"This is an Accepted Manuscript of an article published by Taylor & Francis in Expert Review of Proteomics *on 20/12/2016, available online: http://www.tandfonline.com/10.1080/14789450.2017.1265449."*

A review article

'The progress and potential of proteomic biomarkers for type 1 diabetes in children'

Robert Moulder[1], Santosh D. Bhosale[1], Riitta Lahesmaa[1], David R. Goodlett[1,2]
[1]Turku Centre for Biotechnology, University of Turku, Turku Finland
[2]School of Pharmacy, University of Maryland, Baltimore, MD USA

Robert Moulder Ph.D, is a senior scientist at Turku Centre for Biotechnology
Santosh D. Bhosale M.Pharm.is a Ph.D student at Turku Centre for Biotechnology
Riitta Lahesmaa M.D. Ph.D. is an Academy of Finland Professor at Turku Centre for Biotechnology
David R. Goodlett is a Professor in Pharmaceutical Sciences at University of Maryland's School of Pharmacy in Baltimore

Abstract

Introduction: Although it is possible to identify the genetic risk for type 1 diabetes (T1D), it is not possible to predict who will develop the disease. New biomarkers are needed that would help understand the mechanisms of disease onset and when to administer targeted therapies and interventions.

Areas Covered: An overview is presented of international study efforts towards understanding the cause of T1D, including the collection of several extensive temporal sample series that follow the development of T1D in at risk children. The results of the proteomics analysis of these material are presented, which have included bodily fluids, such as serum or plasma and urine, as well as tissue samples from the pancreas.

Expert Commentary: Promising recent reports have indicated detection of early proteomic changes in the serum of patients prior to diagnosis, potentially providing new measures for risk assessment. Similarly, there has been evidence that post-translationally modifications may result in the recognition of islet cell proteins as autoantigens; proteins thus modified

could be used as targets for immunomodulation to overcome the threat of autoimmune response.

Key words:

autoantibodies, beta cells, biobanks, biomarkers, human leukocyte antigen, Islets of Langerhans, type 1 diabetes,

.

1 INTRODUCTION.

1.1 Genetic risk and the Environment

Type 1 diabetes (T1D) is the most common form of diabetes in children and results from autoimmune destruction of the insulin producing β-cells in the pancreas (1)(2). By the time diagnosis is established a large proportion of beta cells have been lost or are no longer functional (3), at which point the body can no longer produce sufficient insulin to maintain blood sugar levels. The condition is fatal unless the patient receives regular supplementation of insulin. Even whilst treated, T1D leads to serious long term complications, which can include blindness and physical disabilities.

The risk for T1D is characterized by a strong genetic component, and has in particular been associated with certain variants of the human leukocyte antigen (*HLA*) gene (4-6). Amongst the *HLA* encoded genes that are associated with immune function, several alleles and haplotypes of the *HLA class II* genes and the *HLA-A* and *-B* alleles have the strongest T1D associations, in particular certain HLA-DR-DQ haplotypes(6). Many other genetic loci have been linked with T1D, amongst which *INS* and *PTPN22* are notable (7).

In addition to genetic factors there is a large body of evidence supporting the influence of the living environment, in particular hygiene, early childhood diet and infections (8). The incidence varies drastically from country to country, with rates less than 1 per 100,000 in parts of South East Asia, in contrast to levels that have reached as much as 60 per 100,000 in Finland (9, 10). In the western world T1D is the most common metabolic-endocrine disorder in children and is particularly common in Northern Europe, with incidence rates typically 15-30 per 100,000 in England, Germany and the Scandinavian countries (9). Notably, the incidence has increased drastically in modern times, as exemplified by a five-fold increase in

Finland during the past seventy years (10, 11). Current estimates suggest an annual global rate of increase of 3 % (12).

Currently the detection of blood-borne autoantibodies against islet cell proteins are the only established indications of the onset of the autoimmune attack that results in this disease (7). The events that lead to the initiation and determine the duration of this pre-disease phase remain to be established and biomarkers are thus required to help understand the etiology of the disease, improve risk assessment and develop potential therapies. As proteins represent the functional molecules in the cell and can be modified in different ways after synthesis, determination of their relative abundances and temporal profiles has the potential to establish markers of disease and provide insight into the changes leading to its onset.

1.2 The Pancreas and the Beta cell loss

Beta cells are located in the Islets of Langerhans in the pancreas, which are small conglomerates of cells dispersed in the exocrine tissue and make about 4.5% of the mass of the pancreas (13). In their recent analysis of the composition and anatomy of an adult pancreas, Ionescu-Tirgoviste and co-workers estimated that it contained in the order of three million islets (13). In addition to the alpha and delta cells, the beta cells constitute 65-80% of the cells in the islets. The beta cells contain insulin secretory granules, which are subcellular organelles that are compacted with insulin (14). Notably, amongst the autoantibodies detected with the onset of T1D related autoimmunity are several that are directed towards components of the insulin secretory granules and autoantibodies against insulin are often the first detected in children (15). Current hypotheses as to why insulin or other islet proteins become recognized as antigens include the formation of fusion peptides (16) and post-translational modification (17).

The destruction of beta cells is associated with infiltration of cytoxic CD8+ cells and macrophages (18-20). Although the precise mechanism of this remains unclear, the following model, based on that described by Crevecoeur *et al.*, is presented to illustrate how a cascade of T cell mediated events would perpetuate this destruction once there is a loss of self-tolerance (21)(22). Metabolic stress, inflammation or viral infection would result in cytokine secretion and activate and attract innate and adaptive immune cells in the islets. This environment could lead beta cells to increase expression of beta cell derived antigens on their major histocompatibility (MHC) class I molecules, cause ER stress and result in the initiation of apoptosis. Subsequent phagocytosis by antigen presenting cells would lead to the presentation of beta cell antigens by MHC class II molecules, which would stimulate the maturation of CD8+ T lymphocytes to cytotoxic CD8+ cells that would recognize beta cell antigens (presented by MHC class I molecule on the beta cell surface), leading to further beta cell death. The cycle is perpetuated by cytokine release and creation of an inflammatory environment resulting in further ER stress, apoptosis, epitope spreading, cell death, progressively reducing the capacity to produce insulin. In the normal situation, self-regulation limits this cascade and prevents the propagation to autoimmunity, whereas in T1D this element of control and self-recognition is somehow lost. Notably the strongest genetic association with T1D is through the human leukocyte antigen genes, which encode class I and II MHC proteins that present peptides derived from the degraded proteins. Consequently, characterization of the peptides that are presented by MHC I and II molecules to the CD4+ and CD8+ T cells has been the focus of T1D research (23)(24).

1.3 Prospective Sampling of TID Risk Cohorts

In view of the frequency and increasing incidence of T1D in specific geographic regions, a number of prospective studies have been established during the past few decades. Typically,

at risk children have been recruited on the basis of genotyping and family history for surveillance and periodic sample collection, e.g. serum and whole blood. These studies include the German BabyDiab study (25), the USA-based Diabetes Auto Immunity Study in the Young (DAISY), the Finnish Diabetes Prediction and Prevention study (DIPP) (26), the Swedish Diabetes Prediction in Skåne (DIPIS) study (27), and The Environmental Determinants of Diabetes in the Young (TEDDY), which has combined sample collections from the USA, Finland, Sweden and Germany (28). Prevention studies have also been included in these projects, comprising the use of nasal and oral insulin, probiotics and hydrolyzed formula milk amongst others (29)(30). T1D associated autoantibodies are measured from the serum samples, including islet cell autoantibodies (ICA) and antibodies against insulin (IAA), glutamic acid decarboxylase 65 (GADA), IA-2 protein (IA-2A) and zinc transporter-8 (ZnT8)(31). Detection of persistently high titers of any of these autoantibodies is regarded as an indication of the onset of autoimmune process, and the appearance of multiple autoantibodies generally precedes progression to T1D. Recent evaluation of data from the major clinical centers has shown that the majority of children that are multiple autoantibody positive will progress to T1D within 15 years. Whilst these autoantibodies have to date been the only true markers of increased risk and T1D onset, recent reports have also shown that other late phase markers are detectable. Research from the DIPP study has recently shown that the first phase insulin response is decreased as early as 4-6 years before diagnosis (32). Addressing HbA1c levels, measurements from the DIPP study have also demonstrated its utility as a biochemical marker for predicting the time to diagnosis of T1D in children with multiple autoantibodies (33).

