TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

PROTEOMICS INSIGHTS INTO TYPE 1 DIABETES DEVELOPMENT

Investigating changes in serum proteome and the role of persistent enterovirus infection

M. Karoliina Hirvonen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1876 | MEDICA – ODONTOLOGICA | TURKU 2025

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

PROTEOMICS INSIGHTS INTO TYPE 1 DIABETES DEVELOPMENT

Investigating changes in serum proteome and the role of persistent enterovirus infection

M. Karoliina Hirvonen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1876 | MEDICA – ODONTOLOGICA | TURKU 2025

University of Turku

Faculty of Medicine
Institute of Biomedicine
Medical Microbiology and Immunology
Turku Doctoral Programme of Molecular Medicine

Supervised by

Professor Riitta Lahesmaa, MD, Ph.D.
Turku Bioscience Centre
University of Turku and
Åbo Akademi University
Turku, Finland

Niina Lietzén, Ph.D.
3PBiovian
Turku, Finland

Reviewed by

Professor Tuula Nyman, Ph.D.
Department of Immunology
University of Oslo and
Oslo University Hospital
Oslo, Norway

Professor Vesa Hytönen, Ph.D.
Faculty of Medicine and Health Technology
Tampere University
Tampere, Finland

Opponent

Professor Jochen Schwenk, Ph.D.
Department of Protein Science
SciLifeLab
KTH Royal Institute of Technology
Stockholm, Sweden

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-952-02-0136-4 (PRINT)
ISBN 978-952-02-0137-1 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama, Turku, Finland 2025

To My Family and Friends

UNIVERSITY OF TURKU
Faculty of Medicine
Institute of Biomedicine
Medical Microbiology and Immunology
KAROLIINA HIRVONEN: Proteomics Insights into Type 1 Diabetes Development – Investigating Changes in Serum Proteome and the Role of Persistent Enterovirus Infection
Doctoral Dissertation, 156 pp.
Turku Doctoral Programme of Molecular Medicine
May 2025

ABSTRACT

Type 1 diabetes is a chronic disease that leads to gradual destruction of insulin producing beta cells in the pancreas. By the time of diagnosis, most of the beta cells have already been destroyed. There is a need to improve our understanding of the molecular events governing the disease, from early signs to diagnosis and beyond. Identifying molecules that change at the various stages of disease could improve clinical decisions, assist in monitoring, and help classify individuals into subgroups.

In the work presented in this dissertation, mass spectrometry (MS)-based proteomics methods were employed to investigate type 1 diabetes-associated changes in serum protein levels and to explore how persistent enterovirus infection, which has been linked to the disease, alters protein expression and secretion in pancreatic ductal cells (PANC-1). Longitudinal proteomics analysis of pre-onset serum samples from children who developed type 1 diabetes at a young age revealed decreased levels of apolipoprotein C1 (APOC1) early in the disease, which remained lower until diagnosis. Targeted MS analysis of serum samples from newly diagnosed (ND) youth and unaffected family members (UFMs) further identified lower serum APOC1 levels in ND youth and revealed 12 additional proteins with differing serum levels between the two groups. Eleven proteins were also found to be associated with fasting C-peptide/glucose ratios, which served as a measure of functional beta cell mass. A validation study in a second cohort confirmed serum level differences for 10 proteins, including APOC1, between ND and UFMs and replicated negative association with the C-peptide/glucose ratios for three proteins. Persistent enterovirus infection in PANC-1 cells, using two different coxsackievirus B1 (CVB1) strains, led to significant changes in protein expression and secretion, especially affecting mitochondrial functions and the regulated secretory pathway. Notably, the two CVB1 strains also showed different effects on proteins related to the antiviral immune response. Several of the significantly altered proteins in the CVB1 models were also detected in the serum samples.

This research identified a set of proteins linked to the progression of type 1 diabetes and those modulated during persistent CVB1 infection. While a specific CVB1 signature was not found in the serum analyses, a group of overlapping proteins identified in the studies may contribute to the CVB1 infection signature.

KEYWORDS: Type 1 diabetes, Enterovirus, Proteomics, Mass spectrometry

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Biolääketieteen laitos
Lääketieteellinen mikrobiologia ja immunologia
KAROLIINA HIRVONEN: Tyypin 1 diabetekseen liittyvät muutokset proteiinien ilmentymisessä – Havaitut erot seerumiproteomissa sekä kroonisten enterovirusinfektioiden rooli sairauden kehittymisessä
Väitöskirja, 156 s.
Molekyylilääketieteen tohtoriohjelma
Toukokuu 2025

TIIVISTELMÄ

Tyypin 1 diabetes on sairaus, jossa kehon insuliinia tuottavat haiman beetasolut tuhoutuvat. Diagnoosivaiheessa sairastuneet ovat jo menettäneet suurimman osan insuliinia tuottavista beetasoluista. On tärkeää saada lisätietoa sairauteen liittyvistä molekyylitason muutoksista alkaen sairauden alkuvaiheesta aina diagnoosiin ja siitä eteenpäin. Molekyylit, joissa havaitaan muutoksia sairauden eri vaiheissa, voivat tuoda merkittävää lisätietoa kliiniseen päätöksentekoon, edesauttaa sairauden kulun seuraamista sekä helpottaa yksilöiden jakamista sairauden eri alaryhmiin.

Tässä väitöskirjassa hyödynnettiin massaspektrometriaan perustuvaa proteomiikkaa tyypin 1 diabetekseen liittyvien proteiinimuutosten havaitsemiseen seerumissa sekä tarkasteltiin kroonisen enterovirusinfektion aiheuttamia muutoksia haiman duktaalisoluissa. Proteiinien ilmentymiseen liittyviä muutoksia seerumissa tarkasteltiin ensin tyypin 1 diabetekseen sairastuneiden lasten sekä verrokkilasten välillä. Tässä pitkittäistutkimuksessa havaittiin apolipoproteiini C1 (APOC1) -tasojen laskevan sairauden varhaisessa vaiheessa ja pysyvän tällä tasolla diagnoosin saakka. Matalammat seerumin APOC1-tasot havaittiin myös vasta diagnosoiduilla nuorilla verrattaessa heidän näytteitään sairastumattomista perheenjäsenistä koostuvan verrokkijoukon näytteisiin. Ryhmien välillä löytyi eroja myös 12 muun proteiinin ilmentymisessä. Lisäksi 11 proteiinin kohdalla havaittiin yhteys proteiinin seerumitason ja C-peptidi/glukoosisuhteen välillä. Myöhemmässä tutkimuksessa, vahvistettiin erot 10 proteiinien kohdalla vastadiagnosoitujen ja verrokkiryhmän välillä, mukaan lukien APOC1. Lisäksi kolmen proteiinin kohdalla toistettiin yhteys proteiinin seerumitason ja C-peptidi/glukoosisuhteen välillä. Kroonisen enterovirusinfektion solumalleissa havaittiin laajoja muutoksia proteiinien ilmentymisessä. Eri enteroviruskantojen välillä löydettiin myös eroja erityisesti immuunivasteeseen liittyvissä proteiineissa. Osa proteiineista, joissa havaittiin muutoksia infektoiduissa solumalleissa, saatiin mitattua myös seeruminäytteistä.

Tässä väitöskirjassa esitettiin joukko proteiineja, jotka ovat yhteydessä tyypin 1 diabeteksen kehittymiseen sekä proteiineja, joissa havaittiin muutoksia kroonisen enterovirusinfektion aikana. Proteiinit, joista löytyi muutoksia sekä infektoiduissa soluissa että seerumissa, saattavat osaltaan viitata kroonisen infektion olemassaoloon.

AVAINSANAT: Tyypin 1 diabetes, Enterovirus, Proteomiikka, Massaspektrometria

Table of Contents

Abbreviations

1AAb+	Single autoantibody-positive
A1BG	Alpha-1B-glycoprotein
A2M	Alpha 2-macroglobulin
AAb-	Autoantibody-negative
ALDH1A3	Aldehyde dehydrogenase 1a3
ALDOA	Fructose-bisphosphate aldolase A
APOA4	Apolipoprotein A4
APOB	Apolipoprotein B
APOC1	Apolipoprotein C1
APOM	Apolipoprotein M
Arg	Arginine
BMI	Body mass index
C2	Complement C2
CAR	The coxsackievirus and adenovirus receptor
CHGB	Chromogranin-B
COL1A1	Collagen alpha-1(I) chain
CPE	Carboxypeptidase E
CVB	Coxsackie B virus
CVB1	Coxsackievirus B1
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DIPP	The Finnish diabetes prediction and prevention study
DTT	Dithiothreitol
ESI	Electrospray ionisation
F2	Prothrombin
FDR	False discovery rate
FFPE	Formalin-fixed paraffin-embedded
GADA	Glutamic acid decarboxylase 65 autoantibodies
GPX3	Glutathione peroxidase 3
HLA	Human leukocyte antigen
IAA	Insulin autoantibodies

IA-2A	Tyrosine phosphatase-like insulinoma antigen 2 autoantibodies
ICA	Islet-cell autoantibodies
IFIH1	Interferon-induced helicase C domain-containing protein 1
IFN	Interferon
IFNL1	Interferon lambda-1
IGF	Insulin-like growth factor
IGFBP2	Insulin-like growth factor-binding protein 2
IGFBP3	Insulin-like growth factor-binding protein 3
IGHM	Immunoglobulin heavy constant mu
INNODIA	Innovative approaches to understanding and arresting type 1 diabetes
IPA	Ingenuity pathway analysis
IRES	Internal ribosome entry site
LFQ	Label-free quantification
LMM	Linear mixed effects model
LRP1	Prolow-density lipoprotein receptor-related protein 1
Lys	Lysine
MALDI	Matrix-assisted laser desorption/ionization
MCAM	Cell surface glycoprotein MUC18
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
m/z	Mass-to-charge ratio
ND	Newly diagnosed
nLC	Nano-flow liquid chromatography
PANC-1	Pancreatic duct cell line
PRM	Parallel reaction monitoring
PTM	Post-translational modification
QC	Quality control
RT-qPCR	Quantitative reverse transcription polymerase chain reaction
SCG3	Secretogranin-3
SERPING1	Plasma protease C1 inhibitor
SRM	Selected reaction monitoring
TEDDY	The environmental determinants of diabetes in the young
TGFBI	Transforming growth factor-beta-induced protein ig-h3
TOF	Time-of-flight
UFM	Unaffected family member
ZnT8	Zinc transporter 8

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I **Hirvonen MK**, Lietzén N, Moulder R, Bhosale SD, Koskenniemi J, Vähä-Mäkilä M, Nurmio M, Orešič M, Ilonen J, Toppari J, Veijola R, Hyöty H, Lähdesmäki H, Knip M, Cheng L[#], Lahesmaa R[#]. Serum APOC1 levels are decreased in young autoantibody positive children who rapidly progress to type 1 diabetes. *Sci Rep*, 2023; 13(1): 15941. doi: 10.1038/s41598-023-43039-4.

II Moulder R*, Välikangas T*, **Hirvonen MK**, Suomi T, Brorsson CA, Lietzén N, Bruggraber SFA, Overbergh L, Dunger DB, Peakman M, Chmura PJ, Brunak S, Schulte AM, Mathieu C, Knip M, Elo LL[#], Lahesmaa R[#]; INNODIA consortium. Targeted serum proteomics of longitudinal samples from newly diagnosed youth with type 1 diabetes distinguishes markers of disease and C-peptide trajectory. *Diabetologia*, 2023; 66(11):1983-1996. doi: 10.1007/s00125-023-05974-9.

III Moulder R*, **Hirvonen MK***, Välikangas T*, Suomi T, Overbergh L, Peakman M, Mathieu C, Knip M, Elo LL[#], Lahesmaa R[#]. Targeted serum proteomics of longitudinal samples from newly diagnosed youth with type 1 diabetes affirms markers of disease. *Diabetologia*, 2025; *In press*. doi: 10.1007/s00125-025-06394-7.

IV Lietzén N*, **Hirvonen K***, Honkimaa A*, Buchacher T, Laiho JE, Oikarinen S, Mazur MA, Flodström-Tullberg M, Dufour E, Sioofy-Khojine AB, Hyöty H[#], Lahesmaa R[#]. Coxsackievirus B Persistence Modifies the Proteome and the Secretome of Pancreatic Ductal Cells. *iScience*, 2019; 19:340-357. doi: 10.1016/j.isci.2019.07.040.

*, [#] Authors contributed equally

1 Introduction

Type 1 diabetes results from a selective and severe immune-mediated destruction of insulin-producing beta cells in the pancreas, forcing these individuals to be dependent on exogenous insulin for the rest of their lives. Genetic susceptibility and environmental factors, such as viral infections, are known to play a role in the initiation of the disease. By the time of diagnosis, individuals have already lost the majority of their functional beta cell mass. Understanding the molecular events throughout the disease progression is critical for stratifying patients into different subgroups, identifying the optimal timing for intervention and therapy, and preventing acute complications.

Protein expression and secretion are altered in nearly all diseases. To detect these disease-associated changes from clinical samples, a large-scale study of proteins, i.e. proteomics, provides insights into disease development and may lead to discoveries of biomarkers that can be used for the prediction and monitoring of the disease. Mass spectrometry-based proteomics has developed into a highly advanced technology that enables quantitative analysis of hundreds to thousands of proteins simultaneously with very high specificity. Depending on the biological question, it is possible to analyse cells, tissues, or body fluids such as blood serum and plasma, which are readily available.

In the work presented in this dissertation, mass spectrometry-based proteomics methods were employed to study type 1 diabetes-associated changes in protein expression levels. The main focus of the analyses was on longitudinal plasma and serum sample series obtained from prospective cohort studies. Additionally, to gain a broader perspective on disease etiology, one study explores how persistent enterovirus infection, that is associated with type 1 diabetes, modulates the protein expression and secretion of pancreatic ductal cells. Altogether, the aims of this dissertation are (1) to define protein patterns associated with the progression of type 1 diabetes, (2) to gain knowledge of the persistent enterovirus-induced changes in protein expression and secretion, and (3) to investigate whether signs of virus-induced changes are detected in plasma/serum of children who progress to type 1 diabetes.

2 Review of the Literature

2.1 Type 1 diabetes

Type 1 diabetes is a chronic disease that results from a gradual destruction of insulin-secreting beta cells in the pancreas leading to lifelong need for treatment with exogenous insulin. Insulin is a vital hormone that regulates the levels of body's blood sugar that is glucose. In response to dietary intake, blood glucose levels increase, which promptly triggers beta cells to secrete insulin. At insulin-sensitive tissues (e.g. liver, muscles, adipose tissue), the presence of insulin stimulates cells to uptake and store carbohydrates, lipids and amino acids, whereas during fasting, these are used for maintaining the body's energy balance. If the beta cells cannot produce enough insulin, glucose levels rise, which ultimately leads to hyperglycaemia. In addition, to maintain energy balance, the body starts breaking stored lipids and accumulates acids called ketones in the blood. If patients are not treated with exogenous insulin, dehydration caused by hyperglycaemia and excessive acid production can, in severe cases, lead to death from diabetic ketoacidosis (Boulpaep, 2017a).

2.1.1 Pathogenesis

Up until just a century ago, type 1 diabetes was a lethal disease with no treatment. The discovery of insulin in 1921 was a revolutionary achievement that has since changed the lives of individuals with type 1 diabetes (Banting et al., 1922). Although, the treatment and understanding of the disease have improved notably over the decades, there is still a lot to be learnt and crucially important discoveries to be made including, answer(s) to the burning question: "What triggers type 1 diabetes?"

Type 1 diabetes results from a selective immune destruction of insulin-producing beta cells in the pancreatic endocrine glands called the islets of Langerhans. These islets contain different types of secretory cells, among which beta cells are the most abundant, secreting insulin, proinsulin, C-peptide and amylin (**Figure 1**) (Boulpaep, 2017a; Cabrera et al., 2006). Insulin synthesis first includes the transcription of preproinsulin, which after cleavage of a leader sequence results proinsulin. In the *trans* Golgi, proinsulin is packed into secretory granules, where a C-domain that connects A and B domains, is enzymatically cleaved. The end products are insulin

molecule that consist of the joined A and B chains and the C-peptide. The insulin granules are stored inside the cell, and upon stimulus, the components inside are secreted into the extracellular space including a 1:1 molar ration of insulin and C-peptide. C-peptide remains in the blood longer than insulin and is therefore used as a marker of endogenous insulin secretion (Boulpaep, 2017a).

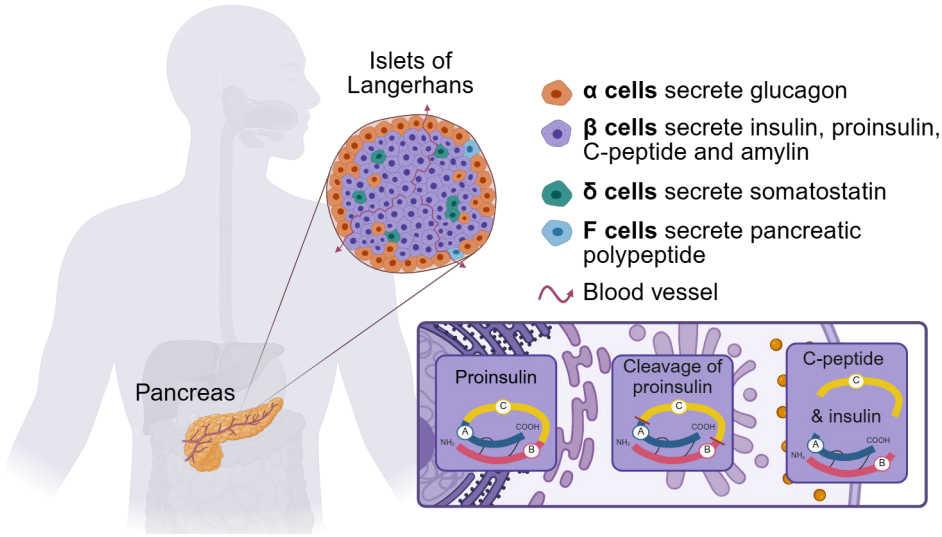

Figure 1. Illustration of the pancreatic islets and insulin processing.

Based on the latest estimates, globally more than 8 million people are living with type 1 diabetes, and annually, around half a million new cases are diagnosed (Gregory et al., 2022; Ogle et al., 2022). Although most diagnoses are made in adulthood, the disease commonly affects children and adolescents and is more prevalent in males than females (Gregory et al., 2022; Östman et al., 2008). Although there are large geographical differences, the worldwide incidence has increased in the past 35 years and is expected to continue to rise (DIAMOND Project group, 2006; Gregory et al., 2022; Patterson et al., 2019). The highest frequency of the disease is observed in Finland, where around 60 cases per 100,000 per year are diagnosed (Harjutsalo et al., 2013). Interestingly, in Finland and other high-incidence countries, the incidence trend has plateaued or even reduced in the past decade (Patterson et al., 2019). Moreover, in the coming years, the relatively largest increase in prevalent cases is expected to occur in lower-income countries (Gregory et al., 2022).

Being born into a family with a type 1 diabetes-affected family member increases the risk for the disease, but the level of risk depends on who has been diagnosed

(Bonifacio and Ziegler, 2010). For example, a young age at diagnosis of the father is associated with a high recurrent risk (Lorenzen et al., 1998). Also, after one identical twin is diagnosed, the cumulative incidence of type 1 diabetes for the other twin is 65% by the age of 60 years (Redondo et al., 2008). Certain genetic risk variants are known to associate with type 1 diabetes. These are mostly located in the major histocompatibility complex (MHC) in chromosome 6 (6p21.3), known as the human leukocyte antigen (HLA) region in humans (Pociot et al., 2010). The molecules encoded in the HLA region can be divided into class I, class II, and class III based on the structure and function of the gene products. The molecules most strongly associated with type 1 diabetes are the highly polymorphic HLA class II cell surface receptors HLA-DR and HLA-DQ. The combinations of these haplotypes can be associated with varying levels of susceptibility or protection (Noble and Valdes, 2011). In particular, haplotypes HLA-DR3–DQ2 and HLA-DR4–DQ8 are associated with a high disease risk. In addition, there are around 60 non-HLA regions reported to affect the risk of the disease, including the insulin gene region (Gootjes et al., 2022). By incorporating multiple risk loci into a genetic risk score, it has been possible to improve identification of at-risk individuals from the general population (Bonifacio et al., 2018).

Consideration of the genetic predisposition alone cannot explain why the worldwide incidence and prevalence of type 1 diabetes have increased over the past decades (Mobasseri et al., 2020; Patterson et al., 2009). Therefore, much attention has been paid to the changes in the environment and lifestyle (Rewers and Ludvigsson, 2016). Indeed, it has been shown that despite an origin in a low-incidence region, growing up in a high-incidence country increases the risk of type 1 diabetes (Oilinki et al., 2012; Söderström et al., 2012). Several environmental factors during pregnancy and early childhood have been studied, including the hygiene hypothesis (Kallionpää et al., 2015; Strachan, 2000), gut microbiota (De Goffau et al., 2013; Vatanen et al., 2018), dietary factors such as infant nutrition (Patelarou et al., 2012) and vitamin D (Infante et al., 2019; Miettinen et al., 2020), chemicals (Predieri et al., 2020) and viral infections, including enteroviruses that are discussed in more detail later in the text.

The combination of genetic predisposition and an environmental trigger initiates the development of islet autoimmunity, which is defined by the persistent presence of islet autoantibodies directed against specific beta cell proteins (Eisenbarth, 1986; Insel et al., 2015). These include biochemical autoantibodies against insulin (IAA), glutamic acid decarboxylase 65 (GADA), tyrosine phosphatase-like insulinoma antigen 2 (IA-2A) and zinc transporter 8 (ZnT8). These islet beta cell antigens localise within the secretory granule, apart from glutamic acid decarboxylase 65, which is found primarily in the membrane of secretory microvesicles (Arvan et al., 2012). Also, classical islet-cell autoantibodies (ICA) can be tested, but these lack

predictive value since the risk of developing type 1 diabetes is similar for ICA-positive individuals and those with no islet autoantibodies (Siljander et al., 2009). Detection of two or more biochemical autoantibodies against any of these targets marks a point at which progression to type 1 diabetes is nearly inevitable (Ziegler et al., 2013). Nevertheless, individuals with a single autoantibody have a much lower risk to develop the disease, and the autoantibody might also appear transiently. It has been shown that, after two years from seroconversion, children who remain positive for a single islet autoantibody have 30% risk of developing type 1 diabetes over the next 15 years, and the risk is only 12% if the child reverts to islet autoantibody negative within two years (Anand et al., 2021).

Because autoantibodies hold predictive value and can be detected even before the clinical onset of the disease, preclinical staging of type 1 diabetes has been introduced (**Figure 2**) (Insel et al., 2015). At stage 1, individuals have developed multiple type 1 diabetes-associated autoantibodies but remain normoglycemic. Progression to stage 2 takes place when multiple autoantibodies are present and dysglycemia is evident. Finally, stage 3 represents the clinical manifestation of type 1 diabetes accompanied by hyperglycemia. The purpose of this preclinical staging is to provide a framework for risk screening that can later be applied in clinical practice and subject stratification for therapeutic interventions. It is noteworthy that this staging is likely to be refined in the future due to the heterogenic nature of the disease.

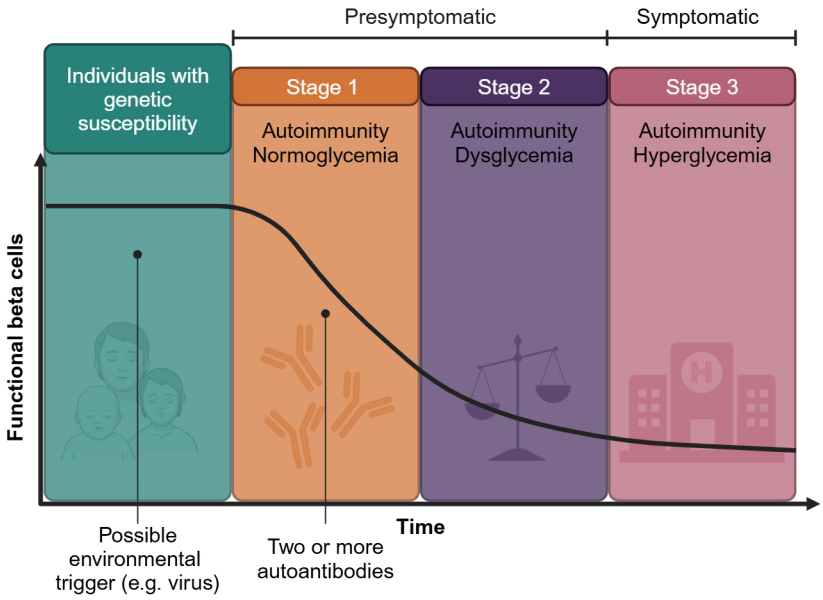

Figure 2. Preclinical staging of type 1 diabetes. Modified from Insel et al. 2015.

At the time of diagnosis, individuals often have symptoms of hyperglycemia including fatigue, weight loss, excessive thirst, and frequent urination. The diagnosis of diabetes is made if the fasting glucose measured in venous plasma exceeds 7.0 mmol/L or if the two-hour plasma glucose during an oral glucose tolerance test (OGTT) or the random plasma glucose exceeds 11.1 mmol/L (Sacks et al., 2023). Alternatively, the diagnosis can be made based on an elevated haemoglobin A1c (HbA1c) concentration (≥ 48 mmol/mol), which indicates the amount of glucose attached to haemoglobin over a time period of 2-3 months. To differentiate type 1 diabetes from other diabetes subtypes, autoantibody and C-peptide measurements have proven useful (Leighton et al., 2017; Sacks et al., 2023). To this day, exogenous insulin has remained the principal treatment for type 1 diabetes.

2.1.2 Heterogeneity

A combination of multiple genetic risk variants and environmental factors contributes to the heterogeneity of the disease process in type 1 diabetes. This includes differences in the type and timing of the first appearing autoantibody, the longitudinal autoantibody pattern, the rate of progression, and the responses to intervention (Endesfelder et al., 2019; Ghalwash et al., 2024). As the first appearing autoantibody, variations of both a single islet autoantibody and multiple autoantibodies are detected. Furthermore, when observing longitudinal autoantibody profiles of individuals, several autoantibody combinations are possible and may include the presence of transient autoantibodies.

