the progression of disease, before the actual onset of T2D. Therefore we identified proteins, which change their abundance in the plasma during the progression of disease.

The most prominently regulated one was MFAP4, being significantly up-regulated (+0.46 fold up). MFAP4 was also found in the shotgun experiment at the 16 weeks time point. There it was quantified for being more abundant in the HFD group compared to its corresponding controls also at additional time points, namely at 2, 4, 6, and 16 weeks. At the 12 weeks time-point, no signal could be quantified (figure 20A).

The MRM measurements reproduced the trend seen with the unbiased shotgun approach, but due to its higher accuracy the elevation in the HFD group is significant for each time-point (figure 20B). In contrast the abundance of the peptide, when quantified in the ob/ob mouse model was found to be significantly higher in the obese ob/ob mice compared to its controls. This is true for both genders and significant at all time-points excluding the 4 weeks time-point in the female group and the 20 weeks time-point in the male group (figure 20 C+D).

Another protein being regulated significantly at the 16 weeks time-point in the DIO model, was found to be different in all models (DIO and ob/ob) and by both approaches (shotgun and MRM): phospholipid protein PLTP. It was found to be highly significant up-regulated after 12 and 16 weeks in the HFD group of the DIO model quantified with SuperHirn (figure 21A). The synthetic peptide used for the MRM measurements is different from the peptide used for quantification with the unbiased approach. In the ob/ob model the MRM analysis showed
Results

significant higher values for the ob/ob group in both genders at all time-points except for 20 weeks in the female group and 12 weeks in the male group.

Vasorin (VSN) was found to be reduced in its abundance in all control groups. This was seen as general trend over all experiments and all models, but was not significant in any of the cases. In the DIO model using the shotgun approach, we found VSN only at the 16 weeks time-point down-regulated (figure 22A). This was confirmed with the MRM approach at this particular time-point, but also at the 2, 4, and 12 weeks time-points (figure 22B). In the female ob/ob model the same trend was seen, but only significant at 8 and 12 weeks (figure 22C). The male ob/ob model showed a significant down-regulation at all time-points, except of the 4 weeks time-point, even though a clear reduction was detected (figure 22D).

Desmoglein 2 (DSG2) is the fourth candidate biomarker, which could be quantified significantly at all time-points of the DIO mouse model using the MRM approach (figure 23B), but not in the in the ob/ob mouse model (figure 23C+D). There the 20 weeks age time-point was missing in both genders, and only at 8 and 12 weeks of age in the female group a significant down-regulation was found in the ob/ob group. Still, the general trend of down-regulation in the disease samples can be seen. With the shotgun approach, only the 12 and 16 weeks time-point could be quantified, whereas at 12 weeks DSG2 was significantly down regulated (figure 23A).
Figure 18. Dendrogram of the results derived from the shotgun quantification of the high fat diet experiment at the (A) 2 weeks time-point as a representative example for the time-points, where the two groups could not be separated from each other and (B) 16 weeks time point of high fat diet experiment, where it was possible to separate the two groups. Red HF (high fat diet group); blue LF (normal chow group/control group).
Results

Figure 19. Beanplot presentation of statistically significantly (p>0.05) regulated proteins in the plasma of mice after 16 weeks of feeding (DIO) using a quantitative shotgun MS approach. Only proteins are taken into account, which are regulated more than 0.20 fold log2 ratio and a p>0.05. Red (normal chow/control group), blue (high fat diet group).
**Figure 20.** Abundance of MFAP4 in blood plasma samples derived from the DIO experiment (A, B) and in the ob/ob experiment (C, D). Peptide sequence used for the quantification in (A): R.VDLEDFEN[115.0269]NTAYAK and for (B, C, D): VDLEDFEDNTAYAK. Error bars are indicated as ±SEM. ** p ≤ 0.001; * p ≤ 0.01
Figure 21. Abundance of PLTP in blood plasma samples derived from the DIO experiment (A) and in the ob/ob experiment (B, C). Peptide sequence used for the quantification in (A): R.GAFFPLKEDN[115.0269]WSLPNR and for (B, C): GHFYYDISDVR. Error bars are indicated as ±SEM. ** p ≤ 0.003; * p ≤ 0.03
Results

**Figure 22.** Abundance of VASN in blood plasma samples derived from the DIO experiment (A) and in the ob/ob experiment (B, C). Peptide sequence used for the quantification in (A): K.LHEISN[115.0269]ETFR and for (B, C, D): LHEISDETFR. Error bars are indicated as ±SEM. ** p ≤ 0.01; * p ≤ 0.05
Figure 23. Abundance of DSG2 in blood plasma samples derived from the DIO experiment (A) and in the ob/ob experiment (B, C). Peptide sequence used for the quantification in (A): R.YVQN[115.0269]GTYTAK and for (B, C, D): YVQDGTYTAK. Error bars are indicated as ±SEM. ** $p \leq 0.01$; * $p \leq 0.05$
6. Discussion

6.1. The N-glyco proteomes of beta-cells and islets

In this study we showed that the highly selective method of Cell Surface Capturing (CSC) combined with deep sequencing enabled us to discover the surface N-glyco proteome of the murine MIN6 pancreatic beta-cell line. We defined a high quality, surface specific catalogue of 259 N-glycosylated proteins, each identified by a minimum of 2 N-glyco peptides with a peptide probability of more than 99%. Bioinformatics analysis proved the accuracy of the approach, with the finding that 84.4% of the proteins have either more than one TMD or a GPI-anchor predicted. In addition 88.5% of the proteins are GO annotated for membranes and 52.3% are annotated for plasma membranes. These results indicate on one hand a high accuracy for surface proteins, but also a lack in bioinformatics predictions for TMD, and surface abundance. This handicap in predicting a surface abundance for a protein is probably caused by the missing information on surface proteomes, which we address in this study with our experimental settings and results. In comparison to obtained datasets from others, which had a similar experimental setting in other mammalian cells (Wollscheid et al., 2009) as well as in insect cells (Schiess et al., 2008) our data outcomes are consistent. The percentages of predicted surface proteins as well as the accuracy in term of N-glyco site specificity are similar. However, to our knowledge, our database (biodata.ethz.ch) contains the most comprehensive and accurate surface proteome of a cell line known and available at the moment.

Furthermore we defined the whole cell N-glyco proteome catalogues of the MIN6 cell line and primary human pancreatic islets. Since these catalogues do not reflect the surface proteome of the cells, but all N-glycosylated proteins, the total number of proteins is higher according to the strict criteria mentioned above. There is a reduced abundance of TMD and GPI-anchored proteins, which is due to increased capturing of proteins that are associated with vesicles and not having these features. These might be soluble proteins, secreted later on during their vesicle processing. When we compared the mouse MIN6 CSC, the mouse MIN6 WCC, and the human pancreatic islet WCC proteome we found 104 proteins in common. For these proteins one can assume that they exist on the beta-cell surface in mouse as well as in humans. They are also independent of the cellular model, since MIN6 cells is an immortalized cell line, whereas the human pancreatic islets are isolated primary cells. It cannot be excluded that the overlapping proteins are beta-cell surface specific, since the human pancreatic islets...
comprise of more than only one cell type. In addition none of the proteins we identified is known to be specifically expressed on beta-cells, and reported for the first time ever. To define a particular cell type it is highly likely that a single surface protein is not enough, but rather a pattern of proteins might define it like a fingerprint. In that context it is necessary to continue measurements of surface proteins, to be able to distinguish, between generally abundant surface proteins and the ones which are as a pattern unique to a certain cell type or status of a cell. The proteins which make the pattern unique to the cell type are not generally of high abundance and probably also not restricted to the cell surface. Previous proteomics approaches using islets or different beta-cell lines were not able to identify many surface proteins. The major drawback with these approaches was the separation of proteins using 2D gel electrophoresis, which makes it very difficult, in fact virtually impossible to identify membrane proteins due to their hydrophobic nature. Our measurements allowed us to identify very low abundant proteins like G-protein coupled receptors, ion channels, peptidases as well as kinases. These are the proteins, which are essential for the cell to sense its environment and to integrate external stimuli into intracellular signaling events. The knowledge of the surface proteome of beta-cells provides a new basis for further studies of surface mediated signaling events and potential targets for treatment of diabetes.

6.2. Surface protein dynamics under glucose treatment

In the following chapter about the quantitative measurements on surface proteins the term up-/down-regulation of a surface protein is used for the purpose of convenience. It indicates the abundance of the protein at the cell surface, but does not mean expression changes either on transcriptional nor on translational level. Transcriptional/translational changes are specifically mentioned. The reasons for up-/down-regulation of the abundance of a particular protein include de-glycosylation, cleavage of the extracellular domain, internalization followed by recycling or degradation. All values given for a deregulation of any proteins are indicated as log2 ratios.

To elucidate the dynamics of the beta-cell surface proteome we made use of the surface catalogue and applied a focused label-free quantification approach. This approach allowed us to observe global changes in the abundance of proteins at the cell surface, which are not accessible with any other methodology. The analysis of the surface proteome after one hour of stimulation with 25 mM glucose enabled us to gain a unique insight into protein dynamics at the surface. Within the 15 significantly changed proteins we found kinases, peptidases and receptors being regulated. The protein found to be most up-regulated was the Integrin alpha-2
(ITGA2). It is known to act as a receptor for laminin, collagen, collagen C-propeptides, fibronectin and E-cadherin. This receptor recognizes the proline-hydroxylated sequence ‘G-F-P-G-E-R’ in collagen and is responsible for adhesion of platelets and other cells to collagens (Surin et al., 2008; Sweeney et al., 2008). Due to the organizational function of ITGA2 at the newly synthesized extracellular matrix, it can be speculated that the higher abundance of the protein at the surface is caused by an increase of the cell membrane surface area. One reason could be the increased number in vesicle fusion events, leading to an increase surface area and plasticity. Although insulin vesicle exocytosis can either being a ‘kiss and run’ event or a full vesicle fusion or a combination of both (MacDonald and Rorsman, 2007), it might be possible that some protein parts of the vesicle remain in the plasma membrane (see below). Another possible reason for the up-regulation of ITGA2 could be the mitogenic effect of glucose in beta-cells, which results in growth, would therefore increase the surface area and makes the re-organization of newly formed matrix proteins necessary.

The LDL receptor (LDLR), which was found to be up-regulated +0.30 fold after one hour of glucose stimulation in our MIN6 cell experiments, is known to be up-regulated upon insulin stimulation on a transcriptional level in HepG2 cells (David P. WADE et al., 1989). Since the used beta-cells secrete insulin upon glucose stimulation, the up-regulation at the surface of the LDLR found in our experiments is probably due to the autocrine function of insulin on the beta-cell. Since we did not see a difference in the LDLR mRNA levels after 1 hour of stimulation, one can assume that the regulation of the surface abundance is either on a translational level or on a receptor recycling level. This open question cannot be answered without doing a Whole Cell Capturing (WCC) experiment under the same conditions and see if the total LDL- receptor generally changes on the protein level.

PVRL1 is also known as nectin-1, and was found to be +0.29 fold up-regulated. It is a cell-cell adhesion molecule, and when mutated in humans it causes the cleft lip/palate-ectodermal dysplasia syndrome, also known as Zlotogora–Ogur syndrome. The nectins are known to be involved in cell migration, and cytoskeleton remodeling (Ogita et al., 2008). Due to high insulin vesicle secretion events upon glucose stimulation, and the increased cellular traffic, there is an ongoing remodeling of the beta-cell cytoskeleton, whereas nectin-1 surface abundance might positively influence, and improve the ability of secretion events. To investigate the influence of nectin-1 in beta-cell cytoskeletal remodeling it would be necessary to undertake further experiments, which could include the knock down of nectin-1 and see the effects on insulin secretion pulsatility.
Another interesting protein found to be up-regulated is the SLC6A6, the taurine transporter (+0.25 fold). This transporter was shown to improve the glucose responsiveness of the body, when administered systemically (Ribeiro et al., 2009). When taurine was given to mice, they showed improved glucose tolerance, and higher insulin sensitivity compared to the controls. The islets of taurine treated animals also secreted more insulin in response to high concentrations of glucose. In addition taurine seems to have effects in regulating swelling of cells in respect to hormone secretion or cell growth (Shennan, 2008). One potential function of SLC6A6 could therefore be to help the beta-cell keeping control over its volume during insulin secretion.