Although some specific differences exist between the aforementioned prospective studies, the sample collection strategies are quite similar. As an example, we illustrate the sample collection process for the DIPP study. Newborns are screened for T1D risk, and at risk

children/families invited to participate in the study. The children periodically visit the clinic where blood samples are collected every three months for the first two years of life and then every six months until the age of 15. If a child seroconverts, which is defined as appearance of T1D associated autoantibodies, sample collection continues at or reverts to three month intervals. While children diagnosed with T1D may be monitored for a short time, they generally leave the study after diagnosis. The sample collected in these Finnish centers has provided a rich resource for the research towards biomarkers of T1D (34). In addition to proteomics measurements (35), transcriptomics (36)(37), metabolomics (38) and metagenomics (39) analyses of these samples have been made.

Another recent study of note is the Finnish coordinated DIABIMMUNE study, which has addressed the influence of modern and developing living environments in comparison to a more traditional rural setting upon the incidence of T1D and allergy. The study has collected samples from newborns from the towns of Espoo (Finland), Tartu (Estonia) and Petrozavodsk (Russia) during the first three years of life, providing samples from 3, 6, 12, 18, 24, 36 months. So far, the data supports the contrasting rates of T1D incidence between these regions, and implicates the plasticity in the developing immune system towards environmental factors that may begin *in utero* (40-42).

2 PROTEOMIC TARGETS

Although a large body of T1D research has been made using non obese diabetic (NOD) murine and cellular models from other organisms (43), the type of sample material available from humans, and in particular children, is somewhat more restricted. Whilst blood, serum and urine are accessible, pancreatic tissue is only available from autopsies or occasionally from surgery on young adults and is thus very rare. A number of biobanks of pancreatic tissue have, however, now been established (*vide infra*). To address the different tissues that have been used to study T1D in man, this review is divided in terms of these targets, as illustrated in **Figure 1**. The subdivision is further represented in **Table 1**, which include examples of targets of proteomic analyses in T1D research.

Amongst the proteomic research collaborations that have targeted diabetes research, the Human Diabetes Proteome Project, has aimed to establish mass spectrometry and antibody-based methods and collect knowledge in this field. The latter have included extended proteome databases on diabetes-related cells/fluids/tissues, firstly concentrating on the creation of the human islet proteome (44), in addition to establishing lists of proteins relevant to diabetes-related conditions (45).

2.1 Pancreatic Tissue, islets and beta cells

The pancreas has both endocrine and exocrine functions. Most pertinent to the condition of T1D is the loss of the endocrine secretion of insulin that results from the destruction of the beta cells in the Islets of Langerhans in the pancreatic islets. For a detailed overview of proteomic studies of beta cell function in T1D, the reader is referred in particular to the review of Crevecoeur *et al.*, (22).

In order to establish disease related differences in the pancreatic proteome, it has been necessary to establish the nature of the healthy tissue or islet cell. Although much of the earlier work used murine models (46)(22) and 2DE gels to characterize pancreatic cells and tissue, human pancreatic tissue is sometimes available from cadavers or from organ resection during surgical intervention, and the depth and specificity of proteomic coverage has progressively grown as technology has improved. As an example of an earlier study, Hu *et al.* used 2DE and MALDI-TOF to characterize human pancreatic tissue, identifying 302 proteins (47). Similarly, Ahmed and co-workers were first to study the proteome of human islets, and with their 2DE MALDI-TOF/TOF approach they identified 66 islet cell specific proteins (48). Metz *et al*. used a 2D LC-MS/MS approach to characterize the islet proteome of mixed samples from a biobank, identifying in the order of 3,600 proteins (49).

Sample banks of pancreatic tissue have been specifically established from patients with diabetes. Recently, Burch *et al.* used samples from the *Network for Organ Donors with Diabetes* (nPOD), analyzing human pancreatic tissue lysates from organ donors with a range of diagnoses and conditions, including T1D-associated autoantibodies in the absence of diabetes, T1D, T2D and without diabetes (50). The average age of the T1D patients was 22.5 years (n =5), with the youngest 19 years. For this study they used a label free LC-MS/MS approach with a Q-Exactive (quadrupole-Orbitrap) mass spectrometer. They reported the differential abundance of approximately 350 of the 1140 proteins detected, out of which there were 60 significantly upregulated tissue proteins with functional relevance to T1D that have been implicated in inflammation, metabolic regulation, and autoimmunity (**see Table 1**). Liu and coworkers analyzed similar material from nPOD and reported the identification of over 5,000 proteins (51). For these analyses they labelled the samples with 10-plex TMT reagents and performed 2D LC-MS/MS with a Q-Exactive, with the additional fractionation likely accounting for the larger number of detected proteins. They reported that the T1D cadaveric

human pancreata exhibited a unique exocrine tissue proteomic profile and that many of the observed differences were consistent with the reported literature for beta cell destruction and the pathological processes that are believed to take place in the T1D pancreas. In particular, these differences were functionally related to immune response, viral infection and cell apoptosis (**Table 1**).

Notably, much of the data concerning insulitis and islet infiltration in association with T1D has come from post-mortem samples. For example, Wilcox and co-workers analyzed pancreatic material from 29 deceased recent onset T1D subjects (aged 1-42 years, median 11.7 +/- 1.6 years), observing that CD8+ cells were the most abundant islet infiltrating cells (19). Similarly, Coppieters *et al.* used 45 cadaveric T1D pancreatic samples to screen for insulin deficient islets, CD8+ insulitis and HLA class I hyper expression. Islet reactive CD8+ T cells were found in the islets of both recent onset and longstanding T1D patients (20).

The Diabetes Virus Detection study (DiViD) has performed pancreatic tail resection by laparoscopy on newly diagnosed T1D adult patients (24-35 years) and aims at a compherensive charaterisation of these unique samples with a range of state-of-the-art methods (52). The first RNA-seq analysis of the samples was recently reported (53). The achievement of effective proteomic analysis of such samples will require sensitive and reproducible technology for sample preparation analysis, such as has been recently demonstrated using PCT technology (54). As an alternative strategy to study pancreatic tissue, Green-Mitchell and coworkers used MALDI-MS imaging to compare pancreatic tissue from healthy subjects and T1D patients (55). They reported the detection of higher levels of insulin B chain in samples from healthy tissue and that their method with on-tissue reduction of inter-domain disulphide bonds resulted in localization of the pancreatic islets of Langerhans.

In addition to direct characterization of the tissue proteome, researchers have isolated islet-infiltrating CD4+ T cell clones from deceased organ donors with T1D. As an example, Delong *et al.* tested whether human CD4+ T cells from organ donors with T1D would recognize hybrid insulin fusion peptides (HIPs). The analyses were made as a follow up to experiments in which LC-MS/MS was used to characterize HIPs as epitopes for pathogenic CD4+ T cells in a mouse model (16).