To date, two distinct patterns of autoimmunity have been described. These are distinguished by the first appearing autoantibody, which is hereafter referred to as seroconversion (Giannopoulou et al., 2015; Ilonen et al., 2013; Lernmark et al., 2023). These two patterns, initiated with either IAA or GADA, are associated with specific HLA DR-DQ haplotypes as well as non-HLA variants (Ilonen et al., 2019). Individuals with IAA first, seroconvert at a relatively younger age, with a peak of incidence between one and two years of age, and carry more often the HLA-DR4-DQ8 haplotype. In contrast, individuals with GADA as the first autoantibody have the peak incidence later, between two and five years of age, and are more commonly carriers of the HLA-DR3-DQ2 haplotype (Giannopoulou et al., 2015; Ilonen et al., 2013; Lernmark et al., 2023). Additionally, childhood-onset type 1 diabetes is frequently characterised by the presence of IAA. However, in adult-onset type 1 diabetes, multiple disease-associated autoantibodies are detected less frequently, especially as age increases, but GADA is often present (Leslie et al., 2021).

Histopathological findings of human pancreatic samples from individuals with type 1 diabetes support the separation of these suggested disease subtypes, which also seem to express distinct immunological profiles (Redondo and Morgan, 2023).

The autoimmune attack tends to be more aggressive in young children as illustrated by a substantial presence of islet-infiltrated lymphocytes in the islets. In contrast, in adult patients, only mild lymphocytic infiltration has been detected.

Such heterogeneity and the existence of well-defined disease subtypes may indicate distinct triggering mechanisms. Hence, associations between different environmental determinants and the type of the first appearing autoantibody have been explored, including dietary factors, exposures, life events, and infections (Ilonen et al., 2019; Lernmark et al., 2023). To date, most of the reported risk associations have been related to IAA as the first appearing autoantibody and include, for example, coxsackie B virus (CVB) infections (Ilonen et al., 2019; Sioofy-Khojine et al., 2018).

2.1.3 Enteroviruses

Enteroviruses belong to the large family of *Picornaviridae* viruses, that are common human pathogens causing a wide spectrum of diseases, ranging from mild seasonal childhood infections to life-threatening conditions such as myocarditis (Tuthill et al., 2010; Zhang et al., 2021). These non-enveloped, positive-sense single-stranded RNA viruses have a diameter of approximately 25–30 nm. The four structural proteins, VP1-4, are used to construct the icosahedral viral capsid, inside which the viral genome is stored. The non-structural proteins vary slightly depending on the genus. Enteroviruses express seven non-structural proteins: 2A-C and 3A-D, which are involved in the processing of the polyprotein, manipulation of the host cell environment, and viral genome replication (Tuthill et al., 2010).

Enterovirus infections have been associated with type 1 diabetes development and islet autoimmunity in many studies (Nekoua et al., 2022; Wang et al., 2021). For example, enteroviral RNA has been found in the pancreatic islets of type 1 diabetes patients and autoantibody-positive individuals more often than in the healthy population (Geravandi et al., 2021; Ylipaasto et al., 2004). In addition, viral capsid protein VP1 has been detected more frequently in the insulin-containing islets of the patients compared to non-diabetic controls (Richardson et al., 2013). Also, in a unique Norwegian study, where pancreatic biopsies from six living patients with newly diagnosed type 1 diabetes were studied, both enteroviral RNA and VP1 were found in the islets (Krogvold et al., 2022, 2015). Further supporting the involvement of enterovirus infection, the presence of enteroviruses has also been detected in other clinical samples, such as in the blood (Oikarinen et al., 2011) and intestine (Oikarinen et al., 2008) of type 1 diabetes patients, as well as in stool samples of children collected prior to or at seroconversion (Honkanen et al., 2017; Kim et al., 2019).

Among enterovirus species, coxsackie B virus (CVB) serotypes, in particular, have been associated with type 1 diabetes. (Nekoua et al., 2022). There are six different CVB serotypes, CVB1-6, which are considered either risk factors or protective (Laitinen et al., 2014). In many studies, especially CVB1 infection has been linked to the pathogenesis of type 1 diabetes (Oikarinen et al., 2014; Sioofy-Khojine et al., 2018). CVBs are considered as cytolytic viruses, which after replication release the viral particles through lysis of the host cell. Interestingly, CVBs can also establish persistent infections, of which there are two types: steady-state persistence and carrier-state persistence (Frisk, 2001). In steady-state persistence, nearly all cells are infected, but cytolysis is not induced during viral replication cycle. In the more frequently described carrier-state persistence, a low number of cells is infected, but high titres of virus particles are produced. Carrier-state persistence of CVB has been reported *in vitro* in human non-insulin producing pancreatic beta cells (Nekoua et al., 2020) and ductal cells (Alidjinou et al., 2017).

The gastrointestinal tract and the respiratory tract are the primary replication sites of enteroviruses (van der Linden et al., 2015). However, from there, the virus can spread to other organs, such as the pancreas via the blood circulation. The presence of Coxsackie and adenovirus receptor (CAR) in the pancreatic beta cells might partly explain why enteroviruses target these cells (Ifie et al., 2018). Based on studies in pancreatic cell models and CVB-infected primary human pancreatic islets, CVBs are likely to establish carrier-state persistent infections in the islets (Chehadeh et al., 2000). It is suggested that these viruses may use alternative, non-lytic ways, such as vesicles or cellular protrusions for viral egress (Netanyah et al., 2020; Owusu et al., 2021). CVB-infected ductal epithelium cells have also been frequently reported and may contribute to virus spread to beta cells (Geravandi et al., 2021; Ylipaasto et al., 2004).

2.1.4 Prospective cohort studies

Considering the heterogeneity of type 1 diabetes, the disease likely exhibits diverse pathogenesis and therefore, an optimal therapy varies among individuals. To study the pathogenesis of the disease and to identify and direct the right individuals for certain clinical interventions, several type 1 diabetes-associated prospective cohort studies have been initiated, which provide clinical samples and data for research use. Depending on the study, the time period of interest has been set on the preclinical period, starting from birth, the post-diagnosis period, or both.

Among the first preclinical population-based studies, the Finnish Type 1 Diabetes Prediction and Prevention (DIPP) study was established in 1994 and is still ongoing (Kupila et al., 2001). The DIPP study screens new-born babies in the cities of Turku, Tampere, and Oulu, and recruits individuals carrying the HLA risk alleles

to participate in the study and attend frequent follow-up visits. Similar country-specific population-based studies have been initiated, for example, in Sweden (Larsson et al., 2004) and Germany (Ziegler et al., 1999). Global consortiums have launched large studies, such as the Environmental Determinants of Diabetes in the Young (TEDDY) study (Rewers et al., 2008), which investigates the cause of type 1 diabetes and the innovative approaches to understanding and arresting type 1 diabetes (INNODIA) study, which examines samples from newly diagnosed individuals and aims to undertake clinical intervention studies (Dunger et al., 2022).

These prospective studies have enabled the longitudinal screening of individuals and the analysis of the clinical data and samples using multiple approaches, including genomics, transcriptomics, proteomics, lipidomics, and metabolomics (Long and Linsley, 2024). These technologies have provided deeper insights into the etiology of the disease and have the potential to reveal patterns that can be used as biomarkers to predict and monitor individuals. As our understanding of type 1 diabetes subgroups increases and several clinical trials are ongoing, these technologies are likely to play a critical role in identifying the right individuals for specific clinical interventions, determining optimal therapeutic windows, and monitoring treatment outcomes.

2.2 Mass spectrometry (MS)-based proteomics

Proteins are functional molecules that serve essential roles in the living organisms. The concept of "proteome" was first introduced three decades ago and represents all the proteins expressed at a specific time by a certain biological system such as cell or tissue (Wilkins et al., 1996). Compared to the genome, which is nearly constant throughout an organism, the proteome is an entity that adjusts under different conditions, such as cell type, developmental stage, or state of health. The interdisciplinary field of proteomics, defined as the large-scale study of proteins, is an exciting area of science that measures the dynamic system of proteins to determine interactions, investigate protein modifications, define their roles in biological processes and cellular functions, as well as understand disease mechanisms (Aebersold and Mann, 2016). Over the past decades and still today, mass spectrometry (MS) has been one of the most widely used methods in the field of proteomics.

2.2.1 MS-based analysis of the proteome

There are around 20,000 genes in the human genome that are predicted to be protein coding (Nurk et al., 2022). However, it is extremely challenging to estimate the entire number of proteins expressed by the human body, since various factors contribute to

the complexity of proteins. These include allelic variations, alternative-splicing that enables production of several proteins from a single gene, addition of post-translational modifications (PTMs), such as glycosylation or phosphorylation that can influence protein function or activity, as well as somatic recombination that refers to gene rearrangement, which for example, enables the diverse antibody repertoire (Aebersold et al., 2018). These distinct molecular forms arising from a single gene can be called proteoforms (Smith et al., 2013).

Proteins are the end result of the genetic information that is expressed and put into action. They are formed by long, unbranched polypeptide chains made from amino acids, of which there are 20 types (**Figure 3**). The chemical character of each amino acid is defined by its side group, whereas the core structure is identical and connects the adjacent amino acids through a peptide bond. The side chain can be non-polar and hydrophobic (water-fearing) or polar and hydrophilic (water-loving), with or without a charge. Each protein is built from a unique sequence of amino acids, which defines its final conformation into the most energetically favourable three-dimensional structure that is stabilised by chemical bonds. Because each amino acid has a specific mass, apart from leucine and isoleucine, which are isobaric, and each protein has a unique sequence of amino acids, this information can be used to detect and distinguish proteins by mass spectrometry.

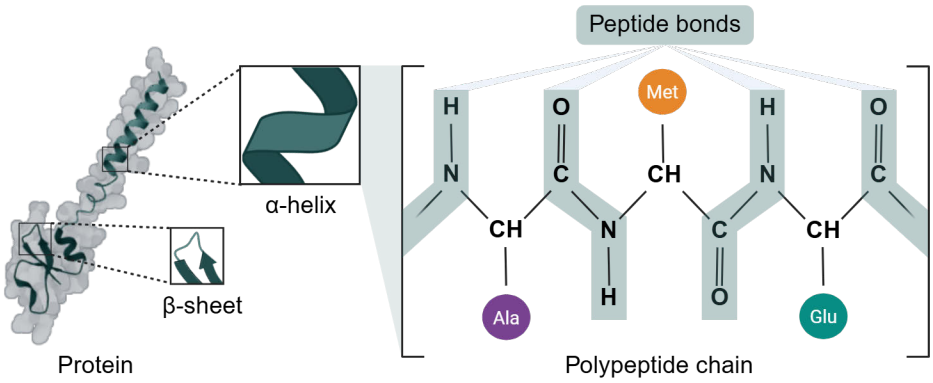

Figure 3. Protein structure. Ala = Alanine, Met = Methionine, Glu = Glutamic Acid.

Mass spectrometry (MS)-based proteomics is a rapidly developing technology for protein detection and quantification (Guo et al., 2025). There are two major approaches used to measure proteins: “bottom-up” and “top-down” proteomics. In bottom-up proteomics, proteins are enzymatically or chemically digested into shorter peptides before MS analysis. Peptides are generally easier to separate, ionise, and fragment than intact proteins, making this approach more sensitive (Aebersold and

Mann, 2016; Switzar et al., 2013). As a result, bottom-up proteomics is currently applied more widely. However, by digesting the proteins, much of the proteoform-level information is lost, because only a limited number of peptides can be detected, and it is challenging to assign a certain peptide to the original proteoform (Dupree et al., 2020). Conversely, in top-down proteomics, intact proteins are directly analysed by MS. This provides comprehensive and more accurate identification and quantitation of different proteoforms. Nevertheless, technical improvements, for example in sensitivity and better informatics tools, are required before it can be fully harnessed for optimal protein detection. Therefore, top-down proteomics approaches continue to develop (Roberts et al., 2024). The third, not as well-known approach is "middle-down" proteomics, which focuses on the analysis of middle-range peptides that are larger than those detected in bottom-up proteomics but smaller than intact proteins (Cristobal et al., 2017).

The earliest MS-based proteomics analyses were based on "in-gel digestion" protocols, where proteins are first separated by gel electrophoresis, then visualised by staining the polyacrylamide gels, and finally enzymatically digested into peptides prior to MS analysis (Shevchenko et al., 1996). Nowadays, other protocols have mostly displaced the use of gel-based methods. Typically, sample preparation starts with the isolation of proteins, which, in the case of cells or tissue, requires a lysis step (**Figure 4A**). Next, a reduction step is needed to break inter- and intramolecular disulfide bonds. Subsequently, the resulting free thiol groups of cysteine residues are alkylated using an alkylating agent, such as iodoacetamide, to prevent the reformation of the disulfide bonds. The proteins are then digested into peptides using sequence-specific enzymes or chemicals (Switzar et al., 2013). In MS-based proteomics, the most commonly used digestion reagent is trypsin, which cleaves the polypeptide chain exclusively at the carboxyl side of arginine (Arg) and lysine (Lys) residues (Burkhart et al., 2012). This results in an optimal average peptide length of around 14 amino acids and at least two positive charges: one at the N-terminus and the other at the C-terminal Arg/Lys. These tryptic peptides are well-suited for MS analysis and their fragmentation pattern is often more reproducible and interpretable.

In the next step, peptides are separated in a high-performance liquid chromatography (HPLC) system to limit the number of peptides entering the MS simultaneously (**Figure 4B**). Different HPLC modes can be applied for peptide separation, which take advantage of the intrinsic characteristics of peptides, including size, net charge, or hydrophobicity (Mant et al., 2007). The most widely used HPLC mode in MS-based proteomics is reversed-phase liquid chromatography, due to its speed, efficacy, and compatibility with MS (Lenčo et al., 2022). Typically, so-called nano-flow columns are used with flow rates in the order of hundreds of microlitres per minute, which improve the sensitivity of the MS analysis. The columns are packed with silica-based support that forms a hydrophobic stationary

phase. The mobile phase often contains a small proportion of an acid, such as formic acid, and a linearly increasing concentration of a strong solvent, such as acetonitrile. The peptides are eluted based on their hydrophobicity. In practice, the increasing concentration of the strong solvent in the mobile phase gradually weakens the affinity of peptides for the hydrophobic stationary phase. Consequently, the most hydrophilic peptides pass through the column first, followed by peptides with hydrophobic side chains in the order of their increasing hydrophobicity.

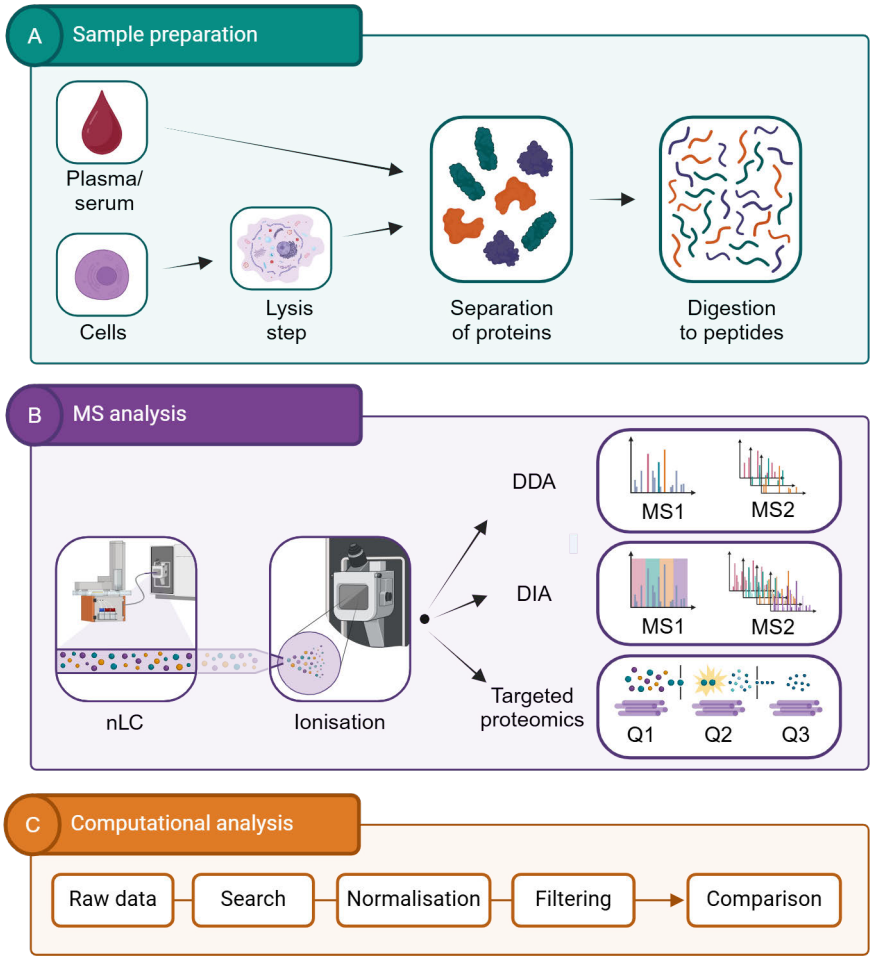

Figure 4. Typical bottom-up proteomics workflows. (A) Proteins are first separated and subsequently digested into peptides. (B) Next, peptide samples are injected into an nLC column, which separates the peptides before ionisation and entry into the mass spectrometer. Samples can be analysed using either discovery proteomics methods, such as data-dependent acquisition (DDA) or data-independent acquisition (DIA), or a targeted proteomics approach. (C) The raw data is then preprocessed and analysed using computational approaches.

Before peptides can enter the mass spectrometer, they need to be ionised. The most commonly used ionisation methods are electrospray ionisation (ESI) and matrix-assisted laser desorption/ionization (MALDI), both of which are considered soft ionisation techniques that do not cause fragmentation of the molecules (Fenn et al., 1989; Tanaka et al., 1988). ESI is particularly favorable because it enables the direct interfacing of nano-flow liquid chromatography (nLC) and thereby the continuous analysis of peptides in solution (**Figure 4B**). In an ESI ion source, peptides are sprayed directly from the capillary column of the nLC, which is held at high voltage to form charged droplets. By solvent evaporation, the charged droplets will shrink to smaller and smaller droplets, ultimately leading to the formation of free gas-phase ions that can enter the MS.

Mass spectrometers are highly sophisticated instruments, which can resolve ions based on their mass-to-charge ratio (m/z). In bottom-up proteomics, mass (m) is the molecular weight of the ionised peptide in Daltons, and charge (z) is the number of elementary charges present on the ionised molecule. Inside the mass spectrometer, ions are introduced into vacuum, which prevents them colliding with other molecules. There are various types of mass spectrometers, which are built using different configurations for ion transfer, filtering, accumulation, and fragmentation. However, by default, all MS instruments are equipped with an ion source, a mass analyser, and a detector.

Just as there are different ion sources, there are also several types of mass analyser. The most commonly used mass analysers are quadrupoles, ion traps, and time-of-flight (TOF), which all have their advantages and disadvantages (Haag, 2016). Quadrupole mass analysers are versatile and can be operated to filter ions based on specific m/z, to transfer ions from one part of the mass spectrometer to another or for ion fragmentation (March, 1997). Therefore, they are often used in tandem, as in triple quadrupole mass spectrometers, or they can serve a task in other mass spectrometers with a different main mass analyser. Despite providing excellent sensitivity, quadrupole analysers have a limited mass range and relatively low resolution. Ion traps differ notably in their fundamentals and design, but as the name suggests, they all trap ions inside electromagnetic fields. One branch of ion traps is Orbitrap, which has two hollow outer electrodes with an inner electrode located within them (Nolting et al., 2019). Accumulated ion packets are injected into the trap through a narrow hole, and due to an electric field between the electrodes, ions start oscillating back and forth of the axis of the inner electrode. The frequency of the oscillating ion ring defines the ions' m/z and can be detected with high mass accuracy. TOF analysers are often combined with MALDI ion source. In the TOF analyser, m/z is measured based on the time it takes from accelerated packet of ions to fly through a flight tube and reach the detector (Mamyrin, 2001). In principle, ions with different m/z will arrive at the detector at different times and can thus be

separated from one another. Orbitrap and TOF analysers are often used in high-resolution mass spectrometers, which perform particularly well during untargeted analysis of complex samples. Compared to quadrupole analysers, these instruments have a higher resolution, a wider mass range, and better mass accuracy (Mirzaei and Carrasco, 2016).

In MS-based proteomics, where accurate detection and quantification of peptides and proteins is required, tandem mass spectrometry (MS/MS) is commonly used. In MS/MS analysis, the entering precursor ions with a given m/z are first analysed (MS1 level), followed by fragmentation of the ions in a collision cell and subsequent analysis of the fragmented ions (MS2 level). During fragmentation, the imparted energy breaks the peptide backbone apart. Most commonly, this happens at the amide bond between the amino acids, resulting in b- and y-ions depending on whether the charge is retained at the amino- or carboxy-terminal part, respectively (Mann and Steen, 2004). Depending on the type of analysis, mass spectrometers can be operated in different modes. Discovery or shotgun proteomics aims to identify as many proteins in the sample as possible without prior knowledge, while targeted proteomics focuses on the analysis of a predefined set of peptides (**Figure 4B**) (Aebersold and Mann, 2016).

For a long time, data-dependent acquisition (DDA) has been the preferred method for discovery proteomics experiments. In DDA, a full spectrum of peptide ions is first measured, followed by a selection of the most intense peptide ions, subsequent fragmentation of these ions, and acquirement of the fragmentation spectra (Aebersold and Mann, 2016). However, there are often missing values observed across multiple measurements, which compromises the quantitative accuracy and the reproducibility of the analysis. Nowadays, an upward trend in the proteomics field has been the replacement of the DDA methods by data-independent acquisition (DIA) methods. In DIA, using sequential, marginally overlapping precursor isolation windows, peptide ions in the entire m/z range (usually 400-1200 m/z) are measured and fragmented (Fröhlich et al., 2024). As a result, more complete proteome coverage is achieved with lower number of missing values. Additionally, more precise protein quantitation is gained, as MS2-level information is also used for quantification. However, compared to DDA, the fragmentation spectra are far more complex, which has been a major hurdle. Recently, drastic developments in DIA analysis software tools have overcome this obstacle, paving the way for deep protein coverage and accurate quantitation (Fröhlich et al., 2024; Mann et al., 2021).

Targeted proteomics is commonly used for the verification of discovery proteomics results due to the high sensitivity, quantitative accuracy, and reproducibility of the method (Borras and Sabido, 2017). These qualities also make it an ideal approach for the detection of specific low-abundant proteins that are often difficult to capture with DDA and DIA methods. However, only a limited number of

target peptides can be analysed in a single run, ranging from tens to hundreds. The two main targeted approaches are selected reaction monitoring (SRM), also called multiple reaction monitoring (MRM), and parallel reaction monitoring (PRM). In both approaches, an inclusion list is first created that defines the target peptides with known m/z and chromatographic retention time (RT). SRM experiments are performed in triple quadrupole mass spectrometers. The first quadrupole filters the listed precursor ions, the second quadrupole acts as a collision cell to fragment these ions, and the third quadrupole filters the selected fragment ion for detection. In SRM experiments, the inclusion list also includes m/z of the fragment ions, whereas in PRM experiments the whole fragment spectra of the precursor is measured in parallel. Therefore, PRM experiments are performed in high resolution mass spectrometers, such as in Orbitrap instruments (van Bentum and Selbach, 2021).

The raw data from mass spectrometers containing peptide and fragment ion spectra is then processed using proteomics software packages that enable peptide identification and quantification, such as MaxQuant (Cox and Mann, 2008) and Skyline (MacLean et al., 2010) (**Figure 4C**). Often these software have built-in search engine that uses sequence database, commonly in a form of FASTA file, for peptide identification. Spectral features from peptides representing a specific protein are then combined to calculate a single intensity value for that protein. The more commonly used approach for protein quantification is relative quantification, where the abundances of the same protein are compared between samples (Sinha and Mann, 2020). An alternative approach is absolute quantification, which defines the protein concentration in a sample and can be used to compare the same protein between samples or different proteins within a sample. However, absolute quantification is still very limited in its use due to analytical and technical challenges, which relate, for example, to reproducibility and the non-quantitative nature of ESI-MS (Nosti et al., 2022). Relative quantification can either be done using a label-free quantification (LFQ) or a label-based quantification strategy (Lindemann et al., 2017). In LFQ, the extracted raw data is normalised and then protein levels are compared between different conditions. LFQ is a cost-efficient and straightforward strategy that works well for large sample numbers. However, compared to label-based quantitation, it is more prone to quantitative variation if not carefully monitored (Sinha and Mann, 2020). Label-based quantification is commonly performed using stable isotopes. As these labelled peptides are chemically identical, their behaviour is exactly the same, but their mass is different, thus allowing differentiation of samples. In targeted analysis, synthetic heavy reference peptides can be spiked into the samples and used for peptide quantification, as well as for data normalisation (van Bentum and Selbach, 2021). There are multiple software tools available for the downstream processing and statistical analyses of proteomics data. (Chen et al., 2020).

2.2.2 The blood plasma/serum proteomics

Blood is the primary clinical specimen because it is easily collected and contains diagnostically relevant information (Anderson, 2010). Blood has four main components: plasma, red blood cells, white blood cells and platelets. Plasma is watery fluid that is rich in proteins but also consists of electrolytes, carbohydrates, and lipids. The difference between plasma and serum is the presence of an anticoagulant. To obtain plasma, an anticoagulant needs to be added, whereas serum is obtained after clotting and therefore lacks fibrinogen and other related coagulation factors (Boulpaep, 2017b).

The total protein concentration of plasma is approximately 60-80 μg/ml, and it represents a complex proteome comprising more than 10,000 proteins (Nanjappa et al., 2014; Tirumalai et al., 2003). These proteins are derived from different cells and tissues throughout the body. As most diseases alter the endogenous protein expression, changes observed in the plasma proteome can be indicative of the disease. In addition, proteins of pathogenic origin can be found (Anderson and Anderson, 2002). The most dominating protein in plasma is albumin, which covers more than 50% of the total protein content. In addition, there are another 21 proteins, including immunoglobulins, transferrin, haptoglobulin, complement factors, and apolipoproteins, which together with albumin cover nearly 99% of the total protein content (Tirumalai et al., 2003). The remaining 1% represents the diverse group of low abundant proteins. This group is often considered the most interesting part of the plasma proteome, as it is the most promising source of new biomarkers (Millioni et al., 2011). The deeper the plasma proteome can be explored, the more likely it is to detect clinically significant proteins, such as cellular ligands, signal proteins, or tissue leakage proteins that, as a result of disease, are found in plasma (Geyer et al., 2017; Millioni et al., 2011). The dynamic range, which exceeds 10 orders of magnitude between albumin and the rarest proteins, is one of the major challenges in plasma proteomics (Anderson and Anderson, 2002). It is currently possible to detect around 1,000 proteins from neat plasma using state-of-the-art mass spectrometers (Niu et al., 2025). Although analysis of neat plasma is cost-efficient and straightforward, there is a strong interest in achieving much deeper coverage of the proteome.