Insulin-like growth factor-2 receptor (IGF2R) was found to be up-regulated in the glucose stimulation experiment, and is known to be important during embryonic development, where insulin-like growth factor-2 (IGF2) participates in the regulation of islet growth and differentiation. When IGF2 is applied on beta-cells in culture it induces proliferation (Milo-Landesman and Efrat, 2002), whereas the systemic over-expression of IGF2 in mice (Petrik et al., 1999) or specifically in pancreatic beta-cells (Devedjian et al., 2000; Okamoto et al., 2006) promotes islet cell hyperplasia and causes a type 2 diabetes phenotype. Therefore the authors conclude that IGF2 has major impact in islet hyperplasia during the onset of the disease, and that hyperplasia and hypersecretion of insulin might already occur early in disease (Devedjian et al., 2000). Therefore one can speculate from the finding of an up-regulation of IGF2R at the beta-cell surface caused by glucose stimulation, that the beta-cell becomes more sensitive to IGF2 and therefore support the growth effect on beta-cells upon glucose. One could further speculate that high local concentrations of insulin as a consequence of glucose induced insulin secretion might activate the IGF2R signaling and therefore cause the beta-cell hyperplasia. Moreover, since IGF2 is not expressed in adult rodents (Nakae et al., 2001) there is most likely another ligand binding to the IGF2R, which activates intracellular IGF2 signaling and as a consequence leading to hyperplasia. Furthermore one can consider that this up-regulation of the receptor occurs in vivo as well, and is a major contributor to the beta-cell hyperplasia in type 2 diabetes.

To elucidate the underlying mechanisms of the IGF2R on the beta-cell hyperplasia and progression of type 2 diabetes one could think about a mouse experiment, where the IGF2R fl/fl mouse (Wylie et al., 2003) would be crossed with the tamoxifen inducible RIP-cre (Nir et al., 2007) mouse. The knockout would be induced after the animal has reached adulthood to ablate the beta-cells, then a high fat diet would be applied. The non-induced controls should
develop beta-cell hyperplasia after about 16 weeks of high fat diet feeding. In case the IGF2 receptor has an impact on hyperplasia, the induced animals should not show this beta-cell phenotype upon high fat diet feeding.

One very well studied beta-cell surface protein PTPRN, which is one of the autoantigens in T1D (Notkins et al., 1998) was found to be up-regulated +0.23 fold. It was shown to be an insulin secretory granule (ISG) protein (Brunner et al., 2007) and more recently described to bind the insulin receptor, thereby stabilizes IRS2 and as a results drives beta-cell growth and proliferation (Torii et al., 2009). Therefore one can conclude that the up-regulation of PTPRN2 at the surface seen in our experiments is due to increased fusions of insulin containing vesicles with the membrane. This is probably a part of an autocrine feedback loop of insulin on the beta-cell, which drives beta-cell growth and survival via IRS2 signaling.

Three members of the tetraspanin family proteins were identified to be regulated upon stimulation by glucose and in combination with GLP-1: TSPAN15 (discussed below) TSPAN3, and CD63, the first characterized tetraspanin ever. A tetraspanin consists of four transmembrane domains (up to 350 amino acids) and the protein family consists of 33 members in humans and mouse. The function of the tetraspanins includes the involvement in physiological processes ranging from egg-sperm fusion (Vjugina and Evans, 2008), immunological responses (Levy and Shoham, 2005b) and tissue differentiation (Hemler, 2005). Experiments established that tetraspanins interact with one another and form a structural platform for the assembly of a novel class of microdomains called tetraspanin-enriched microdomains (TERM, TEM). These tetraspanin webs are proposed to regulate the spatial juxtaposition of associated transmembrane receptors (e.g. integrins, receptor tyrosine kinases) on the plasma membrane via a network of homotypic and heterotypic interactions, which results in coordination of signaling pathways (Levy and Shoham, 2005a). Originally CD63 was discovered in activated blood platelets as cell surface glycoprotein 40 and in human melanoma. Since CD63 is a tetraspanin, it interacts with many different proteins either directly or indirectly. Interaction partners include integrins, other tetraspanins, cell surface receptors, kinases, adaptor proteins and other proteins, including TIMP-1, H⁺/K-ATPase, MT1-MMP and syntenin-1 (Pols and Klumperman, 2009). The binding to synthenin-1 in particular might be very interesting, since it is known to bind syntaxin-1a (Ohno et al., 2004), which plays an essential role in the regulation of insulin secretion in beta-cells (Dong et al., 2007). Therefore one could speculate that the tetraspanin CD63 serves as a platform for the priming of insulin granules in the pancreatic beta-cell via the mediator synthenin-1. This could
be tested using confocal microscopy, where one could show co-localization of CD63 and syntenin-1 at the primed insulin granules at the cell surface. In addition it would be necessary to prove direct binding via immunoprecipitating syntenin-1 and probe for CD63 and syntaxin-1a. To further elucidate the functionality of this interaction, siRNA knockdown experiments have to be undertaken, to investigate the influence of CD63 on insulin secretion in beta-cells.

TMEM16F was +0.22 fold up-regulated and is known as a member of the anoctamin/TMEM16 family, which are Ca\(^{2+}\)-activated Cl\(^-\) channels (CaCCs). These channels perform many functions in cell physiology including secretion of fluids from acinar cells of secretory glands, amplification of olfactory transduction, regulation of cardiac and neuronal excitability, mediation of the fast block to polyspermy in amphibian oocytes, and regulation of vascular tone (Hartzell et al., 2009). Still there is not too much known about the functionality of TMEM16F, and due to its potential activation by Ca\(^{2+}\) and being a Cl\(^-\)-channel, there is a lot of potential when studying protein, to discover a major modulator of insulin secretion. Further investigations could involve knockdown experiments in combination with electrophysiological measurements of the membrane with stimulation.

ENPP4 is up-regulated +0.21 fold and belongs to the ectonucleotide pyrophosphatase/phosphodiesterase family, one of its members namely ENPP1 is known to inhibit insulin receptor signaling and has recently emerged as a key player in the development of insulin resistance (Abate et al., 2006; Bacci et al., 2007). Only little information is available about ENPP4, however GNF SymAtlas database (Su et al., 2004) expression patterns show a higher expression of ENPP4 in pancreatic islets compared to whole pancreas. This indicates an enrichment of ENPP4 in islets or even an exclusive expression of the protein in islets. Its expression in other tissues is low in general, though there is higher than median expression in immune cells and brain. When looking at protein expression using the human protein atlas (Berglund et al., 2008) does not reveal a clear picture, since the antibody staining is not specific. The abundance changes at the cell surface after glucose stimulation found in our experiments, the very specific expression pattern and the known function of its family member ENPP1 suggest a relevant functionality of ENPP4 in pancreatic islets and in insulin secretion. To further elucidate the functionality of ENPP4 one could immunoprecipitate the insulin receptor in beta-cells and/or pancreatic islets and probe for ENPP4 to see if there is an interaction, as seen for ENPP1 in insulin responsive tissues (Abate et al., 2006; Bacci et al.,
Discussion

2007). Therefore one could assume a potential role of ENPP4 in regulating insulin receptor functionality in pancreatic beta-cells.

The interleukin-1 receptor 1 (IL1R1) was found to be down-regulated -0.23 fold, which is probably not due to internalization, since it is known that the receptor undergoes intramembrane proteolytic processing, mediated by presenilin-dependent gamma-secretase activity (Elzinga et al., 2009b). Elzinga et al. showed a metalloprotease-dependent proteolytic liberation of the soluble IL-1R1 ectodomain, which produces an approximately 32-kDa large C-terminal domain fragment. This fragment remains in the plasmamembrane and serves as a substrate for subsequent gamma-secretase cleavage, which generates an approximately 26-kDa intracellular domain, which finally activates MAPK signaling and cytokine secretion. These findings suggest a regulated intramembrane proteolysis being a control mechanism for IL-1R1-mediated signaling (Elzinga et al., 2009b). With these findings one could speculate that the down-regulation observed in our experimental setting is not due to the internalization or degradation but rather due to release from the membrane. However, since this mechanism is not known for beta-cells it would be necessary to prove this mechanism in beta-cells using the same experimental approach as Elzinga et al.

We found the H2-Q1 MHC protein -0.24 fold down-regulated after one hour of glucose stimulation. Unlike all other mammalian species, which have only one class I region, rat and mouse possess a second class I region on the centromeric side of the MHC. Neither the regulation nor the function of the H2-Q1 protein is known in any context, besides being an MHC class I protein. In human type 1 diabetic (T1D) patients the hyperexpression of class I MHC by all endocrine cells in numerous islets is a well recognized phenomenon, and characteristic of the T1D disease (Foulis, 2008). Nevertheless its function in beta-cells, in particular in glucose induced insulin secretion can only be elucidated by biochemical characterization of the protein and functional analysis. The relevance of this protein in non-rodent mammalians is doubtful, since it seems to be very specific to mice and rats.

Unlike all other mammalian species, which have only one class I region, rat and mouse possess a second class I region on the centromeric side of the MHC. Neither the regulation nor the function of the H2-Q1 protein is known in any context, besides being a member of the MHC proteins and sharing MHC features.

The down-regulation of ENPEP (-0.28 fold) also known as Aminopeptidase A (APA) is not studied in beta-cells. For brain APA is known to be responsible for the conversion of
Discussion

angiotensin (ANG) II to ANG III, which leads via different pathways to a reduction of blood pressure, reduction of vasopressin release, reduction of thirst, a reduction of sodium appetite and an increase in the baroreceptor reflex (Ferreira and Raizada, 2008).

Also in the field of Alzheimer’s disease APA seems to play an important role. There are several indications showing that N-terminally truncated forms of amyloid-beta (Abeta) peptides are likely the earliest and most abundant species, which are immunohistochemically detectable in Alzheimer's disease-affected brains (Sevalle et al., 2009). The data of Sevallee et al. establish that APA contributes to the N-terminal truncation of Abeta in a calcium dependant manner, and suggest that this cleavage is likely to be abrogated as a protective function.

The tumor-associated calcium signal transducer 1 (TASCD1) also known as Epithelial Cell Adhesion Molecule (EPCAM) is a pan-epithelial homotypic cell adhesion molecule that is expressed on almost all carcinomas and was found to be down-regulated -0.30 fold. It is linked to the Cadherin-Catenin pathway and the WNT pathways, responsible for intracellular signaling and polarity. It has been shown to modulate the expression target genes like C-MYC and VEGF. It has been used as an immunotherapeutic target in the treatment of gastrointestinal, urological and other carcinomas (Chaudry et al., 2007). EPCAM is known to be expressed all over the pancreas (Seeberger et al., 2009). Since there is no function for EPCAM described in beta-cells with respect to glucose induced insulin secretion, it remains unclear what the function of the down-regulation of a cell adhesion molecule might be. Rather than a down-regulation, an up-regulation would make more sense, since EPCAM activation would promote C-MYC and VEGF to improve beta-cell functionality under stimulated conditions in an in vivo situation where growth of beta-cells and vascularization become limiting after long term high metabolic situations.

TSPAN 15 is -0.41 fold down-regulated after one hour of stimulation, which makes the tetraspanin superfamily members to be the most frequent down-regulated proteins measured in our experiments. As mentioned above, tetraspanin proteins have a coordinative function in diverse signaling pathways. This coordinational function is possibly also the role of TSPAN15 in the case of glucose stimulated insulin secretion. But since no potential binding partners are known for TSPAN15, there is no basis for speculations. To elucidate the function of the down-regulation of TSPAN15, it makes sense to find its binding partners. An immunoprecipitation of TSPAN15 followed by mass spectrometric analysis would allow first
insights into signaling pathways, for which TSPAN15 based micro domains might serve as platform.

