uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Karen Powell

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biochemistry)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Identification of Type 1 Diabetes-related Proteins

TITRE DE LA THÈSE / TITLE OF THESIS

Fraser Scott

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Bill Cameron

Mary-Ellen Harper

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Identification of type 1 diabetes-related proteins

Karen L. Powell

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.Sc. degree in Biochemistry

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-25821-7
Our file Notre référence
ISBN: 978-0-494-25821-7

ABSTRACT

Type 1 diabetes (T1D) is a T-cell mediated autoimmune disease, whereby the insulin-secreting β-cells of the pancreas are targeted and destroyed by the patient's immune system. Although genetic susceptibility is required for the onset of this autoimmune disorder, environmental factors, including infectious agents and diet, also play an important but poorly understood role. Recently, WP5212 was identified as the first candidate diabetes-related wheat protein. To identify candidate molecular mimics of this wheat protein in target tissues, a novel approach was developed. Antibodies raised against two short peptides of this dietary antigen cross-react with self proteins: an exocrine pancreatic 33 kDa protein and a 28 kDa protein in the islets, spleen and mesenteric lymph nodes. In order to identify additional putative targets, a broad spectrum polyclonal serum was generated. In addition, a second antigenic wheat protein was identified by screening a wheat cDNA expression library with antibodies from a highly wheat-sensitive patient with both T1D and celiac disease. Further characterization of candidate molecular targets and antigenic wheat proteins may reveal novel pathways related to environmentally-induced T1D.

ACKNOWLEDGEMENTS

I would like to thank Dr. Fraser Scott for providing me with the opportunity to complete my M.Sc. in his laboratory. I would like to express my appreciation to Dr. F. Scott, Dr. Alexander Strom and Dr. Amanda MacFarlane for their mentorship and guidance throughout my research. I would also like to thank Charles Gyamera and Dr. M. Mbikay for their collaboration on the generation of the high-titer polyclonal anti-sera. I would like to express my gratitude to Dr. S. Bennett, Dr. A. Kumar and Dr. M. Mbikay, my thesis advisory committee, for their valuable advice. Funding for this work was granted to Dr. Scott by the Juvenile Diabetes Research Foundation and the Canadian Institutes of Health Research. I am grateful to the Ontario Ministry of Training, Colleges and Universities for granting me an Ontario Graduate Scholarship and to the University of Ottawa for awarding me scholarships.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AP: alkaline phosphatase
APC: antigen presenting cells
APS: ammonium persulfate
BBdp: BioBreeding diabetes-prone
BCIP: 5-bromo-4-chloro-3-indolyl phosphate
BLAST: basic local alignment search tool
BSA: bovine serum albumin
CD: Celiac disease
cDNA: complement deoxyribonucleic acid
CTLA-4: cytotoxic T lymphocyte associated 4
CVB: coxsackie B virus
DAISY: Diabetes AutoImmunity Study in the Young
DC: dendritic cell
DNA: deoxyribonucleic acid
DIPP: Diabetes Prediction and Prevention Study
DiMe: Childhood diabetes in Finland study group
DTT: dithiothreitol
EDTA: ethylene diamine tetraacetic acid
ELISA: enzyme-linked immunosorbent assay
EEV: eukaryotic expression vector
EV: enterovirus
FBS: fetal bovine serum
GAD65: glutamic acid decarboxylase 65
GALT: gut-associated lymphoid tissue
GbssI: granule-bound starch synthase I
Glb1: storage globulin 1
HC: hydrolyzed casein
HEK: human embryonic kidney
HLA: human leukocyte antigen
HMW: high molecular weight
HRP: horse radish peroxidase
IA-2: insulinoma-associated protein 2
IEF: isoelectric focusing
IFN-γ: interferon-γ
Ig: immunologlobulin
IL-2: interleukin-2
IPTG: isopropyl β-thiogalactopyranoside
LB: luria broth
MALDI-TOF MS: matrix assisted laser desorption/ionization time of flight mass spectrometry
MHC: major histocompatibility complex
MLNs: mesenteric lymph nodes
NBT: nitroblue tetrazolium
NK: natural killer
NOD: non-obese diabetes
NZY: NZ amine
OD: optical density
ORF: open reading frame
PBS: phosphate buffered saline
PBS-T: phosphate buffered saline + Tween
pfu: plaque forming unit
PMSF: phenylmethylsulfonyl fluoride
PVDF: polyvinylidene fluoride
SD: standard deviation
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sf21: *Spodoptera frugiperda*
T1D: type 1 diabetes
TBS: tris buffered saline
TBS-T: tris buffered saline + Tween
TCR: T cell receptor
TEMED: N,N,N',N'-tetramethylethylenediamine
TRIGR: Trial to Reduce Insulin-Dependent Diabetes in the Genetically at Risk
tTG: tissue transglutaminase
VNTR: variable number tandem repeats
WP: wheat protein
WPran: randomly selected wheat cDNA clone
WP5012: Wheat cDNA clone 5012
WP5021: Wheat cDNA clone 5021
WP5212: Wheat cDNA clone 5212
WP32112: Wheat cDNA clone 32112
X-gal: 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Autoimmunity

Autoimmune disease affects between three and five percent of the general population (1). To date more than forty different autoimmune diseases are known, the most common include thyroditis, rheumatoid arthritis, Graves disease, celiac disease, type 1 diabetes, multiple sclerosis, and systemic lupus (1). Autoimmunity is antigen-specific and occurs when the host's immune system mistakenly attacks self-proteins, damaging a specific organ or acting systemically. The immune system possesses several mechanisms to delete or silence self-reactive lymphocytes, including negative selection and peripheral tolerance. Normally, immature T and B cells whose receptors interact strongly with self proteins receive death signals and are consequently deleted from the lymphocyte repertoire in the thymus and bone marrow, respectively (2). Tolerance also extends into the periphery whereby naïve T and B cells which encounter antigens in the absence of the required co-stimulation become inactivated and non-responsive, a state known as anergy (2). A defect in any one of these mechanisms may lead to a breakdown in tolerance to self proteins and result in the development of autoimmunity. It is believed that autoimmunity is triggered by a single antigen and, as the disease progresses, the immune response expands and targets additional antigens/epitopes by determinant spreading (1). Depending on the disease, either T cells or autoantibodies can mediate the damage to the tissues. Furthermore, autoantibodies may serve as immunological markers of disease progression in T cell-mediated autoimmunity.

Both genetic susceptibility and environmental insults are believed to modulate the development of autoimmunity by affecting the "overall reactivity of the immune system, the specific antigen and its presentation, and the target tissue" (1). It is hypothesized that autoimmunity generally arises from an abnormal immune reaction against antigens normally present in the environment. However, the environmental trigger(s) of a chronic dysregulated immune reaction are extremely difficult to elucidate due to a lag time of anywhere from days or months to years between the triggering environmental insults and the development of autoimmunity. Furthermore, the timing and duration of antigen exposure may be important. A number of different approaches has been developed in the attempt to uncover the underlying antigens in autoimmune disease(s). Epidemiological studies are commonly used to link environmental insults and autoimmunity; however, these studies do not ascertain causality. Case studies of unique individuals can provide invaluable insight and new direction. Antibodies isolated from individuals with autoimmune diseases are frequently used to identify autoantigens present in the target tissues. Animal models of autoimmunity are valuable tools to investigate the effect of suspected environmental triggers. The identification of triggers and/or promoters involved in the development of autoimmunity is crucial to the understanding of pathology in autoimmune disease.

Type 1 diabetes

Over two million Canadians are affected by diabetes, and 10% of these individuals have type I diabetes (T1D; www.diabetes.ca), the most severe form of diabetes, which occurs mainly in children and young adults. T1D is a chronic autoimmune disease which targets and destroys the insulin producing β-cells of the islets of Langerhans in the pancreas. Full-blown diabetes

occurs when the majority (80-90%) of the β-cell mass has been destroyed (3) resulting in insufficient insulin production and secretion to control blood-glucose levels. T1D patients are required to adopt a strict routine to monitor their blood-glucose levels and, in order to prevent ketosis, coma and death, must inject themselves daily with insulin. Most patients will experience severe complications which reduce both their quality of life and life expectancy, including heart disease, kidney disease, eye disease, nerve damage, erectile dysfunction and amputations (www.diabetes.ca).

The pancreas is located deep in the abdomen and is susceptible to damage during biopsy. The difficulty of studying the pancreas in humans necessitates reliance on various animal models to study the characteristics of spontaneous autoimmune diabetes. The two most extensively studied animal models are the non-obese diabetic (NOD) mouse and the diabetes-prone BioBreeding (BBdp) rat, both of which display clinical characteristics which resemble human diabetes. These animal models exhibit signs of insulitis prior to the development of diabetes which occurs around adolescence in the rat and young adulthood in the mouse. Islet-specific autoantibodies can also be detected. It is also important to consider that while these animal models are mostly inbred, humans represent a far wider spectrum of genetic variation. While these models have been pivotal in the advancement of our knowledge of diabetes and are crucial to explore potential treatments, it is important to acknowledge that different gene-environment interactions may be responsible for the initiation or promotion of autoimmune diabetes in various species. For example, to date there are more than 192 ways to prevent diabetes in the NOD mouse (4), but how these relate to diabetes in human remains to be determined.

T1D Pathogenesis

The pre-clinical phase of diabetes is characterized by an asymptomatic period during which progressive β-cell destruction and the development of autoantibodies occur. In this pre-diabetic phase mononuclear cells infiltrate the islets of Langerhans, a process known as insulitis. Macrophages and dendritic cells are the first to infiltrate the islet, followed by T and B lymphocytes and nature killer cells (5). Several complementary studies in the NOD mouse, including adoptive transfer experiments and analyses using T cell inhibitors, demonstrate that T1D is a T cell-mediated disease in this animal model (6). CD4+ and CD8+ T cells have been shown to play an important role in the progression of T1D in the NOD mouse by acting as recruiters and final effectors, respectively (5). Evidence also suggests that human T1D is T cell-mediated (6).

While the cause of the initiation of the immune attack remains to be ascertained, evidence from animal models supports the theory that once specific β-cell autoimmunity has developed, a vicious, irreversible cycle ensues. It is thought that the naive β-cell autoreactive T cells do not initially encounter the antigen at the target tissue, but rather at a distant location (7). Antigen presenting cells (APCs), comprising macrophages, dendritic cells and B cells, take up and present self-antigens to naive β-cell autoreactive CD4+ T cells in the peripheral lymph nodes (5). These self-reactive T cells must have escaped the negative selection and peripheral tolerance in the immune system (8). Once activated, these CD4+ T cells expand and migrate to islets where further activation occurs as they encounter additional β-cell autoantigens. This stimulates the secretion of Th1 cytokines, including interleukin (IL)-2 and interferon (IFN)-γ, which in turn recruit and activate CD8+ cytotoxic

T cells and more macrophages (5). The vicious cycle continues as these additional cytokines are released, activating new immune cells.

The actual mechanisms by which β-cell death is induced are not clearly understood. Animal models provide evidence that apoptosis plays an important role. It is probable that the destruction of β-cells is a consequence of a combination of all aspects of the inflammatory response (8). Although controversial, results from NOD mouse experiments indicate that β-cells can be killed directly by cytotoxic T cells upon cell-cell contact through the Fas/FasL or perforin/granzyme pathways (6). Soluble mediators, including pro-inflammatory cytokines, free radicals and reactive oxygen intermediates, may be directly cytotoxic to β-cells, an aspect that is still under debate (6). Macrophages and even β-cells themselves, through the release of soluble death mediators, may be involved in the destruction of the β-cells (6). However, how these death effector systems translate from observations in animal models to humans remains to be determined.

Autoantibodies – markers of disease progression

Autoantibodies can be detected in the blood during the long pre-diabetic phase prior to the appearance of clinical symptoms (9). The activation of autoreactive B cells and repeated antigen exposure result in the production of these autoantibodies which undergo subsequent isotype switching from immunoglobin (Ig) M to IgG. While the autoantibodies are not believed be pathogenic, they reflect disease progression and can serve as immunological markers to predict development of T1D in at-risk individuals (9).

The three best-characterized autoantibodies in the prediction of disease risk are those against insulin, 65 kDa glutamic acid decarboxylase (GAD65) and insulinoma-associated protein 2 (IA-2). Insulin was the first and only β-cell specific autoantigen identified in T1D (10). Up to 70% of T1D patients have insulin autoantibodies (10) and these are often the earliest autoantibodies detected during infancy (11). While GAD65 and IA-2 are not β-cell specific, as many as 70% and 66% of diabetics respectively possess antibodies against these proteins (10). The appearance of several of these autoantibodies has been shown to indicate an increased risk of disease progression. Sera from diabetic patients may possess antibody reactivity against additional autoantigens, including heat shock protein-60, islet cell autoantigen of 69 kDa, and carboxyl peptidase H (10).

The activation of B cells, and the subsequent isotype switching of their immunoglobulins to the secreted IgG form, requires stimulatory signals from T cells whose receptors bind a peptide from the same protein or protein complex as the B cell epitope (2). Consequently, the presence of autoantibodies in the serum of diabetic patients has led to the search for T cell proliferation responses against the candidate autoantigens. Therefore, the identification of autoantibodies is an important step in the discovery of potential initiating T cell antigens of T1D. Recent studies indicate that insulin may be a crucial target of the autoimmune response and insulin-reactive T cells have been detected in some human patients (12, 13). The question of which autoantigens are initial targets and which ones arise from determinant spreading remains to be answered. It is possible that islet autoimmunity is activated against a lone autoantigen (3) and the immune attack then becomes diversified, as shown by the

appearance of multiple autoantibodies (14) with the release of sequestered antigens following the initial destruction of the target.

A genetic predisposition is required to develop T1D [1]

The genetics of T1D is extremely complex; it is a polygenic disease with over 20 loci believed to be associated with increased risk (15). The genes for major histocompatibility complex (MHC) class II region, known as human leukocyte antigen (HLA) in humans, are implicated in most autoimmune diseases. These genes code for proteins expressed on the surface of antigen presenting cells (APCs) which are responsible for presenting antigens to T cells. In diabetes, HLA genes account for up to ~40% of the genetic risk and include HLA-DQ and HLA-DR, and more specifically HLA-DR3 and -DR4 (15). Non-HLA loci associated with this disease impart only modest susceptibility. Approximately 10% of genetic risk can be attributed to the variable number of tandem repeats (VNTR) region located upstream from the insulin gene (15). This region is responsible for insulin expression in the thymus, and lower expression levels may be linked to less efficient negative selection of autoreactive lymphocytes (15). The insulin gene is further implicated in T1D due to an association of *INS-23* HphI polymorphisms with the presence of diabetes-related autoantibodies (16). The cytotoxic T lymphocyte associated 4 (CTLA-4) gene is a third locus which has been extensively studied. It has been suggested that low levels of soluble CTLA-4 result in immune responses which are not well controlled (15). Other recently identified diabetes-associated genes include *SUMO-4* (17), *T-bet* (18) and Toll-like receptor (19). Identification of additional candidate risk genes may provide new insight into T1D pathogenesis.

[1] The following six pages reflect sections I contributed to the review article (26)

T1D is influenced by environment

While individuals susceptible to developing T1D carry a genetic predisposition, increasing evidence from epidemiological analyses indicates that environment plays a major role. Firstly, studies of twin pairs reveal incomplete disease penetrance in monozygotic twins, with the concordance rate varying from 18-70% (20) which depends on the population and the age of initial diagnosis. Siblings of a diabetic patient are at a greater risk of developing diabetes than the general population (20). In addition, the annual incidence of T1D has been steadily increasing, for example 2.8% in New South Wales (21) and 4.9% in Franche-Comté, France (22). These rates cannot be attributed to changes in genetic risk alone. Finally, the prevalence of T1D also exhibits dramatic geographic discrepancy, with annual incidence ranging from as low as 0.1 per one hundred thousand in China and Venezula up to ~37-50 per one hundred thousand in Finland and Sardinia (23). The Avalon Peninsula in Newfoundland has the highest annual diabetes incidence observed in Canada, at 36 per one hundred thousand, a rate that is amongst the highest in the world (24). Together these observations implicate environmental factors as key modulators of disease progression.

As previously mentioned, the identification of environmental factors that influence chronic diseases is difficult. Furthermore, the destruction of the β-cells is irreversible and consequently the removal of environmental factors will not likely alleviate symptoms (25). In light of this, several studies are ongoing in both high-risk, first-degree relatives of T1D patients and individuals from the general population, including the Trial to Reduce IDDM in the Genetically at Risk (TRIGR), Diabetes Prediction and Prevention study (DIPP) and Diabetes AutoImmunity Study in the Young (DAISY) (26). These studies aim to determine

whether the onset of diabetes is influenced by exposure to environmental agents such as viruses and foods.

Viruses

Viral infections have been implicated in the development of T1D for more than a century. Enteroviruses (EV) are of particular interest as an early study (27) demonstrated that a group B coxsackievirus (CVB) strain isolated from the pancreas of a diabetic child induced diabetes in a mouse model following infection. Furthermore, in recent studies CVB have been shown to either accelerate or prevent diabetes depending on the levels of two cytokines, IL-2 and IFN-gamma, in the NOD mouse model (28). Human serological evidence examining the presence of antibodies and viral RNA suggests that multiple EV infections may increase the risk of developing T1D (reviewed in (29, 30)). The DiME and DIPP studies indicate that these infections occur immediately prior to the detection of autoantibodies (29), markers of beta-cell autoimmunity. However, other epidemiological studies (26, 31) have found no difference between patient and control groups. It is believed that point mutation affecting the virus' ability to infect certain tissue types may be an important factor in determining whether an EV strain will be diabetogenic (29, 30). Other viral infections which may be implicated in T1D include mumps, congenital rubella syndrome (CRS), human cytomegalovirus (CMV) and reteroviruses (30). Additional viruses potentially associated with this disease continue to be identified. For example, a recent case study of a 66-year old Taiwanese male who received a blood transfusion suggests a link between beta cell autoimmunity and hepatitis C virus infection (32). However, the role that viruses play in the development of T1D remains controversial due to contradictory results

and a lack of evidence that these viral infections directly cause β-cell autoimmunity and T1D in humans.

Diet

Over 25 years ago, surveys conducted by Helgason and his colleagues implicated a common food additive, N-nitroso compounds, in the development of diabetes (33, 34). This provided the first evidence to suggest that diet can modulate onset of autoimmune diabetes. Extensive analyses in animals, investigating micronutrients, carbohydrate, fat, fibre, protein and caloric intake, have shown the amino acid source to be the major factor responsible for the dietary modification of diabetes (35, 36). These studies have identified both diabetes-retardant and diabetes-promoting protein sources. Hydrolyzed casein (HC), composed of peptides of between 2 to 27 amino acids in length, has been shown to prevent the development of diabetes in both NOD mice and diabetes-prone BioBreeding rats and has been adopted as the standard control. Studies have demonstrated that this diet can both inhibit insulitis (37) and promote islet neogenesis (38). Whether HC is protective or lacks diabetes-promoting agents remains under debate. Diabetes-promoting protein sources include, but are not limited to, milk and wheat.

Evidence supporting wheat as a modulator of T1D

It was demonstrated in early studies in the BBdp rat (39-41) that dietary components can influence the onset of diabetes. Analysis of the major components of the standard NIH-07 cereal-based diet determined its major diabetogenic ingredient to be wheat gluten (40). The diabetogenicity of wheat has been confirmed in several studies where animals fed a semi-

purified diet supplemented with wheat protein have been shown to develop diabetes at a significantly higher frequency than animals fed a protective hydrolyzed casein diet (42-44). Wheat proteins are also potent diabetes-promoting dietary agents in NOD mice (44-47). However, contradictory studies in this animal model suggest the addition of wheat gluten to a gluten-free diet does not alter diabetes incidence (48, 49). Consequently, while the diabetes-promoting effect of wheat is well established in the BB rat, it is inconsistent in the NOD mouse. Based on earlier studies of various crude fractions of wheat gluten, the low molecular weight glutenin fraction, which can contain albumin and globulin contaminants (42), is believed to be one candidate protein source with diabetes-promoting activity.

There is mounting evidence that wheat proteins play a role in the pathogenesis of type 1 diabetes in humans. The frequency of this disease is increased in developed countries where refined wheat flour, a major component of which is gluten, is an integral part of the daily diet. Several independent studies have investigated a link between the timing of gluten introduction to infant diet and the risk of developing T1D. The BABYDIAB study provided evidence that children of T1D patients exposed to gluten at less than three months of age have an increased risk of developing islet autoantibodies, unlike babies of T1D patients exposed to gluten after six months of age (50). Similarly, the DAISY study revealed a window for gluten introduction, between ages 4 – 7 months, associated with lower risk of islet autoantibody development (51). These studies suggest that the timing of gluten introduction can affect T1D onset. The BABYDIET study (52) is designed to determine whether the risk of developing islet autoimmunity in susceptible children can be reduced by delaying the introduction of gluten until 12 months.

Furthermore, Klemetti and colleagues (53) have demonstrated that T cells isolated from newly-diagnosed T1D patients proliferate in response to wheat gluten, indicating a primed immune response against wheat. This proliferation has recently been confirmed in a larger subset of diabetic patients (54). While it has been shown that a gluten-free diet does not alter diabetes-related autoantibody titers in diabetic patients, it may restore the loss of first phase insulin secretion and insulin sensitivity (55, 56). Thus, a gluten-free diet does not reverse autoimmunity, but implementing this diet may promote regeneration of islet mass.

WP5212 – a novel candidate diabetes-related wheat protein

Recently, the first diabetes-related wheat protein was identified by screening a wheat cDNA expression library using sera from diabetic BB rats. The nucleotide sequence of clone WP5212 shares 90% homology at the nucleotide level and 80% identity at the amino acid level with the *Triticum aestivum* wheat storage protein, Glb1 protein (42). BBdp rats exhibited significantly higher antibody reactivity against this protein than both asymptomatic and control animals. Furthermore, antibody reactivity correlated with pancreatic insulitis in this animal model of spontaneous diabetes (42). A recent case study of a highly wheat-sensitive individual with both T1D and CD illustrates that this patient possesses both antibody reactivity and T cell proliferation against WP5212 (57). This is an encouraging development, demonstrating that this wheat protein, first discovered using an animal model, can be antigenic in humans. Further analysis of a sample of diabetic patients indicates that a subset of these patients has increased antibody reactivity (58) and/or T cell proliferation (54) against this protein. This evidence strongly suggests that WP5212, a contaminant of wheat gluten, may be related to diabetes in humans.

A link between T1D and celiac disease

Wheat has already been established as a trigger of an autoimmune-based disease. Celiac disease (CD), a gluten-sensitive enteropathy, is a chronic inflammatory disease in which ingestion of gluten results in severe damage to the gut epithelium. The suggested link between wheat and T1D is further strengthened by the considerably higher incidence of celiac disease within the diabetic population (1-8%) than in the general population (0.4-1.0%) (59). A majority of patients with both diseases develop CD first, suggesting the standard treatment for CD, a gluten-free diet, might prevent the onset of diabetes in some individuals. Furthermore, a number of newly-diagnosed diabetic children (30%) have been found to possess the serological markers of CD (60, 61), and children of diabetic patients may also develop these antibodies (62). It is also known that both T1D and CD share the HLA haplotype DQ2-DR3. Together these factors support the association between wheat and T1D.

Implications of a role for gut dysfunction in autoimmune diseases

The gut epithelium is a single layer of cells which separates the internal and external environments. It is at this interface where environmental factors, including diet and viruses, are sampled by the immune system (63). The gut acts as a gatekeeper, regulating these interactions through tight junctions which control gut permeability by synchronizing rapid responses triggered by food components and viruses (64). The gut-associated lymphoid tissue (GALT) is responsible for maintaining the delicate balance between mucosal tolerance to food antigens and commensal bacteria and immunity against pathogens (65). An

abnormal immune response to non-pathogenic foreign molecules exposed to the immune system via the gastrointestinal tract can lead to disease.

Celiac disease and Crohn's disease are two examples of immune dysregulation diseases mediated through the gastrointestinal tract. Celiac disease results from an immune intolerance to ingested gluten which contains celiac-toxic peptides (66). As these gliadin peptides cross the gut epithelium, they are deamidated by tissue transglutaminase. APCs take up these deamidated peptides and present them on HLA-DQ2 and HLA-DQ8 to CD4+ T cells (67). In susceptible individuals, this leads to the activation of a series of events which results in intestinal epithelium damage, including mucosal inflammation, destruction of villi, crypt hyperplasia (68). Similarly, the gut inflammation and damage which occur in Crohn's disease are believed to be a result of an abnormal immune reaction in response to the commensal bacteria (69). In both these diseases, the gut damage affects permeability, increasing the interactions between the immune system and the triggering components of the lumen, and intensifying the vicious immune reaction.

The gastrointestinal tract is associated with two major immune dysregulation diseases and evidence is emerging that it may also play a role in T1D in both animal models and humans. The characterization of celiac disease and Crohn's disease may lead to the development of an important model of gut-mediated autoimmunity. The techniques and strategies used to unravel the intricate gene-environment relationships of these diseases may be employed to begin to understand the complexities of other autoimmune diseases which appear to be modulated through the gastrointestinal tract.

Identification of antigenic molecular structures

The identification of key environmental factors, let alone their antigenic molecular structures, is exceedingly difficult in chronic diseases. A common theory which attempts to explain the association between environmental insults and the pathogenesis of autoimmune disease is molecular mimicry. It is believed to occur when a lymphocyte fails to discriminate between a foreign molecule and a self protein which shares sequence or structural similarity. As a result, once the immune system is primed to attack this foreign protein, it can also recognize and mistakenly target the self protein. Therefore, in order to explore the potential involvement of molecular mimicry in the pathogenesis of diabetes, the characterization of the antigenic structures of environmental factors is crucial.

To facilitate the identification of the antigenic structures present in candidate environmental factors, it is of interest to examine techniques which have been developed to identify food allergens. These techniques exploit the generation of immunoglobulin E (IgE) by the immune system in response to the corresponding dietary allergen. The most common method to detect immunogenic proteins in food is to separate them by two-dimensional gel electrophoresis and probe a replicate membrane with IgE in serum from patients with the food allergy of interest. Proteins of interest are identified by mass spectrometry analysis or N-terminal sequencing (70). This technique has been used recently to identify novel isoforms of the major cherry allergen, Pru av 1 (71). A variation of this technique is the use of non-denaturing conditions, which allows for the detection of conformational epitopes within the native structure of allergens (72). A second tool used for the identification of allergens is cDNA expression libraries. These libraries are created by cloning cDNA,

synthesized using RNA extracted from the food of interest as a template, into bacteriophage. IgE in serum of patients with food allergy is used to screen the library and identify positive clones. The nucleotide sequence of the insert is analyzed and the corresponding amino acid sequence can be deduced. This method has been used to identify several food allergens, including a major oyster allergen (73).

Despite the fact that T1D is recognized as a T cell mediated disease, antibodies present in the serum of diabetic individuals can reveal crucial information regarding molecular structures involved in this autoimmune disease. Recent studies have demonstrated that the techniques used to identify food allergens can be adapted to identify candidate immunogenic dietary wheat proteins related to T1D (42). Furthermore, by screening pancreatic proteins with serum from diabetic individuals allowed for the identification of key autoantigens implicated in T1D, including insulin, GAD65 and IA-2 (9). This suggests that it is possible to use similar approaches to identify novel T1D-related antigenic molecular structures associated with wheat.

Although T1D research has been ongoing for decades, the trigger that causes the immune system to mistakenly attack the β-cells remains unknown and the details surrounding the death of the β-cells are unclear. The search continues for antigens which may initiate the process; however, due to the complexity of this disease which involves several genes, multiple organ systems, and numerous external agents, a considerable number of hurdles remain to be overcome. The aim of this research was to identify candidate antigenic wheat

proteins and to explore the potential mechanisms by which wheat proteins may modulate the development of diabetes.

HYPOTHESIS

The hypothesis of this research is that an abnormal immune reaction against specific wheat proteins, including the recently identified WP5212, may be associated with the development of T1D. If this is the case, antibodies produced by this immune reaction may be used to discover additional diabetes-related wheat proteins. If these wheat proteins act by molecular mimicry, antibodies raised against these wheat proteins could lead to the identification of key molecules in target tissues involved in wheat-related T1D.

RESEARCH QUESTIONS

Research Question 1: What are the immunodominant epitopes of the first identified diabetes-related wheat protein, WP5212, a homologue of the wheat storage protein Glb1?

Aim and Approach: The aim was to identify the antigenic peptides of WP5212 and determine which of these peptides are immunodominant epitopes. A peptide library consisting of 64 overlapping peptides spanning the entire WP5212 amino acid sequence was screened using sera from WP5212 antibody-reactive diabetic patients. Antigenic peptides of interest were screened using sera from control subjects, lacking anti-WP5212 antibodies, to determine which peptides were related to T1D.

Rationale: Recently, increased antibody reactivity against WP5212 has been associated with diabetes in two animal models and in a large subset of patients. However, how this protein is involved in the pathogenesis of diabetes remains to be ascertained. The identification and

analysis of immunodominant epitopes may help expose mechanisms through which dietary antigens are involved/participate in the development of diabetes.

Research Question 2: Can antibodies raised against WP5212 be used to develop a methodology to identify candidate molecular targets?