To overcome this challenge and improve the detection of low-abundance proteins, several strategies have been developed. In MS-based proteomics, immunodepletion of high-abundance proteins and sample fractionation have commonly used. Nevertheless, by depleting the samples, cross-reactivity of antibodies to similar epitopes and proteins bound to the depletion targets may cause bias in the results (Bellei et al., 2011; Geyer et al., 2016). Sample fractionation, however, is often laborious and can compromise throughput, which might diminish its applicability. In recent years, alternative enrichment approaches have entered the

market. These include Seer's nanoparticles, which, through their physicochemical properties, induce the assembly of a corona of enriched proteins on their outer surface that can be analyzed (Blume et al., 2020). With this technology it is possible to detect up to 4000 proteins from plasma (Beimers et al., 2025; Huang et al., 2025). Alternatively, there are magnetic bead-based enrichment methods, such as PreOmics® ENRICH-iST technology, which also targets low-abundant proteins, and Mag-Net by ReSyn Biosciences™ (Wu et al., 2024), which focuses on the enrichment of extracellular vesicles. With these technologies it is possible to detect over 2000 proteins or even up to 5,000 proteins from plasma (Beimers et al., 2025; Heil et al., 2023; Wu et al., 2024). Interesting results have also been obtained with an affordable method that uses perchloric acid for selective protein precipitation to increase low abundant proteins (Viode et al., 2023).

Additionally, there are non-MS-based proteomics technologies available. These include antibody- and aptamer-based affinity proteomics technologies, such as Olink® Explore, Olink® Reveal and SomaScan by SomaLogic, respectively. These technologies enable targeted detection of several thousand proteins in plasma and serum (Cronje et al., 2023; Sun et al., 2023). The technology of choice often depends on multiple factors such as cost, availability, throughput and accuracy.

Another challenge arises from pre-analytical variability. Like all proteomes, the plasma proteome is dynamic and represents the proteins present at the specific moment when the sample is taken, potentially causing variability in the data. This variability might be due to recent exercise or food intake. In addition, sample handling before and during blood processing, sample storage, and the number of freeze-thaw cycles are all sources of variation and thus may cause alterations in plasma proteome composition (Ignjatovic et al., 2019). These factors may affect protein identification, quantitation, reproducibility, PTMs, and the overall data quality. Furthermore, understanding the difference between plasma and serum is important, as the type of sample can affect the levels of many proteins. (Geyer et al., 2019). Therefore, detailed documentation, standardised pre-analytical procedures, and sufficient quality controls are key elements in producing high-quality data containing relevant information.

2.2.3 Type 1 diabetes-associated plasma/serum proteomics studies

The published type 1 diabetes-associated plasma and serum proteomics studies (**Table 1**) have primarily focused on identifying disease-associated changes by comparing samples from already diagnosed individuals (De Oliveira et al., 2018; Metz et al., 2008; Zhang et al., 2013; Zhi et al., 2011) or from genetically susceptible children before the onset of the disease (Liu et al., 2018; Moulder et al., 2015;

Nakayasu et al., 2023; von Toerne et al., 2017) to matched controls. These studies have provided important information on type 1 diabetes-associated protein alterations in the plasma and serum, although the reported results between studies are somewhat inconsistent.

The earlier type 1 diabetes-related proteomics publications have focused on recently diagnosed individuals (Metz et al., 2008; Zhang et al., 2013) or individuals who have been living with type 1 diabetes for a longer time (Zhi et al., 2011). In these studies, pooled cross-sectional serum and/or plasma samples that were either fractionated (Metz et al., 2008) or immunodepleted (Zhang et al., 2013; Zhi et al., 2011) prior to discovery MS analysis were used. For the validation of the discovery results in the studies by Zhi *et al*. (2011) and Zhang *et al*. (2013), individual samples were analysed, as was also done in the later published study by De Oliveira *et al*. (2018). While comparing the study results, no overlap between the significantly altered proteins in these post-diagnosis studies was found. By extending the comparison to significant proteins identified in the discovery studies, six overlapping proteins with similar trends reported in at least two separate studies were found. These included zinc-alpha-2-glycoprotein (AZGP1), corticosteroid-binding globulin (SERPINA6) and lumican (LUM), which were increased in diagnosed individuals as well as clusterin (CLU), complement C3 and complement C4 that were decreased.

The latter studies of type 1 diabetes progression, have mostly analysed longitudinal pre-onset serum or plasma sample series obtained from prospective cohort studies. Considering the heterogeneity in type 1 diabetes development, some studies have analysed rather heterogeneous populations (Liu et al., 2018; Moulder et al., 2015), whereas others have focused on more homogeneous groups of individuals (Nakayasu et al., 2023; von Toerne et al., 2017). In these studies, individual samples, which were immunodepleted before MS analysis were measured, except in the study by Nakayasu *et al*. (2023), where pooled samples in the discovery phase were used. Similarly to post-onset studies, the significant results between the studies were relatively inconsistent. By compering the significantly altered proteins in these studies, 22 proteins were found to be reported as significant in at least two out of the four studies. Among these, only nine proteins show consistent trends in at least two studies. These included biotinidase (BTD) and complement component C8 alpha chain (C8A) that were were increased in those individuals who developed islet autoantibodies and/or progressed to type 1 diabetes. The decreased proteins included three apolipoproteins (APO), namely APOC4, APOA4, and APOC2, as well as mannose-binding protein C (MBL2), peptidase inhibitor 16 (PI16), vinculin (VCL), and complement C3. As the proteomics technologies continue to develop and the study designs improve, more accurate disease-associated protein signatures are likely to be acquired.

Table 1. Type 1 diabetes-associated MS-based proteomics studies. Case = individual diagnosed with type 1 diabetes, Ctrl = control, AAb+ = autoantibody-positive individual, SC = seroconversion

Study	Pre/post diagnosis	Sample type	Study cohort	Samples	Age range	Key results
Metz T *et al*. 2008	Post	Plasma and Serum	10 case-ctrl pairs	Coss-sectional, pooled (2)	< 30 years	5 proteins with significant differences
Zhi W *et al*. 2011	Post	Serum	Discovery: 30 case-ctrl pairs, Validation: 1139 case and 848 ctrl	Coss-sectional, Discovery: pooled, Validation: Immunoassay, individuals	0-80+ years	6 validated proteins with significant differences
Zhang Q *et al*. 2013	Post	Serum (and plasma)	Discovery: 50 case and 100 ctrl, Targeted: 10 case-ctrl pairs	Coss-sectional, Discovery: pooled, Targeted: individual	Discovery: 10-29 years, Targeted: 5-23 years	6 validated proteins with significant differences
De Oliveira *et al*. 2018	Post	Serum	30 case-ctrl pairs	Cross-sectional, individual	18-59 years	8 proteins with significant differences
Moulder *et al*. 2015	Pre	Serum	19 case-ctrl pairs	Longitudinal, 5-11 samples/ individual	3 months - 12 years	38 proteins with significant differences
von Toerne *et al*. 2016	Pre	Serum	Discovery: 15 rapid AAb+, 15 slow AAb+ and 15 ctrl, Targeted: 70 slow AAb+ - ctrl pairs	Discovery: 2 samples/ individual Targeted: cross-sectional	Discovery: Median age 2.1-5.1 years Targeted: Median age 3.2 years	26 validated proteins with significant differences
Liu *et al*. 2018	Pre	Plasma	11 case and 10 ctrl	Longitudinal, 9 samples/ individual	9 months-14 years	13 proteins with significant differences
Nakayasu *et al*. 2023	Pre	Plasma	Discovery: 46 case-ctrl pairs, 46 AAb+ -ctrl pairs, Targeted: 94 case-ctrl pairs, 401 AAb+ -ctrl pairs	Discovery: pre vs. post SC, pooled, Targeted: longitudinal, 7 samples/ individual	9 months - 6 years	83 validated proteins with significant differences

3 Aims

The aim of this doctoral dissertation was to provide molecular insights into the development of type 1 diabetes, starting from its early signs to diagnosis and beyond. An additional goal was to understand how persistent enterovirus infection, which is associated with type 1 diabetes in many studies, modulates the protein expression and secretion of pancreatic cells. The key technology used in all these studies was mass spectrometry-based proteomics.

The specific aims of this dissertation were:

1. To identify early type 1 diabetes-associated molecular changes by studying longitudinal plasma and serum samples of young children who progressed to type 1 diabetes before the age of 5 years (**Study I**).
2. To evaluate and validate a panel of proteins previously associated with type 1 diabetes using longitudinal serum samples from newly diagnosed youth collected during the first year after diagnosis (**Studies II** and **III**).
3. To understand how persistent enterovirus infection modulates protein expression and secretion in pancreatic ductal cells (**Study IV**).
4. To investigate whether the virus-induced proteomic changes are detectable in the plasma/serum samples of children who progress to type 1 diabetes.

4 Materials and Methods

Figures 1-5 and 7 presented in this doctoral dissertation were created with Biorender.com. In **Figure 3**, the protein structure (beta-defensin 1) was obtained from AlphaFold Protein Structure Database (Jumper et al., 2021; Varadi et al., 2024, 2022).

4.1 Ethical considerations

In this dissertation, human serum and plasma samples were analysed, which were obtained either from the Finnish Type 1 Diabetes Prediction and Prevention (DIPP) study (**Study I**) or from the pan-European Innovative approaches to understanding and arresting type 1 diabetes (INNODIA) study (**Studies II** and **III**). In both studies, the guidelines of the Declaration of Helsinki for research on human participants were followed. The ethical committees of the participating hospitals approved the study protocols. Either the parents or the participants gave their written informed consent for the participation.

4.2 Samples

4.2.1 Human plasma and serum

In **Study I**, longitudinal non-fasting plasma and serum samples from 92 DIPP children (≤ 5 years of age) with HLA-conferred genetic susceptibility for type 1 diabetes were analysed. These samples were collected between 1996 and 2015 in the cities of Tampere (plasma) and Turku (serum). The plasma samples from 30 DIPP children were used for the discovery proteomics analysis and included up to four follow-up samples per child, collected between 3 and 36 months of age. The cohort included 10 multiple autoantibody positive children who had progressed to type 1 diabetes before the age of four years (progressors) and their matched single autoantibody-positive (1AAb+) and autoantibody-negative (AAb-) children. The matching was based on HLA risk group, sex, region and date of birth. For the subsequent targeted proteomics analysis, 62 DIPP children were selected. From each

child a median of nine follow-up samples was received, collected between 2 and 60 months of age. This cohort included samples from 31 progressors, who had developed type 1 diabetes before the age of five years and their matched AAb- children.

In **Studies II** and **III**, longitudinal serum samples were measured from "the first 100" (**Study II**) and "the next 150" (**Study III**) newly diagnosed (ND) INNODIA participants and their first-degree relatives. The INNODIA study centres around Europe used tightly controlled sample handling protocol while collecting samples between 2018 and 2022 (Dunger et al., 2022). The samples from ND were collected within 6 weeks of diagnosis and then 3, 6, and 12 months after. Cross-sectional samples from autoantibody-negative unaffected family members (UFMs) served as reference. In both studies, the data used in the comparisons was restricted to participants ≤ 18 years old resulting 86 ND (288 samples in total) and 194 UFMs in **Study II** and 146 ND (560 samples in total) and 272 UFMs in **Study III**. In addition, three samples were specifically selected and used as quality controls. From all ND youth, fasting C-peptide and fasting serum glucose measurements were taken. The fasting C-peptide/glucose ratios were used to estimate the amount of functional beta cells and as an indication of the rate of disease progression.

4.2.2 Human pancreatic ductal cells (PANC-1)

In **Study IV**, the effects of persistent enterovirus infection in human pancreatic ductal cells (PANC-1, from Professor Didier Hober's laboratory in France) was studied. The persistent infection model was established in Professor Heikki Hyöty's laboratory at Tampere University (Honkimaa et al., 2020). In brief, PANC-1 cells were infected with two different CVB1 strains: prototype strain ATCC (American type culture collection, strain Conn-5) and wild type strain 10796, originally isolated in 1983 in Argentina (Hämäläinen et al., 2014). A low virus titre was used for the establishment of the primary infection, which instantly caused a strong cytopathic effect and extensive cell death (**Study IV**/Figure 1A). The surviving cells started to recover and were maintained for one year post infection. Alongside non-infected control cells were grown. The presence of virus was monitored at regular basis using quantitative reverse transcription polymerase chain reaction (RT-qPCR) as described previously (Honkanen et al., 2013). At day 362 after initial infection, the presence of viral capsid protein VP1 was confirmed by immunohistochemistry using formalin-fixed paraffin-embedded (FFPE) cell samples stained with VP1 detecting monoclonal antibody (M7064, clone 5-D8/1, Agilent Dako), as previously described (Krogvold et al., 2015).

For the proteomics analysis, three biological replicates of infected and non-infected cell samples and cell culture supernatants were collected 300-322 days after

the initial infection. Four days before sample collection, the cells were transferred to serum-free DMEM medium supplemented with antibiotics, and two days before collection the cell viability was estimated by ToxiLight™. Scraped cells were pelleted by centrifugation and the dry cell pellets were stored at -80 °C. The cell culture supernatants were cleared by centrifugation and stored at -80 °C.

4.3 Sample preparation

4.3.1 Plasma proteomic profiling

In total 101 plasma samples from 30 DIPP children were prepared for the discovery proteomics analysis in **Study I**. The samples were prepared and analysed batch-wise in a blinded fashion. First, 8 µl of plasma was taken, from which the 12 most abundant plasma proteins were depleted using Pierce™ top 12 spin columns (Thermo Scientific). The lower abundance proteins were then precipitated with ice-cold acetone and stored at -20 °C overnight. The precipitate was then dissolved in 150 µl of 8 M urea, followed by reduction with dithiothreitol (DTT) and alkylation with iodoacetamide. Before overnight trypsin digestion, the urea concentration was adjusted to 1.4 M. The samples were then acidified and desalted doubly with Sep-Pak C18 cartridges (50 mg, Waters), dried and stored at -20 °C. Before MS analysis the samples were reconstituted into 2% formic acid, 2% acetonitrile. The protein concentration in each sample was estimated using a NanoDrop-1000 UV spectrophotometer (Thermo Scientific), after which retention time standard peptides (iRT, Biognosys), used as quality control, were spiked into each sample.

4.3.2 Targeted serum proteomics

In total 524 serum samples in **Study I**, 482 serum samples in **Study II** and 832 serum samples in **Study III** were prepared for targeted proteomics analyses. The samples were prepared and analysed batch-wise in a blinded fashion. First, 4 µl of undepleted serum was directly diluted with 8 M urea. Otherwise, sample preparation protocol followed the same steps as described above in the discovery proteomics paragraph with minor changes. The samples were prepared in 96 well plates. To ensure higher throughput, reduce human error, and maintain consistency in sample preparation, a semi-automated sample preparation protocol using Biomek^{NX} robot was developed for **Study III**. Desalting was performed using 96 well Sep-Pak C18 plates (100 mg, Waters). After drying and storing at -20 °C, the samples were reconstituted either with 2% formic acid, 2% acetonitrile (**Studies I** and **II**) or with 0.1% formic acid, 2% acetonitrile (**Study III**). Synthetic peptide analogues labelled with heavy

isotopes (PEPotec, Thermo Fischer Scientific) were spiked into the samples (~10 fmol/µl) together with retention time standard peptides (MSRT1, Sigma-Aldrich).

4.3.3 Cell lysates

The cell pellets were lysed in 200 µl of 4% SDS, 0.1 M DTT in Tris HCl. To inactivate the virus, the sample tubes were incubated at 95 °C for 5 min. Subsequently, the samples were sonicated to reduce viscosity of the cells. The samples were then centrifuged, after which the protein containing supernatant was collected. Filter-aided sample preparation (FASP) protocol with slight modifications was followed for detergent removal and protein digestion (Wiśniewski et al., 2009). In brief, around 60 µg aliquots of the lysates were first mixed with 400 µl of 8 M urea in Microcon-30 kDA filter units (YM-30, Millipore) and centrifuged. Two rinsing steps with 400 µl of 8 M urea were then performed. Subsequently, 300 µl of 0.05 M iodoacetamide was added to each filter and incubated. After this, the filters were centrifuged. The samples were then washed two times with 400 µl of 8 M urea. Prior to overnight trypsin digestion, the buffer was changed by washing the samples twice with 400 µl of 10 % acetonitrile in 25mM Tris/HCl. In the following day, the peptide containing filtrates were collected by centrifuging the samples and rinsing the filters twice with 250 µl of the digestion buffer. Before desalting, the samples were dried and reconstituted with 1 ml of 0.1% trifluoroacetic acid. Desalting was performed using 96 well Sep-Pak C18 plates, after which the samples were dried. Prior to MS analysis, the samples were reconstituted with 2% formic acid, 2% acetonitrile and the protein concentration was determined with NanoDrop-1000 UV spectrophotometer (Thermo Scientific).

4.3.4 Cell culture media

Cell culture media was first concentrated by centrifuging 7 ml of media in 10 kDa Amicon Ultra-15 Centrifugal filter units (Merck Millipore). The virus was then inactivated by heating the concentrated samples at 95 °C for 5 min. A second concentration step was performed by centrifuging the sample in 3 kDa Amicon Ultra-0.5 Centrifugal filter units (Merck Millipore), after which buffer exchange was performed by adding 250 µl of 8 M urea in the filter and centrifuging. The protein containing concentrate inside the filter was then transferred into a new tube. After that, reduction and alkylation steps were performed. Before overnight trypsin digestion, the urea concentration was reduced to 1.5 M. In the next day, the samples were acidified and then desalted using 96 well Sep-Pak C18 plates and dried. Prior to MS analysis, the samples were reconstituted and protein concentration was measured, as described earlier.

4.4 LC-MS analyses

4.4.1 Data-dependent acquisition (DDA)

The peptide samples were analysed in triplicate using a high resolution Q Exactive Hybrid Quadrupole-Orbitrap (**Study I**) or Q Exactive HF Hybrid Quadrupole-Orbitrap (**Study IV**) mass spectrometer (Thermo Fisher Scientific), both coupled with EASY-nLC 1000 liquid chromatographs (Thermo Fisher Scientific) and a nano-electrospray ion sources (Thermo Fisher Scientific). In **Study I**, the peptide samples were first loaded into a 2 cm pre-column (inner diameter 100 µm, packed with 5 µm C18 silica particles from Michrom Bioresources) in buffer A (0.1% formic acid, 2% acetonitrile), followed by separation using an in-house made 15 cm analytical column (inner diameter 75 µm, 5 µm C18 silica beads by Michrom Bioresources). In **Study IV**, in-house made 40 cm analytical column (inner diameter 75 µm, 1.9 µm ReproSil-Pur C18 beads from Dr. Maisch) was used and maintained at 60 °C throughout the runs using an in-house made column oven. The peptides were separated using a linear 90 min (**Study I**) or 110 min **(Study IV**) binary gradient (buffer A and buffer B: 0.1% formic acid, 95% acetonitrile). The most intense peptide precursor ions (10 in **Study I**, 12 in **Study IV**) were selected for fragmentation using a high-energy C-trap dissociation. To avoid repeated selection of identical precursor ions, the dynamic exclusion was set to 20 s.

4.4.2 Selected reaction monitoring (SRM)

Preselected peptides presenting 12 proteins in **Study I**, 98 proteins in **Study II** and 21 proteins in **Study III** were measured using the SRM approach. A TSQ Vantage Triple Quadrupole mass spectrometer (Thermo Fisher Scientific) interfaced with EASY-nLC 1000 (Thermo Fisher Scientific) (**Studies I** and **II**) or Evosep One (Evosep) (**Study III**) liquid chromatograph and a nano-electrospray ion source (Thermo Fisher Scientific) was used to obtain the SRM measurements. To develop a retention time-scheduled data acquisition methods, Skyline software was used. In **Studies I** and **II**, the peptide samples were separated using the same two-column setup as described in the DDA-approach of **Study I**. In **Study III**, the peptide samples were loaded onto Evotip Pure disposable trap columns and separated using an 8 cm analytical column (EV1094 Endurance column, inner diameter 100 µm, 3 µm C18 silica beads by Dr. Maisch). For peptide separation linear 60 min (**Study I**), 45 min (**Studies I** and **II**) or 24 min (**Study III**) binary gradient was used.

4.5 Preprocessing of the raw data

4.5.1 MaxQuant

The raw data obtained from the high resolution Orbitrap mass spectrometers (**Studies I** and **IV**) was preprocessed with MaxQuant software version 1.5.5.1 (Cox and Mann, 2008) integrated with Andromeda search engine (Cox et al., 2011). The peptide lists were searched against combined SwissProt human and TrEMBL enterovirus protein sequence database with added iRT peptide sequences and a common contaminants database. Label-free quantification (LFQ) was selected with the LFQ minimum ratio count of 2. Enzyme specificity was set to trypsin and a maximum of two missed cleavages was allowed. N-terminal acetylation and methionine oxidation were selected as variable modifications and cysteine carbamidomethylation as a fixed modification. A false discovery rate (FDR) was set to 1% at peptide and protein level and was determined by searching a reverse database. A minimum length of seven amino acids was required. To transfer peptide identifications between two raw files, "match between runs" feature was enabled. Otherwise MaxQuant default settings were used.

4.5.2 Skyline

Skyline software (MacLean et al., 2010) was used to manually inspect the raw data from the triple quadrupole mass spectrometers (**Studies I, II** and **III**) by comparing the light (endogenous) and the heavy (synthetic) peptide elution profiles. In **Study I**, Skyline was also used for data normalisation by dividing the total peak area of each peptide with the sum of peak areas of two peptides representing endogenous alpha-1B-glycoprotein (A1BG). In our previous studies, A1BG has been one of the most stable serum proteins detected in longitudinal samples from children and was therefore selected as a global standard (Lietzén et al., 2018; Moulder et al., 2015).

4.6 Data analysis

4.6.1 Perseus software platform

The "ProteinGroups.txt" output file from MaxQuant was analysed with Perseus software platform, which was developed for the analysis of quantitative proteomics data (Tyanova et al., 2016). The LFQ normalised protein intensities were selected and the data was first filtered by removing proteins identified only with one variable modification site or belonging to reverse data base. Then potential contaminants were manually removed and only proteins identified with minimum of two razor

and/or unique peptides were kept in the data matrix. The data from three technical replicates was combined by calculating the median intensity value with the requirement of minimum of two valid values per protein. Finally, only proteins that were quantified in at least half (**Study I**) or one third (**Study IV**) of the samples were kept. In **Study IV**, unpaired Student's t-test was used to find significant differences in protein expression levels between the samples. Proteins with an FDR < 5% and a fold change > 1.5 were considered significantly differentially expressed.

4.6.2 An additive Gaussian process regression model

In **Study I**, the bioinformatics analyses were carried out in collaboration with Associate Professor Harri Lähdesmäki and Dr. Lu Cheng at Aalto University. An additive Gaussian process regression model, LonGP, developed for statistical analysis of longitudinal data, was used for statistical analysis of the discovery and targeted proteomics data (Cheng et al., 2019). Continuous covariates included age (days from birth to sampling date), "sero" (days from seroconversion date to sampling date) and "t1d" (days from type 1 diabetes diagnosis date to sampling date), which were all converted to the unit of months. Discrete covariates included sex, group (progressor, 1AAb+, AAb-), pair id and individual id. Interaction flags were set to false for sero and t1d. For targeted proteomics data, additional discrete covariate "batch" was used as the data was analysed in three batches. The default LonGP parameters were used. To be considered statistically significant effect, the effect had to be included in the final cross-validated model and explained variation had to be more than 1%.

4.6.3 Linear mixed effects model

The bioinformatics analyses in **Studies II** and **III** were conducted in collaboration with Professor Laura Elo, Dr. Tommi Välikangas and Associate Professor Tomi Suomi at University of Turku. To normalise the $\log_2$-transformed data and adjust for batch effects, peptide-wise linear mixed effects models (LMM) were used. There was no need for imputation of the missing values since the proportion of missing data was low (~0.3%). The technical reproducibility of the frequently measured QC samples was good, with a mean coefficient of variation around 3% across all the quantified proteins and the mean Pearson correlation of 0.98 between the QC samples. The statistical analyses were conducted using programming language R version 4.0.0 (R Core Team, 2018), with the R package lme4 version 1.1-27.1 and lmerTest version 3.1-3 to compose the LMMs (Bates et al., 2015). In the analysis of peptide intensity differences and their associations with beta cell function, an FDR

threshold of 0.05 was applied after multiple hypothesis correction using the Benjamini–Hochberg procedure.

Prior to LMM analysis, the fasting C-peptide/glucose ratios were transformed to natural logarithm to facilitate regression analysis. Fixed effects included sex, height and body mass index (BMI) score ("age based BMI" expressed as standard deviation score [BMI-SDS]). Individual and study centre were defined as random effects, with individual nested under study centre. Height and BMI-SDS were applied to control for individual differences in body size while simultaneously controlling for possible age-related effects.

LMMs were also applied to search for associations between peptide slopes and the fasting C-peptide glucose ratios at 24 months, while considering the effects of sex, height, BMI score and study centre. Only ND youth with at least three study visits including the first 6 week visit and the fasting C-peptide/glucose measurement at 24 months were included in the analysis (n=33). In addition, associations between the peptide slopes and the changes between the fasting C-peptide/glucose ratios at 24 months and at 6 weeks (24 months/6 weeks) were studied.

For the comparison of ND youth and UFMs, peptide-wise LMMs were used. These models included age and sex as fixed effects and individual and study centre as random effects, with individual nested under the study centre. In this comparison, BMI score was not used.

In **Study III**, the peptide data was also combined protein-wise by calculating meta p-values using the sum of z (Stouffer's) method (metap 1.10) and adjusted for multiple correction by Benjamini-Hochberg procedure.