Specific analysis of beta cells has similarly been mostly limited to animal models. In order to evaluate the utility of animal models in the study of islet infiltration and beta cell death, Jörns and coworkers compared pancreatic tissue from four major animal models to that from human. On the basis of cytokine expression profiles, immunohistochemistry and the composition of immune cells from the islets they found that three out of the four animal models mirrored the situation in human T1D (56). Rondas *et al.* investigated the protective effects of Glucagon-like peptide 1 on human islets. They analyzed the protein expression profiles of human islets of Langerhans organ donors (60 +/- 15 years, with no disease specificity) treated with cytokines (IL-1β and IFN-γ) (57). The expression profiles were compared using 2D-DIGE, with MALDI TOF/TOF for protein identification. In a subsequent study they reported that cytokine induced inflammation of mouse beta cells post-translationally citrullinated GRP78 to become a T1D autoantigen (17). Citrullination of autoantigens has previously been detected in relation to human rheumatoid arthritis (58).

As indicated earlier, the recognition of major histocompatibility complex (MHC)-bound peptides may be a pivotal stage in the autoimmune destruction of the pancreatic beta cells (23). The concept of post-translational modification in T1D has recently come into focus with the concept that immune system may be ignorant rather than tolerant of the post-translationally modified (PTM) islet autoantigens (59). In their recent detailed review, McGlaughlin and coworkers report several examples of post-translational modifications in

T1D, including deamidation, disulfide bridges, citrullination and oxidative modification (59). They also consider the origin of these neoantigens and the challenges and importance of their identification.

2.2 *Blood samples and cellular profiles:*

The type of samples that have been used for the study of diabetes in man have otherwise been restricted to body fluids such as urine, saliva, serum /plasma. Whilst for the most part these fluids can provide metabolic signatures of the biological processes associated with the disease onset and disease status, functional differences may be more directly gained from the proteomics analysis of cell populations. For instance, researchers have isolated PBMCs and specific cell populations, such as CD4+ cells, T_{reg} and NK cells, from whole blood (60), although to our knowledge there has not yet been any detailed proteomics analysis of these cell populations published.

To investigate the roles of T cells in the autoimmune response, Mannering and co-workers isolated proinsulin specific CD4+ cells from a patient with T1D. They identified a HLA DR4 restricted epitope in human insulin that was recognized by CD4+ T cell clones isolated from an HLA DR4(+) child with autoantibodies to insulin, but not in the two healthy controls. Using proteomics technology they further defined a novel post-translational modification that was required for T cell recognition of the insulin A-chain in T1D (61). Such a modification could be considered the basis of the shift in the immune systems response against insulin as an antigen, as similarly reported with insulin fusion peptides (16). Similarly, van Lummel *et al*. investigated the PTM of islet autoantigens and their role in T1D. They found that 28 out of the 31 candidate islet epitopes detected were post-translationally modified. Here they also isolated PBMCs from a 11 year old girl newly diagnosed with T1D and compared the results

of stimulation assays with deamidated proinsulin with native proinsulin. They found CD4+ cells that could become cross reactive against unmodified islet autoantigens after activation with a post translationally modified insulin peptide epitope (62). For further details on post translational modifications and the pathogenesis of T1D, the reader is referred in particular to the recent review of McLaughlin and co-workers (59).

2.3 *Serum and Plasma:*

Many of the serum proteomics studies of T1D have analyzed samples from the diagnosed condition rather than the pre-diagnosis phase. Metz *et al.* performed one of the first major LC-MS/MS studies of the serum proteome of T1D patients and later extended their methods to analyze a larger study group (63, 64). Their results highlighted a panel of 24 serum proteins associated with innate immunity that were changed with the disease status. Although the samples were primarily from adult diabetics, in a later published follow up study and validation of these findings, Zhang *et al.* included the analysis of samples from children in the age range of 5-15 years. They demonstrated that peptides from platelet basic protein and plasma protease C1 inhibitor achieved high sensitivity and specificity for classification of the T1D samples. Their proteomic strategy was based on the LC-MS/MS accurate mass and time tag (AMT) method (65) with selected reaction monitoring used in the validation. To enable deeper proteomic coverage and populate the AMT library, the pooled serum was depleted of the higher abundance proteins using the tandem immunoaffinity columns and fractionated using strong cation exchange (SCX).

Zhi *et al.* evaluated the use of proteomic technologies for the discovery of T1D biomarkers, proposing that serum or plasma was the specimen of choice due to its richness of biological information and easy availability. They stressed the importance of sample preparation to

enable characterization of the low- to medium-abundance proteins and favored the use of random hexapeptide library beads (Proteominer, PM) over the traditional immune-depletion methods. Using a LTQ linear ion trap mass spectrometer they reported detection of in the order of 1000 serum proteins from the samples prepared with the PM beads. As a cautionary note they suggested further evaluation of the reproducibility of the PM beads was needed and indicated that the method required large serum/plasma volumes (66). Following on from this evaluation, they used a similar 2D LC-MS/MS approach with spectral counting to analyze pooled samples from T1D patients and controls (three pools of 10 from cases and controls) (67). They reported the detection of almost 2,500 proteins and found significant differences in 21 serum proteins that had associations with inflammation, oxidation, metabolic regulation, and autoimmunity. Extensive immunoassay-based validations were made in a cohort of T1D patients and controls, including 1,139 patients and 848 controls, with both the case and control groups including over one hundred children under 15 years of age. The validations confirmed the T1D patients displayed significantly higher serum levels of adiponectin, insulin-like growth factor binding protein, C-reactive protein, serum amyloid protein A, and significantly lower transforming growth factor beta-induced, as well as myeloperoxidase. In comparison with the study of Zhang *et al.* (see above) (64) the overlap of the two panels of biomarkers was, however, marginal. These findings are also summarized in **Table 1**.

Purohit *et al.* used the surface enhanced laser desorption- time of flight SELDI-TOF technique to profile serum proteins from T1D patients and healthy autoantibody-negative controls (68). These samples were collected from the Prospective Assessment in Newborns for Diabetes Autoimmunity (PANDA) study in the Southeastern US (69), with a mean subject age of 14.3 years (range 0.9 to 42). They assessed the reproducibility of SELDI to identify putative biomarkers for T1D and reported 146 protein/peptide peaks that showed

statistical differences in expression levels between T1D patients and controls. Although the use of a multivariate model allowed them to observe promising specificity and sensitivity, they concluded that due to the poor day-to-day reproducibility of the technique, as well as the absence of direct protein identification, the platform was not suitable for the discovery and validation of biomarkers when only small proteomic changes were expected.

Albrethsen and coworkers also evaluated SELDI for the analysis of serum from T1D patients for the detection of biomarkers for disease prediction and disease monitoring (70). They compared samples from 270 patients with T1D from 18 different pediatric centers collected 1, 6, and 12 months after diagnosis (766 samples in total). The average age at clinical diagnosis was 9.1 +/- 3.7 years (mean +/- SD). From this analysis they found that members of the apolipoprotein family (APOC1, APOC3) were increased with time from diagnosis. They concluded that their results supported the use of serum cohorts for further proteomic studies and illustrated the potential of SELDI for high-throughput protein profiling to evaluate serum cohorts prior to proteomics biomarker research.

Massa *et al*. used the immunoproteomic Serological Proteome Analysis (SERPA) method to identify new autoantigens in children with T1D who tested negative for the autoantibodies currently used in clinical screening. After separation by a 2D gel, human islet proteins were blotted on a membrane and then exposed to sera of T1D patients (71). They identified 11 proteins as putative autoantigens, demonstrating the feasibility of immunoproteomics as a strategy for the identification of new candidate T1D autoantigens and evaluating risk in patients negative for the autoantibodies currently used in clinical practice.

Miersch and co-workers also screened for autoantibodies, although here they used a nucleic acid programmable array to compare the immunoreactivity of serum from T1D cases and controls against 6000 human proteins (72). The statistical analysis of the data revealed 26

novel and one known T1D-associated autoantigen. The median ages for the cases and control subjects studied (~170 vs 140) was in the order of 15 years. Amongst the novel putative autoantigens detected, they presented data that promoted DYRK2 as a minor autoantigen (36% sensitivity at 98% specificity). Their application supported the use of protein microarrays as a high-throughput platform for profiling of serum antibodies to a large number of protein antigens.