Aim and Approach: The aim was to develop an approach to identify candidate molecular mimics that may play a role in wheat-mediated diabetes pathogenesis. Antibodies were previously raised against two antigenic peptides of WP5212. These antibodies were used to screen proteins isolated from target tissues of the diabetes-prone BioBreeding rat. Candidate proteins were subsequently isolated by two dimensional gel electrophoresis and identified by mass spectrometry analysis and/or N-terminal sequencing. Proteins of interest were expressed in *E. coli* and antibody cross-reactivity was verified by western blot analysis. In addition, polyclonal anti-WP5212 sera were generated by DNA immunization to include antibodies which recognize new linear and conformational epitopes.

Rationale: One hypothesized mechanism responsible for triggering and/or promoting the development of T1D is molecular mimicry. The development of a methodology which can be used to identify potential candidate molecular mimics is crucial to explore this hypothesis. The identification of proteins in target tissues which share sequence or structural similarities with diabetes-related environmental insults may help elucidate the role(s) these proteins play in the pathogenesis of T1D.

Research Question 3: Are there additional candidate wheat proteins related to type 1 diabetes?

Aim and Approach: The aim was to identify additional diabetes-related wheat proteins. This was accomplished by immunoscreening a wheat cDNA library with immunoglobulin G (IgG) serum from an extremely wheat-sensitive subject with both T1D and CD. To determine whether the identified proteins are related to either T1D or CD, recombinant proteins were screened with sera from patients with T1D, CD or both diseases and antibody reactivity against these candidate proteins were compared to healthy controls.

Rationale: Increasing evidence indicates that wheat can modify the development of diabetes in animal models. Furthermore, numerous family and case control studies and preliminary data from prospective natural history trials (reviewed in (26)) also implicate wheat as a modulator of T1D in humans. Recently WP5212 has been identified as a candidate diabetes-related protein, by screening a wheat cDNA expression library with antibodies from diabetic BioBreeding rats (42). This demonstrates proof of concept that antibodies from diabetic individuals can be used to discover diabetes-related wheat proteins. The identification of additional putative diabetes-related wheat proteins may provide insight into the mechanism(s) through which wheat affects the development of T1D.

MATERIALS AND METHODS

Patients

Case study subject. The antibody reactivity to wheat proteins of a highly wheat-sensitive 28 year old Caucasian female with both type 1 diabetes and celiac disease was studied. The patient carries the HLA haplotypes DRB1*03/*04 and DQB1*0201/*0302 (DQ2/DQ*), which are associated with T1D and CD.

Patient study groups. Caucasian patients with either type 1 diabetes, celiac disease or both diseases and healthy control subjects, with no significant medical history, were recruited for the study in agreement with the ethical guidelines established by the Ottawa Hospital Research Ethics Board and the Children's Hospital of Eastern Ontario Ethics Board. For Study 1, the age of diabetic patients (n = 8) and control subjects (n = 8) ranged from 8-41 and 21-29 years, respectively. A different set of patients was used for Study 5. Diabetic patients (n = 9) ranged from 19-34 years of age, celiac patients (n = 5) ranged from 25-33 years of age and patients with both diseases (n = 6) ranged from 12-43 years of age. Healthy control subjects (n = 9) had an age range of 22-28 years.

Samples. Blood samples were obtained with informed consent from the subjects for the isolation of serum/plasma. Serum was tested for IgA antibodies to tissue transglutaminase, the autoantigen associated with celiac disease.

Tissue preparation

Tissues (kidney, spleen, heart, mesenteric lymph nodes (MLNs), liver, pancreas and gut) were harvested from BBc (control) rats and homogenized for 25 s in HSN buffer (refer to **Appendix 1** for the composition of buffers and media). Following a 15 min incubation on ice, samples were centrifuged at 19000 x*g*. Supernatants were transferred to fresh tubes and pellets were resuspended in HSN buffer. Protein concentrations were quantified by spectroscopy, using a bovine serum albumin (BSA) standard curve. Islets were isolated by Dr. L.M. Kauri following the protocol previously described in (74).

Sucrose density gradient centrifugation

Proteins from the supernatant fraction of the spleen and pancreas were further separated by sucrose gradient centrifugation. Briefly, a continuous 20-60% sucrose gradient in sucrose gradient buffer was formed using a variable flow mini-pump (VWR) and a gradient former (Bio-Rad). The 20% and 60% sucrose solutions were placed in the mixer and reservoir chambers, respectively. The solution was transferred from the bottom of the gradient former into ultra-clear centrifuge tubes (14 x 89 mm; Beckman). Following an incubation at 4°C overnight, the gradient was overlaid with the cystolic fraction from the tissue and centrifuged at 150000 x*g* (SW 41 Ti, Beckman Coulter) (4°C) for ~16 hrs. The bottom of the tube was pierced and 0.5-1.0 ml fractions were collected.

Immunoscreening of the wheat cDNA expression library

Construction of wheat ZAP Express Custom cDNA library. Total RNA was isolated from the hard red spring wheat by Dr. Karolina Burghardt following the method previously described in (75). The ZAP Express Custom cDNA library was constructed at Stratagene by inserting wheat cDNA into the *Eco*RI/*Xho*I cloning site of pBK-CMV, the ZAP Express vector (Stratagene).

Titering the library. *Escherichia coli* XL1-Blue MRF' cultures were grown in LB supplemented with 0.2% [w/v] maltose and 10 mM $MgSO_4$ to late log phase (~6-7 hrs). Cells were harvested by centrifugation at ~1000 x*g* for 12 min and resuspended in 10 mM $MgSO_4$ to an OD_{600} ~0.5. 600 µl of the culture was infected with 1 µl of serial dilutions (10^{-1} to 10^{-5}) of wheat ZAP Express Custom cDNA phage library (Stratagene) and incubated at 37°C for 15 min to enable the phages to infect the bacteria. The infected bacteria were combined with 8 ml of pre-warmed (~48°C) NZY top agar including 2 mg X-gal, and the mixture was poured over warm NZY agar plates (150 X 50 mm). Plates were inverted and incubated at 37°C overnight. Plaque counts were performed the following day on plates which contained between 30 and 300 plaque forming units (pfu). White plaques indicated the presence of an insert, while blue plaques denoted empty vectors.

Serum preparation. A blood sample from the case study subject, with both type 1 diabetes and celiac disease, was obtained with informed consent. Serum was pre-absorbed with *E. coli* cell lysate to bind antibodies recognizing bacterial peptides. Briefly, 100 ml of *E. coli* culture, grown to saturation in LB medium, was harvested by centrifugation. Cells were

resuspended in 3 ml of 50 mM Tris-Cl, 10 mM EDTA (pH 8.0). The suspension was frozen and thawed four times and sonicated six times for 20 s each time. The extract was centrifuged at 12000 x*g* for 10 min and the supernatant was collected. Serum was diluted 1/10 in 5% [w/v] skim milk powder in TBS-T. An equal volume of *E. coli* cell lysate was added to the diluted serum and incubated overnight with end-over-end shaking at 4°C. The pre-absorbed serum was further diluted to 1/800 and 0.01% sodium azide was added.

Primary screening of the wheat cDNA phage library. E. coli XL1-Blue MRF' cultures were prepared as described above. 600 µl of E. coli cells were infected with 3.5×10^4 phage from the wheat ZAP Express Custom cDNA library (Stratagene) and incubated at 37°C for 15 min. Protein expression was induced by the addition of 15 µl of 2 M isopropyl-1-thio-β-D-galactopyranoside (IPTG) to the culture, which was then incubated for an additional 15 min at 37°C. The infected bacteria were combined with 8 ml of pre-warmed (~48°C) NZY top agar and the solution was poured over warm NZY agar plates (150 X 50 mm). The plates were inverted and incubated at 37°C overnight.

Nitrocellulose membranes were placed over the agar surface. Following a three-hour incubation at 37°C and a one-hour incubation at 4°C, the nitrocellulose membranes were peeled off the plate, washed 5x 5 min in TBS-T and blocked overnight at 4°C in 5% skim milk powder in TBS-T to block free binding sites on the membranes. Membranes were incubated with the above-mentioned serum for 1 hr at room temperature and washed 5x 5 min in TBS-T. Following a one-hour incubation in the secondary antibody, alkaline phosphatase-conjugated goat anti-human IgG (Sigma; diluted 1/20000 in 5% skim milk in

TBS-T blocking solution), the membranes were washed 3x 5 min in TBS-T and 2x 5 min in TBS. Antibody binding was detected using AP developer solution. Dark purple plaques indicated putative positive clones, which were subsequently cored from the agar, placed in 500 µl of SM buffer containing 20 µl of chloroform and stored at 4°C.

Secondary and tertiary screening of the wheat cDNA phage library. All putative positive clones from the primary screening were screened again to ensure positive antibody reactivity to the expressed protein and to achieve phage clonality. 1 µl of the cored phage stock was incubated with 200 µl of XL1-Blue-MRF' *E. coli* for 15 min at 37°C. 3.75 µl of 2 M IPTG were added to the sample, which was then incubated for another 15 min at 37°C. The mixture was added to 2 ml of pre-warmed NZY top agar and poured over a warm NZY plate. The plates were incubated overnight at 37°C. The plates were overlaid with nitrocellulose membranes and processed as described above. This screening process was repeated until phage clonality was achieved.

Single phagemid excision. Overnight cultures of XLOLR *E. coli* and XL1-Blue-MRF' *E. coli* were grown in NZY broth and LB supplemented with 0.2% [w/v] maltose and 10 mM $MgSO_4$, respectively. Both cultures were harvested by centrifugation at 1000 x*g* for 15 min and re-suspended in 10 mM $MgSO_4$ to an OD_{600} of 1.0. 200 µl of XL1-Blue MRF' cells were combined with 250 µl of phage stock and 1 µl of ExAssist helper phage and incubated at 37°C for 15 min. The culture was added to 3 ml of NZY broth and incubated at 37°C for 3 hrs. The mixture was heated to 65°C for 20 min and then centrifuged at 1500 x*g* for 15 min. The supernatant containing the excised phagemid vector was decanted into a fresh sterile

tube. 200 µl of XLOLR cells were mixed with either 10 µl or 100 µl of phage supernatant and incubated at 37°C for 15 min. 300 µl of NZY broth was added to the samples. Following a 45 min incubation, 200 µl of the culture was plated on LB agar plates containing 50 µg/ml kanamycin, which were incubated at 37°C overnight.

Phagemid DNA preparation and sequencing. Phagemid DNA was isolated from 5 ml bacterial cultures using a Wizard Plus SV minipreps DNA kit (Promega), following the manufacturer's protocol. To verify the presence of an insert, the phagemid vectors were digested with *Eco*RI and *Xho*I or *Xba*I. The DNA fragments were separated by 1% agarose gel electrophoresis in TAE buffer. The cDNA inserts were sequenced at the Ontario Genomics Innovation Centre (Ottawa, Canada) on a 3730 DNA Analyzer (Applied Biosystems) using standard T3 forward and T7 reverse primers.

Expression of recombinant his-tagged proteins in Escherichia coli

Recombinant construct [pET-30a(+)].

Polymerase chain reaction (PCR). Forward and reverse primers, containing *Eco*RI and *Not*I restriction sites respectively, were designed to amplify the complete nucleotide sequence of the desired protein. PCR reactions were performed in a final volume of 25 µl containing 2 mM $MgCl_2$, 1x PCR buffer, 0.2 mM dNTPs, 0.2 mM each primers (see **Table 1**) and 1 unit taq DNA polymerase (Invitrogen). The PCR was performed on the Mastercycler (Eppendorf) using the thermal parameters indicated in **Table 2**. PCR products were separated by 1% agarose gel electrophoresis. Buffers, enzymes, primers and dNTPs were

removed from the PCR products using Montage PCR columns (Millipore) following the manufacture's protocol.

Table 1. PCR primers and templates

	WP32112	**PAN33**
Forward primer	5'-TACAGGAATTCAACCCATACTTCTCC-3'	5'-TACAGGAATTCTTCCCTTTGGAAGAT-3'
Reverse complement primer	5'-TACAGGCGGCCGCACCTCTTTCCTCTCT-3'	5'-TACAGGCGGCCGCGAGGCAAATGAACT-3'
Template DNA	pBK-CMV-WP32112 vector	Pancreatic acinar/ducts cDNA*

*prepared by Dr. Lisa Kauri

Table 2. PCR thermal parameters

	Temperature	**Time**	
Initial denaturation	94°C	60 s	
Denaturation	94°C	60 s	35 cycles
Annealing	55.5°C	90 s	
Elongation	72°C	90 s	
Final elongation	72°C	7 min	

Ligation. The PCR product and pET-30a(+) vector (Novagen), which contains a his-tag sequence at the N-terminal of the multiple cloning site, were digested with *Eco*RI and *Not*I in React3 buffer (Invitrogen) for 1 hr at 37°C. Restriction enzymes were removed using Micropure-EZ (Millipore) following the standard protocol. The amplified sequence was cloned into the *Eco*RI/*Not*I restriction sites of the pET-30a(+) vector. The ligation was performed in a final volume of 20 µl containing 70 ng of linearized vector, ~ 210 ng of digested PCR, 4 µl 5x buffer and 1 µl T4 DNA ligase (Roche). The sample was mixed gently and incubated for 1 hr at room temperature.

Transformation. The ligation sample of interest was transformed into $CaCl_2$ competent BL21 *E. coli* cells. Briefly, 20 µl of ligation reaction was added to 200 µl of cells and gently mixed. The samples were placed on ice for 30 min, heat shocked by a 90 s incubation at 42°C, and then placed on ice for an additional 5 min. 750 µl of LB medium was added to the cell/ligation mixture and the cells were grown at 37°C for 30 min. 200 µl of transformation was spread on antibiotic selective LB agar containing 50 µg/ml kanamycin. The plates were inverted and incubated overnight at 37°C. Clones were randomly selected and grown to log phase (~5 hrs). Plasmid DNA from 5 ml overnight cultures in selective media was purified using Wizard Plus SV Miniprep kit (Promega) following the standard protocol. Plasmids were digested with *Eco*RI and *Not*I and analyzed by 1% agarose gel electrophoresis in TAE buffer to verify the presence of the insert.

Expression of his-tagged proteins. A BL21 *E. coli* culture containing the desired vector was grown, in antibiotic selective LB broth containing 50 µg/ml kanamycin, at 37°C until log phase. To induce expression of the recombinant protein, IPTG was added to a final concentration of 2 mM. Cells were harvested by centrifugation at 10,000 x*g* for 10 min, resuspended in 100 µl dH_2O and lysed in 10 µl 10% SDS and 20 units DNase. Following a 10 min incubation, 300 µl sample buffer was added to the lysed cells.

Semi-purification of his-tagged proteins. His-tagged proteins were semi-purified using Ni-Sepharose-G pull-down under denaturing conditions. A 50 ml bacterial culture was induced to express the his-tagged protein as previously described. Cells were harvested and the pellet was resuspended in 20 ml lysis buffer. Cell lysate was incubated with end-over-end rotation

for 2 hrs at room temperature and centrifuged at 10000 x*g* for 30 min. Meanwhile 200 µl Ni-Sepharose-G (Amersham Biosciences) was washed three times in PBS. The cell lysate supernatant was transferred to the tube containing the Ni-Sepharose-G and incubated with end-over-end rotation at room temperature for 1 hr. The beads were pelleted by centrifugation at 3000 x*g* for 1 min and washed three times in 1 ml PBS. The his-tagged proteins were eluted by adding 400 µl sample buffer to Ni-Sepharose-G and incubating the sample at 95°C for 10 min. Samples from each step were resolved by SDS-PAGE and proteins were visualized with Gelcode blue (Pierce) to verify the purity of the sample.

SDS Polyacrylamide gel electrophoresis (PAGE) and Western blotting

Preparation of 10% and 12.5% SDS-PAGE gels. One dimensional SDS-PAGE was performed using Hoefer mini VE electrophoresis units. A 10% separating gel was prepared by mixing 2.5 ml of 40% acrylamide:bis solution [37.5:1] (BioRad), 2.5 ml of lower Tris (pH 8.8) and 5 ml H_2O. 50 µl of 10% ammonium persulfate (APS) and 10 µl of N,N,N',N'-tetramethylethylenediamine (TEMED) were added and the solution was mixed and poured immediately. A 12.5% separating gel was prepared as above; however, the volume of the acrylamide:bis solution was adjusted to 3.2 ml, and H_2O to 4.3 ml. A 6% stacking gel was prepared by mixing 0.5 ml of 40% acrylamid:bis solution, 0.83 ml upper Tris (pH 6.8) and 2 ml H_2O. 16.7 µl of 10% APS and 3.3 µl of TEMED were added and the solution was mixed and poured immediately.

1D SDS-PAGE. The protein samples were diluted in 6X SDS sample buffer and denatured by heating for 5 mins at 95°C. The BenchMark pre-stained protein ladder (Invitrogen) and MagicMark protein ladder (Invitrogen) were used as protein standards. Protein samples were resolved by electrophoresis at 75 V for 30 min then at 200 V in running buffer.

2D Gel electrophoresis. IPG strips (7cm; BioRad) of the corresponding pH range were rehydrated overnight in 125 µl of sample solution containing the proteins of interest in ReadyPrep Sequential Extraction Reagent 3 (SEB, BioRad) and 1.25 µl Tributyl phosphine (TBP, BioRad). The IPG strips were isoelectrically focused on the PROTEAN IEF Cell (BioRad) at 250 V for 15 min then the voltage was increased to 5000 V until focusing was complete at 30000 Vhrs. The strips were flash frozen in liquid nitrogen and stored at -80°C until use. Strips were reduced and alkylated by sequential 15 min incubations at room temperature in 2% [w/v] dithiothreitol, 0.3% [w/v] SDS and 2.5% [w/v] iodoacetamide, 0.3% [w/v] SDS in equilibration buffer, respectively. Excess buffer was drained and the strips were placed in to the well of a 12.5% SDS-PAGE gel. MagicMark protein ladder (Invitrogen) was used as the protein standard. Proteins were resolved by electrophoresis in running buffer at 75V for 30 min and then voltage was increased to 200V until complete.

Electrotransfer. Proteins were electrotransferred from gels to nitrocellulose membranes or polyvinylidene fluoride (PVDF) membranes at a constant current of 350 mA for 1 hr in transfer buffer. Membranes were stained with Ponseau-S to ensure protein transfer, followed by four washes in PBS-T to destain the membranes.

Immunoblotting. Nitrocellulose membranes containing transferred proteins were used for Western blot analyses. Membranes were blocked in 5% [w/v] skim milk powder in PBS-T overnight at 4°C. After a 1 hr incubation at room temperature in primary antibody diluted in blocking solution (**Table 3**), the membranes were washed 5x 5 min in PBS-T. Following a 1 hr incubation in secondary antibody diluted in blocking buffer (**Table 4**) and 5x 7 min washes in PBS-T, the membranes were incubated in the ECL substrate for 1 min and exposed to Hyperfilm (Amersham Biosciences).

Table 3. Primary antibodies

Primary antibody	Source	Study	Dilution
anti-WP5212 (aa 336-352, 572-586) serum	Sigma-Genosys	2	1/500
monoclonal mouse anti-his antibodies	Serotec	2, 5	1/2000
case study subject serum	N/A	5	1/800
patient serum	N/A	5	1/200
polyclonal anti-WP5212 serum	DNA immunization	3	1/500

Table 4. Secondary antibodies

Conjugated secondary antibody	Company	Dilution
HRP-conjugated goat-anti-rabbit-Ig	Dako	1/20000
HRP-conjugated rabbit-anti-mouse-Ig	Dako	1/15000
HRP-conjugated goat-anti-human-IgG	Sigma	1/20000

HRP = horse radish peroxidase

Silver stain. Gels were washed 3x 5 min in ddH_2O and fixed overnight in 50% ethanol, 5% acetic acid. Following a 10 min wash in 50% ethanol and a 10 min wash in ddH_2O, the gels were placed in sensitizer solution (0.02% [w/v] sodium thiosulphate) for 2 min. The gels were washed 2x 1 min in ddH_2O and incubated in a 0.1% [w/v] silver nitrate solution for 30

min. Following a 1 min rinse in ddH_2O, proteins were visualized by adding the developer (0.04% formalin in 2% [w/v] sodium carbonate). The developed gels were rinsed briefly in ddH_20 and placed in stop solution (5% [v/v] acetic acid) for 5 min and stored in 1% [v/v] acetic acid solution. Spots of interest were cored and placed in 1% [v/v] acetic acid solution.

Matrix assisted laser desorption/ionization time of flight mass spectroscopy (MALDI-TOF MS) analysis

MALDI-TOF MS analyses of the isolated proteins of interest were performed at the Ontario Genomics Innovation Centre (Ottawa, ON). Samples were trypsin digested prior to analysis on the QSTAR XL Hybrid LC/MS/MS system (Applied Biosystems). The MALDI/MS spectra were analyzed and matched to all rat proteins using Mascot searches (Matrix Science).

N-terminal protein sequencing

Proteins of interest were excised from Ponseau-S stained electrotransferred PVDF membranes and sent to the Eastern Quebec Proteomics Core Facility (Laval, QC) for sequencing. N-terminal Edman degradation was performed on a ABI Procise Clc, followed by PTH analysis on a HPLC column. Bioinformatic analyses of the identified sequence were performed.

Construction and immunoscreening of WP5212 peptide library

PEPscreenTM WP5212 peptide library. A custom PEPscreen™ peptide library spanning the WP5212 protein sequence, consisting of 64 17-mer peptides overlapping by 8 amino acids, was constructed at Sigma Genosys (note peptide 397-413 aa failed quality control). Peptides were re-suspended in 10% acetic acid or 8% acetic acid/20% acetonitrile or 7% acetic acid/18% acetonitrile/10% DMSO, based on their solubility (**Appendix 2**).

Immunoscreening. 1 μg of each peptide was dotted on Ultrabind US450 modified polyethersulfone affinity membrane (0.45 μm; PALL Gelman Laboratory). Membranes were blocked overnight in 5% skim milk powder in PBS-T. Plasma from eight type 1 diabetic patients (containing WP5212-reactive antibodies) or eight healthy controls was pooled and diluted in 5% skim milk in PBS-T to a final plasma dilution of 1/200 for each patient. Membranes were incubated in this plasma dilution for 1 hr at room temperature. Following 5x 5 min washes with PBS-T, membranes were incubated with the secondary antibody, HRP-conjugated goat anti-human IgG (Sigma) diluted 1/20,000 in blocking solution, for 1 hr at room temperature. Membranes were washed 5x 5 min in PBS-T. Antibody binding was detected using ECL substrate and exposing the membrane to Hyperfilm. Peptides bound by antibodies in the pooled plasma, were screened with plasma from these individual diabetic patients and healthy control subjects (diluted 1/100 in blocking buffer), as describe above.

Generation of high-titer polyclonal rabbit anti-WP5212 sera

Vector Construction. The full length WP5212 cDNA sequence was excised from the pBK-CMV ZAP Express vector (Stratagene) by sequential restriction digests using *Eco*RI and *Not*I. The pCI-neo mammalian expression vector (Promega) was also digested with these enzymes. The digested samples were separated by 1% agarose gel electrophoresis in TAE buffer and the bands of interest were visualized with UV light and excised. DNA was purified from the gel using the QIAEX II Gel Extration Kit (Qiagen) following the manufacture's protocol. The WP5212 sequence was cloned into the *Eco*RI/*Not*I restriction sites of the pCI-neo vector using T4 DNA ligase (Roche) (ratio insert:vector = 3:1). Electrocompetent XLI-Blue *E. coli* were transformed with the pCI-neo-WP5212 construct and plated on LB agar containing ampicillin (100 μg/ml). Clones were randomly selected and 5 ml cultures were grown overnight in selective media. Plasmids were isolated using the Wizard Plus SV miniprep system (Promega) following the standard manufacture's protocol. Restriction mapping of the vector was performed by digestion with *Nco*I and separation by 1.5% agarose gel electrophoresis in TAE buffer. DNA inserts were sequenced at the Ontario Genomics Innovation Centre (Ottawa, Canada) on a 3730 DNA Analyzer (Applied Biosystems) using the standard T3 and the T7 eukaryotic expression vector (EEV) promoter primers. To verify the full sequence of the insert, internal primers for WP5212 were used (forward primer: 5'-ACCACGGGTTCGTCAAGG-3'; reverse complement: 5'-AACACCTCCTGCACCTCC-3'). Large quantities of plasmid DNA, free of endotoxin contamination, were isolated using the Endofree Plasmid Giga kit (Qiagen) following the manufacture's protocol.

Expression in mammalian cells. In order to confirm the potential of the cDNA vaccine, the expression of WP5212 was verified in transiently transfected mammalian cells. Six-well plates were seeded with $8x10^5$ human embryonic kidney (HEK) 293 cells/well and grown in DMEM supplemented with 10% FBS overnight at 37°C with 5% CO_2 until ~90% confluent. For each transfection, 12 µl of lipofectamine reagent (Invitrogen) was added to 88 µl of Opti-MEM and incubated at room temperature for 5 min. The lipofectamine reagent mixture was then added to 100 µl Opti-MEM containing 2 µg of pCI-neo-WP5212 plasmid DNA and incubated at room temperature for 2 mins. Medium was replaced with fresh DMEM and the lipofectamine/plasmid solution was added. Non-transfected cells were used as a negative control. Following a 5 hr incubation at 37°C with 5% CO_2, the medium was replaced with 2 ml DMEM supplemented with 10% FBS and gentamycin and the cells were incubated for 48 hrs. Cells were washed twice with PBS and collected in 1 ml of cell lysis buffer. Cellular debris was removed by centrifugation at 14000 x*g* for 20 min at 4°C. The supernatant was collected and stored at -80°C.

Verification of mRNA expression. RNA was isolated from the cell lysates using the RNeasy kit (Qiagen) following manufacture's instructions. Reverse transcription (RT) reactions were performed in a final volume of 40 µl containing 2 µg of total RNA, 2 µl $Oligo(dT)_{12-18}$ primer, 0.5 mM dNTPs, 1x first-strand buffer, 10 mM dithiothreitol (DTT), 2 µl RNase OUT and 400 units M-MLV reverse transcriptase (Invitrogen). Samples were incubated at 37°C for 50 min, followed by 94°C for 5 min. PCR was performed in a final volume of 25 µl containing 5 µl cDNA, 2 mM $MgCl_2$, 1x PCR buffer, 0.2 mM dNTPs, 0.2 mM WP5212 primers (forward: 5'-ACCACGGGTTCGTCAAGG-3'; reverse complement: 5'-

AACACCTCCTGCACCTCC-3') and 1 unit taq DNA polymerase (Invitrogen). Template DNA from the pCI-neo:WP5212 vector was used as a positive control for the PCR reaction. The PCR was performed on the Mastercycler (Eppendorf) using the following thermal parameters: Initial denaturation: 5 min at 94°C; Cycle (30x): denaturation 30 sec at 94°C, annealing 45 sec at 57°C, elongation 2 min at 72°C; Final elongation: 7 min at 72°C. PCR products were separated by 1% agarose gel electrophoresis in TAE buffer.

Immunization. Two female New Zealand White rabbits were kept at the University of Ottawa animal facility in accordance with the institutional guidelines and ethical approval. The immunization protocol was modeled after the procedure described in (76). Briefly, pre-immune sera were collected from the two rabbits prior to the immunizations. The rabbits received three immunizations consisting of four intradermal injections of 80 μg plasmid DNA in 200 μl of sterile 0.15 N NaCl at two-week intervals. Following these injections, blood was collected every week for eight weeks. A final immunization was given on week 19 and the final blood sample was collected eleven days later. Blood samples were centrifuged at ~5,000 x*g*, sera were collected, aliquoted and stored at -80°C. Rabbit sera were analyzed by ELISA and Western blotting using recombinant WP5212.

Enzyme-linked immunosorbent assay (ELISA)

96-well plates (Nunc) were coated with 1.0 μg of purified recombinant his-tagged WP5212 (prepared by Majid Mojibian and Alexander Strom) diluted in ELISA binding buffer. Wells incubated with ELISA binding buffer alone were used as negative controls. Wells were blocked with 3% BSA in PBS and incubated for 1 hr at 37°C. Following two washes in

PBS-T (0.05%), the wells were incubated for 1 hr at room temperature with serum samples diluted 1/500 in 1% BSA in PBS. The plates were washed seven times in PBS-T prior to a one-hour incubation with the secondary antibody (HRP-conjugated goat-anti-rabbit Igs diluted 1/2000 in 1% BSA in PBS). 100 µl of TMB (3,3',5,5'-tetramethylbenzidine) substrate reagent (PharMingen) was added to each well. The colourimetric reaction was allowed to proceed for 5 min and then stopped with 100 µl 1M H_2SO_4. Absorbance (450 nm) was read on the Multiskan Ascent 96-well plate reader (Thermo Labsystems).

Bioinformatics

Homology searches. Nucleotide and translated Basic Local Alignment Search Tool (BLAST) searches (77) of the wheat clones were performed against the GenBank database using default settings. When searching for the identity of the 33 kDa pancreatic protein, protein-protein BLAST searches were performed using the PAM 30 matrix which identifies short, nearly exact matches.

Three-dimensional modeling. The translated amino acid sequence of WP32112 was submitted to SWISS-MODEL (http://swissmodel.expasy.org/) for three-dimensional modeling. The ensuing predicted structure of WP32112 was viewed using Deep View-SwissPDB Viewer (http://ca.expasy.org/spdbv/).