4.6.4 Functional analyses and SignalP

In **Study IV**, DAVID functional annotations (Huang et al., 2009) and Ingenuity Pathway Analysis (IPA, Qiagen) were used for functional enrichment analyses. To be considered significant, FDR < 0.05 was required for Gene Ontology classes with DAVID and p-value < 0.01 in IPA pathways.

The SignalP 4.1 server was used to predict the presence of signal peptides among the detected proteins (Petersen et al., 2011). Both SignalP-TM and SignalP-noTM were applied to distinguish between signal peptides and transmembrane regions.

4.7 Further experiments with PANC-1 cells

4.7.1 Quantitative real time-PCR

In **Study IV**, some of the proteomics findings were verified at transcriptional level using real time-PCR. These experiments were carried out at the University of Turku and Karolinska institute, and are described in detail in the manuscript.

4.7.2 Western Blot

For western blots, the cell lysate samples prepared for the proteomics analysis in **Study IV** were used. The samples were first boiled with 6 x loading dye (330 mM Tris-HCl, pH 6.8, 330 mM SDS, 6% β-ME; 170 µl Bromophenol blue; 30% glycerol). The samples were then loaded to Mini-PROTEAN TGX Precast Protein Gels (BioRad Laboratories) and subsequently transferred to PVDF membranes (Trans-Blot Turbo Transfer Packs, BioRad Laboratories). The used primary antibodies were eIF4G (#2617, Cell Signaling Technology) and beta-actin (A5441, Sigma-Aldrich).

4.7.3 In situ hybridisation

QuantiGene ViewRNA *in situ* hybridisation-technique was used for the detection of a viral pattern recognition receptor IFIH1/MDA5 (#VA1-13031) in the FFPE samples collected in **Study IV** at day 102 and 228 post infection. This experiment was conducted at Tampere University following a standard protocol as previously described (Laiho et al., 2015).

4.7.4 Mitochondrial visualisation

The visualisation of mitochondrial network morphology in **Study IV** was conducted at Tampere University and is described in detail in the publication. A minimum of three biological replicates of live cells persistently infected with CVB1 (ATCC and 10796) and non-infected controls were used for the experiment.

After imaging, the cells were categorised based on their mitochondrial network morphology into four groups: elongated filaments, filaments, intermediate, and fragmented network. Elongated filaments represented cells with long filaments as mitochondrial network. The filaments group had shorter filamentous mitochondrial networks and very few to no fragmented mitochondria. The intermediate group was defined as cells with a clear mixture of elongated and fragmented mitochondria. Cells categorised as fragmented had a large majority of a fragmented mitochondrial

network. Out-of-focus and uncategorised cells were also detected and included in the analysis. The proportion of uncategorised cells was 15% in non-infected cells, 3% in CVB1 ATCC cells and 4% in CVB1 10796 cells. For calculation of the percentages in each category, the measurements from each biological replicates were averaged and a standard deviation was calculated separately for each condition. Statistically significant results were defined using z-test.

4.8 Data availability

The raw data from the discovery proteomics analyses have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol et al., 2022) with the data set identifiers PXD033744 (**Study I**), PDX012153 and PDX012154 (**Study IV**). The raw SRM data and Skyline documents of **Study I** are available through Panorama Public (Sharma et al., 2018) (https://panoramaweb.org/APOC1_rapidT1D.url) with the dataset identifier PXD033946.

5 Results

5.1 Plasma/serum proteomics

Longitudinal plasma and/or serum proteomes were studied in the pre-onset samples collected from children who progressed to type 1 diabetes before the age of five years (progressors) (**Study I**) and in samples obtained from youth newly diagnosed (ND) with type 1 diabetes, collected during the first year after diagnosis (**Studies II** and **III**) (**Figure 5**). In all studies, the longitudinal protein/peptide patterns were compared with similar patterns either from matched single autoantibody-positive (1AAb+) and/or autoantibody-negative (AAb-) children (**Study I**) or from autoantibody-negative unaffected family members (UFMs) (**Studies II** and **III**).

In the discovery proteomics analysis of **Study I**, 269 protein groups (hereafter referred as proteins) were identified and quantified with an average of 254 (±17) proteins per sample. Among the selected 98 target proteins (including disease-associated and control proteins) analysed in **Study II**, 79 proteins were detected in **Study I**.

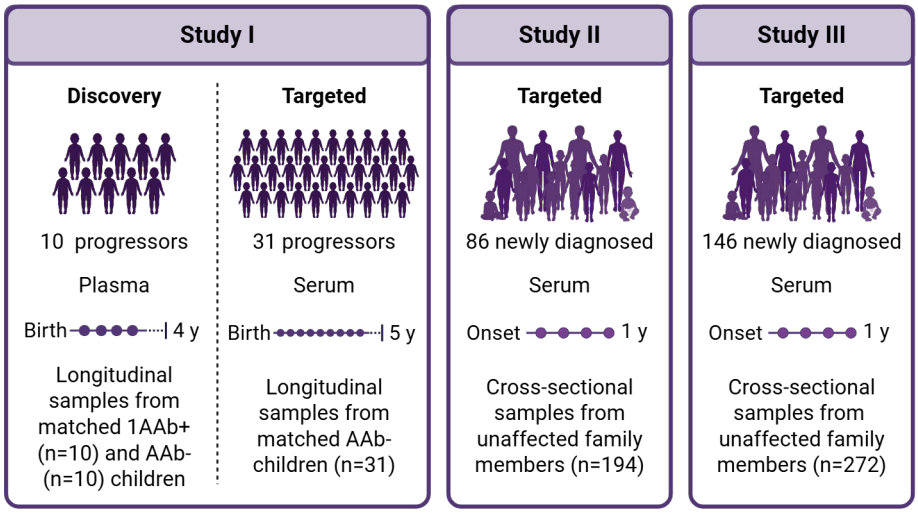

Figure 5. Summary of the Studies I, II and III. Progressors = children who progressed to type 1 diabetes, 1AAb+ = single autoantibody-positive, AAb- = autoantibody-negative.

5.1.1 The effect of age on plasma/serum proteome

Longitudinal data modelling by LonGP identified 115 statistically significant proteins with age-associated trends in the discovery proteomics data from **Study I**, and five additional proteins in the more extensive targeted proteomics data. These age-associated proteins clustered into two separate groups, with either an increasing or decreasing age-associated trend (**Figure 6**). The group showing an increasing age trend was enriched with proteins associated with humoral immune response according to Gene Ontology classification, whereas proteins related to developmental process were enriched in the group with a decreasing age trend.

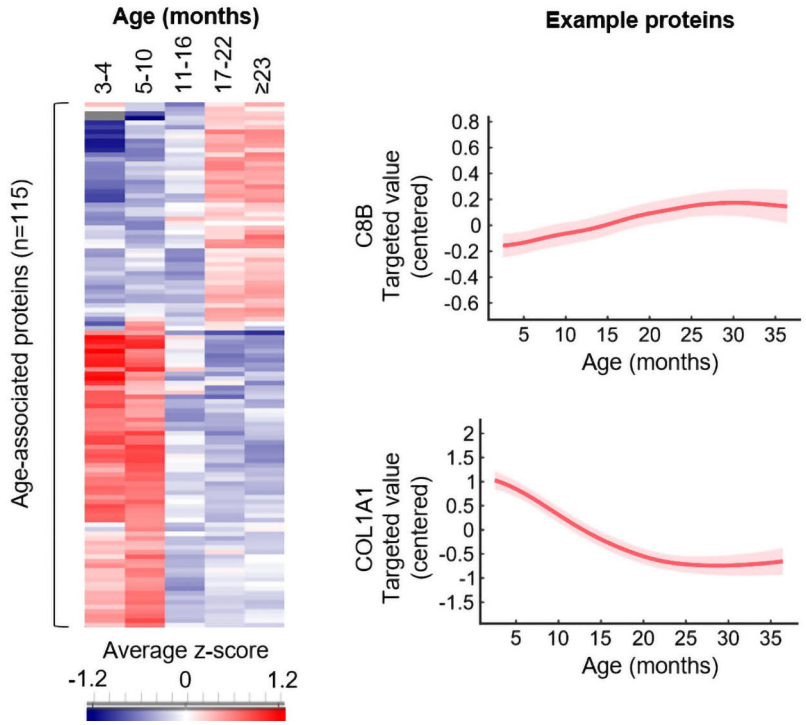

Figure 6. Hierarchical clustering of the 115 statistically significant proteins with age-associated changes in Study I. On the left side, example proteins complement component C8 beta chain (C8B) and collagen alpha-1(I) chain (COL1A1) representing increasing and decreasing age-associated trends, respectively.

5.1.2 Disease-associated changes before the onset

In **Study I**, the longitudinal plasma protein profiles of ten progressors who had developed multiple autoantibodies and progressed to type 1 diabetes before the age of four years were compared with the profiles of matched 1AAb+ and AAb- children

using LonGP (**Figure 5**). To investigate proteins for which the plasma levels changed around the time of seroconversion, the protein profiles of each progressor and 1AAb+ child were aligned based on the seroconversion date, and the profiles of the AAb- chidren were aligned based on the age of the matched progressor. In total, 14 proteins with statistically significant seroconversion-associated changes were identified (**Study I**/Table 1), seven of which were verified using the targeted proteomics approach (**Table 2**).

Another interest of the study was to identify proteins that altered prior to the onset of the disease. For this purpose, the protein profiles of each progressor were aligned based on the diagnosis date, and the profiles of each 1AAb+ and AAb- child were aligned based on the age of the matched progressor. Statistically significant changes were found in three proteins (**Study I**/Table 1). Among these, cysteine-rich secretory protein 3 (CRISP3) and apolipoprotein C1 (APOC1) were selected for the targeted proteomics analysis (**Table 2**).

Lastly, the plasma protein profiles of the three study groups were compared to identify proteins that would distinguish the groups from one another. Based on LongGP, two proteins, namely immunoglobulin heavy constant mu (IGHM) and alpha-2-macroglobulin (A2M) were noted, among which IGHM was verified using targeted proteomics (**Table 2**). Additionally, CD5 antigen like (CD5L), which binds to IGHM (Hiramoto et al., 2018), was selected for the targeted proteomics analysis due to its strikingly similar longitudinal protein profile with IGHM (**Table 2**).

Based on the data and LonGP results, a total of 11 proteins were selected for verification using a targeted serum proteomics approach and a separate cohort of DIPP children. This cohort included 31 children who had developed type 1 diabetes before the age of five years and their matched AAb- children (**Figure 5**). Altogether, 24 peptides representing the 11 targets and an internal standard were included in the analysis. For LonGP analysis, the normalized peptide level data was used. Based on the results, the levels of two peptides representing APOC1 decreased after seroconversion and remained lower in the progressors (**Study I**/Figure 3).

Table 2. Proteins selected for SRM analysis in Study I. CD5L, marked with an asterisk, was not among the significantly altered proteins but selected based on its strikingly similar protein profile with IGHM.

Protein ID	**Gene**	**Discovery:** Explained variation of "age" (%)	**Discovery:** Disease-associated effect (LonGP term)	**Discovery:** Explained variation of the disease-associated effect (%)	**Targeted:** Explained variation of "age" (%)	**Targeted:** Disease-associated effect (LonGP term)	**Targeted:** Explained variation of the disease-associated effect (%)
P01591	IGJ	-	"sero"	16.1	2 peptides: 11.8 and 18.2	-	-
P06727	APOA4	-	"sero"	15.2	3 peptides: 52.4, 47.0 and 49.0	-	-
P02774	GC	-	"sero"	4.4	2 peptides: 0.2 and 4.3	-	-
P08519	LPA	8.7	"sero"	4.3	2 peptides: 2.1 and 2.0	-	-
P18065	IGFBP2	47.7	"sero"	1.7	2 peptides: 33.9 and 56.7	-	-
P02787	TF	-	"sero"	1.4	2 peptides: 4.7 and 3.2	-	-
P00748	F12	-	"sero"	1.2	2 peptides: 0.6 and 0.9	-	-
P54108	CRISP3	-	"t1d"	3.9	1 peptide: 0.6	-	-
P02654	APOC1	-	"t1d"	2.1	2 peptides: 3.9 and 2.5	"sero"	2 peptides: 4.6 and 4.9
P01871	IGHM	24.2	"group" and "age*group"	9.8 and 4.2	2 peptides: 10.1 and 11.0	-	-
O43866	CD5L*	26.5	-	-	2 peptides: 11.1 and 14.4	-	-

5.1.3 Disease-associated changes during the first year from diagnosis

In **Study II**, longitudinal peptide profiles representing the 98 target proteins, including 85 type 1 diabetes-associated proteins and 13 reference proteins, were compared between ND youth (a total of 288 samples) and UFMs (**Figure 5**). Statistically significant differences (FDR ≤ 0.05) were observed in the levels of 18 peptides, representing 13 proteins (**Table 3**). These proteins were further validated

in a similar, however, separate, subsequently recruited larger group of ND youth (560 samples in total) and UFMs using the SRM approach (**Study III**) (**Figure 5**). For some proteins, additional peptides were measured (**Table 3**). In this follow-up study, the changes in ten out of the 13 target proteins found in **Study II** were replicated. From the remaining three proteins, which all belonged to the insulin-like growth factor (IGF) family, the validations were limited by the poor signal intensities and quality of the data obtained.

Table 3. Proteins with significant differences between ND and UFMs. Protein meta p-values were calculated by combining the peptide p-values using the sum of z (Stouffer's) method (R package, metap 1.10). Protein FDRs were calculated using the Benjamini–Hochberg method.

Study	Protein ID	Gene	Peptide	Peptide FDR	Peptide p-value	Effect size[a]	Protein meta p-value	Protein FDR
Study II	P68871	*HBB*	SAVTALWGK	4.7E−02	0.005	0.40	-	-
Study III	P68871	*HBB*	SAVTALWGK	0.002	7.7E-04	0.35	2.60E-06	4.4E-06
Study III	P68871	*HBB*	VNVDEVGG EALGR	0.001	5.2E-04	0.36		
Study II	P00734	*F2*	TATSEYQTF FNPR	8.5E−09	1.1E-10	0.26	-	-
Study III	P00734	*F2*	TATSEYQTF FNPR	1.7E-24	6.4E-26	0.42	3.2E-17	1.9E-16
Study II	P04196	*HRG*	DGYLFQLLR	1.7E−02	0.001	0.16		
Study III	P04196	*HRG*	DGYLFQLLR	0.021	0.012	0.10	0.004	0.005
Study III	P04196	*HRG*	ADLFYDVEAL DLESPK	0.083	0.064	0.08		
Study II	Q04756	*HGFAC*	LEACESLTR	1.3E−02	9.6E-04	0.14	-	-
Study III	Q04756	*HGFAC*	LEACESLTR	6.8E-10	7.6E-11	0.30	2.3E-10	9.0E-10
Study III	Q04756	*HGFAC*	VANYVDWINDR	0.014	0.008	0.14		
Study II	P06681	*C2*	AVISPGFD VFAK	1.2E−05	4.9E-07	0.13	-	-
Study II	P06681	*C2*	HAFILQDTK	3.9E−02	0.004	0.08		
Study III	P06681	*C2*	AVISPGFD VFAK	4.1E-05	9.6E-06	0.14	8.3E-10	2.5E-09
Study III	P06681	*C2*	HAFILQDTK	4.1E-05	1.1E-05	0.14		

Study	Protein ID	Gene	Peptide	Peptide FDR	Peptide p-value	Effect size[a]	Protein meta p-value	Protein FDR
Study II	P22352	*GPX3*	FLVGPDGIP IMR	5.5E−03	3.71E-04	0.11	-	-
Study III	P22352	*GPX3*	FLVGPDGIPI MR	0.006	0.003	0.09	2.5E-09	6.1E-09
Study III	P22352	*GPX3*	QEPGENSEIL PTLK	8.6E-04	2.9E-04	0.09		
Study III	P22352	*GPX3*	NSCPPTSELL GTSDR	1.4E-04	4.0E-05	0.14		
Study II	P43652	*AFM*	DADPDTFFAK	5.5E−04	3.4E-05	−0.13	-	-
Study II	P43652	*AFM*	AESPEVCFNE ESPK	7.1E−05	3.2E-06	−0.16		
Study II	P43652	*AFM*	GQCIINSNK	2.6E−04	1.3E-05	−0.20		
Study III	P43652	*AFM*	DADPDTFFAK	3.1E-06	5.8E-07	-0.14	1.9E-07	3.7E-07
Study III	P43652	*AFM*	AESPEVCF NEESPK	0.049	0.037	-0.08		
Study III	P43652	*AFM*	GQCIINSNK	0.025	0.016	-0.19		
Study II	Q15582	*TGFBI*	LTLLAPLNS VFK	4.0E−04	2.3E-05	−0.22	-	-
Study III	Q15582	*TGFBI*	LTLLAPLNS VFK	0.025	0.016	-0.08	0.016	0.018
Study II	P02654	*APOC1*	EFGNTLEDK	8.5E−09	2.7E-09	−0.42	-	-
Study II	P02654	*APOC1*	EWFSETFQK	1.2E−07	1.4E-10	−0.46		
Study III	P02654	*APOC1*	EFGNTLEDK	0.004	0.002	-0.16	6.4E-06	9.6E-06
Study III	P02654	*APOC1*	EWFSETFQK	0.001	4.7E-04	-0.26		
Study II	P02766	*TTR*	AADDTWEPF ASGK	8.5E−09	1.2E-10	−0.44	-	-
Study II	P02766	*TTR*	TSESGELHGLT TEEEFVEGIYK	4.2E−07	1.2E-08	−0.46		
Study III	P02766	*TTR*	AADDTWEPF ASGK	1.9E-10	1.4E-11	-0.28	3.1E-18	3.7E-17
Study III	P02766	*TTR*	TSESGELHGLT TEEEFVEGIYK	9.6E-08	1.4E-08	-0.30		
Study II	P08833	*IGFBP1*	AQETSGEEISK	8.0E−06	2.7E-07	0.38	-	-
Study II	P17936	*IGFBP3*	ETEYGPCR	3.0E−02	0.003	−0.11	-	-
Study II	P05019	*IGF1*	APQTGIVDEC CFR	3.8E−02	0.003	−0.35	-	-

5.1.4 Associations between serum proteome and fasting C-peptide/glucose ratios

To evaluate whether the alterations in the peptide levels were associated with beta cell function, the longitudinal peptide profiles of ND youth were compared with the concurrent fasting C-peptide/glucose ratios in **Study II**. Statistically significant associations (FDR ≤ 0.05) were found for 12 peptides, representing 11 proteins (**Table 4**). Of these, three proteins were negatively associated and eight proteins were positively associated with the fasting C-peptide/glucose ratios. The shape of the fasting C-peptide/glucose profile over time was hyperbolic, first increasing and then decreasing (**Study II**/Figure 2).

These 11 significant proteins were selected for validation in **Study III,** using a similar approach and a subsequently recruited group of ND youth, as described in the previous paragraph. Here, negative association with the fasting C-peptide/glucose ratios was confirmed for three proteins: apolipoprotein B (APOB), apolipoprotein M (APOM), and glutathione peroxidase 3 (GPX3) (**Table 4**). Notably, the hyperbolic fasting C-peptide/glucose curve was not as prominent as observed in **Study II** (**Study III**/Figure S2, Figure S5).

Furthermore, for some of the ND individuals in **Study II** (n=33), fasting C-peptide glucose measurements were also available at 24 month time point. The longitudinal peptide slopes from these ND youth were compared with the fasting C-peptide/glucose ratios at 24 months. Based on the modelling, peptide profiles of GPX3 during the first year from diagnosis indicated the rate of decline in fasting C-peptide/glucose (FDR ≤ 0.05) (**Study II**/Figure 3). The fasting C-peptide/glucose ratios were better preserved in individuals with decreasing GPX3, whereas in those individuals with increasing GPX3, the loss of C-peptide/glucose was more pronounced. However, this result was not replicated in **Study III**.

Table 4. Proteins with significant associations with the fasting C-peptide/glucose levels. Protein meta p-values were calculated by combining the peptide p-values using the sum of z (Stouffer's) method (R package, metap 1.10). Protein FDRs were calculated using the Benjamini–Hochberg method.

Study	Protein ID	Gene	Peptide	Peptide FDR	Peptide p-value	Effect size[a]	Protein meta p-value	Protein FDR
Study II	P22352	GPX3	FLVGPDGIPIMR	0.003	≤ 0.001	−1.04	-	-
Study III	P22352	GPX3	FLVGPDGIPIMR	0.096	0.025	-0.20	0.002	0.007
Study III	P22352	GPX3	QEPGENSEILPTLK	0.096	0.024	-0.27		
Study III	P22352	GPX3	NSCPPTSELLGTSDR	0.450	0.139	0.16		
Study II	O95445	APOM	AFLLTPR	0.048	0.003	−0.52	-	-
Study III	O95445	APOM	AFLLTPR	0.005	3.2E-04	-0.45	1.3E-05	7.0E-05
Study III	O95445	APOM	SLTSCLDSK	0.038	0.006	-0.25		
Study II	P04114	APOB	EVGTVLSQVYSK	0.045	0.002	−0.22	-	-
Study III	P04114	APOB	EVGTVLSQVYSK	0.096	0.025	-0.13	1.2E-06	0.025
Study III	P04114	APOB	NIQEYLSILTDPDGK	0.005	3.9E-04	-0.23		
Study III	P04114	APOB	ITENDIQIALDDAK	0.019	0.002	-0.26		
Study II	P05155	SERPING1	LLDSLPSDTR	0.024	0.001	−0.64	-	-
Study II	P07360	C8G	SLPVSDSVLSGFEQR	0.035	0.002	−0.53	-	-
Study II	P04196	HRG	DGYLFQLLR	0.048	0.003	−0.42	-	-
Study II	O00187	MASP2	VLATLCGCQESTDTER	0.035	0.001	−0.36	-	-
Study II	P18065	IGFBP2	LIQGAPTIR	0.024	0.001	−0.28	-	-
Study II	P18065	IGFBP2	LEGEACGVYTPR	0.024	0.001	−0.27	-	-
Study II	O14791	APOL1	VTEPISAESGEQVER	0.048	0.003	0.51	-	-
Study II	P05019	IGF1	GFYFNKPTGYGSSSR	0.01	≤ 0.001	0.31	-	-
Study II	P02452	COL1A1	ICVCDNGK	0.035	0.002	0.22	-	-

5.2 Persistent Coxsackievirus B1 infection

In **Study IV**, persistent CVB1 infections using the ATCC and 10796 strains were successfully established at Tampere University by Dr. Anni Honkimaa and maintained for one year alongside the non-infected control cells. Exactly one year after the initial infection, the cells were stained against viral capsid protein VP1 (**Study IV**/Figure 1B). Both CVB1 strains showed characteristics of carrier-state persistence, with only a low proportion of positively stained cells. A few weeks before the viral VP1 protein staining, the cells and cell culture media for proteomics analyses were collected.

5.2.1 Distinct changes in protein expression

In total, 5,130 proteins were identified and quantified from the PANC-1 cell lysate samples after global proteome profiling using LC-MS/MS. Based on data analysis, both CVB1 strains heavily modified the expression of the host proteins but also separated them from each other (**Study IV**/Figure 2A-B). In total, 688 proteins were significantly upregulated in CVB1-infected cells compared to controls, among which 85 proteins were upregulated in both CVB1 infection models. By contrast, significant downregulation was observed in 2,001 proteins, with 520 proteins downregulated in both CVB1 infection models.

Viral peptides were detected in the cells and cell culture supernatants of both CVB1 models indicating the presence of actively replicating virus. The coxsackievirus and adenovirus receptor (CAR), which mediates viral entry into the host cell, was only detected in the non-infected cells (**Study IV**/Figure 3D). This is in line with previous studies on carrier-state persistent CVB infections, where downregulation of CAR has been observed (Alidjinou et al., 2017; Pinkert et al., 2011). Moreover, based on Ingenuity Pathway Analysis (IPA) analysis, the virus entry via endocytic pathways was downregulated in CVB1-infected cell models (**Study IV**/Figure 3E). This restriction likely limits the spread of the virus to the surrounding cells.

In both CVB1 cell models, canonical translation initiation seemed to be impaired and internal ribosome entry site (IRES)-mediated translation was favoured. The supporting findings included the upregulation of Poly(rC)-binding protein 2 (PCBP2) in both CVB1 models, which is a critical protein for IRES-mediated translation and viral replication (Sean et al., 2009). Additionally, based on IPA analysis, proteins involved in cholesterol biosynthesis and mevalonate pathway were highly upregulated in the CVB1-infected cells (**Study IV**/Figure 3B), where efficient CVB replication through accumulation of cholesterol in the viral replication organelles is required (Albulescu et al., 2015). Furthermore, a number of proteins responsible for protein nucleocytoplasmic trafficking were downregulated in both

CVB1 models (**Study IV**/Figure 3C). Finally, western blot analysis confirmed the impairment of host translation initiation factor eIF4G, which is an essential protein for canonical translation initiation (Borman et al., 1997). In both CVB1 models, cleavage of eIF4G was observed (**Study IV**/Figure 3A).

A dominant feature in both CVB1 cell models was a strong downregulation of mitochondrial proteins, more specifically, proteins involved in mitochondrial energy metabolism and oxidative phosphorylation (**Study IV**/Figure 3E). Additionally, significant downregulation was noted for proteins included in related pathways such as fatty acid beta-oxidation, citric acid cycle, and leucine and valine degradation (**Study IV**/Figure 3E). To further investigate the mitochondrial changes, MitoTracker staining was used to visualise the mitochondrial network that readily responds to existing conditions. A fragmented mitochondrial network was observed in CVB1 10796-infected cells, whereas a filamentous network was present in the CVB1 ATCC-infected and non-infected cells (**Study IV**/Figure 4). Further supporting the latter findings, mitofusin 1 (MFN1), mitofusin 2 (MFN2) and dynamin-like 120-kDa mitochondrial protein (OPA1), which all play a key role in mitochondrial fusion processes, were downregulated in CVB1 10796 samples.

5.2.2 Alterations in protein secretion

Similarly to the cell lysates, extensive alterations in protein secretion was observed in the cell culture supernatants from which 3,181 protein groups were identified and quantified (**Study IV**/Figure 2C-D). Here, 919 proteins were significantly inreased in the supernatants of the CVB1-infected cells compared to the control cells, among which 166 were increased in both CVB1 models. Conversely, 630 proteins were significantly decreased in the supernatant samples of CVB1-infected cells, of which 197 were decreased in both CVB1 models.

