THE UNIVERSITY OF CHICAGO

INFLUENCE OF DIETARY PROTEIN SOURCE AND THE MICROBIOTA ON TYPE-1 DIABETES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON IMMUNOLOGY

BY
MATTHEW CARROLL FUNSTEN

CHICAGO, ILLINOIS
DECEMBER 2022

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Sasha. I came into your lab very green, but you trusted me to learn, taught me to design experiments (with the proper controls), and how to look for the forest rather than the trees. Not everything succeeded, but you always pushed me to keep trying. I learned from you that success only comes with hard work, and eventually my projects and experiments started to progress. Some took unexpected twists and turns, LPS was certainly a surprise, but you taught me to keep an open mind and to follow the data. Thank you for your patience, for always taking the time to listen to my ideas, and for your offbeat sense of humor. I am the scientist that I am today because of you, and for that, I will always be grateful.

I would also like to thank Tanya. You taught me to always know *exactly* what I was talking about, how to field difficult questions, and how to admit that I did not know the answer. Thank you for being a candid source of advice, for your encouragement, and for the legendary holiday parties.

I would also like to thank my committee: Pete Savage, Bana Jabri, Marisa Alegre, and Daria Esterhazy. You all helped me to narrow my focus to one project from the three that I initially started with, and encouraged me throughout my PhD. Pete, thank you for always being someone I could turn to whenever I needed advice and for being a great teacher. Bana, thank you for always challenging me in presentations and meetings. Marisa, thank you for suggesting that we colonize TLR4 KO mice with *E. faecalis*. This ended up being a key *in vivo* experiment that solidified my project. And Daria, thank you for your immediate engagement in COI from your first day and excellent questions in committee meetings. Watching your lab grow has been inspiring.

I would also like to acknowledge past and present members of the Chervonsky and Golovkina labs who I had the pleasure of spending graduate school with. Jean Lee, I learned so much from working with you and always felt safe asking you for advice. I know that you have a bright future ahead and I am looking forward to seeing where you go. Renee DePooter,

thank you for teaching me (and so many more COI graduate students) flow cytometry and for always being someone who I could turn to for help with experiments. Andrey Kuznetsov, thank you for being a great research partner with all of the *E. faecalis* experiments. Jessica Spring, Emily Cullum, and Helen Beilinson, it's been great to work and to learn alongside you. Thank you for all of the tubes and plates that I borrowed over the years and for the microbiology advice. Lonya Yurkoveyskiy and Camilla Hansen, thank you for starting this work and for leaving behind a rewarding project. To our technicians, Grace Ryan, Yunisse Gonzalez, Katie Senter, Nicole Garcia, and Clark Halpern: Your work kept the lab and our massive mouse colony afloat. I could not have done this work without your help. And finally, to our new undergraduates Katie Fox, Laura Sams, and Brandon Hawkins, thank you for being great students and for jumping into research head first this summer.

I owe a big thank you as well to our bioinformatics collaborators, Aly Khan and Derek Reiman. You both taught me so much during our frequent zoom meetings to discuss the three single-cell RNA sequencing datasets that I produced for this project. I owe much of what I know about computational analysis to you two.

I also owe a big thank you to my close friends from graduate school: Dave Klawon, Manisha Krishnan, Lari DeWet, Ryan Brown, Avi De Leon, Celisa Chan, Peter Flynn, and Nicholas Venturelli. I am so thankful to have found all of you and have so many fond memories of game nights, hanging at the point, and exploring Chicago together. Dave, thank you so much for being someone I could always talk about immunology with, even if all of our other friends were completely uninterested. I would also like to thank the Histones a cappella group for being a refuge for me on Wednesday nights and a place where I could have a musical outlet during graduate school. Thank you to Connor Tatooles, Grace Peng, Imran Kazmi, Jessica Dekker, Keely Renwick, and our heroic DM Rohit Agrawal. You all kept me sane throughout the pandemic on our weekend misadventures throughout Pranim.

Finally, I would like to thank my family. Mom and Dad, thank you for always encouraging my interest in science from letting me watch Jurassic Park when I was probably way too

young, to encouraging me to go to graduate school when I knew that I wanted to pursue science but was unsure about whether I was talented enough to do it. You believed in me more than I believed in myself. Jenna and Cole, thank you for being the best younger sister and brother I could hope for. You both have grown up so much since I moved to Chicago; I am so excited to see where your lives take you. Thank you to Finn for being handsome without even trying; you are an inspiration to us all. Gabi, Rick, and Camille, thank you for making me part of your family. You have been so welcoming to me over the years and I appreciate so much all of the support you have given to me and Katya since we met. And lastly, thank you to my wonderful wife Katya. I am so lucky to have you in my life. You have always encouraged me even when experiments failed or when my paper got difficult reviews back and I have always been inspired by your path through science and the success that you have had. Exploring Chicago these last 6 years and traveling across Europe and the United States with you has been absolutely incredible. I don't know what the future holds, but I am so excited to spend it with you. I love you.

ABSTRACT

Both diet and commensal microbes impact the development of autoimmune diseases such as Type-1 diabetes (T1D). Since commensals are highly sensitive to changes in diet and can modify dietary components, it was unclear whether diet's effect on autoimmunity was microbe independent or microbe dependent. We found that a diet containing hydrolyzed casein (HC) as the sole amino source protected non-obese diabetic (NOD) mice in both specific pathogen free (SPF) and germ free (GF) conditions. HC did not affect the generation of autoimmune effector cells. Instead, HC reduced β-cell stress by reducing insulin secretion, which was associated with reduced activation of autoreactive T cells. Adding gluten (a wheat storage protein complex associated with celiac disease) to the HC diet (HC+4%gluten) restored high T1D incidence to SPF housed NOD mice, but not to NOD mice raised in GF conditions. Gluten did not restore high β-cell stress but rather promoted inflammation and the expansion of diabetogenic T cells in the islets of Langerhans. To promote T1D, gluten had to be proteolytically digested by microbes. T1D was able to develop in mice fed HC+4%gluten and monocolonized with *Enterococcus faecalis* (*E. faecalis*) secreting gluten digesting proteases, but not with mutant bacteria incapable of secreting these proteases. Bacterial digestion liberated T cell activating peptides from gluten as well as lipopolysaccharide (LPS) capable of activating innate immunity. LPS released from microbially-digested gluten promoted T1D development. Consistent with this, gnotobiotic NOD mice lacking the LPS sensor TLR4 and monocolonized with E. faecalis were resistant to T1D when fed HC+gluten diet. Together, these results indicate that dietary influence on autoimmunity can either be microbe independent or microbe dependent and that microbial digestion can make an otherwise innocuous dietary component diabetogenic by modifying its ability to stimulate the innate immune system. Ultimately, these findings are important for developing dietary interventions to prevent autoimmunity.

CHAPTER 1

INTRODUCTION

In this work, the non-obese diabetic (NOD) mouse model of type-1 diabetes (T1D) was used to address outstanding questions concerning diet's influence on the development of autoimmune disease. These questions were: Do different dietary proteins rely on the microbiota to influence T1D development? How do diet induced changes to β-cell physiology influence T1D development? And how do microbial modifications of diet influence autoimmunity?

1.1 Early history of diabetes mellitus

Diabetes mellitus, hereafter referred to as diabetes, is a group of metabolic diseases that result in hyperglycemia. Hyperglycemia occurs either due to failure of the body to produce insulin or due to an acquired insensitivity to insulin [1, 2]. Regardless of the cause, the consequences of hyperglycemia are severe and can lead to chronic complications such as blindness, neuropathy, and nephropathy caused by microvascular damage, or acute complications such as diabetic ketoacidosis which can result in coma or death [3]. Managing the disease to avoid these complications is highly burdensome requiring constant monitoring of blood glucose concentrations, injections of insulin, and significant lifestyle changes. At a population level the cost is also significant with countries estimated to spend between 5-20% of total healthcare expenditures on diabetes and diabetes complications [4]. As there is currently no cure, and the disease is multifactorial in nature, much more research needs to be done on factors that influence diabetes progression and the development of cost effective prophylactic treatments.

Diabetes has been known since ancient times. In 1552 BCE the Egyptian physician Hess-Ra described a disease characterized by excessive urination and emaciation in the *Papyrus Ebers* [5]. Physicians from India at around the same time observed that the urine from diabetes patients attracted ants and flies thereby earning the disease the name "madhumeha"

literally translating to "honey urine" [6]. Indians of the fifth century CE also were the first to break diabetes into two distinct subtypes, congenital and late onset [7]. The term diabetes, from the Greek word meaning to run-through or siphon, is usually attributed to Araetus of Cappodocia (81-133 CE) who explained the emaciation seen in diabetic patients as flesh "liquefied into urine" [8]. Other references to diabetes come from Chinese and Japanese classical sources and Islamic sources from the middle ages who all described patients that had sweet urine and were more prone to infection, boils, gangrene, and mental troubles [8].

Up until the 18th century, diabetes was presumed to be a disease of the kidneys due to excessive urination observed in patients. That began to change in 1776 when Matthew Dobson discovered that both dried urine and blood from diabetic patients contained substantial sugar content suggesting that diabetes was a more systemic condition [8]. The seminal work linking diabetes to the pancreas came in 1889 from Joseph von Mering and Arthur Minkowski [9]. Following complete pancreatectomy, the dogs in the study developed glycosuria, intense thirst, and loss of weight despite adequate food consumption. Later work from Minkowski and from Edouard Hedon found through engraftment experiments that the anti-diabetic factor secreted by the pancreas was produced internal to the pancreas and did not travel through the exocrine pancreatic ducts thus separating the digestive role of the pancreas from its endocrine role [10, 11].

The next step was the recognition of the islets of Langerhans as the production site of the anti-diabetic factor. Following their discovery in 1869 [12], Gustave Laguesse suggested in 1893 that the islets might be responsible for regulating blood glucose levels [13]. This was confirmed experimentally several years later in the early 20th century. In 1901, hyaline degeneration in the pancreatic islets of diabetic patients was described [14]. This was followed by the independent discoveries by two labs that secretions from the islets of Langerhans could reduce blood glucose concentrations [15, 16]. Both groups are credited with coining the term insulin, from the Latin word for island *insula*, to describe the theoretical islet produced hormone controlling blood glucose [17]. It would take another seven years for insulin to be

discovered. During these years, three separate groups found that preparations from alcohol extracts of pancreatic tissue could reduce blood glucose concentrations in dogs following pancreatectomy [18–20]. However, these studies were abandoned following initial trials in patients due to the toxicity of the preparations [21]. Success finally came in 1922 when Frederick Banting, Charles Best, and James Collip refined the extraction of insulin from the pancreas to produce a much purer product free of toxic side effects. Importantly, when tested in several diabetic human patients, the purified extracts dramatically reduced blood glucose concentration [22]. This work earned Banting and his mentor John Macleod the Nobel Prize in Medicine in 1923 and within a few years, insulin was being produced by the drug manufacturer Eli Lilly and available to the general public.

The discovery and commercialization of insulin turned a diagnosis of diabetes from a death sentence into a disease that could be managed. Widespread treatment of diabetes patients also made it possible both to effectively distinguish between, what would soon be named, type-1 and type-2 diabetes (T1D and T2D). It was recognized early on that some patients responded to insulin treatment while others did not [23]. The patients that responded tended to be young and lean whereas the patients that failed to respond tended to be older and overweight. In addition, insulin injected into young lean diabetic patients cleared blood glucose just as quickly as healthy controls whereas older overweight patients treated similarly failed to clear blood glucose [24]. It was thus hypothesized that diabetes in patients who responded to insulin was caused by an inability to produce insulin (this was confirmed in 1952 [25]) and diabetes in patients who failed to respond was caused by an insensitivity to insulin driven by an unknown mechanism. The names, type-1 and type-2, which refer to insulin sensitive and insulin insensitive diabetes, were coined in 1951 [26] and are the terms that we use today.

Why was there no insulin in patients with T1D? One possibility that emerged in the 1950's was autoimmunity. Antibodies reactive to thyroid extracts were identified in patients with Hashimoto's disease in 1956 [27] and one year later antibodies reactive to the adrenal ex-

tracts were identified in patients with Addison's disease [28]; both patients also had immune infiltrates in their respective diseased organs. In addition, animals immunized with thyroid tissue from their species developed thyroiditis highly reminiscent of histological changes in Hashimoto's disease [29]. These observations led to the hypothesis that both diseases were immune mediated and the possibility that other, then unexplained diseases might also be.

The first immune component to T1D that was recognized was infiltration of the islets by immune cells, named insulitis in 1940 [30]. Insulitis had been observed sporadically in young diabetic patients since the beginning of the 20th century but was thought to be a rare occurrence rather than a marker of diabetes [31]. This view began to change when first, previously healthy animals immunized with insulin and complete Freund's adjuvant (CFA) developed insulitis [32] mirroring the previous experiments done with thyroid extracts [29] and second, the finding that insulitis was in fact, quite common in young patients who died shortly after a diabetes diagnosis [33]. We know now that this is because insulitis is present well before diagnosis of diabetes and declines as the islet tissue is progressively destroyed. Other immune components recognized during this time period were the genetic linkage to particular HLA alleles [34, 35] and the discovery of islet reactive antibodies in type-1 diabetic patient sera [36–38]. The autoimmune dogma of T1D development was finally solidified in the late 1970's with the development of animal models, the the Bio Breeding (BB) rat [39] and later the non-obese diabetic (NOD) mouse [40]. These models of spontaneous diabetes driven by immune destruction of the islets have allowed the study of early stages of autoimmune diabetes in much greater detail.

Overall, the early work on diabetes transformed our understanding of the disease from a single malady involving the liquification of the body to two distinct conditions either driven by a deficiency in insulin or an insensitivity to insulin. In the case of insulin-deficient diabetes (T1D), insulin deficiency was shown to be caused by the destruction of the insulin-producing cells in the islets of Langerhans by an infiltrating immune system. Subsequent sections will focus on our current understanding of the disease process in T1D and the different genetic

and environmental factors that contribute to disease progression.

1.2 T1D disease progression and the role of the immune system

The development of hyperglycemia in T1D is the endpoint of a long pre-clinical period that follows predictable progressive stages: First, autoreactive T cells are activated in the pancreas draining lymph nodes (PLNs) and shortly thereafter, islet-reactive autoantibodes appear. These initial activation steps are followed by the infiltration of the islets by diverse cell types from the immune system, and hyperglycemia appears following the destruction of the islets by the infiltrating cells. As the only observable pre-clinical marker for developing T1D in humans is the appearance of islet reactive antibodies [41], much of the knowledge that we have of the early stages of T1D development and the involvement of the immune system comes from studying disease progression in BB rats and in NOD mice.

In NOD mice, leukocyte infiltration in the islets begins at about 3 weeks of age whereas hyperglycemia starts to appear by 13 weeks of age in most colonies, though this varies considerably from colony to colony [42]. The earliest $CD45^+$ cells to appear in the islets are of myeloid origin with the largest numbers of cells coming from the conventional dendritic cell (cDC) and macrophage subsets. In most strains of mice, cDCs and macrophages are observable in the islets during the first month of life however, in NOD mice and in NOD mice lacking T cells and B cells due to the scid mutation (NOD.SCID), cDCs and macrophages are present in the islets at much higher numbers [43]. It is hypothesized that this early infiltration of cDCs into the islets is precipitated by a wave of physiologic β-cell apoptosis that occurs at 2 weeks of age in mice of all strains [44]. In NOD mice, this wave of β-cell death presumably provides self-antigens to the infiltrating myeloid cells that are then able to traffic back to the PLNs and present the antigen to autoreactive T cells, which become activated thereby breaking self-tolerance, though how this break in tolerance occurs is still unknown. In any event, the activation of T cells rapidly changes the composition of islet infiltrates; multiple subsets of DCs, macrophages, T cells [including cytotoxic T lymphocytes

(CTLs)], and B cells are all present in the islets by 4 weeks of age and steadily increase in numbers as time goes on [45].

T cells are thought to be the key drivers of T1D development and make up the vast majority of the $CD45^+$ cells infiltrating the islets. NOD mice lacking T cells do not develop insulitis or diabetes but both conditions will readily develop if T cells from another NOD mouse are injected [46]. As the transfer of only major histocompatibility complex class I (MHCI)-restricted $CD8^+$ T cells or MHC class II (MHCII)-restricted $CD4^+$ T cells fails to induce disease in NOD.SCID mice but the transfer of both subsets together succeeds, it is thought that both $CD4^+$ and $CD8^+$ T cells are necessary for T1D to develop [47, 48]. Many of the T cells that develop in mice and in humans react to peptides from the β-cell proteins glutamic acid decarboxylase (GAD), zinc transporter 8 (ZnT8), and insulin which are all also the targets of autoantibodies, as well as islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP), and chromogranin A [49]. Experimental evidence has shown that many of these antigens are important for disease development. Transgenic T cells carrying T cells receptors (TCRs) that recognize insulin, IGRP, and chromogranin A epitopes are all capable of transferring diabetes to T cell-deficient mice on their own [50–52].

B cells play a major supporting role in the progression of T1D. NOD mice lacking B cells have much lower disease incidence with some institutions reporting no diabetes at all [53, 54]. Interestingly, these mice have reduced but still detectable insulitis suggesting that B cells are not required for the earliest stages of disease development, but may accelerate the autoimmune process [55]. This acceleration is not thought to occur through the production of autoantibodies, despite autoantibodies being an early marker of developing diabetes, as NOD mice that contain B cells that cannot secrete antibodies still develop diabetes [55]. Therefore, it is likely that B cells participate in T1D development in their capacity as antigen presenting cells (APCs) and NOD mice with B cells that lack MHCII are protected from T1D [56]. Among the antigen-presenting cells, B cells have the unique abilities both to concentrate antigen specific to their B cell receptor (BCR) [57] and to proliferate in response

to toll-like receptor (TLR) stimulation and activation by T cells reactive to the same antigen [58, 59]. These features together suggest that B cells may accelerate the development of diabetes by serving as an expanding pool of APCs presenting islet antigens following the initial activation of a few B cells by T cells activated very early in the autoimmune process. This expanding pool could theoretically be capable of accelerating further T cell activation.

β-cell killing in the islets is primarily directed by T cells [60]. $CD8^+$ T cells have been shown to kill β-cells directly via Fas-FasL interactions and through the use of perforin. $CD4^+$ T cells, in contrast, kill β-cells by directing myeloid cells to secrete tumor necrosis factor (TNF)-α. Killing all of the β-cells is not necessary for diabetes to develop. Treatment of recently diabetic NOD mice with α-CD3 antibodies which deplete T cells results in the reversal of T1D and recovery of pancreatic insulin secretion [61]. This recovery means that not all β-cells are killed by the infiltrating immune cells but, given the lack of insulin production during diabetes, that the surviving β-cells have stopped producing insulin either as a defense mechanism or due to silencing by T cells.

The study of NOD mice over the last 30+ years has greatly expanded our knowledge of the early stages in T1D development. Unanswered questions in the field include how tolerance to β-cell antigens is broken early on in disease, how islet infiltration progresses to β-cell destruction, and how many of the cell types that infiltrate the islets, even if not mentioned here, contribute to disease development.

1.3 Genetic control of T1D

T1D is primarily a disease controlled by genetic factors. While incidence in the general population is about 0.4%, incidence rises to 6-7% among siblings of T1D patients [62] and monozygotic twins have about a 50% concordance rate [63]. A majority of the genetic risk for T1D comes from the HLA locus containing the MHC genes, specifically HLA-DR and HLA-DQ, which has been linked to T1D for forty years [64]. In NOD mice, the MHC locus is also strongly linked to diabetes and is sufficient to induce insulitis in C57BL/6J mice

congenic for the NOD locus [65]. Besides the HLA locus, several other immune related genes have been linked to T1D through GWAS studies. These include *PTPN22* and *CTLA4* which are both negative immune regulators and *IL2RA* which affects T cell sensitivity to IL-2 and affects Treg maintenance [66]. Non-immune related loci include *VNTR* where allelic variation controls insulin transcription levels in the thymus affecting the negative selection of T cells [67], and *GLIS3* which affects β-cell apoptosis during periods of endoplasmic reticulum (ER) stress [68].

Of particular interest is the genetic link between T1D and celiac disease. Celiac disease is thought to be caused by immune responses against gluten antigens that result in the destruction of villi in the small intestine [69]. The two diseases are linked by epidemiological data, which shows that patients diagnosed with T1D are much more prone to develop celiac disease than the general population [70]. T1D and celiac disease also share a genetic linkage for the HLA class II alleles HLA-DQ2 and HLA-DQ8 along with the immune genes *CTLA4* and *IL2RA* [71]. Whether gluten also contributes to T1D as it does in celiac disease is controversial and discussed in greater detail below.

1.4 The role of environmental factors in T1D

Epidemiological data show that T1D rates have been rising 3-4% per year over the last five decades [72]. Despite strong evidence that T1D is driven by genetic susceptibility, disease incidence has risen far too quickly to be accounted for by changes in human genetics. This is especially true in the context of decreasing frequencies of high-risk HLA alleles [73] and differing incidence rates between bordering and genetically similar countries [74]. Other evidence for non-genetic contributions to T1D development include the lack of concordance in monozygotic twin studies [63] and high variation in diabetes rates in NOD mouse colonies housed at different facilities [42]. It should also be noted that this rise in incidence cannot be explained by better diagnostic tools; the severe physical consequences of untreated T1D typically make diagnosis obvious. These facts suggest that environmental factors are driving

higher T1D incidence rates globally.

Environmental factors are defined as any non-genetically encoded factors that can influence an organism's phenotype. These include physical insults, chemicals, toxins, exposure to radiation, infections, community structure of the commensal microbiota, diet, and stress, all of which have been linked to the development of autoimmune disease. Environmental factors of particular interest in T1D have been pathogenic and non-pathogenic microorganisms, and diet due to the prevalence of the hygiene hypothesis. Initially formulated to explain the higher autoimmune incidence rates in the historically industrialized countries of Western and Northern Europe (including the US) and lower incidence rates in historically non-industrialized countries in the global south, the hygiene hypothesis states that the current rise in autoimmune disease including T1D, is due to the decline in infectious disease burden during the latter half of the 20th century as a result of better hygiene and medical care [75]. Supporting this hypothesis, there is a strong negative correlation between T1D rates and the burden of tuberculosis infection and of childhood diarrheal diseases for individual countries [75, 76].

While initially devised to focus on the relationship between infection and autoimmunity, the hygiene hypothesis can be expanded to also include diet and the commensal microbiota. Food consumption patterns have changed massively since World War II with diets in western countries progressively becoming denser in carbohydrates and fats. This has accompanied changes to food production including increased use of chemical fertilizers and antibiotics. These changes correlate well with rising T1D incidence rates, and diet has been proven to influence the development of T1D in mice and rats [77]. More recently appreciated is the contribution of the commensal microbiota, defined as all of the non-pathogenic microorganisms that live on or within a host organism. In exchange for shelter and nutrients derived from the host diet, these microbes aid their host in the digestion of complex carbohydrates, produce nutrients such as vitamins that are essential for host health, and provide protection from pathogens through colonization resistance and by training the host immune system

[78]. In addition to these functions, the use of sterile mice and rats has shown that the microbiota has a profound effect on the development of autoimmunity including T1D [79] and is also suspected to influence human diabetes development [80]. As the microbiota is highly sensitive to dietary changes [81, 82] and to antibiotics which have dramatically increased in use in the past half century, it is highly likely that it has changed in response to changes in food production and consumption. The hygiene hypothesis is therefore incomplete without a thorough understanding of how diet, commensal microbes, and pathogenic microbes all contribute to the development of autoimmunity.

1.4.1 Diet's influence on T1D development

Diet is an environmental factor that has long been of interest in T1D research. Geographic variability in diet composition may contribute to differing T1D rates in different regions. In addition, modifications to diet have profound effects on the diabetes incidence rates in animal models and have thus been extensively studied in rodents and humans.

Both micronutrient (vitamins and minerals) and macronutrient (carbohydrates, lipids and proteins) effects on T1D incidence have been researched. The most extensively studied micronutrient has been vitamin D. Due to reduced direct sunlight and fewer hours of sunlight in countries with the highest rates of T1D, it was hypothesized that vitamin D deficiency may contribute to disease development. Supporting this, vitamin D deficiency in NOD mice early in life promoted T1D incidence [83]. In reverse experiments, oral supplementation of NOD mice with high doses of 1,25-dihydroxyvitamin D(3), the biologically active form of vitamin D, completely prevented diabetes development in the treated animals [84]. However, there have been conflicting results. Vitamin D supplementation of complex diets did not prevent T1D in NOD mice, although vitamin D was still present in the control diets [85]. In addition, NOD mice lacking the vitamin D receptor (*Vdr*) did not develop diabetes at accelerated rates compared to *Vdr* sufficient animals [86]. Finally, in a study of Finnish newborns, vitamin D supplementation was associated with a decreased incidence of T1D while infants suspected

of having rickets, which is caused by vitamin D deficiency, had a higher chance to develop T1D [87]. While some results with vitamin D are promising, conflicting studies indicate a need for more research.