Post-translational modification prediction. The amino acid sequence for pancreatic trypsin 1 was submitted to various algorithms listed below in **Table 5** to identify putative posttranslational modifications.

Table 5. Computer algorithms used to predict post-translational modifications

Algorithm	Predicted post-translational modification	Reference
NetAcet	Acetylation sites	(78)
NetOGlyc	O-linked *N*-acetylgalactosamine (O-GalNAc) gylcosylation sites	(79)
YinOYang	O-linked beta *N*-acetylglucosamine (O-beta-GlcNAc) attachment sites	(80)
NetPhos	Serine, threonine and tyrosine phosphorylation sites	(81)
PrePS	Prenylation sites	http://mendel.imp.ac.at/sat/PrePS/index.html
Sulfinator	Tyrosine sulfation sites	http://ca.expasy.org/tools/sulfinator

RESULTS

Study 1 – Epitope mapping of the diabetes-related wheat protein WP5212

There are multiple IgG-binding regions throughout the WP5212 protein. A peptide library, consisting of 64 overlapping peptides representing the entire amino acid sequence of WP5212, was generated to examine which regions are recognized by WP5212-reactive antibodies. Each peptide was 17 amino acids in length, overlapping the previous peptide by eight amino acids. One peptide, corresponding to aa 397-413, failed quality control and consequently was not included in the study. The peptide library was probed with pooled plasma from eight WP5212 antibody-positive T1D patients and eight healthy controls. The characteristics of the study groups are listed below in **Table 6**. Eight IgG-binding regions were identified (**Figure 1**), corresponding to amino acids 1-17, 163-180, 217-234, 262-279, 281-296, 388-404, 478-495 and 505-521.

Table 6. Characteristics of the human study groups (Study 1)

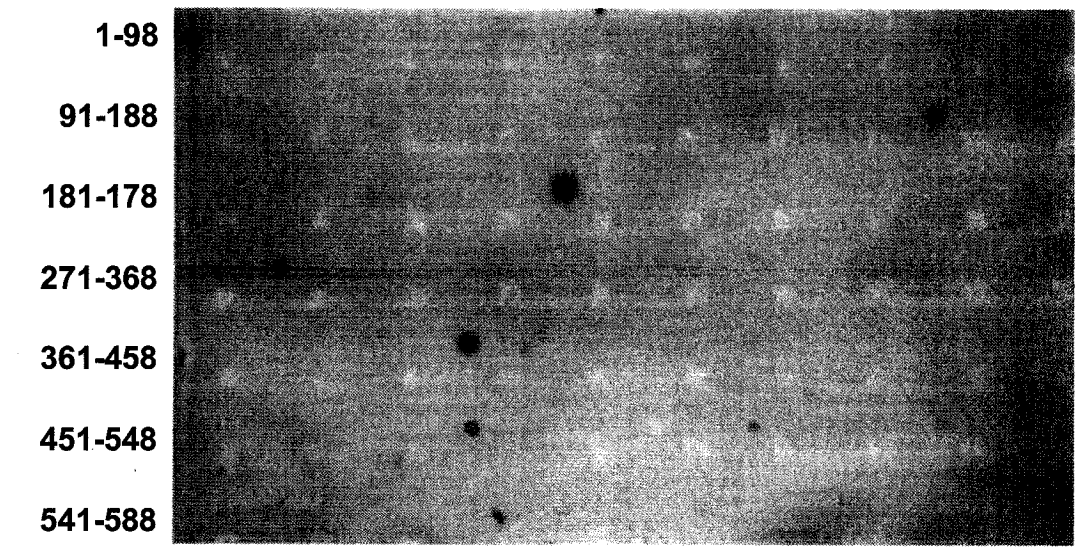
Subject Type	N	Sex Ratio (M/F)	Mean age ± SD (years)	Mean age ± SD at T1D diagnosis (years)
T1D	8	3/5	23.5 ± 9.5	14.6 ± 10.8*
Control	8	5/3	25.8 ± 3.0	n/a

* age at T1D diagnosis missing for one patient

Identification of immunodominant epitopes. To determine if any of these peptides were immunodominant, each set of eight peptides was probed with plasma from individual T1D patients considered WP5212 antibody positive and from healthy control subjects. The characteristics of the individuals are listed below in **Table 7**. There were no significant differences in sex ratio or age at sampling between the study groups. Antibodies from all the

Figure 1. There are multiple IgG-binding regions throughout the WP5212 protein. *Panel A*: Epitope mapping was performed on the WP5212 protein by creating a peptide library, containing sixty four 17-mer overlapping peptides. 1 μg of each peptide was probed with pooled serum from nine diabetic patients with anti-WP5212 antibody reactivity. *Panel B*: The sequence of WP5212 protein is shown. Highlighted peptides correspond to the IgG-binding regions shown in panel A.

A.



B.

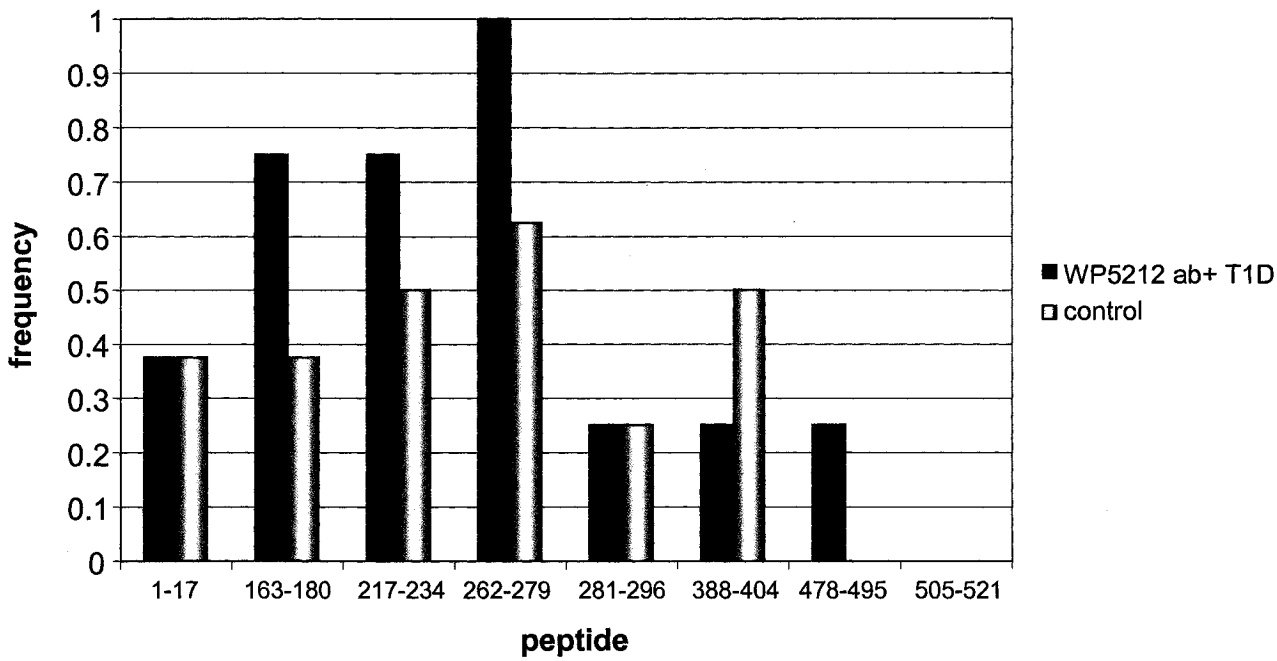
```
[redacted]SLLFAAAVSASHDEEEDRRGGRSLQRCVQRCQQ   50
DRPRYSHARCVQECRDDQQQHGRHEQEEQGRGHGRHGEGEREEEQGRGRG   100
RRGQGEREEEQGRGRGRRGEGERDEEHGDGRRPYVFGPRSFRRIIRSDHG   150
FVKALRPFDEVS[redacted]PRAFVVPGLTDADGVGYVAQG   200
EGVLTVIENGEKRSYT[redacted]ANTDGRRKLVIAKILHT   250
ISVPGKFQYFS[redacted]K[redacted]EEEK   300
SISIVRASEEQLRELRRQASEGDQGHHWPLPPFRGDSRDTFNLLEQRPKI   350
ANRHGRLYEADARSFHALAQHDVRVAVANITPGSMTA[redacted]   400
[redacted]GEVEIVCPHLGRDSERREQEHGKGRWRSEEEEDDRRQQRRRGSGSE   450
SEEEQDQQRYETVRARVSRGSAFVVPP[redacted]VVCFEI   500
NAER[redacted]DPAQELAFGRPAREVQEVFRAKDQQDEGF   550
VAGPEQQQEHERGDRRRGDRGRGDEAVEAFLRMATAAL   588
```

T1D patients bound to multiple peptides. Surprisingly, antibodies from healthy control subjects also bound to the same epitopes. The frequencies of antibody reactivity against the peptides for the two groups are shown in **Figure 2.** Antibodies from 75% (6/8) of the patients recognized peptide aa163-180 and aa217-234. Antibodies from 100% (8/8) bound to peptide aa262-279 of the patients. While these frequencies were higher compared to the control subjects (aa163-318: 37.5%; aa217-234: 50%; aa262-279: 62.5%), the difference was not significant (Fisher Exact Test). Interestingly, antibodies from diabetic patients detected lower concentrations of serially diluted recombinant WP5212 in comparison to control subjects (not shown).

Table 7. Characteristics of the individual subjects in the human study (Study 1)

Disease status	Subject	Age	Sex	Age at T1D diagnosis	Anti-WP52112 status
Type 1 diabetes	D1	25	F	18	+
	D2	25	M	no data	+
	D3	23	F	17	+
	D4	18	F	11	+
	D5	29	M	11	+
	D6	19	M	5	+
	D7	8	F	4	+
	D8	41	F	36	+
Healthy control	C1	29	M	n/a	-
	C2	28	M	n/a	-
	C3	21	F	n/a	-
	C4	24	F	n/a	-
	C5	28	F	n/a	-
	C6	26	M	n/a	-
	C7	28	M	n/a	-
	C8	22	M	n/a	-

Figure 2. Frequency of antibody reactivity against selected IgG-binding peptide of WP5212. Peptides were probed with sera from WP5212 antibody positive T1D patients (n=8) and healthy control subjects (n=8) and antibody reactivity was scored as positive or negative. The differences in the frequencies of antibody reactivity between groups are not significant (Fisher Exact Test).

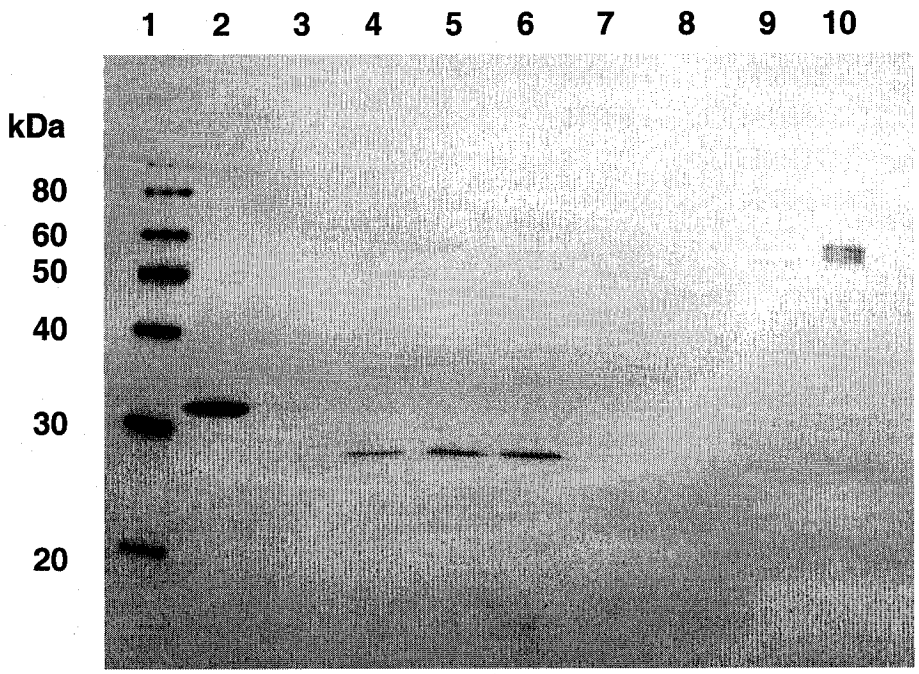

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
frequency
1-17
163-180
217-234
262-279
281-296
388-404
478-495
505-521
peptide
WP5212 ab+ T1D
control

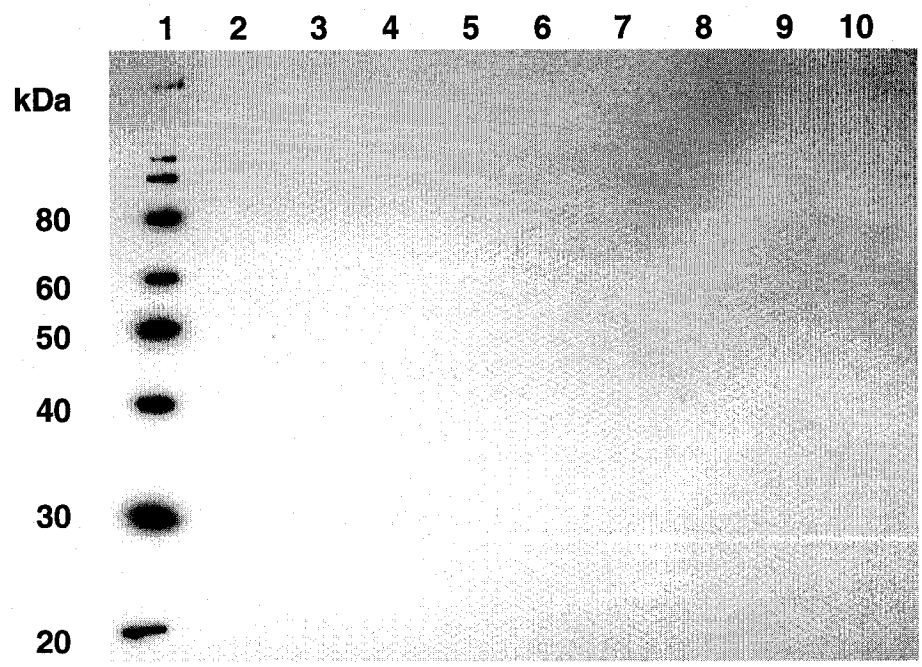
Study 2 – Identification of candidate molecular mimics in target tissues

Antibodies raised against WP5212 bind to proteins in target tissues. Rabbit polyclonal antibodies were previously raised against two short peptides of WP5212 (aa 336-352 and 572-586) predicted to be immunogenic by Sigma-Genosys (Mississauga, ON). In order to identify potential candidate molecular mimics, 40 μg of protein isolated from target tissues from the BioBreeding rat were resolved by 1D SDS-PAGE, transferred to nitrocellulose membranes and probed using the anti-WP5212 serum. Western blot analysis revealed WP5212 antibodies bound to a 33 kDa protein present in the exocrine pancreatic tissue and to a 28 kDa protein present in purified islets, MLNs and spleen, while no antibody binding was observed in liver, heart or kidney tissues (**Figure 3**). Antibody reactivity against these two proteins was not present in Western blots using control, pre-immune serum (**Figure 3**).

Isolation of pancreatic 33 kDa protein. In order to determine the identity of the 33 kDa protein expressed in the exocrine pancreatic tissue, 2D gel electrophoresis was performed followed by MALDI-TOF MS analysis at the Ottawa Institute of Systems Biology facility (Ottawa, ON). Proteins from the pancreatic tissue homogenate were partially purified by sucrose density gradient centrifugation to increase the electrophoretic resolution for identification. The 20-60% gradient was run and seventeen 0.8 ml-fractions were collected. Western blot analyses using anti-WP5212 antibodies were performed to define the fractions containing the proteins of interest. These fractions were subsequently pooled, dialyzed against dH_2O and concentrated. Proteins were separated by isoelectric focusing using IEF strips pI 4→7, resolved in the second dimension by 10% SDS-PAGE and then electrotransferred to nitrocellulose membranes. Western blot analysis revealed anti-WP5212

Figure 3. Antibodies raised against WP5212 cross-react with a 33 kDa pancreatic protein and a 28 kDa protein expressed in the pancreatic islets, spleen and MLNs. Rat tissue extracts were separated by SDS-PAGE and western blots were probed with rabbit serum raised against two WP5212 peptides. Prestained benchmark ladder (lane 1); target tissues: pancreas (lane 2), pancreatic acinar/duct cells (lane 3), isolated islets (lane 4), MLNs (lane 5) and spleen (lane 6); control tissues: liver (lane 7), kidney (lane 8), heart (lane 9); positive control: WP5212 (lane 10).


1 2 3 4 5 6 7 8 9 10
kDa
80
60
50
40
30
20


1 2 3 4 5 6 7 8 9 10
kDa
80
60
50
40
30
20

antibody binding to two 33 kDa proteins with pIs of approximately 5 (**Figure 4**). A parallel 2D run was silver stained (**Figure 4**) to reveal the two protein spots of interest.

Identification of pancreatic 33 kDa protein by MALDI-TOF MS analysis. To identify these proteins, the two spots of interest were excised from the gel and analyzed by MALDI-TOF MS. Sequence database searches were performed using Mascot software (82) limited to *Rattus norvegicus*. The search yielded several candidate proteins, including 78 kDa glucose-regulating protein precursor, tropomyosin and pancreatic trypsin 1 (**Table 8**).

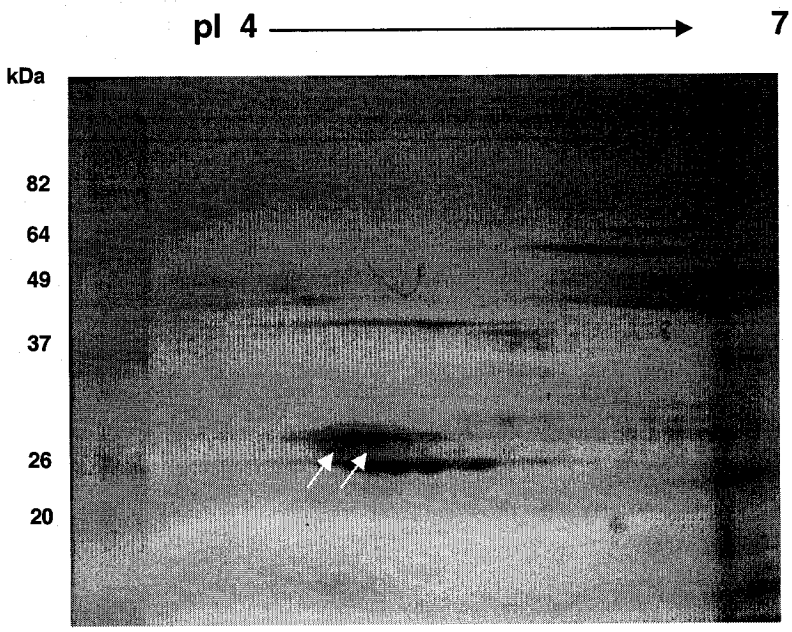
Table 8. Peptides identified by MALDI-TOF MS analysis

Protein	Accession No.	Score	Theoretical kDa/pI	Peptide matches
78 kDa glucose-regulating protein precursor	gi|121570	114	72.4/5.07	LTPEEIER ITITNDQNR ELEEIVQPIISK
tropomyosin	gi|7447397	65	28.7/4.69	MELQEIQLK + oxidation (M) KLVIIEGDLER
pancreatic trypsin 1	gi|6981420	39	26.0/4.71	LGEHNINVLEGDEQFINAAK

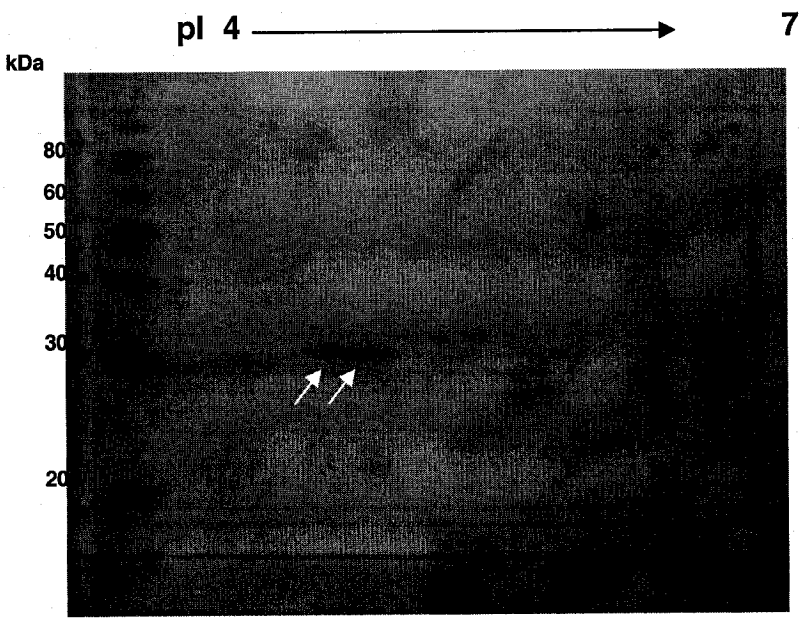
N-terminal sequencing identifies the pancreatic 33 kDa protein as pancreatic trypsin 1. While the 78 kDa glucose-regulating protein precursor, a member of the HSP70 family, reported the best peptide matches, this protein is twice the size of the pancreatic 33 kDa protein. In addition, anti-HSP70 antibodies failed to bind to proteins of the expected size (not shown). Therefore, pancreatic proteins separated by 2D gel electrophoresis were electrotransferred to PVDF membrane and stained with Ponseau-S. The spots of interest were excised and 19 residues were identified by N-terminal sequencing (Eastern Quebec Core Proteomics Facility, Laval University). The resulting identified amino acids indicate the presence of more than one major sequence (**Table 9**). To identify the 33 kDa protein,

Figure 4. Two-dimensional gel electrophoresis analysis of proteins isolated from the pancreas. *Panel A*: Pancreatic proteins were resolved by two-dimensional gel electrophoresis and proteins were visualized with silver stain. *Panel B:* Western blot analysis of these proteins reveal anti-WP5212 antibody binding to two 33 kDa proteins. White arrows show the corresponding spots.

A.



B.



BLAST searches for short, nearly exact matches using the possible amino acid combinations were performed, limiting the searches to *Rattus norvegicus*. The most probable candidate peptide is FPLEDDDKIVGGYTCPE, which shares 15/17 identities with amino acids 16-32 of pancreatic trypsin 1. Additional analysis of pancreatic trypsin 1 revealed the N-terminus contains a signal peptide from amino acid 1-15 and MALDI-TOF MS analysis matched one 20mer peptide with pancreatic trypsin 1 (**Table 8**); this further supports the identification of the 33 kDa protein as pancreatic trypsin 1.

Table 9. N-terminal sequence results

Amino acid Residue	**Identified amino acids**	
	major	**minor**
1	Phe	Pro, Val
2	Leu , Pro	Gly, Glu
3	Leu, Glu	Asp, Asn, Val, Gly
4	Asp, Glu	N/A
5	Asp	N/A
6	Asp	N/A
7	Asp	Lys
8	Lys, Ile	Val
9	Ile	Val
10	Val, Gly	N/A
11	Gly, Val	N/A
12	Gly, Tyr	Leu
13	Tyr, Thr	N/A
14	Tyr, Glu	Leu
15	Glu, Thr	Pro
16	Pro, Glu	N/A
17	Pro, Glu	N/A
18	Pro, Asn	N/A
19	Asn	N/A

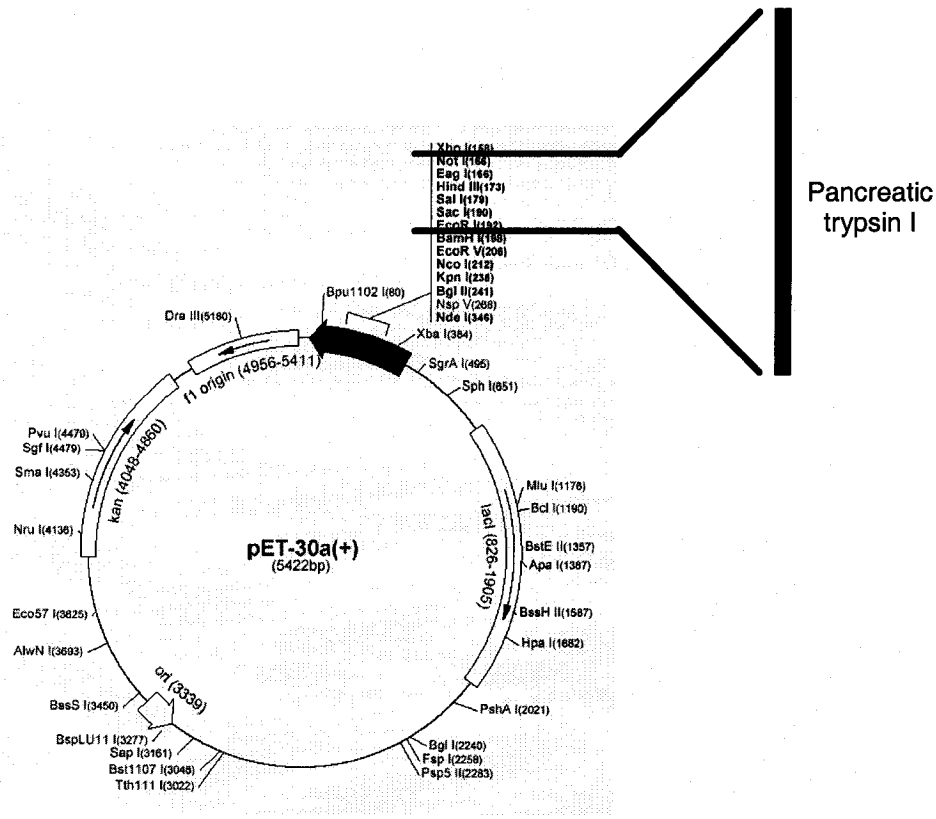
Anti-WP5212 antibodies do not bind to recombinant pancreatic trypsin 1 expressed in E. coli. Pancreatic trypsin 1 DNA was amplified by PCR from pancreatic acinar/duct cDNA,

using primers flanked by *Eco*RI and *Not*I restriction sites. The amplified ~750 bp sequence was ligated into the *Eco*RI/*Not*I cloning site of the pET-30a(+) vector (**Figure 5**). BL21 *E. coli* were transformed with the construct and plated on selective LB plates containing 50 μg/ml kanamycin. Eight clones were randomly selected and plasmid DNA was isolated from an overnight culture. The presence of the insert was verified by restriction digests of plasmid DNA with *Eco*RI and *Not*I (**Figure 5**). The orientation and integrity of the translated amino acid sequence was established by sequencing (**Figure 6**). To confirm pancreatic trypsin 1 protein expression, 1D SDS-PAGE and immunoblotting were performed. Coomassie blue staining revealed over-expression of a 40 kDa protein, the expected molecular weight of the his-tagged pancreatic trypsin 1 protein, compared to control *E. coli* (**Figure 7**). Western blot analysis indicated binding of anti-his antibodies to the over-expressed 40 kDa protein, supporting the expression of his-tagged pancreatic trypsin 1. However, anti-WP5212 antibodies did not recognize and bind to the recombinant pancreatic trypsin 1 expressed in *E. coli* (**Figure 7**).

Pancreatic trypsin 1 contains several predicted post-translational modifications. The amino acid sequence of *Rattus norvegicus* pancreatic trypsin 1 was submitted to several bioinformatics algorithms in order to predict the putative post-translational modifications of this protein. There is one predicted acetylation site (78), six predicted O-beta-GlcNAc attachment sites (80) and two predicted phosphorylation sites (81). There are no predicted O-GalNAc glycosylation sites (79), prenylation sites (http://mendel.imp.ac.at/sat/PrePS/index.html) or tyrosine sulfation sites

Figure 5. Construction of the pET-30a(+)-pantry1 expression vector. *Panel A*: The pancreatic trypsin 1 nucleotide sequence was amplified by PCR and cloned into the NotI/EcoRI cloning site of the pET-30a(+) vector (Novagen). *Panel B*: Restriction digest of the pET-30a(+)-pantryI vector using *Not*I/*Eco*RI. The pET-30a(+) vector is 5.4 kb and the pancreatic trypsin 1 insert is 800 bp.

A.


Pancreatic trypsin I
Xho I(158)
Not I(166)
Eag I(166)
Hind III(173)
Sal I(179)
Sac I(190)
EcoR I(192)
BamH I(198)
EcoR V(206)
Nco I(212)
Kpn I(236)
Bgl II(241)
Nsp V(268)
Nde I(346)
Bpu1102 I(80)
Xba I(384)
SgrA I(496)
Sph I(651)
Mlu I(1178)
Bcl I(1190)
BstE II(1357)
Apa I(1387)
BssH II(1587)
Hpa I(1682)
PshA I(2021)
Bgl I(2240)
Fsp I(2258)
Psp5 II(2283)
Tth111 I(3022)
Bst1107 I(3048)
Sap I(3161)
BspLU11 I(3277)
BssS I(3450)
AlwN I(3693)
Eco57 I(3825)
Nru I(4136)
Sma I(4353)
Sgf I(4479)
Pvu I(4470)
Dra III(5180)
pET-30a(+)
(5422bp)
f1 origin (4956-5411)
kan (4048-4860)
lacI (826-1905)
ori (3339)

B.