To investigate alterations in the trafficking of extracellular vesicles, the presence of tetraspanin proteins, which are highly enriched within extracellular vesicles, was explored. Three members of the family were found in the data, namely CD63, CD81 and CD9 (**Study IV**/Figure 5A). Decreased levels of CD63 was observed in the supernatants of the CVB1 ATCC-infected cells, whereas CD9 was decreased in the supernatants from the CVB1 10796 model. No significant alterations were detected in CD81 protein levels. Overall, persistent CVB1 infection seem to reduce the extracellular vesicle trafficking to some extent, although there were strain-specific differences.

Additionally, attention was paid to proteins carrying a signal peptide, as these proteins are likely to be secreted through the classical ER-Golgi pathway, which, in previous studies, has been reported to be blocked by coxsackievirus (Cornell et al., 2006; de Jong et al., 2006). Indeed, among the 371 differentially expressed proteins

with signal peptides based on SignalP (Petersen et al., 2011), around 60% of the proteins detected in the supernatants of each CVB1 infection model were decreased.

Interestingly, several proteins with critical roles in the regulated secretory pathway, which is responsible for the on-demand secretion of proteins, such as hormones like insulin, were decreased or absent in the cell culture supernatants of both CVB1-infected cells (**Study IV**/Figure 5B-C). These included five members of the granin family, namely chromogranin-B (also called secretogranin-1) (CHGB), secretogranin-2 (SCG2), secretogranin-3 (SCG3), neurosecretory protein VGF (VGF), and ProSAAS (PCSK1N), as well as two important enzymes: peptidylglycine α-amidating monooxygenase (PAM) and carboxypeptidase E (CPE). Further exploration using RT-qPCR, confirmed decreased levels of CHGB and SCG3 already at transcriptional level (**Study IV**/Figure 5D-E). However, CPE was significantly increased at the mRNA level in the supernatant of CVB1 10796-infected cells and was not detected in the supernatant of the CVB1 ATCC model (**Study IV**/Figure 5F).

5.2.3 Proteins associated with beta cells

Among the most strongly upregulated proteins in CVB1 ATCC-infected cells was aldehyde dehydrogenase 1a3 (ALDH1A3) and matricellular protein SPARC (**Study IV**/Figure 6A-B). ALDH1A3 is a marker of dedifferentiated beta cells (Cinti et al., 2016), whereas SPARC is secreted by the pancreatic endocrine and exocrine cells and is associated with a negative impact on beta cell growth and survival (Ryall et al., 2014). Neither of these proteins was detected in the cell lysates of CVB1 10796 infection model, but the level of SPARC was strongly decreased in the cell culture supernatants. Furthermore, the sushi domain-containing protein 2 (SUSD2), which is involved in the maturation of the endocrine pancreas during the human foetal pancreatic differentiation (Ramond et al., 2017), was one of the most upregulated proteins in cell lysates of CVB1 10796 model, but significantly downregulated in the CVB1 ATCC-infected cells (**Study IV**/Figure 6C). Finally, downregulation of heparan sulfate proteoglycan core protein (HSPG2) (**Study IV**/Figure 6D), also known as perlecan, and laminin subunit alpha-5 (LAMA5) was observed in the cells infected with the CVB1 10796 strain. These proteins are the principal components of the peri-islet basement membrane, which serves as a protective barrier surrounding the islets (Korpos et al., 2013).

5.2.4 Effect on antiviral immune response

Alterations in interferons (IFNs) were explored to better understand host cell responses towards the two CVB1 strains: ATCC and 10796. The only detected IFN

in the proteomics analysis was IFN-lambda 1 (IFNL1), which was exclusively present in the supernatant of the cells persistently infected with the CVB1 ATCC strain (**Study IV**/Figure 7A). This result was further confimed by RT-qPCR, where the strongest IFNL1 mRNA expression was observed in cells infected with the CVB1 ATCC strain (**Study IV**/Figure 7B). Based on IPA analysis, there were 26 downstream targets of IFNL1 detected in the data. Among these, 14 targets we significantly upregulated in the cell lysates samples of the CVB1 ATCC model, but not in the CVB1 10796 model (**Study IV**/Figure 7C-D). However, in CVB1 10796-infected cells, 20 downstream targets were significantly downregulated or not detected. Investigation of interferon-induced helicase C domain-containing protein 1 (IFIH1) expression by in situ hybridization, was in line with these findings, as high expression was only observed in the CVB1 ATCC infected cells (**Study IV**/Figure 7E). Several other proteins involved in antiviral immune response were also detected in the cell lysates and cell culture supernatants, many of which were increased in CVB1 ATCC-infected cells and decreased in cells infected with the CVB1 10796 strain.

5.3 Overlapping results among all studies

To explore the results of **Studies I**, **II**, **III**, and **IV**, the proteins with statistically significant differences were compared, as shown in **Figure 7**. In the comparison, the results between **Study IV** cell lysates and cell culture supernatants were not compared, as this was done in the publication. Additionally, in **Study IV**, the significant proteins in cell lysates and supernatants also included those that were not detected in any of the three replicates of one condition (ATCC, 10796, and non-infected cells).

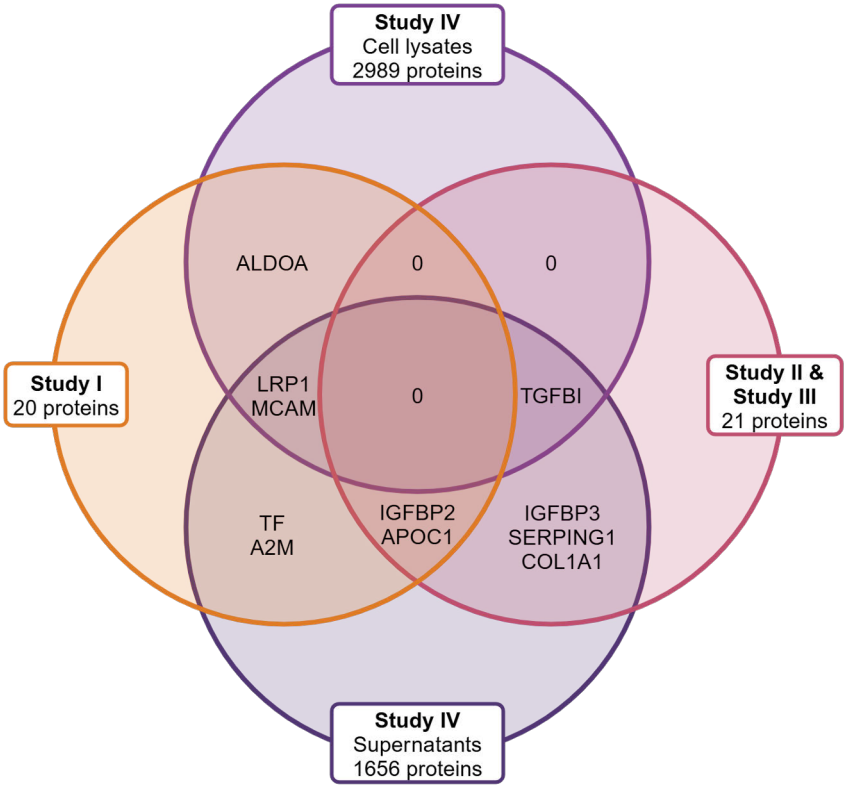

Figure 7. Significantly altered proteins overlapping between Studies I, II, III and IV.

Based on the comparison, there was a total of 11 proteins with overlapping results. These included fructose-bisphosphate aldolase A (ALDOA) and prolow-density lipoprotein receptor-related protein 1 (LRP1), which were downregulated in the CVB1 10796-infected cell lysates and decreased after seroconversion in children who progressed to type 1 diabetes or remained autoantibody-positive in **Study I**. Additionally, the levels of LRP1 were also decreased in the secretome of both CVB1 infection models. Cell surface glycoprotein MUC18 (MCAM) and APOC1 showed opposite results, as APOC1 was increased in the secretome of the CVB1 ATCC and decreased in progressors after seroconversion in **Study I** and in ND youth in **Studies II** and **III**, whereas MCAM was increased in the secretome of the CVB1 10796 and decreased after seroconversion in progressors and 1AAb+ children. Serotransferrin (TF) was not detected in the secretome of the CVB1 10796 model and was increased after seroconversion in progressors and 1AAb+ children, while alpha 2-macroglobulin (A2M) was only detected in the secretome of the CVB1 ATCC and had higher levels in progressors and 1AAb+ children compared to AAb- children in

Study I. Decreased levels of the insulin-like growth factor-binding proteins 2 and 3 (IGFBP2 and IGFBP3) were detected in the secretome of both CVB1 infection models. Additionally, lower levels of IGFBP2 were found in progressors and 1AAb+ children after seroconversion in **Study I,** whereas a negative association between IGFBP2 and C-peptide/glucose ratios was observed in **Study II**. Furthermore, lower levels of IGFBP3 were observed in ND youth compared to UFMs in **Study II**. Decreased levels of transforming growth factor-beta-induced protein ig-h3 (TGFBI) were observed in the cell lysates of the CVB1 ATCC model, as well as in the secretome samples of both CVB1 infection models, although it was not detected in the CVB1 10796 cell lysates. In **Studies II** and **III**, lower TGFBI levels were observed in ND youth compared to UFMs. Plasma protease C1 inhibitor (SERPING1) was detected only in one technical replicate of the secretome sample from the CVB1 10796 model and was absent in the secretome samples of the CVB1 ATCC model. In **Study II**, a negative association between SERPING1 and C-peptide/glucose ratios was observed. Lastly, lower levels of collagen alpha-1(I) chain (COL1A1) were detected in the secretome of the CVB1 ATCC model, whereas in **Study II**, a positive association between COL1A1 and C-peptide/glucose ratios was observed in ND youth.

6 Discussion

Type 1 diabetes is a complex disease that involves genetic risk variants and environmental factors. The disease develops silently and is often diagnosed at the time when the majority of functional beta cells have already been lost. A deeper understanding of molecular changes at each stage of the disease progression is important in order to predict and monitor islet autoimmunity, characterise disease subgroups, find the right therapeutic windows, and prevent acute complications.

In the work presented in this dissertation, longitudinal pre-onset plasma and serum proteomes were studied in a group of children who developed type 1 diabetes at a young age, along with age and sex matched 1AAb+ and AAb- children. From these discovery proteomics analyses, 11 protein targets, were selected for verification using the targeted SRM approach. From **Study I** it was concluded that serum APOC1 levels are decreased after seroconversion in children who progressed to type 1 diabetes.

In **Study II**, a panel of proteins previously associated with type 1 diabetes was evaluated. These targets were studied in longitudinal serum samples from ND youth collected during the first year after diagnosis and compared with cross-sectional samples from UFMs. To investigate protein associations with beta cell function, the longitudinal peptide patterns were compared with the fasting C-peptide/glucose measurements. The levels of 13 proteins differed between ND and UFMs, and in 11 proteins, significant associations with changes in fasting C-peptide and glucose were found. These 21 significant proteins, including two proteins that showed significant results in both comparisons, were selected for validation using a similar targeted proteomics approach and a separate cohort of subsequently recruited ND youth and UFMs (**Study III**). In **Study III**, the levels of 10 proteins differing between ND and UFMs were validated, and the negative association with the fasting C-peptide/glucose ratios was confirmed for APOB, APOM and GPX3.

In **Study IV**, the proteome and secretome of PANC-1 cells persistently infected with the CVB1 strain ATCC and 10796 were investigated alongside the non-infected cells. Prominent changes were observed in the protein expression and secretion of each CVB1 persistent infection model, both of which showed features of carrier-state persistence. In these CVB1-infected cells, signs of actively replicating virus

were detected, along with silencing of the viral receptor CAR. Additionally, in both CVB1 models, proteins involved in key mitochondrial functions were downregulated, and morphological changes were observed. Investigation of the secretomes revealed decreased levels of proteins carrying a signal peptide and proteins involved in the regulated secretory pathway in both CVB1 infection models. Finally, differences between the two strains were observed in the antiviral responses, as CVB1 ATCC triggered the antiviral signalling in the infected cells, whereas CVB1 10796 shut it down.

6.1 Strengths and limitations

As discussed in the review of the literature, the study of type 1 diabetes has been challenged by the heterogeneity of the disease. In previous plasma/serum proteomics studies, this heterogeneity, particularly regarding to age and the rate of progression, has frequently not been adequately acknowledged. **Study I** benefitted from the specifically selected cohort, which included children under five years old who, within months after seroconversion, progressed to type 1 diabetes, along with their well-matched 1AAb+ and AAb- children. Although the discovery study included the interesting group of 1AAb+ children, this group turned out to be rather heterogeneous, as some children developed transient autoantibodies and some were later diagnosed with multiple autoantibodies. Another strength of **Study I** was the verification cohort, from which comprehensive longitudinal data was collected, with a median of nine samples per child. As mentioned earlier, plasma and serum proteomes have distinct characteristics (Geyer et al., 2019). Therefore, in **Study I**, the discovery proteomics analysis was strictly conducted in plasma samples, whereas serum samples were exclusively used in the verification study. Additionally, the proteins selected for SRM were not among the proteins that are known to be significantly altered between plasma and serum. Finally, using LonGP analysis, the dominant effect of age on plasma/serum of young children could be taken into account while extracting the disease-associated trends from the data. Still, most disease-associated changes were unsuccessfully replicated. This might be due to relatively small sample sizes, dissimilarities in sample series and longitudinal autoantibody profiles, or differences in sample handling before and during blood processing, sample storage, and the number of freeze-thaw cycles.

In **Studies II** and **III**, the pre-analytical sample processing was tightly controlled by the master protocol, which likely impacted the success rate of the validation. These studies also benefited from longitudinal sampling and set sampling times immediately after diagnosis, as well as associated measurements reflecting beta cell function. Although SRM analysis provides sensitive peptide measurements, the study was restricted to the exclusive detection of the predefined target proteins.

Additionally, since the participants were recruited on the basis of diagnosis, in **Study III**, the proportion of males was higher than females. Furthermore, as described in the results, the hyperbolic curve for the fasting C-peptide/glucose ratios was less pronounced in **Study III**, which might partly explain the low success rate of validating these results. By observing the individual-specific C-peptide/glucose curves, the proportion of individuals with rapid decline in beta cell function was more pronounced (**Study III**/Figure S5). Lastly, additional limiting factors in these studies were the relatively wide age range, varying autoantibody patterns, and a relatively small number of individuals.

In **Study IV**, persistent coxsackievirus infections using two CVB1 strains were successfully established in the human pancreatic ductal cell line PANC-1. As the presence of CVB is also detected in human ductal cells, through which the virus might spread to beta cells (Geravandi et al., 2021; Nekoua et al., 2022; Ylipaasto et al., 2004), this infection model is highly relevant in the context of type 1 diabetes. At the time **Study IV** was published, persistent infection models had not been published in an insulin-producing human pancreatic cell line, which would have been a more relevant model. Some insulin-producing cell lines excist, such as commercially available clonal beta cell line 1.1B4, which is a hybrid cell line generated by the fusion of primary human beta cells with a pancreatic ductal carcinoma cell. In this cell line, persistent CVB infection has since been successfully established (Honkimaa et al., 2020). However, later research has demonstrated that these cells are not entirely of human origin, but a heterogeneous mixture of rodent and human cells (Chaffey et al., 2021). Other insulin-producing cell lines are EndoC-β1 and the more recently generated EndoC-β2 (Scharfmann et al., 2014). However, only acute CVB infections have been studied with these cell lines (Netanyah et al., 2020). Interestingly, a recent study established stem-cell derived cells that produce insulin and closely mimic pancreatic islets (Balboa et al., 2022).

6.2 Central findings

Proteins with significant changes or associations with fasting C-peptide glucose in **Studies I**, **II**, and **III** mainly belonged to apolipoproteins, the IGF family and coagulants, or were involved in oxidative stress or the retention of beta cell function and integrity. In these studies, the most interesting findings were the 12 significant proteins, which were first detected in **Study II** and later validated in **Study III**. Among these, APOC1 was initially highlighted in **Study I**, where it showed decreased serum levels after seroconversion in children who progressed to type 1 diabetes compared to AAb- children. Furthermore, in **Studies II** and **III**, a similar trend was observed, as APOC1 levels were lower in ND compared to UFMs. It is quite exceptional how well the significant differences in protein levels between ND

and UFMs were replicated in **Studies II** and **III**, particularly considering that proteomics analyses of human biofluids often face challenges in validating identified biomarkers (Nakayasu et al., 2021). Most likely, the critical factors in these studies that led to succession in validation were the existence of a master protocol for sample collection, as well as the QC samples included in the experimental design.

In **Study IV**, extensive changes in protein expression and secretion were observed in PANC-1 cells persistently infected with two different CVB1 strains: ATCC and 10796. Complementary to this study, a transcriptomics study was later published, in which the same cell samples, collected at the same time as the proteomics samples, were used (Buchacher et al., 2021). The results from the transcriptomics study widely supported the proteomics findings, but also highlighted changes in the pancreatic microenvironment and lysosomal function, as well as provided a deeper view of the secretory pathway and the strain-specific differences in the activation/suppression of the antiviral immune response. In the context of defining accessible biomarkers of type 1 diabetes, it would be interesting to determine whether these virus-associated changes can be detected in the serum samples.

In the last chapter of the Results section the plasma and serum proteomics findings of **Studies I**, **II**, and **III** were compared to the significantly altered proteins detected in the cell lysates and secretome samples of the CVB1-infected PANC-1 cells in **Study IV**. In total, 11 proteins were found in this comparison, as illustrated in **Figure 7**. Among those was also APOC1, which was increased in the secretome of ATCC infected cells. Although this increase is contrary to the trend observed in the serum proteome, it is unlikely that this reflects the serum levels, as APOC1 is primarily synthesised in the liver, despite being expressed in many other tissues (Rouland et al., 2022).

An interesting set of proteins highlighted in these studies includes TGFBI, IGFBP2 and IGFBP3. These secreted proteins were all decreased in the secretome of both CVB1-infected cells, and similar decrease or absence of the proteins was also observed in the transcriptomics data (Buchacher et al., 2021). TGFBI levels were decreased in ND youth compared to UFMs in **Study II** and this observation was later confirmed in **Study III**. Based on the Human Protein Atlas (proteinatlas.org, Uhlén et al., 2015) Single Cell resource, TGFBI is enhanced in pancreatic endocrine cells and more specifically in the ductal cells. TGFBI is an important protein for islet integrity, survival, and function, and certain single nucleotide polymorphisms (SNPs) in this gene are associated with type 1 diabetes (Han et al., 2014). Moreover, in a cross-sectional serum proteomics study, lower TGFBI levels were observed in individuals with type 1 diabetes (Zhi et al., 2011). Similarly, in a pre-onset study that compared longitudinal serum profiles of children who progressed to the disease and matched autoantibody-negative children, lower TGFBI levels were reported in

progressors (Moulder et al., 2015). Both of these studies are in accordance with the trends observed in **Studies II** and **III**.

IGFBP2 and IGFBP3 bind to and modulate the actions of insulin-like growth factors (IGFs), which have a similar structure to insulin and are important proteins for cell growth and survival, also within the pancreatic islets. Interestingly, in another enterovirus study, suppressed IGFBP2 expression was observed in a CVB1-infected human lung cancer cell line. Treatment with the drug vemurafenib, which inhibits acute and chronic enterovirus infection, retained the expression of IGFBP2 (Ianevski et al., 2020; Laajala et al., 2023). IGFBP2 is also an interesting protein in the context of diabetes since according to the Human Protein Atlas Tissue resource, its expression is enhanced in the pancreas. In **Study I**, IGFBP2 levels decreased after seroconversion in children who progressed to type 1 diabetes or remain 1AAb+ compared to AAb- children. Similarly, in previous serum and plasma proteomics studies, lower IGFBP2 levels have been observed in longitudinal pre-onset proteomes of children who progressed to type 1 diabetes (Liu et al., 2018; Moulder et al., 2015). However, in individuals living with type 1 diabetes, higher IGFBP2 levels were reported (Zhi et al., 2011). In **Study II**, a negative association with fasting C-peptide/glucose ratios was observed in two peptides representing IGFBP2, which is in line with a previously published study where IGFBP2 inversely correlated with insulin levels (Rajpathak et al., 2012).

IGFBP3 is the most abundant IGFBP in circulation, and based on the Human Protein Atlas, enhanced IGFBP3 expression is observed in the adipose tissue. Lower serum IGFBP3 levels were observed in ND youth compared to UFMs in **Study II**. Similarly, other studies have also observed lower IGFBP3 levels in individuals with type 1 diabetes (Bereket et al., 1995; Hedman et al., 2004). IGFBP3 can bind to LRP1, which is an endocytic and signalling receptor, thus regulating intracellular signalling (Baxter, 2013). In **Study I**, lower plasma LRP1 levels were observed after seroconversion in progressors and 1AAb+ children. However, there were multiple missing values present in the data. In **Study IV**, decreased levels of LRP1 were observed in the secretome of both CVB1-infected cells and in the cell lysates of 10796-infected cells. Although LRP1 is detected in plasma, it is primarily located in the plasma membrane of cells, with additional presence in the nucleoli and the cytosol, according to the Human Protein Atlas Subcellular resource. Furthermore, an activated form of the proteinase inhibitor A2M can also bind to LRP1 and similarly influence the regulation of cellular events (Dato and Chiabrando, 2018). In **Study I**, despite being a depletion target, higher plasma levels of A2M were observed in progressor and 1AAb+ children compared to controls, whereas in **Study IV**, A2M was only detected in the cell lysates of the ATCC infection model.

6.3 A panel of interesting type 1 diabetes-associated proteins

At the same time **Studies I** and **II** were published, an interesting TEDDY proteomics study was released. This study investigated longitudinal plasma samples from children diagnosed with type 1 diabetes or who remained autoantibody-positive by the age of six years, as well as their respective control groups (Nakayasu et al., 2023). Since the results were not compared at the time, this section compares the findings from the TEDDY proteomics study with the significant results from the studies presented in this dissertation, while also revisiting earlier plasma and serum proteomics studies.

The age range of participants in TEDDY proteomics study (Nakayasu et al., 2023) was rather similar to that in **Study I**. Interestingly, similar plasma patterns between these two studies were observed for APOC1, APOA4, ALDOA, and alpha-2-HS-glycoprotein (AHSG), the plasma levels of which decreased after seroconversion in children who progressed to type 1 diabetes or remained autoantibody-positive compared to controls. Likewise, a similar increase in plasma levels of vitamin D-binding protein (GC) was observed after seroconversionin in both studies.

In **Study III**, the levels of 10 out of the 13 proteins, which differed between ND and UFMs, were validated and association between peptide levels and C-peptide/glucose ratios was confirmed for three (APOB, APOM and GPX3) out of the 11 proteins. All of these proteins were also detected in the TEDDY proteomics study (Nakayasu et al., 2023). In addition to APOC1, the TEDDY study confirmed increased plasma levels of prothrombin (F2) after seroconversion in children who progressed to type 1 diabetes compared to controls. In **Studies II** and **III**, increased F2 levels were observed in ND youth compared to UFMs. Opposite results were also observed between the studies. For example complement C2 (C2) was increased in ND youth compared to UFMs, whereas in the TEDDY study (Nakayasu et al., 2023), C2 levels were decreased after seroconversion in children who progressed to type 1 diabetes. Likewise, afamin (AFM) was decreased in ND youth compared to UFMs but increased in children who progressed to type 1 diabetes or remained autoantibody-positive in the TEDDY study.

Based on the results obtained from **Studies I**, **II**, **III**, and **IV**, and their overlap with the TEDDY proteomics study (Nakayasu et al., 2023) and earlier type 1 diabetes plasma and serum proteomics studies, **Table 5** lists proteins that should be carefully monitored when performing type 1 diabetes-associated proteomics experiments, especially on plasma and serum. These include the 12 validated proteins in **Studies II** and **III**, proteins with similar trends in **Study I** and the TEDDY proteomics study (Nakayasu et al., 2023), as well as the most interesting proteins overlapping between **Studies I**, **II**, **III**, and **IV** (**Figure 7**).

Table 5. Proteins with similar trends in the studies presented in this dissertation, including supporting findings in other type 1 diabetes-associated proteomics studies.

Protein ID	Gene	Study I	Study II	Study III	Study IV	Supporting findings
P22352	GPX3		x	x		
O95445	APOM		x	x		
P04114	APOB		x	x		
P68871	HBB		x	x		
P04196	HRG		x	x		
Q04756	HGFAC		x	x		
P06681	C2		x	x		
P43652	AFM		x	x		
Q15582	TGFBI		x	x	x	Zhi et al. 2011, Moulder et al. 2015
P02766	TTR		x	x		Zhang et al. 2013
P00734	F2		x	x		De Oliveira et al. 2018, Nakayasu et a. 2023
P02654	APOC1	x	x	x		Nakayasu et a. 2023
P06727	APOA4	x				Moulder et al. 2015, von Toerne et al. 2017, Nakayasu et a. 2023
P04075	ALDOA	x				Nakayasu et a. 2023
P02765	AHSG	x				Nakayasu et a. 2023
P02774	GC	x				Nakayasu et a. 2023
Q07954	LRP1	x			x	
P18065	IGFBP2	x	x		x	Moulder et al. 2015, Liu et al. 2018
P17936	IGFBP3		x		x	
P01023	A2M	x			x	De Oliveira et al. 2018

6.4 Future perspectives

In recent years, as our understanding of the type 1 diabetes pathogenesis has increased, medicine developers have initiated therapeutic approaches to delay, cure and prevent type 1 diabetes. An important achievement was made in 2022, when the U.S. Food and Drug Administration (FDA) approved teplizumap, a humanized

monoclonal antibody that binds CD3 on T cells, as the first drug to delay the onset of type 1 diabetes (Herold et al., 2023). These different therapeutic approaches are often directed to individuals at certain stages of type 1 diabetes. Alternatively, they aim to prevent the triggering of islet autoimmunity, such as CVB vaccines (Dunne et al., 2019; Hyöty et al., 2024). In all these situations, it is important to know whom to treat and when.