6.3. Surface protein dynamics under glucose stimulation in combination with GLP-1 treatment

The protein found to be most up-regulated (+0.32 fold) after 1 hour of stimulation with 25 mM glucose plus 20 nM GLP-1 was leucyl/cystinyl aminopeptidase (LNPEP). LNPEP is a zinc-dependent aminopeptidase that cleaves various peptides (Keller, 2003). The protein can be secreted in maternal serum, reside in intracellular vesicles with the insulin-responsive glucose transporter GLUT4, or form a type II integral membrane glycoprotein. The protein catalyzes the final step in the conversion of angiotensinogen to angiotensin IV (AT4), and is also a receptor for AT4. Alternative splicing results in multiple transcript variants, which encode different isoforms. The other names of the protein (insulin-regulated aminopeptidase; insulin-responsive aminopeptidase) suggest a role in the insulin feedback loop in beta-cells. In non-stimulated fat cells for example more than 90% of LNPEP localizes to small intracellular vesicles and the endosomal recycling compartment, whereas less than 10% of LNPEP is found at the plasma membrane (Karylowski et al., 2004; Ross et al., 1998). The distribution of LNPEP is altered within 5 minutes after insulin stimulation, and as a result 50% of LNPEP is found at the cell surface (Ross et al., 1997). In muscle and fat cells the relative subcellular distribution of LNPEP under basal, and insulin-stimulated conditions is the same as for the insulin-responsive glucose transporter GLUT4 (Ross et al., 1998). In addition LNPEP and GLUT4 are unique among the proteins found in the GLUT4 vesicles. No other GLUT4 vesicle protein exhibits the same efficient intracellular sequestration in non-stimulated cells, and marked translocation to the cell surface in response to insulin (Kandror et al., 1995). Since we found LNPEP up-regulated after induced insulin secretion in beta-cells and the expression of GLUT4 in pancreatic islets is known (Kobayashi et al., 2004), one can speculate that the same regulation of LNPEP/GLUT4 does exist in pancreatic beta-cells. The function might be to increase the glucose up-take to a maximum for optimal insulin secretion, which is mainly driven by intracellular glucose metabolism. For further proof it would be interesting to investigate the behavior of LNPEP/GLUT4 in MIN6 cells under low glucose culturing conditions, and treated with different amounts of insulin. When performing the same N-glyco peptide capturing experiments one would expect to see an insulin dependent abundance of LNPEP at the surface of the beta-cell. In parallel it would be necessary to monitor the mRNA and protein levels of GLUT4 within the cells and at the cells surface. The abundance of
Discussion

GLUT4 at the surface has to be determined using antibodies, since we were not able to detect GLUT4 with our technology at the cell surface. In case GLUT4 is not regulated by LNPEP in beta-cells, one might unravel the regulation of the protein by looking at proteins regulated in a similar manner at the plasma membrane as ENPEP after different insulin stimuli. It still remains unclear if insulin itself has an effect on insulin secretion, whether directly or indirectly, positive or negative, even though the wealth of data, both historical and those gathered over the past 10 years, provide strong evidence that the pancreatic beta-cell indeed is a target for insulin action. However, additional work is needed to solve the existing controversies, especially with regard to insulin’s action upon its won secretory process (Leibiger et al., 2008).

The protein PTPRN2 was found to be up-regulated +0.30 fold upon the stimulation with glucose and GLP-1. As mentioned before, PTPRN2 was found to be an insulin secretory granule (ISG) protein (Brunner et al., 2007). It is further known, that PTPRN2 binds the insulin receptor, thereby stabilizing IRS2 and as a result driving the beta-cell growth and proliferation (Torii et al., 2009). Therefore we speculate that the up-regulation of PTPRN2 at the surface seen in our experiments is due to increased fusions of insulin containing vesicles with the membrane. An additional hint to support the higher abundance caused by fusion events is the fact, that we found an increase in abundance also across the different types of stimulation, even though it was only significant for PTPRN2 with the glucose/GLP1 stimulation. PTPRN2 could be monitored in both stimulation experiments and was found to be higher in glucose/GLP-1 (+0.30 fold) compared to only glucose (+0.20 fold). An even higher up-regulation (+0.40 fold) was observed, when the MIN6 cells were treated with a secretagogue cocktail. This indicates a correlation between the amounts of secreted insulin/fused vesicles to PTPRN2 surface abundance. This is probably a part of an autocrine feedback loop of insulin secretion on the beta-cell, which drives beta-cell growth and survival via IRS2 signaling.

Ectonucleotide pyrophosphatase/phosphodiesterase 3 (ENPP3) was found to be up-regulated +0.27 fold after one hour of stimulation with glucose/GLP-1. This protein belongs to a group of ectoenzymes that are involved in hydrolysis of extracellular nucleotides. These ectoenzymes possess ATPase and ATP pyrophosphatase activities and are type II transmembrane proteins. In basophiles ENPP3 is a very specific marker for cell activation and allergy diagnosis. The associated signaling events with ENPP3 up-regulation seem to involve PI3K signaling and show a dependency on Ca²⁺ (Buhring et al., 2004). Since only little is
known about ENPP3 and nothing is known in respect to beta-cells, only biochemical characterization of the protein can provide insight in its involvement in stimulated insulin secretion.

The probably best known reference surface protein being strongly regulated by glucose stimulation is the GLUT2 glucose transporter. This transporter is very specifically expressed in pancreatic beta-cells and is essential for glucose-stimulated insulin secretion, thereby controlling blood glucose homeostasis in response to diet. It was shown that the GLUT2 residency on the beta-cell surface is dependent on its N-glycosylation. In case the N-glycosylation is lost, the GLUT2 transporter is endocytosed and redistributed into endosomes and lysosomes (Ohtsubo et al., 2005). Even though the receptor is N-glycosylated we were not able to measure it, either due to limitations of the methodology applied in our experiments or because GLUT2 is not expressed in the MIN6 cells used for the study. In contrast we found the GLUT1 glucose transporter +0.27 fold up-regulated after one hour with glucose/GLP-1 stimulation. Since we found both GLUT1 and GLUT2 expressed on mRNA level in the MIN6 cells used in our experiments one would assume relatively more GLUT1 than GLUT2 due to the cancer cell line character of the MIN6 cells, it might well be that GLUT1/3 compensates for lower amounts of GLUT2, which was seen by loss of 1st phase insulin secretion after knockdown of GLUT2 in beta-cell (Guillam et al., 2000). This compensation is neither regulated on a transcriptional nor on a translational level, since the total amounts of GLUT1/3 protein did not change (Guillam et al., 2000). Therefore it is likely that the up-regulation of GLUT1 observed in our experiments after stimulation is caused by translocation of the protein to the plasma membrane. In fact this could explain the previous observation by Guillam et al. in their GLUT2 knockdown cells. Maybe the very different $K_m$-values (Uldry and Thorens, 2004) of GLUT1 ($K_m$ ~3 mM) and GLUT2 (~17 mM) could also explain the up-regulation of GLUT1 after glucose/GLP-1 induced insulin secretion necessary, since there is a high demand of intracellular glucose.

The tetraspanin CD63 was similarly up-regulated as in the glucose stimulation experiment namely +0.26 fold up. Therefore one can assume a similar function for CD63 also after the glucose/GLP-1 stimulation (discussed above). Furthermore one can assume that the regulation of CD63 is glucose dependent, but cannot be further enhanced by GLP-1.

The protein suppression of tumorigenicity 14 (colon carcinoma) (ST14) also known as matripase was found to be up-regulated +0.24 fold after one hour of stimulation with glucose/GLP1. Matripase is an epithelial-derived, integral membrane serine protease. This
protease forms a complex with the Kunitz-type serine protease inhibitor hepatocyte growth factor activator inhibitor (HAI-1), and is found to be activated by sphingosine1-phosphate. Matripase has been shown to cleave and activate hepatocyte growth factor/scattering factor, and urokinase plasminogen activator, which suggest the function of this protease as an epithelial membrane activator for other proteases and latent growth factors (Szabo et al., 2008). Its function in pancreatic beta-cells is so far not elucidated, but since it is known that GLP-1 has a positive impact on growth factor signaling in beta-cells (Holz and Chepurny, 2005), it might well be that ST14 up-regulation plays a role in the activation of this pathways via activating other surface proteases to further support the growth effect mediated by GLP-1 (Holz and Chepurny, 2005).

LTBR the lymphotoxin beta receptor (TNFR superfamily, member 3), was +0.23 fold up-regulated. It is known to be expressed on the surface of most cell types, including cells of epithelial and myeloid lineages, but not on T and B lymphocytes. In lymphoid tissue the activation of LTBR by its ligands activates the classical and noncanonical NF-kappaB signaling pathways (Madge et al., 2008). A similar effect could be induced by the up-regulation of the receptor in beta-cells. Increased NF-kappaB signaling would be a possible explanation for the known positive effect of GLP-1 on beta-cells survival via the anti-apoptotic pathways.

In contrast to the down-regulation of TSPAN15 when cells were stimulated with glucose only, one hour stimulation with glucose/GLP-1 lead to a +0.22 fold up-regulation of TSPAN15. As already discussed tetraspanins build microdomains, which serve as platforms for other types of proteins like integrins or receptors tyrosine kinases. Therefore it might well be, that the down-regulation of TSPAN15 after glucose stimulation might be an adaptive effect of the cells, which might influence the insulin secretion in a negative feedback loop. Since GLP-1 activates other pathways to further increase the insulin secretion, it might well be that the protein associated with TSPAN15 might be important to activate a pathway, which has the ability to further enhancing the insulin secretion. Since the protein associated with TSPAN15 is not known, this conclusion will remain highly speculative until an experimental proof has shown the involved pathways.

Unc-5 homolog A (UNC5A) was found to be up-regulated +0.21 fold after glucose/GLP-1 stimulation. It belongs to the family of UNC5H netrin-1 receptors, which are transmembrane proteins containing 2 immunoglobulin (Ig)-like domains and the 2 type I thrombospondin motifs in the extracellular region. The ligands of UNC5A are netrins, which are secreted
proteins that direct axon extension and cell migration during neural development. They act as attractants for some cell types and as repellents for others, which makes them bifunctional proteins. These opposite actions are thought to be mediated by two classes of receptors, whereas the UNC5 family receptors mediate the repellent response to netrin (Arakawa, 2004). UNC5 receptors are known to be pro-apoptotic since they have a caspace-3 cleavage site at their intracellular domain. This cleavage site is covered or blocked after receptor dimerization caused by netrin binding (Hong et al., 1999). Under such circumstances UNC5A is anti-apoptotic (Paradisi et al., 2008). De Brueck et al. found netrin-1 expressed in late gestation in the fetal pancreas, which is the period of islet morphogenesis. However, in adult pancreas no netrin-1 expression could be detected. The expression re-appeared, when they applied a duct ligation, which is known to induce islet neogenesis. These observations suggest that netrin-1 plays a role in pancreatic morphogenesis, both prenatally and in the regenerating adult rat pancreas (De Breuck et al., 2003). The up-regulation of the netrin receptor UNC5A suggests a potential involvement of netrin re-expression in pancreatic islets induced by GLP-1. To test this hypothesis it would be necessary to measure mRNA levels of netrin-1 in beta-cells before and after stimulation with glucose/GLP-1. The expression of netrin-1 is known to be regulated by NF-kappaB, and since we found up-regulation potential NF-kappaB signaling activators (LTBR, discussed above), there is evidence for a re-expression of netrin-1 caused by glucose/GLP-1.

Another member of the transmembrane 4 superfamily TSPAN3 was found to be up-regulated (+0.21 fold) after stimulation with glucose/GLP-1. Only little is known about the function of TSPAN3/OAP-1 except that it forms a complex with OSP/claudin-11 and beta-1 integrin which regulates proliferation and migration in oligodendrocyte progenitors (Tiwari-Woodruff et al., 2001). One can hypothesize that TSPAN3 has a similar function in beta-cells, and therefore assists the proliferative effect of glucose/GLP-1.

The up-regulation (+0.20 fold) of ectonucleotide pyrophosphatase/phosphodiesterase 1 (ENPP1) upon glucose/GLP1 stimulation has not yet been described for beta-cells, but for insulin target tissues it is a known observation. Overexpression of ENPP1 in target tissues causes an early intrinsic defect observed in human insulin resistance (Bacci et al., 2007). Cohort studies in humans looking for a K121Q polymorphism showed a correlation between this mutation, and defects in glucose homeostasis already in young individuals (Baratta et al., 2008). A delay in total insulin release during an oral glucose tolerance test was also found in this group of patients. Furthermore ENPP1 is thought to be involved in peripheral insulin
Discussion

resistance in T2D, where it reduces IR/IRS insulin signaling (Goldfine et al., 2008). In hepatic cell lines in vitro, and in an animal model of diabetes (db/db mice), suppression of ENPP1 expression improves insulin sensitivity via PKB/AKT phosphorylation supporting the proposition that ENPP1 inhibition is a potential therapeutic approach for the treatment of type 2 diabetes (Zhou et al., 2009). Most of the studies concerning ENPP1 investigate its function only in peripheral tissues in respect to diabetes. Therefore no function of the protein for beta-cells is known; only the expression was demonstrated in pancreas and pancreatic islets (Goding et al., 2003). One can now speculate based on the mentioned findings that the function of ENPP1 might control the insulin sensitivity of beta-cells by modulating pathways involved in insulin signaling via PKB/AKT phosphorylation. Even though ENPP1 was not found to be regulated in our glucose stimulation experiments, ENPP1 might still have an impact during glucose/GLP-1 treatment. To further investigate the influence of ENPP1 on glucose/GLP-1 stimulated insulin secretion one could test the effects of down-regulation, and overexpression in MIN6 cells using siRNA, and viral vector systems. In addition it would be possible to isolate islets from ENPP1 knockout mice (Sali et al., 1999), and compare their functionality to the corresponding controls.

The homologue to the *Drosophila* sidekick protein (SDK1) was found to be down-regulated (-0.21 fold) upon glucose/GLP-1 stimulation. The sidekick protein is a member of the immunoglobulin superfamily, and is required for pattern formation in the Drosophila eye (Nguyen et al., 1997). Together with DSCAMs, sidekick directs lamina-specific synaptic connections in vertebrates retina (Yamagata and Sanes, 2008). All studies on SDK1 and its homologue do not elucidate any functionality or signaling, which could give any basis for speculation about the protein’s function in beta-cells. Only biochemical characterization including siRNA and overexpression experiments can elucidate the impact of SDK1 on stimulated insulin-secretion.