Bian and co-workers have recently reported the use of proteome arrays for the immunoproteomic profiling of antiviral antibodies in new onset T1D cases and controls (median subject age 12.7 years, 42 vs 42) (73). They observed that Epstein-Barr virus (EBV) antibodies were significantly higher in case than control subjects, suggesting a potential role of EBV in T1D development. To date, the strongest associations of autoimmunity and T1D with viral infection have mostly been linked with the cocksackie virus, the Coxsackie B1 strain in particular (74). In the study of Bian *et al.*, however, they found no history of infection with the E2 strain of the Coxsackie B virus that was considered in their analyses.

Recently, researchers used a combination of immunoprecipitation and LC-MS/MS, amongst other biochemical techniques to establish Tetraspanin-7 as the identity of the autoantibody hitherto known as glima (75). The presence of the autoantibody was established in sera from young diabetic patients (age 12 -26 years). Other proteomic studies of plasma and sera have sought to establish the nature of the IgG complex in patients with T1D. For example, Roveri and coworkers recently reported the detection of circulating glucose-regulated protein 94-IgG complexes (glucose-regulated protein 94 is also known as Endosplasmin) in the plasma of T1D subjects (76).

As the previous examples exemplify, studies that have addressed the comparison of pre-diabetic serum have been less common (35). Nevertheless, the publication of further reports

of this nature can be anticipated in the not too distant future as the prospective sample collections that have grown sufficiently provide the critical mass for temporal comparisons between individuals that develop T1D and matched controls. McGuire and co-workers used SELDI to compare the cord blood of children that subsequently developed T1D, comparing 54 cases with 108 controls. Although they were able to distinguish the T1D developing subjects the identity of the distinguishing features was not established (77).

In our own studies we have used samples from the DIPP project and charted the temporal changes from early infancy to seroconversion and diagnosis (35). Participation and adherence to the DIPP study has been such that for many of the children diagnosed with T1D there are samples available for healthy children that can be closely matched by date of birth and sample collection, as well as gender, risk group and geographic region. We have used samples from 19 paired subjects (38 individuals) and performed LC-MS/MS with both iTRAQ and label free quantification of serum immunoaffinity depleted of the most abundant proteins. The samples (266 in total) spanned from early infancy through to seroconversion and diagnosis (median age at diagnosis 4.3 +/- 2.8). In addition to the detection of differentially abundant proteins, time trends and correlated changes were interpolated from the data. Most recently, von Toerne *et al.* have analyzed depleted serum samples from German children prospectively recruited in the BABYDIAB study (78). They used label free LC-MS/MS to compare serum from children that progressed to T1D within 3.5 years of seroconversion (n = 15), children that progressed to T1D in 9.5 years or more after seroconversion (n = 15), and children that remained autoantibody negative (n=15). They found differences in 26 proteins and performed SRM validation with the comparison of autoantibody positive and negative children (70 vs. 70). They found that the measurement of peptides from apolipoprotein M and apolipoprotein C-IV (APOC-IV) could be used to discriminate between the autoantibody-positive and autoantibody-negative children. They

equated these differences to lipid-associated metabolic processes and homeostasis, and suggested that changes in lipid metabolism could occur early in the autoimmunity process. A notable overlap between these observation and our own study was that we similarly observed lower levels of APOC-IV levels in the children that subsequently developed T1D.

As with many studies of the serum proteome the results of the study of von Toerne, as was the case with our own, only detailed the quantification of several hundred proteins. This overview of the medium abundance proteins could miss important disease related proteins. In a recent study of cardiovascular risk using isobaric labelling of extensively depleted plasma, quantification of in the order of 5,000 proteins has been reported (79). If sufficient sample and subject numbers were analyzed with this level of detail, such methodology might provide important details of the proteomic changes associated with the emergence of T1D.

2.4 Saliva:

The flow and composition of saliva are important in the maintenance of health in the oral cavity and have both been found altered in subjects with T1D (80). Moreover, saliva reflects the balance of the mucosal immune system (81), as shown by Caseiro and co-workers with their iTRAQ based analysis of the salivary proteome and peptidome profile of T1D in adults. To investigate this association in children, Cabras *et al*. used LC-MS/MS to analyze the acidic soluble fraction of whole saliva of T1D patients aged 11.7 ± 3.6 years (mean age ±SD) with diabetes onset at 5.1 ± 3.3 years (n = 22), in comparison with sex- and age-matched control subjects (n = 22) (82). They observed that statherin, proline-rich peptide PB, salivary acidic proline-rich phosphoprotein 1/2 and histatin 1, were significantly less abundant, and that isoforms of S100A9 were more abundant in the saliva of diabetic children. They interpreted these differences could be a reflection of a compromise in the defense provided by the repertoire of saliva peptides in the oral cavity of children with T1D. In more recent

studies in adults, cytokine measurements from 114 T1D subjects indicated an increased inflammatory burden in the saliva that was associated with decreased glycemic control and gingival condition (83).

2.5 Urine:

Meir and coworkers used CE-MS to identify urinary protein patterns for the early recognition of the development of diabetic nephropathy. They compared samples from 44 T1D patients (aged 14.8 +/- 2.8 years) with nine healthy controls (aged 14.1 +/- 2.2 years) (84). Although the data indicated differences in the patterns of polypeptides detected, these were not characterized. Su *et al.* have compared the urinary proteome of siblings discordant for T1D from a group of children aged around 12 years (+/- 3.5) (85). The samples were concentrated using ultracentrifugation and digested using a filter-aided sample preparation method (FASP) (86) protocol. The digests were fractionated by reverse phase chromatography (pH 6.5) prior to analysis using electrospray LC-MS/MS. On average in the order of one thousand protein groups were identified from the analysis of samples from 81 subjects. Among the differentially abundant proteins detected in these analyses were a number of lysosomal enzymes and proteins with functional roles in vascular permeability and adhesion. These changes were interpreted to reflect hyperglycaemia-associated inflammation in the kidney vasculature.

3 CONCLUSION

Overall, research towards the identification of proteomic biomarkers of T1D in children has been historically limited by the available methodology and sample material, both of which are constantly emerging and growing. Proteomics technology has improved in terms of reproducibility and sample requirement, having developed to the level that thousands of proteins can be detected per hour (87). These improvements may translate to improve the potential for discovery in the analysis of samples from the biobanks of T1D developing children and provide detailed insight into the changes associated with disease progression. With such goals borne in mind, proteomic analysis of specific post-translational modifications, or simple biofluids remain targets of interest. Newer and emerging methodology, such as data independent analysis, targeted SRM and PRM could strengthen these investigations and facilitate translation from discovery to validation (88, 89). As a noteworthy example of targeted mass spectrometry, researchers at the Institute of Systems Biology in Seattle have recently reported SRM assays for 99.7% of the annotated human proteome (90). In terms of the perspective of improving the prediction of T1D, new panels of biomarkers could redefine risk groups, although the ultimate goal would nevertheless be prevention. In this respect, additional biochemical information might enable better selection for preventative trials. Similarly, the identification of new autoantigens, for instance resulting from post-translational modification, could help define targets for immunomodulation and re-education of the immune system.

4 ACKNOWLEDGEMENT:

The National Technology Agency of Finland (Finland Distinguished Professor Programme, grant 40398/11), the Juvenile Diabetes Research Foundation (JDRF, 17-2011-586), the Academy of Finland (Centre of Excellence in Molecular Systems Immunology and Physiology Research (grant 250114), and Biocenter Finland (Turku Proteomics Facility) are thanked for their financial support.