Of the macronutrients, carbohydrates, lipids, and proteins, only protein source has been consistently shown to influence T1D development in animal models. In spite of the key role insulin plays in carbohydrate metabolism, no connection between simple sugar intake and T1D progression has been reported in NOD mice or BB rats. In the diabetes autoimmunity study in the young (DAISY), children with a higher glycemic index or who reported more sugar intake were at higher risk from developing T1D following a positive test for anti-islet antibodies suggesting that sugar may not affect the early stages of autoimmunity but could hasten β-cell death [88, 89]. There has also been no consistent connection between the complex carbohydrate fiber and T1D development or protection. While increased dietary fiber was found to protect NOD mice in one study [90] it had no reported effect on BB rats in another [91]. In humans, elevated dietary fiber was correlated with increased risk of T1D, but its effect could not be separated from other dietary components found to also be correlated with T1D in the same study [92]. These conflicting results could potentially be explained by the presence or absence of gut residing fiber fermenting microbes. Bacteria produce short-chain fatty acids (SCFAs) by fermenting fiber and NOD mice fed diets supplemented with the SCFAs acetate and butyrate are protected from T1D [93, 94]. SCFA-producing microbes have also been correlated with protection from T1D in humans [80] underscoring the need for additional research. Besides SCFAs, other dietary lipids (fish oil, soy oil, and omega-3 fatty acids) have not been linked to the promotion or attenuation of T1D in rodents or humans [72, 85, 95].

In contrast to carbohydrates and lipids, protein source has been shown to have a larger effect on T1D incidence, with individual proteins classified as either diabetes retardant or diabetes promoting. Semi-purified diets using hydrolyzed casein (HC) as the sole amino acid source consistently inhibit T1D development in NOD mice and BB rats [96–99]. Previously

suggested protective mechanisms by HC include the presence of bioactive peptides in HC [100], enhancement of pancreatic secretion [101], and reduction of gut permeability [102, 103] though none of these have been experimentally tied to T1D. Interestingly, HC also protects BB rats raised in germ free (GF) conditions [99] meaning it does not rely on the microbiota for its protective effects and potentially explains its consistent effects at many different institutions over several decades. The mechanism put forward by Patrick *et al.* was the restoration of high regulatory T cell frequencies in the lamina propria of HC fed BB rats which typically have very low regulatory T cell frequencies due to chronic lymphopenia [104]. Currently, it is not clear whether HC protects NOD mice via a similar mechanism.

Protein sources typically labeled T1D promoting include gluten, cow's milk, and soy, the most consistent promoter in animal models being gluten [77]. In contrast to HC however, gluten's impact on diabetes incidence in animal models has been much less consistent. Several studies found that gluten free diets protected NOD mice from T1D [85, 105] and that adding gluten back to a gluten free diet restored high incidence [106]. However, two other studies showed no effect [97, 98]. Confusing matters further, Funda *et al.* found that both removing gluten and increasing the gluten content of a normally diabetogenic diet protected NOD mice from T1D [107]. Similarly conflicting data comes from human studies where gluten was found to either increase risk for developing diabetes [92, 108] or had no effect [109]. Cow's milk (including studies of intact rather than hydrolyzed casein) has also been suggested to be diabetogenic, though again based on very mixed evidence. Early studies in BB rats and in NOD mice found that supplementing synthetic diets with cow's milk promoted T1D [98, 110]. These studies were followed up by a human study that found a correlation between early introduction of dairy products and later T1D development [111] though the influence of other dietary products introduced at the same time cannot be excluded. These contrast with later studies that found that diets supplemented with milk or intact casein were as protective as HC diets suggesting that milk itself or its intact proteins may have no effect on T1D [97, 112, 113].

Explaining these inconsistencies is a challenge. Timing of introduction and maintenance on these different diets was not consistent across the studies. Additionally, many of the studies used different dietary formulations as their background for evaluating the impact of gluten on T1D; alimental components present in one study may be absent from another. Finally, as will be discussed in greater detail later, the animal studies took place in different facilities with different microbes present. Therefore, it is possible that dietary influence on T1D may depend on the microbiota.

1.4.2 Influence of infections on T1D development

Pathogenic and non-pathogenic microorganisms are another environmental factor that influence T1D development. Discordance in disease development in monozygotic twins, significant geographic variability, and rising incidence over time has led some to suspect that an environmental trigger is required to initiate T1D in genetically susceptible humans [114–116]. Infections, particularly viral infections, have long been suspected to fill this role due to the early observation that T1D onset in many populations is seasonal [117]. As such, a number of viruses have been studied and correlated with T1D infection including enteroviruses, cytomegalovirus, rotavirus, rubella, and mumps [118]. However, moving beyond simple correlation to a mechanistic link has proven elusive. Congenital rubella and mumps infections were initially suspected to trigger rapid onset fulminant diabetes, but are no longer suspected to do so as successful vaccination campaigns have largely eliminated the viruses from developed countries without impacting rising T1D rates [119, 120]. Stronger evidence for a viral role in T1D pathogenesis comes from the study of enterovirus infection. Exposure to enteroviruses (particularly coxsackieviruses) have been consistently associated with T1D development [121, 122] including in recent data from the environmental determinants of diabetes in the young (TEDDY) prospective study that found a correlation between fecal shedding of enterovirus and the development of islet autoantibodies [123]. In addition, infection of mice with coxsackie B virus initiated tissue destruction of the pancreatic islets

and a small subset of infected animals developed diabetes [124]. Despite this evidence, conflicting data shows that the link between enterovirus infection and T1D is far from clear. A well-documented epidemic of coxsackie virus in Alaska failed to raise the incidence of diabetes in the formerly infected [125]. Coxsackie B virus infection of NOD mice has also produced conflicting results; infection accelerated disease progression in one study [126] and protected animals in two others [127, 128]. These discrepancies make it difficult to ascribe a causal relationship between viral infection on a broad scale and the development of T1D. If viruses play a role in provoking diabetes, it is probably a minor one at the population level superseded by genetic susceptibility and other environmental factors.

Far stronger evidence suggests that infections protect against rather than promote T1D. Rising autoimmune and allergy incidence has long been associated with falling rates of major infections and of parasitic diseases [129]. In animal models, T1D incidence rises significantly when NOD mice bred in conventional facilities are decontaminated and moved to specific pathogen-free facilities (SPF) where disease burden is reduced [130]. In addition, a large number of pathogens have been experimentally shown to protect NOD mice from T1D [129]. Generally, protective mechanisms triggered by these pathogens are poorly understood and strongly depend on the age of the host. Most of the proposed mechanisms involve the activation of systemic immune regulatory mechanisms, upregulation of transforming growth factor β (TGFβ) [131] or the co-inhibitory receptor PD-1 on CD8$^+$ cells for example [128]. These regulatory mechanisms can also be triggered by systemic administration of microbial products without an active infection. Injection of NOD mice with complete Freund's adjuvant (CFA) (which is composed of dead mycobacteria), lipopolysaccharide (LPS) from Gram-negative bacteria, or polyinosine-polycytidylic acid [poly(I:C)], double-stranded RNA) protected NOD mice from T1D [132, 133]. These products are all innate immune agonists suggesting that the systemic activation of innate immunity may trigger regulatory mechanisms that are sufficient to halt T1D development and that major infections are protective for this very reason. As these protective mechanisms cannot be activated as easily from

inside the gut (orally delivered LPS for example [134]), it is possible that a limiting infection localized to the gut may have different effects on T1D development.

1.4.3 Influence of the commensal microbiota on T1D development

The non-pathogenic microorganisms that occupy various niches within a hosts body are known as the commensal microbiota. In exchange for shelter and for nutrients derived from the host diet, the microbiota provides the host with three key functions. First, the microbiota provides the host with essential nutrients such as vitamins and short chain fatty acids that are derived from the host diet [135]. These nutrients are essential and need to be supplemented into the diets of GF animals to prevent death [136]. Second, the microbiota provides the host a layer of defense against pathogens called colonization resistance. The mere presence of a dense layer of microorganisms (10^{12} bacteria per ml in the mammalian gut) makes it very difficult for invading pathogens to find space and nutrients to establish infection as they must compete with the commensal bacteria [137]. Third, the microbiota can influence the normal development of its host. In some invertebrates, such as *Drosophila*, the microbiota is required for normal body development [138]. In vertebrates, the lack of microbes results in abnormal development of the gastrointestinal tract and the immune system, the relative dearth of Th17 cells and colonic Tregs in GF compared to SPF housed mice being key examples [139, 140].

Our firmest evidence that the microbiota is involved in T1D development comes from animal studies. In Wen *et al.* 2008 [141], it was shown that NOD mice lacking myeloid differentiation primary response 88 (MyD88) protein, a key adaptor for toll-like receptor (TLR) signaling, were completely protected from T1D when housed in SPF conditions. However, when raised in GF conditions without exposure to microbes, MyD88 deficient NOD mice developed T1D at the same rates as their wild type (WT) counterparts. Importantly, the incidence in GF conditions for both groups was *higher* than that observed for WT NOD mice raised in SPF conditions. From this we can draw two important conclusions. First,

the microbiota is not required for the development of T1D and in most cases appears to be protective. Lower incidence in NOD mouse colonies at most SPF facilities supports this conclusion [42]. And second, the microbiota was responsible for protecting MyD88 deficient animals from T1D.

The lack of diabetes in MyD88 deficient NOD mice could be explained by two hypotheses put forward in 2013 [79]: Either specific microbial taxa expanded or contracted in MyD88 KO animals resulting in protection of the mice (specific lineage hypothesis), or microbes triggered unidentified protective pathways that MyD88 signaling was capable of overcoming. In the second case, the overall balance of pro-diabetic and anti-diabetic signals triggered by microbes determined the microbial influence on T1D in a particular host (balanced signal hypothesis). Followup experiments showed that multiple distinct microbial consortia were capable of protecting NOD mice that lacked MyD88 [142]. The same study also identified TIR-domain-adaptor-inducing-interferon-β (TRIF) and TLR4 (which signals both through MyD88 and TRIF) as protective pathways that are triggered by the microbiota. Importantly, MyD88 (probably diabetes promoting) and TRIF and TLR4 (protective pathways) are all activated directly or downstream of receptors that respond to conserved microbial features shared by pathogens, suggesting that commensals and pathogens may protect from diabetes *via* the same mechanisms. Cumulatively, these data supported the balanced signal hypothesis over the specific lineage hypothesis as particular taxa did not seem to matter, but changing the host's genetic sensitivity to different microbial signals did modulate diabetes incidence.

Studying the influence of the microbiota on T1D in humans has been much more challenging. T1D is usually not evident in humans until the development of hyperglycemia at the very end stages of disease. Microbial differences observed between diabetics and controls therefore are probably not causal but a consequence of the altered metabolism caused by diabetes. To overcome this, large prospective cohorts of genetically susceptible people must be tracked for many years to determine whether microbial differences predict later T1D. This second task is confounded by a second set of issues. Compared to laboratory rodents,

humans have much more diverse genetics, eat a more mixed diet that varies considerably between cultures and regions, and engage in different lifestyles (for example living with or without a pet) resulting in considerable microbial diversity between individuals [143]. This diversity might mask any microbial changes that contribute to diabetes. While the large prospective TEDDY study failed to find bacterial taxa strongly associated with T1D [144], it did find greater short-chain fatty acid production genes in diverse taxa from patients [80]. This finding underscores both the problem and a potential solution. When studying a disease from a microbe-centric point of view, multiple microbes present in different hosts may perform the same function for their hosts. Therefore, identifying the function that the microbes perform is much more useful than identifying the microbes that perform it.

1.4.4 Diet's influence on the commensal microbiota

Over large timescales the composition of the microbiota is remarkably stable. Fecal samples from humans collected regularly over a five-year period found that on average 60% of the microbial strains identified in each individual at the beginning of the study were present in that individual at the end [145]. Many of these strains (defined as 96% genome similarity) were also found in mothers or siblings suggesting that early maternally inherited colonizers may be especially capable of persisting long term. This similarity between family members is also explained by genetic control of microbiota composition. Mice from the same genetic background generally share more similar microbial compositions to each other than to mice from other strains [146] and human monozygotic twins have more similar microbial compositions than do dizygotic twins or unrelated individuals [147]. If genetics was the sole factor that influenced microbiota composition however, one would expect even more similar microbial communities in genetically identical individuals strongly suggesting that the environment heavily influences the microbiota. A long-term study of the microbiota in two individuals over the course of a year illustrates these influences [81]. While the microbiota stayed mostly stable over the course of a year in both individuals, large changes in community

structure occurred following a major gastrointestinal infection and following major changes to diet. In addition, mice fed the same diet had much more similar microbial communities than did mice fed a different diet regardless of the genetic strain [148]. Therefore, diet has an enormous effect on the composition and stability of the gut microbiota that trumps other factors such as genetics.

Diet induced changes to the microbiota have deep effects on host physiology. Probably the most pronounced example comes from the obesity field. The transfer of microbes from obese mice into lean GF mice caused the GF mice to become obese whereas the transfer of microbes from lean mice failed to do so [149]. Outside of obesity, diet induced changes to the microbiota can also effect the immune system. Feeding mice a ketogenic diet has been shown to deplete bifidobacterial species resulting in a loss of intestinal Th17 cells [150]. In another example, high fat diet induced reductions in Lactobacillaciae that were tied to reductions in gut *Reg3*γ antimicrobial peptide expression [151]. Whether diet induced changes to the microbiota influence T1D is an open question. Reciprocally, microbial changes to diet can also influence the host. Microbial fermentation of dietary fiber produces SCFAs which influence immunity, digestive function, hormone secretion and other aspects of host physiology [152] and, as mentioned above, there is growing evidence that SCFAs may protect rodents and humans from T1D [80, 94]. Whether, other microbial changes to diet influence autoimmunity is unknown and requires further study.

1.5 Aims of this study

The development of T1D is both influenced by diet and the microbiota. At the same time, dietary change strongly affects microbiota community composition and the metabolites it can produce resulting in physiologic changes for the host. Therefore, a given diet's influence on disease progression may be microbe-independent or microbe-dependent. To understand mechanistically how different diets affect T1D development, animals raised both with and without microbes must be fed these different diets and studied.

This thesis addresses the following questions:

1. Can diet affect T1D development independent from the microbiota or do diets rely on microbes to influence disease progression?

2. What effect do microbial modifications to diet have on T1D?

CHAPTER 2

MATERIALS AND METHODS

2.1 Animals

NOD/ShiLtJ (Stock No. 001976), NOD.Cg-Prkdcscid/J (NOD.SCID, Stock No. 001303), NOD.129P2(C)-Tcratm1Mjo/DoiJ (NOD.TCRa KO, Stock No. 004444), and NOD.Cg-Tg(TcraBDC2.5,TcrbBDC2.5)1Doi/DoiJ (NOD.BDC2.5, Stock No. 004460) mice were purchased from The Jackson Laboratory (Bar Harbor, ME;) and bred in house. B6.MyD88 KO mice were a gift from S. Akira and backcrossed to NOD/ShiLtJ mice for more than 10 generations as previously described [141] (NOD.MyD88 KO) and are currently bred in house. C57BL/10ScN mice were originally purchased from NIH (now at Jackson Labs, Stock No. 003752) and backcrossed to NOD/ShiLtJ for more than 10 generations as previously described [142] (NOD.TLR4 KO) and are currently bred in house. NOD.TGM2 KO mice were produced using CRISPR/Cas9 technology. Tgm2 exons 1 and 2 were targeted by the guides 5'-GTCGCCGCTAGCCTGG-3' and 5'- GATTTGGAGATTCAGGC-3' respectively to produce a dropout mutation between the two cut sites. Lack of TGM2 protein expression in the knockout mice was confirmed by western blotting (Fig. S3B, see methods below). All animals were kept under SPF conditions and NOD/ShiLtJ, NOD.MyD88 KO and NOD.TLR4 KO were kept under GF conditions at the University of Chicago Animal Resource Center. GF status was monitored by aerobic and anaerobic fecal cultures and PCR amplification of bacterial 16S rRNA genes from fecal DNA as previously described [153].

Gnotobiotic NOD mice were derived at 4 weeks of age from GF mice by introduction of bacteria via gastric gavage of cecal content taken from an SPF NOD/ShiLtJ mouse. Mice were placed on different diets and their cecal and small intestine contents were isolated after 8 weeks for 16S rRNA gene analysis. For colonization with E. faecalis, mating pairs were placed on Casein+4%gluten diet and gavaged with wild-type or mutant bacteria. Bacteria were transferred to the progeny naturally from the mother, and the efficiency of colonization

of the progeny was confirmed by PCR targeting the 16S rRNA encoding gene (using fecal DNA) followed by sequencing.

Purified diets were prepared by Envigo (Madison, WI) utilizing hydrolyzed (TD.120338) and intact (TD.120337) caseins from AMCO (Burlington, NJ). Gluten-containing diets (TD.150366 and TD.130340) were based on TD.120338 and contained 4% of gluten protein from wheat gluten supplied by ADM (WhetPro 80) (Gluten 1) or by Mallindra Milling corporation (Gem of the West) (Gluten 2), respectively. TD.170233 diet contained 21.7% egg-white solids completely replacing hydrolyzed casein in TD.120338. Diets were vacuum packaged and irradiated (20-50 kGy range) and tested for a microbial panel before being given to mice in germ-free and SPF housing. Standard chow diet in our facilities was NIH-31 modified open formula 7013 from Envigo. This diet contains 35.5% wheat (contains about 10% protein) and 10% wheat middlings (contain about 16% protein). Total wheat protein is about 5.2% in the chow diet. Assuming gluten constituting between 70% and 75% of wheat protein [154], its contents in chow diet is around 4%. That is similar to HC+4%gluten diet TD.130340 diet. In all experiments, the respective diets were administered to mothers throughout pregnancy and until pups were weaned and then they were kept on the same diets until 30 weeks of age.

Unless noted otherwise, female mice were used in all experiments. All experiments were performed in accordance with both The University of Chicago Animal Care and Use Committee and national guidelines.

2.2 Microbial strains

OG1RF - parental strain [155], (gelE+, sprE+); TX5266-OG1RF ΔfsrB, (gelE-, sprE-, direct mutagenesis) [156], TX5243 [157] (gelE+, sprE-, direct mutagenesis), a generous gift from Dr. Barbara Murray (UT Health Science Center, Houston), TX5264 [158] (gelE-, sprE+, direct mutagenesis) (JRC105-OG1RF Δ(gelE-sprE) (direct mutagenesis) [159] were grown anaerobically or aerobically in Bacto Brain Heart Infusion (BHI) (Becton, Dickenson

& Company, Sparks, MD) broth at 37°C with shaking for 16 – 18 hrs.

TX5264 mutant bacteria transformed with either empty shuttle vector containing nisin-inducible promoter pMSP3535 [160], or gelE gene in the same plasmid [161] were used to test the role of GelE in liberation of antigenic peptides from gluten. The cultures were supplemented with 10µg/ml Erythromycin +/- (Sigma) and 25ng/ml nisin (Sigma).

2.3 Diabetes testing

Diabetes development was monitored by weekly testing of urine glucose with Diastix strips (Bayer, Elkhart, IN).

2.4 Splenocyte transfer

Single cell suspensions were prepared from non-diabetic 10–13-week-old NOD female mice on either HC or regular chow diet. Red blood cells were lysed by suspension of cell pellets in 900μl sterile, distilled water (Gibco) for 10 seconds followed by addition of 100μl 10X PBS. Cells were then spun down and resuspended in PBS. 2x10^7 cells were injected intravenously into recipient mice. Recipient mice were monitored for diabetes development weekly by measuring urine glucose with Diastix (Bayer, Elkhart, IN).

2.5 T cell proliferation *in vivo*

CD4 T cells were enriched using anti-CD4 magnetic beads (Miltenyi Biotec, Auburn, CA) from splenocytes of NOD mice carrying the transgene for BDC2.5 TCR. T cells were labeled with 5μM carboxyfluorescein succinimidyl ester (CFSE) (Molecular Probes, Eugene, OR) for 15 minutes at 20°C followed by two washes in Click's Medium (Irvine Scientific, Santa Ana, CA) containing 5% FCS (Atlanta Biologicals, Lawrenceville, GA) followed by resuspension in PBS. 1.5-2x10^6 T cells were transferred intravenously to recipient mice on different diets. Pancreatic and non-draining lymph nodes were harvested 72 hours post transfer and analyzed

for CFSE dilution by FACS on an LSR-II flow cytometer and analyzed using FlowJo software (Treestar).

2.6 ELISPOT Analysis

2 x 10^5 PLN cells from NOD females fed either HC or HC+4% gluten and 4 x 10^5 NOD splenocytes, depleted of CD3ϵ cells (Miltenyi Biotec, Auburn, CA) according to the manufacturer's instructions, were incubated overnight with 1μg/ml NRPA7 mimic peptide recognized by 8.3 TCRs [162], 1μg/ml mimic peptide recognized by AI4 TCRs [163], or 10μg/ml InsB 15-23 peptide recognized by G9C8 TCRs [164] in 96-well Immobilon-P membrane plate (Millipore) pre-treated with anti-IFN-γ antibodies (BD Bioscience). Plates were treated and spots revealed according to the manufacturer's instructions. Background spots produced (no peptide control) were subtracted from the total spot count in wells with peptide. Results are an average of three parallels. PLN cells from each mouse were also stained with α-CD3ϵ, -TCRβ, -CD4, and -CD8α, antibodies (Biolegend) to calculate the frequency of peptide-specific cells per 10^6 CD8+ $\alpha\beta$T cells.

2.7 Intraperitoneal Glucose Tolerance Test (IPGTT) and Intraperitoneal Insulin Tolerance Test (IPITT)

IPGTT was performed as follows: mice were fasted for 12 h and tested for baseline of blood glucose levels prior to glucose administration. Mice were then injected i.p. with a 20% glucose solution in PBS (2 g/kg body weight) and blood glucose levels measured by glucometer before and 15, 30, 60, 90 and 120 minutes after injection. Data are presented as the difference in glucose levels between each timepoint and time 0. IPITT was performed by fasting mice for 4h prior to the test. Baseline glucose measurement was made before i.p. injection of 1U/kg insulin (Sigma) followed by measurement of glucose at 15-, 30-, and 60-min. Data are expressed as percent of starting glucose levels.

2.8 Measurement of serum insulin levels following glucose challenge

Mice were fasted for 8 hours and injected i.p. with 20% glucose solution in PBS (2 g/kg body weight). After 5 minutes, cheek blood was collected in sterile collection tubes. Serum was spun down at 8,000 g. for 8 minutes at 4°C to remove remaining cellular debris. Insulin concentration was measured using the Mouse Insulin ELISA Kit (Alpco) according to the manufacturer's instructions.

2.9 Isolation of murine pancreatic islets of Langerhans

Mice were killed by cervical dislocation and the pancreata were perfused with 5 ml of collagenase P solution [0.5 mg/ml, (Roche) in HBSS (Gibco) supplemented with 0.1 M HEPES, 1 mM $MgCl_2$, 5 mM D- glucose, 1.25 mM $CaCl_2$, and 0.02% BSA] via the common bile duct. Perfused pancreata were then dissected away from the surrounding tissue and incubated for 10 minutes at 37°C. Digestion was halted by washing the tissue twice with supplemented HBSS. The tissue was then applied to a Histopaque-1119 (Sigma) gradient, and the interface was collected, washed in supplemented HBSS, and then placed in Click's medium with 5% fetal bovine serum (FBS). Islets were handpicked using an inverted microscope. For single-cell RNA sequencing experiments, handpicked islets were dissociated by incubation in Trypsin-EDTA (Gibco) at 37° C for 2 minutes followed by washing in Click's media containing 10% FBS. For flow cytometry of islet infiltrates, handpicked islets were dissociated by incubation in enzyme free cell dissociation buffer (Gibco) for 3 minutes at 37° C followed by washing in Click's media containing 5% FBS.

2.10 Insulitis scoring from histology sections

Pancreata from 13 week old mice fed different dets were dissected with fixed overnight in Telly's fixative at room temperature before being transferred to 70% ethanol for longer term storage. Fixed pancreata were dehydrated, embedded in paraffin blocks, sectioned into 20 5μm sections spaced 40μm apart, and stained with hemotoxylin and eosin by the University of Chicago Human Tissue Resources Center. Insulitis was graded in a blinded fashion using the following scoring system: 0, no visible infiltration; I, periinsulitis; II, insulitis with $<$50%; and III, insulitis with $>$50% islet infiltration. At least 10 islets per section and at least 100 islets per animal were scored.