The down-regulation (-0.21 fold) of neogenin (NEO1) at the plasma membrane in beta-cells has not yet been reported, even though there are numerous studies investigating the receptor’s function. Neogenin is a multifunctional transmembrane receptor belonging to the immunoglobulin superfamily. It displays identical secondary structure to the Netrin-binding axon guidance receptor deleted in colorectal cancer (DCC), and is involved in axon guidance, and cell survival (Matsunaga et al., 2004). Like in DCC, neogenin is able to transduce signals elicited by netrin. These neogenin-netrin interactions have been implicated in tissue morphogenesis, angiogenesis, myoblast differentiation, and also in axon guidance. Neogenin
is also involved in neuronal differentiation, apoptosis, and repulsive axon guidance. Numerous studies have been started to elucidate the in vivo functions of neogenin, and its role in multiple aspects of development, and homeostasis (Wilson and Key, 2007). One could now hypothesize that the binding of netrin-1 to the neogenin should prevent apoptosis similar as discussed for UNC5A. However, if the function of neogenin would be to prevent the beta-cells from apoptosis, one would expect it to be up-regulated or unchanged upon glucose/GLP-1 stimulation. But since we observed neogenin to be down-regulated, the change in receptor abundance might have a different function in beta-cells than in other tissues. Most of the studies on NEO1 focused on its cleavage by caspase-3, and on the binding of its ligands (netrin and RGM), but not on its surface abundance (Wilson and Key, 2007). Therefore only experimental investigations will be able to elucidate the function of NEO1 in beta-cells. Such experiments should include the down-regulation of NEO1 by siRNA followed by glucose stimulation to elucidate the effect of lower protein abundance on insulin secretion.

The transmembrane protease, serine 4 (TMPRSS4) is known to be overexpressed in pancreatic carcinoma (Choi et al., 2008), and found to be down-regulated (-0.25 fold) in our glucose/GLP-1 stimulated MIN6 cells. Proteases play important roles in the development, and homeostasis of an organism, and therefore, even though the functions of TMPRSS4 in this respect are still unclear, the down-regulation of this protease might have an important influence on beta-cells when stimulated with glucose/GLP-1. To elucidate these functions it will be necessary to experimentally investigate the loss of TMPRSS4 in beta-cells, and the impact of the down modulation on the beta-cell functionality.

The Zinc-transporter SLC39A14 transports Zn\textsuperscript{2+} over the plasma membrane in a temperature dependent manner (Taylor et al., 2005), and it was found to be down-regulated -0.26 fold after glucose/GLP-1 stimulation. In the pancreatic beta cell, Zn\textsuperscript{2+} is crucial during the insulin granule maturation process. The insulin is crystallized with Zn\textsuperscript{2+} and calcium in the form of dense-core granules during this process (Howell et al., 1978). Furthermore it is well established, that Zn\textsuperscript{2+} is essential for insulin production, and functionality (Emdin et al., 1980). Since the rate of insulin production and secretion is very high when cells are stimulated with glucose/GLP-1, one would expect higher demand of Zn\textsuperscript{2+}, and therefore an increased abundance of SLC39A14 on the cell surface. Zinc in general is crucial for the function of more than 300 enzymes, and it is very important for cellular processes like cell division, and apoptosis (Truong-Tran et al., 2000; Vallee and Falchuk, 1993). The exact effect
Discussion

of the down-regulation of SLC39A14 in beta-cells, and its impact on cellular processes remains to be explored.

The cell adhesion molecule 4 (CADM4) is a homologue of the human lung cancer tumor suppressor TSLC1/CADM1 (Fukami et al., 2003), and was found to be down-regulated (-0.26 fold) at the plasma membrane after glucose/GLP-1 stimulation. Loss of CADM4 expression was found in several cancers thus suggesting a function as tumor suppressor. CADM4 was also demonstrated to be expressed on the plasma membrane in murine cells, especially at the cell-cell attached site (Fukami et al., 2003). Based on this, one could speculate about a role as a potential driver of beta-cell proliferation upon glucose/GLP-1 stimulation. Since the down-regulation on the plasma membrane might indicate its inactivation as a tumor suppressor.

CD164 is a receptor/adhesion molecule, was found to be down-regulated -0.27 fold. It is implicated in mediating or regulating hematopoietic precursor cell adhesion to stroma, and may serve as a potent negative regulator of hematopoietic progenitor cell proliferation (Jorgensen-Tye et al., 2005). Also in Drosophila a proliferative effect after down-regulation of the CD164 homologue was found, which indicates a conserved function of CD164 on proliferation (Zhou et al., 2006). Therefore its down-regulation (-0.27 fold) from the plasma membrane observed in our experiments might implicate a pro-proliferative effect on the beta-cells caused by the glucose/GLP-1 stimulus. But since there is no experimental evidence for a function of CD164 in beta-cells, only functional analysis will be able to elucidate the impact of CD164 on beta-cell proliferation under stimulated as well as under un-stimulated conditions.

SLC1A4 is a glutamate/neutral amino acid cotransporter, and was found to be -0.33 fold down-regulated. Polymorphisms in SLC1A4 are known to associate with schizophrenia, whereas no functional mechanism is known behind this finding (Deng et al., 2008). In intestinal epithelial cells an inhibition of SLC1A4 was found after LTD4 treatment of the cells. This effect was secondary to a decrease in affinity of the cotransporter for alanine without a significant change in cotransporter numbers, but it was not secondary to an alteration in the Na⁺ extruding capacity of the cells. Our observations show that SLC1A4 is down-regulated upon glucose/GLP-1 stimulation as secondary effect of the cells to react on Na⁺ extrusion changes in beta-cells. The Na⁺/Ca²⁺ exchanger system is activated under stimulation conditions, and plays an important role in the function, growth, and demise of the beta-cell (Herchuelz et al., 2002).
Human sodium/inositol cotransporter 2 (SMIT2/SLC5A3) imports inositols in exchange with Na⁺, and was found to be down-regulated -0.36 fold. It was shown that the receptor abundance is changed in peripheral tissues of type 2 diabetes patients, and this is thought to contribute to the insulin resistance (Lin et al., 2009). Still there is a lot of information missing concerning the function of SLC5A3 in these tissues, which has to be elucidated in future experiments. In our experiments the down-regulation could be explained by another pathway which might involve the autocrine effect of insulin. However, only further experiments can define the exact function of SLC5A3 in beta-cells.

The *Drosophila* FAT is a tumor suppressor essential for the control of cell proliferation during Drosophila development. We found its homologue FAT1 -0.40 fold down-regulated after glucose/GLP-1 stimulation. FAT1 is a member of the cadherin superfamily, a group of integral membrane proteins characterized by the presence of cadherin-type repeats. In addition to 34 tandem cadherin-type repeats, FAT1 has five epidermal growth factor (EGF)-like repeats and one laminin A-G domain. This gene is expressed at high levels in a number of fetal epithelia. Its product probably functions as an adhesion molecule and/or signaling receptor, and is likely to be important in developmental processes, and cell communication.

FAT1 is an important regulator of actin dynamics, and controls cell-cell interactions through the activity/binding to ENA/VASP proteins (Tanoue and Takeichi, 2005). In FAT1 knockdown cells the junctional actin cytoskeleton does not form properly. Nevertheless this phenotype cannot be only explained by dysfunction of ENA/VASP proteins. Therefore other signaling pathways downstream of FAT1 must have an influence. Several proline-rich sequences, and a PDZ-domain binding motif can be found in the FAT1 protein (Tanoue and Takeichi, 2005). For speculations about the function of FAT1 in beta-cells, and why it is down-regulated upon glucose/GLP-1 stimulation, the identification of molecules, which interact with these consensus sequences and motifs, will be crucial.

The hepatocyte growth factor activation inhibitor SPINT2, also known as HAI-2, is a serine protease inhibitor and a favorable prognosis marker in cervical cancer. It inhibits cell growth through the apoptotic pathway (Nakamura et al., 2009). The hepatocyte growth factor (HGF) regulates cell growth, cell motility, and morphogenesis by activating a tyrosine kinase signaling cascade after binding to the proto-oncogenic c-MET receptor (Funakoshi and Nakamura, 2003). Therefore one could speculate, that its down-regulation (-0.42 fold) after glucose/GLP-1 stimulation, as we found it in beta-cells, would promote cell growth. This could happen via abolishing the inhibition of HGF-signaling. To test this it would be
necessary to study the influence of HGF on beta-cells in combination with and without SPINT2 knockdown. This would further elucidate the function of SPINT2 not only in the context of beta-cells, but also in respect to its function in cancer.

Cell adhesion molecule 1 (CADM1/IGSF4) is a novel pancreatic-islet cell adhesion molecule that mediates nerve-islet cell interactions. The strong correlation between CADM1 expression and hormonally functional phenotypes suggests that CADM1 is involved in hormone secretion from islet cell tumors (Koma et al., 2008). CADM1 is found to be down-regulated in several tumors, and therefore its thought to function as a tumor suppressor (Murakami, 2005). The down-regulation (-0.42 fold) of the protein in our experimental setting, enables the speculation that this negative regulation might inactivate its tumor suppressor function to drive beta-cells proliferation and influence the insulin secretion as it was suggested by Koma et al.

The CD24A antigen is well studied in antigen-presenting immune cell including Langerhans cells (LC) and dermal dendritic cells (dDC) (Stuttle et al., 2008), and we found it -0.49 fold down-regulated. Still only little is known about the regulation of CD24a expression even in LC and DC cells. Stuttle et al. only noticed that an increased expression of CD24a was found in LCs and dDCs after their sensitization and migration to lymph nodes. Therefore increased CD24a expression might be related to the migratory or homing behavior of LCs/dDCs (Stuttle et al., 2008). In the literature there is no hint about a putative function of CD24a in beta-cells, therefore it is difficult to interpret the observed down-regulation of the protein from the surface upon glucose/GLP-1 stimulation.

The adrenergic alpha 2A receptor (ADRA2A) was down-regulated -0.58 fold upon glucose/GLP-1 stimulation for 1 hour. The receptor belongs to the alpha-2-adrenergic receptors, which are members of the G-protein coupled receptor superfamily, comprising of 3 highly homologous subtypes: alpha2A, alpha2B, and alpha2C. These receptors have a critical role in regulating neurotransmitter release from sympathetic nerves and from adrenergic neurons in the central nervous system. Studies in mouse revealed that both the alpha2A and alpha2C subtypes were required for normal presynaptic control of transmitter release from sympathetic nerves in the heart and from central noradrenergic neurons; the alpha2A subtype inhibits transmitter release at high stimulation frequencies, whereas the alpha2C subtype modulates neurotransmission at lower levels of nerve activity. Generally alpha-2-adrenergic receptors are known to have a negative effect on adenylate cyclase, increasing the activity of K+-channels and decreasing Ca2+-channel abundance, and increase ERK/MAP kinase
Discussion

signaling. As a result, cAMP levels are reduced, a hyperpolarisation of the plasma membrane occurs and the Ca\(^{2+}\) influx is reduced (Sanders et al., 2008). All this counteracts for an effective insulin secretion by the pancreatic beta-cell. It was shown that mice with beta-cell specific knockout of ADRA2A have reduced blood glucose levels, increased insulin levels and improved glucose tolerance (Savontaus et al., 2008). In addition, Fagerholm et al. showed that alpha2a-adrenoreceptor antagonism increases insulin secretion and synergistically augments the insulinotropic effect of glibenclamide in mice (Fagerholm et al., 2008). These results clearly show that a lower abundance/functionality of the ADRA2A in beta-cells improve the insulin secretion and beta-cell functionality. Therefore our observation, of the down-regulation of ADRA2A upon glucose/GLP-1 stimulation makes sense, since it is well known that glucose/GLP-1 increases the insulin secretion of beta-cells. However, all previous studies on ADRA2A beta-cell knockouts conclude, that in wild type animals the receptor signaling is regulated only by signals coming from peripheral tissues. We present here for the first time evidence, that the beta-cell itself is able to control the input from nerves by actively down-regulating the ADRA2A receptor from the surface to reduce its signaling, which results in improved insulin secretion.