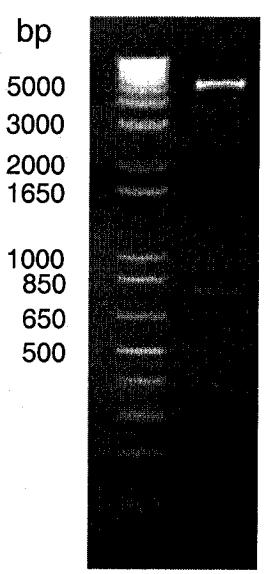

bp
5000
3000
2000
1650
1000
850
650
500

Figure 6. Verification of the cloned pancreatic trypsin I amino acid sequence integrity. Sequence alignment of the pancreatic trypsin I translated sequence (insert) cloned into the pET-30a(+) vector with the known pancreatic trypsin I amino acid sequence. The alignment was performed by ClustalW (www.ebi.ac.uk/clustalw/).

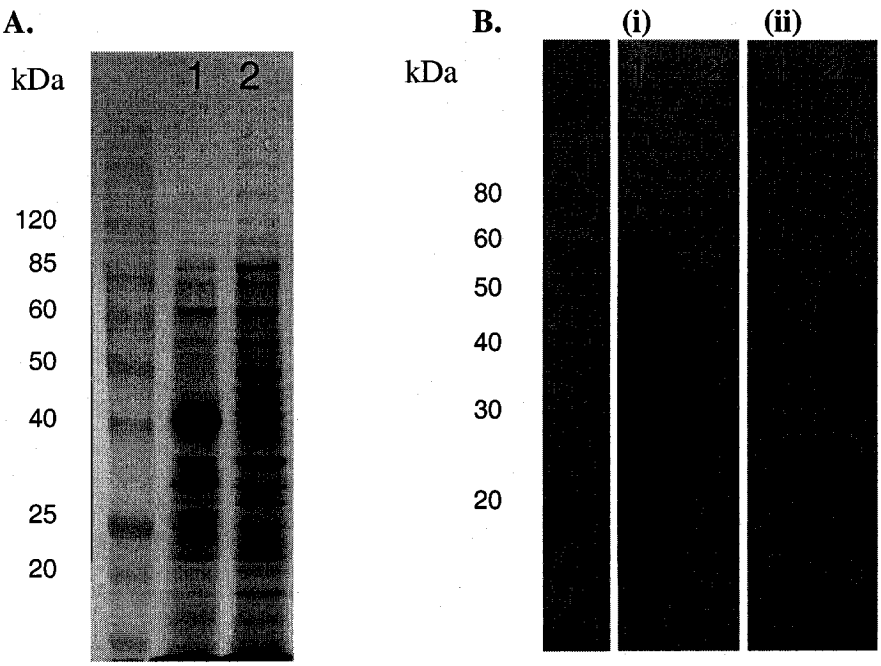
```
Insert:   53  FPLEDDDKIVGGYTCPEHSVPYQVSLNSGYHFCGGSLINDQWVVSAAHCYKSRIQVRLGE  112
PantryI:  16  ************************************************************  75

Insert:  113  HNINVLEGDEQFINAAKIIKHPNYSSWTLNNDIMLIKLSSPVKLNARVAPVALPSACAPA  172
PantryI:  76  ************************************************************  135

Insert:  173  GTQCLISGWGNTLSNGVNNPDLLQCVDAPVLSQADCEAAYPGEITSSMICVGFLEGGKDS  232
PantryI: 136  ************************************************************  195

Insert:  233  CQGDSGGPVVCNGQLQGIVSWGYGCALPDNPGVYTKVCNFVGWIQDTIAAN  283
PantryI: 196  **************************************************  246
```

Figure 7. Expression of his-tagged pancreatic trypsin I protein. *Panel A*: Proteins from *E. coli* transformed with the pET-30a(+)-pantrypI vector (lane 1) and control *E. coli* (lane 2) separated by SDS-PAGE and visualized by Coomassie blue staining revealed overexpression of a 40 kDa protein. Prestained benchmark ladder (Invitrogen) was used as the protein standard. *Panel B*: Western blot analysis indicates that anti-his antibodies (i) bind to the overexpressed protein; however there is no anti-WP5212 antibody binding to this protein (ii). MagicmarkX (Invitrogen) was used as the protein standard.


A.
kDa
1
2
120
85
60
50
40
25
20
B.
(i)
(ii)
kDa
80
60
50
40
30
20

(http://ca.expasy.org/tools/sulfinator/). The predicted post-translational modifications are listed in **Table 10**.

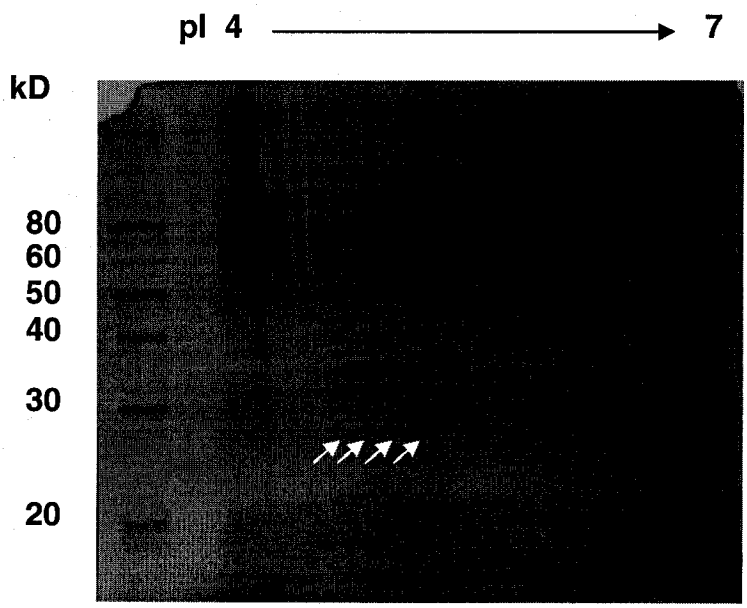
Table 10. Predicted post-translational modification for pancreatic trypsin 1

Post-translational modification	Amino acid	Amino acid residue
acetylation	serine	2
O-GlcNAc	serine	34
O-GlcNAc	serine	40
O-GlcNAc	serine	60
O-GlcNAc	threonine	137
O-GlcNAc	serine	167
O-GlcNAc	threonine	242
phosphorylation	serine	34
phosphorylation	serine	167

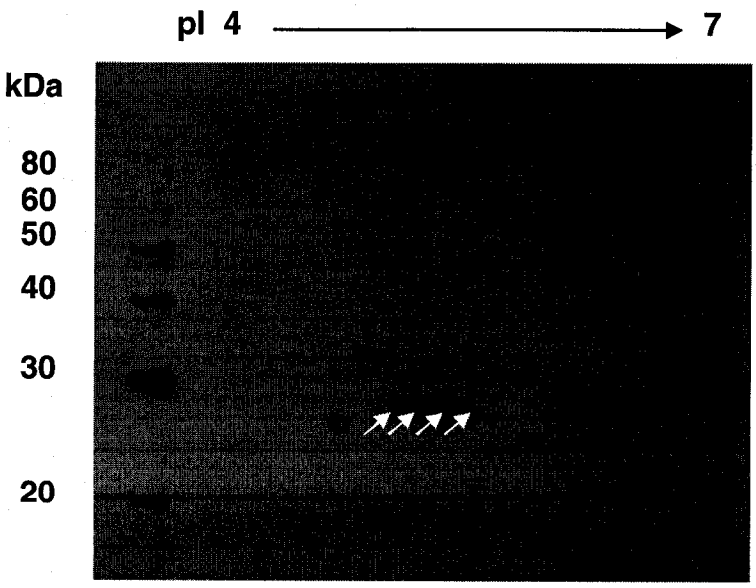
Isolation of the 28 kDa protein expressed in pancreatic islets, spleen and MLNs. To identify the 28 kDa protein recognized by anti-WP5212 antibodies, 2D gel electrophoresis and western blot were performed using the proteins of the spleen. The spleen was chosen because it is the largest of the three tissues (spleen, MLNs and islets) and therefore there is more available material. Proteins from spleen homogenate were separated by sucrose density gradient centrifugation. Fractions containing the protein of interest, identified by western blot analysis, were pooled, dialyzed and concentrated. Proteins were separated by IEF (pI 4→7) in the first dimension and 12% SDS-PAGE in the second dimension. Following 2D gel electrophoresis, the proteins were electrotransferred to nitrocellulose membranes. Western blot analysis indicated anti-WP5212 antibody binding to four 28 kDa proteins with pIs of approximately 5.5 (**Figure 8**). A parallel 2D run, visualized by silver stain, revealed numerous faint protein spots in the area of interest (**Figure 8**). Consequently, it was difficult to determine which spots corresponded to the proteins of interest.

Figure 8. Two-dimensional gel electrophoresis analysis of proteins isolated from the spleen. *Panel A*: Proteins from spleen homogenate were resolved by two-dimensional gel electrophoresis and proteins were visualized with silver stain. *Panel B:* Western blot analysis of these proteins reveal anti-WP5212 antibody binding to four 28 kDa proteins. White arrows show the corresponding areas.

A.



B.



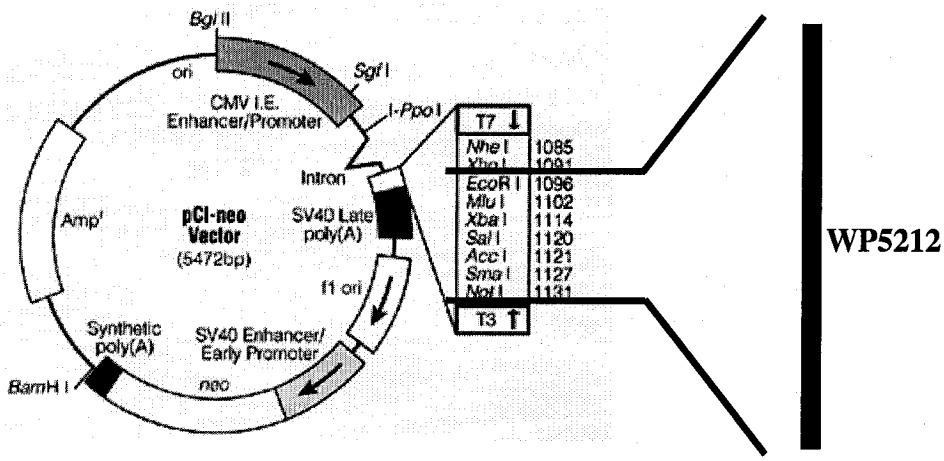
Study 3 – Generation of rabbit polyclonal high-titer broad-spectrum anti-WP5212 sera

The polyclonal WP5212 antibodies, raised against two short antigenic amino acid sequences, facilitated the identification of candidate molecular mimics in target tissues. However, as previously mentioned, these antibodies only recognize a small portion of the WP5212 molecule. In order to obtain broader spectrum high-titer polyclonal WP5212 antibodies, polyclonal anti-WP5212 sera was produced using DNA immunization. Serum from immunized rabbits was collected and verified for the presence of WP5212 antibodies. These antibodies should recognize conformational and linear epitopes and can be used to identify additional molecular mimics by immunoblotting, immunoprecipitation and immunohistochemical staining of target tissues.

Generation of pCI-neo-WP5212 construct. The WP5212 cDNA sequence was excised from the pBK-CMV-WP5212 vector (83) using *Eco*RI and *Xho*I and cloned into the multiple cloning site of the pCI-neo mammalian expression vector (Promega) which contains a CMV promoter the recombinant protein expression. XL1-blue *E. coli* were transformed with pCI-neo:WP5212 and plated on selective LB media containing 100 μg/ml ampicillin. Plasmid DNA was isolated from 10 randomly selected clones and digested with *Eco*RI and *Xho*I to ensure the presence of the insert. Restriction digest analysis revealed all ten clones contained an insert of the expected size. Clone #2 was selected and the orientation of the insert was confirmed by restriction mapping with *Nco*I, which cuts the vector four times and the insert twice (**Figure 9**), followed by sequencing (**Figure 10**).

Figure 9. The pCI-neo-WP5212 construct used for immunization. *Panel A*: The WP5212 nucleotide sequence was amplified and cloned into the *EcoRI/NotI* cloning site of the pCI-neo mammalian expression vector (Promega) *Panel B:* Restriction mapping of the pCI-neo-WP5212 mammalian expression vector with *Nco*I. The pCI-neo vector contains four *Nco*I and WP5212 possess two. Digestion with this enzyme produces six DNA fragments of the predicted sizes: approximately 3.0 kb, 1.8 kb, 1.1 kb, 700 bp, 600 bp and 300 bp.

A.


Bgl II
ori
Sgf I
CMV I.E.
Enhancer/Promoter
I-Ppo I
T7
Nhe I 1085
Xho I 1091
EcoR I 1096
Mlu I 1102
Xba I 1114
Sal I 1120
Acc I 1121
Sma I 1127
Not I 1131
T3
Intron
Amp
pCI-neo
Vector
(5472bp)
SV40 Late
poly(A)
f1 ori
Synthetic
poly(A)
SV40 Enhancer/
Early Promoter
BamH I
neo
WP5212

B.

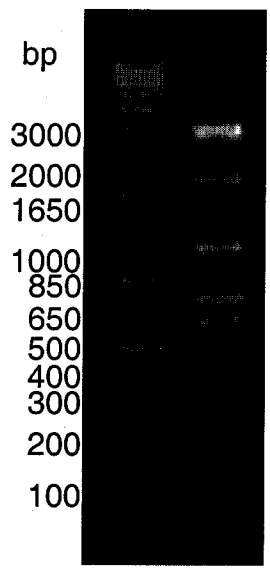

bp
3000
2000
1650
1000
850
650
500
400
300
200
100

Figure 10. Verification of the cloned WP5212 amino acid sequence integrity. Sequence alignment of the translated WP5212 sequence cloned into the pCI-neo vector with the known amino acid sequence of the diabetes-related Glb1 homologue. The alignment was performed by ClustalW (www.ebi.ac.uk/clustalw/).

```
WP5212      MATRGRATIPLLFLLGTSLLFAAAVSASHDEEEDRRGGRSLQRCVQRCQQDRPRYSHARC 60
insert      ************************************************************

WP5212      VQECRDDQQQHGRHEQEEQGRGHGRHGEGEREEEQGRGRGRRGQGEREEEQGRGRGRRGE 120
insert      ************************************************************

WP5212      GERDEEHGDGRRPYVFGPRSFRRIIRSDHGFVKALRPFDEVSRLLRGIRNYRVAIMEVNP 180
insert      ************************************************************

WP5212      RAFVVPGLTDADGVGYVAQGEGVLTVIENGEKRSYTVRQGDVIVAPAGSIMHLANTDGRR 240
insert      ************************************************************

WP5212      KLVIAKILHTISVPGKFQYFSAKPLLASLSKRVLTAALKTSDERLGSLLGSRQGKEEEEK 300
insert      ************************************************************

WP5212      SISIVRASEEQLRELRRQASEGDQGHHWPLPPFRGDSRDTFNLLEQRPKIANRHGRLYEA 360
insert      ************************************************************

WP5212      DARSFHALAQHDVRVAVANITPGSMTAPYLNTQSFKLAVVLEGEGEVEIVCPHLGRDSER 420
insert      ************************************************************

WP5212      REQEHGKGRWRSEEEEDDRRQQRRRGSGSESEEEQDQQRYETVRARVSRGSAFVVPPGHP 480
insert      ************************************************************

WP5212      VVEIASSRGSSNLQVVCFEINAERNERVWLAGRNNVIAKLDDPAQELAFGRPAREVQEVF 540
insert      ************************************************************

WP5212      RAKDQQDEGFVAGPEQQQEHERGDRRRGDRGRGDEAVEAFLRMATAAL 588
insert      **************************************************
```

HEK 293 cells transfected with pCI-neo-WP5212 express WP5212 mRNA. To verify that the pCI-neo-WP5212 construct will induce expression of WP5212 in mammalian cells, HEK 293 cells were transiently transfected with the construct. RT-PCR analysis of total mRNA from transfected cells using WP5212 internal primers revealed a band of the expected size (~1.1 kb), not present in untransfected cells (**Figure 11**).

DNA immunization results in anti-WP5212 antibodies in the rabbit sera. Two rabbits received three sets of intradermal injections with 320 μg of the pCI-neo-WP5212 construct at two week intervals. Sera were collected prior to immunizations, weekly during weeks 5-12 and week 19. The rabbits were immunized with a final set of intradermal injections on week 19 and blood samples were collected 11 days later. Anti-WP5212 antibody activity was determined by ELISA and western blot analyses. The anti-WP5212 antibody activity of the sera was monitored throughout the immunization period by ELISA (**Figure 12**). Following the first three immunizations, the antibody titer of Rabbit #1 increased until week eleven, when the anti-WP5212 activity reached a plateau. The antibody titer did not increase after the final immunization. Rabbit #2 produced a higher titer antiserum in response to the first three WP5212 DNA immunizations, which did not drop and remained stable until the final sample collection. Western blot analyses demonstrated that antibodies in the sera isolated from the final bleed of both rabbits bound to recombinant his-tagged WP5212 (**Figure 13**). Serum from Rabbit #2 had higher anti-WP5212 antibody reactivity than the serum from Rabbit #1. No antibody binding was detected in the pre-immunization sera from either rabbit (**Figure 13**).

Figure 11. Expression of WP5212 mRNA in transfected HEK 293 cells. HEK 293 cells were transiently transfected with the pCI-neo-WP5212 vector. Total mRNA was isolated and expression of WP5212 mRNA was determined by RT-PCR. Analysis reveals amplification of a 1.1 kb band in transfected cells (lane 1), not present in control cells (lane 2). pCI-neo-WP5212 plasmid DNA was used as the template for the PCR positive control (lane 4), while the negative control lacked a DNA template (lane 3).

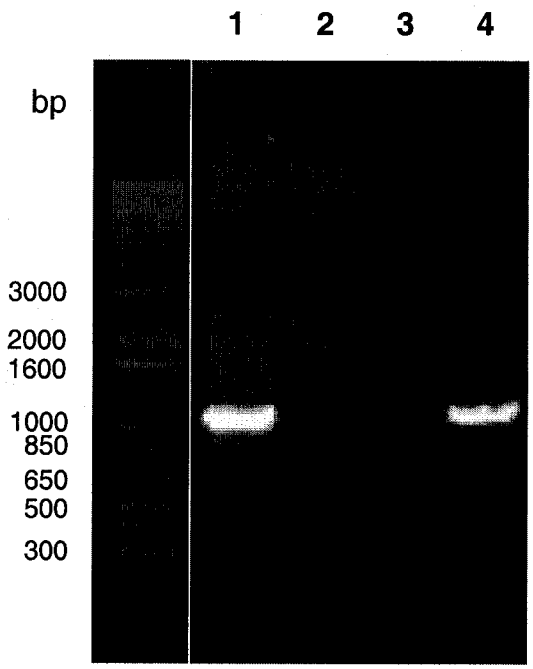
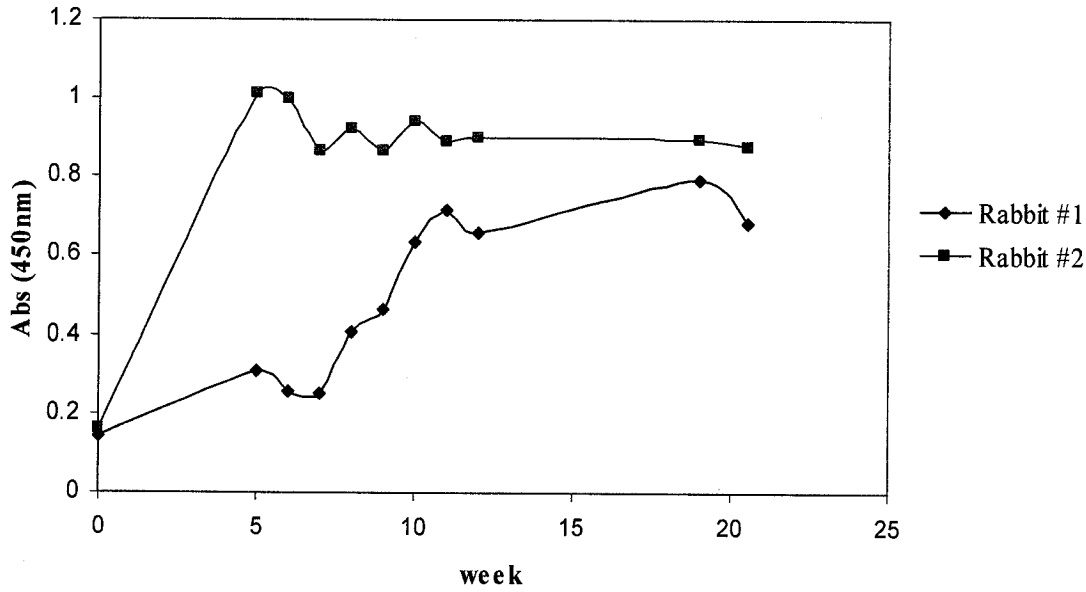

1
2
3
4
bp
3000
2000
1600
1000
850
650
500
300

Figure 12. DNA immunization generated high anti-WP5212 antibody reactivity, as determined by ELISA. *Panel A:* The sera show increased antibody titer over the length of the experiment with respect to the pre-immunized sera, with Rabbit #2 having higher anti-WP5212 antibody reactivity than Rabbit #1. *Panel B:* A dilution curve of the final bleed of the two rabbits.

A.


1.2
1
0.8
0.6
0.4
0.2
0
Abs (450nm)
0
5
10
15
20
25
week
Rabbit #1
Rabbit #2

B.

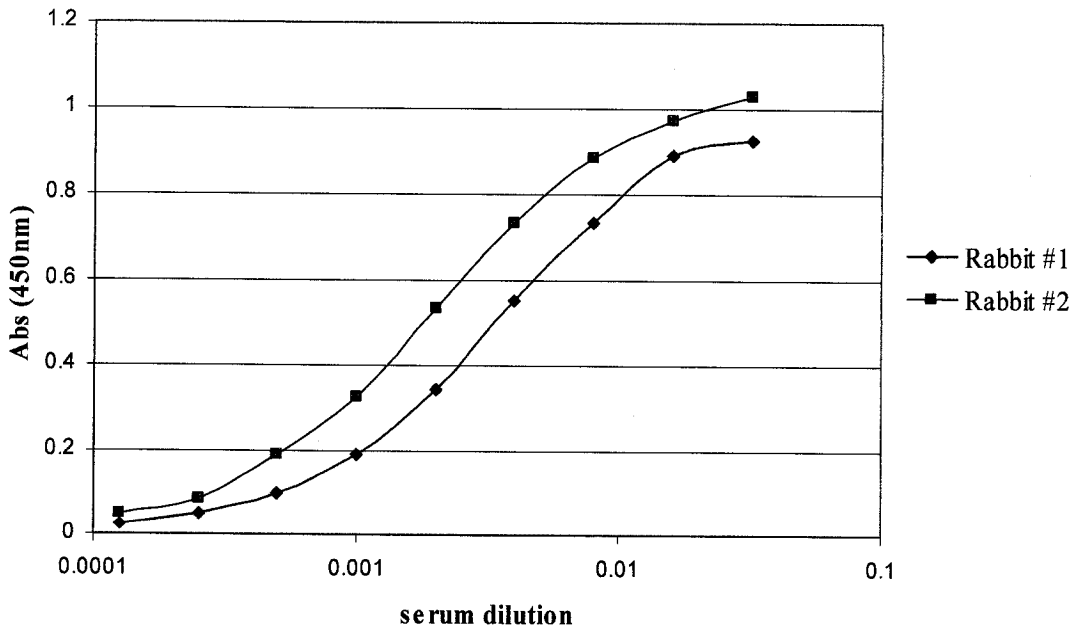
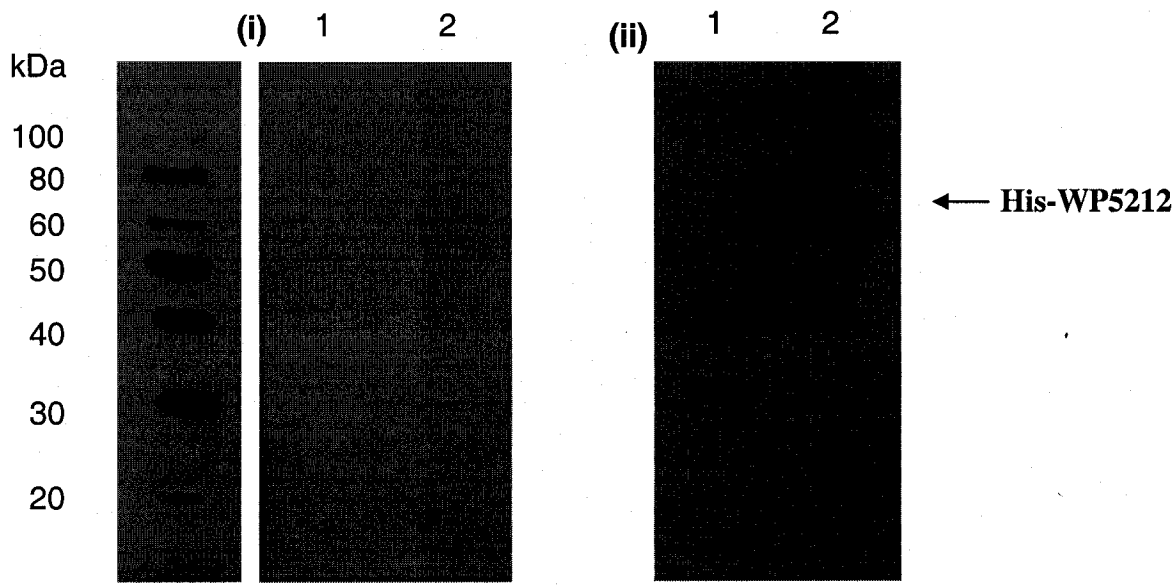

1.2
1
0.8
0.6
0.4
0.2
0
Abs (450nm)
0.0001
0.001
0.01
0.1
serum dilution
Rabbit #1
Rabbit #2

Figure 13. Western blot analysis of rabbit sera collected prior to and following DNA immunization. Antibodies from the final bleed (lane 2) detect recombinant his-tagged WP5212, whereas the pre-immunized sera (lane 1) fails to bind to this protein. Rabbit #1 (i) has slightly less anti-WP5212 antibody reactivity than Rabbit #2 (ii). The MagicmarkX (Invitrogen) was used as the molecular mass standard.


(i)
1
2
kDa
100
80
60
50
40
30
20
(ii)
1
2
His-WP5212

Study 4 – Immunoscreening a wheat cDNA expression phage library to identify additional candidate diabetes-related wheat proteins.

Increasing evidence indicates that wheat can modify the development of diabetes in animal models. Furthermore, recent epidemiological studies also implicate wheat as a modulator of T1D in humans (50, 51). Recently, a wheat storage protein WP5212 (a Glb1 homologue) was identified as the first candidate diabetes-related protein, by screening a wheat cDNA expression library with antibodies from diabetic BioBreeding rats (42). This demonstrated proof of concept that antibodies from diabetic individuals can be used to discover diabetes-related wheat proteins. In order to identify additional putative diabetes-related wheat proteins, the wheat cDNA library was screened with serum from a highly-wheat sensitive patient with both type 1 diabetes and celiac disease.

Characterization of the wheat cDNA expression library. The unamplified ZAP Express wheat cDNA expression library was comprised of approximately 1.0 $x10^6$ recombinant phages. The titer of the amplified library was approximately 8.5 x 10^7 pfu/ml, of which 1.4% were non-recombinant (**Table 11**). To characterize the wheat cDNA expression library used to identify WP5212, 30 clones were randomly selected. Phagemid DNA was isolated, restriction digests were performed to verify the presence of an insert and inserts were subsequently sequenced. These clones were termed WPran1 through WPran30. Twenty-nine clones were successfully sequenced, while only one clone (WPran29) failed the identification process.

Table 11. Titer of the amplified wheat cDNA expression library

Trial	Titer (x 10^7 pfu/ml)	% non-recombinant
1	5.5	1.8
2	8.6	0
3	9.1	0
4	10.6	2.8
Mean ± SD	8.45 ± 2.14	1.4 ± 1.3

The BLAST homology search results are summarized in **Table 12**. Approximately one third (10/29) of the sequenced clones share nucleic acid sequence homology with high molecular weight (HMW) glutenins. Two clones shared homology with gliadins. Nine clones shared homology with plant proteins of low molecular weights, including storage proteins, structural proteins and enzymes. Eight clones shared nucleic acid sequence homology with DNA which code for unidentified plant proteins. Two clones share low homology with mammalian proteins (E-values: 0.22 and 5e-5). Sequencing of these randomly selected clones indicates that both high and low molecular weight proteins are well represented in this library.