2.11 Single cell RNA Sequencing

Cells were loaded onto a Chromium Controller (10X Genomics) [165] and processed using the Chromium Single Cell 3' Library & Gel Bead Kit (10X Genomix, v3) according to the manufacturer's protocol. cDNA and library quality were assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Samples were sequenced using the Ilumina HiSeq 4000 (Ilumina, San Diego, CA). All single cell barcoding, library preparation, quality assessment, and sequencing were performed at the Functional Genomics Facility at the University of Chicago. For experiments using NOD wild-type mice, sequencing was done on one pool per biological condition with three 15-week-old mice in each pool. For experiments using NOD.scid mice, three individual 8-10-week-old mice per biological condition were sequenced. One HC fed NOD.scid sample was removed from the analysis due to very low numbers of reads per cell. Raw sequencing files were processed through the CellRanger pipeline (version 4.0.0) to conduct all alignment and gene expression quantification steps using the mm10 reference mouse genome.

2.12 Single-cell data processing

Single cell gene expression data were input into the Seurat Package (version 3.1.5) [166] using R (version 3.6.3). Doublets and dead cells were removed from further analysis by filtering out cells with detectable gene expression in less than 200 genes or greater than 7500, and mitochondrial gene content represented greater than 25% of all genes expressed. Gene expression count matrices were normalized using a global-scale normalization function LogNormalize. Variable features for downstream analysis were identified using the FindVariableFeatures function using the default of 2000 variable features. Cell clustering was performed by first transforming the data with the ScaleData function, performing linear dimensional reduction using RunPCA, and then FindClusters and RunUMAP were used to finish the cell clustering and project the visualization into a two-dimensional space. Cell types were annotated using variable genes that drove separation between different clusters (*Ins1*- β-cells; *Gcg*- α-cells; *Sst*- δ-cells; *Cd3d*- T cells; *Ms4a1*- B cells; *Adgre1*- Macrophages; *Cd209a*- pDC's; *Cdh5*- Endothelial cells; *Col1a1*- stromal cells; also see **Figure 4.4B**). Due to low cell numbers, DC's and macrophages were grouped together as myeloid cells for subsequent analysis. Seurat clusters defined by canonical markers from multiple cell types (*Cd3e* and *Cd79a* for example) were deemed aggregates of multiple cells and removed from the analysis. For comparison of *Ins1* High and *Ins1* Low β-cells in chow fed NOD.SCID mice, *Ins1* High and Low β-cells were defined as β-cells with *Ins1* levels of above 9.0 or below 8.5 log total sum scaled counts respectively. Kernel density estimation was used to define the probability density function of beta cells based on *Ins1* and *Ins2* expression levels within each diet.

2.13 Differential gene expression analysis of single-cell RNA sequencing data

Differential gene expression analysis was preformed using the scanpy Python package (version 1.4.5.1) [167]. Batch effects were removed using scanpy's implementation of ComBat, using the sample sequencing run dates as the batch keys. Differential expression was tested for using the rank_gene_groups function with the "t-test_overestim_var" method. P-values were adjusted using the Benjamini-Hochberg method and significant genes were selected as those with a log fold change magnitude greater than 0.5 and an adjusted p-value less than 0.05.

For further analysis of *Ins1*$^{\text{High}}$ vs *Ins1*$^{\text{Low}}$ β-cells, genes upregulated in *Ins1*$^{\text{High}}$ cells (log2 fold change > 0.4, adjusted p-value < 0.05) were input into the STRING database [168], which analyzes known and predicted gene-gene interactions. Genes with no connections were removed from the analysis. From this analysis, 3 clusters were identified that contained genes involved in 1) Insulin processing and secretion; 2) Unfolded protein response and endoplasmic reticulum (ER) associated protein degradation; and 3) Protein processing in the ER. GSEA analysis was carried out using GSEA software v4.1.0 [169, 170] using the publicly available Hallmark gene sets (h.all.v7.4.symbols.gmt).

2.14 TCR sequencing

Single cell suspensions of islets from 3 HC fed and 3 HC+4%gluten fed 16 week old NOD mice were washed in 1X PBS supplemented with 1% FBS and 0.05% NaN_3. To block Fc receptors, cells were incubated in 1X PBS supplemented with 1% FBS and 0.05% NaN_3 and anti-mouse CD16/32 antibodies (Biolegend, 101301) for 10 minutes at 4° C. Individual samples were then incubated with 2.5 μg/ml of a unique Total-SeqTM-C antibody (Biolegend; see Key Resources Table) for 30 minutes at 4° C and then washed again with 1X PBS supplemented with 1% FBS and 0.05% NaN_3. Cells were loaded onto a Chromium

Controller (10X Genomics) [165] and processed using the Chromium single cell mouse TCR amplification kit and the Chromium single cell 5' library construction kit (10x Genomix) according to the manufacturer's protocol. cDNA and library quality were assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Samples were sequenced using the Ilumina NovaSeq 6000 (Ilumina, San Diego, CA). All single cell barcoding, library preparation, quality assessment, and sequencing were performed at the Functional Genomics Facility at the University of Chicago. Gene expression data was imported using the scanpy package [167] in Python for two rounds of sequencing. Cell counts from cells captured in both sequencing events were aggregated together and the rest were discarded. Any cell whose second largest hashtag read count was at least 75% of the largest hashtag read count was filtered as well. The remaining cells were annotated to mice using the hashtag with the largest count value. Expression values were total sum scaled to 10,000 and then log normalized using a pseudocount of 1.

VDJ data from the two runs were combined and integrated into the single cell expression data using the scirpy package [171] in Python. We retained only cells with high confidence VDJ calls containing a productive single productive chain. Using the subset of cells taken from the islet, we separated the set of cells into CD4 and CD8 populations. We then identified expanded CDR3 alpha and CDR3 beta sequences within each population by identifying sequences found in three or more cells within at least one mouse. Cells expanded based on CDR3 alpha sequences or CDR3 beta sequences were aggregated respectively based on diet. Chi-squared tests were used to identify sequences with significantly different enrichment proportions between HC and HC+4% Gluten. The distributions of all expanded sequence clone sizes were then visualized using the seaborn package in Python.

Clonotype clusters for CDR3 alpha and CDR3 beta sequences were identified using pairwise alignment distance of the amino acid sequence using the BLOSUm62 scoring matrix [172]. Clusters were identified using a distance threshold of 5. The network of clonotype clusters was generated and visualized using clonotype_network from the scirpy package in

Python. This analysis was done on CD4 and CD8 populations separately.

2.15 Flow cytometry

Single cell suspensions were incubated with Zombie Near IR (Biolegend, 423105) in 1X PBS for 10 minutes at room temperature for Live/Dead discrimination. Cells were then washed in 1X PBS supplemented with 1% FBS and 0.05% NaN_3. To block Fc receptors, cells were incubated in 1X PBS supplemented with 1% FBS and 0.05% NaN_3 and anti-mouse CD16/32 antibodies (Biolegend, 101301) for 10 minutes at 4° C. For surface staining, cells were incubated for 30 minutes at 4° C with fluorescently tagged antibodies. For analysis of islet infiltrates, cells were stained with α-CD45 (30-F11), -CD4 (GK1.5), -CD19 (1D3), -FR4 (12A5), -PD-L1 (MIH5), -CD62L (MEL-14), -CD80 (16-10A1), PD-1 (J43), and -CD11c (N418) all from BD Bioscience, and α-CD44 (IM7), -CD103 (2E7), -CD69 (H1.2F3), -CD11b (M1/70), -F4/80 (BM8), -CD86 (PO3), -TCRβ (H57-597), and -CD8α (53-6.7) all from Biolegend. For analysis of stimulated PLN macrophages, cells were stained with α-CD19 (1D3) from BD Bioscience, and α-CD11c (N418), -CD11b (M1/70), -F4/80 (BM8), and -TCRβ (H57-597) from Biolegend. Cells were then fixed and permeabilized using the Foxp3/Transcription Factor Staining Buffer Set (ThermoFisher) according to the manufacturer's instructions. Cells were then stained with fluorescently tagged intracellular antibodies overnight at 4° C. For analysis of islet infiltrates, cells were intracellularly stained with α-FoxP3 (FJK-16S) from Thermo/Fisher, α-CTLA-4 (UC10-4B9), and -Ki67 (16A8) from Biolegend, and α-Insulin (182410) from R&D Systems. For analysis of stimulated PLN macrophages, cells were intracellularly stained with α-TNFα (MP6-XT22) from Biolegend. Flow cytometry on islet infiltrates and stimulated pancreatic lymph node macrophages was done on a Cytek Aurora Instrument (Cytek). Raw data was unmixed using single color controls made using UltraComp ebeads^{TM} Compensation beads (ThermoFisher) prepared according to the manufacturer's instructions or cells. Data was analyzed using FlowJo software (Treestar) with the UMAP plugin [173].

2.16 Western Blotting

Western blotting was performed using thymic lysates in RIPA buffer (Stem Cell Technology) supplemented with protease inhibitor cocktail (Pierce Thermo-Fisher) and separated by gel electrophoresis (20μg of protein/lane). Blots were stained with anti-TGM2 primary antibodies (R&D Systems,) or mouse anti-actin (clone C4, Santa Cruz Biotechnology) followed by secondary HRP-anti-sheep Ig conjugate (R&D Systems) or by HRP-sheep-anti-mouse-Ig conjugate (Amersham) respectively and developed using Super Signal West Femto Maximum Sensitivity Substrate (Thermo-Fisher).

2.17 Bacterial DNA sequencing and analysis

Small intestine and cecal contents of female mice were collected using sterile instruments into cryovials and snap frozen in liquid nitrogen until processing. Samples were processed and sequenced at the Environmental Sample Preparation and Sequencing Facility at Argonne National Laboratory. DNA was extracted from the samples using the DNeasy PowerSoil HTP 96 Kit (QIAGEN, Germantown, MD) according to the manufacturer's instructions. The V4 region of the 16S rRNA-encoding gene (515F-806R) was PCR amplified to survey the total bacterial community in the extracted samples using the Illumina MiSeq platform as described [174, 175]. PCR reactions were carried out in triplicate for each sample using sterile, DNase-free 96 well plates with appropriate (DNA template-free) negative controls using the 5 PRIME MasterMix (5 PRIME, Gaithersburg, MD). PCR reactions were conducted using an initial denaturation step of 95C for 3 minutes, followed by 35 cycles at 95°C for 30 seconds, 55°C for 45 seconds, then 72°C for 1.5 minutes. A single extension step at 72^{o}C for 10 minutes was used at the end. Triplicate PCR reactions were then pooled together, primer dimers were removed from the pooled products using the UltraClean 96 PCR Cleanup Kit (QIAGEN) and total DNA quantified using the PicoGreen® dsDNA Assay (Life Technologies) was resuspended at 2 ng/μl. Amplicons were sequenced on an Illumina

MiSeq using 151×151 base pair paired-end sequencing. Raw sequence data was processed using the QIIME 1.9 analysis pipeline [176]. Initially, paired-end reads were joined and then aligned to a database of reference sequences using PyNAST [177]. These were clustered into operational taxonomic units (OTUs) using UCLUST at 97% similarity [178] and a consensus taxonomy from the Greengenes reference database [179] was assigned to each sequence using the UCLUST taxonomy assigner.

Statistical analysis of the 16S dataset was carried out with the MATLAB (MathWorks, Natick, MA, USA) software package. Principal component analysis (PCA) of the microbiomes was performed using the relative abundance of genera from each sample. The percent variance explained by each principal component was calculated and reported in corresponding figures. Genera with the highest relative abundance were identified based on mean expression across all mice in small intestine and cecum. The top ten genera were visualized using a box plot to demonstrate differences among diets.

2.18 Detection of gluten digesting bacteria

Serial 10-fold dilutions (from 10^{-3} to 10^{-7}) of mouse small intestine contents were plated on BHI-gliadin agar plates (Sigma) and on minimal medium-gluten agar plates [180] and the plates were incubated anaerobically for 2 – 6 days at 37°C. Gliadin-digesting colonies were detected by transparent halos formed around them in the cloudy gliadin-agar layer [181]. Colonies with transparent halos were then subcultured onto agar plates supplemented with gliadin ethanol extracts to confirm enzymatic activity. For identification of bacteria, DNA was extracted from individual colonies and 16S rDNA amplified and analyzed by Sanger sequencing.

2.19 Preparation of gluten digests

Gluten powder (3%, w/vol) was suspended in the appropriate liquid phase: BHI, Dulbecco's Modification of Eagle's medium (DMEM, Gibco), or bacterial culture supernatants grown in the above media. When required, 1% by volume of the log phase bacterial culture was added to the mix. The samples were incubated at 37°C for 18-20 hours with sufficient shaking to maintain gluten particles in suspension. Samples were then centrifuged to precipitate the non-soluble matter, filter-sterilized with 0.22μm membrane filters and frozen at -80°C. Resulting supernatants were used for further analysis. For heat inactivation experiments, samples were incubated at 85°C for 30 minutes. For trypsin experiments, trypsin-EDTA (Gibco) was applied to the samples at a final concentration of 0.05% and samples were subsequently incubated overnight at 37°C. For polymyxin B treatment, 100 μl of polymyxin B-Agarose beads (Sigma) per ml of sample were incubated with prepared supernatants at 4°C overnight with gentle rotation. Beads were removed by centrifugation at 800g for 10 minutes followed by passing the supernatant through a 0.22μm membrane filter.

2.20 Activation of gliadin specific T cells *in vitro*

Mouse T cell line 578_BV7_AV12 was made by engineering murine BW58α-β cells devoid of endogenous TCR to express human CD4 and TCR from a DQ2.5-glia-2–reactive CD4 T cell clone (TCC 578) or a DQ8-glia-α1-reactive CD4 T cell clone (TCC 489) as previously described [182, 183]. IL-2 secretion was measured after overnight incubation of 5x10^4 T cells with 5-10x10^3 human lymphoblastoid cell line expressing HLA-DQ2.5 (a generous gift from Dr. Carole Ober, The University of Chicago) in 200μl of Click's medium with 5% fetal calf serum. IL-2-secretion was elicited by addition of 10% (v/v) or less of bacterial digests of wheat gluten. IL-2 levels were determined by ELISA using a kit (Biolegend) according to manufacturer's instructions.

2.21 Stimulation of macrophages with gluten digests

Peritoneal macrophages were isolated 4 days after intraperitoneal administration of 1.5 ml of Thioglycolate (Difco). Macrophages were plated in Click's medium at 10^5 cells per well in a 96 well flat-bottom plate. Sterile filtered supernatants prepared as above were applied to the macrophages at a maximal concentration of 10% of the final culture volume unless specified otherwise. Supernatants were collected at hour 6 to assay TNF production and after overnight culture (18-20) hours to assay IL-6 production. TNF and IL-6 were measured using ELISA kits (BD and Biolegend) according to the manufacturers' instructions.

2.22 Stimulation of pancreatic lymph node macrophages

Pancreatic lymph nodes from 6–10-week-old NOD mice were incubated with Collagenase D (1.78 mg/ml, Roche 11088858001) and DNase I (0.1 mg/ml, Sigma DN25) in Click's media with 10% FBS for 30 minutes at 37°C and then passed through a nylon mesh to release macrophages. Single cell suspensions were then plated at 10^6 cells per well of a non-TC treated, round bottom 96 well plate and incubated for 6 hours with GolgiPlug only (1:500 vol/vol, BD555029), GolgiPlug and E. faecalis gluten digest (10% final culture volume), or GolgiPlug and 1 ng/ml of LPS. Cells were then washed and stained with fluorescently labeled antibodies.

2.23 Limulus amebocyte lysate (LAL) assay

LPS concentrations were determined with kinetic-chromogenic test kit Endochrome-K (Charles River Endosafe, Charleston, SC) according to the manufacturer's protocol. All reagents and components used for the assay were endotoxin-free. The LAL Reagent was rehydrated with Endotoxin-Specific Buffer Solution to prevent interference of β-D-Glucans with LAL-reactive material during testing. *E. coli* Control Standard Endotoxin (Charles River Endosafe) was used for the standard calibration curve dilutions. 0.1 ml of each sample

was transferred in a well of a LAL test-certified 96-well microplate (Greiner Bio-One GmbH, Frickenhausen, Germany) and then 0.1 ml of LAL reagent was quickly added to each well. The contents of the wells were mixed by tapping and the plate was placed in the VERSAmax microplate reader (Molecular Devices, Sunnyvale, CA). The 450 nm absorbance values were recorded for each well every 15s for 1 hour at room temperature. Onset times at 0.200 OD units were used to calculate the LPS standard curve and the sample LPS concentration.

2.24 Production and screening of gluten reactive T cell hybridomas

To generate bacterially digested gluten reactive T cells, gluten digests were prepared as above using *E. faecalis* mutant TX5243 (gelE+ sprE-) [157] and subsequently passed through a 10kD Amicon Ultra-4 centrifugal filter unit to remove proteases, intact bacterial cell wall components, and large undigested proteins. The digest's 10kD filtrate was then emulsified 1:1 with CFA and injected in a 50 μl volume into each hind foot pad of NOD mice fed a 20% casein diet devoid of gluten. 10 days later, popliteal lymph nodes were taken from the mice, dissociated between the frosted ends of 2 sterile microscope slides, and restimulated in culture with irradiated splenocytes and the 10kD gluten digest filtrate (used at 10% of total culture volume).

On restimulation day 3, the T cells were fused with BW5147 (BW) cells. In short, restimulated T cells and BW cells were washed 2 times each in 37°C serum free Click's media and then mixed together at a 2:1 T cell:BW cell ratio in autoclaved scintillation vials. The cells were then spun down at 2,000 rpm for 5 minutes to form a thin layer of cells at the bottom of each vial. The vials were then incubated at 37°C for 20 minutes after which the supernatants were decanted with care taken not to disturb the layer of cells at the bottom of each vial. Cells were fused by slowly applying 1 ml of 37°C polyethylene glycol-1500 (PEG-1500, Sigma) to the cell layers and incubating them for exactly 70 seconds. 10 ml of 37°C

serum free Click's media was then added in a slow gentle stream to each cell layer to wash out the PEG-1500 and each vial was spun down again at 2,000 rpm for 3 minutes. Supernatants were again carefully decanted and the cell layers were washed 2 additional times with serum free Click's media. After the third wash, 5ml 37°C Click's media with 5% FBS was added to each vial, again taking care not to disturb the cell layers. The cells were incubated at 37°C for 45 minutes after which, they were gently resuspended, counted, and plated at a cell density of 30,000 cells per well of a 96-well flat bottom plate in Click's media with 5% FBS and incubated at 37°C overnight. The next morning hypoxanthine (100 μM) aminopterin (0.4 μM) thymidine (16 μM) (HAT, final concentrations, Gibco) selection media was added to each well and the hybrids were allowed to grow out. Viable hybridomas were transferred to individual wells of 24 well, flat bottomed plates in Click's media with 5% FBS and HAT selection media.

To screen hybridomas for specificity to bacterially digested gluten, individual hybridomas were stimulated to produce IL-2 overnight with 10^5 NOD splenocytes either in media alone or with supernatants from *E. faecalis* grown with or without gluten applied at 10% of the culture volume. IL-2 concentrations in the supernatants was measured using an IL-2 ELISA kit (Biolegend) according to the manufacturer's instructions. Hybrids that failed to react to any stimulation were characterized as not reactive. Hybrids that reacted to splenocytes without added antigen were characterized as self reactive. Hybrids that reacted to *E. faecalis* grown without gluten and with gluten were characterized as *E. faecalis* reactive. Finally, hybrids that only reacted to *E. faecalis* digested gluten were characterized as gluten reactive.

2.25 Quantification and statistical analysis

Statistical analysis of diabetes incidence, cell proliferation, blood glucose levels, flow cytometry, and ELISA results were performed using Prism 9 (GraphPad). All bar-graphs and dot-plots are expressed as Means ± sem. The statistical difference between any two groups was determined by Student's t test. For multiple groups the statistical difference

was determined using the one-way analysis of variance (ANOVA) with post-hoc Tukey test. Statistics for T1D survival data were determined by Mantel-Cox long-rank test using Prism 9 (GraphPad). Incremental Area Under the Curve (iAUC) for IPGTT results were calculated in Prism 9 from curves drawn from the difference in glucose levels at each timepoint and time 0. p-values or adjusted p-values < 0.05 were considered statistically significant.

Table 2.1: Antibodies used for Flow Cytometry Experiments

Antigen	Color	Clone	Manufacturer	Dilution
Islet Infiltrates				
CD45	Brilliant UltraViolet 395	30-F11	BD	1:400
CD4	Brilliant UltraViolet 496	GK1.5	BD	1:400
CD19	Brilliant UltraViolet 563	1D3	BD	1:400
FR4	Brilliant UltraViolet 737	12A5	BD	1:400
PD-L1	Brilliant Violet 480	MIH5	BD	1:400
CD44	Brilliant Violet 510	IM7	Biolegend	1:200
CD62L	Brilliant UltraViolet 805	MEL-14	BD	1:400
CD103	Brilliant Violet 605	2E7	Biolegend	1:100
PD-1	Brilliant Violet 650	J43	BD	1:200
CD80	Brilliant Violet 711	16-10A1	BD	1:400
CD69	Brilliant Violet 785	H1.2F3	Biolegend	1:400
CD11c	Brilliant Blue 515	N418	BD	1:200
CD11b	PerCP-Cy5.5	P84	Biolegend	1:400
F4/80	PE-Dazzle 594	BM8	Biolegend	1:400
CD86	PE-Cy7	PO3	Biolegend	1:400
TCRβ	AlexaFluor 700	H57-597	Biolegend	1:200
CD8α	APC-Fire 750	53-6.7	Biolegend	1:400
FoxP3	efluor-450	FJK-16S	Thermo/Fisher	1:300
CTLA-4	PE	UC10-4B9	Biolegend	1:50
Insulin	APC	182410	R&D Systems	1:100
Ki67	AlexaFluor 647	16A8	Biolegend	1:3200
Stimulation of PLN Macrophages				
CD19	Brilliant UltraViolet 563	1D3	BD	1:400
CD11c	APC	N418	Biolegend	1:200
CD11b	FITC	P84	Biolegend	1:400
F4/80	PE-Dazzle 594	BM8	Biolegend	1:400
TCRβ	AlexaFluor 700	H57-597	Biolegend	1:200
TNFα	Brilliant Violet 421	MP6-XT22	Biolegend	1:400

Table 2.2: Genotyping and 16S Amplicon Sequencing Primers

Primer Name	Sequence	Use
NOD BDC2.5 F	5'-ATGCACAGCCTCCTGGGGTTGTTG-3'	NOD.BDC2.5 Genotyping
NOD BDC2.5 R	5'-TGTTCCTGTCCCGAAGGTAAGT-3'	NOD.BDC2.5 Genotyping
TCRα KO F	5'-GCACTCCTGATAGCCATAAG-3'	NOD.TCRα KO Genotyping
TCRα KO R	5'-TCTATCGCCTTCTTGACGAG-3'	NOD.TCRα KO Genotyping
TGM2 KO Ex1F	5'-GTGATCCTCGCTTGAGTGTC-3'	NOD.TGM2 KO Genotyping
TGM2 KO Ex1R	5'-GAATCTGGGCTTGGAGAACAC-3'	NOD.TGM2 KO Genotyping
TGM2 KO Ex2F	5'-CACTGTGCTTCACTCACATTG-3'	NOD.TGM2 KO Genotyping
TGM2 KO Ex2R	5'-CAGTACACACAGGCTTGCAG-3'	NOD.TGM2 KO Genotyping
8F	5'-AGAGTTTGATCCTGGCTCAG-3'	Bacterial 16S sequencing
1391R	5'-GACGGGCGGTGWGTRCA-3'	Bacterial 16S sequencing

CHAPTER 3

MICROBE INDEPENDENT PROTECTION FROM TYPE-1 DIABETES BY DIETARY REDUCTION OF β CELL STRESS.

3.1 Preface

Diabetes incidence in NOD mice fed different diets was monitored by L.A. Yurkovetskiy and A.V. Chervonsky. Spleen transfer experiments and *in vivo* proliferation assays with BDC2.5 T cells were started by L.A. Yurkovetskiy and completed by myself. L.A. Yurkovetskiy and C.H.F. Hansen performed glucose and insulin tolerance tests on mice fed different diets in this chapter. Glucose tolerance tests performed on mice from different genetic backgrounds and single-cell RNA-sequencing data collection, cell clustering analysis, identification of *Ins1*-High and *Ins1*-Low β-cells, and STRING network analysis of genes upregulated in *Ins1*-High β-cells were performed by myself. Lastly, differentially expressed genes were identified in collaboration with D. Reiman and A.A. Khan.

3.2 Introduction

T1D is a debilitating autoimmune disease caused by the progressive destruction of the islets of Langerhans by the immune system. While the disease is primarily driven by genetic factors, rapidly rising incidence globally strongly suggests that environmental factors are also involved [72]. One such environmental factor thought to contribute is diet. Dietary modifications have been shown to influence disease incidence in animal models of disease and are suspected to do so in humans [77, 88, 89, 92, 97, 112]. Another environmental factor that contributes to T1D is the commensal microbiota [80, 141, 142]. Given that the microbiota rapidly adapts to dietary changes [81, 148, 184], it is possible that diet may influence T1D by changing microbial community composition. If that were true, the diversity in human microbiota composition would make the development of prophylactic

dietary inventions impossible. Therefore, the identification of protective diets that function independent from the microbiota is of utmost importance.