As a proof of concept of our experimental settings, it was of high interest to monitor the abundance of the GLP-1 receptor on the surface, since GLP1R surface abundance is known to be changed upon stimulation with its ligand GLP-1 (Widmann et al., 1995). Due to low abundance of the receptor we were only able to monitor and quantify the abundance of GLP1R in a single experiment, in which we used 3 times more cells compared to the other experiments. In this single experiment, we monitored and quantified 2 different N-glyco peptides matching the receptor (figure 10B). Both were down-regulated -0.90 fold which fits the published data (Widmann et al., 1995). Importantly, here we were able to absolutely quantify the abundance of the receptor on the surface of the murine MIN6 beta-cells after one hour of stimulation with glucose/GLP-1. In our experimental settings, this regulation did not happen at the mRNA level, as it was shown in vivo for rat islets, where GLP-1 mRNA levels were down regulated when the rats were exposed to hyperglycemia over a period of 4 weeks (Xu et al., 2007). This difference which we observed in our experimental approach, compared to the published data can be explained by a more acute effect in our model compared to a long-term effect in the animals. Xu et al. also observed a translational down-regulation of the receptor in their experiments. This we could not show either, since the available antibody did not work in our hands. Still the internalization of the GLP-1 receptor was so far neither shown for the endogenous protein nor for a pancreatic beta-cell and is a novel finding.
In summary, we demonstrated the feasibility of a label-free quantitative proteomics approach to monitor abundance changes of proteins at the cell surface of adherent mammalian cells, which was not successfully undertaken before. The monitoring of overall surface proteome changes gives deep insights into the origin of beta-cell surface signaling upon stimulation with glucose alone or together with GLP-1. Even though the function of the deregulation of these proteins, and the interplay of potential signaling pathways involved remains to be elucidated, the observations we made provide a basis for further functional studies in the complex interplay which influences beta-cell functionality, proliferation, growth, apoptosis and survival.

6.4. Candidate Biomarker evaluation in mouse models of T2D disease

Quantifying changes in protein abundance at the surface of a beta-cell in culture upon different stimuli provides major insights into surface signaling events, but does not really reflect the in vivo situation in a complex organism. In that situation there are many more stimuli than only one at the same time which influence the behavior of a beta-cell in its natural environment embedded in the pancreas as a component of a functional islet. The system is much more complex compared to the controlled cell culture experiment discussed before.

To tackle such a complex problem, one can choose different ways. One possibility is to take the already defined MIN6 cell surface reference map to specifically quantify label-free N-glyco peptides of interest. This approach reduces the loss of information due to technical limitations of the MS run, such as the selection for fragmentation ions. Another advantage using this approach is the ability of high throughput measurements, since the machine does not run in scanning mode, which saves a lot of time. When the signal-to-noise ratio is good for the acquired experiments, it is possible to measure greater than two fold changes, which are statistical significant. The biggest advantage of this focused label-free approach is that it can be performed with samples from any source and that there is no labeling necessary. A clear disadvantage is that multiple possible errors can occur during sample processing and MS analysis, for example, if the method is sensitive enough to the remaining detergents or abundant background peptides in some of the samples. The overall normalization of the single runs, and spiking in reference peptides in the samples are ways to minimize these limitations (Meng et al., 2007) (this issue will be discussed later). Finally we decided not to use the focused label-free quantification with the beta-cell proteome as reference, but decided to use...
the mass-spectrometer in the scanning mode. This would provide additional information about the abundance of proteins, which were not found in the beta-cell N-glyco proteome. Anyway, most of the proteins of the beta-cell proteome will be quantified using the MRM approach (discussed later) since the potential biomarker candidates are contained in the beta-cell proteome.

For the MRM approach, we used public available data sources like gene expression data (Barrett et al., 2007), protein expression data (Berglund et al., 2008), protein-protein interaction data (Chatr-Aryamontri et al., 2008; Kerrien et al., 2007), gene ontology groups (Harris et al., 2004) and the KEGG database (Kanehisa and Goto, 2000) in combination with the beta-cell proteome to reveal potential biomarker candidates. For normalization reasons, ‘house keeping’ proteins were selected to avoid bias caused by protein amount or processing variations. Since each tryptic digestion of a protein will produce a large number of peptides, it makes sense to only focus on one to three representative proteotypic peptides per protein. The selection of these peptides is again very crucial for the success of the MRM experiment, since for most of the generated peptides derived from a tryptic digest only a small number is observed in a routine measurement (Kuster et al., 2005). Therefore it makes sense to use databases containing peptides, which have been previously observed in other experiments (Deutsch et al., 2008). In addition other peptide features should be taken into account like the uniqueness of the peptide. Missed cleavages and non-tryptic cleavage should be avoided; post-translational modifications and chemical induced modifications should be avoided as well. The final critical step is the optimization of transitions, since it is essential to select specific transitions for the most intense fragments of a peptide to obtain a high-sensitivity result. Since the total number of transitions per LC-MS run is limited, the assay is also limited to the analysis of a few peptides. Taking all these parameters into account, we chose 317 synthetic peptides to be screened as candidates for potential biomarkers. This is very little relative to the total number of potentially existing peptides, which is 749’163 unique tryptic peptides within a mass range of 500-5000 based on IPI database protein entries. Note that this number does not consider possible sequence variation and posttranslational modifications of each peptide. Even though only 52,442 unique peptides (7.0%) contain potential N-linked glycosites (Zhang et al., 2006), the set of peptides we chose is minor. Still it might be sufficient to cover the beta-cell relevant markers due to our bioinformatics efforts.

Another critical point in such kind of an in vivo model experiment is the choice of the mouse model. It is necessary to ensure, no matter which metabolic model for T2D one uses that the
metabolic difference of the distinct groups is only due to either the treatment (HFD) or the genetic background (ob/ob), but not caused by other genetic defects. The example of the C57BL/6J mice used for diet induced obesity experiments demonstrates this very clear: Recently it was observed that the original C57BL/6J strain derived from Jackson laboratories shows already a metabolic phenotype even under an unchallenged situation. This phenotype was shown to be caused by a mutation in the nicotinamide nucleotide transhydrogenase (Nnt) gene (Freeman et al., 2006a; Freeman et al., 2006b; Toye et al., 2005). The C57BL/6J mice used in our experiments are derived from a colony located at Janvier laboratories, which was tested negative for the relevant Nnt mutation by the company Biolytix. For this reason we assume, that the metabolic phenotype in our experiments with the C57BL/6J animals is caused by the high fat chow, but not by deficiency in Nnt.

This metabolic phenotype of the DIO mice was monitored for each time-point with several relevant physiological parameters, and after 16 weeks of high fat diet feeding the HFD group showed a dramatic hyperglycemia even after overnight starvation (figure 13B). This indicates a successful induction of T2D after 16 weeks of feeding. In addition, the IPGTT performed after 12 weeks (figure 14) indicates the successful induction of T2D. When then analyzing the results obtained from the shotgun approach, and looking at the overall data exclusively using the N-glyco peptides with a probability greater than 0.99, we were able to separate the groups in a dendrogram at the 16 weeks time-point (figure 18), which indicates a clear difference comparing the two groups of mice (control vs. HFD) - another support of the successful induction of the obesity caused T2D in this model.

Having a closer look at the ratios of significant over-threshold regulated peptides; we found 13 proteins being deregulated. 10 of them were up-regulated and 3 were down-regulated in their abundance in the DIO blood plasma samples. One group of them matches to the GO category “response to stress” and consists mostly of proteins involved in coagulation signaling (Krupiczkojc et al., 2008). For example the coagulation factor F10 was up-regulated (+0.40 fold) in the HFD group, and is known to be synthesized in the liver, wherefore vitamin K is required (Menegatti and Peyvandi, 2009). It is a serine endopeptidase (protease group S1), which becomes activated into Xa by both factor IX (also found to be up-regulated, +0.20 fold) and factor VII dependent coagulation factor of the blood coagulation cascade. All of them are parts of the domain which is structurally homologous to those of the other haemostatic serine proteases. The mature factor is activated by the cleavage of the activation peptide by factor IXa (in the intrinsic pathway), or by factor VIIa (in the extrinsic pathway).
Discussion

The activated coagulation factor F10 then converts prothrombin to thrombin in the presence of factor Va, Ca$^{2+}$, and phospholipid during blood clotting. Mutations of this gene result in factor X deficiency, a hemorrhagic condition of variable severity (Krupiczojc et al., 2008). Besides other mechanisms, the increased activity of the coagulation cascade, the decreased activity of the fibrinolytic cascade, increased inflammation, increased oxidative stress, endothelial dysfunction, disturbances of lipids and glucose intolerance are all associated with the metabolic syndrome (Darvall et al., 2007).

Therefore one can speculate that with our unbiased approach we partially monitor the activation of the coagulation cascade, leading to thrombosis and inflammation, both known to be associated with obesity.

When looking at other proteins deregulated, the microfibrillar-associated protein 4 (MFAP4) is the most up-regulated one. This is not only the case at the 16 weeks time-point, but also at the earlier ones, which was confirmed by the MRM analysis of the same samples. In contrast to the DIO mouse model, MFAP4 is down-regulated in the ob/ob mouse model, which seems to be confusing at first sight, but might be explained by the following speculations: MFAP4 is also known as MAGP-36, and was originally discovered in porcine aorta (Kobayashi et al., 1989). MAGP-36 has been localized on the surface of elastic fibers or laminae in immunohistochemical studies. During development and growth, an increase of MAGP-36 on the aorta surface and in elastic fibers was found in chicken (Toyoshima et al., 2005). In this study they could roughly correlate the increase in abundance of MAGP-36 with an increase in blood pressure. Therefore the question arises: is there a difference in blood pressure of the DIO HFD animals and the ob/ob animals? As generally known, one would expect an increased blood pressure in a mouse model for the metabolic syndrome, as hypertension is one of the hallmarks of the syndrome (Huang, 2009). This is partially true for the DIO mouse model, where a slightly increased blood pressure was found for male C57BL/6J mice fed with a high fat diet for 16 weeks (Williams et al., 2003). The authors argue with the major increased plasma leptin levels in the HFD group, which influence the sympathetic nervous system leading to the changes in blood pressure. Even though the changes in blood pressure are minor, and mice in general seem to be resistant to obesity-induced hypertension, it is notable that the maximum mean blood pressure is increased under high-fat diet of about 15 mmHg (Williams et al., 2003). Importantly, in the ob/ob mouse model the leptin is missing and the animals show a lower blood pressure compared to their lean controls (Mark et al., 1999),
therefore one might speculate about MFAP4 as a potential biomarker candidate for hypertension.

In addition, recent studies in humans identified MFAP4 as a potential marker for liver cirrhosis (Molleken et al., 2009). Since non fatty liver disease is also known to be related to obesity and to the metabolic syndrome, it might well be that the lack of leptin in the ob/ob mouse might be protective against liver cirrhosis and that DIO animals develop mild liver cirrhosis from the very beginning of the HFD feeding. It would be necessary to prove the measurements of the opposing results for MFAP4 in the two models using a MFAP4 specific antibody onset of liver cirrhosis in the HFD animals, but not in the ob/ob mice.

Phospholipid transfer protein (PLTP) is a lipid transfer glycoprotein, and was found in both the HFD animals and the ob/ob mice elevated at any time-point, even though not always significant. In earlier studies, a positive role of PLTP in high-density lipoprotein (HDL) metabolism was shown in vivo and in vitro. The over-expression of human PLTP in C57BL/6 mice lowered plasma HDL levels, probably because of increased HDL catabolism (Foger et al., 1997). Unexpectedly, plasma HDL concentration was also found to be decreased in PLTP knockout mice (Yan et al., 2004). These intriguing observations can be explained by recent studies that indicate that PLTP is not only involved in remodeling of HDL sub-fractions but also in VLDL turnover. The role of PLTP in atherogenesis and VLDL synthesis was demonstrated in transgenic mouse models with increased susceptibility for development of atherosclerosis, bred into PLTP knockout mice. The data clearly show that PLTP can be pro-atherogenic. As mentioned above, however, PLTP may have antiatherogenic potential in wild-type C57BL/6 mice. Coincidently, high levels of PLTP and CETP have been found in the plasma of obese patients (Tzotzas et al., 2009). PLTP activity and abundance have been reported to be abnormally elevated in type 2 diabetes (T2D) and insulin-resistant states. This elevation is frequently associated with hypertriglycerideridemia and obesity (Albers and Cheung, 2004). Our results are consistent with previous findings on PLTP in obesity and diabetes models. Therefore one can say that our methodology applying MRM as well as unbiased shotgun technology to quantify proteins changes in plasma is a successful strategy to find biomarker candidates for the metabolic syndrome. The protein vasorin (VASN) was found to be higher in all control animals of the DIO and the ob/ob mouse models compared to the disease. For some time points this difference was found to be significant. Not much is known about vasorin. It is thought to be a transforming growth factor beta-binding protein expressed in vascular smooth muscle cells, where it modulates the arterial response to injury in vivo.
Discussion

(Ikeda et al., 2004). In zebrafish VASN is known under the name SLITL2, and thought to be involved in central neural and vascular morphogenesis (Chen et al., 2005). Neither GNF (Su et al., 2004) nor the human protein atlas (Berglund et al., 2008) show a specific expression for a particular organ. In GNF the highest expression of VASN is in osteoblasts at day 5, 14, and 21. In the human protein atlas the strongest signal was found in the salivary gland, but the antibody showed modest expression in most of the other tissues excluding smooth muscle, cerebellum, reproductive and sexual organs. The protein atlas results do not match the findings of Ikeda et al., (et al in italics always) which show the expression of the protein in vascular smooth muscle cells. Therefore one can only speculate about the origin of the protein and its function. Maybe, its regulation observed in our experiments is caused by changes in the vasculature in the context to the higher incidences of cardiac diseases in relation to obesity and T2D, which would lead into the same direction as MFAP4 and VASN.