To better support the decision-making, proteomics approaches can be utilised. However, it is critical to carefully select the study cohorts and the protocols used. The heterogeneity of type 1 diabetes development is a major challenge and likely a significant reason why it has been difficult to validate type 1 diabetes biomarkers. In cohort selection, attention should be paid to defining groups that are expected to be more homogeneous, such as individuals with IAA or GADA as the first appearing autoantibody. It is also important to understand the differences in the disease process between children and adults, as this is currently less well studied (Leslie et al., 2021). As presented in **Study I** and other studies (Bjelosevic et al., 2017; Lietzén et al., 2018; Liu et al., 2017; Mikus et al., 2021), the effect of age is a dominant factor affecting the serum proteome, especially in young children. Therefore, longitudinal sampling is important, with a sufficient number of sampling points.

Currently, most samples used in proteomics analysis are acquired from different screening programmes. Establishing a uniform sample collection protocol, which would be used at all sites of the study, is the first step. The staff working at the sites should be trained for the job. Based on the protocol, thorough documentation should be provided, which includes a detailed description of how the samples are handled, processed, and stored. Ideally, this documentation would also include information on fasting status, the level of physical activity before the sample collection, and the time and date of the collection.

In proteomics workflows, QCs should be included to monitor critical steps in sample preparation and technical variations in sample analysis. Recent advancements in instrumentation, analytical methods, and sample preparation enable the analysis of larger cohorts, which, due to increased statistical power, likely provide more accurate results (Bader et al., 2023). To prepare samples for such large experiments, automated or semi-automated sample preparation needs to be implemented in order to minimise human error and ensure good reproducibility. Most untargeted proteomics analyses today are conducted using DIA methods, which ensure better data completeness, deeper proteome coverage, and enhanced sensitivity (Fröhlich et al., 2024). The state-of-the-art mass spectrometers are fast, precise, and sensitive. For example, with the Orbitrap Astral mass spectrometer (Thermo Scientific), it is possible to detect over 1,000 proteins in neat plasma (Niu et al., 2025). Moreover, by using a bead-based vesicle enrichment protocol and the same instrument, it is possible to detect up to 5,000 plasma proteins with a 60-min

gradient (Heil et al., 2023). From a technological perspective, one should also keep in mind the alternative proteomics technologies such as Olink® Explore and SomaScan, as they may provide specific advantages, particularly for high-throughput sample analysis.

To better study the enterovirus signature in plasma/serum, samples from individuals with acute enterovirus infection could first be studied. This would likely indicate which proteins are altered by the virus infection. Presumably, persistent enterovirus infection in the pancreas would result in a slightly different or less profound signature. Additionally, as shown in **Study IV**, different strains are likely to produce slightly different signatures. Therefore, it would be important to study samples from a large number of individuals. Samples received from prospective cohort studies could be screened for acute and past enterovirus infections using, for example, PCR or neutralising antibodies, which can identify different enterovirus serotypes. Using a proteomics approach, signs of the enterovirus signal could then be traced from these samples.

7 Conclusions

The aim of the work presented in this doctoral dissertation was to provide a deeper understanding of the molecular changes underlying the pathogenesis of type 1 diabetes, as well as to investigate how persistent enterovirus infection, which has been associated with type 1 diabetes development, alters the protein expression and secretion in pancreatic ductal cells. The principal technology used in all studies was mass spectrometry-based proteomics. Based on the results presented in this dissertation, the following conclusions can be drawn:

1. In children who progressed to type 1 diabetes, serum levels of APOC1 decreased after seroconversion and remain lower until diagnosis compared to autoantibody-negative children. Similarly, in newly diagnosed youth during the first year following diagnosis, the serum levels of APOC1 are lower than those in unaffected family members.
2. A panel of 21 proteins, showing significant associations with type 1 diabetes, was defined. These included 13 proteins, whose serum levels differed during the first year from diagnosis between newly diagnosed youth and unaffected family members, and 11 proteins with either positive or negative associations with the fasting C-peptide/glucose ratios. Further validation of these 21 proteins in a new cohort confirmed the serum level differences in 10 proteins between newly diagnosed youth and unaffected family members, as well as replicated a negative association with the fasting C-peptide/glucose ratios for APOB, APOM, and GPX3.
3. Two coxsackievirus B1 strains, ATCC and 10796, were able to establish a carrier-state persistent infection in the human ductal cell line PANC-1, which caused extensive changes in the host's protein expression and secretion. In the context of type 1 diabetes, one of the most interesting findings was the downregulation of proteins with critical roles in the regulated secretory pathway, which is responsible for the on-demand secretion of hormones and other molecules, including insulin. Another interesting feature of the data was the differences between the two CVB1

strains. The most striking difference was observed in the antiviral immune response, as in ATCC-infected cells this response was clearly activated, whereas in 10796- infected cells, it appeared to be suppressed.

4. A total of 11 significantly altered proteins in the CVB1-infected cells and cell culture supernatants also showed significant alterations in plasma and/or serum samples of individuals progressing to or diagnosed with type 1 diabetes. Although these overlapping proteins were found, no viral signature was confirmed. The potential contribution of these proteins to a CVB1 signature in plasma or serum remains to be explored in a more appropriate cohort.

Acknowledgements

I would like to express my deepest gratitude to my PhD supervisors, Riitta Lahesmaa and Niina Lietzén. You have provided excellent guidance and mentorship throughout these years. Thank you Riitta for giving me the opportunity to work in your lab and for allowing me explore and try new methods and techniques. You have always encouraged me and pushed me forward when I was cautious. Thank you Niina for being my mentor in proteomics and sharing all the knowledge you have in this field. I learned so much while working with you. I am particularly grateful for the advice and feedback you gave, which helped me grow scientifically.

I would like to thank my pre-examiners, Tuula Nyman and Vesa Hytönen, for their efficient and thorough review of the book, as well as for providing valuable comments and suggestions to improve it. I would also like to thank Jochen Schwenk for accepting the invitation to be the opponent, and I look forward to your visit and our discussions.

This book and these studies would not exist without successful collaborations. I would like to thank Heikki Hyöty from Tampere University for being a member of my follow-up committee and, together with his team, especially Anni Honkimaa and Amirbabak Sioofy-Khojine, for providing their knowledge on enteroviruses and CVB1-infected cells. Thank you to Lu Cheng and Harri Lähdesmäki for your bioinformatics expertise in Study I, as well as to Laura Elo, Tommi Välikangas, and Tomi Suomi for your bioinformatics expertise in Studies II and III. I would also like to thank the people behind the DIPP study and the INNODIA consortium for the opportunity to analyse these unique samples and for providing valuable feedback on the manuscripts. The Turku Proteomics Facility has been irreplaceable. Thank you Pekka, Arttu, Mirva, Johannes and Otto for your advice and help with running the instruments and troubleshooting. I would also like to thank Michael Courtney from the Turku Screening Unit for helping to set up the automated sample preparation.

For the people from the lab, I would like to especially thank Robert Moulder for the great conversations, guidance, and mentorship. I am still amazed by your outstanding knowledge of MS-based proteomics. Being a member of the proteomics team has been a joy. In addition to Niina and Robert, I would like to thank Santosh and Kedar for the great teamwork and brainstorming. I would like to thank Marjo

and Sarita for helping us with the sample preparations, and Anne for your efficient assistance on multiple different things. For years, I have had the privilege of sharing an office with two brilliant scientists, Tanja and Ankitha. You are wise and kind, and I am truly grateful for our friendship. I also cherish my friendship with Inna and Essi. It has always been a joy to have lunch with you and engage in colorful discussions about life in and outside the lab. Thank you Inna for sharing the PhD journey with me. The Lahesmaa lab has changed over the years, but it has always been a place where I enjoy working, and the major reason for that is the people I have worked with. Therefore, I would like to warmly thank the current and former members of the lab for all the wonderful memories.

I was fortunate to grow up in a lively family, surrounded by siblings, grandparents, cousins, aunts, and uncles. I love the gatherings with my extended childhood family: Mum, Dad, Eetu, Roosa, Eve, Jesse, and little Frida. I am extremely grateful to my Mum and Dad, who have always believed in me and supported me no matter what. While I was carrying out my doctoral studies, my parents, as well as Mikael's parents, helped greatly, especially by taking care of the kids. I owe my warmest thanks to my friends Laura, Anniina, Eeva, Essi, Sofia, Sonja, and Mona, as well as to Mikael's side of the family Satu, Pekka, Sanna, Kim, Lauri, Heidi, and the kids. Thank you for those moments in life when everything feels just a little bit lighter. Most importantly, I would like to thank the greatest teachers of my life, Hilla and Frans, with whom life is just so much more fun. A gentle rub also goes to my furry friend Dino, who has an amazing ability to calm me down. Last but not least, I would like to thank Mikael for your love, understanding and endless support. Love you all!

This PhD dissertation was conducted at the Turku Centre for Biotechnology, University of Turku, and Åbo Akademi University. I would like to acknowledge the Turku Doctoral Programme of Molecular Medicine (TuDMM) for providing fundamental training, guidance, and support throughout my studies. I would also like to thank the InFLAMES Doctoral Module for great gatherings and interesting events. A big and warm thank you goes to all the sources from which I have received funding over these years to conduct this PhD work. Thank you to TuDMM, the Diabetes Research Foundation, the Finnish Cultural Foundation, the Kyllikki and Uolevi Lehikoinen Foundation, the Maud Kuistila Memorial Foundation, the Orion Research Foundation, the Päivikki and Sakari Sohlberg Foundation, the Turku University Foundation, and the Yrjö Jahnsson Foundation.

April 2025
Karoliina Hirvonen

References

Aebersold, R., Agar, J.N., Amster, I.J., Baker, M.S., Bertozzi, C.R., Boja, E.S., Costello, C.E., Cravatt, B.F., Fenselau, C., Garcia, B.A., Ge, Y., Gunawardena, J., Hendrickson, R.C., Hergenrother, P.J., Huber, C.G., Ivanov, A.R., Jensen, O.N., Jewett, M.C., Kelleher, N.L., Kiessling, L.L., Krogan, N.J., Larsen, M.R., Loo, J.A., Ogorzalek Loo, R.R., Lundberg, E., Maccoss, M.J., Mallick, P., Mootha, V.K., Mrksich, M., Muir, T.W., Patrie, S.M., Pesavento, J.J., Pitteri, S.J., Rodriguez, H., Saghatelian, A., Sandoval, W., Schlüter, H., Sechi, S., Slavoff, S.A., Smith, L.M., Snyder, M.P., Thomas, P.M., Uhlén, M., Van Eyk, J.E., Vidal, M., Walt, D.R., White, F.M., Williams, E.R., Wohlschlager, T., Wysocki, V.H., Yates, N.A., Young, N.L., Zhang, B., 2018. How many human proteoforms are there? Nat. Chem. Biol. 14, 206–214. https://doi.org/10.1038/nchembio.2576

Aebersold, R., Mann, M., 2016. Mass-spectrometric exploration of proteome structure and function. Nature 537, 347–355. https://doi.org/10.1038/nature19949

Albulescu, L., Wubbolts, R., van Kuppeveld, F.J.M., Strating, J.R.P.M., 2015. Cholesterol shuttling is important for RNA replication of coxsackievirus B3 and encephalomyocarditis virus. Cell. Microbiol. 17, 1144–1156. https://doi.org/10.1111/cmi.12425

Alidjinou, E.K., Engelmann, I., Bossu, J., Villenet, C., Figeac, M., Romond, M.B., Sané, F., Hober, D., 2017. Persistence of Coxsackievirus B4 in pancreatic ductal-like cells results in cellular and viral changes. Virulence 8, 1229–1244. https://doi.org/10.1080/21505594.2017.1284735

Anand, V., Li, Y., Liu, B., Ghalwash, M., Koski, E., Ng, K., Dunne, J.L., Jönsson, J., Winkler, C., Knip, M., Toppari, J., Ilonen, J., Killian, M.B., Frohnert, B.I., Lundgren, M., Ziegler, A.-G., Hagopian, W., Veijola, R., Rewers, M., 2021. Islet Autoimmunity and HLA Markers of Presymptomatic and Clinical Type 1 Diabetes: Joint Analyses of Prospective Cohort Studies in Finland, Germany, Sweden, and the U.S. Diabetes Care 44, 2269–2276. https://doi.org/10.2337/DC20-1836

Anderson, N.L., 2010. The Clinical Plasma Proteome: A Survey of Clinical Assays for Proteins in Plasma and Serum. Clin. Chem. 56, 177–185. https://doi.org/10.1373/clinchem.2009.126706

Anderson, N.L., Anderson, N.G., 2002. The human plasma proteome: history, character, and diagnostic prospects. Mol. Cell. Proteomics 1, 845–867. https://doi.org/10.1074/mcp.R200007-MCP200

Arvan, P., Pietropaolo, M., Ostrov, D., Rhodes, C.J., 2012. Islet autoantigens: Structure, function, localization, and regulation. Cold Spring Harb. Perspect. Med. 2. https://doi.org/10.1101/CSHPERSPECT.A007658

Bader, J.M., Albrecht, V., Mann, M., 2023. MS-Based Proteomics of Body Fluids: The End of the Beginning. Mol. Cell. Proteomics 22. https://doi.org/10.1016/J.MCPRO.2023.100577

Balboa, D., Barsby, T., Lithovius, V., Saarimäki-Vire, J., Omar-Hmeadi, M., Dyachok, O., Montaser, H., Lund, P.E., Yang, M., Ibrahim, H., Näätänen, A., Chandra, V., Vihinen, H., Jokitalo, E., Kvist, J., Ustinov, J., Nieminen, A.I., Kuuluvainen, E., Hietakangas, V., Katajisto, P., Lau, J., Carlsson, P.O., Barg, S., Tengholm, A., Otonkoski, T., 2022. Functional, metabolic and transcriptional maturation of human pancreatic islets derived from stem cells. Nat. Biotechnol. 2022 407 40, 1042–1055. https://doi.org/10.1038/s41587-022-01219-z

Banting, F.G., Best, C.H., Collip, J.B., Campbell, W.R., Fletcher, A.A., 1922. Pancreatic Extracts in the Treatment of Diabetes Mellitus. Can. Med. Assoc. J. 12, 141.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Baxter, R.C., 2013. Insulin-like growth factor binding protein-3 (IGFBP-3): Novel ligands mediate unexpected functions. J. Cell Commun. Signal. 7, 179–189. https://doi.org/10.1007/S12079-013-0203-9

Beimers, W.F., Overmyer, K.A., Sinitcyn, P., Lancaster, N.M., Quarmby, S.T., Coon, J.J., 2025. A Technical Evaluation of Plasma Proteomics Technologies. bioRxiv Prepr. Serv. Biol. https://doi.org/10.1101/2025.01.08.632035

Bellei, E., Bergamini, S., Monari, E., Fantoni, L.I., Cuoghi, A., Ozben, T., Tomasi, A., 2011. High-abundance proteins depletion for serum proteomic analysis: Concomitant removal of non-targeted proteins. Amino Acids 40, 145–156. https://doi.org/10.1007/S00726-010-0628-X/TABLES/4

Bereket, A., Lang, C.H., Blethen, S.L., Gelato, M.C., Fan, J., Frost, R.A., Wilson, T.A., 1995. Effect of insulin on the insulin-like growth factor system in children with new-onset insulin-dependent diabetes mellitus. J. Clin. Endocrinol. Metab. 80, 1312–1317. https://doi.org/10.1210/jcem.80.4.7536205

Bjelosevic, S., Pascovici, D., Ping, H., Karlaftis, V., Zaw, T., Song, X., Molloy, M.P., Monagle, P., Ignjatovic, V., 2017. Quantitative Age-specific Variability of Plasma Proteins in Healthy Neonates, Children and Adults. Mol. Cell. proteomics MCP 16, 924–935.

Blume, J.E., Manning, W.C., Troiano, G., Hornburg, D., Figa, M., Hesterberg, L., Platt, T.L., Zhao, X., Cuaresma, R.A., Everley, P.A., Ko, M., Liou, H., Mahoney, M., Ferdosi, S., Elgierari, E.M., Stolarczyk, C., Tangeysh, B., Xia, H., Benz, R., Siddiqui, A., Carr, S.A., Ma, P., Langer, R., Farias, V., Farokhzad, O.C., 2020. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. Nat. Commun. 2020 111 11, 1–14. https://doi.org/10.1038/s41467-020-17033-7

Bonifacio, E., Beyerlein, A., Hippich, M., Winkler, C., Vehik, K., Weedon, M.N., Laimighofer, M., Hattersley, A.T., Krumsiek, J., Frohnert, B.I., Steck, A.K., Hagopian, W.A., Krischer, J.P., Lernmark, Å., Rewers, M.J., She, J.X., Toppari, J., Akolkar, B., Oram, R.A., Rich, S.S., Ziegler, A.G., 2018. Genetic scores to stratify risk of developing multiple islet autoantibodies and type 1 diabetes: A prospective study in children. PLoS Med. 15. https://doi.org/10.1371/journal.pmed.1002548

Bonifacio, E., Ziegler, A.G., 2010. Advances in the prediction and natural history of type 1 diabetes. Endocrinol. Metab. Clin. North Am. 39, 513–525. https://doi.org/10.1016/j.ecl.2010.05.007

Borman, A.M., Kirchweger, R., Ziegler, E., Rhoads, R.E., Skern, T., Kean, K.M., 1997. elF4G and its proteolytic cleavage products: Effect on initiation of protein synthesis from capped, uncapped, and IRES-containing mRNAs. RNA 3, 186–196.

Borras, E., Sabido, E., 2017. What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry. Proteomics (Weinheim) 17, n/a. https://doi.org/10.1002/pmic.201700180

Boulpaep, E.L., 2017a. The Endocrine Pancreas, in: Boron, W.F., Boulpaep, E.L. (Eds.), Medical Physiology. Elsevier, 2017, Philadelphia, pp. 1035–1053. https://doi.org/http://dx.doi.org/10.1016/B978-1-4557-4377-3.00018-5

Boulpaep, E.L., 2017b. Blood, in: Boron, W.F., Boulpaep, E.L. (Eds.), Medical Physiology. Elsevier, 2017, Philadelphia, pp. 429-446.e1. https://doi.org/http://dx.doi.org/10.1016/B978-1-4557-4377-3.00018-5

Buchacher, T., Honkimaa, A., Välikangas, T., Lietzén, N., Hirvonen, M.K., Laiho, J.E., Sioofy-Khojine, A.B., Eskelinen, E.L., Hyöty, H., Elo, L.L., Lahesmaa, R., 2021. Persistent coxsackievirus B1 infection triggers extensive changes in the transcriptome of human pancreatic ductal cells. iScience 25, 103653. https://doi.org/10.1016/J.ISCI.2021.103653

Burkhart, J.M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., Zahedi, R.P., 2012. Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality

on MS-based proteomics. J. Proteomics 75, 1454–1462. https://doi.org/10.1016/j.jprot.2011.11.016

Cabrera, O., Berman, D.M., Kenyon, N.S., Ricordi, C., Berggren, P.O., Caicedo, A., 2006. The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. Proc. Natl. Acad. Sci. U. S. A. 103, 2334–2339. https://doi.org/10.1073/pnas.0510790103

Chaffey, J.R., Young, J., Leslie, K.A., Partridge, K., Akhbari, P., Dhayal, S., Hill, J.L., Wedgwood, K.C.A., Burnett, E., Russell, M.A., Richardson, S.J., Morgan, N.G., 2021. Investigation of the utility of the 1.1B4 cell as a model human beta cell line for study of persistent enteroviral infection. Sci. Rep. 11. https://doi.org/10.1038/s41598-021-94878-y

Chehadeh, W., Kerr-Conte, J., Pattou, F., Alm, G., Lefebvre, J., Wattré, P., Hober, D., 2000. Persistent Infection of Human Pancreatic Islets by Coxsackievirus B Is Associated with Alpha Interferon Synthesis in β Cells. J. Virol. 74, 10153–10164. https://doi.org/10.1128/jvi.74.21.10153-10164.2000

Chen, C., Hou, J., Tanner, J.J., Cheng, J., 2020. Bioinformatics methods for mass spectrometry-based proteomics data analysis. Int. J. Mol. Sci. 21, 2873. https://doi.org/10.3390/ijms21082873

Cheng, L., Ramchandran, S., Vatanen, T., Lietzén, N., Lahesmaa, R., Vehtari, A., Lähdesmäki, H., 2019. An additive Gaussian process regression model for interpretable non-parametric analysis of longitudinal data. Nat. Commun. 10. https://doi.org/10.1038/s41467-019-09785-8

Cinti, F., Bouchi, R., Kim-Muller, J.Y., Ohmura, Y., Sandoval, P.R., Masini, M., Marselli, L., Suleiman, M., Ratner, L.E., Marchetti, P., Accili, D., 2016. Evidence of β-cell dedifferentiation in human type 2 diabetes. J. Clin. Endocrinol. Metab. 101, 1044–1054. https://doi.org/10.1210/jc.2015-2860

Cornell, C.T., Kiosses, W.B., Harkins, S., Whitton, J.L., 2006. Inhibition of Protein Trafficking by Coxsackievirus B3: Multiple Viral Proteins Target a Single Organelle. J. Virol. 80, 6637–6647. https://doi.org/10.1128/JVI.02572-05

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V, Mann, M., 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 10, 1794–1805.

Cristobal, A., Marino, F., Post, H., Van Den Toorn, H.W.P., Mohammed, S., Heck, A.J.R., 2017. Toward an Optimized Workflow for Middle-Down Proteomics. Anal. Chem. 89, 3318–3325. https://doi.org/10.1021/ACS.ANALCHEM.6B03756/ASSET/IMAGES/LARGE/AC-2016-03756G_0003.JPEG

Cronje, H.T., Mi, M.Y., Austin, T.R., Biggs, M.L., Siscovick, D.S., Lemaitre, R.N., Psaty, B.M., Tracy, R.P., Djousse, L., Kizer, J.R., Ix, J.H., Rao, P., Robbins, J.M., Barber, J.L., Sarzynski, M.A., Clish, C.B., Bouchard, C., Mukamal, K.J., Gerszten, R.E., Jensen, M.K., 2023. Plasma Proteomic Risk Markers of Incident Type 2 Diabetes Reflect Physiologically Distinct Components of Glucose-Insulin Homeostasis. Diabetes 72, 666–673. https://doi.org/10.2337/DB22-0628

Dato, V.A., Chiabrando, G.A., 2018. The Role of Low-Density Lipoprotein Receptor-Related Protein 1 in Lipid Metabolism, Glucose Homeostasis and Inflammation. Int. J. Mol. Sci. 19. https://doi.org/10.3390/IJMS19061780

De Goffau, M.C., Luopajärvi, K., Knip, M., Ilonen, J., Ruohtula, T., Härkönen, T., Orivuori, L., Hakala, S., Welling, G.W., Harmsen, H.J., Vaarala, O., 2013. Fecal microbiota composition differs between children with β-cell autoimmunity and those without. Diabetes 62, 1238–1244. https://doi.org/10.2337/db12-0526

de Jong, A.S., Visch, H.-J., de Mattia, F., van Dommelen, M.M., Swarts, H.G., Luyten, T., Callewaert, G., Melchers, W.J., Willems, P.H., van Kuppeveld, F.J., 2006. The coxsackievirus 2B protein increases efflux of ions from the endoplasmic reticulum and Golgi, thereby inhibiting protein trafficking through the Golgi. J. Biol. Chem. 281, 14144–14150.