To determine whether dietary interventions can succeed regardless of microbial influence, we studied NOD mice fed a diet containing hydrolyzed casein (HC) as the sole amino acid source. HC diets have been previously shown to protect NOD mice in SPF conditions and BB rats in both SPF and GF conditions from T1D [96–99] suggesting that its protective effect may be microbe independent. Confirming this, we found that HC also protects NOD mice raised in the absence of microbes in GF conditions. Instead of eliminating autoreactive T cells, HC's protective effect was based on reductions in β-cell stress due to reduced insulin secretion. Overall, this work illustrates the importance of targeting organ health in autoimmune pathogenesis and can serve as a model for how to identify protective diets for use in the prevention of T1D.

3.3 Results

3.3.1 Casein diet protects NOD mice from Type-1 Diabetes independent from the microbiota

To test whether HC's protective effect was microbe independent, female NOD mice in SPF or GF conditions were fed a diet containing HC, the same diet containing intact casein (IC), or standard chow and their offspring were monitored for T1D development while maintained on the same diets **(Figures 3.1A and B)**. As previously shown, NOD mice fed HC or IC in SPF conditions had significantly reduced incidence of T1D and disease onset was delayed in mice who did develop hyperglycemia. In GF conditions, both HC and IC fed mice were also protected compared to their chow fed counterparts. Therefore, protection from T1D by casein was microbe independent.

To determine whether protection by HC was due to changes to the immune system, splenocytes from chow-fed or HC-fed mice were transferred into chow-fed immunodeficient

NOD.SCID **(Figure 3.1C)** or NOD.*Tcr*αKO **(Figure 3.1D)** recipients and T1D development was monitored. In both groups, mice receiving splenocytes from HC-fed donors or from chow-fed donors progressed to T1D at the same rate. Thus, HC did not produce an irreparable defect in the autoimmune effector cells that drive β-cell destruction. In reverse experiments, splenocytes from chow-fed mice were transferred into NOD.*Tcr*αKO mice fed either chow or HC **(Figure 3.1E)**. Compared to the chow-fed recipients, T1D incidence was reduced and onset was delayed in the HC-fed recipients. In addition, islet reactive TCR-transgenic $CD4^+$ BDC2.5 T cells [51, 185] transferred into WT NOD mice proliferated less in the pancreatic draining lymph nodes of mice fed HC diet **(Figure 3.1F)**. We therefore concluded that HC does not change the ability of NOD mice to generate diabetogenic effector T cells but rather renders the islets or their local environment less stimulatory to the immune system.

3.3.2 Control of glucose tolerance in NOD mice by genetics and diet

How could HC make the islets less stimulatory to the immune system? One possibility was that HC affects β-cell insulin secretion. Insulin is a major autoantigen driving the development of T1D. Numerous diabetogenic $CD4^+$ T cell clones [186] and a $CD8^+$ T cell clone [50, 164] isolated from islets react to peptides in the insulin B chain between amino acids 9 and 23. Also, replacing the native *Ins1* and *Ins2* genes in NOD mice with an insulin transgene containing a tyrosine to alanine substitution at insulin B chain position 16 completely protected NOD mice from diabetes, insulitis, and the development of autoantibodies [187]. In addition, NOD mice secrete significantly more insulin in response to glucose challenge than non-diabetes prone C57BL/6J (B6) mice, a phenotype that was independent of islet infiltration [188]. To further understand the genetic control of insulin secretion in NOD mice, we crossed immunodeficient NOD.SCID mice with B6 mice carrying both the NOD MHC locus and SCID mutation (B6g7.SCID) and subjected the offspring to intraperitoneal glucose tolerance tests (IPGTT) (**Figure 3.2A, B**). In IPGTT, faster clearance of blood

glucose (high glucose tolerance) serves as a proxy for high insulin secretion. Female F1 mice produced from either cross, [F1(female NOD x male B6g7) or F1(female B6g7 x male NOD)] had high glucose tolerance similar to female NOD mice meaning that high glucose tolerance in NOD mice was controlled by a dominant, non-mitochondrial gene, not located within the NOD MHC locus. To identify the number of genes controlling the trait, F1.SCID mice were backcrossed to B6g7.SCID mice to produce N2 mice. These offspring were then subjected to IPGTT (**Figure 3.2B**). Approximately, 50% of the N2 hybrids had high glucose tolerance like NOD mice and the other 50% had low glucose tolerance like B6g7 mice indicating that high glucose tolerance in NOD mice may be controlled by a single gene.

It should be noted that glucose tolerance in mouse strains operates on a spectrum rather than on a high-low binary with some strains exhibiting high glucose tolerance like NOD mice and other strains exhibiting low glucose tolerance like B6 mice [189, 190]. In our hands **(Figure 3.2C, D)**, NOD mice were similarly glucose tolerant to the non-diabetic strain BALB/cJ but was significantly more glucose tolerant than FVB/NJ mice. High glucose tolerance (and by proxy high insulin secretion) is therefore not a unique trait to NOD mice but may be one of many traits that contributes to progression to islet autoimmunity.

To test whether HC diet affected insulin secretion in NOD mice, chow fed and HC fed NOD mice were subjected to IPGTT **(Figure 3.3A, B)**. Compared to chow fed NOD mice, HC fed NOD mice were significantly less glucose tolerant. The reduction in glucose tolerance imposed by HC was not due to reduced insulin resistance. HC fed mice injected with insulin showed a *more rapid* reduction in blood glucose levels than chow fed mice **(Figure 3.3C, D)** indicating that HC fed mice were simultaneously less glucose tolerant and more insulin sensitive than chow fed NOD mice. The reduction in glucose tolerance was also not due to reduced damage to the islets as HC fed NOD.$TCR\alpha$ KO mice still had reduced glucose tolerance compared to chow fed NOD.$TCR\alpha$ KO mice **(Figure 3.3E, F)**. Finally, direct measurements of plasma insulin concentrations 5 minutes after glucose challenge revealed that NOD mice fed HC secreted significantly less insulin in response to glucose than NOD

mice fed chow **(Figure 4.2C)**. Therefore, HC reduced the genetically predetermined high glucose tolerance in NOD mice by reducing islet insulin secretion.

3.3.3 Dietary reduction in insulin secretion leads to a reduction in β-cell stress

High insulin secretion may play a role in the pathogenesis of both type-1 and type-2 diabetes due to its association with β-cell stress. In a model of T2D progression, rising insulin resistance leads to rising demand for insulin production; eventually the β-cells become dysfunctional due to the chronically high insulin demand and begin to die [191]. β-cell failure during high demand for insulin may be due to endoplasmic reticulum (ER) stress caused by the unfolded protein response (UPR) which becomes upregulated in β-cells in insulin resistant mice [192]. Interestingly, NOD mice contain variants of two genes, *Glis3* and *Xrcc4*, that render their β-cells highly susceptible to apoptosis during the UPR [68]. This means NOD mice, independent of autoimmune infiltration, may have β-cells that are particularly sensitive to ER stress, especially given the strain's genetic predisposition for high insulin secretion.

To assess whether HC diet reduces β-cell stress, we performed single-cell RNA-sequencing (scRNA-seq) analysis on islets from NOD.SCID mice fed either chow or HC. SCID mice were used to exclude the influence of immune infiltrates on islet gene expression. Zeroing in on the β-cells **(Figure 3.4A, B)**, we were initially surprised to find that 0 genes were differentially regulated between β-cells from chow fed and HC fed mice **(Figure 3.4C)**. We reasoned that this could be due to high transcriptional variability within the β-cells. When we looked more closely at the single-cell expression levels of both mouse insulin encoding genes, *Ins1* and *Ins2*, we saw a shift in the distribution of insulin RNA expression: Chow fed mice contained more β-cells with very high insulin expression than HC fed mice **(Figure 3.5A, B)**. Comparing gene expression between Ins^{high} and Ins^{low} cells **(Figure 3.5C)** we found that Ins^{high} cells were enriched for networks of genes related to insulin processing and secretory pathways, protein processing in the ER, UPR genes, and genes related to protein

retrotranslocation **(Figure 3.5D)** all indicative of greater ER stress. Therefore, HC reduced the frequency of β-cells with very high insulin expression thereby reducing the frequency of β-cells undergoing ER stress.

Overall we think that HC reduces stress induced β-cell intrinsic damage by reducing insulin secretion which together reduce the activation of the adaptive immune system and damage caused by autoimmune infiltration and attack. HC's initial reduction in β-cell intrinsic damage therefore reduces the overall damage done to the islets preventing or delaying the development of hyperglycemia **(Figure 3.6)**.

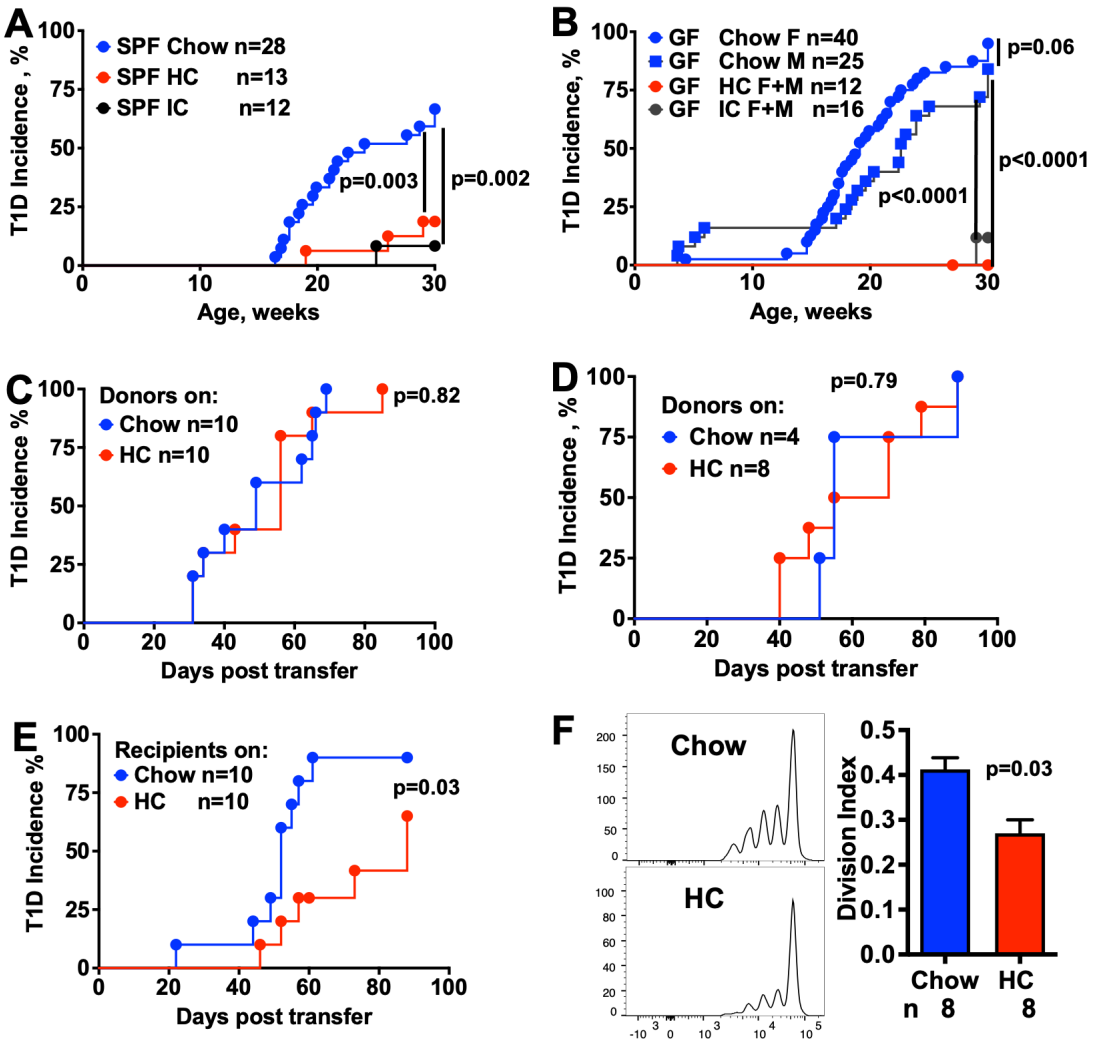

Figure 3.1: Protection from T1D by HC diet is independent from the microbiota. **(A)** Type 1 diabetes (T1D) incidence in SPF female NOD mice fed regular chow, hydrolyzed casein (HC), and intact casein (IC) diets. **(B)** T1D in germ-free (GF) NOD mice on the same diets. **(C, D)** Transfer of 2 x 10^7 splenocytes from pre-diabetic SPF NOD mice fed either chow or HC into NOD.SCID **(C)** or NOD.*Tcr*αKO **(D)** recipients fed regular chow (data combined from 2 experiments). p values were estimated using Mantel-Cox long-rank test. **(E)** Transfer of 2 x 10^7 splenocytes from SPF NOD mouse fed regular chow into NOD.*Tcr*αKO mice either on chow or HC diets (data combined from 2 experiments). p values in A-E were calculated using Mantel-Cox long-rank test. **(F)** Proliferation of CFSE-labeled BDC2.5 T cells 3 days after transfer in the pancreatic lymph nodes of recipient SPF NOD mice fed chow or HC diets (8 mice per group, combined from two independent experiments, mean±sem. p value calculated using Student's t test.

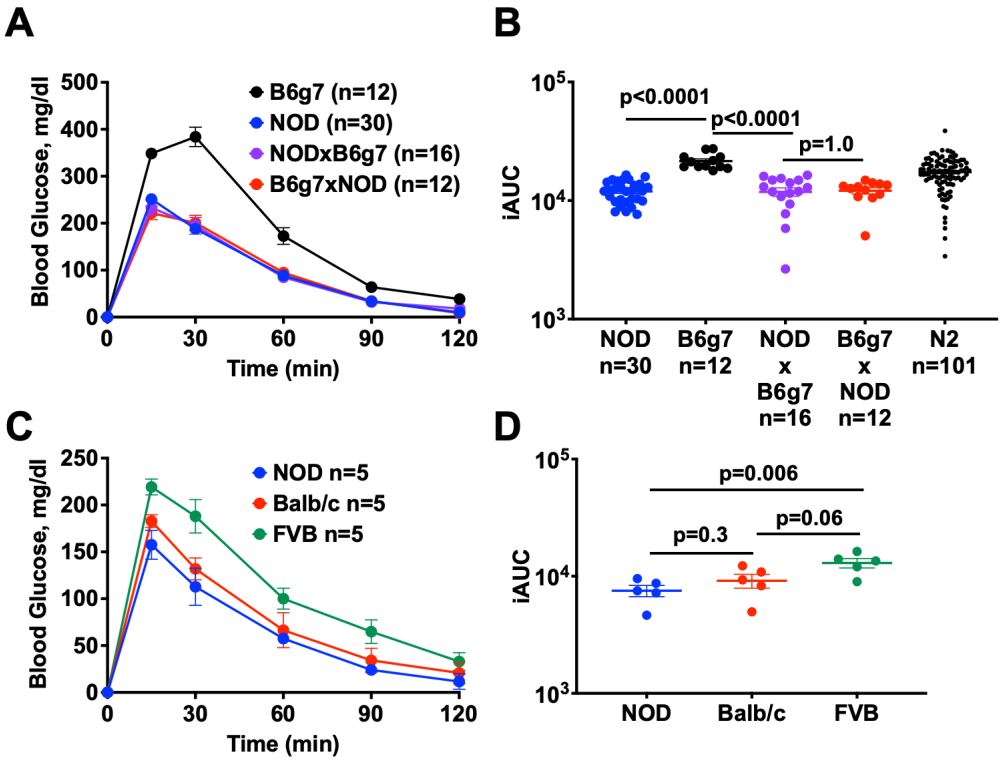

Figure 3.2: Genetic control of high glucose tolerance in NOD mice. (A) Intraperitoneal glucose tolerance test (IPGTT) results between 7–10-week-old female NOD.SCID, B6g7.SCID, and F1 hybrids crossed from both directions (NOD x B6g7 and B6g7 x NOD all carrying the SCID mutation). Data are combined from 7 independent experiments. Mean±sem. **(B)** Data from (A) shown as incremental area under the curve (iAUC). IPGTT results from N2 hybrids [F1(NOD.SCID x B6g7.SCID) x B6g7.SCID] combined from 8 independent experiments are included. Mean±sem. p values calculated using one-way ANOVA with post-hoc Tukey test. **(C)** IPGTT results between NOD, BALB/c, and FVB mice (all WT). Data are from a single experiment. Mean±sem. **(D)** Data from (C) shown as iAUC. Mean±sem. p values calculated using Student's t test.

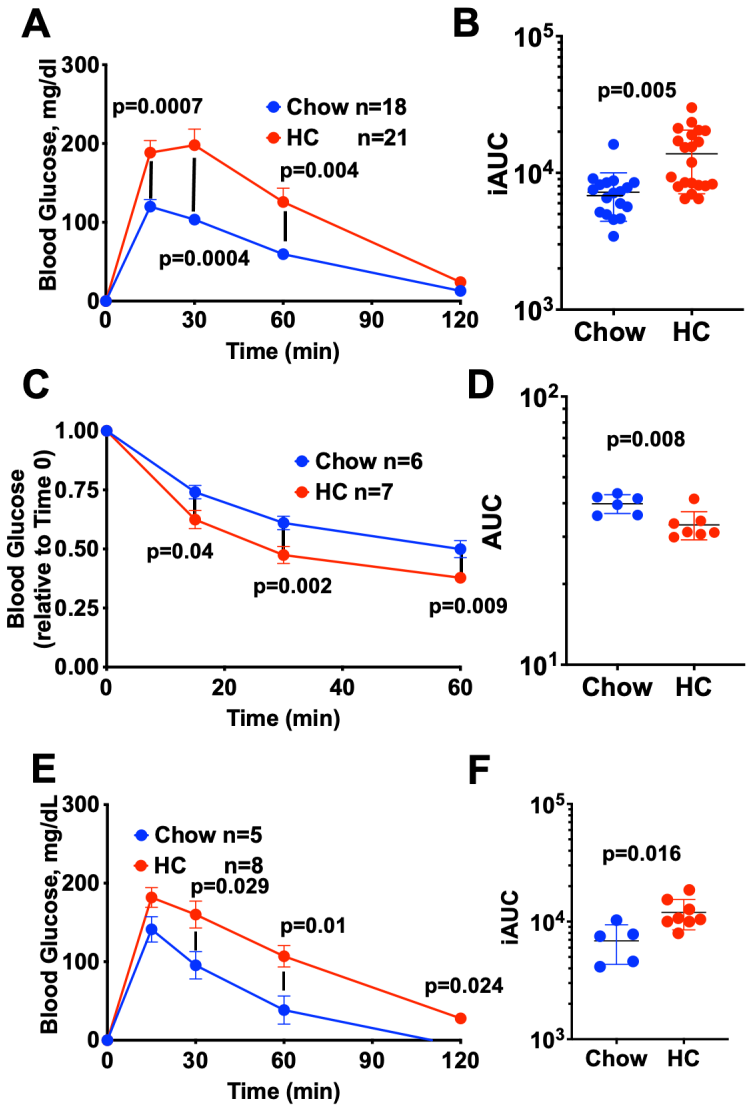

Figure 3.3: HC diet reduces high glucose tolerance in NOD mice. **(A)** IPGTT results between 6–8-week-old female NOD mice fed either chow or HC diets. Data are combined from 5 independent experiments. Mean±sem. **(B)** Data from (A) shown as iAUC. Mean±sem. p values calculated using Student's t test. **(C)** Clearance of blood glucose following intraperitoneal injection of insulin in 6–8-week-old female NOD mice fed either chow or HC diets. Mice were fasted for 6 hours prior to insulin injection. Data are combined from 2 independent experiments. Mean±sem. **(D)** Data from (C) shown as AUC. Mean±sem. p values calculated using Student's t test. **(E)** IPGTT results between 6–8-week-old female NOD.*Tcra*KO mice fed either chow or HC diets. Data are combined from 2 independent experiments. Mean±sem. **(F)** Data from (E) shown as iAUC. Mean±sem. p values calculated using Student's t test.

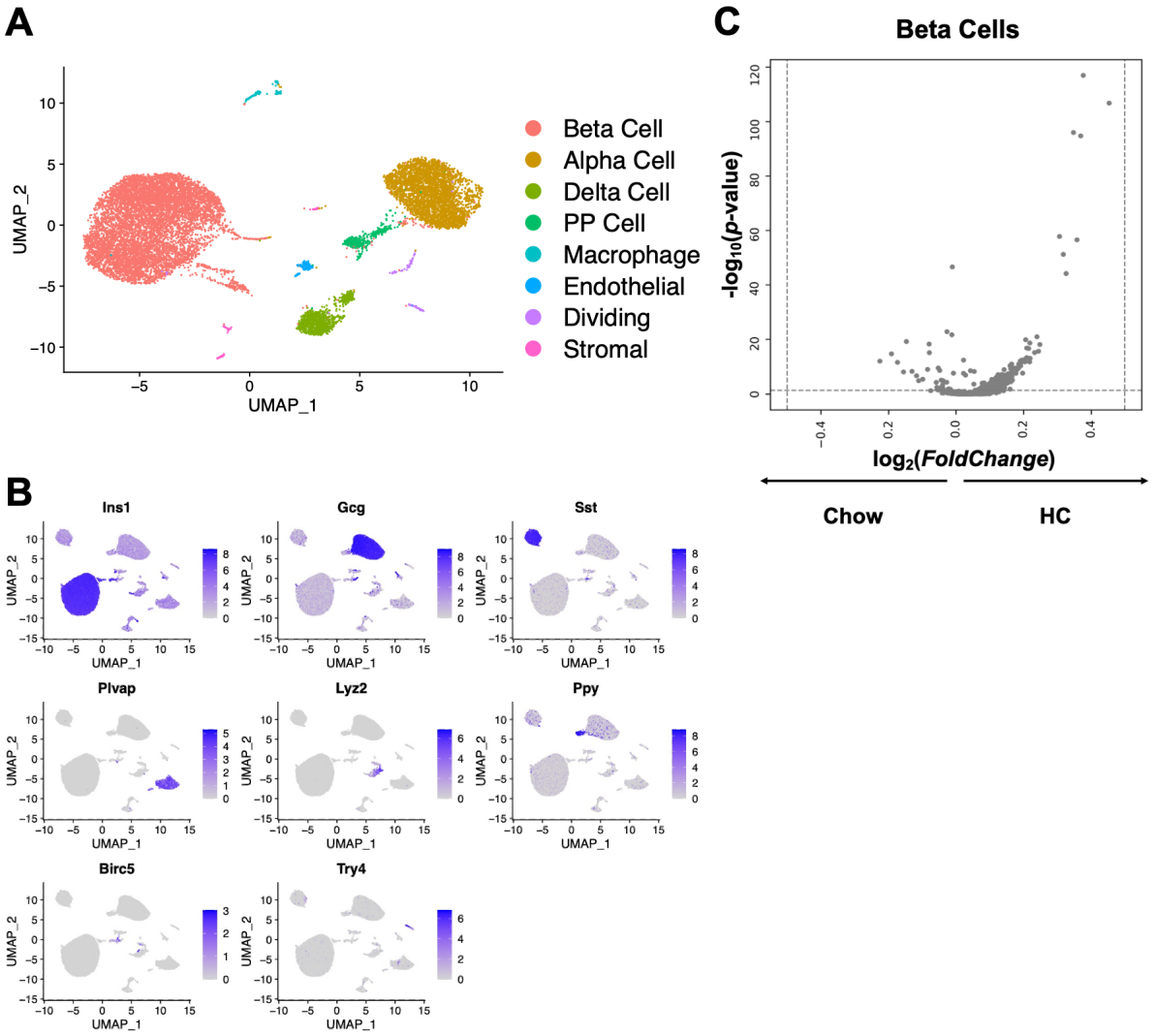

Figure 3.4: Cell populations of the non-inflamed pancreatic islets. **(A)** Uniform Manifold Approximation and Projection (UMAP) dimension reduction plot of islet cells from 8-10-week-old NOD.SCID mice fed chow of HC diets (3 mice per diet). Cell clusters based on similarities in transcriptional profiles obtained through single-cell RNA-Sequencing analysis. **(B)** Expression of canonical endocrine, endothelial, myeloid, and acinar cell genes in clusters of cells. **(C)** Volcano plot showing differential gene expression between β-cells from islets of chow and HC fed NOD.SCID mice. No genes were differentially expressed.