Desmoglein 2 (DSG2) is according to GNF SymAtlas (Su et al., 2004) highly expressed in bronchial epithelial cells, whereas the human protein atlas shows a broad expression of the protein excluding brain and liver. It was recently shown to be a substrate of kallikrein 7 involved in pancreatic cancer (Ramani et al., 2008). In general DSG2 is known to be a calcium-binding transmembrane glycoprotein component of desmosomes. Desmosomes are cell-cell junctions between for example epithelia, and myocardial cells (Garrod et al., 2008). Why DSG2 is significantly down-regulated in the obese animals of both models remains speculative. One could think of an involvement in cardiac diseases as a missense variant in desmoglein-2 predisposes to dilated cardiomyopathy (Posch et al., 2008), as we discussed it for the other 3 potential biomarkers. This is also supported by its high expression in cardiac tissue, making it an attractive Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy (ARVD/C) candidate gene. Furthermore most of the publications about DSG2 indicate a strong impact on cardiac function and a relation to cardiac diseases (Awad et al., 2008).

As final remark we can conclude that we were not able to identify a specific biomarker for the progression of T2D or the amount of remaining functional beta-cells in the pancreas. However, we successfully demonstrated that the approaches we applied are feasible tools to identify potential biomarkers in mouse models of disease, since we identified 4 potential biomarkers for the progression of the metabolic syndrome. Even though we are not able to define the origin or cause of the protein abundance changes in the blood samples we observed in our experiments, we definitely found biomarkers for the state of the distinct disease models we investigated. However, the proteins we identified do not necessarily reflect the state of
T2D, but could also be markers for other systemic defects caused by the obesity of the mice, for example cardiac or arterial related diseases.
7. Material and Methods

7.1. Animal models
The ob/ob mice (B6.V/BomTac-Lep ob, obtained from Taconic) were maintained at the RC-HCI, ETH-Zurich in a specific pathogen-free facility. Male C57BL/6J mice (obtained from Elevage Janvier) were fed either with normal or high fat chow (Research Diets, D12331) starting at the age of 5 weeks, one week after arrival in the facility. The maintenance as well as all experimentation was in accordance with the Swiss Federal Veterinary Office (BVET) guidelines and approved by the Veterinary Office of Zürich (Switzerland).

7.2. Mouse blood plasma sampling
After 12 hours of food deprivation the animals were euthanized and blood samples were taken with a 25G needle connected to a 1 mL syringe (both heparinized) using cardiac puncture. Blood samples were centrifuged with 3000 x g for 5 minutes at 4°C. The supernatant was transferred into a new tube and the centrifugation step was repeated. The remaining supernatant was stored in aliquots for further use at -80°C.

7.3. Mouse blood plasma analysis
The blood glucose levels were determined after overnight food deprivation from a blood droplet derived from the tail vein using the Roche Accu-Chek® Aviva glucometer. All other standard blood plasma parameters (insulin, total cholesterol, fructosamine, triglycerides) were commissioned to Hoffmann-La Roche, Basel, Switzerland.

7.4. Intraperitoneal Glucose Tolerance Testing (IPGTT)
After 12 hours food deprivation 1.0 g D-glucose (in 0.9% NaCl solution) per kg of body weight was injected i.p., blood glucose levels were determined using a Roche Accu-Chek® Aviva glucometer at 0, 15, 30, 60, 90, and 120 min post injection.

7.5. Mass Spectrometry
The LC-MS and LC-MS/MS analyses to characterize the N-glycosite samples were performed on a hybrid linear ion trap FT-ICR instrument (Thermo, Bremen, Germany). Data were processed and visualized using in-house software Pep3D (Li et al., 2004). In parallel, peptide identification was performed on the linear ion trap instrument in a data-dependent mode (selection of the three most abundant peaks). The identification was performed using Bioworks (version 3.2.) using the International Protein Index (IPI) human protein database.
All the MRM experiments described were carried out on a hybrid quadrupole/linear ion trap mass spectrometer (4000 QTrap, Applied Biosystems/MDS Sciex, Concord, Canada) using a nanoelectrospray source. The spectrometer is capable of operating the final quadrupole as either a conventional transmission radiofrequency/direct current resolving quadrupole mass filter or as an axial ejection linear ion trap mass spectrometer. The instrument was operated in triple quadrupole mode with Q1 set on the specific precursor m/z value (Q1 is not scanning), and Q3 was set to the specific m/z value corresponding to a specific fragment of that peptide. The LINAC collision cell minimizes the cross-talking effect in very fast MRM experiment (up to 5 ms per transition). During a single reaction monitoring measurement one precursor ion was mass-selected by the first quadrupole being set in such a way that ions were transmitted within a narrow window (typically 0.7 Da). The precursor ion of interest underwent fragmentation in the collision cell generating product ions. The collision energy was tuned to optimize the intensity of the fragment ions of interest, one of which was selected and transmitted through the second analyzer (set with a mass window of 0.7 Da). In the multiple reaction monitoring mode, a series of single reactions (precursor/fragment ion transitions) were measured sequentially, and the cycle (typically 1–2 s) was looped throughout the entire time of the HPLC separation. MRM transitions were determined from the MS/MS spectra of the existing peptides. Typically doubly charged precursors (or triply charged in some instances) were selected. Two transitions per peptide (corresponding to high intensity fragment ions) were selected, and the collision energy was optimized to maximize signal strength. The instrument parameters, e.g. declustering potential and collision energy, can be optimized automatically using automation software to get maximal intensity of MRM transitions. For the scheduled MRM experiments a version of the Analyst software (version 1.4.1) was used.

7.6. LC-MS/MS-Analysis

All shotgun analyses were carried out on a high performance LTQ-FT-ICR mass spectrometer equipped with a nanoelectrospray ion source (both Thermo Electron, Bremen, Germany) coupled online to a Tempo 1D-plus nanoLC (Applied Biosystems/MDS Sciex, Foster City, CA). Peptides were separated on a RP-LC column (75 μm x 15 cm) packed in-house with C18 resin (Magic C18 AQ 3 μm; Michrom BioResources, Auburn, CA, USA) using a linear gradient from 98% solvent A (98% water, 2% acetonitrile, 0.15% formic acid) and 2% solvent B (98% acetonitrile, 2% water, 0.15% formic acid) to 30% solvent B over 60 or 90 minutes at a flow rate of 0.3 μL/min. In DDA mode, each MS1 scan (acquired in the ICR cell) was followed by collision induced dissociation (CID, acquired in the LTQ part) of the three (for
quantification) and five (for identification) most abundant precursor ions with dynamic exclusion for 30 seconds. Only MS1 signals exceeding 150 counts were allowed to trigger MS2 scans. Total cycle time was approximately 1 s. For MS1, 106 ions were accumulated in the ICR cell over a maximum time of 500 ms and scanned at a resolution of 100’000 FWHM (at 400 m/z). MS2 spectra were acquired using the normal scan mode, a target setting of 104 ions and accumulation time of 250ms. Single charged ions and ions with unassigned charge state were excluded from triggering MS2 events. The normalized collision energy was set to 32%, and one microscan was acquired for each spectrum.

7.7. Database search

MS2 spectra were searched against the mouse/human IPI protein database (v3.26, 66,440/53,474 entries) that also contained the protein sequence of pig trypsin and human keratins using the SEQUEST search algorithm (Yates et al., 1995). The database search parameters were set as follows: semi-tryptic specificity was required (cleavage after lysine or arginine residues at one peptide termini, unless followed by proline); 2 missed cleavages were allowed; carbamidomethylation (C) was set as fixed modification; oxidation (M) and, if required, deamidation (N) were applied as variable modifications; mass tolerance of the precursor ion and the fragment ions was 15 ppm and 0.8 Da, respectively. A statistical analysis of the identified peptides was performed through the mass spectrometry Transproteomic Pipeline 3.5 (TPP) (Keller et al., 2005), a uniform proteomics MS/MS analysis platform utilizing open XML file formats (Keller et al., 2005). Here, the database search results were validated using the PeptideProphet software (Keller et al., 2002a), which uses various SEQUEST scores and a number of other parameters to calculate a probability score for each identified peptide. N-glycosylation motif information and accurate mass binning were used in PeptideProphet. The peptides were then assigned for protein identification using the ProteinProphet software (Nesvizhskii et al., 2003). ProteinProphet allowed filtering of large-scale data sets with assessment of predictable sensitivity and false positive identification error rates. In this study, we used a PeptideProphet probability score ≥0.9, and a ProteinProphet probability score ≥0.9. Additionally, these peptides were filtered to remove non N-X-S/T motif-containing peptides. Since it is known that the majority of N-linked glycosylation occurs at a consensus N-X-S/T sequence (where X is any amino acid except proline) (Bause, 1983). This resulted in an overall false discovery rate of less than 1.0%.
7.8. Cell Surface Capturing (CSC)

Cell Surface Capturing was performed as described by B. Wollscheid et al. (Wollscheid et al., 2009) MIN6-B1 cells were grown in 15 cm dishes until ~80% confluency, starved for 4 h in D-MEM containing 2.5 mM D-glucose, supplemented with 15% fetal calf serum (FCS, Brunschwig), sodium pyruvate, 71 mM 2-mercaptoethanol, 2 mM L-glutamine, 100 U/mL penicillin and 100 mg/L streptomycin. MIN6 cells were harvested by spinning them down at 450 rcf for 5 min in a centrifuge. After 4 h of starvation 5x10⁷ cells were stimulated with D-MEM culturing medium containing either 25 mM D-glucose, 20 nM GLP-1 plus 25 mM D-glucose or the exocytoses cocktail. After defined time points cells were harvested at 4°C with precooled solutions as following: cells were washed with labeling buffer (PBS pH 6.4, 0.1% FCS), subsequently incubated with a (0.61 mM) sodium meta periodate (Pierce) solution in labeling in the dark at 4°C for 15 minutes and washed again with labeling buffer, followed by an incubation with a 2.5 mM biocytin hydrazide (Bioticum) solution in labeling buffer at 4°C for 1 hour on a horizontal shaker.

Cells were harvested, pelleted with 800 rpm at 4°C and resuspended in detergent free, hypotonic lysis buffer containing Complete Protease Inhibitor Cocktail EDTA free (Roche, Switzerland) and transferred to a tight dounce homogenizer. After swelling on ice for 10 minutes, cells were homogenized with 30 strokes. Lysate was centrifuged with 2,800 x g at 4°C for 10 minutes and repeated for the resulting supernatant. The pellet of the following ultra centrifugation with 100,000 x g at 4°C for 1 hour was dissolved in 0.1 M ammoniumbicarbonate (ABC) buffer containing 0.2% of RapiGest (Waters) using ultrasonication. After adding TCEP to a final concentration of 5 mM, disulfide bonds were first cleaved at 60°C for 30 minutes and then alkylated for 30 minutes by addition of 10 µl of an aqueous 0.4 M iodacetamide solution. Proteins were proteolysed at 37°C over night using trypsin at an enzyme-to-substrate ratio (w/w) of 1/50 (Promega) followed by enzyme inactivation using Complete Protease Inhibitor Cocktail EDTA free (Roche, Switzerland). N-glycosites were allowed to bind to streptavidin beads (Pierce) for 1 hour and washed extensively with 10 ml of the following aqueous solutions; 5 M sodium chloride, PBS containing 0.5 % Triton, 0.1 M sodium carbonate and 0.1 M ABC. Finally, N-glycosites were cleaved of the beads by endoglycosidase PNGase F (10,000 units, New England) at 37°C over night and desalted using reverse phase chromatography (microspin columns, Havard). The dried peptides were dissolved in 15 µl aqueous 0.1 % trifluor acidic acid (TFA) and 4 µl injected per LC-MS/MS analysis.
7.9. Whole Cell Capturing (WCC)

Capturing of N-glycosites was carried out as recently specified. In brief, MIN6 cells were grown according to the previous method (CSC), washed twice with labeling buffer, harvested and pelleted at 800xg at 4°C. Aliquoted cell pellets, each consisting of 3x10^7 cells, were dissolved using ultrasonication in 200 µl of either buffer: (i) 8 M urea/Rapigest/0.1 M ammonium bicarbonate (ABC); (ii) 6M guanidinium-HCl (Gua)/0.1 M ABC or (iii) 50% trifluorethanol (TFE)/0.1% Rapigest/0.1 M ABC. The total protein concentration of each protein mixture was determined by BCA assay (LIT). After cleaving and alkylating the disulfide bonds as described above, the mixtures were diluted with 0.1 M ABC buffer to: (i) 2 M urea, (ii) 0.5 M Gua and (iii) 10% TFE and subjected to enzymatic cleavage at 37°C over night using trypsin at a substrate-to-enzyme ratio (w/w) of 50 to 1. The sample solutions were acidified to pH <2 using 5% TFA and centrifuged at 10,000g for 15 minutes. Subsequently, peptides were extracted from the supernatant by reverse phase chromatography (RP-LC) (Waters, columns), oxidized in a 10 mM sodium periodate/0.1 M sodium acetate buffer (pH 4-5) for 15 minutes in the dark at 4°C, again extracted by RP-LC and coupled to hydrazine beads (Biorad) over night using a over head shaker. Non-bound peptides present in the supernatant were collected and the N-glycosites containing beads washed several times with 5 ml of the following aqueous solutions: 5 M sodium chloride, 80 % acetonitrile, pure methanol, water and 0.1 M ABC buffer. Finally, N-linked glycopeptides were removed from the beads and collected after incubation with 100,000 units of endoglycosidase PNGase F at 37°C over night. The released peptides were once more RP-LC purified (Microspin columns, Harvard) before 1/3 of the sample was subjected to each LC-MS/MS run.