Table 12. Identification of random clone by BLAST homology searches

Clone	DNA homology	Accession Number	%identity/no. bp	E value
WPran1	*Triticum turgidum* HMW-glutenin locus	AY368673.1	89%/412	e-113
WPran2	*T. turgidum* HMW glutenin locus	AY494981.1	58%/167	2e-56
WPran3	*Ades atropalpus* transposon	AY079070.1	96%/52	2e-14
WPran4	*T. aestivum* cDNA, mRNA sequence		89%/39	
WPran5	*T. turgidum* HMW glutenin locus	AY368673.1	85%/286	2e-58
WPran6	*T. turgidum* HMW glutenin locus	AY494981.1	84%/509	e-105
WPran7	*T. monococcum* BAC clones 116F2 & 115G1	AF459639.1	82%/372	6e-53
WPran8	*T. turgidum* HMW-glutenin locus	AY369673.1	88%/472	e-135
WPran9	*T. aestivum* Gamma-gliadin	AF234646.1	81%/678	0.0

WPran10	*T. turgidum* HMW-glutenin locus *Aegilops tauschii* transposon	AY368673.1 SEG_AY534122S2	91%/501 91%/502	e-175 e-180
WPran11	*T. aestivum*	TC196237 AL816276	76%/350 91%/422	9.7e-36 7.1e-76
WPran12	*T. turgidum* HMW glutenin locus	AY494981.1	86%/160	1e-28
WPran13	*Homo sapiens* Dolichyl-phosphate beta-glucosyltransferase	AF102850.1	100%/27	5e-4
WPran14	*T. monococcum* Actin (ACT-1)	AF326781.1	85%/106	1e-15
WPran15	*T.aestivum* mRNA for puroindoline-b	TAFRIABIL	99%/563	0.0
WPran16	*T. monococcum* phosphatidylserine decarboxylase	AY485644.1	96%/911	0.0
WPran17	*T. aestivum* Itrl-2 mRNA for CMx	X75609.1|YACMX 2	93%/573	0.0
WPran18	*T. turgidum* HMW-glutenin locus	AY368673.1	85%/416	5e-98
WPran19	*Mus musculus* Chromosome 3, clone RP24-199C12	AC102259.7	100%/22	0.22
WPran20	Rice mRNA for aspartic protease	D32144.1 RICAPA	81%/266	2e-27
WPran21	*T. aestivum* Beta amylase mRNA	AF470353.1	99%/501	0.0
WPran22	*T. turgidum* HMW- glutenin gene locus	AY494981.1	91%/468	e-172
WPran23	*Hordem vulgare* Calreticulin (CRH2) mRNA	BLYCRH2A	93%/193	6e-66
WPran24	*Aegilops tauschii* transposon Caspar	SEG_AY534122S2	97%/500	0.0
WPran25	*Oryza sativa* cDNA clone: J023111G13	AK100655.1	83%/476	1e-83
WPran26	*T. monococcum* Actin (ACT-1) gene	AF326781.1	92%/297	e-115
WPran27	*H. vulgare* mRNA	Z69631.1	85%/333	1e-74
WPran28	*T. turgidum* HMW-glutenin locus	AY494981.1	92%/535	0.0
WPran30	*T. aestivum* Alpha-gliadin storage protein gene	U51303.1	99%/285	e-151

An antigenic wheat clone was isolated from the wheat cDNA expression library. Primary screening of the wheat cDNA expression library using serum from the case study subject, who is highly wheat-sensitive and has both T1D and CD, revealed 62 positive clones. Through secondary screening, three of these clones were designated positive compared to non-specific antibody binding. These clones were screened repeatedly until they reached clonality. Restriction digests of phagemid DNA revealed three different size inserts,

approximately 1.6 kb, 1.3 kb and 0.9 kb (**Figure 14**). Upon sequencing the cDNA inserts, it became apparent that these DNA sequences code for portions of the same protein, granule-bound starch synthase I (GbssI). Nucleotide and translated BLAST searches of the GenBank were performed (**Table 13**).

Table 13. Identification of the inserts of the three positive clones by homology searches

Clone	DNA homology (% identity/no. bp)	Amino acid homology (% identity/no. amino acids)	ORF length (bp)	Putative protein length (amino acids)
32112	*T. aestivum* (wheat) granule-bound starch synthase I (GbssI) gene 99%/1454	*T. aestivum* (wheat) granule-bound starch synthase I (GbssI) 99%/394	1182	394
5012	*T. aestivum* (wheat) granule-bound starch synthase I (GbssI) gene 94%/1003	*T. aestivum* (wheat) granule-bound starch synthase I (GbssI) 96%/303	909	303
5021	*T. aestivum* (wheat) granule-bound starch synthase I (GbssI) gene 97%/611	*T. aestivum* (wheat) granule-bound starch synthase I (GbssI) 99%/202	606	202

Clone WP32112 contains a 1184 bp open reading frame (ORF) which is in frame with the β-galactosidase ORF. The WP32112 ORF shares 99% homology across 1184 nucleotides with the *Triticum aestivum* granule-bound starch synthase I (GbssI) gene (accession number AF286320.1, **Figure 15**). The WP32112 ORF translates into an expected amino acid sequence of 394 amino acids which shares 99% homology with *T. aestivum* Gbss I (accession number AAG27624.1, **Figure 16**). An amino acid substitution occurs at amino acid 299 M→V and a second substitution at amino acid 309 H→R.

Figure 14. Antibodies from the highly-wheat reactive case study subject with both T1D and celiac disease bind to three clones in the wheat cDNA expression library. Restriction digest analysis of the isolated clones using *Eco*RI and *Xba*I reveals three distinct insert sizes, approximately 1.6 kb, 1.3 kb and 0.9 kb. The 4.6 kb band represents the pBK-CMV plasmid.

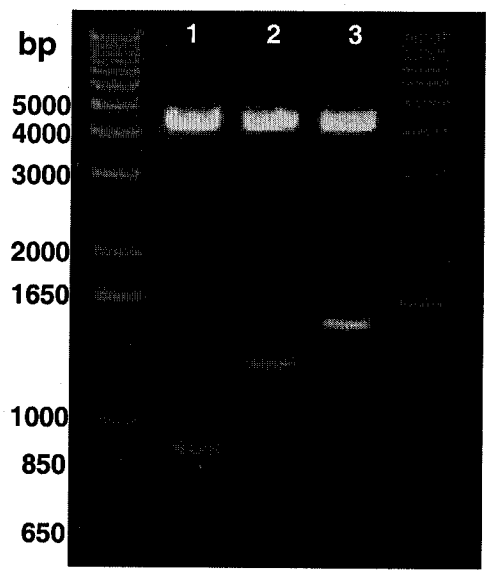

bp
1
2
3
5000
4000
3000
2000
1650
1000
850
650

Figure 15. DNA sequence homology between WP32112 and GbssI. *Panel A*: the ORF of clone WP32112 shares 99% homology with T. aestivum GbssI gene. Homologous nucleotides in the GbssI DNA sequence (AF286320.1) are denoted by dots. The open reading frame of WP32112 is indicated by the burgundy letters. *Panel B*: regions of homology between the nucleotide sequences are indicated by the blue bars. Alignments were preformed using ClustalW (www.ebi.ac.uk/clustalw/).

A.

```
WP32112        1    gcaacccatacttctccggaccctacggggaagacgtggtgttcgtgtgcaatgactggcacacgggccttctggcctgctacctcaagagcaactaccagtccagtggcatctatagga
AF286320.1     835  a.....................................................c.....................................................

WP32112        121  cggccaaggtagcgttctgcatccacaacatctcgtatcagggccgcttctccttcgacgacttcgcgcagctcaacctgcccgacaggttcaagtcgtccttcgacttcatcgacggct
AF286320.1     955  .....................................................................................................

WP32112        241  acgacaagccggtggaggggcgcaagatcaactggatgaaggccgggatcctgcaggccgacaaggtgctcacggtgagcccctactacgcggaggagctcatctccggcgaagccaggg
AF286320.1     1075 .....................................................................................................

WP32112        361  gctgcgagctcgacaacatcatgcgcctcacgggcatcaccggcatcgtcaacggcatggacgtcagcgagtgggaccccgccaaggacaagttcctcgctgccaactacgacgtcacca
AF286320.1     1195 .....................................................................................................

WP32112        481  ccgcgttggaggggaaggcgctgaacaaggaggcgctgcaggccgaggtggggctgccggtggaccggaaggtgccctggtggccttcatcggcaggctggaggagcagaagggccccg
AF286320.1     1315 .....................................................................................................

WP32112        601  acgtgatgatcgccgccatcccggagatcttgaaggaggaggacgtccagatcgttctcctgggcaccgggaagaagaagtttgagcggctgctcaagagcgtggaggagaagttcccga
AF286320.1     1435 .....................................................................................................

WP32112        721  gcaaggtgagggccgtggtcaggttcaacgcgccgctggctcaccagatgatggccggcgccgacgtgctcgccgtcaccagccgcttcgagccctgcggcctcatccagctccaggggga
AF286320.1     1555 .....................................................................................................

WP32112        841  tgcgctacggaacgccgtgcgcgtgcgcgtccaccggcgggctcgtcgacacgatcgtggagggcaagaccgggttccacatgggccgcctcagcgtcgactgcaacgtggtggagccgg
AF286320.1     1675 .....................................................a.....................a...............................

WP32112        961  ccgacgtgaagaaggtggtgaccaccctgaagcgcgccgtcaaggtcgtcggcacgccagcctaccatgagatggtcaagaactgcatgatccaggatctctcctggaaggggccagcca
AF286320.1     1795 .....................................................................................................

WP32112        1081 agaactgggaggacgtgcttctggaactgggggtcgaggggagcgagccaggggtcatcggcgaggagattgcgccgctcgccatggagaacgtcgccgctccctgaagagaggaaagag
AF286320.1     1915 .....................................................................................................

WP32112        1201 gtttttggtgcatggagcattcatccatcttcagggttcccgtgtgaggagatagccgcttgttgtagtgaagggccgatatatacatatttatatatagactaataagtacttatgctt
AF286320.1     2035 .....................................................................................................

WP32112        1321 tgttgtgccgcttgcctctttattacaaac-aaaaaaaaggtaagaagttaggggttgtgcttgtgagagtgtgtacttaatcgtgcttgcattttggtgtggtatattgcaaataaaca
AF286320.1     2155 ..............................  .....................................................................

WP32112        1441 aaggatttgttatgt
AF286320.1     2275 ...............
```

B.

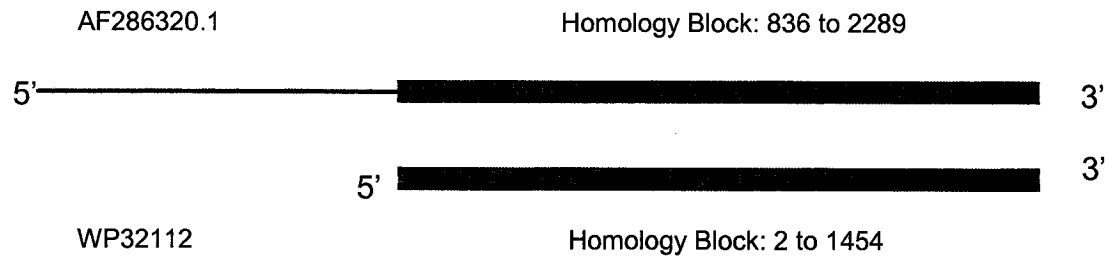


Figure 16. Homology between the predicted amino acid sequence of the isolated clones and the GbssI protein. Amino acid substitutions from the GbssI protein sequence are indicated in red. All the three clones share high sequence homology with GbssI, and consequently clone WP32112 was used for all additional analysis and characterization. Alignments were performed using ClustalW (www.ebi.ac.uk/clustalw/).

```
WP5021      ------------------------------------------------------------
WP5012      ------------------------------------------------------------
WP32112     ------------------------------------------------------------
GbssI       MAALVTSQLATSGTVLGITDRFRRAGFQGVRPRSPADAPLGMRTTGASAAPKQQSRKAHR 60

WP5021      ------------------------------------------------------------
WP5012      ------------------------------------------------------------
WP32112     ------------------------------------------------------------
GbssI       GTRRCLSMVVRATGSAGMNLVFVGAEMAPWSKTGGLGDVLGGLPPAMAANGHRVMVISPR 120

WP5021      ------------------------------------------------------------
WP5012      ------------------------------------------------------------
WP32112     ------------------------------------------------------------
GbssI       YDQYKDAWDTSVVSEIKVADEYERVRYFHCYKRGVDRVFVDHPCFLEKVRGKTKEKIYGP 180

WP5021      ------------------------------------------------------------
WP5012      ------------------------------------------------------------
WP32112     ----------------------------NPYFSGPYGEDVVFVCNDWHTGLLACYLK 29
GbssI       DAGTDYEDNQLRFSLLCQAALEAPRILDLNNNPYFSGPYGEDVVFVCNDWHTGLLACYLK 240

WP5021      ------------------------------------------------------------
WP5012      ------------------------------------------------------------
WP32112     SNYQSSGIYRTAKVAFCIHNISYQGRFSFDDFAQLNLPDRFKSSFDFIDGYDKPVEGRKI 89
GbssI       SNYQSSGIYRTAKVAFCIHNISYQGRFSFDDFAQLNLPDRFKSSFDFIDGYDKPVEGRKI 300

WP5021      ------------------------------------------------------------
WP5012      --MKAGILQADKVLTVSPYYAEELISGEARGCELDNIMRLTGITGIVNGMDVSEWDPIKD 58
WP32112     NWMKAGILQADKVLTVSPYYAEELISGEARGCELDNIMRLTGITGIVNGMDVSEWDPAKD 149
GbssI       NWMKAGILQADKVLTVSPYYAEELISGEARGCELDNIMRLTGITGIVNGMDVSEWDPAKD 360

WP5021      ------------------------------------------LEEQKGPDVMIAAIPEI 17
WP5012      KFLTVNYDVTTALEGKALNKEALQAEVGLPVDRKVPLVAFIGRLEEQKGPDVMIAAIPEI 118
WP32112     KFLAANYDVTTALEGKALNKEALQAEVGLPVDRKVPLVAFIGRLEEQKGPDVMIAAIPEI 209
GbssI       KFLAANYDVTTALEGKALNKEALQAEVGLPVDRKVPLVAFIGRLEEQKGPDVMIAAIPEI 420

WP5021      LKEEDVQIVLLGTGKKKFERLLKSIEEKFPSKVRAVVRFNAPLAHQMMAGADVLAVTSRF 77
WP5012      VKEEDVQIVLLGTGKKKFERLLKSVEEKFPTKVRAVVRFNAPLAHQMMAGADVLAVTSRF 178
WP32112     LKEEDVQIVLLGTGKKKFERLLKSVEEKFPSKVRAVVRFNAPLAHQMMAGADVLAVTSRF 269
GbssI       LKEEDVQIVLLGTGKKKFERLLKSVEEKFPSKVRAVVRFNAPLAHQMMAGADVLAVTSRF 480

WP5021      EPCGLIQLQGMRYGTPCACASTGGLVDTIVEGKTGFHMGRLSVDCNVVEPADVKKVVTTL 137
WP5012      EPCGLIQLQGMRYGTPCACASTGGLVDTIVEGKTGFHMGRLSVDCNVVEPADVKKVVTXL 238
WP32112     EPCGLIQLQGMRYGTPCACASTGGLVDTIVEGKTGFHMGRLSVDCNVVEPADVKKVVTTL 329
GbssI       EPCGLIQLQGMRYGTPCACASTGGLVDTIMEGKTGFHMGHLSVDCNVVEPADVKKVVTTL 540

WP5021      KRAVKVVGTPAYHEMVKNCMIQDLSWKGPAKNWEDVLLELGVEGSEPGVIGEEIAPLAME 197
WP5012      KRAVKVVGTPAYHEMVKNCMIQDLSWKGPXKNWEDVLLELGVEGSEPGIVGEEIAPLALE 298
WP32112     KRAVKVVGTPAYHEMVKNCMIQDLSWKGPAKNWEDVLLELGVEGSEPGVIGEEIAPLAME 389
GbssI       KRAVKVVGTPAYHEMVKNCMIQDLSWKGPAKNWEDVLLELGVEGSEPGVIGEEIAPLAME 600

WP5021      NVAAP 202
WP5012      NVAAP 303
WP32112     NVAAP 394
GbssI       NVAAP 605
```

Clone WP5012 has an ORF of 909 bp and, although it is not in frame with the β-galactosidase ORF, a start codon is present in the position 4-6 of the insert. This ORF shares 95% homology across 909 nucleotides with *T. aestivum* GbssI gene (accession number AF286320.1). Clone WP5012 has an expected product size of 303 amino acids, sharing 96% homology with *T. aestivum* GbssI (accession number AAG27624.1) and 96% homology with the expected amino acid sequence of clone WP32112. The protein alignment reveals there are eleven amino acid substitutions throughout the sequence (**Figure 16**).

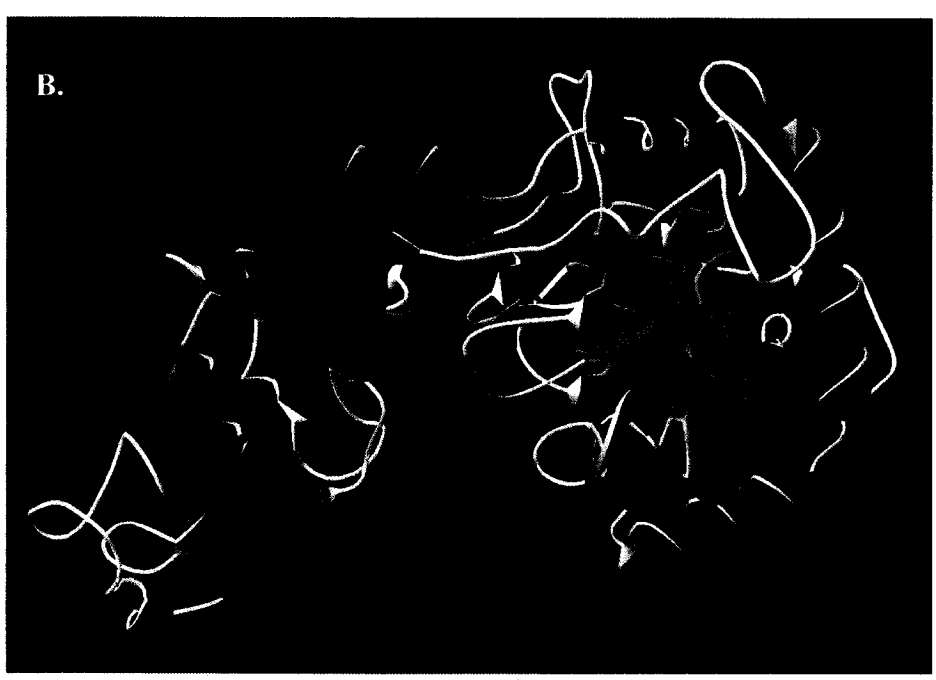
Clone WP5021 has an ORF of 611 bp coding for a product of 202 amino acids. The nucleotide sequence shares 97% homology across 608 nucleotides with *T. aestivum* GbssI gene (accession number AF286320.1). There is 99% identity between the translated amino acid sequence and *T. aestivum* GbssI (accession number AAG27624.1). The translated WP5021 nucleic acid sequence shares 99% homology across 202 amino acids with WP32112 (**Figure 16**). An additional amino acid substitution occurs at position 42 from V→I. Consequently WP32112, WP5012 and WP5021 are all truncated forms of the GbssI protein. All subsequent analyses were performed using WP32112.

The translated amino acid sequence of WP32112 was submitted to SWISS-MODEL to solve the putative three-dimensional structure of this protein. The amino acid sequence for WP32112 was aligned with the template three-dimensional structures of glycogen synthase alone and in complex with ADP, from *Agrobacterium tumefaciens* (PDB identification 1RZV and 1RZU, respectively). The solved three-dimensional structure of WP32112 is illustrated **in Figure 17** overlaid on the known structure of the *A. tumefaciens* glycogen

synthase. The predicted structure of WP32112 contains 16 α-helices and 11 β-sheets interspersed with random coils (**Figure 17**).

Figure 17. Three-dimensional modeling of the WP32112 predicted amino acid sequence. *Panel A*: the predicted three-dimensional model of WP32112 superimposed on known structure of glycogen synthase from *A. tumefaciens*. *Panel B*: α-helices (pink) and β-sheets (blue) are separated by random coils. The translated amino acid sequence of WP32112 was submitted to SWISS-MODEL (http://swissmodel.expasy.org/) and the predicted three-dimensional structure was viewed with Deep View-SwissPDB Viewer (http://ca.expasy.org/spdbv/).


A.

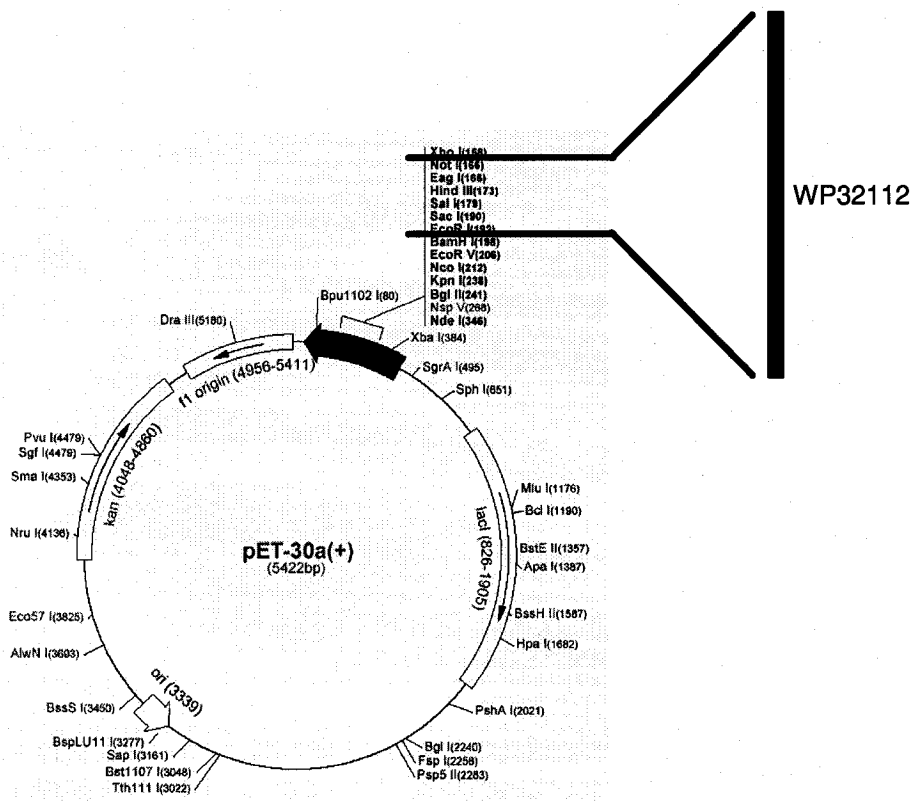

B.

Study 5 – Antibody reactivity to the candidate diabetes-related wheat protein in T1D, CD and T1D/CD patients compared to healthy controls.

Expression of his-tagged WP32112 in BL21 E. coli cells. The WP32112 cDNA sequence was amplified by PCR (**Figure 18**) using primers flanked by *Eco*RI at 5' and *Not*I at 3' restriction sites. The pBK-CMV phagemid DNA from the WP32112 clone was used as DNA template. This 1.2 kb sequence was cloned into the multiple cloning site of the pET-30a(+) vector using the aforementioned restriction sites. BL21 *E. coli* were transformed with the construct and plated on selective LB plates containing 50 μg/ml kanamycin. Eight random clones were selected and plasmid DNA was isolated from an overnight culture. The presence of the insert was verified by restriction digests of plasmid DNA with *Eco*RI and *Not*I. Clone #4 contained the insert. The orientation of the insert was verified by restriction mapping with *Nco*I, which cuts both the vector and the insert once (**Figure 18**). The integrity of the translated amino acid sequence was confirmed by sequencing (**Figure 19**). To verify WP32112 protein expression, SDS-PAGE and immunoblotting of whole cell lysate were performed. Coomassie staining revealed over-expression of a 50 kDa protein, the expected molecular weight of the his-tagged WP32112 protein, compared to control *E. coli* (**Figure 20**). Both monoclonal mouse anti-his antibody and antibodies from the case study subject bound to this protein (**Figure 20**), supporting the identity of the over-expressed 50 kDa protein as WP32112. The protein was subsequently purified from cell lysate using Ni-Sepharose-G. SDS-PAGE and Western blot analysis of the purification steps indicated that the his-tagged WP32112 protein was semi-purified (>70% pure) (**Figure 21**).

Figure 18. Construction of the pET-30a(+)-32112 expression vector. *Panel A*: The WP32112 nucleotide sequence was amplified by PCR and cloned into the NotI/EcoRI cloning site of the pET-30a(+) vector (Novagen). *Panel B*: Restriction mapping of the pET-30a(+)-32112 vector using *Nco*I. The pET-30a(+) vector and the WP32112 insert each contain one *Nco*I restriction site, and digestion with this enzyme produces two DNA fragments of the predicted sizes, 5.5 kb and 1.2 kb.

A.


WP32112
Xho I(158)
Not I(166)
Eag I(166)
Hind III(173)
Sal I(179)
Sac I(190)
EcoR I(193)
BamH I(198)
EcoR V(206)
Nco I(212)
Kpn I(238)
Bgl II(241)
Nsp V(288)
Nde I(346)
Bpu1102 I(80)
Dra III(5180)
Xba I(384)
SgrA I(495)
Sph I(651)
f1 origin (4956-5411)
Pvu I(4479)
Sgf I(4479)
Sma I(4353)
kan (4048-4860)
Nru I(4136)
pET-30a(+)
(5422bp)
lacI (826-1905)
Mlu I(1176)
Bcl I(1190)
BstE II(1357)
Apa I(1387)
BssH II(1587)
Hpa I(1682)
Eco57 I(3825)
AlwN I(3693)
ori (3339)
BssS I(3450)
BspLU11 I(3277)
Sap I(3161)
Bst1107 I(3048)
Tth111 I(3022)
PshA I(2021)
Bgl I(2240)
Fsp I(2258)
Psp5 II(2283)

B.

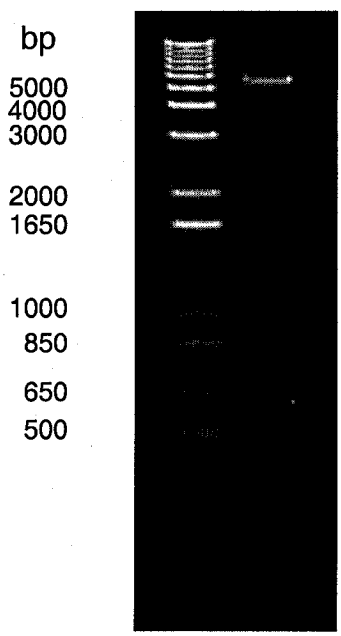

bp
5000
4000
3000
2000
1650
1000
850
650
500

Figure 19. Verification of the cloned WP32112 amino acid sequence integrity. Sequence alignment of the WP32112 translated sequence cloned into the pET-30a(+) vector with the original WP32112 translated sequence in the pBK-CMV vector of the wheat cDNA expression library. The alignment was performed by ClustalW (www.ebi.ac.uk/clustalw/).

```
Insert:   53  NPYFSGPYGEDVMFVCNDWHTGLLACYLKSNYQSSGIYRTAKVAFCIHNISYQGRFSFDD 112
WP32112:   1  ************V********************************************** 60

Insert:  113 FAQLNLPDRFKSSFDFIDGYDKPVEGRKINWMKAGILQADKVLTVSPYYAEELISGEARG 172
WP32112:  61 ************************************************************ 120

Insert:  173 CELDNIMRLTGITGIVNGMDVSEWDPAKDKFLAANYDVTTALEGKALNKEALQAEVGLPV 232
WP32112: 121 ************************************************************ 180

Insert:  233 DRKVPLVAFIGRLEEQKGPDVMIAAIPEILKEEDVQIVLLGAGKKKFERLLKSVEEKFPS 292
WP32112: 181 ******************************************T***************** 240

Insert:  293 KVRAVVRFNAPLAHQMMAGADVLAFTSRFEPCSLIQLQGMRYGTPCACASTGGLVDTIVE 352
WP32112: 241 ************************V*******G*************************** 300

Insert:  353 GKTGFHMGRLSVDCNVVEPADVKKVVTTLKRAVKVVGTPAYHEMVKNCMIQDLSWKGPAK 412
WP32112: 301 ************************************************************ 360

Insert:  413 NWEDVLLELGVEGSEPGVIGEEIAPLAMENVAAP 446
WP32112: 361 ********************************** 394
```

Figure 20. Expression of his-tagged WP32112 protein in *E. coli*. *Panel A*: Proteins from *E. coli* transformed with the pET-30a(+)-WP32112 vector (lane 2) and control *E. coli* (lane 1) separated by SDS-PAGE and visualized by Coomassie blue staining revealed over-expression of a 50 kDa protein. Prestained benchmark ladder (Invitrogen) was used as the protein standard. *Panel B*: Western blot analysis indicates that both anti-his antibodies (i) and serum from the case study subject (ii) bind to the over-expressed protein. MagicmarkX (Invitrogen) was used as the protein standard.