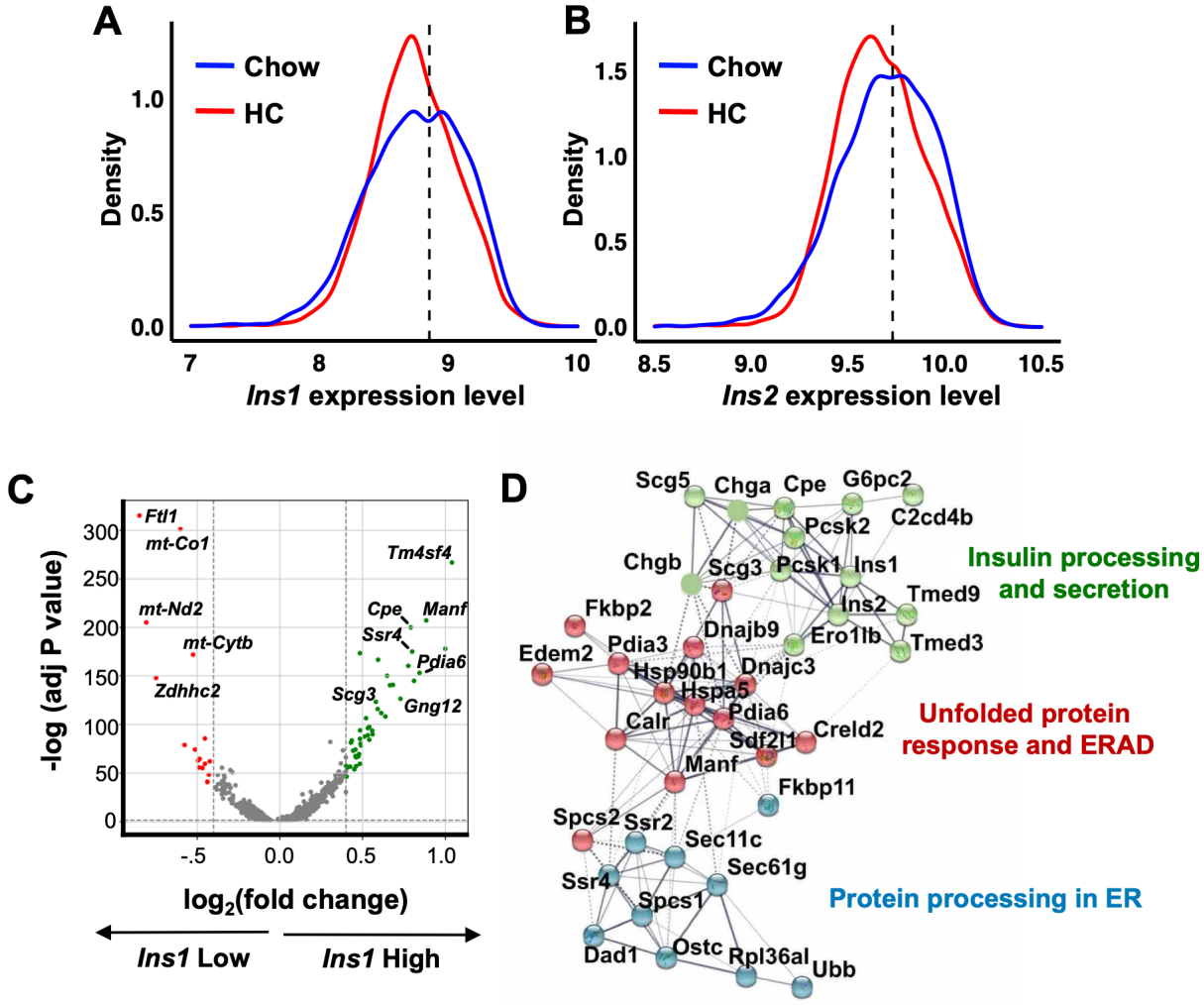

Figure 3.5: HC diet reduces insulin RNA expression and β-cell stress. (A-B) Distribution of relative *Ins1* and *Ins2* gene expression levels in β-cells from 8–10-week-old NOD.SCID mice fed regular chow (blue line) or HC diet (red line). Data is based on single-cell RNA sequencing of NOD.SCID islets and is plotted as a kernel density estimation function. Dashed vertical line indicates the border between low and high expressors of insulin in islets of chow fed mice. **(C)** Volcano plot showing differential gene expression between *Ins1*-High and *Ins1*-Low beta cells isolated from the islets of NOD.SCID mice fed regular chow. Genes indicated have adjusted p-values less than 0.05 and log2 fold change magnitude greater than 0.5. **(D)** Gene networks built using genes upregulated in β-cells with high insulin expression compared to β-cells with low insulin expression using STRING database [168]

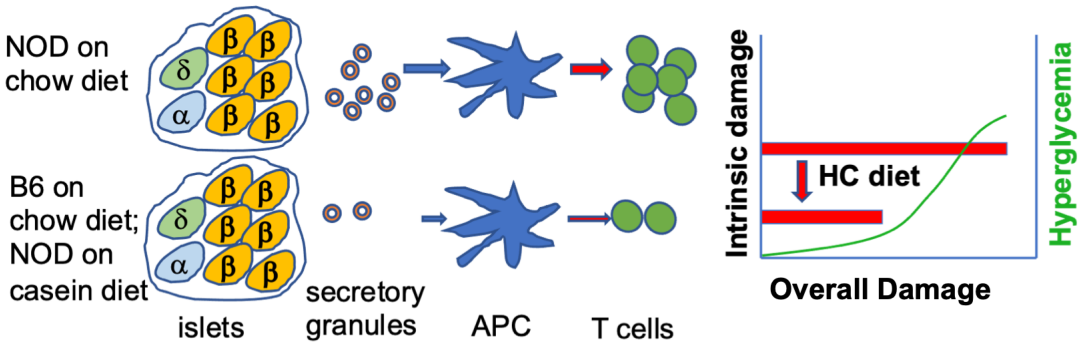

Figure 3.6: A model explaining how HC diet attenuates T1D pathogenesis. HC diet reduces genetically determined high insulin secretion by β-cells thereby reducing β-cell damage caused by high endoplasmic reticulum stress. This reduction in β-cell intrinsic damage reduces activation of the adaptive immune system compounding to reduce overall damage which is caused by the combination of intrinsic cellular stress and autoimmune attack. In the case of HC, overall damage is reduced enough to prevent or delay hyperglycemia.

3.4 Discussion

In this chapter we identified that a diet containing HC as the sole amino acid source protected NOD mice from T1D independent from the microbiota. We showed that HC does not impact the immune system's ability to generate islet-reactive effector cells but rather renders the islets or their environment less stimulatory to the immune system. Examination of β-cell physiology led us to discover that the genetically determined propensity for high insulin secretion in NOD mice compared to B6 mice is a dominant, non-mitochondria linked trait that may be controlled by a single gene. HC diet reduces this high insulin secretion by NOD β-cells which in turn reduces the frequency of β-cells that are under high levels of ER stress.

Several groups [193, 194] including our own [195] have suggested that β-cell health may influence T1D progression. In NOD mice, ER stress was of particular interest because their genetic predisposition for high insulin secretion places a high metabolic burden on their islets. ER stress can promote apoptosis [196], increasing the amount of β-cell antigen available to the immune system. This could create a positive feedback loop where β-cell apoptosis promotes immune infiltration thereby exposing live β-cells to pro-inflammatory cytokines that promote more ER stress [197] and β-cell upregulation of MHC class I [198] further exposing them to the immune system. This suggests that even slight reductions in β-cell stress could greatly reduce overall β-cell death and prevent the development of T1D by slowing down β-cell exposure to the immune system. Our data fit this hypothesis well: HC reduced the genetically controlled high insulin secretion in NOD mice thus reducing the frequency of very stressed β-cells **(Figure 3.5A-D)**. This reduction in insulin secretion was independent from damage done by the immune system suggesting that diet could directly improve the health of the target organ. The reduction in cellular stress may have also reduced β-cell exposure to the immune cells, as HC fed mice also had a reduced capacity to activate autoreactive T cells **(Figure 3.1E, F)**. Whether β-cell stress and high insulin secretion directly impact antigen presentation, upregulation of costimulatory receptors on

professional APCs, or other aspects of the autoimmune process remain open questions and are avenues for future study.

It is also possible that reduction of insulin secretion by HC reduces the activation of autoimmunity *directly* rather than *via* the reduction of β-cell stress. A recent theoretical model [199] posited that hormone hypersecretion may be linked to autoimmunity. Certain endocrine organs both proliferate and secrete hormone in response to the targets that they regulate, for example β-cells and glucose. This leaves a vulnerability as any mutant endocrine cells that are hypersensitive to their targets would tend to both secrete too much of their hormone and proliferate creating a mass of hypersecreting cells. The immune system, under this model, controls this process by removing hypersecreting mutants. However, this also leaves a vulnerability as an overaggressive immune response might kill too much of the target organ leading to autoimmunity. Our data supports this hypothesis: a reduction of insulin secretion by HC corresponded with a reduction in the activation of autoimmunity.

Regardless of the precise mechanism behind HC protection, control of β-cell stress, insulin hypersecretion, or both, we have shown that diet can apply healthy physiologic changes to an endocrine organ that protect genetically at-risk animals from autoimmune disease. This protection occurs independent from an individual's microbiota composition meaning prophylactic use of HC in genetically at-risk individuals may have a reasonable chance of success. However, diets are rarely composed of a single protein source and many different common dietary proteins have been shown to be protective or exacerbatory to T1D, even in HC-based diets [77, 112]. Understanding how these other proteins influence disease progression is critically important. Protective proteins insensitive to microbial influence besides casein may provide new avenues for the development of dietary interventions for the treatment of autoimmune disease. Alternatively, diabetes accelerating proteins may counteract any benefit from the protective proteins and therefore must be studied in depth so that their mechanism of disease promotion can be understood.

CHAPTER 4

GLUTEN REVERSES PROTECTION OFFERED BY CASEIN BY STIMULATING THE IMMUNE SYSTEM.

4.1 Preface

Diabetes incidence in NOD mice fed different diets was monitored by L.A. Yurkovetskiy and A.V. Chervonsky. Glucose tolerance tests were carried out by myself. Direct measurement of serum insulin was started by L.A. Yurkovetskiy and C.H.F. Hansen and completed by myself. *In vivo* proliferation assays were carried out by L.A. Yurkovetskiy. Insulitis from histology sections was scored by A. Kuznetsov. Single-cell RNA-sequencing data collection and cell clustering analysis, TCR-sequencing data collection, flow cytometry on islet infiltrates, and ELISPOT experiments were carried out by myself. Lastly, identification of differentially expressed genes from single-cell RNA-sequencing data and TCR sequencing analysis was performed in collaboration with D. Reiman.

4.2 Introduction

Despite promising results in NOD mice and BB rats (chapter 3 and [99]), HC diets have not proven to be effective at protecting genetically at-risk humans from T1D. In the trial to reduce insulin dependent diabetes mellitus in the genetically at risk (TRIGR), 2159 newborns with high-risk HLA genotypes and one first degree relative with T1D were randomly placed on an HC supplemented formula or a conventional non-hydrolyzed cow's milk formula for use during infancy and monitored for the development of islet antibodies and T1D for a median of 12 years [200]. While early data found that the infants placed on HC supplemented formula had lower rates of islet antibody positivity [201], at the end, overall T1D incidence was not different between the two groups [113]. There are multiple explanations for HC's lack of efficacy in these studies. Intact cow's milk has been shown to be as protective as HC in some

studies [97, 112] (see also **Figure 3.1A,B**) and pro-diabetic in others [98, 111]. Therefore, cow's milk formula may have been a poor control diet. The study also suffered from inevitable inconsistencies that are likely to plague any human study. Both the age of introduction and duration of feeding varied considerably between different infants. Additionally, other dietary products introduced either during or following the formula feeding period could have influenced cumulative T1D development.

One dietary product of interest to us was gluten. Gluten is a catch all term for storage proteins in wheat and is composed of aggregates of gliadin and glutenin proteins which make up 30 and 50% of total wheat grain protein respectively [202]. As wheat makes up about 20% of globally consumed calories [203], gluten should be considered an enormous contributor to human diets. In the case of T1D, gluten has been suspected to influence T1D development in humans [92] and a gluten-free diet was anecdotally shown to reverse T1D in a recently diagnosed infant [204]. Some experimental evidence from NOD mice and BB rats also suggests that gluten may promote T1D [85, 105, 106], though there is conflicting evidence in the literature [97, 98]. Gluten also drives the pathogenesis of celiac disease which is observed in a large fraction of patients with T1D and shares many genetic risk factors with T1D [70, 205].

Here we found that the addition of gluten to HC diet restored high incidence of T1D to NOD mice. The restoration was not due to increases in insulin secretion but was due to the activation of the immune system. Rather than increasing the quantity of islet infiltration, gluten boosted the inflammatory states of multiple cell types including myeloid cells, $\alpha\beta$T cells, and Tregs. Gluten also promoted the expansion of highly diabetogenic IGRP-reactive T cells in both the PLNs and the islets. Overall, this work shows that enhanced inflammation can overcome beneficial changes to target organ physiology to promote T1D.

4.3 Results

4.3.1 Gluten reverses protection by HC diet

To test whether gluten could reverse the protection afforded by HC, we replaced 1/5th of the casein in the HC diet with gluten from two different manufacturers resulting in a 16%HC+4%gluten diet, hereafter referred to as HC+4%gluten. The addition of gluten from either manufacturer to the HC diet restored diabetes incidence to similar levels to that observed in chow fed mice **(Figure 4.1)**.

As HC protected NOD mice from T1D by reducing β-cell stress *via* the reduction of high insulin secretion **(Figure 3.3 and 3.5)**, we performed IPGTT tests to determine whether gluten restored high insulin secretion to the HC fed mice. NOD.SCID mice fed HC+4%gluten remained as glucose intolerant as mice fed HC whereas NOD.SCID mice fed chow had higher glucose tolerance **(Figure 4.2A,B)**. Direct measurements of plasma insulin five minutes following an injection of glucose confirmed that the reduced glucose tolerance observed in HC and HC+4%gluten mice was due to reduced insulin secretion **(Figure 4.2C)**. Interestingly, transferred BDC2.5 T cells proliferated equally well in the PLNs of chow and HC+4%gluten fed mice **(Figure 4.3)** whereas they had proliferated more poorly in PLNs from HC fed mice **(Figure 3.1F)**. Therefore, gluten likely did not restore T1D by restoring high levels of insulin secretion but promoted the activation of autoreactive cells by another mechanism.

4.3.2 Activation of islet immune infiltrates by gluten

To test whether gluten promoted immune activation, we performed scRNA-seq and spectral flow cytometry to characterize the islet infiltrates from HC and HC+4%gluten fed mice. scRNA-seq revealed extensive immune infiltration in islets from both HC and HC+4%gluten fed, 15-week old NOD mice composed of myeloid cells, T cells, and B cells as well as the α-, β-, and δ-endocrine cells expected from the pancreatic islets **(Figure 4.4A, B)**. Gluten

in the diet had a substantial effect on the β-cells. β-cells typically make up 60-80% of the endocrine cells in the islets with α-cells making up 15-20%, δ-cells making up <10%, and other endocrine cells like pancreatic polypeptide-secreting cells and ϵ-cells making up the remainder [206]. Mice fed HC+4%gluten had a reduction in the expected proportion of β-cells to other endocrine cells whereas HC fed mice did not **(Figure 4.5A)**. The remaining β-cells in HC+4%gluten fed mice had upregulated inflammatory response genes, many of them downstream of NFκB including *Atf3, Nfkbiz* and members of the AP-1 transcription factor *Jund, Fosb, Fos, Jun, and Junb* **(Figure 4.5B)**. The cumulative gene signature of β-cells from HC+4%gluten fed mice compared to HC fed mice was of TNFα receptor signaling as revealed by gene set enrichment analysis (GSEA) [169, 170]. All of this pointed to evidence of elevated inflammation in the islets of mice exposed to gluten.

Elevated islet inflammation could either be explained by more islet infiltration or by pro-inflammatory phenotypic differences in the infiltrates themselves in mice fed gluten. Using flow cytometry, we found no evidence for elevated immune infiltration: the frequency (% of total live cells) and absolute numbers of islet CD45$^+$ cells recovered from HC+4%gluten fed mice was not greater than that observed in HC fed mice **(Figure 4.6A,B)**; if anything it was lower, though enormous variability in frequencies within groups was also observed. Frequencies of major leukocyte populations within the islet infiltrating CD45$^+$ compartment also remained unchanged when gluten was present in the diet **(Figure 4.6C,D**. Finally, insulitis scored by histological methods also revealed few differences in the infiltration burdens between HC and HC+4%gluten fed mice **(Figure 4.7)**. Whereas, more scored HC fed mice had a high percentage of non-infiltrated islets, the overall frequencies of infiltrated islets between the two groups was similar. Therefore, elevated levels of insulitis could not account for the increased response to inflammation observed in islets from HC+4%gluten fed mice.

While the numbers and frequencies of infiltrating cells did not appear to be different, gluten did influence the inflammatory phenotypes of the infiltrating cells. Differentially expressed gene analysis of the myeloid compartment **(Figure 4.8A)**identified in the scRNA-

seq dataset revealed elevated expression of the cytokines *Tnf and Cxcl2* as well as other genes downstream of inflammatory stimuli in HC+4%gluten fed mice. In contrast, the islet myeloid cells from HC fed mice had negative immune regulators *Tnfaip8l2* and *Slfn2* upregulated. Using flow cytometry, we also found that DC's from the islets had elevated expression of the inducible costimulatory receptor CD86 **(Figure 4.8B,C)**. T cells were also affected. T cells recovered from the islets from HC+4%gluten fed mice were more proliferative (based on Ki67 staining) than T cells recovered from HC fed mice **(Figure 4.8D,E)**, and Tregs (TCRβ^+ CD4$^+$ FoxP3$^+$) from the same group expressed higher levels of genes associated with tissue residency **(Figure 4.9A)** such as the epithelial binding integrin CD103 [207, 208] **(Figure 4.9B,C)**. The elevation of tissue residency markers on Tregs has been previously used as a sign of elevated inflammation in that organ [209].

Cumulatively, these data suggested that gluten promoted inflammation in the islets by upregulating inflammatory gene expression programs on the infiltrating leukocytes rather than by increasing the numbers of infiltrating leukocytes themselves.

4.3.3 Expansion of diabetogenic T cell clones in HC+4%gluten fed mice.

To test if gluten promoted the expansion of diabetogenic T cell clones, we used single cell TCR sequencing to analyze the TCR repertoires of cells collected from the islets of HC-and HC+4%gluten-fed mice. Comparison of CDR3α sequence frequency revealed that gluten reduced clonal diversity in the islet CD8$^+$ T cells **(Figure 4.10A)**. Compared to CD8$^+$ T cells from HC fed mice, there was a lower frequency of unexpanded clones (cell count per CDR3α sequence = 1) and a higher frequency of expanded clones (cell count per CDR3α sequence $>$ 31). Analysis of the CDR3α sequences themselves **(Figure 4.10B)**revealed an expansion of sequences that were identical to or very similar to the CDR3α sequence from the highly diabetogenic clone NY8.3 which reacts to islet-specific glucose-6-phosphate related subunit (IGRP) [210, 211]. Sequence similarity analysis using the BLOSUm62 scoring matrix [172] further revealed additional CDR3α sequences with high probability of conferring the

same specificity **(Figure 4.10C)**.

We also found reduced islet $CD8^+$ T cell clonal diversity in gluten-fed mice using CDR3β sequences **(Figure 4.11A)** though, unlike with CDR3α, this reduced diversity was not due to the dominance of a particular CDR3β sequence **(Figure 4.11B,C)**. TCR sequencing of $CD4^+$ T cells from islets did not reveal evidence for clonal expansion driven by gluten or the dominance of particular CDR3α or CDR3β clonotypes **(Figure 4.12 and 4.13)**.

To determine whether islet-reactive T cell clones also expanded in the PLNs of mice fed gluten, we performed an enzyme linked immunospot (ELISPOT) analysis of interferon-γ (IFN-γ) production by PLN T cells in response to three peptides recognized by known diabetogenic $CD8^+$ T cell clones [164, 212, 213]. PLNs from HC+4%gluten fed mice had more activated diabetogenic T cell precursors than PLNs from HC fed mice **(Figure4.14A and B)**. As individual NOD mice vary in their responses to different peptides [214], overall $CD8^+$ T cell reactivities are shown in **Figure4.14A**. PLNs from HC+4%gluten-fed mice also had more $CD8^+$ T cells reactive to a peptide recognized by 8.3 T cells **(Figure 4.14B)**, mirroring the expansion of these same T cells in the islets.

Overall, gluten enhanced the immune activation of multiple cell types and the expansion of diabetogenic 8.3-like T cells in the islets and pancreas draining lymph nodes. This resulted in damage to the insulin-secreting β-cells overcoming the protection offered by HC diet.

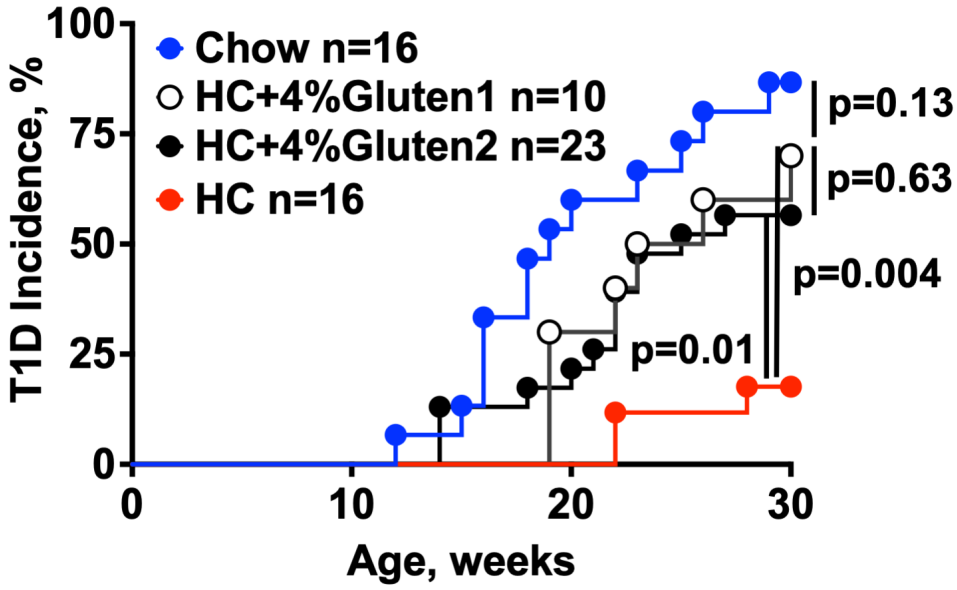

Figure 4.1: Gluten reverses protection from T1D afforded by HC diet. Gluten from two different manufacturers was added to HC diet replacing 1/5th of hydrolyzed casein weight (or 4% of the total diet weight, thus termed 'HC-4%gluten diet') and T1D incidence in female NOD mice was observed in the same experiment. p values were calculated using Mantel-Cox long-rank test.

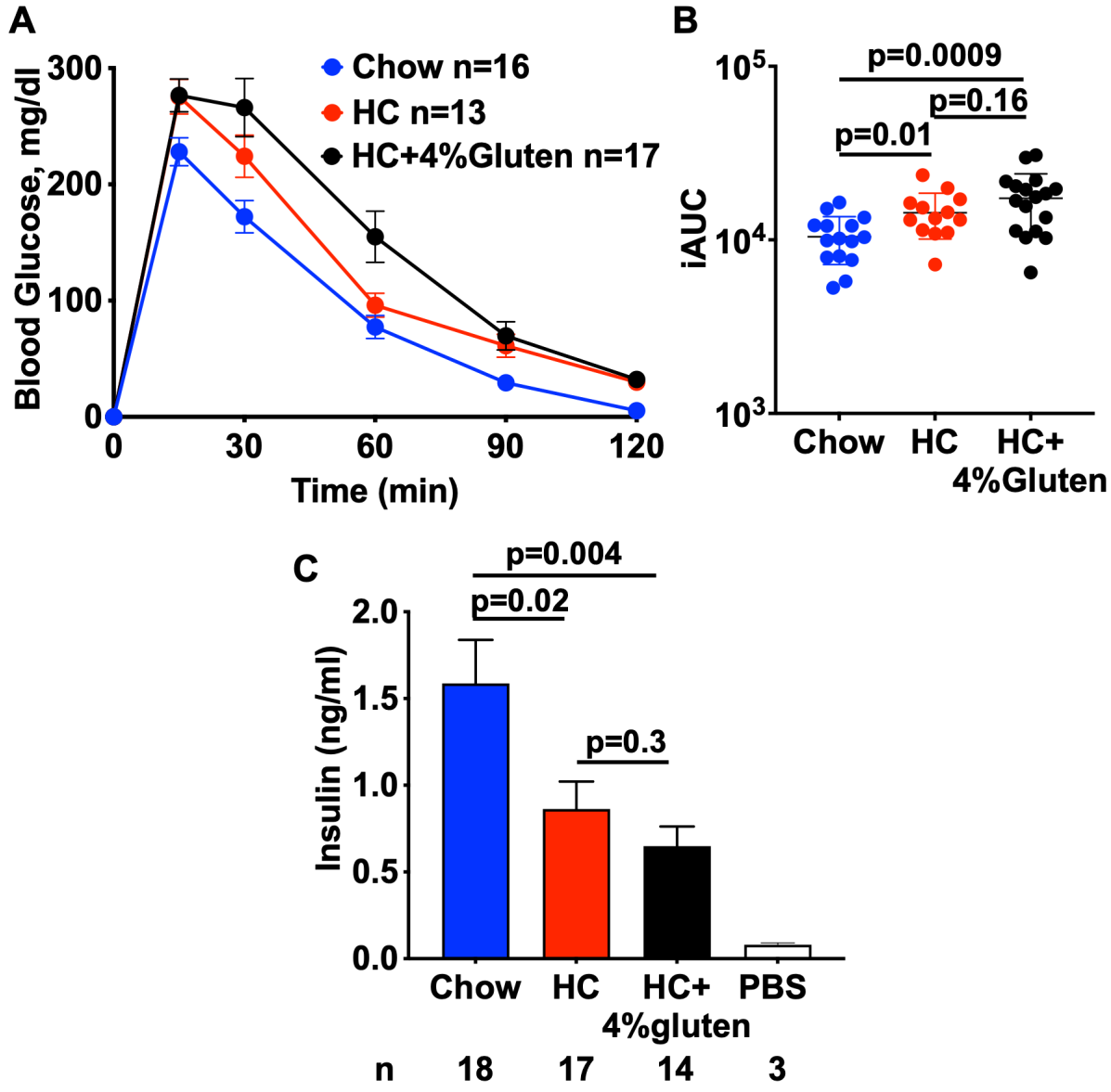

Figure 4.2: Gluten does not restore high glucose tolerance and insulin secretion to HC fed mice. **(A)** IPGTT results comparing NOD.SCID mice fed either chow, HC or HC+4%gluten diets. Data are combined from 5 independent experiments. Mean±sem. **(B)** Data from (A) shown as iAUC. Mean±sem. p values were calculated using Student's t test. **(C)** Direct measurements of plasma insulin 5 minutes after glucose challenge. Data are combined from 4 independent experiments. PBS control – base line insulin in animals injected PBS instead of glucose. Mean±sem. p values were calculated using Student's t test.