5 EXPERT COMMENTARY

Whilst T1D is associated with genetic risk it remains unpredictable who will develop the disease. There is a need for detailed characterization of the biochemical pathways and signatures of the disease, such that the threat and progression can be identified. However, due to the complexity of the disease etiology, different risk groups, and the influence of early diet and environment, wide scale surveillance and screening is required.

In research towards establishing proteomic signatures of T1D and its risk in children, the opportunities for discovery stand to benefit due to the insight and perseverance of the researchers and funding agencies that have established longitudinal sample collections from children with conferred genetic risk. In these collections, serum/plasma samples have been collected on the scale of a million. Efficient systems for sample management, storage and accessibility have been employed, although sample stability may be a concern in the analysis of aged sample collections, some of which are now over 20 years. For instance, reagent manufacturers sometimes indicate that ELISA reagents are intended for fresh samples and might fail with aged misfolded or degraded proteins. Here at least peptide-centric measurements, such as targeted mass spectrometry, of samples of a similar age should be less problematic.

In addition to disease research with serum, there has been the collection of other biofluids, cell populations and tissue sample from biopsies. Pancreatic tissue, nevertheless remains scarce and it can be difficult to find suitable control material.

Future research holds the potential to identify new markers that could stratify risk and identify subjects where intervention would be beneficial. The ultimate goal will be to establish targets that would enable disease prevention. If this new information could be attained from the comparisons of childhood serum profiles, it would be important to overcome some technical barriers, i.e. to achieve both sufficient depth in the analysis and throughput, so as to make best advantage of the detailed serum collections that have been acquired. Typically, in many disease targeting studies, serum profiles have to date quite often not extended beyond the level of hundreds of proteins. Nevertheless, quantitative identification of thousands of serum proteins has been demonstrated in disease orientated research, although the issue of throughput on a scale suitable for multiple samples from large

study groups remains a concern. Moreover, due to heterogeneity of the disease itself and serum in general, pooling is not necessarily the best option.

If the detection of proteins on the scale of thousands is seen as an obstacle with serum analysis, the analysis of tissue and cellular lysates is not limited in this way. Although here, on the contrary, achieving suitable sample size or cell populations may present a challenge.
In the past two decades proteomic technology has developed through the stages of gel-based methodology to data dependent shotgun proteomics and semi-quantitative chemical labelling, through to data independent methods and targeted methods using selected reaction monitoring (SRM) and parallel reaction monitoring. Data independent approaches promise an archive of the total detectable peptide fragment profiles of complex samples, offering new scope for discovery. Notably, with the wide-scale of adoption of SRM methods for validation, and recent development of SRM methods including proteotypic peptides for 97% of the annotated human proteome, confident targeted proteomics of protein panels identified in relation to T1D risk could be determined to establish the emerging threat of disease or the efficacy of therapies.
The concept of post translational modification altering the status of an endogenous protein as an antigen has gained interest as a potential cause of autoimmunity. The characterization of such changes could provide targets for immunotherapy and provide the means for prevention in certain individuals.

5.1 Five-year view

The next five years will see the emergence of new serum proteomics data sets from the prospective childhood T1D studies in Europe and the US. Collective analysis of these data and validation across the geographic regions will be used to establish related findings. The combination of this information with other omics technologies will help establish new risk factors. Characterization of proteomic changes in relation to different infections, whether in serum, lymphocytes or in islet cells will also address quantitative traits in T1D developing subjects. Identification of antigen creating post translational modifications in islet cell proteins will be used to establish whether there exit targets for new immunotherapies for prevention.

5.2 Key issues

- Although there are a number of genetic traits that are associated with the risk of T1D, it is not possible to predict in whom and when the disease will develop.
- Currently the development of a T1D associated autoimmune response is associated with the detection of a number of islet cell specific autoantibodies. However, the triggers to this are not yet clearly established and the duration of this seroconverted state prior to diagnosis is highly variable, ranging from months to a decade.
- The sample materials that are available for the proteomic analysis of T1D related markers in children are most frequently blood, blood serum and urine. Although a number of diabetes-targeted collections have provided pancreatic tissue samples for proteomic analysis, these are rare.
- Several international studies have collected major temporal sample series of blood, serum and stools, from children with a conferred T1D risk. These series follow the changes from early infancy to adolescence or diagnosis. Whilst detailed, the extensive nature of these series are somewhat of a challenge to proteomics in terms of the throughput required.
- Serum proteomics for T1D risk needs to reach sufficient throughput and depth of coverage to establish markers in wide sample series.
- The concept of post translational modification in T1D has recently come into focus with the concept that immune system may be ignorant rather than tolerant of the post translationally modified (PTM) islet autoantigen. The detection of novel PTM sites remains a challenge in this field

6 REFERENCES

1. Gepts, W. (1965) *Pathologic anatomy of the pancreas in juvenile diabetes mellitus.* Diabetes 14, 619-633

2. Bluestone, J. A., Herold, K., and Eisenbarth, G. (2010) *Genetics, pathogenesis and clinical interventions in type 1 diabetes.* Nature 464, 1293-1300

3. Leete, P., Willcox, A., Krogvold, L., *et al.* (2016) *Differential insulitic profiles determine the extent of beta-cell destruction and the age at onset of type 1 diabetes.* Diabetes 65, 1362-1369

4. Singal, D. P. and Blajchman, M. A. (1973) *Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus.* Diabetes 22, 429-432

5. Nerup, J., Platz, P., Andersen, O. O., *et al.* (1974) *HL-A antigens and diabetes mellitus.* Lancet 2, 864-866

6. Noble, J. A. and Erlich, H. A. (2012) *Genetics of type 1 diabetes.* Cold Spring Harb Perspect. Med. 2, a007732

7. Polychronakos, C. and Li, Q. (2011) *Understanding type 1 diabetes through genetics: Advances and prospects.* Nat. Rev. Genet. 12, 781-792 ***Provides a good overview of type 1 diabetes research**

8. Knip, M. and Simell, O. (2012) *Environmental triggers of type 1 diabetes.* Cold Spring Harbor Perspectives in Medicine 2, a007690

9. Purohit, S. and She, J. (2008) *Biomarkers for type 1 diabetes.* International Journal of Clinical and Experimental Medicine 1, 98-116

10. Harjutsalo, V., Sund, R., Knip, M., *et al.*. (2013) *Incidence of type 1 diabetes in finland.* JAMA 310, 427-428

11. Harjutsalo, V., Sjoberg, L., and Tuomilehto, J. (2008) *Time trends in the incidence of type 1 diabetes in finnish children: A cohort study.* Lancet 371, 1777-1782

12. Tuomilehto, J. (2013) *The emerging global epidemic of type 1 diabetes.* Curr. Diab Rep. 13, 795-804

13. Ionescu-Tirgoviste, C., Gagniuc, P. A., Gubceac, E., *et al.* (2015) *A 3D map of the islet routes throughout the healthy human pancreas.* Sci. Rep. 5, 14634

14. Orci, L., Ravazzola, M., Amherdt, M., *et al.* (1986) *Conversion of proinsulin to insulin occurs coordinately with acidification of maturing secretory vesicles.* J. Cell Biol. 103, 2273-2281

15. Ilonen, J., Hammais, A., Laine, A. P., *et al.* (2013) *Patterns of beta-cell autoantibody appearance and genetic associations during the first years of life.* Diabetes 62, 3636-3640

16. Delong, T., Wiles, T. A., Baker, R. L., *et al.* (2016) *Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion.* Science 351, 711-714

17. Rondas, D., Crevecoeur, I., D'Hertog, W., *et al.*. (2015) *Citrullinated glucose-regulated protein 78 is an autoantigen in type 1 diabetes.* Diabetes 64, 573-586

18. Roep, B. O. (2003) *The role of T-cells in the pathogenesis of type 1 diabetes: From cause to cure.* Diabetologia 46, 305-321