De Oliveira, V. do N., Lima-Neto, A.B.M., van Tilburg, M.F., Monteiro-Moreira, A.C. de O., Lobo, M.D.P., Rondina, D., Fernandes, V.O., Montenegro, A.P.D.R., Montenegro Júnior, R.M., Guedes,

M.I.F., 2018. Proteomic analysis to identify candidate biomarkers associated with type 1 diabetes. Diabetes, Metab. Syndr. Obes. Targets Ther. 11, 289. https://doi.org/10.2147/DMSO.S162008

DIAMOND Project group, 2006. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. Diabet. Med. 23, 857–866. https://doi.org/10.1111/j.1464-5491.2006.01925.x

Dunger, D.B., Bruggraber, S.F.A., Mander, A.P., Marcovecchio, M. Loredana, Tree, Timothy, Chmura, P.J., Knip, Mikael, Schulte, A.M., Mathieu, Chantal, Mathieu, C., Gillard, P., Casteels, K., Overbergh, L., Dunger, D., Wallace, C., Evans, M., Thankamony, A., Hendriks, E., Bruggraber, S., Peakman, M., Tree, T., Morgan, N., Richardson, S., Todd, J., Wicker, L., Mander, A., Dayan, C., Alhadj Ali, M., Pieber, T., Eizirik, D., Cnop, M., Brunak, S., Pociot, F., Johannesen, J., Rossing, P., Legido Quigley, C., Mallone, R., Scharfmann, R., Boitard, C., Knip, M., Otonkoski, T., Veijola, R., Lahesmaa, R., Oresic, M., Toppari, J., Danne, T., Ziegler, A.G., Achenbach, P., Rodriguez-Calvo, T., Solimena, M., Bonifacio, E., Speier, S., Holl, R., Dotta, F., Chiarelli, F., Marchetti, P., Bosi, E., Cianfarani, S., Ciampalini, P., de Beaufort, C., Dahl-Jørgensen, K., Skrivarhaug, T., Joner, G., Krogvold, L., Jarosz-Chobot, P., Battelino, T., Thorens, B., Gotthardt, M., Roep, B., Nikolic, T., Zaldumbide, A., Lernmark, A., Lundgren, M., Costecalde, G., Strube, T., Schulte, A., Nitsche, A., von Herrath, M., Wesley, J., Napolitano-Rosen, A., Thomas, M., Schloot, N., Goldfine, A., Waldron-Lynch, F., Kompa, J., Vedala, A., Hartmann, N., Nicolas, G., van Rampelbergh, J., Bovy, N., Dutta, S., Soderberg, J., Ahmed, S., Martin, F., Agiostratidou, G., Koralova, A., Willemsen, R., Smith, A., Anand, B., Puthi, V., Zac-Varghese, S., Datta, V., Dias, R., Sundaram, P., Vaidya, B., Patterson, C., Owen, K., Piel, B., Heller, S., Randell, T., Gazis, T., Bismuth Reismen, E., Carel, J.C., Riveline, J.P., Gautier, J.F., Andreelli, F., Travert, F., Cosson, E., Penfornis, A., Petit, C., Feve, B., Lucidarme, N., Cosson, E., Beressi, J.P., Ajzenman, C., Radu, A., Greteau-Hamoumou, S., Bibal, C., Meissner, T., Heidtmann, B., Toni, S., Rami-Merhar, B., Eeckhout, B., Peene, B., Vantongerloo, N., Maes, T., Gommers, L., Marcovecchio, M. L., Vela, J., Latres, E., 2022. INNODIA Master Protocol for the evaluation of investigational medicinal products in children, adolescents and adults with newly diagnosed type 1 diabetes. Trials 23. https://doi.org/10.1186/s13063-022-06259-z

Dunne, J.L., Richardson, S.J., Atkinson, M.A., Craig, M.E., Dahl-Jørgensen, K., Flodström-Tullberg, M., Hyöty, H., Insel, R.A., Lernmark, Å., Lloyd, R.E., Morgan, N.G., Pugliese, A., 2019. Rationale for enteroviral vaccination and antiviral therapies in human type 1 diabetes. Diabetologia 62, 744–753. https://doi.org/10.1007/S00125-019-4811-7/FIGURES/2

Dupree, E.J., Jayathirtha, M., Yorkey, H., Mihasan, M., Petre, B.A., Darie, C.C., 2020. A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of this Field. Proteomes 8, 1–26. https://doi.org/10.3390/PROTEOMES8030014

Eisenbarth, G.S., 1986. Type I diabetes mellitus. A chronic autoimmune disease. N. Engl. J. Med. 314, 1360–8. https://doi.org/10.1056/NEJM198605223142106

Endesfelder, D., zu Castell, W., Bonifacio, E., Rewers, M., Hagopian, W.A., She, J.X., Lernmark, A., Toppari, J., Vehik, K., Williams, A.J.K., Yu, L., Akolkar, B., Krischer, J.P., Ziegler, A.G., Achenbach, P., 2019. Time-resolved autoantibody profiling facilitates stratification of preclinical type 1 diabetes in children. Diabetes 68, 119–130. https://doi.org/10.2337/db18-0594

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M., 1989. Electrospray ionization for mass spectrometry of large biomolecules. Science (80-.). 246, 64–71. https://doi.org/10.1126/science.2675315

Frisk, G., 2001. Mechanisms of chronic enteroviral persistence in tissue. Curr. Opin. Infect. Dis. 14, 251–256. https://doi.org/10.1097/00001432-200106000-00002

Fröhlich, K., Fahrner, M., Brombacher, E., Seredynska, A., Maldacker, M., Kreutz, C., Schmidt, A., Schilling, O., 2024. Data-Independent Acquisition: A Milestone and Prospect in Clinical Mass Spectrometry–Based Proteomics. Mol. Cell. proteomics 23, 100800. https://doi.org/10.1016/j.mcpro.2024.100800

Geravandi, S., Richardson, S., Pugliese, A., Maedler, K., 2021. Localization of enteroviral RNA within the pancreas in donors with T1D and T1D-associated autoantibodies. Cell Reports Med. 2. https://doi.org/10.1016/j.xcrm.2021.100371

Geyer, P.E., Holdt, L.M., Teupser, D., Mann, M., 2017. Revisiting biomarker discovery by plasma proteomics. Mol. Syst. Biol. 13, 942. https://doi.org/10.15252/msb.20156297

Geyer, P.E., Kulak, N.A., Pichler, G., Holdt, L.M., Teupser, D., Mann, M., 2016. Plasma Proteome Profiling to Assess Human Health and Disease. Cell Syst. 2, 185–195. https://doi.org/10.1016/J.CELS.2016.02.015

Geyer, P.E., Voytik, E., Treit, P. V, Doll, S., Kleinhempel, A., Niu, L., Müller, J.B., Buchholtz, M., Bader, J.M., Teupser, D., Holdt, L.M., Mann, M., 2019. Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. EMBO Mol. Med. 11. https://doi.org/10.15252/EMMM.201910427

Ghalwash, M., Anand, V., Ng, K., Dunne, J.L., Lou, O., Lundgren, M., Hagopian, W.A., Rewers, M., Ziegler, A.G., Veijola, R., 2024. Data-Driven Phenotyping of Presymptomatic Type 1 Diabetes Using Longitudinal Autoantibody Profiles. Diabetes Care 47, 1424–1431. https://doi.org/10.2337/dc24-0198

Giannopoulou, E.Z., Winkler, C., Chmiel, R., Matzke, C., Scholz, M., Beyerlein, A., Achenbach, P., Bonifacio, E., Ziegler, A.G., 2015. Islet autoantibody phenotypes and incidence in children at increased risk for type 1 diabetes. Diabetologia 58, 2317–2323. https://doi.org/10.1007/s00125-015-3672-y

Gootjes, C., Zwaginga, J.J., Roep, B.O., Nikolic, T., 2022. Functional Impact of Risk Gene Variants on the Autoimmune Responses in Type 1 Diabetes. Front. Immunol. 13. https://doi.org/10.3389/fimmu.2022.886736

Gregory, G.A., Robinson, T.I.G., Linklater, S.E., Wang, F., Colagiuri, S., de Beaufort, C., Donaghue, K.C., Magliano, D.J., Maniam, J., Orchard, T.J., Rai, P., Ogle, G.D., 2022. Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. Lancet Diabetes Endocrinol. 10, 741–760. https://doi.org/10.1016/S2213-8587(22)00218-2

Guo, T., Steen, J.A., Mann, M., 2025. Mass-spectrometry-based proteomics: from single cells to clinical applications. Nat. 2025 6388052 638, 901–911. https://doi.org/10.1038/s41586-025-08584-0

Haag, A.M., 2016. Mass analyzers and mass spectrometers. Adv. Exp. Med. Biol. 919, 157–169. https://doi.org/10.1007/978-3-319-41448-5_7

Hämäläinen, S., Nurminen, N., Ahlfors, H., Oikarinen, S., Sioofy-Khojine, A.B., Frisk, G., Oberste, M.S., Lahesmaa, R., Pesu, M., Hyöty, H., 2014. Coxsackievirus B1 reveals strain specific differences in plasmacytoid dendritic cell mediated immunogenicity. J. Med. Virol. 86, 1412–1420. https://doi.org/10.1002/jmv.23903

Han, B., Luo, H., Raelson, J., Huang, J., Li, Y., Tremblay, J., Hu, B., Qi, S., Wu, J., 2014. TGFBI (βIG-H3) is a diabetes-risk gene based on mouse and human genetic studies. Hum. Mol. Genet. 23, 4597–4611. https://doi.org/10.1093/HMG/DDU173

Harjutsalo, V., Sund, R., Knip, M., Groop, P.-H., 2013. Incidence of type 1 diabetes in Finland. JAMA 310, 427–428. https://doi.org/10.1001/jama.2013.8399

Hedman, C.A., Frystyk, J., Lindström, T., Chen, J.-W., Flyvbjerg, A., Ørskov, H., Arnqvist, H.J., 2004. Residual beta-cell function more than glycemic control determines abnormalities of the insulin-like growth factor system in type 1 diabetes. J. Clin. Endocrinol. Metab. 89, 6305–6309. https://doi.org/10.1210/jc.2004-0572

Heil, L.R., Damoc, E., Arrey, T.N., Pashkova, A., Denisov, E., Petzoldt, J., Peterson, A.C., Hsu, C., Searle, B.C., Shulman, N., Riffle, M., Connolly, B., MacLean, B.X., Remes, P.M., Senko, M.W., Stewart, H.I., Hock, C., Makarov, A.A., Hermanson, D., Zabrouskov, V., Wu, C.C., MacCoss, M.J., 2023. Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data-Independent Acquisition. J. Proteome Res. 22, 3290–3300. https://doi.org/10.1021/ACS.JPROTEOME.3C00357/SUPPL_FILE/PR3C00357_SI_001.XLSX

Herold, K.C., Gitelman, S.E., Gottlieb, P.A., Knecht, L.A., Raymond, R., Ramos, E.L., 2023. Teplizumab: A Disease-Modifying Therapy for Type 1 Diabetes That Preserves β-Cell Function. Diabetes Care 46, 1848–1856. https://doi.org/10.2337/dc23-0675

Hiramoto, E., Tsutsumi, A., Suzuki, R., Matsuoka, S., Arai, S., Kikkawa, M., Miyazaki, T., 2018. The IgM pentamer is an asymmetric pentagon with an open groove that binds the AIM protein. Sci. Adv. 4. https://doi.org/10.1126/sciadv.aau1199

Honkanen, H., Oikarinen, S., Nurminen, N., Laitinen, O.H., Huhtala, H., Lehtonen, J., Ruokoranta, T., Hankaniemi, M.M., Lecouturier, V., Almond, J.W., Tauriainen, S., Simell, O., Ilonen, J., Veijola, R., Viskari, H., Knip, M., Hyöty, H., 2017. Detection of enteroviruses in stools precedes islet autoimmunity by several months: possible evidence for slowly operating mechanisms in virus-induced autoimmunity. Diabetologia 60, 424–431. https://doi.org/10.1007/s00125-016-4177-z

Honkanen, H., Oikarinen, S., Pakkanen, O., Ruokoranta, T., Pulkki, M.M., Laitinen, O.H., Tauriainen, S., Korpela, S., Lappalainen, M., Vuorinen, T., Haapala, A.M., Veijola, R., Simell, O., Ilonen, J., Knip, M., Hyöty, H., 2013. Human enterovirus 71 strains in the background population and in hospital patients in Finland. J. Clin. Virol. 56, 348–353. https://doi.org/10.1016/j.jcv.2012.11.018

Honkimaa, A., Kimura, B., Sioofy-Khojine, A.B., Lin, J., Laiho, J., Oikarinen, S., Hyöty, H., 2020. Genetic Adaptation of Coxsackievirus B1 during Persistent Infection in Pancreatic Cells. Microorganisms 8, 1790. https://doi.org/10.3390/MICROORGANISMS8111790

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, T., Campos, A.R., Wang, J., Stukalov, A., Díaz, R., Maurya, S., Motamedchaboki, K., Hornburg, D., Saciloto-de-Oliveira, L.R., Innocente-Alves, C., Calegari-Alves, Y.P., Batzoglou, S., Beys-da-Silva, W.O., Santi, L., 2025. Deep, Unbiased, and Quantitative Mass Spectrometry-Based Plasma Proteome Analysis of Individual Responses to mRNA COVID-19 Vaccine. J. Proteome Res. 24, 1265–1274. https://doi.org/10.1021/ACS.JPROTEOME.4C00909/ASSET/IMAGES/LARGE/PR4C00909_0004.JPEG

Hyöty, H., Kääriäinen, S., Laiho, J.E., Comer, G.M., Tian, W., Härkönen, T., Lehtonen, J.P., Oikarinen, S., Puustinen, L., Snyder, M., León, F., Scheinin, M., Knip, M., Sanjuan, M., 2024. Safety, tolerability and immunogenicity of PRV-101, a multivalent vaccine targeting coxsackie B viruses (CVBs) associated with type 1 diabetes: a double-blind randomised placebo-controlled Phase I trial. Diabetologia 67, 811–821. https://doi.org/10.1007/S00125-024-06092-W/FIGURES/2

Ianevski, A., Yao, R., Biza, S., Zusinaite, E., Mannik, A., Kivi, G., Planken, A., Kurg, K., Tombak, E.-M., Ustav, M., Shtaida, N., Kulesskiy, E., Jo, E., Yang, J., Lysvand, H., Løseth, K., Oksenych, V., Aas, P.A., Tenson, T., Vitkauskienė, A., Windisch, M.P., Fenstad, M.H., Nordbø, S.A., Ustav, M., Bjørås, M., Kainov, D.E., 2020. Identification and Tracking of Antiviral Drug Combinations. Viruses 12. https://doi.org/10.3390/v12101178

Ifie, E., Russell, M.A., Dhayal, S., Leete, P., Sebastiani, G., Nigi, L., Dotta, F., Marjomäki, V., Eizirik, D.L., Morgan, N.G., Richardson, S.J., 2018. Unexpected subcellular distribution of a specific isoform of the Coxsackie and adenovirus receptor, CAR-SIV, in human pancreatic beta cells. Diabetologia 61, 2344–2355. https://doi.org/10.1007/s00125-018-4704-1

Ignjatovic, V., Geyer, P.E., Palaniappan, K.K., Chaaban, J.E., Omenn, G.S., Baker, M.S., Deutsch, E.W., Schwenk, J.M., 2019. Mass Spectrometry-Based Plasma Proteomics: Considerations from Sample Collection to Achieving Translational Data. J. Proteome Res. 18, 4085. https://doi.org/10.1021/ACS.JPROTEOME.9B00503

Ilonen, J., Hammais, A., Laine, A.P., Lempainen, J., Vaarala, O., Veijola, R., Simell, O., Knip, M., 2013. Patterns of β-cell autoantibody appearance and genetic associations during the first years of life. Diabetes 62, 3636–3640. https://doi.org/10.2337/db13-0300

Ilonen, J., Lempainen, J., Veijola, R., 2019. The heterogeneous pathogenesis of type 1 diabetes mellitus. Nat. Rev. Endocrinol. 15, 635–650. https://doi.org/10.1038/s41574-019-0254-y

Infante, M., Ricordi, C., Sanchez, J., Clare-Salzler, M.J., Padilla, N., Fuenmayor, V., Chavez, C., Alvarez, A., Baidal, D., Alejandro, R., Caprio, M., Fabbri, A., 2019. Influence of Vitamin D on Islet Autoimmunity and Beta-Cell Function in Type 1 Diabetes. Nutrients 11, 2185. https://doi.org/10.3390/NU11092185

Insel, R.A., Dunne, J.L., Atkinson, M.A., Chiang, J.L., Dabelea, D., Gottlieb, P.A., Greenbaum, C.J., Herold, K.C., Krischer, J.P., Lernmark, Å., Ratner, R.E., Rewers, M.J., Schatz, D.A., Skyler, J.S., Sosenko, J.M., Ziegler, A.-G., 2015. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. Diabetes Care 38, 1964–1974. https://doi.org/10.2337/dc15-1419

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nat. 2021 5967873 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kallionpää, H., Laajala, E., Öling, V., Härkönen, T., Tillmann, V., Dorshakova, N. V., Ilonen, J., Lähdesmäki, H., Knip, M., Lahesmaa, R., Koski, K., Koski, M., Ryhänen, S., Siljander, H., Hämäläinen, A.M., Ormisson, A., Peet, A., Ulich, V., Kuzmicheva, E., Mokurov, S., Markova, S., Pylova, S., Isakova, M., Shakurova, E., Petrov, V., Karapetyan, T., Varlamova, T., Ilonen, J., Kiviniemi, M., Alnek, K., Janson, H., Uibo, R., Salum, T., von Mutius, E., Weber, J., Ahlfors, H., Moulder, R., Nieminen, J., Ruohtula, T., Vaarala, O., Honkanen, H., Hyöty, H., Kondrashova, A., Oikarinen, S., Harmsen, H.J.M., De Goffau, M.C., Welling, G., Alahuhta, K., Korhonen, T., Virtanen, S.M., 2015. Standard of hygiene and immune adaptation in newborn infants. Clin. Immunol. 155, 136–147. https://doi.org/10.1016/j.clim.2014.09.009

Kim, K.W., Horton, J.L., Pang, C.N.I., Jain, K., Leung, P., Isaacs, S.R., Bull, R.A., Luciani, F., Wilkins, M.R., Catteau, J., Lipkin, W.I., Rawlinson, W.D., Briese, T., Craig, M.E., 2019. Higher abundance of enterovirus A species in the gut of children with islet autoimmunity. Sci. Rep. 9. https://doi.org/10.1038/s41598-018-38368-8

Korpos, É., Kadri, N., Kappelhoff, R., Wegner, J., Overall, C.M., Weber, E., Holmberg, D., Cardell, S., Sorokin, L., 2013. The peri-islet basement membrane, a barrier to infiltrating leukocytes in type 1 diabetes in mouse and human. Diabetes 62, 531–542. https://doi.org/10.2337/db12-0432

Krogvold, L., Genoni, A., Puggioni, A., Campani, D., Richardson, S.J., Flaxman, C.S., Edwin, B., Buanes, T., Dahl-Jørgensen, K., Toniolo, A., 2022. Live enteroviruses, but not other viruses, detected in human pancreas at the onset of type 1 diabetes in the DiViD study. Diabetologia 65, 2108–2120. https://doi.org/10.1007/s00125-022-05779-2

Krogvold, L., Skog, O., Sundström, G., Edwin, B., Buanes, T., Hanssen, K.F., Ludvigsson, J., Grabherr, M., Korsgren, O., Dahl-Jørgensen, K., 2015. Function of isolated pancreatic islets from patients at onset of type 1 diabetes: Insulin secretion can be restored after some days in a nondiabetogenic environment in vitro results from the divid study. Diabetes 64, 2506–2512. https://doi.org/10.2337/db14-1911

Kupila, A., Muona, P., Simell, T., Arvilommi, P., Savolainen, H., Hämäläinen, A.M., Korhonen, S., Kimpimäki, T., Sjöroos, M., Ilonen, J., Knip, M., Simell, O., 2001. Feasibility of genetic and immunological prediction of type I diabetes in a population-based birth cohort. Diabetologia 44, 290–297. https://doi.org/10.1007/s001250051616

Laajala, M., Zwaagstra, M., Martikainen, M., Nekoua, M.P., Benkahla, M., Sane, F., Gervais, E., Campagnola, G., Honkimaa, A., Sioofy-Khojine, A.-B., Hyöty, H., Ojha, R., Bailliot, M., Balistreri, G., Peersen, O., Hober, D., Van Kuppeveld, F., Marjomäki, V., 2023. Vemurafenib Inhibits Acute and Chronic Enterovirus Infection by Affecting Cellular Kinase Phosphatidylinositol 4-Kinase Type IIIβ. Microbiol. Spectr. 11. https://doi.org/10.1128/SPECTRUM.00552-23

Laiho, J.E., Oikarinen, S., Oikarinen, M., Larsson, P.G., Stone, V.M., Hober, D., Oberste, S., Flodström-Tullberg, M., Isola, J., Hyöty, H., 2015. Application of bioinformatics in probe design enables detection of enteroviruses on different taxonomic levels by advanced in situ hybridization technology. J. Clin. Virol. 69, 165–171. https://doi.org/10.1016/J.JCV.2015.06.085

Laitinen, O.H., Honkanen, H., Pakkanen, O., Oikarinen, S., Hankaniemi, M.M., Huhtala, H., Ruokoranta, T., Lecouturier, V., André, P., Harju, R., Virtanen, S.M., Lehtonen, J., Almond, J.W., Simell, T., Simell, O., Ilonen, J., Veijola, R., Knip, M., Hyöty, H., 2014. Coxsackievirus b1 is associated with induction of β-cells autoimmunity that portends type 1 diabetes. Diabetes 63, 446–455. https://doi.org/10.2337/db13-0619

Larsson, K., Elding-Larsson, H., Cederwall, E., Kockum, K., Neiderud, J., Sjölad, S., Lindberg, B., Lernmark, B., Cilio, C., Ivarsson, S.A., Lernmark, Å., 2004. Genetic and perinatal factors as risk for childhood type 1 diabetes. Diabetes. Metab. Res. Rev. 20, 429–437. https://doi.org/10.1002/dmrr.506

Leighton, E., Sainsbury, C.A., Jones, G.C., 2017. A Practical Review of C-Peptide Testing in Diabetes. Diabetes Ther. 8, 475–487. https://doi.org/10.1007/s13300-017-0265-4

Lenčo, J., Jadeja, S., Naplekov, D.K., Krokhin, O. V., Khalikova, M.A., Chocholouš, P., Urban, J., Broeckhoven, K., Nováková, L., Švec, F., 2022. Reversed-Phase Liquid Chromatography of Peptides for Bottom-Up Proteomics: A Tutorial. J. Proteome Res. 21, 2846–2892. https://doi.org/10.1021/acs.jproteome.2c00407

Lernmark, Å., Akolkar, B., Hagopian, W., Krischer, J., McIndoe, R., Rewers, M., Toppari, J., Vehik, K., Ziegler, A.G., 2023. Possible heterogeneity of initial pancreatic islet beta-cell autoimmunity heralding type 1 diabetes. J. Intern. Med. 294, 145–158. https://doi.org/10.1111/joim.13648

Leslie, R.D., Evans-Molina, C., Freund-Brown, J., Buzzetti, R., Dabelea, D., Gillespie, K.M., Goland, R., Jones, A.G., Kacher, M., Phillips, L.S., Rolandsson, O., Wardian, J.L., Dunne, J.L., 2021. Adult-onset type 1 diabetes: Current understanding and challenges. Diabetes Care 44, 2449–2456. https://doi.org/10.2337/dc21-0770

Lietzén, N., Cheng, L., Moulder, R., Siljander, H., Laajala, E., Härkönen, T., Peet, A., Vehtari, A., Tillmann, V., Knip, M., Lähdesmäki, H., Lahesmaa, R., 2018. Characterization and non-parametric modeling of the developing serum proteome during infancy and early childhood. Sci. Rep. 8, 5883. https://doi.org/10.1038/s41598-018-24019-5

Lindemann, C., Thomanek, N., Hundt, F., Lerari, T., Meyer, H.E., Wolters, D., Marcus, K., 2017. Strategies in relative and absolute quantitative mass spectrometry based proteomics. Biol. Chem. 398, 687–699. https://doi.org/10.1515/HSZ-2017-0104

Liu, C.-W., Bramer, L., Webb-Robertson, B.-J., Waugh, K., Rewers, M.J., Zhang, Q., 2018. Temporal expression profiling of plasma proteins reveals oxidative stress in early stages of Type 1 Diabetes progression. J. Proteomics 172, 100–110.

Liu, C.-W., Bramer, L., Webb-Robertson, B.-J., Waugh, K., Rewers, M.J., Zhang, Q., 2017. Temporal profiles of plasma proteome during childhood development. J. Proteomics 152, 321–328.

Long, S.A., Linsley, P.S., 2024. Integrating Omics into Functional Biomarkers of Type 1 Diabetes. Cold Spring Harb. Perspect. Med. 14. https://doi.org/10.1101/cshperspect.a041602

Lorenzen, T., Pociot, F., Stilgren, L., Kristiansen, O.P., Johannesen, J., Olsen, P.B., Walmar, A., Larsen, A., Albrechtsen, N.C., Eskildsen, P.C., Andersen, O.O., Nerup, J., Enk, B., Hansen, F.M., Juhl, H., Laursen, H.B., Sølling, K., Vadstrap, S., 1998. Predictors of IDDM recurrence risk in offspring of Danish IDDM patients. Diabetologia 41, 666–673. https://doi.org/10.1007/s001250050966

MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., MacCoss, M.J., 2010. Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26, 966–968. https://doi.org/10.1093/bioinformatics/btq054

Mamyrin, B.A., 2001. Time-of-flight mass spectrometry (concepts, achievements, and prospects). Int. J. Mass Spectrom. 206, 251–266. https://doi.org/10.1016/S1387-3806(00)00392-4

Mann, M., Kumar, C., Zeng, W.F., Strauss, M.T., 2021. Artificial intelligence for proteomics and biomarker discovery. Cell Syst. 12, 759–770. https://doi.org/10.1016/J.CELS.2021.06.006

Mann, M., Steen, H., 2004. The abc's (and xyz's) of peptide sequencing. Nat. Rev. Mol. Cell Biol. 5, 699–711. https://doi.org/10.1038/nrm1468

Mant, C.T., Chen, Y., Yan, Z., Popa, T. V., Kovacs, J.M., Mills, J.B., Tripet, B.P., Hodges, R.S., 2007. HPLC analysis and purification of peptides. Methods Mol. Biol. 386, 3–55. https://doi.org/10.1007/978-1-59745-430-8_1

March, R., 1997. An introduction to quadrupole ion trap mass spectrometry. J. mass Spectrom. 32, 351–369. https://doi.org/10.1002/(SICI)1096-9888(199704)32:4<351::AID-JMS512>3.3.CO;2-P

Metz, T.O., Qian, W.-J., Jacobs, J.M., Gritsenko, M.A., Moore, R.J., Polpitiya, A.D., Monroe, M.E., Camp, D.G., Mueller, P.W., Smith, R.D., 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.

Miettinen, M.E., Niinistö, S., Erlund, I., Cuthbertson, D., Nucci, A.M., Honkanen, J., Vaarala, O., Hyöty, H., Krischer, J.P., Knip, M., Virtanen, S.M., 2020. Serum 25-hydroxyvitamin D concentration in childhood and risk of islet autoimmunity and type 1 diabetes: the TRIGR nested case–control ancillary study. Diabetologia 63, 780–787. https://doi.org/10.1007/s00125-019-05077-4

Mikus, M., Järnbert-Pettersson, H., Johansson, C., Nilsson, P., Scheynius, A., Alm, J., 2021. Protein profiles in plasma: Development from infancy to 5 years of age. Proteomics. Clin. Appl. e2000038. https://doi.org/10.1002/prca.202000038

Millioni, R., Tolin, S., Puricelli, L., Sbrignadello, S., Fadini, G.P., Tessari, P., Arrigoni, G., 2011. High Abundance Proteins Depletion vs Low Abundance Proteins Enrichment: Comparison of Methods to Reduce the Plasma Proteome Complexity. PLoS One 6, e19603. https://doi.org/10.1371/JOURNAL.PONE.0019603

Mirzaei, H., Carrasco, M., 2016. Modern proteomics : sample preparation, analysis and practical applications 533.