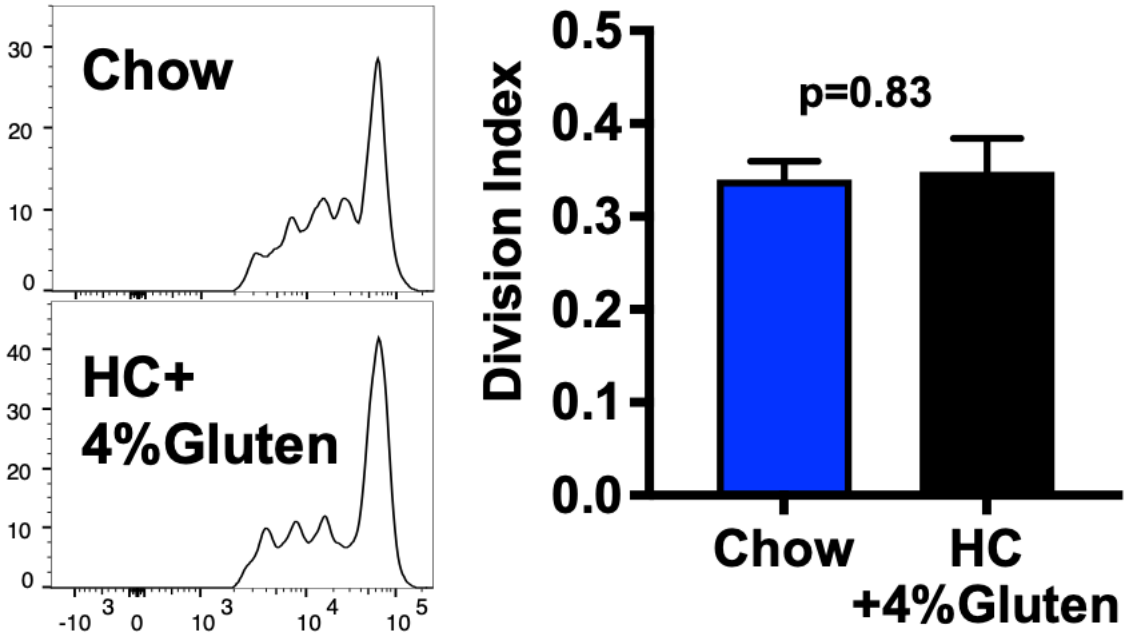

Figure 4.3: Islet reactive T cells proliferate equally well in chow and HC+4%gluten fed mice. Proliferation of CFSE-labeled BDC2.5 T cells in the pancreatic lymph nodes of mice on chow and HC+4%gluten diets, mean±sem. Combined data from 2 experiments. p value was calculated using Student's t test.

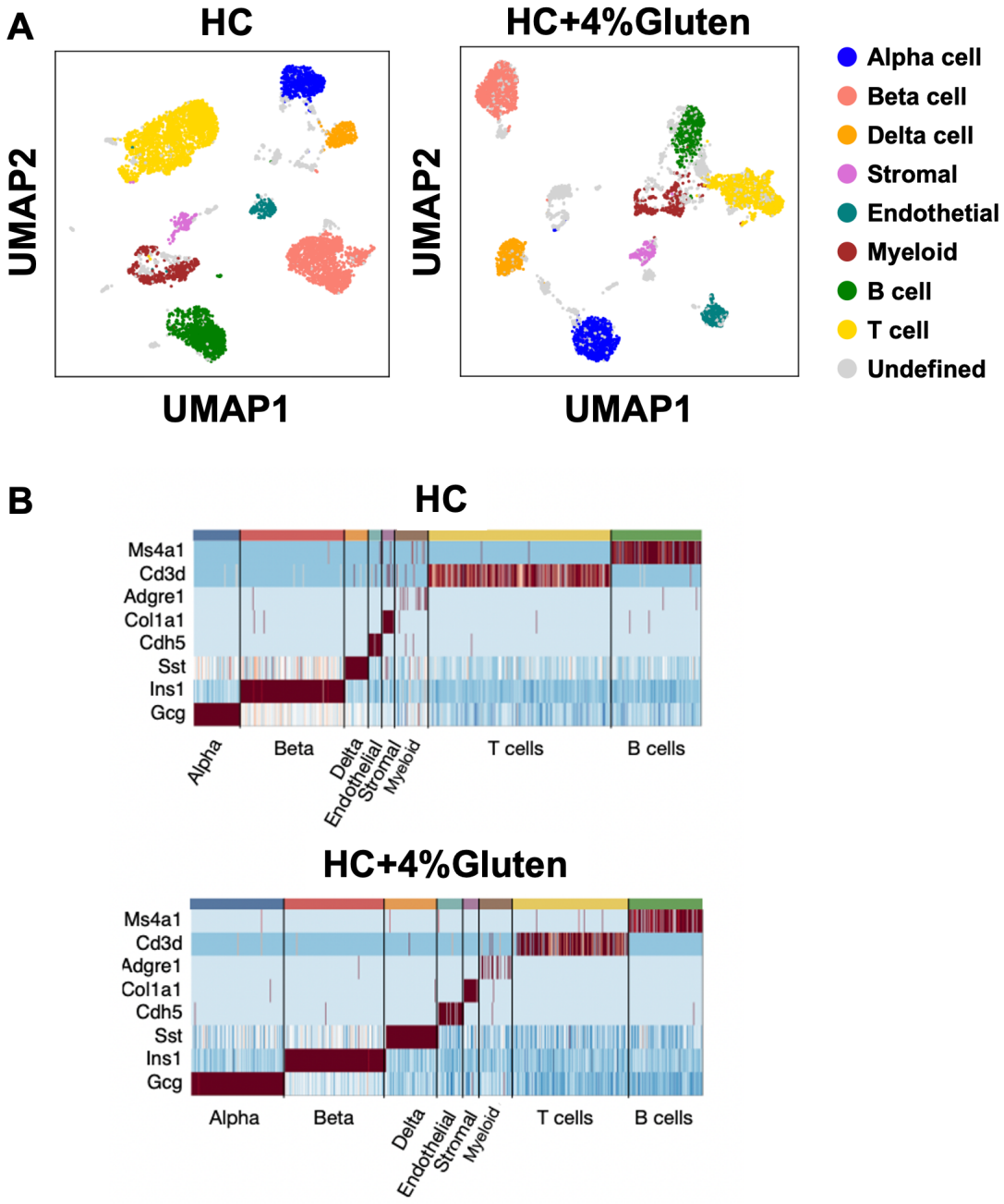

Figure 4.4: Cell populations of the inflamed islets of HC and HC+4%gluten fed mice. (A) Uniform Manifold Approximation and Projection (UMAP) dimension reduction plots of islet cells from 15-week-old NOD mice on indicated diets (SCS analysis, pool of 3 mice per group). **(B)** Heatmap showing selected markers in islet cell populations for UMAP images in (A). Each gene marker is z-score normalized across all cells.

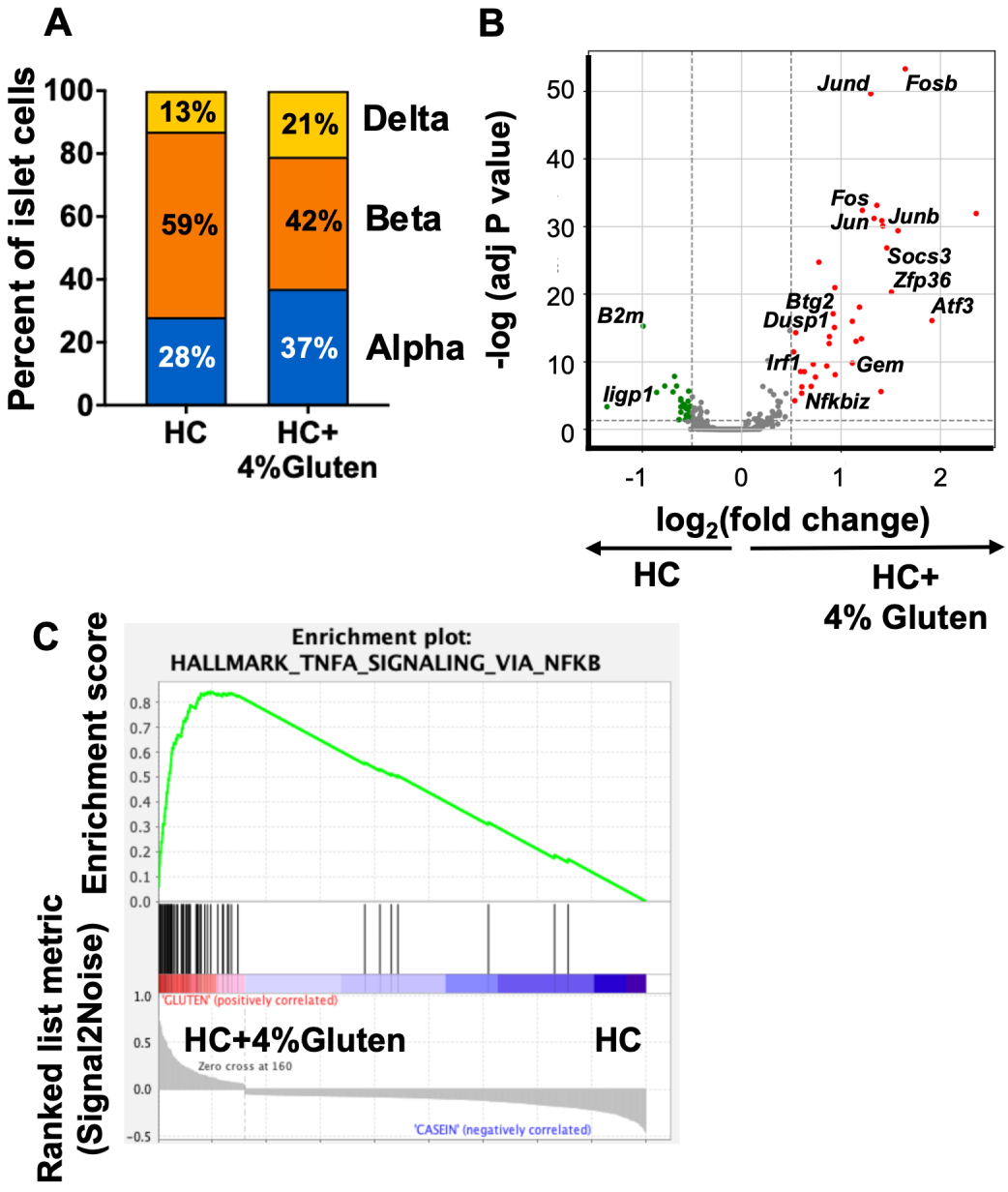

Figure 4.5: Loss and inflammation of β cells in HC+4%gluten fed mice. **(A)** Ratios of α-, β-, and δ-cells in the islets of HC- and HC + 4%gluten-fed NOD mice defined by SCS. **(B)** Volcano plot showing differential gene expression between beta cells from mice on HC vs. HC+4%gluten diets. Indicated genes have adjusted p-values less than 0.05 and log2 fold change magnitude greater than 0.5. **(C)** Gene Set Enrichment Analysis (GSEA) showing enrichment of hallmark TNFα signaling via NFκB in β-cells from HC+4%gluten fed NOD islets. FDR q value < 0.05.

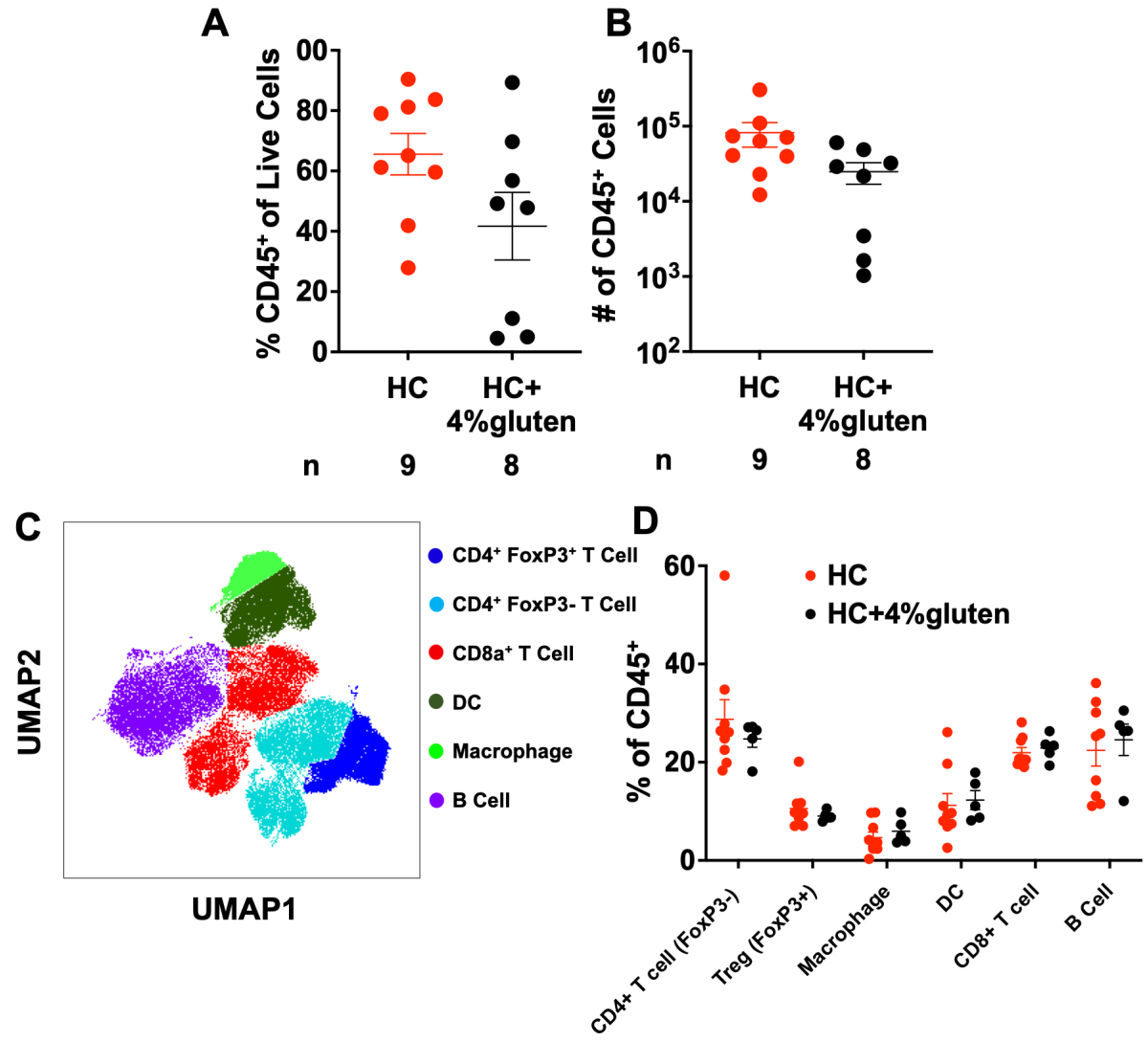

Figure 4.6: Gluten does not alter leukocyte population frequencies in the pancreatic islets. **(A)** Frequency of leukocytes ($CD45^+$) among all live cells in the islets of HC and HC+%gluten fed NOD mice. Mean±sem. **(B)** Raw counts of leukocytes in the islets of HC and HC+4%gluten fed NOD mice. Mean±sem. **(C)** Flow cytometry UMAP projection of live, $CD45^+$ cells from NOD islets. Data collected from islets from HC fed (n=9) and HC+4%gluten fed (n=5) NOD mice pooled together for UMAP projection. It is built on key markers: TCRβ, CD4, CD8α, FR4, FoxP3, CD44, CD62L, PD-1, CD103, CD69, CTLA4, Ki67, CD19, CD11c, CD11b, F4/80, PD-L1, CD80, and CD86. Colors indicate manually defined cell clusters. **(D)** Frequencies of clusters in (C) out of total $CD45^+$ cells. Mean±sem.

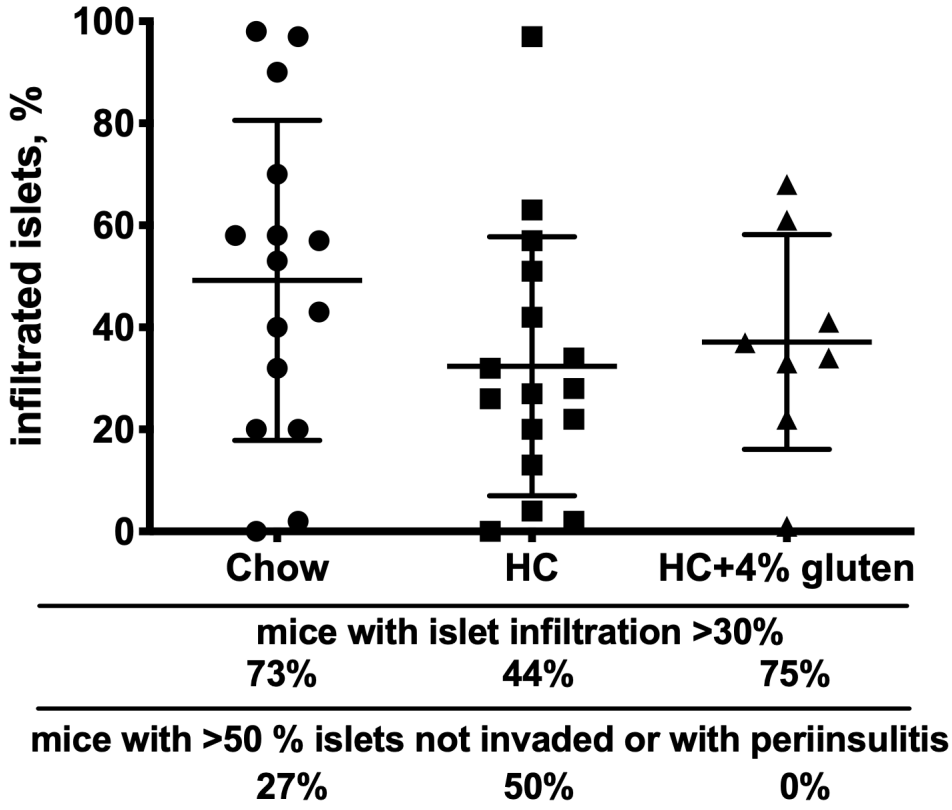

Figure 4.7: Gluten does not strongly impact insulitis. Frequency of islets within each diet group with histological evidence of islet infiltration (histology score of II or III, see Materials and Methods). Each point is an individual mouse, mean±SD. All mice were 13 weeks old and only females were used in this experiment.

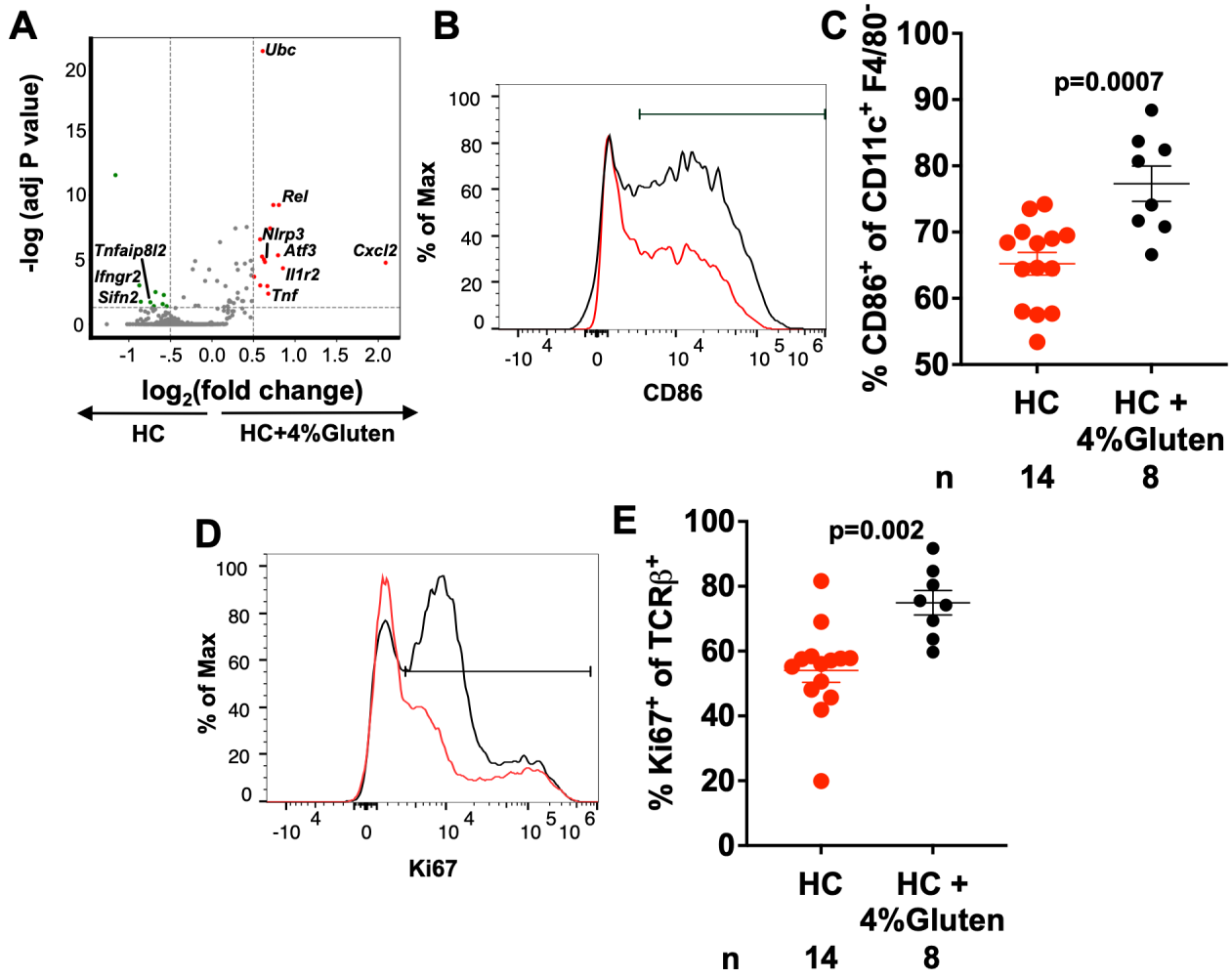

Figure 4.8: Activation of islet myeloid and T cells in HC+4%gluten fed mice. **(A)** Volcano plot showing differential gene expression between myeloid cells from islets of HC and HC+4%gluten fed NOD mice. Genes indicated have adjusted p values less than 0.05 and log2 fold change magnitude greater than 0.5. **(B,C)** Representative histogram of CD86 expression by dendritic cells ($CD45^+$ $CD11c^+$ $F4/80^-$) in the islets of HC and HC+4%gluten fed NOD mice and quantification of their expression. Data combined from 6 experiments. Mean±sem. p value was calculated using Student's t test. **(D,E)**) Representative histogram and quantification of Ki67 expression in $\alpha\beta$T cells ($CD45^+$, $CD11c^-$, $F4/80^-$, $CD19^-$, $TCR\beta^+$) from the islets of HC and HC+4%gluten fed NOD mice. Data combined from 6 experiments. Mean±sem. p value was calculated using Student's t test.