19. Willcox, A., Richardson, S. J., Bone, A. J., *et al.* (2009) *Analysis of islet inflammation in human type 1 diabetes.* Clin. Exp. Immunol. 155, 173-181

20. Coppieters, K. T. (2012) *Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients.* J. Exp. Med. 209, 51-60

21. Gysemans, C., Callewaert, H., Overbergh, L., *et al.* (2008) *Cytokine signalling in the beta-cell: A dual role for IFNgamma.* Biochem. Soc. Trans. 36, 328-333

22. Crevecoeur, I., Rondas, D., Mathieu, C., *et al.* (2015) *The beta-cell in type 1 diabetes: What have we learned from proteomic studies?* Proteomics Clin. Appl. Aug;9(7-8):755-66. **** Provides an excellent review of T1D research of the past years with particular attention to the pancreas and beta cells.**

23. Dudek, N. L. and Purcell, A. W. (2014) *The beta cell immunopeptidome.* Vitam. Horm. 95, 115-144 *** A clear overview of research toward characterizing the beta cell immunopeptidome**

24. van Lummel, M., van Veelen, P. A., de Ru, A. H., Pool, J., *et al.* (2016) *Discovery of a selective islet peptidome presented by the highest-risk HLA-DQ8trans molecule.* Diabetes 65, 732-741

25. Hummel, S. and Ziegler, A. G. (2011) *Early determinants of type 1 diabetes: Experience from the BABYDIAB and BABYDIET studies.* Am. J. Clin. Nutr. 94, 1821S-1823S

26. Kupila, A., Muona, P., Simell, T., *et al.* (2001) *Feasibility of genetic and immunological prediction of type I diabetes in a population-based birth cohort.* Diabetologia 44, 290-297

27. Larsson, K., Elding-Larsson, H., Cederwall, E., *et al.* (2004) *Genetic and perinatal factors as risk for childhood type 1 diabetes.* Diabetes Metab. Res. Rev. 20, 429-437

28. TEDDY Study Group. (2008) *The environmental determinants of diabetes in the young (TEDDY) study.* Ann. N. Y. Acad. Sci. 1150, 1-13

29. Wherrett, D. K. (2014) *Trials in the prevention of type 1 diabetes: Current and future.* Can. J. Diabetes 38, 279-284 ***An interesting insight on ongoing clinical trials that address T1D prevention**.

30. Knip, M., Akerblom, H. K., Becker, D., *et al.*. (2014) *Hydrolyzed infant formula and early beta-cell autoimmunity: A randomized clinical trial.* JAMA 311, 2279-2287

31. Ziegler, A. G., Rewers, M., Simell, O., *et al.* (2013) *Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children.* JAMA 309, 2473-2479

32. Koskinen, M. K., Helminen, O., Matomaki, J., *et al.* (2016) *Reduced beta-cell function in early preclinical type 1 diabetes.* Eur. J. Endocrinol. 174, 251-259

33. Helminen, O., Aspholm, S., Pokka, T., *et al.* (2015) *HbA1c predicts time to diagnosis of type 1 diabetes in children at risk.* Diabetes 64, 1719-1727

34. Heinonen, M. T., Moulder, R., and Lahesmaa, R. (2015) *New insights and biomarkers for type 1 diabetes: Review for scandinavian journal of immunology.* Scand. J. Immunol. 82, 244-253

35. Moulder, R., Bhosale, S. D., Erkkila, T., *et al.* (2015) *Serum proteomes distinguish children developing type 1 diabetes in a cohort with HLA-conferred susceptibility.* Diabetes 64, 2265-2278

36. Elo, L. L., Mykkanen, J., Nikula, T., *et al.* (2010) *Early suppression of immune response pathways characterizes children with prediabetes in genome-wide gene expression profiling.* J. Autoimmun. 35, 70-76

37. Kallionpaa, H., Elo, L. L., Laajala, E., *et al.* (2014) *Innate immune activity is detected prior to seroconversion in children with HLA-conferred type 1 diabetes susceptibility.* Diabetes 63, 2402-2414

38. Oresic, M., Simell, S., Sysi-Aho, M., *et al.* (2008) *Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes.* J. Exp. Med. 205, 2975-2984

39. Giongo, A., Gano, K. A., Crabb, D. B., *et al.*. (2011) *Toward defining the autoimmune microbiome for type 1 diabetes.* ISME J 5, 82-91

40. Kallionpaa, H., Laajala, E., Oling, V., *et al.*. (2014) *Standard of hygiene and immune adaptation in newborn infants.* Clin. Immunol. 155, 136-147

41. Kostic, A. D., Gevers, D., Siljander, H., *et al.*. (2015) *The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes.* Cell. Host Microbe 17, 260-273

42. Vatanen, T., Kostic, A., d'Hennezel, E., *et al.*. (2016) *Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans.* Cell 165, 842-853

43. Sparre, T., Larsen, M. R., Heding, P. E., *et al.*. (2005) *Unraveling the pathogenesis of type 1 diabetes with proteomics: Present and future directions.* Mol. Cell. Proteomics 4, 441-457* **A good review on early proteomics research toward T1D**

44. Topf, F., Schvartz, D., Gaudet, P., *et al.*.. (2013) *The human diabetes proteome project (HDPP): From network biology to targets for therapies and prevention.* Translational Proteomics 1, 3-11

45. Schvartz, D., Bergsten, P., Baek, K.-H., *et al.*. (2015) *The human diabetes proteome project (HDPP): The 2014 update.* Translational Proteomics 8-9, 1-7

46. Bergsten, P. (2009) *Islet protein profiling.* Diabetes Obes. Metab. 11 Suppl 4, 97-117

47. Hu, L., Evers, S., Lu, Z. H., *et al.*. (2004) *Two-dimensional protein database of human pancreas.* Electrophoresis 25, 512-518

48. Ahmed, M., Forsberg, J., and Bergsten, P. (2005) *Protein profiling of human pancreatic islets by two-dimensional gel electrophoresis and mass spectrometry.* J. Proteome Res. 4, 931-940

49. Metz, T. O., Jacobs, J. M., Gritsenko, *et al.*. (2006) *Characterization of the human pancreatic islet proteome by two-dimensional LC/MS/MS.* J Proteome Res. 2006 Dec;5(12):3345-54.

50. Burch, T. C., Morris, M. A., Campbell-Thompson, *et al.*. (2015) *Proteomic analysis of disease stratified human pancreas tissue indicates unique signature of type 1 diabetes.* PLoS One 10, e0135663

51. Liu, C., Atkinson, M. A., and Zhang, Q. (2016) *Type 1 diabetes cadaveric human pancreata exhibit a unique exocrine tissue proteomic profile.* Proteomics 16, 1432-1446

52. Krogvold, L., Edwin, B., Buanes, T., *et al.*. (2014) *Pancreatic biopsy by minimal tail resection in live adult patients at the onset of type 1 diabetes: Experiences from the DiViD study*. Diabetologia 57, 841-843

53. Krogvold, L., Skog, O., Sundstrom, G., *et al.*. (2015) *Function of isolated pancreatic islets from patients at onset of type 1 diabetes; insulin secretion can be restored after some days in a non-diabetogenic environment in vitro. results from the DiViD study.* Diabetes Jul;64(7):2506-12.