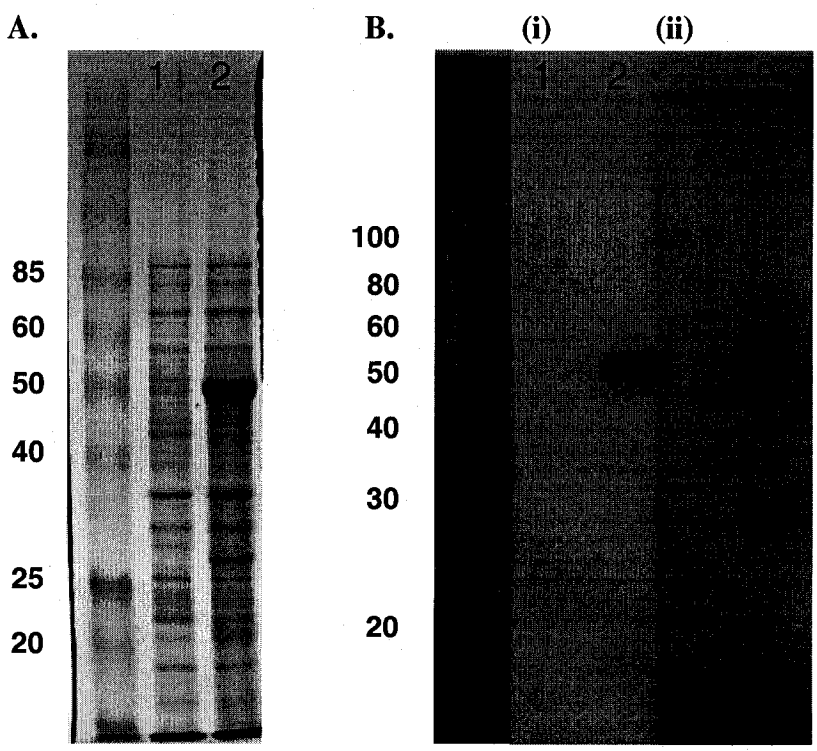
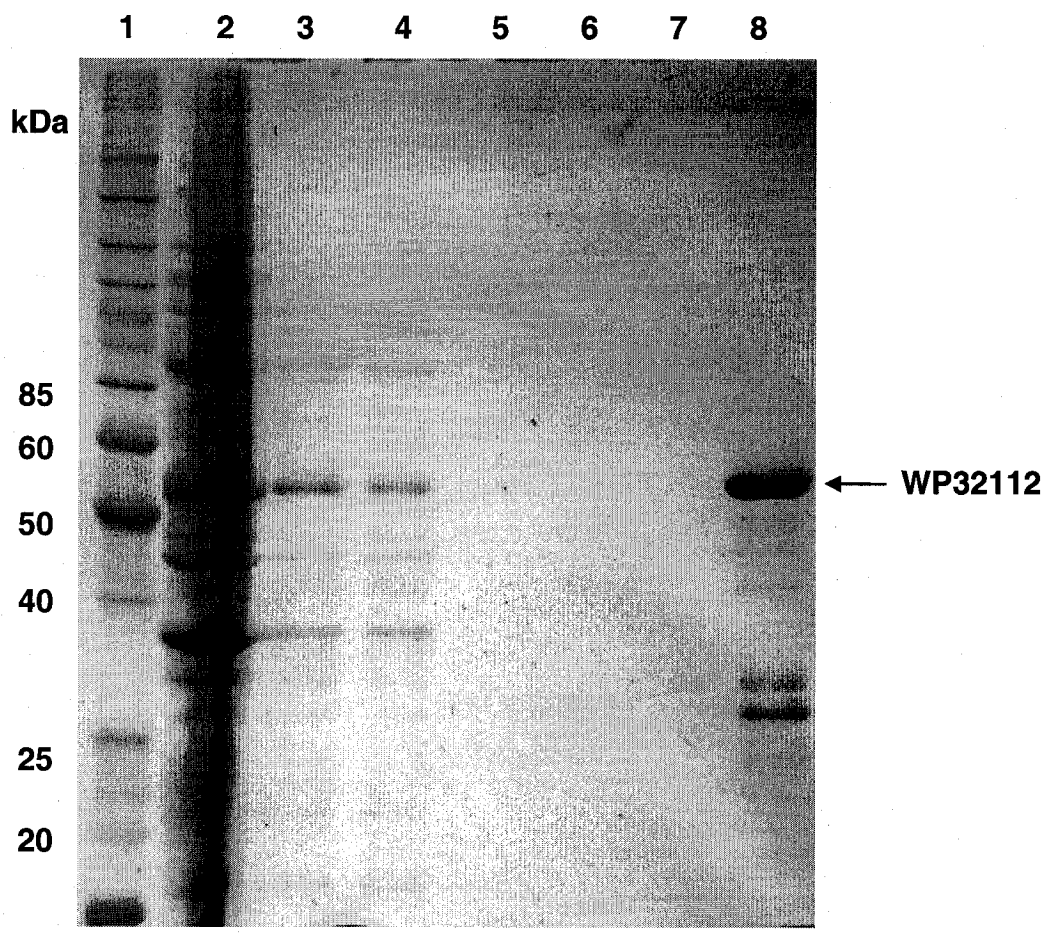

A.
1
2
85
60
50
40
25
20
B.
(i)
(ii)
1
2
100
80
60
50
40
30
20

Figure 21. Semi-purification of recombinant his-tagged WP32112 by Ni Sepharose-G pulldown. Proteins from each purification step were resolved by gel electrophoresis to monitor the purity of the recombinant WP32112. Prestained benchmark ladder (Invitrogen; lane 1), cell lysate pellet (lane 2), cell lysate supernatant (lane 3), flowthrough (lane 4), washes 1-3 (lanes 5-7), semi-purified WP32112 (lane 8).


1
2
3
4
5
6
7
8
kDa
85
60
50
40
25
20
WP32112

Patient reactivity against WP32112. WP32112 was identified using serum from the case study subject with both T1D and CD. Consequently, in order to determine if antibody reactivity against this protein correlated with either or both of the diseases, patients with T1D (n = 9), CD (n = 5) or T1D/CD (n = 7) and healthy controls (n = 9) were randomly selected from our study groups and screened for antibody reactivity against WP32112. The characteristics of these study groups are listed below in **Table 14**. The characteristics of individual subjects, including age, sex, age at T1D diagnosis and tissue transglutaminase autoantibody status are listed in **Table 15**.

Table 14. Characteristics of the human study groups (Study 5)

Subject Type	**n**	**Sex Ratio (M/F)**	**Mean age ± SD (years)**	**Mean age ± SD at T1D diagnosis (years)**
T1D	9	1/8	26.2 ± 5.8	17.6 ± 8.6
CD	5	1/4	27.4 ± 3.2	n/a
T1D/CD	7	1/6	27.0 ± 12.7	13.8 ± 6.5
Control	9	4/5	24.7 ± 2.4	n/a

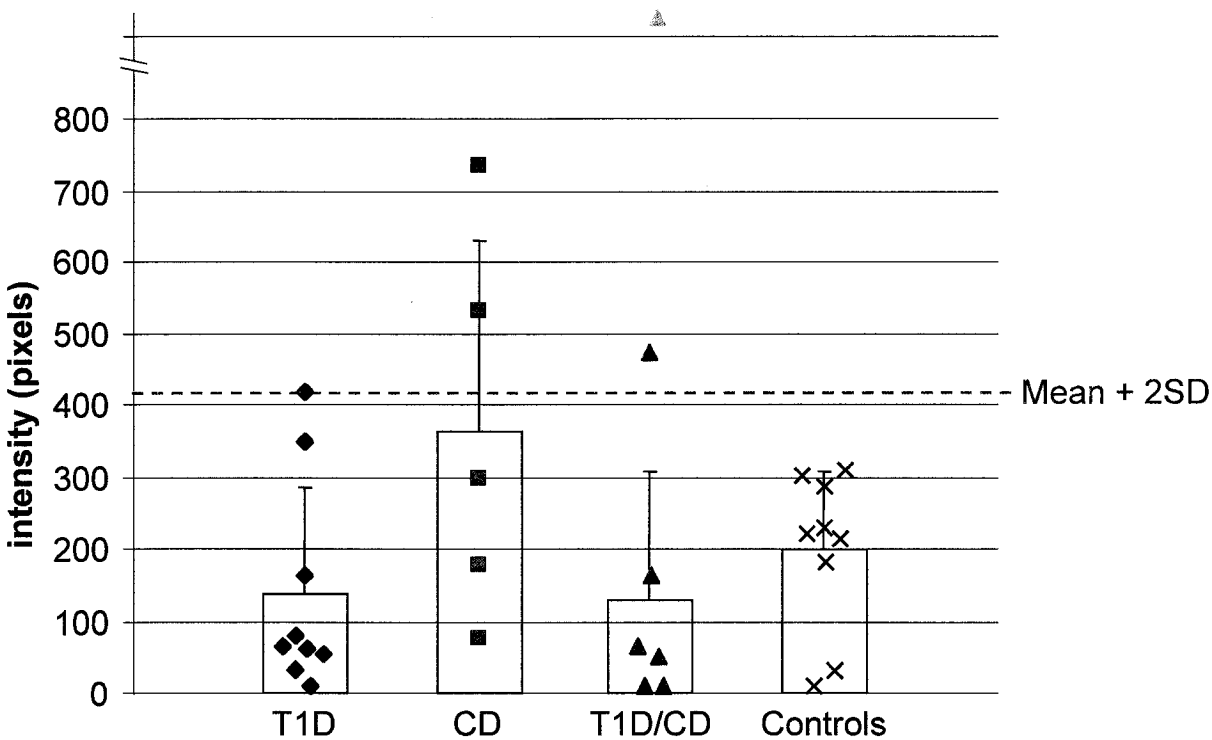
Table 15. Characteristics of the individual subjects in the human study (Study 5)

Disease status	Subject	Age	Sex	Age at T1D diagnosis (years)	tTG level	Anti-WP32112
Type 1 diabetes	T1D1	33	F	25	<1	-
	T1D2	19	F	16	<1	-
	T1D3	20	F	11	<1	-
	T1D4	33	F	26	2	-
	T1D5	34	F	29	<1	+
	T1D6	21	F	4	<1	-
	T1D7	25	F	18	<1	-
	T1D8	25	M	no data	<1	-
	T1D9	26	F	12	<1	-
Celiac disease	CD1	33	M	n/a	<1	-
	CD2	25	F	n/a	6	+
	CD3	27	F	n/a	<1	-
	CD4	26	F	n/a	7	+
	CD5	26	F	n/a	<1	-
Type 1 diabetes & Celiac disease	T1D/CD1	43	F	12	>100	+
	T1D/CD2*	28	F	18	7	+
	T1D/CD3	12	F	no data	16	-
	T1D/CD4	38	F	22	>100	-
	T1D/CD5	16	M	10	15	-
	T1D/CD6	37	F	17	>100	-
	T1D/CD7	15	F	4	no data	-
Healthy control	CON1	28	M	n/a	<1	-
	CON2	24	F	n/a	<1	-
	CON3	22	F	n/a	<1	-
	CON4	23	F	n/a	<1	-
	CON5	26	F	n/a	<1	-
	CON6	22	M	n/a	<1	-
	CON7	26	M	n/a	<1	-
	CON8	23	F	n/a	<1	-
	CON9	28	M	n/a	<1	-

A tissue transglutaminase autoantibody level <4 is negative, 4-10 is weakly positive, >10 is positive
Anti-WP32112 antibody level greater than mean+2SD of the control groups is considered positive
* NB: the antibody reactivity for the case study patient was not included in the graph or average of the T1D/CD as the antibody reactivity did not fall within the linear range

Semi-purified recombinant WP32112 was separated by one-dimensional gel electrophoresis and Western blotting was performed using serum from individual patients or control subjects. Densitometry analysis of the band corresponding to WP32112 was performed using the Kodak Digital Science TM image station 440CF, and is depicted in **Figure 22.**

Figure 22. Densitometry analysis of antibody reactivity against WP32112. 1D SDS-PAGE and Western blots of semi-purified his-tagged WP32112 transformed bacteria cell lysate were probed with serum antibodies from control subjects (n=9), patients with type 1 diabetes (n=9), celiac disease (n=5) or both diseases (n=6). Densitometric analyses of the Western blots were performed to determine antibody reactivity against WP32112. Individuals with antibody reactivity values greater than the mean + 2SD (417.3 mean intensity/pixel; red line) of the control group were designated WP32112 antibody positive. The differences were not significant (t-test). NB: The case study subject is depicted by the orange triangle, however, the densitometry value does not fall within the linear range.


intensity (pixels)
800
700
600
500
400
300
200
100
0
Mean + 2SD
T1D
CD
T1D/CD
Controls

The average antibody reactivity against WP32112 was higher but not significant in patients with CD alone compared with controls ($p = 0.12$). Interestingly, although not statistically significant, control subjects had a higher mean antibody reactivity against this wheat protein compared with patients with T1D ($p = 0.32$) and both diseases (T1D/CD; $p = 0.35$). Using the cut-off value of mean antibody reactivity of the control group plus two standard deviations (mean intensity/pixel + 2SD = 417.3), the following were considered anti-WP32112 antibody positive: 40% (2/5) of the CD patients, 28% (2/7, including the case study subject) of the T1D/CD patients, 11% (1/9) of the T1D patients and 0% (0/9) of the control subjects (**Table 15**, above). These difference were not significant (t-test) because of the small number of subjects ($n = 5$-9).

DISCUSSION

Importance of defining environmental cues associated with type 1 diabetes

For over thirty years, type 1 diabetes has been recognized as an autoimmune disease, as evidenced by studies in the 1970s demonstrating the presence of immune cell infiltrates in the pancreas, of autoantibodies in the blood of patients, and the association of immune system genes with the disease (reviewed (4)). T1D is a severe chronic autoimmune disease that results from insulin deficiency due to the destruction of the pancreatic β-cells. However, underlying triggers of the disease remain to be determined. The complexity of this disease, due to the influence of many risk genes and numerous environmental factors makes the causes extraordinarily difficult to establish. Furthermore, literature examining the importance of the environmental candidates is especially conflicting. Nonetheless, one way to prevent diabetes may simply be the avoidance or alteration of an environmental trigger. In support of this theory, results from a recent study which monitored the development of diabetes in high risk infants fed either casein hydrolysate or conventional cow's milk-based formulas have provided the first evidence that dietary intervention may indeed modulate the development of autoimmune diabetes in humans (84).

Prior to the discovery of insulin by Dr. Frederick Banting and Charles Best in 1921, the standard treatment for T1D was a restrictive diet; this often resulted in excessive weight loss in the patient and even starvation. While insulin is not a cure for diabetes nor does it impede or reverse the immune attack against the β-cells, it is nonetheless a life-saving treatment as it partially restores blood glucose levels to near normal. In recent years, groundbreaking advances have been made in the Edmonton protocol for pancreatic islet transplantation,

making this a desirable treatment option for brittle diabetic patients (85). However, this treatment is far for ideal. The availability of donor islets is limited, immunosuppressive drugs have several severe side effects and only 10% of patients remain free of insulin injections after five years (86). In addition, islets are injected into a milieu where environmental cues have primed the recipient's immune system to attack the new β-cells. Therefore, while these treatments are life-saving, it is important to devise strategies to understand the nature of these environmental cues, how they manipulate the immune system and how they can be modified.

A genetic predisposition is required for the development of T1D; however, fewer than 10% of individuals carrying risk genes will in fact develop diabetes (87). This suggests that environment accounts for a majority of diabetes risk. The identification of environmental factors is crucial to understand their implications in the pathogenesis of diabetes, and conversely, the elucidation of mechanisms involved in environment-associated diabetes will provide new insight into how these environmental insults can be identified. Currently, the identification of candidate environmental factors implicated in the development of human T1D relies heavily on epidemiological studies. Animal models of spontaneous diabetes, including the NOD mouse and diabetes-prone BB rat, are used to identify and confirm the diabetes-promoting capacity of environmental agents. These studies have linked various viral infections and dietary agents to disease development; however, in both cases results are inconclusive, possibly reflecting the multifactorial nature of the disease.

The immune system interacts with both gut pathogens and dietary agents through the gut-associated lymphoid tissue (GALT), which is responsible for eliciting the appropriate response, tolerance or immunity. There is increasing evidence to suggest that intestinal dysfunction and dysregulation of GALT plays an important role in the pathogenesis of diabetes. The diabetes-prone BB rat, a well-characterized model of spontaneous diabetes, shows evidence of histological gut damage, increased gut permeability and increased gut inflammation (reviewed in (88)). NOD mice display signs of mild enteropathy (89) and the gut mucosa of diabetic patients also appears to be abnormally permeable (90). Furthermore, a newly-defined link has been established between pancreatic lymph nodes (PLNs), where immune tolerance to pancreatic proteins is disrupted, and the gut. It has been demonstrated in the NOD mouse that both self-peptides from the islet β-cells and foreign molecules from the gut are presented in PLNs, and gut damage promotes the priming of diabetogenic T cells in the PLNs (91). Consequently, this link provides a potential route by which dietary agents and gut pathogens can influence the development of diabetes via the gastrointestinal tract.

Wheat as the major dietary contributor

Early studies have demonstrated the diabetes-promoting properties of wheat in both the NOD mouse and the diabetes-prone BB rat animal models (40, 44-47). In humans, recent studies have indicated that a window of time exists in which introduction to gluten is associated with a decreased risk of diabetes (50, 51). In addition, there appears to be an unusual immune response against wheat proteins in a subset of diabetic patients, shown by the presence of wheat antibodies and T cell reactivity against wheat proteins (53, 54, 57, 58, 61).

The ingestion of wheat has also been linked to gut damage, which may be a pivotal event in diabetes pathogenesis. A recent study has demonstrated that NOD mice fed a wheat-based diet develop small intestinal enteropathy associated with the development of this disease (89). Following weaning, diabetes-prone BB rats also display signs of immunologically-mediated enteropathy preceding the onset of diabetes; however, the nature of the diet appears to have little to no influence on the development of histological damage (43). Nonetheless, in this animal model, the wheat-based diet evokes alterations in the expression of intestinal mucins, decreasing total and neutral mucin levels (92). It has been demonstrated *in vitro* that exposure to gliadin disrupts barrier function by modifying the expression of apical junction proteins (93). Furthermore, increasing evidence *in vitro* suggests that wheat proteins can induce the signaling pathway of zonulin (94), a protein which modulates intestinal permeability by reversibly opening the tight junctions (95). The up-regulation of this protein has been shown to precede the development of diabetes and inhibition of the zonulin receptor reduced the incidence of diabetes in the BB rat (96). In addition, Sapone and colleagues (87) reported a relationship between the up-regulation of zonulin and increased intestinal permeability in human T1D patients.

The first identified diabetes-related wheat protein, WP5212, which shares homology with the storage protein Glb1, was identified by screening a wheat cDNA expression library with sera from diabetic BB rats (42). Homologous proteins are present in other cereal and vegetable plants, and this protein also shares sequence homology with known allergens and self proteins, including ZO-2. Further analysis in the BB rat indicated that higher antibody titers against this protein correlated with increased insulitis, the first stage of diabetes (42). Additional studies have demonstrated antibodies against this protein to be associated with

diabetes in three distinct species: rat, mouse and human (42, 54, 58). This groundbreaking discovery further strengthens the argument that certain wheat proteins may be unusually antigenic in diabetes-prone individuals and may be associated with disease pathogenesis. The research presented in this thesis aims to i) examine the potential role of the wheat protein WP5212 in the pathogenesis of diabetes and ii) identify additional candidate immunogenic wheat proteins.

Importance of conformational epitopes for the antigenicity of WP5212

The nature of the antigenic molecular structures of immunogenic proteins may provide valuable insight into the mechanism of their involvement in disease. Therefore the first goal of this research was to identify the immunodominant epitopes of WP5212. Immunodominant epitopes are either epitopes recognized by the majority of patients or responsible for the bulk of antibody binding in a single patient (97). Linear or continuous epitopes are generally detected by screening a library consisting of short overlapping peptides spanning the entire protein sequence. Following the identification of antibody-binding peptides, shorter peptides are used to determine the shortest epitope required for antibody recognition (70).

A peptide library comprising 64 peptide sequences overlapping the previous peptide by eight amino acids was constructed at SigmaGenosys. This library was initially screened by dot blots with pooled serum from eight diabetic patients categorized as anti-WP5212 antibody-positive. Antibodies from this patient pool recognized and bound to eight peptides revealing the antigenic portion of WP5212. These eight peptides of interest were screened with serum

from individual patients with T1D and from control subjects in order to determine their immunodominant epitopes. Serial dilutions of WP5212 were also included on the membrane as controls. While a higher frequency of diabetic patients exhibited antibody reactivity against three peptides (amino acids 163-318, 217-234 and 262-279) than control subjects, the differences were not statistically significant possibly due to the small number of subjects. Surprisingly, despite the small variation in antibody reactivity to these eight peptides, there was a remarkable difference in the antibody reactivity against the serial dilution of recombinant WP5212 between the two study groups. Anti-WP5212 activity in the serum of individual diabetic patients could be detected against 2-fold lower concentrations of recombinant WP5212 than in control subjects.

Although this technique failed to identify immunodominant linear epitopes, the results suggest that conformational epitopes or post-translational modifications may play a role in the antigenicity of WP5212. The integrity of the antibody-binding regions of WP5212 must be able to resist the harsh conditions of food processing and the gastrointestinal tract, by being thermostable, stable in low pH and resistant to proteases, as are the eptiopes of many food allergens (98). Conformational epitopes depend on the three dimensional structure of the protein. Consequently, these epitopes are more difficult to define since their identification requires "x-ray crystallography of the allergen-antibody complex" (70). Therefore, further investigation is required to identify these epitopes of the protein.

Molecular mimicry

One popular hypothesis on the process by which environmental insults may influence the development or progression of autoimmune disease is based on molecular mimicry. This theory is based on the redundancy of the lymphocyte receptors and, consequently, lymphocytes activated against a foreign antigen are unable to distinguish sequence or structural similarity from a self-protein. The ultimate result is that the host's immune system mounts an attack against self proteins and causes the destruction of target organ(s). It is possible, though not yet proven, that molecular mimics may be involved in triggering, accelerating or even attenuating autoimmune disease (99).

This elegant hypothesis was originally proposed in relation to diabetes to explain how viral infections may induce the onset of this autoimmune disease. Certain viral proteins share homology with known diabetes-related autoantigens. For instance, GAD65 shares a six amino acid sequence with the non-structural protein (P2C) of CVB 4 (100). However, while T cell cross-reactivity has been detected at bulk culture level (100) there remains a lack of evidence at the clonal level (101). Furthermore, T cell cross-reactivity has been demonstrated between GAD65 and the major DNA-binding protein (pUL57) of human cytomegalovirus (102). The molecular mimicry hypothesis was extended to include dietary proteins when it was discovered that an epitope contained in cow's milk bovine serum albumin, known as ABBOS, shares sequence identity with the pancreatic β-cell surface protein ICA69 (97). In addition, the candidate diabetes-related wheat protein WP5212 shares sequence homology with the tight junction protein ZO-2 (42). Furthermore, an antigenic WP5212 B cell epitope, recently identified by screening an epitope library using

serum from an extremely wheat-sensitive patient with both diabetes and CD, was found to share homology with both intermediate filament proteins desmin and vimentin and with glial fibrial acid protein (MacFarlane et al., unpublished). Molecular mimicry involving dietary agents has been suggested in several other autoimmune diseases, including celiac and Crohn's diseases (103, 104).

A novel approach

In addition to sequence homology, structural similarities may also be an important factor in molecular mimicry. For example, the degeneracy and flexibility of the T cell receptor (TCR) implies that for T cell cross-reaction only a small fraction of the 10-12 amino acid mimetic peptide needs to be shared with the self protein (99). Furthermore, antibodies can recognize both linear and conformational epitopes. Therefore, it is important to devise a strategy to identify corresponding mimics of environmental insults in target tissues which does not rely solely on amino acid sequence identity. The main goal of this research project was to develop a novel approach to identify candidate target proteins involved in wheat-induced diabetes. Antibodies from diabetic individuals have been an indispensable tool, in the identification of candidate environmental triggers/promoters such as viruses (29, 30) and the recently identified wheat protein WP5212 (42). They have also been crucial in the identification of autoantigens, including insulin, IA2 and GAD65, by various methods, including immunoblotting and immunoprecipitation (reviewed in (10)). We have adapted this methodology to investigate whether antibodies raised against a specific environmental insult can cross-react with proteins expressed in target tissue(s) to identify candidate molecular mimics.

Antibodies were generated in rabbits against two peptides of the recently identified diabetes-related wheat protein WP5212 which were considered antigenic by predictive algorithm (Sigma-Genosys). To determine whether self-proteins in target tissues share these predicted immunological epitopes, proteins isolated from target tissues were screened using this rabbit serum by Western blots. Immunoscreening identified antibody binding to a 33 kDa protein present in the exocrine pancreas and to a 28 kDa protein expressed in the pancreatic islets, spleen and MLNs.

Initial results from MALDI-TOF MS/MS analysis suggested that the identity of the 33 kDa protein to be HSP70. However, the lack of anti-HSP70 antibody binding to this protein extract indicated that the 33 kDa protein was not likely a member of the HSP70 family. To reveal the true identity of the 33 kDa protein, N-terminal sequencing by Edman degradation was performed. While sequencing revealed the presence of more than one sequence in the sample, Blast homology searches identified the peptide sequence FPLEDDDKIVGGYTCPE to map to amino acids 16-32 of *Rattus norvegicus* pancreatic trypsin 1. Although this peptide does not represent the N-terminus of newly-synthesized pancreatic trypsin 1, additional bioinformatics analysis revealed the presence of a signal peptide (amino acids 1-15) which would be cleaved. The predicted molecular weight and pI of this protein matches that of the 33 kDa protein. Furthermore, the MALDI-TOF MS analysis matched one 20-amino acid peptide to pancreatic trypsin 1. Although there was no significant amino acid sequence homology between pancreatic trypsin 1 and the two peptides used to generate the anti-sera, if key amino acids are conserved it is possible that the epitope can retain its structure required for antibody recognition. Immunohistochemisty provided further support

of the identification of the 33 kDa protein as a secreted protein. Pancreatic sections probed with the anti-WP5212 peptide antibodies revealed antibody binding to the exocrine portion of the pancreas in a "chicken feet" pattern (Wang et al., unpublished), characteristic of secreted proteins. It remains to be determined if this anti-WP5212 antibody binding co-localizes with pancreatic trypsin 1 antibody binding. Together these results suggest the antibodies generated against two short WP5212 peptides could be cross-reacting with pancreatic trypsin 1.

The antibody cross-reactivity between WP5212 and pancreatic trypsin 1 should be verified. *E. coli*, transformed with the construct encoding his-tagged pancreatic trypsin 1, over-expressed a 40 kDa protein of the expected size of the recombinant pancreatic trypsin 1. However, while anti-his antibodies bound to this over-expressed protein, confirming the presence of a his tag, the anti-WP5212 peptide antibodies did not recognize it. One major drawback associated with expression of eukaryotic proteins in a prokaryotic expression system is the absence of eukaryotic post-translational modifications. In addition, eukaryotic proteins may be misfolded in the prokaryotic system. Either of these factors can alter the immunogenicity of affected epitopes. Several recent studies have demonstrated the importance of post-translational modification for epitope recognition. For example, a disulphide bridge in an insulin A chain T cell epitope is required for T cell activation in diabetes (105), and gliadin needs to be deamidated to be immunogenic in celiac disease (66). Consequently it would be of interest to determine whether recombinant pancreatic trypsin 1 expressed in a eukaryotic system recovers the epitope recognized by anti-WP5212 peptide antibodies. In a future study, to avoid the difficulties associated with the verification of antibody cross reactivity, it would be beneficial to probe the Western blots with purified anti-

WP5212 antibodies. These antibodies can be separated from the serum by column chromatography using purified recombinant WP5212 protein.

Because diabetes is recognized as a T cell-mediated autoimmune disease, further studies are required to determine whether this potential antibody cross-reactivity between the wheat storage globulin WP5212 and pancreatic trypsin 1 reflects immunological cross-reactivity by T cells. Degeneracy at the level of the T cell receptor would strengthen the argument for the involvement of a mimicking structure shared between these two proteins in the pathogenesis of diabetes. This could be accomplished by determining whether T cells isolated from diabetic patients proliferate in response to pancreatic trypsin 1. If such T cells stimulated in response to pancreatic trypsin 1 are detected, it would be important to determine whether they can also proliferate in response to WP5212. The reciprocal experiment would confirm these findings.

Although this methodology is still in the early stages of development, preliminary results have demonstrated proof of concept that antibodies raised against an environmental agent can be used to identify candidate molecules in target tissues which share similar structures. This approach can be applied to identify proteins involved in environment-induced diabetes or other autoimmune diseases, based on the hypothesis that molecular mimicry plays a key role in the induction of autoimmunity. The identification of these candidate targets may help to elucidate novel pathways involved in the pathogenesis of autoimmune disease.

While immunological cross-reactivity is present, the concept of molecular mimicry remains very difficult to prove, and direct evidence supporting this hypothesis in the pathogenesis of

diabetes is lacking. In order to validate this hypothesis, it must be shown that this immunological cross-reactivity actually triggers the onset of disease. Many studies have demonstrated that inoculation with purified mimicking antigens in adjuvant can activate self-reactive lymphocytes and possibly induce autoimmune disease; however, these pathologies are characteristic of acute, not chronic, autoimmunity (99). A recent study has provided compelling evidence that molecular mimicry may be an initiating factor in multiple sclerosis. Croxford and colleagues (106) infected mice with a nonpathogenic virus encoding a mimic peptide, sharing sequence homology with the proteolipid protein epitope associated with encephalitis. The result was the onset of rapid organ-specific T cell-mediated autoimmunity leading to a non-progressive paralytic disease (106). These studies represent one of the few instances of documented molecular mimicry in autoimmune disease.