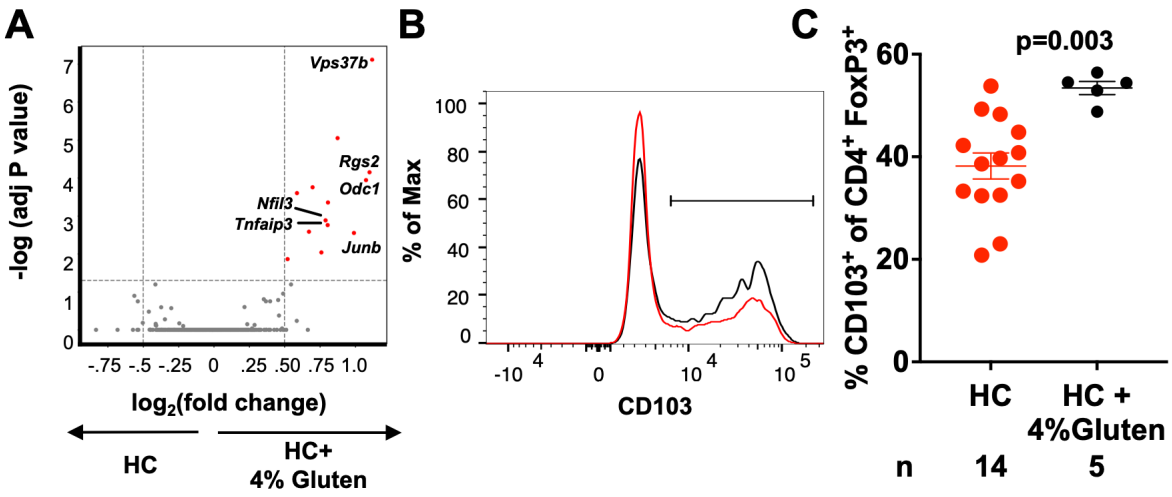

Figure 4.9: Upregulation of tissue residency markers on islet Tregs from HC+4%gluten fed mice. **(A)** Volcano plot showing differential gene expression between Treg ($CD4^+FoxP3^+$) cells from the islets of HC and HC+4%gluten fed NOD mice. Genes indicated have adjusted p-values less than 0.05 and log2 fold change magnitude greater than 0.5. **(B,C)** Representative histogram and quantification of CD103 expression in Tregs ($CD45^+$, $CD11c^-$, $F4/80^-$, $CD19^-$, $TCR\beta^+$, $CD4^+$, $FoxP3^+$) from the islets of HC and HC+4%gluten fed NOD mice. Data combined from 6 experiments. Mean±sem. p value was calculated using Student's t test.

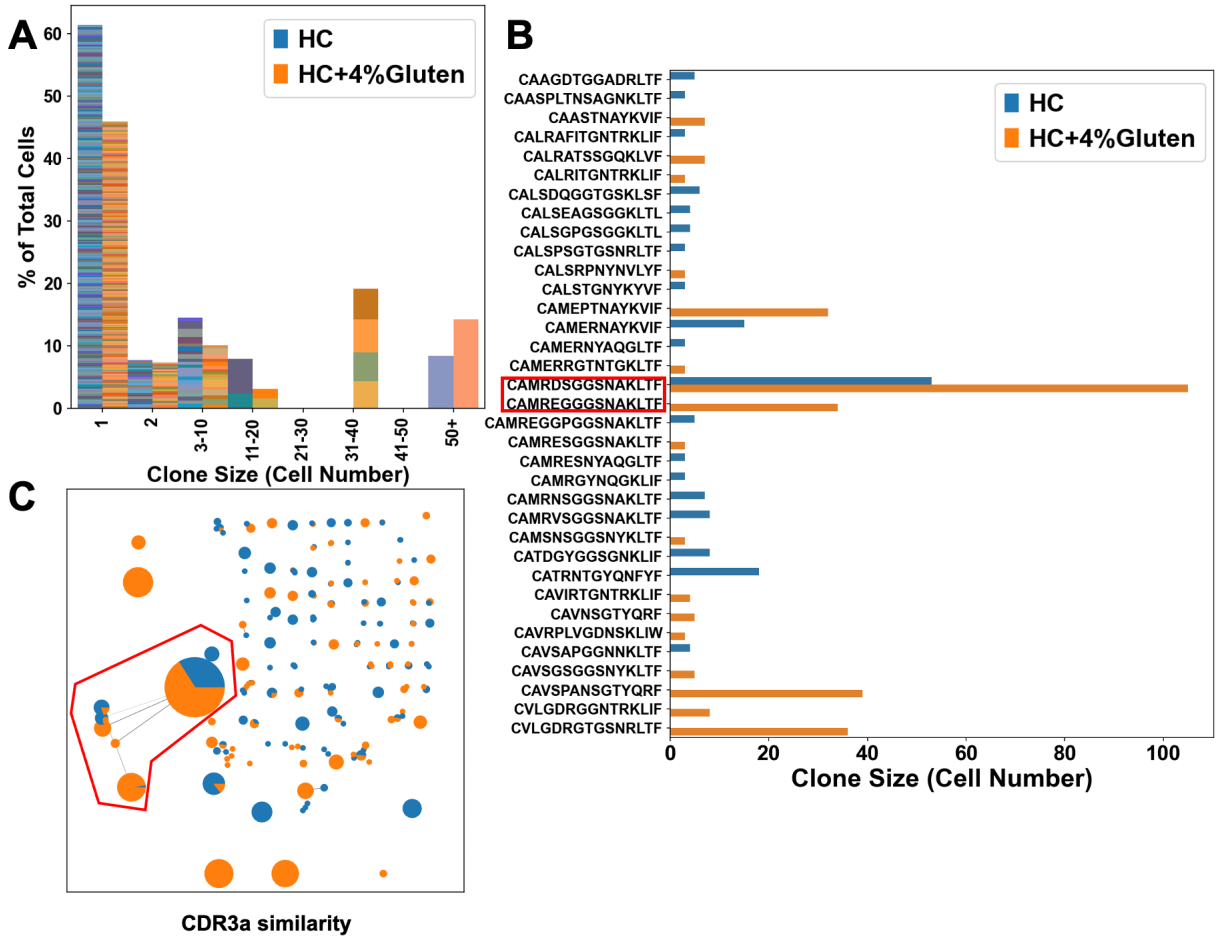

Figure 4.10: Analysis of CDR3α sequences from islet CD8$^+$ T cells. **(A)** Distribution of expanded CDR3α sequences and clonal sizes in islet infiltrates from mice on HC (blue background) and HC+4%Gluten (orange background) diets from single cell sequenced CD8$^+$ cells (three mice per diet). The abscissa represents the size of expanded clonotypes and the ordinate represents the % of individual clonotype of total sequenced CD8$^+$ cells. Individual CDR3α sequences are represented by different colors. **(B)** The number of expanded cells from each clonotype ($\geq$3 cells with identical TCRs) derived from expanded CDR3α sequences from CD8$^+$ cells in the islets of HC (blue) and HC+4%Gluten (orange) fed mice. CDR3α identical and visibly similar to NY8.3 are highlighted with a red rectangle. **(C)** Network of clonotype clusters derived from pairwise sequence alignment of CDR3α sequences for CD8$^+$ cells. Each circle represents a unique CDR3α sequence and the size of the circle is proportional to the size of the clone. Dashed lines represent clusters of sequences whose distance was within a similarity score of 5 based on BLOSUM62 scoring. Clonotypes carrying CDR3α identical or similar to NY8.3 are highlighted with a red line.

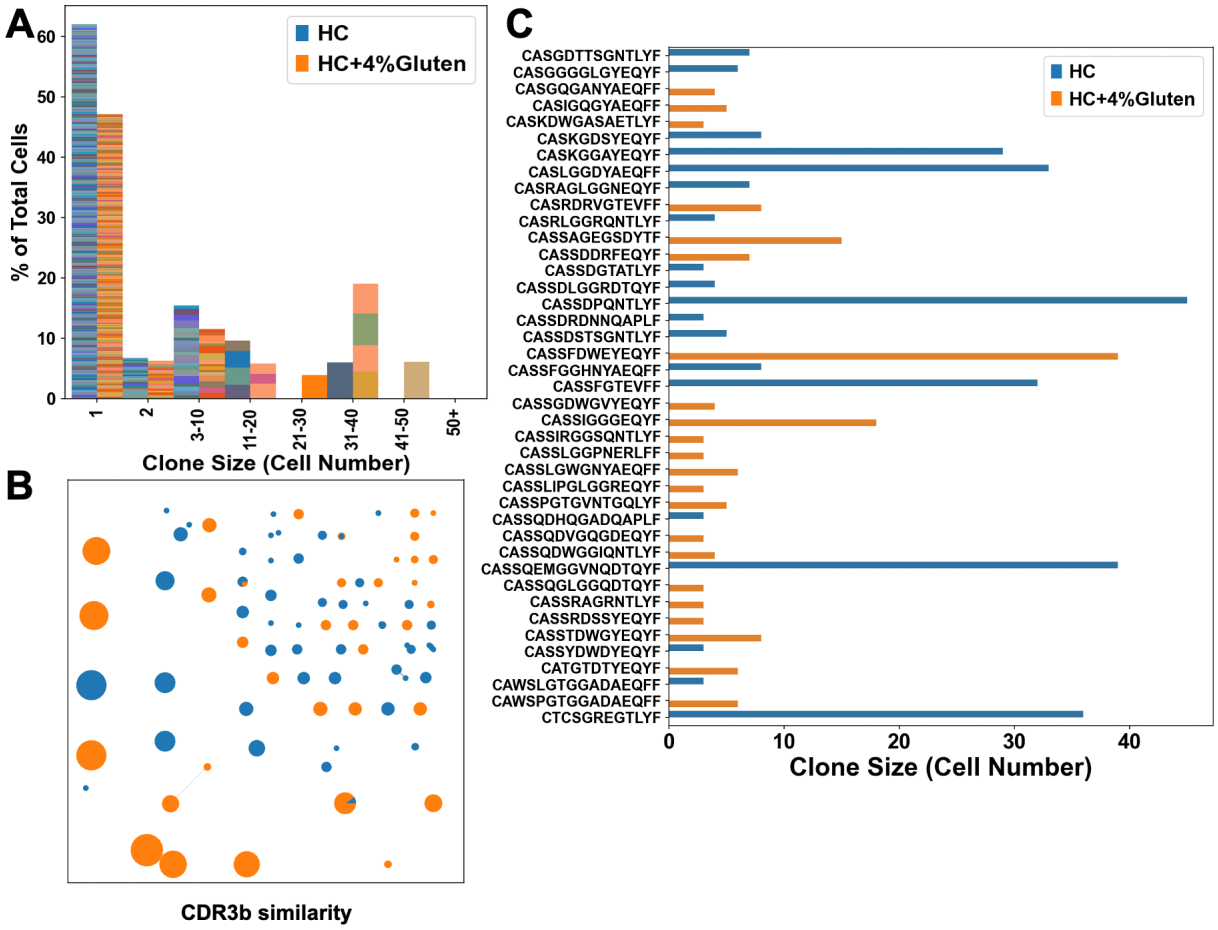

Figure 4.11: Analysis of CDR3β sequences from islet CD8$^+$ T cells. **(A)** Distribution of expanded CDR3β sequences and clonal sizes in islet infiltrates from mice on HC (blue background) and HC+4%Gluten (orange background) diets from single cell sequenced CD8$^+$ cells (three mice per diet). The abscissa represents the size of expanded clonotypes and the ordinate represents the % of individual clonotype of total sequenced CD8$^+$ cells. Individual CDR3β sequences are represented by different colors. **(B)** Network of clonotype clusters derived from pairwise sequence alignment of CDR3β sequences for CD8$^+$ cells. Each circle represents a unique CDR3β sequence and the size of the circle is proportional to the size of the clone. Dashed lines represent clusters of sequences whose distance was within a similarity score of 5 based on BLOSUM62 scoring. **(C)** The number of expanded cells from each clonotype ($\geq$3 cells with identical TCRs) derived from expanded CDR3β sequences from CD8$^+$ cells in the islets of HC (blue) and HC+4%Gluten (orange) fed mice.

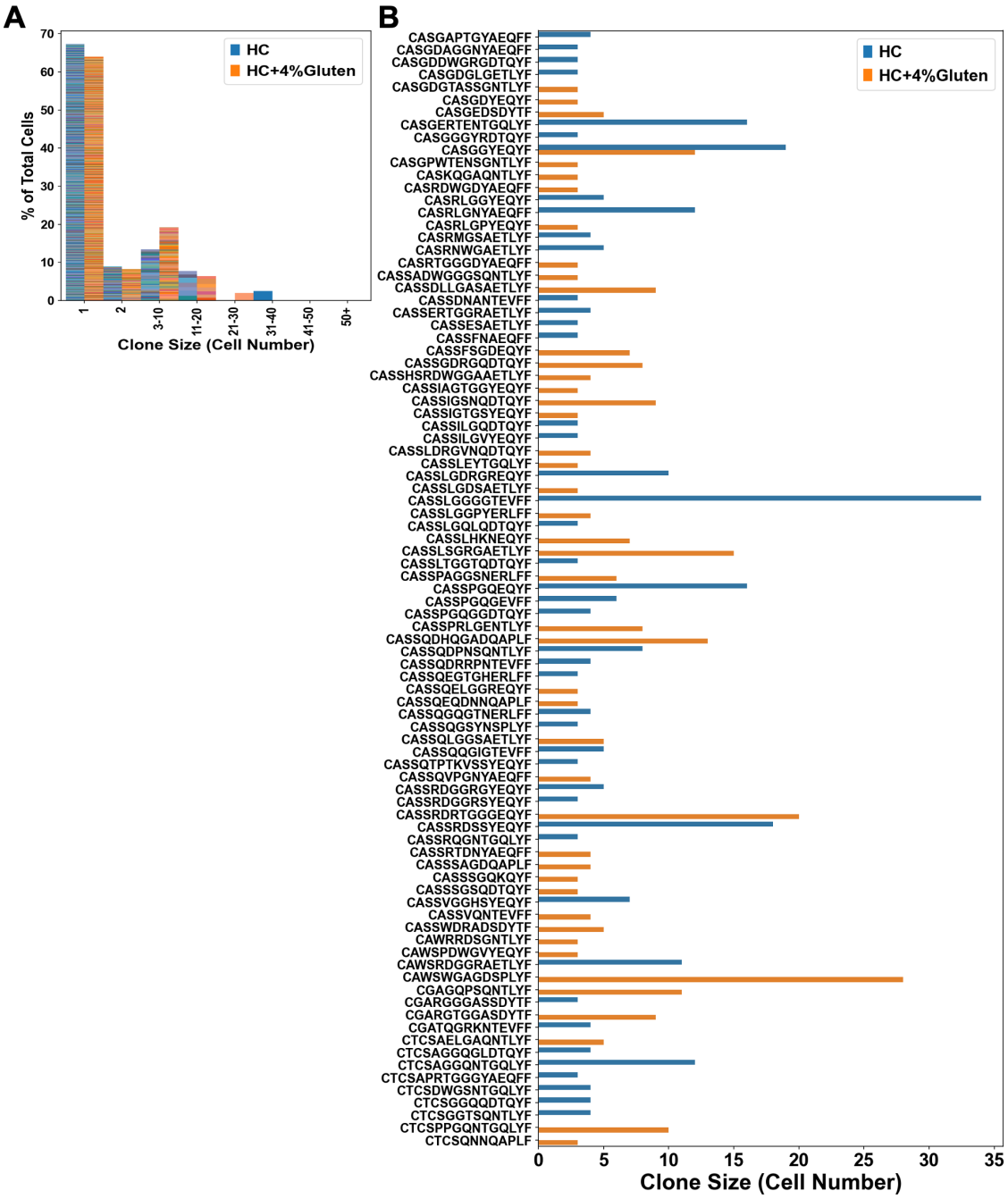

Figure 4.12: Analysis of CDR3α sequences from islet CD4$^+$ T cells. (A) Distribution of expanded CDR3α sequences, and clonal sizes in islet infiltrates from mice on HC (blue background) and HC+4%Gluten (orange background) diets from single cell sequenced CD4$^+$ cells. The abscissa represents the size of expanded clonotypes and the ordinate represents the % of individual clonotype of total sequenced CD4$^+$ cells. Individual CDR3α sequences are represented by different colors. **(B)** The number of expanded cells from each clonotype derived from expanded CDR3α sequences from CD4$^+$ T cells in the islet infiltrates in HC (blue) or HC+4%Gluten (orange) fed mice.

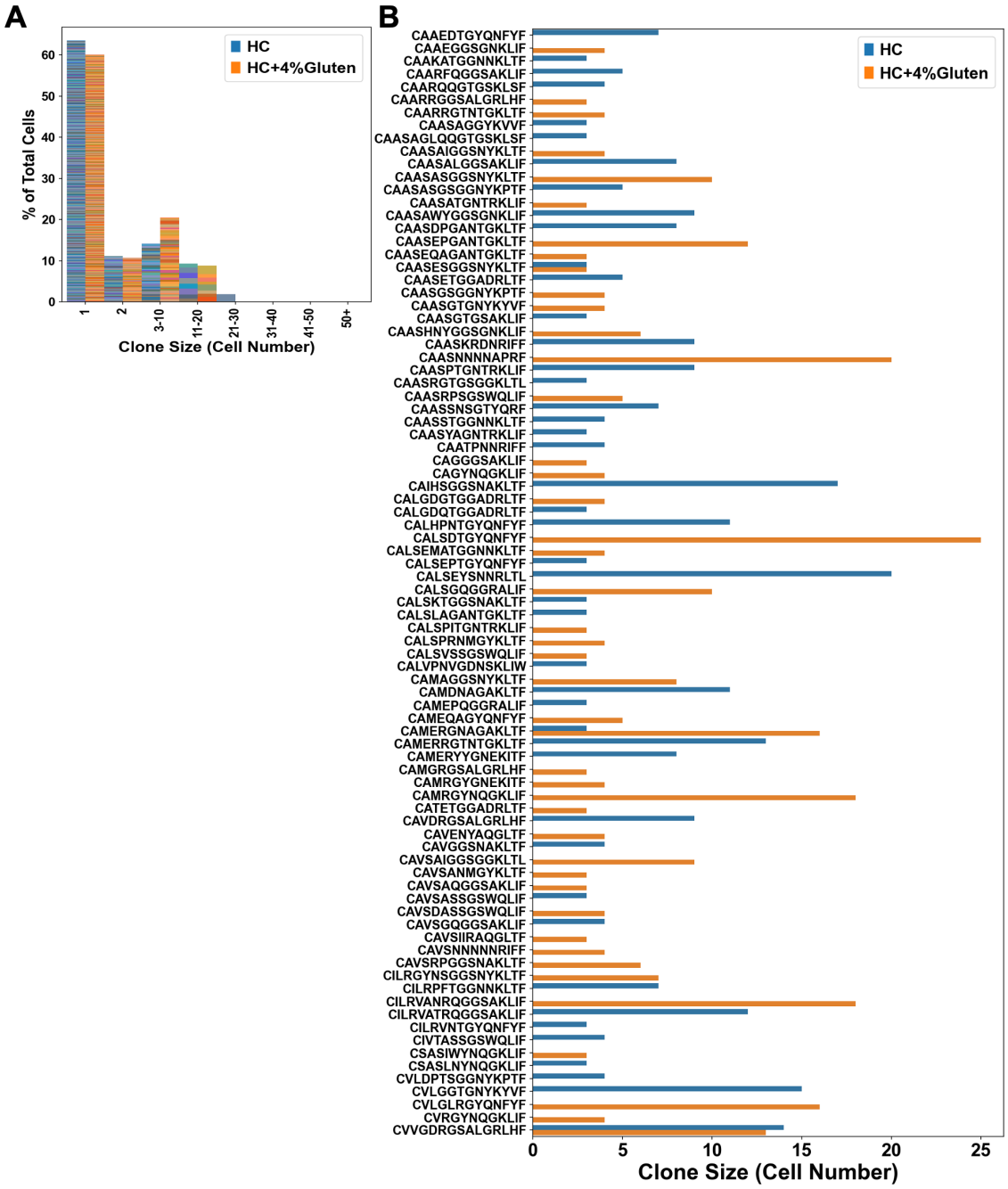

Figure 4.13: Analysis of CDR3β sequences from islet CD4$^+$ T cells. (A) Distribution of expanded CDR3β sequences, and clonal sizes in islet infiltrates from mice on HC (blue background) and HC+4%Gluten (orange background) diets from single cell sequenced CD4$^+$ cells. The abscissa represents the size of expanded clonotypes and the ordinate represents the % of individual clonotype of total sequenced CD4$^+$ cells. Individual CDR3β sequences are represented by different colors. **(B)** The number of expanded cells from each clonotype derived from expanded CDR3β sequences from CD4$^+$ T cells in the islet infiltrates in HC (blue) or HC+4%Gluten (orange) fed mice.

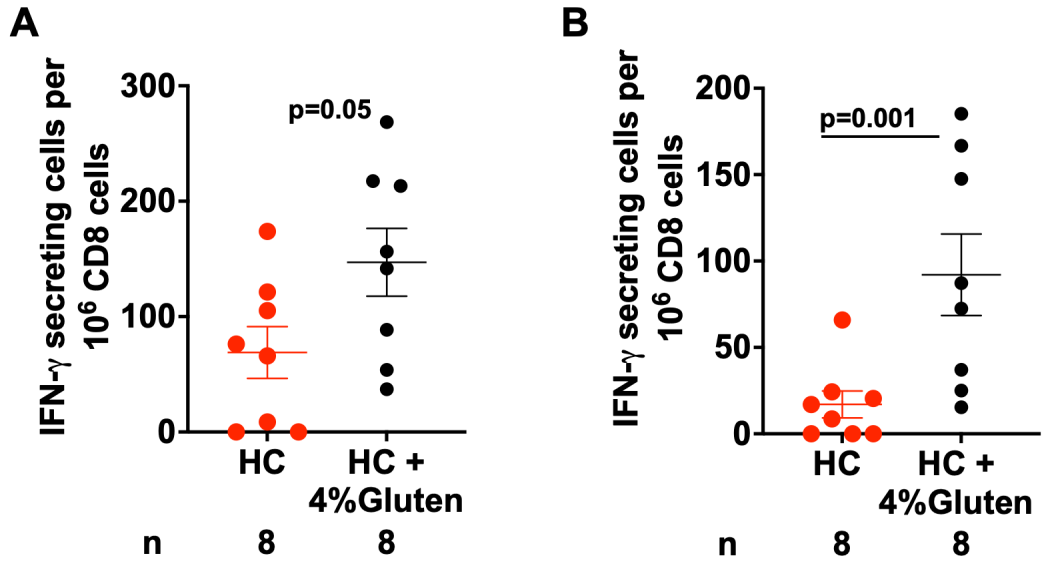

Figure 4.14: Expansion of diabetogenic CD8$^+$ T cell clones in the PLNs of HC+4%gluten fed mice. **(A)** Combined frequencies of CD8$^+$ T cells producing IFN-γ in response to NRPA7 mimic peptide, mimic peptide YFIENYLEL, or InsB 15-23 recognized by the diabetogenic clones 8.3, AI4, and G9C8 respectively in the PLN from HC fed and HC+4%gluten fed NOD mice. Each point is the sum of reactivities to each peptide (averaged from 3 technical replicates) in each mouse normalized to the frequency of CD8$^+$ T cells measured in the PLN. Mean±sem. Data are combined from 3 experiments. p values calculated using Student's t test. **(B)** The frequency of CD8$^+$ T cells producing IFN-γ in response to NRPA7 mimic peptide recognized by the diabetogenic clone NY8.3 in the PLN from HC fed and HC+4%gluten fed NOD mice. Each point is an average of 3 technical replicates per mouse. Mean±sem. Data are combined from 3 experiments. p values calculated using Student's t test.

4.4 Discussion

In this chapter we found that the addition of gluten to a protective HC diet restored T1D to NOD mice. Protection by HC was based on its ability to reduce β-cell insulin secretion and associated ER stress; gluten did not restore high insulin secretion. Instead, gluten promoted islet inflammation resulting in the enhanced activation of myeloid cells, T cells, and Tregs in the islets. $CD8^+$ T cells sharing similar TCRs to the highly diabetogenic, IGRP reactive clone 8.3 were also increased in frequency in the islets and their draining lymph nodes in HC+4%gluten fed NOD mice.