54. Guo, T., Kouvonen, P., Koh, C. C., *et al.*. (2015) *Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps.* Nat. Med. 21, 407-413

55. Green-Mitchell, S. M., Cazares, L. H., Semmes, O. J., *et al.*. (2011) *On-tissue identification of insulin: In situ reduction coupled with mass spectrometry imaging.* Proteomics Clin. Appl. 5, 448-453

56. Jorns, A., Arndt, T., Meyer zu Vilsendorf, A., *et al.*. (2014) *Islet infiltration, cytokine expression and beta cell death in the NOD mouse, BB rat, komeda rat, LEW.1AR1-iddm rat and humans with type 1 diabetes.* Diabetologia 57, 512-521

57. Rondas, D., Bugliani, M., D'Hertog, W., *et al.*. (2013) *Glucagon-like peptide-1 protects human islets against cytokine-mediated beta-cell dysfunction and death: A proteomic study of the pathways involved.* J. Proteome Res. 12, 4193-4206

58. Klareskog, L., Amara, K., and Malmstrom, V. (2014) *Adaptive immunity in rheumatoid arthritis: Anticitrulline and other antibodies in the pathogenesis of rheumatoid arthritis.* Curr. Opin. Rheumatol. 26, 72-79

59. McLaughlin, R. J., Spindler, M. P., van Lummel, M., *et al.*. (2016) *Where, how, and when: Positioning posttranslational modification within type 1 diabetes pathogenesis.* Current Diabetes Reports 16, 1-9** **A concise update on post translational modifications detected in relation to type diabetes and their potential associations in triggering autoimmune reactions.**

60. Walker, L. S. and von Herrath, M. (2016) *CD4 T cell differentiation in type 1 diabetes.* Clin. Exp. Immunol. 183, 16-29

61. Mannering, S. I., Harrison, L. C., Williamson, N. A., *et al.*. (2005) *The insulin A-chain epitope recognized by human T cells is posttranslationally modified.* J. Exp. Med. 202, 1191-1197

62. van Lummel, M., Duinkerken, G., van Veelen, P. A., *et al.* (2013) *Posttranslational modification of HLA-DQ binding islet autoantigens in type 1 diabetes.* Diabetes 63, 237-247

63. Metz, T. O., Qian, W. J., Jacobs, J. M., *et al.* (2008) *Application of proteomics in the discovery of candidate protein biomarkers in a diabetes autoantibody standardization program sample subset.* J. Proteome Res. 7, 698-707

64. Zhang, Q., Fillmore, T. L., Schepmoes, A. A., *et al.* (2013) *Serum proteomics reveals systemic dysregulation of innate immunity in type 1 diabetes.* J. Exp. Med. 210, 191-203***A detailed report on the discovery and validation of serum biomarkers in T1D patients**

65. Pasa-Tolic, L., Masselon, C., Barry, R. C., *et al.* (2004) *Proteomic analyses using an accurate mass and time tag strategy.* BioTechniques 37, 621-4, 626-33, 636 passim

66. Zhi, W., Purohit, S., Carey, C., *et al.* (2010) *Proteomic technologies for the discovery of type 1 diabetes biomarkers.* J. Diabetes Sci. Technol. 4, 993-1002

67. Zhi, W., Sharma, A., Purohit, S., *et al.* (2011) *Discovery and validation of serum protein changes in type 1 diabetes patients using high throughput two dimensional liquid chromatography-mass spectrometry and immunoassays.* Mol. Cell. Proteomics 10, M111.012203 ***Description of the LC-MS/MS discovery of T1D serum biomarkers and extensive ELISA validation.**

68. Purohit, S., Podolsky, R., Schatz, D., *et al.* (2006) *Assessing the utility of SELDI-TOF and model averaging for serum proteomic biomarker discovery.* Proteomics 6, 6405-6415

69. Carmichael, S. K., Johnson, S. B., Baughcum, A., *et al.* (2003) *Prospective assessment in newborns of diabetes autoimmunity (PANDA): Maternal understanding of infant diabetes risk.* Genet. Med. 5, 77-83

70. Albrethsen, J., Kaas, A., Schonle, E., *et al.* (2009) *F* Proteomics Clin. Appl. 3, 383-393

71. Massa, O., Alessio, M., Russo, L., *et al.* (2013) *Serological proteome analysis (SERPA) as a tool for the identification of new candidate autoantigens in type 1 diabetes.* J. Proteomics 82, 263-273

72. Miersch, S., Bian, X., Wallstrom, G., *et al.* (2013) *Serological autoantibody profiling of type 1 diabetes by protein arrays.* J. Proteomics 94, 486-496

73. Bian, X., Wallstrom, G., Davis, A., *et al.* (2016) *Immunoproteomic profiling of antiviral antibodies in new-onset type 1 diabetes using protein arrays.* Diabetes 65, 285-296

74. Laitinen, O. H., Honkanen, H., Pakkanen, O., *et al.* (2014) *Coxsackievirus B1 is associated with induction of beta-cell autoimmunity that portends type 1 diabetes.* Diabetes 63, 446-455

75. McLaughlin, K. A., Richardson, C. C., Ravishankar, A., *et al.* (2016) *Identification of tetraspanin-7 as a target of autoantibodies in type 1 diabetes.* Diabetes Jun;65(6):1690-8.

76. Roveri, A., Zaccarin, M., Pagetta, A., T *et al.* (2015) *Proteomic investigation on Grp94-IgG complexes circulating in plasma of type 1 diabetic subjects.* J. Diabetes Res. 2015, 815839

77. McGuire, J. N., Eising, S., Wagner, A. M., *et al.* (2010) *Screening newborns for candidate biomarkers of type 1 diabetes.* Arch. Physiol. Biochem. 116, 227-232

78. von Toerne, C., Laimighofer, M., Achenbach, P., *et al.* (2016) *Peptide serum markers in islet autoantibody-positive children.* Diabetologia Nov 4. [Epub ahead of print]

79. Keshishian, H., Burgess, M. W., Gillette, M. A., *et al.* (2015) *Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury*. Mol. Cell. Proteomics 14, 2375-2393 ***An outstanding example of deep profiling plasma proteomics**

80. Moore, P. A., Guggenheimer, J., Etzel, K. R., *et al.* (2001) *Type 1 diabetes mellitus, xerostomia, and salivary flow rates.* Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. 92, 281-291

81. Caseiro, A., Ferreira, R., Padrao, A., *et al.* (2013) *Salivary proteome and peptidome profiling in type 1 diabetes mellitus using a quantitative approach.* J. Proteome Res. 12, 1700-1709

82. Cabras, T., Pisano, E., Mastinu, A., *et al.* (2010) *Alterations of the salivary secretory peptidome profile in children affected by type 1 diabetes.* Mol. Cell. Proteomics 9, 2099-2108

83. Kuehl, M. N., Rodriguez, H., Burkhardt, B. R., *et al.* (2015) *Tumor necrosis factor-alpha, matrix-metalloproteinases 8 and 9 levels in the saliva are associated with increased hemoglobin A1c in type 1 diabetes subjects.* PLoS One 10, e0125320

84. Meier, M., Kaiser, T., Herrmann, A., *et al.*. (2005) *Identification of urinary protein pattern in type 1 diabetic adolescents with early diabetic nephropathy by a novel combined proteome analysis.* J. Diabetes Complications. 19, 223-232

85. Suh, M. J., Tovchigrechko, A., Thovarai, V., *et al.* (2015) *Quantitative differences in the urinary proteome of siblings discordant for type 1 diabetes include lysosomal enzymes.* J. Proteome Res. 14, 3123-3135

86. Wisniewski, J. R., Zougman, A., Nagaraj, N., *et al.* (2009) *Universal sample preparation method for proteome analysis.* Nat. Methods 6, 359-362

87. Hebert, A. S., Richards, A. L., Bailey, *et al.* (2014) *The one hour yeast proteome.* Mol. Cell. Proteomics 13, 339-347

88. Chapman, J. D., Goodlett, D. R., and Masselon, C. D. (2014) *Multiplexed and data-independent tandem mass spectrometry for global proteome profiling.* Mass Spectrom. Rev. 33, 452-470

89. Shi, T., Song, E., Nie, S., *et al.* (2016) *Advances in targeted proteomics and applications to biomedical research.* Proteomics Aug;16(15-16):2160-82.