Generation of polyclonal antisera

Our initial experiments provided proof of concept that antibodies raised against WP5212 can be used to detect proteins in target tissues which may be related to diabetes. However, the anti-WP5212 peptide serum used was generated against two peptides, which represents only a small portion of the protein. Therefore, to improve upon this approach and to identify additional candidate targets, a new polyclonal serum was generated by DNA immunization. Antibodies produced in this manner should recognize both linear and conformational epitopes, representing the entire protein in its natural state (76). Furthermore, no adjuvant is required for this immunization procedure. This technique was ideal since, at the time, purified WP5212 was not available, and therefore the only other option was cell-based

immunization which would have generated antibodies to "irrelevant antigens of the injected cells" (76).

The DNA construct used for immunization was generated by cloning the WP5212 sequence into the pCI-neo mammalian expression vector (Promega). Thus, the WP5212 insert was flanked by the cytomegalovirus promoter, which has been shown to induce efficient gene expression (107) and the simian virus SV40 polyA signal to prevent readthrough transcription (107). While DNA can be delivered by several injection methods, including intramuscular, mucosal or by gene-gun (107) the rabbits were immunized by intradermal injection. This delivery method was selected because Chowdhury and colleagues (76) demonstrated a poor humoral response to intramuscular and subcutaneous injection in mice and reported the production of high titer serum in rabbits by DNA immunization delivered intradermally. The skin is an immune-competent organ containing dendritic cells which act as antigen-presenting cells. Following intradermal injection, the dendritic cells take up the DNA and express the protein and present peptides on both MHC class I and class II molecules (108). These cells can also present peptides from proteins taken up through phagocytosis of cellular debris from transfected cells which have undergone necrosis or apoptosis (109). The dendritic cells then migrate to the draining lymph nodes where they activate naïve T cells and induce immunity (109). A Th2 response is required for optimal B cell activation and IgG production (109).

Two New Zealand rabbits were given three intradermal injections of the pCI-neo-WP5212 construct at two week intervals. Anti-WP5212 antibody reactivity was monitored by ELISA, which established that both rabbits possessed anti-WP5212 antibodies subsequent to these

immunizations, with Rabbit #2 having higher antibody titers. In the weeks following these initial immunizations, the antibody titers did not drop as would be expected normally. One possible explanation could be that WP5212, a wheat protein, is likely present in the diet of the rabbits. Consequently, the rabbits would be constantly exposed to the protein, which would continuously stimulate their immune systems. This phenomenon would inhibit the decline of antibody titers following the initial immunizations, preventing the antibody titers to increase following the final boost, as would be anticipated for the secondary antibody response. These anti-WP5212 antibodies present in the sera of both rabbits were shown to recognize and bind to recombinant WP5212 by Western blot. This serum will be an invaluable tool which will improve the approach we have developed, with the potential to identify additional candidate molecular targets in the future.

Identification of an additional antigenic wheat protein

The identification of the first candidate diabetes-related protein, WP5212, was a crucial finding in determining potential environmental insults which may be involved in the pathogenesis of diabetes. This protein, discovered using serum form overt diabetic rats, was subsequently shown to be immunogenic in three distinct species: rat, mouse and human (42). The final objective of this research involved the identification of additional candidate antigenic wheat proteins. To accomplish this aim, the wheat cDNA expression library used to identify WP5212 was exploited for a second time to screen over one million recombinant phages expressing various wheat proteins. Because the expression library was already several years old, it was crucial to characterize it to determine the variability and complexity of the library. The sequencing of twenty-nine randomly selected clones revealed a wide

variety of wheat proteins were represented in this expression library, although the gliadins appeared to be under-represented in this small group (n = 29).

Physicians at the Ottawa Hospital alerted our lab that a 28 year old Caucasian woman, affected by both T1D and CD, was willing to participate in our study. She was highly wheat sensitive and ingestion of wheat induced a number of symptoms, including multiple bowel movements per day, oral ulcerations and severe lip swelling (57). Following her diagnosis with CD, she responded only transiently to the standard gluten-free diet. When a strict specific-carbohydrate, cereal-free diet was adopted, these symptoms were alleviated (57). Due to her extreme sensitivity to wheat, she became our index patient and has been vital in establishing the relationship between the wheat storage globulin WP5212 and diabetes in humans. Antibodies in her serum possess strong anti-WP5212 activity and T cells isolated from her peripheral blood mononuclear cells proliferate in response to this protein (57).

The extreme wheat sensitivity of this patient suggested that her serum should contain antibodies against several wheat proteins. Consequently the wheat cDNA expression library was screened with serum from this volunteer to identify candidate antigenic wheat proteins. Antibodies from this patient recognized and bound to three distinct clones. Subsequent BLAST homology searches identified all three wheat clones to code for truncated forms of the granule-bound starch synthase I protein (GbssI), an enzyme which incorporates glucose molecules into amylose and amylopectins of starch (110). The clone with the longest insert, WP32112, was selected for the ensuing analysis to ensure that all potential antigenic epitopes were included.

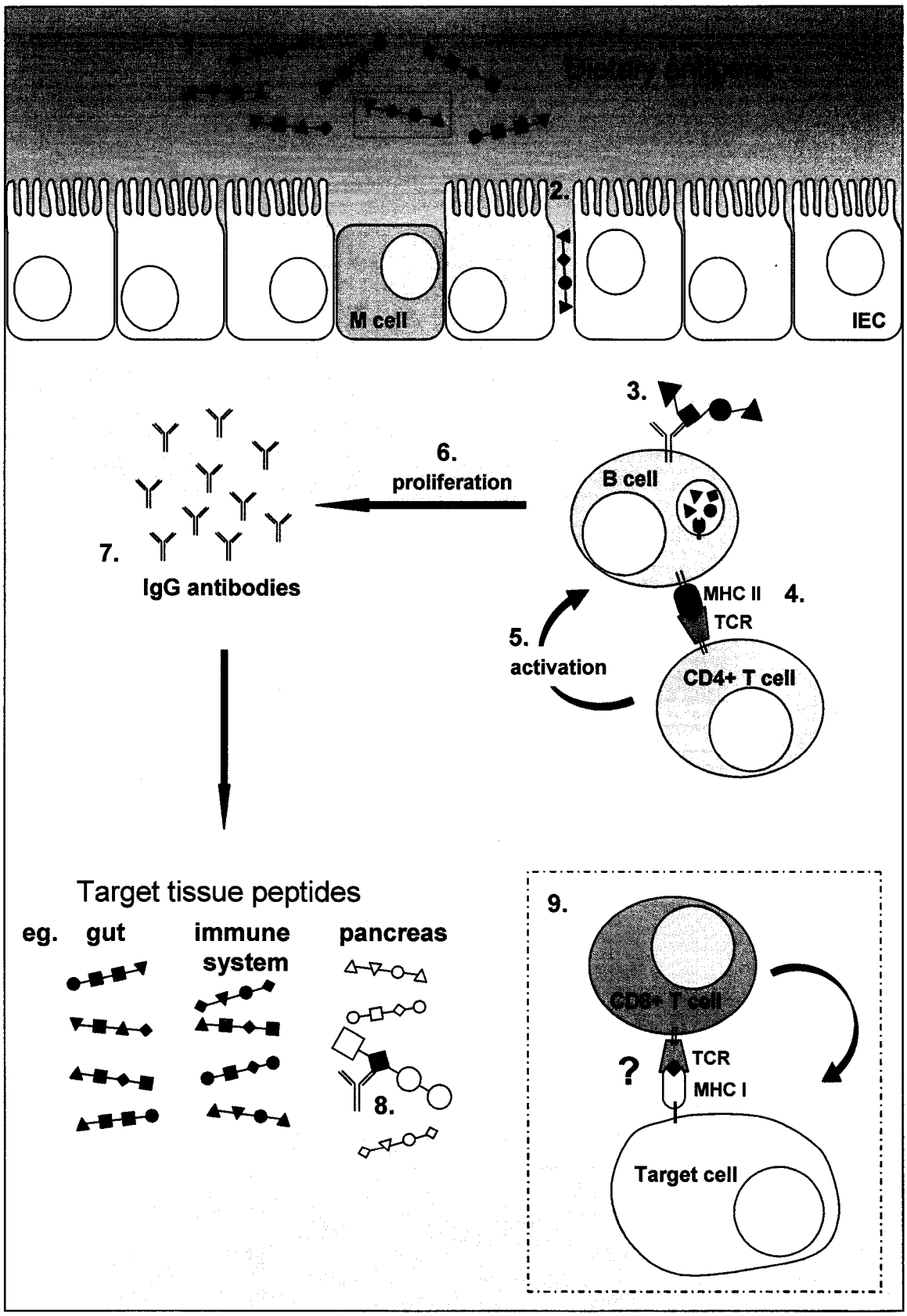
Because this protein was identified using serum from a patient with both diseases, it was necessary to determine whether antibody reactivity against WP32112 was associated with either T1D or CD or both diseases. In order to evaluate antibody titers more accurately, recombinant his-tagged WP32112 was expressed in *E. coli* and semi-purified with Ni-Sepharose G. The identity of the protein was confirmed by Western blot analysis using both anti-his antibodies and the serum from the case study patient. In order to evaluate the individual's reactivity against this protein, equal quantities of WP32112 were separated by 1D SDS-PAGE; this process ensures the standardization of the quantity of protein screened. Preliminary results from Western blot analysis using serum from individual patients with T1D, CD or both diseases and control subjects suggest that, while anti-WP32112 antibody reactivity does not appear to be associated with T1D, it might be related to CD. To verify this potential association, it would be essential to analyze larger sample groups for antibody reactivity to pure WP32112 by ELISA. Unfortunately, purification of the recombinant his-WP32112 from *E. coli* was incomplete and the protein of interest was contaminated by several bacterial proteins. Our lab has recently established a system for the expression and purification of recombinant his-tagged wheat proteins in Sf21 insect cells using a baculovirus delivery vector. Therefore this technique should be used to obtain purified WP32112 for further analysis and characterization. It is possible that antibodies against these antigenic wheat proteins may serve as a marker for gut leakiness and therefore the presence of these antibodies may be exploited to identify individuals with gut barrier dysfunction who may be at higher risk of developing diabetes.

How do these findings integrate into our current model of T1D pathogenesis

Autoimmune diabetes occurs following the breakdown of tolerance to the insulin-producing β-cells of the islets of Langerhans in the pancreas. Wheat proteins have been implicated as a major modulator of diabetes incidence. **Figure 23** highlights the major findings of this work and in this scheme I speculate how these wheat proteins may be incorporated into a model of diabetes pathogenesis. Recent key studies have linked the importance of a dysfunctional gut barrier, regulated by the zonulin pathway, and the presence of a pro-inflammatory state in the gastrointestinal tract in the progression towards the breach of tolerance and subsequent development of diabetes in BB rats (88, 96) and humans (87). The increased permeability of the intestinal epithelium results in a "leaky" gut whereby excessive quantities of dietary proteins (eg. wheat) cross this barrier and interact with the immune system. Dendritic cells and other antigen-presenting cells, including B cells, take up these proteins and present peptides derived from them on MHC class II molecules to CD4+ T cells.

A healthy immune system will not mount an immune reaction against dietary antigens or commensal bacteria. The mechanism by which this occurs is oral tolerance, whereby immune tolerance is induced against these antigens, but not to pathogens present in the gut (2). The expansion of lymphocytes in response to these dietary antigens signifies a breach in oral tolerance and a defect in the immune system. In turn, these activated CD4+ T cells proliferate and can provide the required signals for B cell activation. Repeated exposure to antigen induces B cell proliferation and immunoglobulin isotype switching from IgM to the secreted IgG form. These antibodies against wheat proteins, including WP5212 and possibly WP32112, that are unusually antigenic in a subset of T1D patients and CD patients may

Figure 23. Potential mechanism of aberrant immune response to wheat peptides. Dietary proteins are broken down into peptides composed of several distinct epitopes (depicted by the shapes) within the gastrointestinal tract (**1**). Increased gut permeability allows excess amounts of these dietary antigens to cross through the gut epithelium (**2**) into the lumen. Alternatively, the dietary peptides can pass through the M cells. Once in the lumen, the B cells sample these peptides (**3**) and present shorter peptides on MHC class II molecules to previously activated CD4+ T cells (**4**). These CD4+ T cells provide the signals (**5**) required to activate the B cell which results in B cell proliferation (**6**) and immunoglobulin isotype switching to the secreted IgG form. This results in the presence of antibodies against dietary proteins in the serum of the host (**7**). These antibodies can cross-react with epitopes present in target tissues (**8**). The question remains whether the presence of these cross-reactive antibodies reflects degeneracy at the T cell receptor level too (**9**). Modified from (26).


2.
M cell
IEC
3.
B cell
6.
proliferation
7.
IgG antibodies
MHC II
4.
TCR
5.
activation
CD4+ T cell
Target tissue peptides
eg. gut
immune system
pancreas
8.
9.
CD8+ T cell
?
TCR
MHC I
Target cell

serve as markers of gut damage. Antibodies raised against WP5212 cross-react with epitopes present in target tissues, including the pancreas, MLNs and spleen. Although these antibodies are not suspected to be pathogenic, it remains to be determined whether this antibody cross-reactivity reflects redundancy at the level of T cell receptors. If this is the case, then tolerance to these self-proteins would be broken and T cells would mistakenly target and damage self tissues.

Pancreatic trypsin 1, the inactive precursor of trypsin, has not previously been identified as an autoantigen associated with autoimmune disease. However, the pancreatic trypsin 1 gene (Prss1) is positioned near the rat diabetes susceptibility lymphopenia gene and its human homologue (111). In addition, several missense mutations in the human homologue of pancreatic trypsin 1, cationic trypsinogen, have been linked to hereditary pancreatitis (reviewed in (112)). It is believed that these mutations result in the premature activation of trypsin which leads to autodigestion of the pancreas (112, 113). Furthermore, unexplained elevated levels of serum trypsin have been recently observed in patients with CD (114). For the first time, pancreatic trypsin 1 has been identified as a candidate molecular mimic of the antigen wheat protein WP5212. The question remains whether pancreatic trypsin 1 is involved in diabetes and if so, how this would occur. During development, both the endocrine and exocrine pancreas arise from a common epithelial precursor. Furthermore, it is believed that acinar cells can transdifferentiate into β-cells via a duct-like intermediate in order to restore β-cell mass (Reviewed in (115)). Consequently, if these cells with transdifferentiation potential expressing pancreatic trypsin 1, are additional targets of the

immune attack, pancreas remodeling would be limited and the regenerative ability of β-cell mass would be lost. Further studies are required to explore this possibility.

Future directions

Following the identification of candidate antigenic wheat proteins, the question remains whether WP5212 or other candidate proteins trigger/promote diabetes or whether the immune response against these proteins is simply a marker of gut damage. In order to determine if these candidate wheat proteins are responsible for the increased incidence of diabetes in animals fed a wheat-based diet, it would be of interest to supplement wheat-free diets with individual candidate proteins. However, due to the difficulty purifying the large quantities of protein required for feeding animals for the duration of the study, these studies are not feasible. Alternatively, it would be of interest to induce oral tolerance by feeding candidate antigenic wheat proteins to diabetes-prone neonate. A recent study demonstrated that oral exposure of neonates, ages 4-7 days, to antigens present in diabetes-promoting wheat-based diets delayed the onset of diabetes in BBdp rats (116). While this procedure would also require milligram amounts of the protein, several methods have recently been developed to overcome this obstacle, including expression of antigens conjugated to non-toxic mucosal adjuvant cholera toxin B (CTB) in silkworms (117) or oral administration of virus encoding CTB-antigens (118). The dose of the candidate proteins will be crucial, because it is an important determinant for the initiation of immunity or oral tolerance by the oral administration of antigens (119).

Recent studies have demonstrated that immunization against key autoantigens, including insulin, has altered the incidence of diabetes, confirming their involvement in the development of this disease (120). Therefore, it would be of interest to immunize diabetes-susceptible animals with the candidate antigenic wheat proteins and study the incidence of diabetes and pancreatic damage. While recent studies immunize animals with protein, an alternative is DNA immunization, which eliminates both the time consuming step of purifying the recombinant protein and the requirement of an adjuvant (76). Furthermore, the immune reaction against the candidate protein can be tailored towards a Th1 response, as observed in T1D, by co-immunization with plasmids encoding immunomodulatory molecules (109). Modification of diabetes incidence would suggest a role for these candidate proteins in the onset of the disease.

Finally, as type 1 diabetes is a T cell-mediated disease, it would be of great value to monitor diabetes incidence in adoptive transfer experiments whereby T cells are stimulated with the candidate proteins and injected into nude animals, for example nude BB rats or NOD-SCID mice. Similar studies should also be performed to determine whether identified candidate molecular mimics in target tissues are involved in the development of diabetes. Together these studies will provide crucial insight pertaining to pathways involved in pathogenesis of wheat-modulated autoimmune diabetes.

CONCLUSION

The identity of environmental agents which play a major role in type 1 diabetes pathogenesis is unknown. Epidemiological studies provide initial hypotheses; however, it remains difficult to link past environmental exposures to diabetes development. A recent study identified WP5212 as the first diabetes-related wheat protein and additional analyses indicate that there appears to be an abnormal immune reaction to this dietary antigen in patients, marked by T cell proliferation and antibody reactivity (42, 54, 57, 58). The antigenic molecular structures of environmental factors associated with autoimmune diseases are exceedingly difficult to identify. A similar approach to that which was used to identify WP5212, detected a second antigenic wheat protein. Serum from a highly wheat-sensitive patient with both T1D and CD was used to screen a wheat cDNA expression library. Antibodies from this patient recognized and bound to clone WP32112, which encodes granule-bound starch synthase I. It was determined that antibody reactivity against this protein may be linked to CD; however this association requires further investigation. The presence of antibodies against dietary proteins may be an indication of abnormal gut permeability which could be used as markers for disease prediction (87).

The primary goal of this project was to examine a potential role for molecular mimicry between antigenic wheat proteins and self-proteins in the development of diabetes. In order to detect self-proteins sharing structural similarities with the antigen wheat protein, protein extracts from target tissues were screened by Western blots using serum antibodies raised against WP5212. The results presented in this thesis demonstrate proof of concept for this novel approach. The detection of this antibody cross-reactivity between molecular structures present in self-proteins and the antigenic wheat protein provides further support that an

abnormal immune reaction to dietary antigens could be associated with autoimmunity. This approach identified pancreatic trypsin 1 as a candidate molecular structure involved in diabetes. Interestingly, mutations in this protein have already been associated with pancreatitis (112). Epitope mapping of WP5212 failed to identify immunodominant linear epitopes, however suggesting that post-translational modification and/or conformational epitopes for antibody recognition may be important. As a result, an anti-WP5212 polyclonal serum was generated by DNA immunization to produce antibodies against additional linear and conformational epitopes representing the entire wheat protein. The development of these new tools and approaches will help identify additional candidate T1D-related molecules. The characterization of these molecular structures may offer invaluable insight into the mechanisms related to the pathogenesis of environment-modulated diabetes.

REFERENCES

1. Marrack, P., J. Kappler, and B. L. Kotzin. 2001. Autoimmune disease: why and where it occurs. *Nat Med 7:899.*
2. Janeway, C. A., P. Travers, M. Walport, and M. Shlomchik. 2001. *Immunobiology: The immune system in health and disease.* Garland Publishing, New York.
3. Gianani, R., and G. S. Eisenbarth. 2005. The stages of type 1A diabetes: 2005. *Immunol Rev 204:232.*
4. Atkinson, M. A. 2005. ADA Outstanding Scientific Achievement Lecture 2004. Thirty years of investigating the autoimmune basis for type 1 diabetes: why can't we prevent or reverse this disease? *Diabetes 54:1253.*
5. Yoon, J. W., and H. S. Jun. 2005. Autoimmune destruction of pancreatic beta cells. *Am J Ther 12:580.*
6. Mathis, D., L. Vence, and C. Benoist. 2001. beta-Cell death during progression to diabetes. *Nature 414:792.*
7. Faideau, B., E. Larger, F. Lepault, J. C. Carel, and C. Boitard. 2005. Role of {beta}-Cells in Type 1 Diabetes Pathogenesis. *Diabetes 54 Suppl 2:S87.*
8. Mandrup-Poulsen, T. 2003. Beta cell death and protection. *Ann N Y Acad Sci 1005:32.*
9. Schmidt, K. D., C. Valeri, and R. D. Leslie. 2005. Autoantibodies in Type 1 diabetes. *Clin Chim Acta 354:35.*
10. Lieberman, S. M., and T. P. DiLorenzo. 2003. A comprehensive guide to antibody and T-cell responses in type 1 diabetes. *Tissue Antigens 62:359.*
11. Achenbach, P., K. Koczwara, A. Knopff, H. Naserke, A. G. Ziegler, and E. Bonifacio. 2004. Mature high-affinity immune responses to (pro)insulin anticipate the autoimmune cascade that leads to type 1 diabetes. *J Clin Invest 114:589.*
12. Kent, S. C., Y. Chen, L. Bregoli, S. M. Clemmings, N. S. Kenyon, C. Ricordi, B. J. Hering, and D. A. Hafler. 2005. Expanded T cells from pancreatic lymph nodes of type 1 diabetic subjects recognize an insulin epitope. *Nature 435:224.*
13. Nakayama, M., N. Abiru, H. Moriyama, N. Babaya, E. Liu, D. Miao, L. Yu, D. R. Wegmann, J. C. Hutton, J. F. Elliott, and G. S. Eisenbarth. 2005. Prime role for an insulin epitope in the development of type 1 diabetes in NOD mice. *Nature 435:220.*
14. Sherry, N. A., E. B. Tsai, and K. C. Herold. 2005. Natural History of {beta}-Cell Function in Type 1 Diabetes. *Diabetes 54 Suppl 2:S32.*
15. Pugliese, A. 2004. Genetics of type 1 diabetes. *Endocrinol Metab Clin North Am 33:1.*
16. Hermann, R., A. P. Laine, R. Veijola, T. Vahlberg, S. Simell, J. Lahde, O. Simell, M. Knip, and J. Ilonen. 2005. The effect of HLA class II, insulin and CTLA4 gene regions on the development of humoral beta cell autoimmunity. *Diabetologia 48:1766.*
17. Bohren, K. M., V. Nadkarni, J. H. Song, K. H. Gabbay, and D. Owerbach. 2004. A M55V polymorphism in a novel SUMO gene (SUMO-4) differentially activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus. *J Biol Chem 279:27233.*
18. Sasaki, Y., K. Ihara, N. Matsuura, H. Kohno, S. Nagafuchi, R. Kuromaru, K. Kusuhara, R. Takeya, T. Hoey, H. Sumimoto, and T. Hara. 2004. Identification of a novel type 1 diabetes susceptibility gene, T-bet. *Hum Genet 115:177.*

19. Park, Y., S. Park, E. Yoo, D. Kim, and H. Shin. 2004. Association of the polymorphism for Toll-like receptor 2 with type 1 diabetes susceptibility. *Ann N Y Acad Sci 1037:170.*
20. 1999. Childhood immunizations and type 1 diabetes: summary of an Institute for Vaccine Safety Workshop. The Institute for Vaccine Safety Diabetes Workshop Panel. *Pediatr Infect Dis J 18:217.*
21. Taplin, C. E., M. E. Craig, M. Lloyd, C. Taylor, P. Crock, M. Silink, and N. J. Howard. 2005. The rising incidence of childhood type 1 diabetes in New South Wales, 1990-2002. *Med J Aust 183:243.*
22. Mauny, F., M. Grandmottet, C. Lestradet, J. Guitard, D. Crenn, N. Floret, M. Olivier-Koehret, and J. F. Viel. 2005. Increasing trend of childhood type 1 diabetes in Franche-Comte (France): analysis of age and period effects from 1980 to 1998. *Eur J Epidemiol 20:325.*
23. Karvonen, M., M. Viik-Kajander, E. Moltchanova, I. Libman, R. LaPorte, and J. Tuomilehto. 2000. Incidence of childhood type 1 diabetes worldwide. Diabetes Mondiale (DiaMond) Project Group. *Diabetes Care 23:1516.*
24. Newhook, L. A., J. Curtis, D. Hagerty, M. Grant, A. D. Paterson, C. Crummel, T. Bridger, and P. Parfrey. 2004. High incidence of childhood type 1 diabetes in the Avalon Peninsula, Newfoundland, Canada. *Diabetes Care 27:885.*
25. Scott, F. W. 1994. Food, Diabetes, and Immunology. In *Diet, Nutrition, and Immunity*. R. A. Forse, ed. CRC Press, Inc., Florida, p. 73.
26. Lefebvre, D. E., K. L. Powell, A. Strom, and F. W. Scott. in press. Dietary Proteins as Environmental Modifiers of Type 1 Diabetes Mellitus. *Annual Review of Nutrition.*
27. Yoon, J. W., M. Austin, T. Onodera, and A. L. Notkins. 1979. Isolation of a virus from the pancreas of a child with diabetic ketoacidosis. *N Engl J Med 300:1173.*
28. Serreze, D. V., C. Wasserfall, E. W. Ottendorfer, M. Stalvey, M. A. Pierce, C. Gauntt, B. O'Donnell, J. B. Flanagan, M. Campbell-Thompson, T. M. Ellis, and M. A. Atkinson. 2005. Diabetes acceleration or prevention by a coxsackievirus B4 infection: critical requirements for both interleukin-4 and gamma interferon. *J Virol 79:1045.*
29. Tauriainen, S., K. Salminen, and H. Hyoty. 2003. Can enteroviruses cause type 1 diabetes? *Ann N Y Acad Sci 1005:13.*
30. Knip, M. 2003. Environmental triggers and determinants of beta-cell autoimmunity and type 1 diabetes. *Rev Endocr Metab Disord 4:213.*
31. Graves, P. M., H. A. Rotbart, W. A. Nix, M. A. Pallansch, H. A. Erlich, J. M. Norris, M. Hoffman, G. S. Eisenbarth, and M. Rewers. 2003. Prospective study of enteroviral infections and development of beta-cell autoimmunity. Diabetes autoimmunity study in the young (DAISY). *Diabetes Res Clin Pract 59:51.*
32. Chen, L.-K., Y.-C. Chou, S.-T. Tsai, S.-J. Hwang, and S.-D. Lee. 2005. Hepatitis C virus infection-related Type 1 diabetes mellitus. *Diabetic Medicine 22:340.*
33. Helgason, T., and M. R. Jonasson. 1981. Evidence for a food additive as a cause of ketosis-prone diabetes. *Lancet 2:716.*
34. Helgason, T. 1982. Food additive as a cause of ketosis-prone diabetes. *Lancet 1:286.*
35. Reeves, P. G., F. H. Nielsen, and G. C. Fahey, Jr. 1993. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr 123:1939.*

36. Scott, F. W. 1996. Food-induced type 1 diabetes in the BB rat. *Diabetes Metab Rev 12:341.*
37. Scott, F. W., H. E. Cloutier, R. Kleemann, U. Woerz-Pagenstert, P. Rowsell, H. W. Modler, and H. Kolb. 1997. Potential mechanisms by which certain foods promote or inhibit the development of spontaneous diabetes in BB rats: dose, timing, early effect on islet area, and switch in infiltrate from Th1 to Th2 cells. *Diabetes 46:589.*
38. Wang, G. S., H. Gruber, P. Smyth, O. Pulido, L. Rosenberg, W. Duguid, and F. W. Scott. 2000. Hydrolysed casein diet protects BB rats from developing diabetes by promoting islet neogenesis. *J Autoimmun 15:407.*
39. Elliott, R. B., and J. M. Martin. 1984. Dietary protein: a trigger of insulin-dependent diabetes in the BB rat? *Diabetologia 26:297.*
40. Scott, F. W., G. Sarwar, and H. E. Cloutier. 1988. Diabetogenicity of various protein sources in the diet of the diabetes-prone BB rat. *Adv Exp Med Biol 246:277.*
41. Scott, F. W., R. Mongeau, M. Kardish, G. Hatina, K. D. Trick, and Z. Wojcinski. 1985. Diet can prevent diabetes in the BB rat. *Diabetes 34:1059.*
42. MacFarlane, A. J., K. M. Burghardt, J. Kelly, T. Simell, O. Simell, I. Altosaar, and F. W. Scott. 2003. A type 1 diabetes-related protein from wheat (Triticum aestivum). cDNA clone of a wheat storage globulin, Glb1, linked to islet damage. *J Biol Chem 278:54.*
43. Graham, S., P. Courtois, W. J. Malaisse, J. Rozing, F. W. Scott, and A. M. Mowat. 2004. Enteropathy precedes type 1 diabetes in the BB rat. *Gut 53:1437.*
44. Beales, P. E., R. B. Elliott, S. Flohe, J. P. Hill, H. Kolb, P. Pozzilli, G. S. Wang, H. Wasmuth, and F. W. Scott. 2002. A multi-centre, blinded international trial of the effect of A(1) and A(2) beta-casein variants on diabetes incidence in two rodent models of spontaneous Type I diabetes. *Diabetologia 45:1240.*
45. Funda, D. P., A. Kaas, T. Bock, H. Tlaskalova-Hogenova, and K. Buschard. 1999. Gluten-free diet prevents diabetes in NOD mice. *Diabetes Metab Res Rev 15:323.*
46. Schmid, S., K. Koczwara, S. Schwinghammer, V. Lampasona, A. G. Ziegler, and E. Bonifacio. 2004. Delayed exposure to wheat and barley proteins reduces diabetes incidence in non-obese diabetic mice. *Clin Immunol 111:108.*
47. Hoorfar, J., K. Buschard, and F. Dagnaes-Hansen. 1993. Prophylactic nutritional modification of the incidence of diabetes in autoimmune non-obese diabetic (NOD) mice. *Br J Nutr 69:597.*
48. Elliott, R. B., S. N. Reddy, N. J. Bibby, and K. Kida. 1988. Dietary prevention of diabetes in the non-obese diabetic mouse. *Diabetologia 31:62.*
49. Coleman, D. L., J. E. Kuzava, and E. H. Leiter. 1990. Effect of diet on incidence of diabetes in nonobese diabetic mice. *Diabetes 39:432.*
50. Ziegler, A. G., S. Schmid, D. Huber, M. Hummel, and E. Bonifacio. 2003. Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies. *Jama 290:1721.*
51. Norris, J. M., K. Barriga, G. Klingensmith, M. Hoffman, G. S. Eisenbarth, H. A. Erlich, and M. Rewers. 2003. Timing of initial cereal exposure in infancy and risk of islet autoimmunity. *Jama 290:1713.*
52. Schmid, S., D. Buuck, A. Knopff, E. Bonifacio, and A. G. Ziegler. 2004. BABYDIET, a feasibility study to prevent the appearance of islet autoantibodies in relatives of patients with Type 1 diabetes by delaying exposure to gluten. *Diabetologia 47:1130.*