Mobasseri, M., Shirmohammadi, M., Amiri, T., Vahed, N., Fard, H.H., Ghojazadeh, M., 2020. Prevalence and incidence of type 1 diabetes in the world: A systematic review and meta-analysis. Heal. Promot. Perspect. 10, 98–115. https://doi.org/10.34172/hpp.2020.18

Moulder, R., Bhosale, S.D., Erkkilä, T., Laajala, E., Salmi, J., Nguyen, E. V, Kallionpää, H., Mykkänen, J., Vähä-Mäkilä, M., Hyöty, H., Veijola, R., Ilonen, J., Simell, T., Toppari, J., Knip, M., Goodlett, D.R., Lähdesmäki, H., Simell, O., Lahesmaa, R., 2015. Serum proteomes distinguish children developing type 1 diabetes in a cohort with HLA-conferred susceptibility. Diabetes 64, 2265–2278. https://doi.org/10.2337/db14-0983

Nakayasu, E.S., Bramer, L.M., Ansong, C., Schepmoes, A.A., Fillmore, T.L., Gritsenko, M.A., Clauss, T.R., Gao, Y., Piehowski, P.D., Stanfill, B.A., Engel, D.W., Orton, D.J., Moore, R.J., Qian, W.J., Sechi, S., Frohnert, B.I., Toppari, J., Ziegler, A.G., Lernmark, Å., Hagopian, W., Akolkar, B., Smith, R.D., Rewers, M.J., Webb-Robertson, B.J.M., Metz, T.O., 2023. Plasma protein biomarkers predict the development of persistent autoantibodies and type 1 diabetes 6 months prior to the onset of autoimmunity. Cell reports. Med. 4. https://doi.org/10.1016/J.XCRM.2023.101093

Nakayasu, E.S., Gritsenko, M., Piehowski, P.D., Gao, Y., Orton, D.J., Schepmoes, A.A., Fillmore, T.L., Frohnert, B.I., Rewers, M., Krischer, J.P., Ansong, C., Suchy-Dicey, A.M., Evans-Molina, C., Qian, W.J., Webb-Robertson, B.J.M., Metz, T.O., 2021. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. Nat. Protoc. 2021 168 16, 3737–3760. https://doi.org/10.1038/s41596-021-00566-6

Nanjappa, V., Thomas, J.K., Marimuthu, A., Muthusamy, B., Radhakrishnan, A., Sharma, R., Ahmad Khan, A., Balakrishnan, L., Sahasrabuddhe, N.A., Kumar, S., Jhaveri, B.N., Sheth, K.V., Kumar Khatana, R., Shaw, P.G., Srikanth, S.M., Mathur, P.P., Shankar, S., Nagaraja, D., Christopher, R., Mathivanan, S., Raju, R., Sirdeshmukh, R., Chatterjee, A., Simpson, R.J., Harsha, H.C., Pandey,

A., Prasad, T.S.K., 2014. Plasma Proteome Database as a resource for proteomics research: 2014 update. Nucleic Acids Res. 42, D959–D965. https://doi.org/10.1093/nar/gkt1251

Nekoua, M.P., Alidjinou, E.K., Hober, D., 2022. Persistent coxsackievirus B infection and pathogenesis of type 1 diabetes mellitus. Nat. Rev. Endocrinol. 18, 503–516. https://doi.org/10.1038/s41574-022-00688-1

Nekoua, M.P., Bertin, A., Sane, F., Alidjinou, E.K., Lobert, D., Trauet, J., Hober, C., Engelmann, I., Moutairou, K., Yessoufou, A., Hober, D., 2020. Pancreatic beta cells persistently infected with coxsackievirus B4 are targets of NK cell-mediated cytolytic activity. Cell. Mol. Life Sci. 77, 179–194. https://doi.org/10.1007/s00018-019-03168-4

Netanyah, E., Calafatti, M., Arvastsson, J., Cabrera-Rode, E., Cilio, C.M., Sarmiento, L., 2020. Extracellular Vesicles Released by Enterovirus-Infected EndoC-βH1 Cells Mediate Non-Lytic Viral Spread. Microorganisms 8, 1–13. https://doi.org/10.3390/microorganisms8111753

Niu, L., Stinson, S.E., Holm, L.A., Asp, M., Lund, V., Esmann Fonvig, C., Cobuccio, L., Meisner, J., Juel, H.B., Fadista, J., Thiele, M., Krag, A., Holm, J.-C., Rasmussen, S., Hansen, T., Mann, M., 2025. Plasma proteome variation and its genetic determinants in children and adolescents. Nat. Genet. 2025 1–12. https://doi.org/10.1038/s41588-025-02089-2

Noble, J.A., Valdes, A.M., 2011. Genetics of the HLA region in the prediction of type 1 diabetes. Curr. Diab. Rep. 11, 533–542. https://doi.org/10.1007/s11892-011-0223-x

Nolting, D., Malek, R., Makarov, A., 2019. Ion traps in modern mass spectrometry. Mass Spectrom. Rev. 38, 150–168. https://doi.org/10.1002/mas.21549

Nosti, A.J., Barrio, L.C., Calderón Celis, F., Soldado, A., Encinar, J.R., 2022. Absolute quantification of proteins using element mass spectrometry and generic standards. J. Proteomics 256, 104499. https://doi.org/10.1016/J.JPROT.2022.104499

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M.R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S.J., Diekhans, M., Logsdon, G.A., Alonge, M., Antonarakis, S.E., Borchers, M., Bouffard, G.G., Brooks, S.Y., Caldas, G. V., Chen, N.C., Cheng, H., Chin, C.S., Chow, W., de Lima, L.G., Dishuck, P.C., Durbin, R., Dvorkina, T., Fiddes, I.T., Formenti, G., Fulton, R.S., Fungtammasan, A., Garrison, E., Grady, P.G.S., Graves-Lindsay, T.A., Hall, I.M., Hansen, N.F., Hartley, G.A., Haukness, M., Howe, K., Hunkapiller, M.W., Jain, C., Jain, M., Jarvis, E.D., Kerpedjiev, P., Kirsche, M., Kolmogorov, M., Korlach, J., Kremitzki, M., Li, H., Maduro, V. V., Marschall, T., McCartney, A.M., McDaniel, J., Miller, D.E., Mullikin, J.C., Myers, E.W., Olson, N.D., Paten, B., Peluso, P., Pevzner, P.A., Porubsky, D., Potapova, T., Rogaev, E.I., Rosenfeld, J.A., Salzberg, S.L., Schneider, V.A., Sedlazeck, F.J., Shafin, K., Shew, C.J., Shumate, A., Sims, Y., Smit, A.F.A., Soto, D.C., Sovi, I., Storer, J.M., Streets, A., Sullivan, B.A., Thibaud-Nissen, F., Torrance, J., Wagner, J., Walenz, B.P., Wenger, A., Wood, J.M.D., Xiao, C., Yan, S.M., Young, A.C., Zarate, S., Surti, U., McCoy, R.C., Dennis, M.Y., Alexandrov, I.A., Gerton, J.L., O'Neill, R.J., Timp, W., Zook, J.M., Schatz, M.C., Eichler, E.E., Miga, K.H., Phillippy, A.M., 2022. The complete sequence of a human genome. Science (80-.). 376, 44–53. https://doi.org/10.1126/science.abj6987

Ogle, G.D., Wang, F., Gregory, G.A., Maniam, J., 2022. IDF Atlas Reports : Type 1 Diabetes Estimates in Children and Adults, International Diabetes Federation.

Oikarinen, M., Tauriainen, S., Honkanen, T., Oikarinen, S., Vuori, K., Kaukinen, K., Rantala, I., Mäki, M., Hyöty, H., 2008. Detection of enteroviruses in the intestine of type 1 diabetic patients. Clin. Exp. Immunol. 151, 71–75. https://doi.org/10.1111/j.1365-2249.2007.03529.x

Oikarinen, S., Martiskainen, M., Tauriainen, S., Huhtala, H., Ilonen, J., Veijola, R., Simell, O., Knip, M., Hyöty, H., 2011. Enterovirus RNA in blood is linked to the development of type 1 diabetes. Diabetes 60, 276–279. https://doi.org/10.2337/db10-0186

Oikarinen, S., Tauriainen, S., Hober, D., Lucas, B., Vazeou, A., Sioofy-Khojine, A., Bozas, E., Muir, P., Honkanen, H., Ilonen, J., Knip, M., Keskinen, P., Saha, M.T., Huhtala, H., Stanway, G., Bartsocas, C., Ludvigsson, J., Taylor, K., Hyöty, H., 2014. Virus antibody survey in different

european populations indicates risk association between coxsackievirus B1 and type 1 diabetes. Diabetes 63, 655–662. https://doi.org/10.2337/db13-0620

Oilinki, T., Otonkoski, T., Ilonen, J., Knip, M., Miettinen, P., 2012. Prevalence and characteristics of diabetes among Somali children and adolescents living in Helsinki, Finland. Pediatr. Diabetes 13, 176–180. https://doi.org/10.1111/j.1399-5448.2011.00783.x

Östman, J., Lönnberg, G., Arnqvist, H.J., Blohmé, G., Bolinder, J., Schnell, A.E., Eriksson, J.W., Gudbjörnsdottir, S., Sundkvist, G., Nyström, L., 2008. Gender differences and temporal variation in the incidence of type 1 diabetes: Results of 8012 cases in the nationwide Diabetes Incidence Study in Sweden 1983-2002. J. Intern. Med. 263, 386–394. https://doi.org/10.1111/j.1365-2796.2007.01896.x

Owusu, I.A., Quaye, O., Passalacqua, K.D., Wobus, C.E., 2021. Egress of non-enveloped enteric RNA viruses. J. Gen. Virol. 102. https://doi.org/10.1099/JGV.0.001557

Patelarou, E., Girvalaki, C., Brokalaki, H., Patelarou, A., Androulaki, Z., Vardavas, C., 2012. Current evidence on the associations of breastfeeding, infant formula, and cow's milk introduction with type 1 diabetes mellitus: A systematic review. Nutr. Rev. 70, 509–519. https://doi.org/10.1111/j.1753-4887.2012.00513.x

Patterson, C.C., Dahlquist, G.G., Gyürüs, E., Green, A., Soltész, G., Group, E.S., 2009. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study. Lancet (London, England) 373, 2027–2033.

Patterson, C.C., Harjutsalo, V., Rosenbauer, J., Neu, A., Cinek, O., Skrivarhaug, T., Rami-Merhar, B., Soltesz, G., Svensson, J., Parslow, R.C., Castell, C., Schoenle, E.J., Bingley, P.J., Dahlquist, G., Jarosz-Chobot, P.K., Marčiulionytė, D., Roche, E.F., Rothe, U., Bratina, N., Ionescu-Tirgoviste, C., Weets, I., Kocova, M., Cherubini, V., Rojnic Putarek, N., deBeaufort, C.E., Samardzic, M., Green, A., 2019. Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989-2013: a multicentre prospective registration study. Diabetologia 62, 408–417. https://doi.org/10.1007/s00125-018-4763-3

Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., Walzer, M., Wang, S., Brazma, A., Vizcaíno, J.A., 2022. The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50, D543–D552. https://doi.org/10.1093/nar/gkab1038

Petersen, T.N., Brunak, S., Von Heijne, G., Nielsen, H., 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786. https://doi.org/10.1038/nmeth.1701

Pinkert, S., Klingel, K., Lindig, V., Dörner, A., Zeichhardt, H., Spiller, O.B., Fechner, H., 2011. Virus-Host Coevolution in a Persistently Coxsackievirus B3-Infected Cardiomyocyte Cell Line. J. Virol. 85, 13409–13419. https://doi.org/10.1128/jvi.00621-11

Pociot, F., Akolkar, B., Concannon, P., Erlich, H.A., Julier, C., Morahan, G., Nierras, C.R., Todd, J.A., Rich, S.S., Nerup, J., 2010. Genetics of type 1 diabetes: What's next? Diabetes 59, 1561–1571. https://doi.org/10.2337/db10-0076

Predieri, B., Bruzzi, P., Bigi, E., Ciancia, S., Madeo, S.F., Lucaccioni, L., Iughetti, L., 2020. Endocrine disrupting chemicals and type 1 diabetes. Int. J. Mol. Sci. 21. https://doi.org/10.3390/ijms21082937

R Core Team, 2018. R: A language and environment for statistical computing.

Rajpathak, S.N., He, M., Sun, Q., Kaplan, R.C., Muzumdar, R., Rohan, T.E., Gunter, M.J., Pollak, M., Kim, M., Pessin, J.E., Beasley, J., Wylie-Rosett, J., Hu, F.B., Strickler, H.D., 2012. Insulin-Like Growth Factor Axis and Risk of Type 2 Diabetes in Women. Diabetes 61, 2248–2254. https://doi.org/10.2337/db11-1488

Ramond, C., Glaser, N., Berthault, C., Ameri, J., Kirkegaard, J.S., Hansson, M., Honoré, C., Semb, H., Scharfmann, R., 2017. Reconstructing human pancreatic differentiation by mapping specific cell populations during development. Elife 6. https://doi.org/10.7554/eLife.27564

Redondo, M.J., Jeffrey, J., Fain, P.R., Eisenbarth, G.S., Orban, T., 2008. Concordance for Islet Autoimmunity among Monozygotic Twins. N. Engl. J. Med. 359, 2849–2850. https://doi.org/10.1056/nejmc0805398

Redondo, M.J., Morgan, N.G., 2023. Heterogeneity and endotypes in type 1 diabetes mellitus. Nat. Rev. Endocrinol. 19, 542–554. https://doi.org/10.1038/s41574-023-00853-0

Rewers, M., Ludvigsson, J., 2016. Environmental risk factors for type 1 diabetes. Lancet 387, 2340–2348. https://doi.org/10.1016/S0140-6736(16)30507-4

Rewers, M., She, J.X., Ziegler, A.G., Simell, O.G., Lernmark, Å., Hagopian, W.A., Krischer, J.P., Akolkar, B., 2008. The Environmental Determinants of Diabetes in the Young (TEDDY) Study. Ann. N. Y. Acad. Sci. 1150, 1. https://doi.org/10.1196/ANNALS.1447.062

Richardson, S.J., Leete, P., Bone, A.J., Foulis, A.K., Morgan, N.G., 2013. Expression of the enteroviral capsid protein VP1 in the islet cells of patients with type 1 diabetes is associated with induction of protein kinase R and downregulation of Mcl-1. Diabetologia 56, 185–193. https://doi.org/10.1007/s00125-012-2745-4

Roberts, D.S., Loo, J.A., Tsybin, Y.O., Liu, X., Wu, S., Chamot-Rooke, J., Agar, J.N., Paša-Tolić, L., Smith, L.M., Ge, Y., 2024. Top-down proteomics. Nat. Rev. Methods Prim. 4. https://doi.org/10.1038/S43586-024-00318-2

Rouland, A., Masson, D., Lagrost, L., Vergès, B., Gautier, T., Bouillet, B., 2022. Role of apolipoprotein C1 in lipoprotein metabolism, atherosclerosis and diabetes: a systematic review. Cardiovasc. Diabetol. 21. https://doi.org/10.1186/s12933-022-01703-5

Ryall, C.L., Viloria, K., Lhaf, F., Walker, A.J., King, A., Jones, P., Mackintosh, D., McNeice, R., Kocher, H., Flodstrom-Tullberg, M., Edling, C., Hill, N.J., 2014. Novel role for matricellular proteins in the regulation of islet βcell survival: The effect of SPARC on survival, proliferation, and signaling. J. Biol. Chem. 289, 30614–30624. https://doi.org/10.1074/jbc.M114.573980

Sacks, D.B., Arnold, M., Bakris, G.L., Bruns, D.E., Horvath, A.R., Lernmark, Å., Metzger, B.E., Nathan, D.M., Kirkman, M.S., 2023. Guidelines and Recommendations for Laboratory Analysis in the Diagnosis and Management of Diabetes Mellitus. Diabetes Care 46, e151–e199. https://doi.org/10.2337/dci23-0036

Scharfmann, R., Pechberty, S., Hazhouz, Y., Von Bülow, M., Bricout-Neveu, E., Grenier-Godard, M., Guez, F., Rachdi, L., Lohmann, M., Czernichow, P., Ravassard, P., 2014. Development of a conditionally immortalized human pancreatic β cell line. J. Clin. Invest. 124, 2087–2098. https://doi.org/10.1172/JCI72674

Sean, P., Nguyen, J.H.C., Semler, B.L., 2009. Altered interactions between stem-loop IV within the 5′ noncoding region of coxsackievirus RNA and poly(rC) binding protein 2: Effects on IRES-mediated translation and viral infectivity. Virology 389, 45–58. https://doi.org/10.1016/j.virol.2009.03.012

Sharma, V., Eckels, J., Schilling, B., Ludwig, C., Jaffe, J.D., MacCoss, M.J., MacLean, B., 2018. Panorama public: A public repository for quantitative data sets processed in skyline. Mol. Cell. Proteomics 17, 1239–1244. https://doi.org/10.1074/mcp.RA117.000543

Shevchenko, A., Wilm, M., Vorm, O., Mann, M., 1996. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. 68, 850–858. https://doi.org/10.1021/AC950914H

Siljander, H.T.A., Simell, S., Hekkala, A., Lähde, J., Simell, T., Vähäsalo, P., Veijola, R., Ilonen, J., Simell, O., Knip, M., 2009. Predictive characteristics of diabetes-associated autoantibodies among children with HLA-conferred disease susceptibility in the general population. Diabetes 58, 2835–2842. https://doi.org/10.2337/db08-1305

Sinha, A., Mann, M., 2020. A beginner's guide to mass spectrometry–based proteomics. Biochem. (Lond). 42, 64–69. https://doi.org/10.1042/BIO20200057

Sioofy-Khojine, A.-B., Lehtonen, J., Nurminen, N., Laitinen, O.H., Oikarinen, S., Huhtala, H., Pakkanen, O., Ruokoranta, T., Hankaniemi, M.M., Toppari, J., Vähä-Mäkilä, M., Ilonen, J., Veijola, R., Knip, M., Hyöty, H., 2018. Coxsackievirus B1 infections are associated with the

initiation of insulin-driven autoimmunity that progresses to type 1 diabetes. Diabetologia 61, 1193–1202. https://doi.org/10.1007/s00125-018-4561-y

Smith, L.M., Kelleher, N.L., Linial, M., Goodlett, D., Langridge-Smith, P., Goo, Y.A., Safford, G., Bonilla, L., Kruppa, G., Zubarev, R., Rontree, J., Chamot-Rooke, J., Garavelli, J., Heck, A., Loo, J., Penque, D., Hornshaw, M., Hendrickson, C., Pasa-Tolic, L., Borchers, C., Chan, D., Young, N., Agar, J., Masselon, C., Gross, M., McLafferty, F., Tsybin, Y., Ge, Y., Sanders, I., Langridge, J., Whitelegge, J., Marshall, A., 2013. Proteoform: A single term describing protein complexity. Nat. Methods 10, 186–187. https://doi.org/10.1038/nmeth.2369

Söderström, U., Åman, J., Hjern, A., 2012. Being born in Sweden increases the risk for type 1 diabetes - A study of migration of children to Sweden as a natural experiment. Acta Paediatr. Int. J. Paediatr. 101, 73–77. https://doi.org/10.1111/j.1651-2227.2011.02410.x

Strachan, D.P., 2000. Family site, infection and atopy: The first decade of the "hygiene hypothesis." Thorax 55. https://doi.org/10.1136/thorax.55.suppl_1.s2

Sun, B.B., Chiou, J., Traylor, M., Benner, C., Hsu, Y.H., Richardson, T.G., Surendran, P., Mahajan, A., Robins, C., Vasquez-Grinnell, S.G., Hou, L., Kvikstad, E.M., Burren, O.S., Davitte, J., Ferber, K.L., Gillies, C.E., Hedman, Å.K., Hu, S., Lin, T., Mikkilineni, R., Pendergrass, R.K., Pickering, C., Prins, B., Baird, D., Chen, C.Y., Ward, L.D., Deaton, A.M., Welsh, S., Willis, C.M., Lehner, N., Arnold, M., Wörheide, M.A., Suhre, K., Kastenmüller, G., Sethi, A., Cule, M., Raj, A., Kang, Hyun Ming, Burkitt-Gray, L., Melamud, E., Black, M.H., Fauman, E.B., Howson, J.M.M., Kang, Hyun Min, McCarthy, M.I., Nioi, P., Petrovski, S., Scott, R.A., Smith, E.N., Szalma, S., Waterworth, D.M., Mitnaul, L.J., Szustakowski, J.D., Gibson, B.W., Miller, M.R., Whelan, C.D., 2023. Plasma proteomic associations with genetics and health in the UK Biobank. Nat. 2023 6227982 622, 329–338. https://doi.org/10.1038/s41586-023-06592-6

Switzar, L., Giera, M., Niessen, W.M.A., 2013. Protein digestion: An overview of the available techniques and recent developments. J. Proteome Res. 12, 1067–1077. https://doi.org/10.1021/pr301201x

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., Matsuo, T., 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. Rapid Commun. Mass Spectrom. 2, 151–153. https://doi.org/https://doi.org/10.1002/rcm.1290020802

Tirumalai, R.S., Chan, K.C., Prieto, D.A., Issaq, H.J., Conrads, T.P., Veenstra, T.D., 2003. Characterization of the Low Molecular Weight Human Serum Proteome*S. Mol. Cell. Proteomics 2, 1096–1103. https://doi.org/https://doi.org/10.1074/mcp.M300031-MCP200

Tuthill, T.J., Groppelli, E., Hogle, J.M., Rowlands, D.J., 2010. Picornaviruses 43–89. https://doi.org/10.1007/82_2010_37

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., Cox, J., 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 13, 731–740.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C.A.-K., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., Pontén, F., 2015. Proteomics. Tissue-based map of the human proteome. Science 347, 1260419. https://doi.org/10.1126/science.1260419

van Bentum, M., Selbach, M., 2021. An Introduction to Advanced Targeted Acquisition Methods. Mol. Cell. proteomics 20, 100165. https://doi.org/10.1016/j.mcpro.2021.100165

van der Linden, L., Wolthers, K.C., van Kuppeveld, F.J.M., 2015. Replication and inhibitors of enteroviruses and parechoviruses. Viruses 7, 4529–4562. https://doi.org/10.3390/v7082832

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P.,

Kleywegt, G., Birney, E., Hassabis, D., Velankar, S., 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 50, D439–D444. https://doi.org/10.1093/NAR/GKAB1061

Varadi, M., Bertoni, D., Magana, P., Paramval, U., Pidruchna, I., Radhakrishnan, M., Tsenkov, M., Nair, S., Mirdita, M., Yeo, J., Kovalevskiy, O., Tunyasuvunakool, K., Laydon, A., Žídek, A., Tomlinson, H., Hariharan, D., Abrahamson, J., Green, T., Jumper, J., Birney, E., Steinegger, M., Hassabis, D., Velankar, S., 2024. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. Nucleic Acids Res. 52, D368–D375. https://doi.org/10.1093/NAR/GKAD1011

Vatanen, T., Franzosa, E.A., Schwager, R., Tripathi, S., Arthur, T.D., Vehik, K., Lernmark, Å., Hagopian, W.A., Rewers, M.J., She, J.X., Toppari, J., Ziegler, A.G., Akolkar, B., Krischer, J.P., Stewart, C.J., Ajami, N.J., Petrosino, J.F., Gevers, D., Lähdesmäki, H., Vlamakis, H., Huttenhower, C., Xavier, R.J., 2018. The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. Nature 562, 589–594. https://doi.org/10.1038/s41586-018-0620-2

Viode, A., van Zalm, P., Smolen, K.K., Fatou, B., Stevenson, D., Jha, M., Levy, O., Steen, J., Steen, H., 2023. A simple, time- and cost-effective, high-throughput depletion strategy for deep plasma proteomics. Sci. Adv. 9. https://doi.org/10.1126/SCIADV.ADF9717

von Toerne, C., Laimighofer, M., Achenbach, P., Beyerlein, A., de Las Heras Gala, T., Krumsiek, J., Theis, F.J., Ziegler, A.G., Hauck, S.M., 2017. Peptide serum markers in islet autoantibody-positive children. Diabetologia 60, 287–295.

Wang, K., Ye, F., Chen, Y., Xu, J., Zhao, Y., Wang, Y., Lan, T., 2021. Association Between Enterovirus Infection and Type 1 Diabetes Risk: A Meta-Analysis of 38 Case-Control Studies. Front. Endocrinol. (Lausanne). 12, 706964. https://doi.org/10.3389/fendo.2021.706964

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., Williams, K.L., 1996. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. Biotechnol. Genet. Eng. Rev. 13, 19–50. https://doi.org/10.1080/02648725.1996.10647923

Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., 2009. Universal sample preparation method for proteome analysis. Nat. Methods 6, 359–362. https://doi.org/10.1038/nmeth.1322

Wu, C.C., Tsantilas, K.A., Park, J., Plubell, D., Sanders, J.A., Naicker, P., Govender, I., Buthelezi, S., Stoychev, S., Jordaan, J., Merrihew, G., Huang, E., Parker, E.D., Riffle, M., Hoofnagle, A.N., Noble, W.S., Poston, K.L., Montine, T.J., MacCoss, M.J., 2024. Mag-Net: Rapid enrichment of membrane-bound particles enables high coverage quantitative analysis of the plasma proteome. bioRxiv 2023.06.10.544439. https://doi.org/10.1101/2023.06.10.544439

Ylipaasto, P., Klingel, K., Lindberg, A.M., Otonkoski, T., Kandolf, R., Hovi, T., Roivainen, M., 2004. Enterovirus infection in human pancreatic islet cells, islet tropism in vivo and receptor involvement in cultured islet beta cells. Diabetologia 47, 225–239. https://doi.org/10.1007/s00125-003-1297-z

Zhang, M., Wang, H., Tang, J., He, Y., Xiong, T., Li, W., Qu, Y., Mu, D., 2021. Clinical characteristics of severe neonatal enterovirus infection: a systematic review. BMC Pediatr. 21. https://doi.org/10.1186/s12887-021-02599-y

Zhang, Q., Fillmore, T.L., Schepmoes, A.A., Clauss, T.R.W., Gritsenko, M.A., Mueller, P.W., Rewers, M., Atkinson, M.A., Smith, R.D., Metz, T.O., 2013. Serum proteomics reveals systemic dysregulation of innate immunity in type 1 diabetes. J. Exp. Med. 210, 191–203.

Zhi, W., Sharma, A., Purohit, S., Miller, E., Bode, B., Anderson, S.W., Reed, J.C., Steed, R.D., Steed, L., Hopkins, D., She, J.-X., 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 MCP 10, M111.012203.

Ziegler, A.G., Hummel, M., Schenker, M., Bonifacio, E., 1999. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: The 2-year analysis of the German BABYDIAB Study. Diabetes 48, 460–468. https://doi.org/10.2337/diabetes.48.3.460

Ziegler, A.G., Rewers, M., Simell, O., Simell, T., Lempainen, J., Steck, A., Winkler, C., Ilonen, J., Veijola, R., Knip, M., Bonifacio, E., Eisenbarth, G.S., 2013. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. JAMA 309, 2473–2479. https://doi.org/10.1001/jama.2013.6285

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2025

ISBN 978-952-02-0136-4 (PRINT)
ISBN 978-952-02-0137-1 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)