7.10. Directed LC-MS/MS

To increase the number of identified N-glycoproteins, the recently developed directed LC-MS/MS approach (Schmidt et al., 2008) was applied to extensively map out all detectable N-glycosites extracted from MIN6 cells or human islets. In brief, the data of two initial LC-MS (mapping) runs was converted to the profile mzXML format (Pedrioli et al., 2004) and loaded into the in-house developed software system SuperHirn (Mueller et al., 2007) to detect, integrate and align detected features over multiple LC/MS patterns. Peak intensities were determined by calculating peak areas from extracted ion chromatograms (XICs) of each MS signal. Highly stringent criteria were applied to filter the detected peaks for peptide signals. Only peaks that could be found in both LC-MS runs were considered and singly charged masses as well as features for which a peptide sequence could be assigned from the initial
three LC-MS/MS runs were excluded. Finally, a list of the relevant features was generated and used to build scheduled mass inclusion lists for directed MS-sequencing according to (Schmidt et al., 2008). For directed sequencing, the same MS settings as mentioned above were applied with the following changes: The preview mode option was disabled and the MS1 resolution reduced to 50,000 FWHM (at 400 m/z) and a mass window of ± 10 ppm was used for precursor mass selection. The MS2 data obtained was database searched and mapped back to the initial master map using the SuperHirn (Mueller et al., 2007) software until all detected features were MS sequenced.

7.11. Label-free Quantification

First, a list of proteotypic peptides (PTPs, peptides that have a unique sequence in genome, represent full tryptic cleavage and are therefore suited for accurate protein quantification (Kuster et al., 2005), was generated for each sample and used for targeted inclusion list sequencing. Consequently, more than 700 (500) PTPs could be extracted from the MIN6 surface (human islet) N-glycosite catalogue allowing the consistent screening of more than 300 (300) surface proteins in two MS runs. After screening each individual sample for the PTPs using the same MS settings as for directed LC-MS/MS and database searching, peptide peak extraction, integration, annotation and alignment over all runs of a particular stimulation experiment was done by the SuperHirn (Mueller et al., 2007) software. Subsequently, data filtering, protein ratio and statistic p-value (ANOVA) calculations as well as visual data display were carried out using Spotfire DecisionSite (TIBCO Spotfire, Europe).

7.12. N-glycosite blood sample processing

Samples were processed using the N-glycopeptide isolation protocol as previously described (Zhang et al., 2003), starting with 50 µl of mouse plasma. The resulting mixture of N-glycosite peptides was dissolved in 40 µl of solvent (5% aqueous acetonitrile) prior to LC-MS analysis. Typically 1-2 µl of the solution was used for a single LC-MS run. Briefly blood serum samples (50 µl, containing approximately 4 mg of total protein), human plasma, or serum pools (Sigma-Aldrich) were treated with sodium periodate to oxidize the diol groups of the glycan residues attached to the protein backbone. The coupling to a solid support occurred by reaction of the aldehyde groups thus formed with functionalized hydrazide beads. The immobilized glycoproteins were extensively washed to remove non-covalently bound molecules, and disulfide groups were reduced and alkylated with iodoacetamide on the solid support. The sample was subsequently treated with trypsin, and non-glycosylated peptides were removed by washing. Finally N-glycosites were released by treatment with peptide-N-
glycosidase F, an enzyme that cleaves the covalent bond at the carbohydrate attachment site, thus releasing deglycosylated peptides from the solid support while leaving the glycan moiety ligated to the solid phase. In the process of the enzyme action the asparagine residue of the NX(S/T) consensus sequence is converted into an aspartic acid. The resulting peptide mixture was then analyzed by LC-MS using a reversed phase column.

7.13. Reference Peptides
Peptides used as internal standards were chemically synthesized (Cell Signaling Technology Inc., Danvers, MA and Sigma-Aldrich). They were isotopically labeled by incorporating 15N/13C in one amino acid residue, typically Val, Phe, Arg, or Leu, near the C terminus. The peptide sequences can be found under at http://biodata.ethz.ch/aquapeptides.txt as a plain-text file. Concentrations of the synthetic peptide stock solutions were established by independent amino acid analysis.

7.14. Peptide Separation
HPLC separations were performed either on a Dionex/LC Packings or TempoTM Nano HPLC system (Applied Biosystems/MDS Sciex, Foster City, CA). A C18 capillary column with 75-m inner diameter was used at a flow rate of 200–300 nl/min. A gradient from 2 to 62% acetonitrile (containing 0.1% formic acid) over a 60-min period was used.

7.15. Off-gel electrophoresis
About 1 mg of unbound peptides obtained from both, MIN6 cells and human islets WCC experiments were separated into 24 individual fractions using off-gel electrophoresis applying the following parameters: strip 24 cm, 3-10 pH, 8’000 volts, maximal power 100mW and current 50mA for a total length of 30kVh. Fraction 11-13, 14-16, 17-19, 20-22 were pooled. Thus, the final fraction number reduced to 16. After RP-LC clean up (Macrospin columns, Harvard), 2% of each fraction were LC-MS/MS analyzed.

7.16. Chemicals
Porcine trypsin, modified, sequencing grade, was purchased from Promega (Madison, WI, USA). Tris(2-carboxyethyl)phosphine (TCEP) and iodoacetamide were purchased from Fluka (Buchs, Switzerland). HPLC-grade water and acetonitrile were purchased from Riedel-de Haën (Seelze, Germany), sodium periodate (Pierce), biocytin hydrazide (Biotium, USA), UltraLink immobilized Streptavidin PLUS (Pierce), 15 ml Dounce Tissue Grinder (Wheaton, USA), RapiGest (Waters), PNGase F (NEB), Affi-Prep Hz Hydrazide (BioRad)
7.17. Cell lines and culture conditions
MIN6-B1 cells (Lilla et al., 2003) were cultured in D-MEM (Invitrogen) containing 25 mM glucose, supplemented with 15% fetal calf serum (FCS, Brunswig), sodium pyruvate, 71 mM 2-mercaptoethanol, 2 mM L-glutamine, 100 U/ml penicillin and 100 mg/L streptomycin. Cells were cultured in 15 cm dishes (Gibco) at 37°C with 5% CO2 and grown till ~80% confluency before the experiment. Starvation of cells was performed with D-MEM culturing medium with same supplementation as above, but containing only 2.5 mM glucose.

7.18. Human Pancreatic Islets
The human pancreatic islets were obtained from a collaboration with the University Hospital Zurich. Two batches, each containing 25’000 Islet equivalents (IE) were provided. One donor was female, 67 years old and had a BMI of 24.7. The second donor was female, 37 years old, with a BMI of 23.2. Human pancreatic islets were isolated at the University Hospital Geneva, Switzerland using the two-layer method, as described (Lehmann et al., 2004; Ricordi et al., 1988). Isolated islets were then cultured for 72 h in CMRL 1066 culture medium supplemented with 10% FCS, also as described (reference). The batches of 25 000 IEs were washed with PBS and centrifuged at 800 x g at 4 °C for 5 minutes. After removal of the supernatant, pellets were frozen at -80 °C. Islets were thawed up on ice and processed according to the Whole Cell Capturing (WCC) method described below.

7.19. Insulin secretion assay and insulin content
MIN6 cells (2x10^7 cells/dish) were seeded in 24-well plates 2 days before use. Cells were washed three times with a modified Krebs-Ringer bicarbonate HEPES buffer [KRBH; 125 mM NaCl, 4.74 mM KCl, 1 mM CaCl_2, 1.2 mM KH_2PO_4, 1.2 mM MgSO_4, 5 mM NaHCO_3, 25 mM HEPES (pH 7.4) and 0.1% BSA] supplemented with 2.5 mM glucose and preincubated with this same buffer for 4 h at 37°C. Cells were then incubated for 1 h at 37°C with KRBH containing 2.5 mM glucose, followed by 1 h at 37°C with KRBH containing either 2.5 mM D-glucose, 25 mM D-glucose, 20 nM GLP-1 plus 25 mM D-glucose or the exocytosis cocktail (25 mM D-glucose, 40 mM PMA, 20 mM forskolin, 20 mM arginine). Insulin was measured in 96 well mouse insulin ELISA (Mercodia, Sweden).

7.20. Q-PCR
RNA samples from MIN6 B1 cells were analyzed by quantitative real-time PCR. Reactions were carried out in a volume of 10 μL consisting of 1 μL cDNA, 1x LightCycler enzyme and reaction mix (SYBR Green I dye, TaqDNA polymerase, dNTP; Roche), 1.5 mmol/L MgCl_2,
and 600 nmol oligonucleotide primers. All reactions were performed in a LightCycler (Roche) in which samples underwent 40 cycles of PCR with an annealing temperature of 55°C. The used primers were purchased from Microsynth and designed according to the suggestions derived from the PCR primer bank for quantitative gene expression analysis (Wang and Seed, 2003) and are listed under 7.23. The value obtained for each specific product was normalized to a control gene (18S) and expressed as fold change in relation to the control.

### 7.21. Bioinformatics analysis

The prediction of transmembrane domains of the proteins in the catalogue was performed using the Phobius 1.01 combined transmembrane topology and signal peptide predictor (Kall et al., 2007) and stored a MySQL database. TargetP 1.1 (Emanuelsson et al., 2007) was used to predicts the sub-cellular location and GPI-SOM (Fankhauser and Maser, 2005) predicted GPI-anchoring signal sequences. The IPI protein identifiers provided by the MS analysis pipeline were mapped to sequence and annotation using version 3.43 (30 Apr 2008) of the IPI database (Kersey et al., 2004) downloaded by FTP from the European Bioinformatics Institute. Functional annotation was performed using Release 46.0 (April 1, 2008) of the KEGG ORTHOLOGY (KO) database (Kanehisa et al., 2008; Kanehisa and Goto, 2000). The protein list from KEGG GENES as well as the KO list were downloaded by FTP. Our protein list was matched to KEGG using the BLAST (Altschul et al., 1990) local alignment tool. Graphical data displays were produced using Ploticus software (ploticus.sf.net). Entrez identifier for human and mouse were determined by using the Entrez (Schuler et al., 1996) Programming Utilities (eUtils) on the current version (April 2008) of their database. Orthologous proteins in human and mouse were identified by using the results (p<10^{-10}) of a BLAST search. Automated analysis of the data set using the above mentioned tools was performed using the Python (Sanner, 1999) programming language.