53. Klemetti, P., E. Savilahti, J. Ilonen, H. K. Akerblom, and O. Vaarala. 1998. T-cell reactivity to wheat gluten in patients with insulin-dependent diabetes mellitus. *Scand J Immunol 47:48.*
54. Mojibian, M., D. E. Lefebvre, H. Chakir, and F. W. Scott. 2005. T cell reactivity to wheat proteins in human type 1 diabetes. *Diabetes 54:A79 (suppl).*
55. Pastore, M. R., E. Bazzigaluppi, C. Belloni, C. Arcovio, E. Bonifacio, and E. Bosi. 2003. Six months of gluten-free diet do not influence autoantibody titers, but improve insulin secretion in subjects at high risk for type 1 diabetes. *J Clin Endocrinol Metab 88:162.*
56. Hummel, M., E. Bonifacio, H. E. Naserke, and A. G. Ziegler. 2002. Elimination of dietary gluten does not reduce titers of type 1 diabetes-associated autoantibodies in high-risk subjects. *Diabetes Care 25:1111.*
57. Mojibian, M., H. Chakir, A. J. MacFarlane, D. E. Lefebvre, J. R. Webb, C. Touchie, J. Karsh, J. A. Crookshank, and F. W. Scott. 2006. Immune reactivity to a glb1 homologue in a highly wheat-sensitive patient with type 1 diabetes and celiac disease. *Diabetes Care 29:1108.*
58. MacFarlane, A. J., and F. W. Scott. 2005. Diabetes-associated antibodies to clone WP5212, a Glb1-like wheat storage globulin. *FASEB J 19:A1465.*
59. Schuppan, D., and E. G. Hahn. 2001. Celiac disease and its link to type 1 diabetes mellitus. *J Pediatr Endocrinol Metab 14 Suppl 1:597.*
60. Lampasona, V., R. Bonfanti, E. Bazzigaluppi, A. Venerando, G. Chiumello, E. Bosi, and E. Bonifacio. 1999. Antibodies to tissue transglutaminase C in type I diabetes. *Diabetologia 42:1195.*
61. Catassi, C., A. Guerrieri, E. Bartolotta, G. V. Coppa, and P. L. Giorgi. 1987. Antigliadin antibodies at onset of diabetes in children. *Lancet 2:158.*
62. Hummel, M., E. Bonifacio, M. Stern, J. Dittler, A. Schimmel, and A. G. Ziegler. 2000. Development of celiac disease-associated antibodies in offspring of parents with type I diabetes. *Diabetologia 43:1005.*
63. Solly, N. R., M. C. Honeyman, and L. C. Harrison. 2001. The mucosal interface between 'self' and 'non-self' determines the impact of environment on autoimmune diabetes. *Curr Dir Autoimmun 4:68.*
64. Liu, Z., N. Li, and J. Neu. 2005. Tight junctions, leaky intestines, and pediatric diseases. *Acta Paediatr 94:386.*
65. Fasano, A., and T. Shea-Donohue. 2005. Mechanisms of disease: the role of intestinal barrier function in the pathogenesis of gastrointestinal autoimmune diseases. *Nat Clin Pract Gastroenterol Hepatol 2:416.*
66. Shan, L., O. Molberg, I. Parrot, F. Hausch, F. Filiz, G. M. Gray, L. M. Sollid, and C. Khosla. 2002. Structural basis for gluten intolerance in celiac sprue. *Science 297:2275.*
67. Hamer, R. J. 2005. Coeliac Disease: background and biochemical aspects. *Biotechnol Adv 23:401.*
68. Green, P. H., K. Rostami, and M. N. Marsh. 2005. Diagnosis of coeliac disease. *Best Pract Res Clin Gastroenterol 19:389.*
69. Seibold, F. 2005. Food-induced immune responses as origin of bowel disease? *Digestion 71:251.*

70. Becker, W. M., and G. Reese. 2001. Immunological identification and characterization of individual food allergens. *J Chromatogr B Biomed Sci Appl 756:131.*
71. Reuter, A., D. Fortunato, L. P. Garoffo, L. Napolitano, S. Scheurer, M. G. Giuffrida, S. Vieths, and A. Conti. 2005. Novel isoforms of Pru av 1 with diverging immunoglobulin E binding properties identified by a synergistic combination of molecular biology and proteomics. *Proteomics 5:282.*
72. Hird, H., R. Pumphrey, P. Wilson, J. Sunderland, and P. Reece. 2000. Identification of peanut and hazelnut allergens by native two-dimensional gel electrophoresis. *Electrophoresis 21:2678.*
73. Leung, P. S., and K. H. Chu. 2001. cDNA cloning and molecular identification of the major oyster allergen from the Pacific oyster Crassostrea gigas. *Clin Exp Allergy 31:1287.*
74. Huang, L., H. Shen, M. A. Atkinson, and R. T. Kennedy. 1995. Detection of exocytosis at individual pancreatic beta cells by amperometry at a chemically modified microelectrode. *Proc Natl Acad Sci U S A 92:9608.*
75. Verwoerd, T. C., B. M. Dekker, and A. Hoekema. 1989. A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res 17:2362.*
76. Chowdhury, P. S., M. Gallo, and I. Pastan. 2001. Generation of high titer antisera in rabbits by DNA immunization. *J Immunol Methods 249:147.*
77. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res 25:3389.*
78. Kiemer, L., J. D. Bendtsen, and N. Blom. 2005. NetAcet: prediction of N-terminal acetylation sites. *Bioinformatics 21:1269.*
79. Julenius, K., A. Molgaard, R. Gupta, and S. Brunak. 2005. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology 15:153.*
80. Gupta, R., J. Hansen, and S. Brunak. manuscript in preparation (www.cbs.dtu.dk/services/YinOYang/). Identifying intracellular O-(beta)-GlcNAc `yin-yang' switches in the available human proteome.
81. Blom, N., S. Gammeltoft, and S. Brunak. 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol 294:1351.*
82. Perkins, D. N., D. J. Pappin, D. M. Creasy, and J. S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis 20:3551.*
83. MacFarlane, A. J. 2004. Identification and characterization of the type 1 diabetes related wheat storage globulin WP5212. In *Biochemistry, Microbiology and Immunology*. University of Ottawa, Ottawa.
84. Akerblom, H. K., S. M. Virtanen, J. Ilonen, E. Savilahti, O. Vaarala, A. Reunanen, K. Teramo, A. M. Hamalainen, J. Paronen, M. A. Riikjarv, A. Ormisson, J. Ludvigsson, H. M. Dosch, T. Hakulinen, and M. Knip. 2005. Dietary manipulation of beta cell autoimmunity in infants at increased risk of type 1 diabetes: a pilot study. *Diabetologia 48:829.*
85. Merani, S., and A. M. Shapiro. 2006. Current status of pancreatic islet transplantation. *Clin Sci (Lond) 110:611.*

86. Balamurugan, A. N., R. Bottino, N. Giannoukakis, and C. Smetanka. 2006. Prospective and challenges of islet transplantation for the therapy of autoimmune diabetes. *Pancreas 32:231.*
87. Sapone, A., L. de Magistris, M. Pietzak, M. G. Clemente, A. Tripathi, F. Cucca, R. Lampis, D. Kryszak, M. Carteni, M. Generoso, D. Iafusco, F. Prisco, F. Laghi, G. Riegler, R. Carratu, D. Counts, and A. Fasano. 2006. Zonulin upregulation is associated with increased gut permeability in subjects with type 1 diabetes and their relatives. *Diabetes 55:1443.*
88. Malaisse, W. J., P. Courtois, and F. W. Scott. 2004. Insulin-dependent diabetes and gut dysfunction: the BB rat model. *Horm Metab Res 36:585.*
89. Maurano, F., G. Mazzarella, D. Luongo, R. Stefanile, R. D'Arienzo, M. Rossi, S. Auricchio, and R. Troncone. 2005. Small intestinal enteropathy in non-obese diabetic mice fed a diet containing wheat. *Diabetologia 48:931.*
90. Carratu, R., M. Secondulfo, L. de Magistris, D. Iafusco, A. Urio, M. G. Carbone, G. Pontoni, M. Carteni, and F. Prisco. 1999. Altered intestinal permeability to mannitol in diabetes mellitus type I. *J Pediatr Gastroenterol Nutr 28:264.*
91. Turley, S. J., J. W. Lee, N. Dutton-Swain, D. Mathis, and C. Benoist. 2005. Endocrine self and gut non-self intersect in the pancreatic lymph nodes. *Proc Natl Acad Sci U S A 102:17729.*
92. Courtois, P., G. Nsimba, H. Jijakli, A. Sener, F. W. Scott, and W. J. Malaisse. 2005. Gut permeability and intestinal mucins, invertase, and peroxidase in control and diabetes-prone BB rats fed either a protective or a diabetogenic diet. *Dig Dis Sci 50:266.*
93. Sander, G. R., A. G. Cummins, T. Henshall, and B. C. Powell. 2005. Rapid disruption of intestinal barrier function by gliadin involves altered expression of apical junctional proteins. *FEBS Lett 579:4851.*
94. Clemente, M. G., S. De Virgiliis, J. S. Kang, R. Macatagney, M. P. Musu, M. R. Di Pierro, S. Drago, M. Congia, and A. Fasano. 2003. Early effects of gliadin on enterocyte intracellular signalling involved in intestinal barrier function. *Gut 52:218.*
95. Wang, W., S. Uzzau, S. E. Goldblum, and A. Fasano. 2000. Human zonulin, a potential modulator of intestinal tight junctions. *J Cell Sci 113 Pt 24:4435.*
96. Watts, T., I. Berti, A. Sapone, T. Gerarduzzi, T. Not, R. Zielke, and A. Fasano. 2005. Role of the intestinal tight junction modulator zonulin in the pathogenesis of type I diabetes in BB diabetic-prone rats. *Proc Natl Acad Sci U S A 102:2916.*
97. Burks, A. W., D. Shin, G. Cockrell, J. S. Stanley, R. M. Helm, and G. A. Bannon. 1997. Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity. *Eur J Biochem 245:334.*
98. Breiteneder, H., and E. N. Clare Mills. 2005. Plant food allergens--structural and functional aspects of allergenicity. *Biotechnol Adv 23:395.*
99. Fourneau, J. M., J. M. Bach, P. M. van Endert, and J. F. Bach. 2004. The elusive case for a role of mimicry in autoimmune diseases. *Mol Immunol 40:1095.*
100. Atkinson, M. A., M. A. Bowman, L. Campbell, B. L. Darrow, D. L. Kaufman, and N. K. Maclaren. 1994. Cellular immunity to a determinant common to glutamate decarboxylase and coxsackie virus in insulin-dependent diabetes. *J Clin Invest 94:2125.*

101. Schloot, N. C., S. J. Willemen, G. Duinkerken, J. W. Drijfhout, R. R. de Vries, and B. O. Roep. 2001. Molecular mimicry in type 1 diabetes mellitus revisited: T-cell clones to GAD65 peptides with sequence homology to Coxsackie or proinsulin peptides do not crossreact with homologous counterpart. *Hum Immunol 62:299.*
102. Hiemstra, H. S., N. C. Schloot, P. A. van Veelen, S. J. Willemen, K. L. Franken, J. J. van Rood, R. R. de Vries, A. Chaudhuri, P. O. Behan, J. W. Drijfhout, and B. O. Roep. 2001. Cytomegalovirus in autoimmunity: T cell crossreactivity to viral antigen and autoantigen glutamic acid decarboxylase. *Proc Natl Acad Sci U S A 98:3988.*
103. Natter, S., G. Granditsch, G. L. Reichel, M. Baghestanian, P. Valent, L. Elfman, H. Gronlund, D. Kraft, and R. Valenta. 2001. IgA cross-reactivity between a nuclear autoantigen and wheat proteins suggests molecular mimicry as a possible pathomechanism in celiac disease. *Eur J Immunol 31:918.*
104. Traunmuller, F. 2005. Etiology of Crohn's disease: Do certain food additives cause intestinal inflammation by molecular mimicry of mycobacterial lipids? *Med Hypotheses 65:859.*
105. Mannering, S. I., L. C. Harrison, N. A. Williamson, J. S. Morris, D. J. Thearle, K. P. Jensen, T. W. Kay, J. Rossjohn, B. A. Falk, G. T. Nepom, and A. W. Purcell. 2005. The insulin A-chain epitope recognized by human T cells is posttranslationally modified. *J Exp Med 202:1191.*
106. Croxford, J. L., H. A. Anger, and S. D. Miller. 2005. Viral delivery of an epitope from Haemophilus influenzae induces central nervous system autoimmune disease by molecular mimicry. *J Immunol 174:907.*
107. Henke, A. 2002. DNA immunization--a new chance in vaccine research? *Med Microbiol Immunol (Berl) 191:187.*
108. Rodriguez, F., and J. L. Whitton. 2000. Enhancing DNA immunization. *Virology 268:233.*
109. Larregina, A. T., and L. D. Falo, Jr. 2000. Generating and regulating immune responses through cutaneous gene delivery. *Hum Gene Ther 11:2301.*
110. Wattebled, F., A. Buleon, B. Bouchet, J. P. Ral, L. Lienard, D. Delvalle, K. Binderup, D. Dauvillee, S. Ball, and C. D'Hulst. 2002. Granule-bound starch synthase I. A major enzyme involved in the biogenesis of B-crystallites in starch granules. *Eur J Biochem 269:3810.*
111. Hornum, L., and H. Markholst. 2000. Comparative mapping of the human homologue of the rat diabetes susceptibility gene lyp to a 1.3-Mb segment on HSA7. *Genomics 65:81.*
112. Howes, N., W. Greenhalf, D. D. Stocken, and J. P. Neoptolemos. 2005. Cationic trypsinogen mutations and pancreatitis. *Clin Lab Med 25:39.*
113. Whitcomb, D. C., and C. D. Ulrich, 2nd. 1999. Hereditary pancreatitis: new insights, new directions. *Baillieres Best Pract Res Clin Gastroenterol 13:253.*
114. Carroccio, A., L. Di Prima, C. Scalici, M. Soresi, A. B. Cefalu, D. Noto, M. R. Averna, G. Montalto, and G. Iacono. 2006. Unexplained elevated serum pancreatic enzymes: a reason to suspect celiac disease. *Clin Gastroenterol Hepatol 4:455.*
115. Lardon, J., and L. Bouwens. 2005. Metaplasia in the pancreas. *Differentiation 73:278.*
116. Scott, F. W., P. Rowsell, G. S. Wang, K. Burghardt, H. Kolb, and S. Flohe. 2002. Oral exposure to diabetes-promoting food or immunomodulators in neonates alters gut cytokines and diabetes. *Diabetes 51:73.*

117. Gong, Z., Y. Jin, and Y. Zhang. 2005. Oral administration of a cholera toxin B subunit-insulin fusion protein produced in silkworm protects against autoimmune diabetes. *J Biotechnol 119:93.*
118. Denes, B., V. Krausova, N. Fodor, T. Timiryasova, D. Henderson, J. Hough, J. Yu, I. Fodor, and W. H. Langridge. 2005. Protection of NOD mice from type 1 diabetes after oral inoculation with vaccinia viruses expressing adjuvanted islet autoantigens. *J Immunother 28:438.*
119. Blanas, E., and W. R. Heath. 1999. Oral administration of antigen can lead to the onset of autoimmune disease. *Int Rev Immunol 18:217.*
120. Eckenrode, S. E., Q. G. Ruan, C. D. Collins, P. Yang, R. A. McIndoe, A. Muir, and J. X. She. 2004. Molecular pathways altered by insulin b9-23 immunization. *Ann N Y Acad Sci 1037:175.*

APPENDIX 1 – MEDIA AND BUFFERS

AP development solution	77 mM Tris HCl (pH 9.5), 100 mM NaCl, 5 mM $MgCl_2$ containing 0.015% 5-bromo-4-chloro-3-indolyl phosphate, disodium salt (BCIP) and 0.030% Nitro blue tetrazolium chloride (Nitro-BT)
ECL substrate	100 mM Tris HCl (pH 8.5), 0.2 mM p-Cumaric acid, 1.25 mM luminal, 0.02% [v/v] H_2O_2
Elisa binding buffer	0.1 M Na_2CO_3, 0.1 M $NaHCO_3$, pH 9.6
Equilibration buffer	106 mM Tris HCl, 141 mM Tris base, 2% [w/v] SDS, 10% [v/v] glycerol, 0.51 mM EDTA, 0.22 mM Blue G250, 0.175 mM Phenol Red
HSN buffer	0.2% [w/v] SDS, 0.2% NP-50, protease inhibitor cocktail (1 tablet/10 ml) (Roche), 10 U/ml DNase, 2 mM PMSF
LB agar	LB broth containing 2% [w/v] agar
LB broth	0.5% [w/v] yeast extract, 1% [w/v] NaCl, 1% [w/v] tryptone
Lysis buffer	100 mM NaH_2PO_4, 10 mM Tris-Cl, 8 M urea (pH 8.0)
Lower Tris	1.5 M Tris base (pH 8.8), 14 mM SDS
NZY agar	NZY broth containing 1.5% [w/v] agar
NZY broth	1% [w/v] casein digest, 0.5% [w/v] yeast extract, 0.5% [w/v] NaCl, 0.098% [w/v] $MgSO_4$ anhydrous
NZY top agarose	NZY broth containing 0.7% [w/v] agarose
PBS	120 mM NaCl, 17 mM Na2HPO4, 3 mM KH2PO4, pH 7.3
PBS-T (Western blots) *PBS-T (ELISA)*	PBS containing 0.1% [v/v] Tween20 PBS containing 0.05% [v/v] Tween20
Ponseau-S	0.1% [v/w] Ponceau-S in 1% [v/v] acetic acid
Running buffer	25 mM Tris base, 190 mM glycine, 1% [w/v] SDS, 10 µM phenylmethylsulfonyl fluoride (PMSF)
Sample buffer	100 mM Tris HCl (pH 6.8), 2.5% [w/v] sodium dodecyl sulphate (SDS), 20% [v/v] glycerol, bromophenol blue
SM buffer	38 mM Tris HCl (pH 7.5), 100 mM NaCl, 8 mM $MgSO_4 \bullet 7H_2O$, 0.01% gelatin
Sucrose gradient buffer	50 mM Tris HCl (pH 8.0), 10 mM $MgCl_2$, 10 mM beta-mercaptoethanol 10% [v/v] glycerol
TAE buffer	40 mM Tris-acetate (pH 8.3), 1 mM EDTA
TBS	50 mM Tris-HCl (pH 7.5) 150 mM NaCl

TBS-T	TBS containing 0.1% [v/v] Tween20
Transfer buffer (nitrocellulose)	25 mM Tris base (pH 8.3), 190 mM glycine, 20% methanol
Transfer buffer (PVDF)	10 mM 3-[cyclohexylamino] 1-propane-sulphonic acid (pH 11), 10% [v/v] methanol
Upper Tris	0.5 M Tris base (pH 6.8), 14 mM SDS

APPENDIX 2 – TABLE PEPTIDE LIBRARY

Amino acid	Peptide sequence	Solution
1 - 17	MATRGRATIPLLFLLGT	1
10 - 26	PLLFLLGTSLLFAAAVS	3
19 - 35	LLFAAAVSASHDEEEDR	1
28 - 44	SHDEEEDRRGGRSLQRC	1
37 - 53	GGRSLQRCVQRCQQDRP	3
46 - 62	QRCQQDRPRYSHARCVQ	3
55 - 71	YSHARCVQECRDDQQQH	2
64 - 80	CRDDQQQHGRHEQEEQG	2
73 - 89	RHEQEEQGRGHGRHGEG	1
82 - 98	GHGRHGEGEREEEQGRG	1
91 - 107	REEEQGRGRGRRGQGER	1
100 - 116	GRRGQGEREEEQGRGRG	2
109 - 125	EEQGRGRGRRGEGERDE	1
118 - 134	RGEGERDEEHGDGRRPY	1
127 - 143	HGDGRRPYVFGPRSFRR	1
136 - 152	FGPRSFRRIIRSDHGFV	1
145 - 161	IRSDHGFVKALRPFDEV	1
154 - 170	ALRPFDEVSRLLRGIRN	1
163 - 179	RLLRGIRNYRVAIMEVN	2
172 - 188	RVAIMEVNPRAFVVPGL	2
181 - 197	RAFVVPGLTDADGVGYV	1
190 - 206	DADGVGYVAQGEGVLTV	3
199 - 215	QGEGVLTVIENGEKRSY	3
208 - 224	ENGEKRSYTVRQGDVIV	3
217 - 233	VRQGDVIVAPAGSIMHL	2
226 - 242	PAGSIMHLANTDGRRKL	2
235 - 251	NTDGRRKLVIAKILHTI	2
244 - 260	IAKILHTISVPGKFQYF	2
253 - 269	VPGKFQYFSAKPLLASL	2
262 - 278	AKPLLASLSKRVLTAAL	3
271 - 287	KRVLTAALKTSDERLGS	2
280 - 296	TSDERLGSLLGSRQGKE	1
289 - 305	LGSRQGKEEEEKSISIV	1
298 - 314	EEKSISIVRASEEQLRE	3
307 - 323	ASEEQLRELRRQASEGD	1
316 - 332	RRQASEGDQGHHWPLPP	1
325 - 341	GHHWPLPPFRGDSRDTF	2
334 - 350	RGDSRDTFNLLEQRPKI	3
343 - 359	LLEQRPKIANRHGRLYE	1
352 - 368	NRHGRLYEADARSFHAL	3
361 - 377	DARSFHALAQHDVRVAV	3
370 - 386	QHDVRVAVANITPGSMT	2

Amino acid	Peptide sequence	Solution
379 - 395	NITPGSMTAPYLNTQSF	2
388 - 404	PYLNTQSFKLAVVLEGE	3
397 - 413	LAVVLEGEGEVEIVCPH	1
406 - 422	EVEIVCPHLGRDSERRE	2
415 - 431	GRDSERREQEHGKGRWR	1
424 - 440	EHGKGRWRSEEEEDDRR	1
433 - 449	EEEEDDRRQQRRRGSGS	1
442 - 458	QRRRGSGSESEEEQDQQ	1
451 - 467	SEEEQDQQRYETVRARV	1
460 - 476	YETVRARVSRGSAFVVP	1
469 - 485	RGSAFVVPPGHPVVEIA	3
478 - 494	GHPVVEIASSRGSSNLQ	3
487 - 503	SRGSSNLQVVCFEINAE	2
496 - 512	VCFEINAERNERVWLAG	2
505 - 521	NERVWLAGRNNVIAKLD	2
514 - 530	NNVIAKLDDPAQELAFG	3
523 - 539	PAQELAFGRPAREVQEV	1
532 - 548	PAREVQEVFRAKDQQDE	3
541 - 557	RAKDQQDEGFVAGPEQQ	3
550 - 566	FVAGPEQQQEHERGDRR	1
559 - 575	EHERGDRRRGDRGRGDE	1
568 - 584	GDRGRGDEAVEAFLRMA	2
572 - 588	RGDEAVEAFLRMATAAL	3

Solution 1: 10% acetic acid
Solution 2: 8% acetic acid, 20% acetonitrile
Solution 3: 7% acetic acid, 18% acetonitrile, 10% DMSO

STATEMENT OF CONTRIBUTION OF COLLABORATORS

Dr. Majambu Mbikay and Charles Gyamera-Acheampong
Ottawa Health Research Institute
Ottawa, Ontario, Canada

- Dr. Mbikay provided the pCI-neo mammalian expression vector and Charles Gyamera-Acheampong provided technical assistance regarding DNA immunization

Dr. Gen-Sheng Wang, Dr. Alexander Strom and Majid Mojibian
Lab of Fraser Scott
Ottawa Health Research Institute
Ottawa, Ontario, Canada

- Dr. Wang harvested the requested tissues from BB rats
- Dr. Strom and Majid Mojibian expressed and purified recombinant WP5212

Drs. Karolina Burghardt and Lisa Kauri
Formerly of the lab of Fraser Scott
Ottawa Health Research Institute
Ottawa, Ontario, Canada

- Dr. Burghardt isolated total RNA from hard red spring wheat, AC Barrie caryopses
- Dr. Kauri provided isolated rat islets

Dr. Illimar Altosaar
Department of Biochemistry, Microbiology and Immunology
University of Ottawa
Ottawa, Ontario, Canada

- Dr. Altosaar arranged through Dr. Vern Burrows of Agriculture and Agri-Food Canada to obtain hard red spring wheat caryopses from which RNA was isolated

Drs. Claire Touchie, Jacob Karsh and Erin Keely
Department of Medicine
The Ottawa Hospital
Ottawa, Ontario, Canada

- Drs. Touchie and Karsh recruited the case study subject and Dr. Keely recruited patients for the human study

Dr. Sarah Lawrence
Division of Endocrinology
Children's Hospital of Eastern Ontario
Ottawa, Ontario, Canada

- Dr. Lawrence recruited patients for the human study

KAREN POWELL

EDUCATION

2004 – 2006 University of Ottawa Ottawa ON

- Candidate, Master's of Science in Biochemistry
- Thesis: Identification of type 1 diabetes-related proteins
- Courses: Immunology; Gene expression and protein synthesis

1999 – 2003 Carleton University Ottawa ON

- Bachelor of Science, Highest Honours, Biochemistry and Biotechnology, 2003
- Honours Project: Identification of vegetative incompatibility homologues in *Neurospora crassa*

WORK EXPERIENCE

May 2004 - present University of Ottawa
Master's graduate student Supervisor: Dr. F.W. Scott

- Identification and characterization of proteins related to type 1 diabetes

June 2003 – April 2004 Carleton University & Environment Canada
Jr. Research Technician Supervisor: Dr. M.L. Smith

- Laboratory work related to the generation of data on the survival and persistence of 20 microbial substances listed on the Domestic Substance List (DSL)

Sept 2003 – April 2004 Carleton University
Teaching Assistant Supervisors: Dr. S. Paterson and Dr. M. Providenti

- Cell Physiology and Biochemistry, 2nd year, laboratory component (Fall Term 2003)
- Experimental Microbiology, 3rd year, laboratory component (Winter Term 2004)

Sept 2002 – April 2003 Carleton University
4th Year Student Supervisor: Dr. M.L. Smith

- Identification of vegetative incompatibility homologues in *Neurospora crassa*

Summers 2001, 2002 NRC, Institute of Biological Sciences, PTH Group
Summer Student Supervisor: Dr. J.F. Whitfield

- Laboratory work related to the development and testing of novel analogs of PTH (parathyroid hormone) as potential therapies for postmenopausal osteoporosis

PUBLICATION

Lefebvre, D.E., Powell, K.L., Strom, A., and Scott, F.W. *Dietary Proteins as Environmental Modifiers of Type 1 Diabetes Mellitus.* Annual Review of Nutrition, in press

HONOURS AND AWARDS

2004-2006	Ontario Ministry of Training, Colleges and Universities *Ontario Graduate Scholarship*
2004-2006	University of Ottawa *University of Ottawa Excellence Scholarship*
2004	University of Ottawa *University of Ottawa Strategic Areas of Development Award*
2003	Society of Chemical Industry *Society of Chemical Industry Merit Award*
2003	Carleton University *Senate Medal for Outstanding Academic Achievement*
2000-2003	Carleton University *Dean's List*
2002	Carleton University *Field hockey MVP*
2002	Ontario University Athletics *Field hockey: Second Team All-Star*
2002	Carleton University *Dr. Shakir Sheikh Scholarship in Biochemistry*
2002	Carleton University *Harry H. Southam Scholarship*
2001	Carleton University *Clarence C. Gibson Scholarship*
2000	Carleton University *Murdoch Maxwell MacOdrum Scholarship*
1999	Carleton University Carleton University President's Scholarship
1999	Carleton University Carleton University Faculty of Science Scholarship