90. Kusebauch, U., Campbell, D. S., Deutsch, E. W., *et al.* (2016) *Human SRMAtlas: A resource of targeted assays to quantify the complete human proteome.* Cell Jul 28;166(3):766-78. ****SRM asays developed for 97% of the annotated human proteome**

Figure 1: The proteomic study of type 1 diabetes in man is has included the use of tissue samples, e.g. pancreatic tissue, and bodily fluids such as serum, plasma and urine. In addition to the direct profiling of serum from blood samples peripheral blood mononuclear cell (PBMC) can be isolated to study their response to stimulations. The methodologies applied have spanned the range of shotgun to gel based proteomic approaches, and whilst researchers have frequently searched for differentially abundant proteins the role of post-translational modification have gained attention as cause of autoantigens.

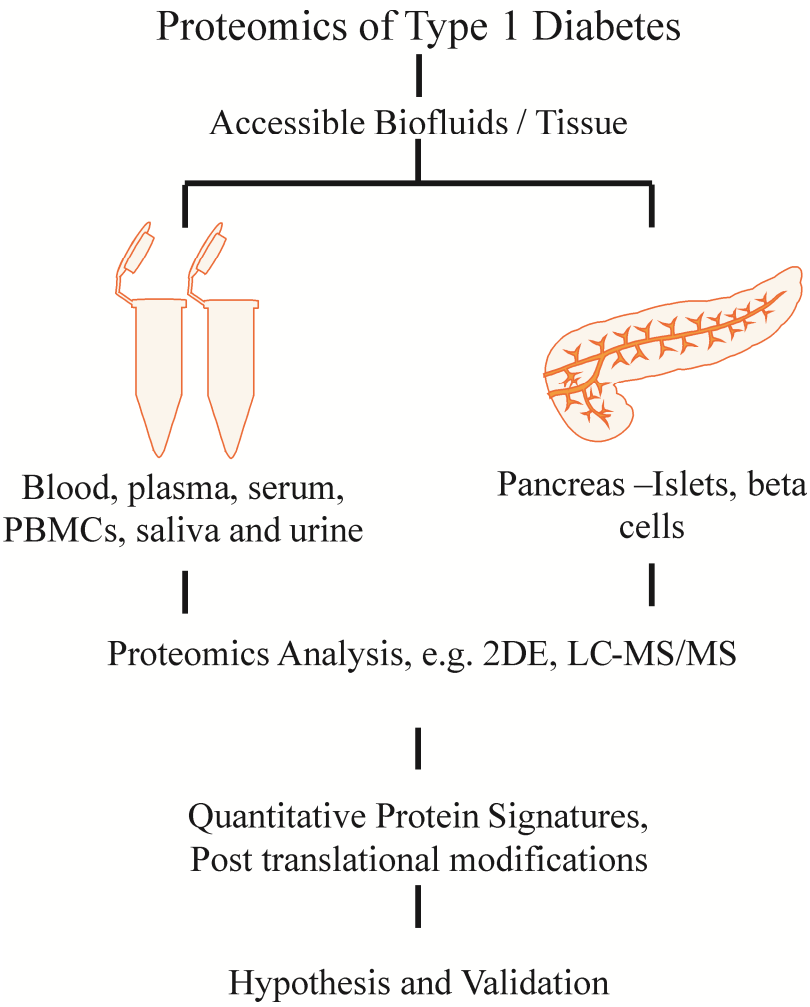

Table 1: Examples of studies of different tissues and bio-fluids for type 1 diabetes research. The studies are identified and described by the following details: i) by the type of sample used (e.g. plasma), ii) the diagnosis status of the subjects compared (ND indicates no disease, AAb+ve indicates that known T1D associated autoantibodies have been detected in the subject's serum), iii) the methods used, iv) the key protein markers detected/reported (the total number of markers indicated in brackets), v) the first author, year of publication, the research aim and reference number, vi) the pathways inferred from the differentially abundant proteins.

Sample	Subject Status	Method	Key proteins & total markers	Author, year, reference, target	Pathways inferred/reported
Plasma	*T1D/ND*	*2DE-LC-MS/MS, AMT*	*ZAG, CBG, LUM, CLUS, TRFE*	*Metz* et al. *2008 (63), biomarkers of T1D*	Not stated
Plasma	*T1D/ND*	2DE, LC-MS/MS	ENPL	Roveri *et al*. 2015 (76), immune complexes	Inflammatory cell signalling, angiogenic-like transformation of vascular cells
Serum	*T1D/ND*	2DE, MALDI-TOF/TOF	GB1B, (11)	Massa *et al*. 2013 (71), A autoantigens	Autoimmunity
Serum	*T1D/ND*	SELDI-TOF-MS	APOC1, APOC3	Albrethsen *et al*. 2009 (70), biomarkers of T1D	T1D disease and/or clinical treatment
Serum	*T1D/ND*	2DE-LC-MS/MS, AMT, SRM	TGFBI, CRP, ADIPO, IBP2, PERM, SAA1, (21)	Zhi *et al*. 2001 (67), biomarkers of T1D	Innate immunity, inflammation and immune response, lymphocyte activation/proliferation and glucose regulation
Serum	*T1D/ND*	LC-MS/MS, spectral counting, ELISA, Luminex	IC1, PGRP2, CXCL7,TTHY, (24)	Zhang *et al*. 2013 (64), biomarkers of T1D	Innate immune response, complement activation cascade, inflammatory response, and blood coagulation
Serum	*Pre-T1D, AAb+ve, ND*	LC-MS/MS, LFQ, iTRAQ	APOC4, APOC2, PFN1,MBL2, FHR5, CO9, BGH3, ADIPO, IGFBP2, (16)	Moulder *et al*. 2015 (35), early biomarkers of risk & progression	Inflammation and immune response
Serum	*Pre-T1D, AAb+ve, ND*	LC-MS/MS, LFQ, SRM	APOM, APOC4, CFH, HGFAC, CP, (26)	von Toerne *et al*. 2016 (78), early biomarkers of risk & progression	Lipid-associated metabolic processes and homeostasis,
Pancreatic tissue	*T1D, AAb+ve, T2D, ND*	LC-MS/MS	OLFM4, ENPP1, REGI3A, (307 & 242)	Burch *et al*. 2015 (50), prognostic/diagnostic biomarkers	Antimicrobial response, Immune Cell Trafficking, Inflammatory Response, Inflammatory Disease, Cell Death and Survival,
Pancreatic tissue	*T1D/ND*	LC-MS/MS, TMT	HNRPK, CUL4A, TSN8, UBE2F, SPRL1, (145)	Liu *et al*. 2016 (51), pathology of T1D	Immune response, viral infection, ubiquitin proteasome cell proliferation and apoptosis
Saliva	*No complications T1D vs. retinopathy, nephropathy, ND*	LC-MALDI/TOF-TOF, ITRAQ	BPI, PAUF , (26)	Caseiro *et al*., 2013 (81), markers from the saliva peptidome and proteome	Metabolic and immune response
Saliva	*T1D/ND*	LC-MS/MS	STAT, SMR3B, PRPC, HIS1, S100A9 (phosphorylated)	Cabras *et al*., 2010, salivary peptides & proteins	Phosphorylation, p38 MAPK pathway
Urine	*T1D-nephropathy/ND*	CE-MS	No identification	Meier *et al., 2005* (84), markers of diabetic nephropathy.	NA
Urine	*T1D/healthy siblings*	2D-LC-MS/MS	TIMP1, GP5, GAS6, KNG1, APOM, FUCA2, NAGA, GSN, ACE2, ALCAM, CDH5, (45)	Suh *et al.* 2015 (85), hyperglycemia & inflammation	Metabolic and Inflammatory Processes, vascular permeability and adhesion