### 7.22. Primer pairs

<table>
<thead>
<tr>
<th>GeneID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLP1R MOUSE (FWD)</td>
<td>ACG GTG TCC CTC TCA GAG AC</td>
</tr>
<tr>
<td>GLP1R MOUSE (REV)</td>
<td>ATC AAA GGT CCG GTT GCA GAA</td>
</tr>
<tr>
<td>ADAR2A MOUSE (FWD)</td>
<td>GTG ACA CTG ACG CTG GTT TG</td>
</tr>
<tr>
<td>ADAR2A MOUSE (REV)</td>
<td>CCA GTA ACC CAT AAC CTC GTT G</td>
</tr>
</tbody>
</table>
Material and Methods

LDLR MOUSE (FWD)  AG TGG CCC CGA ATC ATT GAC
LDLR MOUSE (REV)  CTA ACT AAA CAC CAG ACA GAG GC

7.23. Statistics
Each stimulation experiment (GLP1, HG) was analyzed separately. For every protein, all peptides, and all measured areas under simulated conditions were compared to control from raw data tables, and are normalized by dividing each by the average area size of the corresponding (same peptide and measurement) control. Control values were also normalized by their average. A Mann Whitney U test was performed for mean deviation between stimulated (normalized) and control area sizes of all peptides and measurements. The resulting p-value and the log10 fold change from the average of all normalized stimulated values were stored in the database.
8. Acknowledgement

At first I would like to thank the people which contributed directly to the work as presented here: Katrin Eckhardt\textsuperscript{1,6,7}, Niklaus Fankhauser\textsuperscript{2,6,7}, Vinzenz Lange\textsuperscript{3,4,7}, Alexander Schmidt\textsuperscript{3,6,7}, Sebastien Ribrioux\textsuperscript{2,6}, Reto Ossala\textsuperscript{3}, Lukas Müller\textsuperscript{3,6}, Ralph Schiess\textsuperscript{3,6}, Richard Züllig\textsuperscript{5}, Roger Lehmann\textsuperscript{5}, Giatgen Spinas\textsuperscript{5}, Xavier Gueripel\textsuperscript{4,6}, Holger Moch\textsuperscript{4,6}, Bernd Wollscheid\textsuperscript{3}, Bruno Domon\textsuperscript{2,6}, Rudolf Aebersold\textsuperscript{3,6} and Wilhelm Krek\textsuperscript{1,6}

\textsuperscript{1}Institute of Cell Biology, ETH Zurich, 8093 Zurich, Switzerland; \textsuperscript{2}Institute of Biochemistry, ETH Zurich, Schafmattstrasse 18, CH-8093 Zurich, Switzerland; \textsuperscript{3}Institute of Molecular Systems Biology, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland; \textsuperscript{4}Department of Pathology, Institute of Surgical Pathology, University Hospital Zurich, Schoenbeinstrasse 12, CH-8091 Zurich, Switzerland; \textsuperscript{5}Division of Endocrinology and Diabetes, University Hospital Zurich, CH-8091 Zurich, Switzerland; \textsuperscript{6}Competence Center for Systems Physiology and Metabolic Diseases, Schafmattstrasse 18, CH-8093 Zurich, Switzerland; \textsuperscript{7}This authors contributed equally to the success of this project.

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me. Some of them I would like to mention in particular here:

Above all, I would like to thank my grandparents and parents, which enabled me to find my way to reach this point. I am grateful for their personal and financial support as well as the great unequivocal support throughout, for which my mere expression of thanks likewise does not suffice.

A great thanks goes to my supervisor and mentor Prof. Wilhelm Krek for giving me the opportunity to work on this project. I am grateful for his critical inputs, his guidance, but also the freedom he provided me working in his lab and on the project. I always appreciated all the advices I got and his mentorship during my time in the lab.

I also would like to thank Claes Wollheim for his support during the course of the project, especially as one of my thesis committee member. I am grateful for his critical inputs and his advices, which also contributed to the success of the project.

As member of my thesis committee I would like to thank Ruedi Abersold for his critical inputs in several meetings during the course of the project, and for bringing me back to the ground, when I had to ambitious visions.
As my closest co-worker on the project I would like give a special thanks to Katrin Eckhardt. This collaboration enabled the success of the project. I am grateful for all the fruitful discussions we had in the lab, in meetings, and also while having some beers.

Niklaus Fankhauser I would like to thank for his bioinformatics skills, all his efforts with setting up the database, integrating all the different datasets, and for being such a great colleague.

For the mass spectrometer analysis I would like to thank Vinzenz Lange and Alexander Schmidt. They were great collaborators and fruitful participants in sometimes endless seeming meetings about the next steps of the project.

As my longest and closest colleague during my PhD time, I would like to thank Jens Zehetner for his openness, the interesting discussions, all the undertakings during the last four years and the resulting friendship.

I would like to thank Sreya Basu for her company during the last 4 years. Even though I did not talk to her at the very beginning we became friends at one point, and I would like to thank her for all the discussions about life, and other things, especially during these evenings we spend together with cooking and tasting wine.

Claudia Baltzer I would like to thank for her partnership during the last couple of years, and now even sharing one room with me. We went for wine tasting, we went running, we went climbing, we did a lot of exciting things together. Finally she made my life very rich and I would like to thank her for sharing all this with me.

While being a technician in our lab, I would like to thank Manuela Hitz for introducing me into the proper cleaning procedure of surgical instruments. In private life I would like to thank her, and also Benjamin, for their friendship and support during the last years. Together we crawled in quite some holes, which strengthened our relationship, and made us experience many exciting adventures.

A big thanks goes to Cristiano Migliorini for guiding the beta-cell SystemsX project from the side of F. Hoffmann-La Roche, setting the milestones, putting pressure when it was necessary and showing me how to be a good project manager. I also would like to thank him for all the interesting discussions, and the support I got from him also after the project has ended.
For taking care of my mice I would like to thank Corina Fusaro, Susanne Freedrich, Jörg Fehr, Urs Egli and Thomas Weber. It was really a pleasure to work together with all of you, even though sometimes there were some discrepancies, which we could always solve in a for all satisfying way.

The “reviewer special thanks” goes to Claudia Baltzer, Divya Vats, Sreya Basu, and Nick Fankhauser for critical reading this thesis.

I would like to thank all members of the F-floor groups, which accompanied me during my time in the lab and provided a great atmosphere.

The PhD program of the Competence Center for Systems Physiology and Metabolic Diseases, especially Urte Hinrichs and Davina Rogers, I would like to thank for all the support and offering interesting complementary courses during my time as PhD student.

For any errors or inadequacies that may remain in this work, of course, the responsibility is entirely my own.
9. References:


References


References


References


Nakamura, K., F. Abarzua, A. Hongo, J. Kodama, Y. Nasu, H. Kumon, and Y. Hiramatsu. 2009. Hepatocyte growth factor activator inhibitor-2 (HAI-2) is a favorable prognosis


References

... tumor suppressive activity of the netrin-1 receptors. *Gastroenterology.* 135:1248-1257.


115


10. Curriculum vitae

Education

07/2005-08/2009  PhD Thesis, Institute of Cell Biology, ETH Zurich, Switzerland

- Set up and established methods for the study of Type 2 Diabetes utilizing in vitro and in vivo models
- Explored N-glyco proteome of the murine beta cell line MIN6 and human pancreatic islets
- Quantified changes of N-glyco surface proteins in MIN6 cells upon different stimuli
- Performed mouse studies based on achieved results from in vitro experiments to test potential blood plasma markers for the status of beta cells in disease states
- Member of a collaboration between F. Hoffmann-La Roche, ETH Zurich and University Zurich entitled ‘Systems biology of the beta cell - application to type 2 diabetes progression’

10/1999-01/2005  Biology Studies and Diploma Thesis, Institute of Cell Biology, University of Konstanz, Germany

- Mined raw genome wide database of *P. tetraurelia* for novel homologues of human Syntaxins
- Discovered 19 novel genes, which gave the bases for a publication and funding of a new research project
- Established microinjection into the nucleus of *P. tetraurelia*

11/1998-02/1999  Paramedic Apprentice (Rettungssanitäter)

- Acquired skills in advanced emergency measures
- Managed emergency situations
Curriculum vitae

Professional experience

07/2005-08/2009  Teaching at the Institute of Cell Biology, ETH Zurich, Switzerland

- Lectures in Laboratory courses in Basic Biology and Cell Biology/Molecular Medicine to over 60 students
- Mentored 3 students in laboratory work

03/2006-08/2009  Manager of the rodent facility - Krek group

- Managed and organized transgenic mouse lines of the Krek laboratory (~ 2000 mice)
- Coordinated the relationship between the head of the rodent centre, animal technicians, and scientists.
- Implemented Good Laboratory Practice

06/2006-08/2009  IT - support Krek group

- Maintained hardware and software
- Resolved computer-related problems

03/2000-03/2004  Research assistant at the University of Konstanz, Germany

- Organized biological field work and processed field samples in the laboratory
- Worked with molecular biological, microbiological and cell biological techniques

09/1998-09/1999  Paramedic at DRK ambulance station Friedrichshafen, Germany

- Dealt with emergency situations
- Coordinated the emergency site and operations in the ambulance vehicle
- Responsibility for one emergency assistant
Extracurricular activities

09/2007-06-2009  Founding member and CFO of the NPO “Aiducation International Schweiz”
- Co-established the NPO following a start-up company strategy
- Supported 30 bright but financially distressed secondary school students in Kenya
- Managed a financial budget of ~50'000 CHF

15/03/2007  Head of organizing committee for the PhD Minisymposium, CC-SPMD, Zurich, Switzerland
- “Keeping the Balance – New Insights into Energy Homeostasis”
- Organized symposium with over 100 participants
- Invited 4 international key opinion leaders as speakers
- Managed a financial budget of 12’000 CHF

Professional and advanced training

08-09/2009  Basic Management Skills for Management, Technology, and Economics Master
CO3 AG | Corporate Coaching & Consulting, ETH Zurich, Switzerland

02-04/2008  Entrepreneurship Course “venture challenge”
Innovative Promotion Agency CTI, Zurich, Switzerland

03/2007  Systems Biology of Complex Diseases: Medical Science
CC-SPMD, ETH Zurich, Switzerland

12/2007  Ethics and Technology Transfer Course
CC-SPMD, ETH Zurich, Switzerland

09/2007  CC-SPMD Systems Biology Introductory Course
CC-SPMD, ETH Zurich, Switzerland

07/2006  Proteomics Block Course
IMSB, ETH Zurich, Switzerland
Curriculum vitae

05-06/2006  Current Problems in Protein Analysis
Functional Genomics Center Zurich, Switzerland

02/2006     Proteomics Informatics Course
Institute of Systems Biology, Seattle, USA

10/2005     LTK 1 Lab Animal Course
Institute of Lab Animal Science, University of Zurich, Switzerland

Languages

German     Fluent (mother tongue)
English    Fluent (working language)
French     High school level
Spanish    Level A1.1

IT

- MS Windows OS, Mac OSX, MS Office package, Image processing
- Database mining
- Hardware and Software maintenance

Further personal information

Passionate about photography, travel, cooking as well as mountain and outdoor sports

Born September 21st, 1977 in Ravensburg, Germany
Curriculum vitae

Publications


Presentations


Danzer, C., ’Systems biology of pancreatic beta cells - from cell surface to functionality’. CC-SPMD Medical Science Course; ETH Zurich, Switzerland


Danzer, C., Krek, W. ‘Cell Surface N-Glyco-Proteome of MIN6 cell: a systemic approach’. ICB Lecture Series 2007; ETH Zurich, Switzerland


Danzer, C., Krek, W. ‘Changes in the plasma membrane proteome of pancreatic beta-cells during development of Type 2 Diabetes’. ICB Lecture Series 2006; ETH Zurich, Switzerland
Patents

Biomarker_01 - a novel plasma marker candidate for Type 2 Diabetes
Authors: Alexander Schmidt, Carsten Danzer, Vinzenz Lange, Reto Ossola, Wilhelm Krek and Ruedi Aebersold

Biomarker_02 - a novel plasma marker candidate for Type 2 Diabetes
Authors: Alexander Schmidt, Carsten Danzer, Vinzenz Lange, Reto Ossola, Wilhelm Krek and Ruedi Aebersold

Biomarker_03 - a novel plasma marker candidate for Type 2 Diabetes
Authors: Alexander Schmidt, Carsten Danzer, Vinzenz Lange, Reto Ossola, Wilhelm Krek and Ruedi Aebersold

Biomarker_04 - a novel plasma marker candidate for Type 2 Diabetes
Authors: Alexander Schmidt, Carsten Danzer, Vinzenz Lange, Reto Ossola, Wilhelm Krek and Ruedi Aebersold

Posters


Danzer, C., Lange V., Schmidt, A., Fankhauser, N., Domon, B., Aebersold, R., Krek, W. ‘MRM analysis with blood plasma samples of Type 2 Diabetes mouse models’. ETH Zurich, D-BIOL PhD Summerschool 2007; Stels, Switzerland


References

Prof. Wilhelm Krek, Institute of Cell Biology, Phone: +41 44 633 34 47, Fax: +41 44 633 13 57, email: wilhelm.krek@cell.biol.ethz.ch

Dr. Cristiano Migliorini, F. Hoffmann-La Roche, Phone: +41 61 688 6047, Fax: +41 61 688 1720 email: cristiano.migliorini@roche.com

Prof. Claes Wollheim, CMU/ Dpt PHYM, Phone: +41 22 379 55 48, Fax: +41 22 379 55 43 email: claes.wollheim@medecine.unige.ch

Prof. Ruedi Aebersold, Institute of Molecular Systems Biology, Phone: +41 44 633 31 70, Fax: +41 44 633 10 51, email: aebersold@imsb.biol.ethz.ch