The use of both scRNA-Seq and flow cytometry to evaluate the islet infiltrates of HC and HC+4%gluten fed mice allowed us to use each technique's advantages to compensate for the other's disadvantages. scRNA-Seq allowed us to capture the full gene expression profiles of diverse cell types from the islets in an unbiased fashion. One technical limitation of scRNA-seq experiments is the small number of biological replicates used due to the high cost of the technique; in our experiment we sequenced one pool of mice per group. While sequencing pools of cells from multiple mice is a common practice and has been used by other groups to evaluate infiltrating leukocyte frequency changes in the islets of NOD mice [45], it does not reflect any mouse-to-mouse variability that may exist within a group. Flow cytometry allowed us to compensate for this disadvantage by using a much larger number of biological replicates, albeit evaluated using a much more restricted set of markers. Using flow cytometry, we found that there was great diversity in the overall quantity and frequencies of major subpopulations of infiltrating leukocytes in the islets of HC and HC+4%gluten fed mice **(Figure 4.6A-D)**. This diversity would not have been captured with fewer replicates and likely reflects both the asynchronous nature of T1D development in NOD mice (individual mice do not progress to diabetes at the same rates), and the fact that the phenotypes imposed by the diets are not completely penetrant [1/4 of mice fed HC develop T1D and 1/4 of mice fed HC+4%gluten fail to **(Figure 4.1)**].

Despite this diversity, these techniques allowed us to find that gluten changed multiple

cell types in the islets. Gluten upregulated inflammatory response pathways in β-cells, cytokine production and costimulatory receptor expression in myeloid cells, proliferation in T cells, and tissue residency markers in Tregs **(Figure 4.5, Figure 4.8, and Figure 4.9)**. These factors formed an interconnected system. The inflammatory cytokines produced by myeloid cells such as *Tnf* were the same cytokines that β-cells responded to. In addition, whether myeloid cells were activated up or downstream of T cell activation, upregulation of costimulatory receptors on islet infiltrating DC's would likely contribute to the elevated islet T cell proliferation we observed in HC+4%gluten-fed mice. DC costimulatory receptor upregulation may have also contributed to the expansion of islet CD103$^+$ Tregs. As these Tregs were previously shown to have greater TCR sensitivity and proliferative capacity than Tregs that lacked CD103 [208], they would likely expand as a percentage of total Tregs under conditions of greater APC activation to control said APCs and downstream T cell proliferation. As gluten-fed mice still had more activated myeloid cells and proliferative T cells in their islets, the expansion of islet tissue resident Tregs should be viewed as yet another marker of elevated inflammation and not as evidence for successful suppression of autoimmunity.

We were very interested to find an expansion of IGRP reactive CD8$^+$ T cells in the islets of mice fed gluten. Like in previous studies [215, 216], specificity to IGRP was conferred by similar CDR3α sequences whereas CDR3β sequences were not obviously important. It has been estimated that 20-30% of the CD8$^+$ T cells in the islets of NOD mice recognize peptides from IGRP [162]. Furthermore, a representative IGRP reactive clone, NY8.3, has been shown to be highly diabetogenic in both transgenic NOD mice [211] and in adoptive transfer studies [52, 217]. These cells also expand in NOD mice carrying a gain of function mutation in the gene *Stat3* and are hypothesized to contribute to their accelerated diabetes development [216]. The expansion of IGRP-reactive cells in mice fed gluten may also contribute to their accelerated diabetes development. Whether this expansion is fueled by the elevated myeloid activation observed in the islets or by another unknown mechanism is an open question.

In our hands, replacing only 1/5th of HC with gluten was enough to reverse HC's protective effect and restore high T1D incidence. Gluten did not simply reverse HC's protective effects on β-cell physiology but promoted T1D by stimulating the immune system enough so that autoimmunity could kill the β-cells. This suggests that the protective effect of HC was in reality quite fragile and lends support to the idea that other dietary ingredients may have interfered with HC protection in human trials [113]. As human diets are complex and it would be highly impractical to prescribe strict highly purified diets as prophylactic treatments with no option for the addition of other foods, our findings with HC and gluten underscore the need to study different proteins in isolation and together. To evaluate the combined effects of different dietary components, a similar model to the balanced signal hypothesis [79], normally applied to microbial involvement in T1D, can be used. In a complex diet, the balance of pro- and anti- diabetic signals coming from different dietary components determines the cumulative effect of that diet on diabetes development.

CHAPTER 5

BACTERIAL DIGESTION OF GLUTEN IS REQUIRED FOR STIMULATION OF THE IMMUNE SYSTEM AND PROMOTION OF TYPE-1 DIABETES.

5.1 Preface

T1D incidence was monitored by A.V. Chervonsky, A. Kuznetsov, J. Ratiu, and myself. *Tgm2*.KO mice were made by J. Ratiu and D. Sereze. Western blotting was performed by L.A. Yurkovetskiy. A.A. Khan analyzed 16S rRNA sequencing data. Identification and cultivation of *E. faecalis* was performed by A. Kuznetsov. Digests of gluten were produced by A. Kuznetsov and myself. *In vitro* T cell stimulation assays were performed by A.V. Chervonsky. LAL assays were performed by A. Kuznetsov. Monocolonization experiments were carried out by A. Kuznetsov and myself. All macrophage experiments, ELISAs, flow cytometry, and production and screening of gluten reactive hybridomas were done by myself.

5.2 Introduction

Gluten's ability to promote autoimmunity and reverse protection by HC likely depended on its ability to activate the immune system. How this occurs exactly is an open question. Two hypotheses commonly put forward to explain how a biological variable (such as diet or the microbiota) influences autoimmune activation are molecular mimicry and bystander activation. Molecular mimicry occurs when foreign peptides are sufficiently similar or identical to self peptides and, when presented by professional APCs, lead to cross reactive T cell activation against self tissues. Molecular mimicry has been shown to play a role in rheumatoid fever which is caused by the destruction of myocardial cells due to cross reactivity with streptococcal antigens [218]. Importantly, molecular mimicry does not normally provoke autoimmunity without simultaneous activation of APCs. For example, NOD mice

that transgenically express a Lymphocytic choriomeningitis virus (LCMV) protein under control of the insulin promoter do not progress to T1D even when also carrying a TCR specific for that protein until they are infected with LCMV [219]. Bystander activation on the other hand, is proposed to occur either when an APC activated by a pathogen associated molecular pattern (PAMP) presents both non-self and self-antigens or when T cells activated against non-self antigens secrete cytokines that lower the threshold for further T cell activation. The result of either scenario is the activation of T cells not directly involved in the initial response. One example of bystander activation is the activation of T cells specific to ovalbumin or to gluten following an infection with reovirus [220]. In this case, gluten was the target rather than the driver of bystander activation. Whether gluten could drive immune responses to self antigens, i.e. antigens involved in T1D, is an open question. Lastly, gluten's influence on T1D in animal models has been inconsistent in the literature with some groups finding a pro-diabetic effect [85, 105, 106] and other groups finding no effect [97, 98]. One explanation for these inconsistencies could be microbial differences between facilities. Therefore, the first variable to test was gluten's reliance on the microbiota.

Here we found that gluten's ability to promote T1D and activate the immune system depended on the microbiota and more specifically on the proteolysis of gluten by microbes. Using *Enterococcus faecalis* (*E. faecalis*) as a model commensal microbe, we found that microbial secreted proteases could digest gluten, promoting the activation of both innate and adaptive immunity. Innate immunity was activated by LPS released from digested gluten, increasing macrophage sensitivity to microbial PAMPs. Adaptive immunity could be activated by gluten peptides released by proteolytic cleavage. Finally, LPS released from microbially-digested gluten promoted T1D development. Gnotobiotic NOD mice lacking TLR4 and monocolonized with E. faecalis were resistant to T1D when fed HC+gluten diet. Together, these results indicated that microbial digestion could make an otherwise innocuous dietary component diabetogenic by modifying its ability to stimulate the innate immune system and revealed a novel mechanism for LPS in the promotion of T1D by gluten.

5.3 Results

5.3.1 Promotion of T1D by gluten is microbe dependent.

To test if gluten's promotion of T1D depended on the microbiota, we placed GF NOD mice on HC+4%gluten diet and monitored their cumulative diabetes incidence. Mice fed HC+4%gluten did not develop diabetes in GF conditions **(Figure 5.1)** therefore promotion of T1D by gluten was microbiota dependent.

How could the microbiota be involved? We hypothesized that microbes could activate transglutaminase 2 (TGM2), an enzyme involved in celiac disease that can deamidate or transamidate glutamines in peptides presented to T cells [69]. Using CRISPR-Cas9 technology, we deleted exons 1 and 2 of the *Tgm2* gene from NOD mice **(Figure 5.2A)** resulting in no detectable protein in NOD.*Tgm2*KO animals **(Figure 5.2B)**. NOD.*Tgm2*KO animals had no reduction in T1D incidence either on regular chow or HC+4%gluten diet **(Figure 5.2C)**. In addition, no difference in diabetes incidence was observed when the mice were housed at a different institution **(Figure 5.2D)** suggesting that the lack of protection at the University of Chicago was not due to the presence/absence of particular microbes. Therefore, TGM2 was not involved in the promotion of T1D by gluten and another hypothesis was needed.

Gluten proteins are barely soluble in physiologic salt conditions [202]. We hypothesized that microbes may break gluten down into soluble biologically active fragments capable of promoting the immune activation discussed at length in chapter 4. Testing this hypothesis required the identification of bacteria capable of digesting gluten. To determine the anatomical location where gluten digesting bacteria were likely to reside, we colonized GF mice fed either chow, HC, or HC+4%gluten with cecal contents from a single chow fed NOD mouse housed in SPF conditions, collected SI and cecal contents after 8 weeks, and subjected them to 16S rRNA sequencing. The input microbiota was standardized to ensure that any microbial changes we observed were due to diet and not due to differences in the input microbiota.

Interestingly, SI bacteria from HC+4%gluten fed mice were far more similar to SI bacteria from chow fed than HC fed mice **(Figure 5.3A and C)** whereas the cecal bacteria from HC and HC+4%gluten fed mice were far more similar to each other than to cecal bacteria in chow fed mice **(Figure 5.3B and D)**. As chow and HC+4%gluten diets both promoted T1D while HC diet attenuated T1D, we reasoned that gluten digesting bacteria involved in T1D promotion by gluten likely resided in the SI.

5.3.2 Bacterial proteolysis of gluten activates the innate immune system.

To identify bacteria capable of digesting gluten, SI contents from HC+4%gluten fed mice were plated on BHI-agar plates supplemented with an ethanol extract of gliadin, one of the proteinaceous components of gluten. The gliadin supplemented BHI-agar plates appear cloudy due to gliadin's poor solubility [202] thereby allowing the detection of gliadin digesting bacteria: they form transparent halos around their colonies **(Figure 5.4A)**. Based on 16S rRNA sequencing, most of the gliadin digesting colonies were identified as *Enterococcus faecalis* (*E. faecalis*). *E. faecalis* is a gram-positive commensal microbe present in both mice and humans. The species secretes two proteases, GelE, a zinc metalloprotease, and SprE, a serine protease, whose expression are both controlled by the quorum sensing operon *fsrB* [156]. Luckily, the species has tractable genetics, and several mutants incapable of secreting these proteases were available to us. When plated on BHI-gliadin agar, WT *E. faecalis* (parental strain OG1RF, *gelE+ sprE+* [155]) readily formed halos around its colonies (signaling gluten digestion) **(Figure 5.4B)** while a mutant strain with an *fsrB* operon inactivated by site directed mutagenesis [156] did not **(Figure 5.4C)**.

To test whether bacterially digested gluten could activate the innate immune system, we exposed peritoneal macrophages to sterile filtrates of gluten digested by *E. faecalis* in liquid phase cultures overnight and measured cytokine concentrations in the culture supernatant after a short-term response (TNF, 6 hours) or long-term response (IL-6, overnight). For this assay, *E. faecalis* (with or without gluten) was grown in tissue culture DMEM medium

as BHI activated macrophages on its own. While supernatants from gluten incubated in DMEM without *E. faecalis* triggered cytokine secretion on their own, their stimulatory capacity increased dramatically when incubated with *E. faecalis* **(Figure 5.4D and E)**. The high stimulatory capacity of gluten incubated with bacteria depended on bacterial proteolytic digestion as mutant *E. faecalis* lacking a functional *fsrB* operon or the proteases GelE and SprE could not increase gluten's stimulatory capacity above background. Importantly, cytokine production by macrophages depended entirely on the presence of gluten as *E. faecalis* grown in DMEM was not stimulatory on its own. The increase was also primarily due to the protease SprE as *sprE- E. faecalis* [158] digested gluten elicited less whereas *gelE- E. faecalis* elicited even more TNF secretion from peritoneal macrophages **(Figure 5.4F)**. This is consistent with data showing that GelE negatively regulates SprE [221].

The peritoneal cavity is not the site of diabetes induction in the NOD mouse. It was possible that macrophages in sites more relevant to T1D development that drain antigens from the gut might be more tolerant to innate immune stimulation due to the presence of high concentrations of LPS. Macrophages from the PLN exposed to our bacterial digests of gluten were not insensitive to stimulation and still secreted TNF **(Figure 5.5A and B)**. Innate immune stimulation by bacterial digested protein was also highest when gluten was used as the protein source **(Figure 5.5C)**. Unlike gluten, casein (the other protein in HC+4%gluten diet) and BSA did not elicit high levels of cytokine secretion from peritoneal macrophages following proteolysis by *E. faecalis*.

5.3.3 Activation of the innate immune system by bacterially digested gluten depends on gluten-associated LPS

Gluten was highly capable of activating the innate immune system following digestion by protease secreting *E. faecalis*. As gluten is a complex of different proteins, we reasoned that the stimulatory factor released could be a small peptide. To our surprise, the gluten digests were highly heat sensitive in many cases eliciting more cytokine secretion following

heat treatment at 85°C for 30 minutes **(Figure 5.6A)**. The digests were also insensitive to additional treatment with the enzyme tryspin **(Figure 5.6B)**. Finally, digests passed through a 50kD filter lost all of their stimulatory capacity **(Figure 5.6C)** suggesting the biologically active component was larger than 50kD. Therefore the stimulatory factor was unlikely to be a small peptide.

As the macrophage activating moiety did not appear to be proteinaceous, we treated the digests with the LPS-inactivating antibiotic polymyxin B **(Figure 5.7A)**. Treatment with polymyxin B destroyed the stimulatory capacity of bacterially digested gluten and of undigested gluten. Macrophages lacking MyD88 or the LPS receptor TLR4 were also resistant to activation by bacterially digested gluten **(Figure 5.7B)**. Therefore, bacterial digestion likely released LPS (or another hydrophobic compound with similar activity) that was naturally associated with gluten. Testing several commercial flour sources by the LAL assay revealed LPS present in almost all of them **(Figure 5.7C)**.

If the stimulatory factor is LPS, then why is bacterial digestion required? We envisioned two possibilities: 1. The LPS may be deeply sequestered inside of gluten and require intense proteolytic digestion to be released. 2. Peptides from digested gluten could bind to LPS enhancing its stimulatory capacity. Such enhancement is not unprecedented in the literature. LPS bound to heat-shock proteins [222, 223] or retroviruses [224] is far more potent than LPS added directly to macrophages. If possibility 1 was correct, increasing doses of LPS should increase TNF production by macrophages additively when combined with the LPS released from gluten. If possibility 2 was correct, then digested gluten should be able to amplify (greater than an additive effect) the activating potential of non-stimulatory doses of LPS (calculated empirically in **Figure 5.8A**). When mixed with bacterial digests of gluten or with gluten alone, suboptimal doses of LPS enhanced macrophage TNF secretion beyond response to the sum of the two agonists **(Figure 5.8B)**. The amplification response was always stronger when bacterially digested gluten was used compared to undigested gluten. Interestingly, suboptimal doses of LPS could only be amplified by gluten digests if hexa-

acylated LPS from *Proteobacteria* was used rather than penta-acylated LPS from *Bacteroides* **(Figure 5.8C)**.

In addition to LPS, we also found that bacterially digested gluten could amplify the inflammatory response to a TLR2 agonism. Suboptimal doses of Pam3CysSerLys4 (Pam3CSK4) a synthetic, triacylated lipopeptide agonist of the TLR1/TLR2 heterodimer, **(Figure 5.9A)** mixed with *E. faecalis* digests of gluten or with undigested gluten resulted in a non-linear amplification of TNF secretion **(Figure 5.9B)**; as with LPS, the amplification response was stronger with bacterially digested gluten compared to undigested gluten. In addition, the ability of bacterially digested gluten to amplify TLR2-dependent TNF production strongly depended on LPS from the bacterially digested gluten as the magnitude of the response was significantly diminished when the gluten digest was pretreated with polymyxin B prior to mixing with Pam3CSK4 **(Figure 5.9B, blue line)**. LPS did not account for all of the amplification though. Bacterially digested gluten mixed with suboptimal doses of Pam3CSK4 was still capable of eliciting a non-additive increase in TNF secretion from macrophages that lacked the LPS receptor TLR4 **(Figure 5.9C)**, though the magnitude of TNF secretion was severely diminished. While *E. faecalis* supernatants were not stimulatory when grown without gluten **(Figure 5.4D and E)**, we reasoned that non-stimulatory amounts of peptidoglycan may be shed from the gram-positive bacteria during digestion of gluten potentially activating TLR2 alongside LPS stimulation of TLR4. Simultaneously activating both TLR4 and TLR2 signaling triggers synergistic cytokine production from macrophages [225, 226] and treating macrophages with low doses of both Pam3CSK4 and LPS produced curves similar to the amplification observed using bacterially digested gluten **(Figure 5.9D)**. Therefore, gluten derived LPS may signal alongside other PAMPs to elicit a stronger innate immune response.

Finally, *E. faecalis* was not the only commensal we identified that could digest gluten activating innate immunity. Two other bacteria isolated from the small intestine, *Micrococcus luteus* and *Paenibaccilus motobuensis*, could digest gliadin when grown on BHI-gliadin

agar and could render gluten stimulatory after a short digestion in DMEM **(Figure 5.10)**. Importantly, like *E. faecalis*, neither bug was stimulatory without gluten under our culture conditions and TLR4 KO macrophages were refractory to stimulation by the new bacterial gluten digests.

5.3.4 Bacterial proteolysis of gluten activates adaptive immunity

In addition to LPS, we reasoned that bacterial proteolysis of gluten could release peptides capable of activating adaptive immunity. To test this, we added bacterial digests of gluten to a mixture of the mouse lymphoma line 578_BV7_AV12 (.578) carrying a human T cell receptor against a gliadin-ω2-derived peptide [227] and and human transformed B cell lines carrying HLA-DQ2.5 MHC class II molecules. The addition of bacterially digested gluten and APCs triggered IL-2 production from the responder T cells **(Figure 5.11A)**. The effect here was antigen specific as a second cell line carrying a TCR reactive to gliadin-α1-peptide/HLA-DQ8 complex, TCC489.2.14 (.489) [182] failed to respond to the digest but did respond to its cognate peptide **(Figure 5.11A)**. Liberation of peptides from gluten also relied on proteolytic digestion by *E. faecalis*. While both WT (strain OG1RF [155]) and cell free supernatants from WT *E. faecalis* could liberate peptides from gluten to trigger IL-2 production from gliadin-reactive T cells, undigested gluten and gluten incubated with *E. faecalis* containing a mutated *fsrB* operon incapable of secreting the gelE and sprE proteases failed to do so **(Figure 5.11B)**. Unlike with LPS release, the liberation of peptides was entirely reliant on the protease gelE as gluten digested by *E. faecalis* mutant TX5264 [157] which lacks *gelE* due to site directed mutagenesis failed to stimulate gliadin-reactive T cell IL-2 secretion **(Figure 5.11C)**. Reconstitution of *gelE* under a heterologous promoter in the same strain restored the ability to liberate activating peptides **(Figure 5.11C)**. Therefore, in addition to activating innate immunity, bacterial digestion of gluten can also activate adaptive immunity *via* the liberation of cognate peptides.

Known T cells that drive the pathogenesis of T1D react to β-cell antigens. So how could

gluten reactive T cells play a role in the development of diabetes? We envisioned two possibilities: First, some gluten reactive T cells may be cross reactive to β-cell antigens; their activation by APCs activated by gluten derived LPS and presenting gluten peptides could therefore promote T1D. Second, activated gluten reactive T cells could produce cytokines promoting activation of professional APCs and *bona fide* autoreactive T cells. Testing either possibility *in vivo*, requires the characterization of T cells from NOD mice that respond to gluten derived peptides. To identify these T cells, we adopted a hybridoma strategy. NOD mice (the mice were maintained on HC diet to avoid prior exposure to gluten and potential loss of gluten reactive TCRs due to peripheral tolerance mechanisms) were immunized *via* footpad injection with bacterially digested gluten passed through a 10kD filter and emulsified in CFA. 10kD filtrates were used to limit the immunization to small protein fragments and peptides and to exclude intact bacterial proteases and cell wall components. After 10 days, the popliteal lymph nodes from the mice were collected, the cells were restimulated in culture for a further 3 days using splenocytes as APCs and bacterially digested gluten as a source of antigen before being fused with the T cell thymoma line BW5147. Restimulation of individual clones with bacterially digested gluten and NOD splenocytes as APCs revealed that a vast majority of the screened hybridomas produced IL-2 in response to bacterially digested gluten while the remainder either reacted to *E. faecalis* grown without gluten, splenocytes with no additional antigen, or failed to react **(Figure 5.12A, B)**. These hybridomas will serve as a valuable tool to address gluten-β-cell cross reactivity and bystander activation by gluten reactive T cells in future experiments which will be addressed in detail in this chapter's discussion and in Chapter 6: Future Directions.

5.3.5 Proteolytic digestion of gluten by microbes promotes T1D development

To test of bacterial proteolytic digestion of gluten could contribute to the development of T1D *in vivo*, we monocolonized GF NOD mice fed HC+4%gluten with either *E. faecalis* parental strain OG1RF [155] or a strain with a directly mutagenized *fsrB* operon controlling

gelE and sprE secretion [156]. While GF mice fed HC+4%gluten were protected from T1D, monocolonization with WT but not with mutant *E. faecalis* greatly increased T1D incidence **(Figure 5.13A)**. Incidence in the monocolonized mice was lower than that observed in SPF mice also fed HC+4%gluten, however, subsequent analysis revealed that *E. faecalis* isolated from the SI of mice that did not develop T1D lost the ability to digest gluten **(Figure 5.13B)**. Therefore, proteolytic digestion of gluten by bacteria is probably responsible for the promotion of T1D by gluten. To test whether gluten derived LPS contributes to T1D development, we used TLR4 deficient NOD mice, which ordinarily develop T1D when fed a standard chow diet in either SPF or GF conditions [142]. TLR4 deficient NOD mice monocolonized with *E. faecalis* and fed HC+4%gluten had greatly reduced T1D incidence compared to WT controls **(Figure 5.13C)** suggesting that LPS from gluten contributes to the promotion of T1D. Finally, to test which secreted protease from *E. faecalis* renders gluten diabetogenic, we monocolonized NOD mice fed HC+4%gluten with mutant *E. faecalis* lacking *gelE* [157]. Preliminary results show high T1D incidence in these mice **(Figure 5.13D)** suggesting that rather than gelE, *E. faecalis* protease sprE may be responsible for increasing gluten's diabetogenicity.

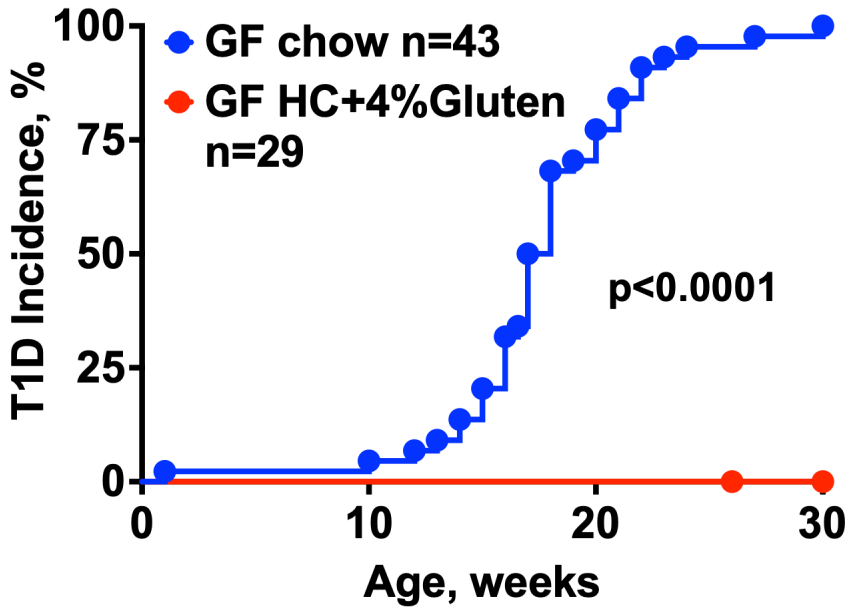

Figure 5.1: Promotion of T1D by gluten requires the microbiota. Comparison of T1D incidence in GF mice fed chow or HC+4%gluten diet. n=number of animals per group. p values were calculated using Mantel-Cox long-rank test.

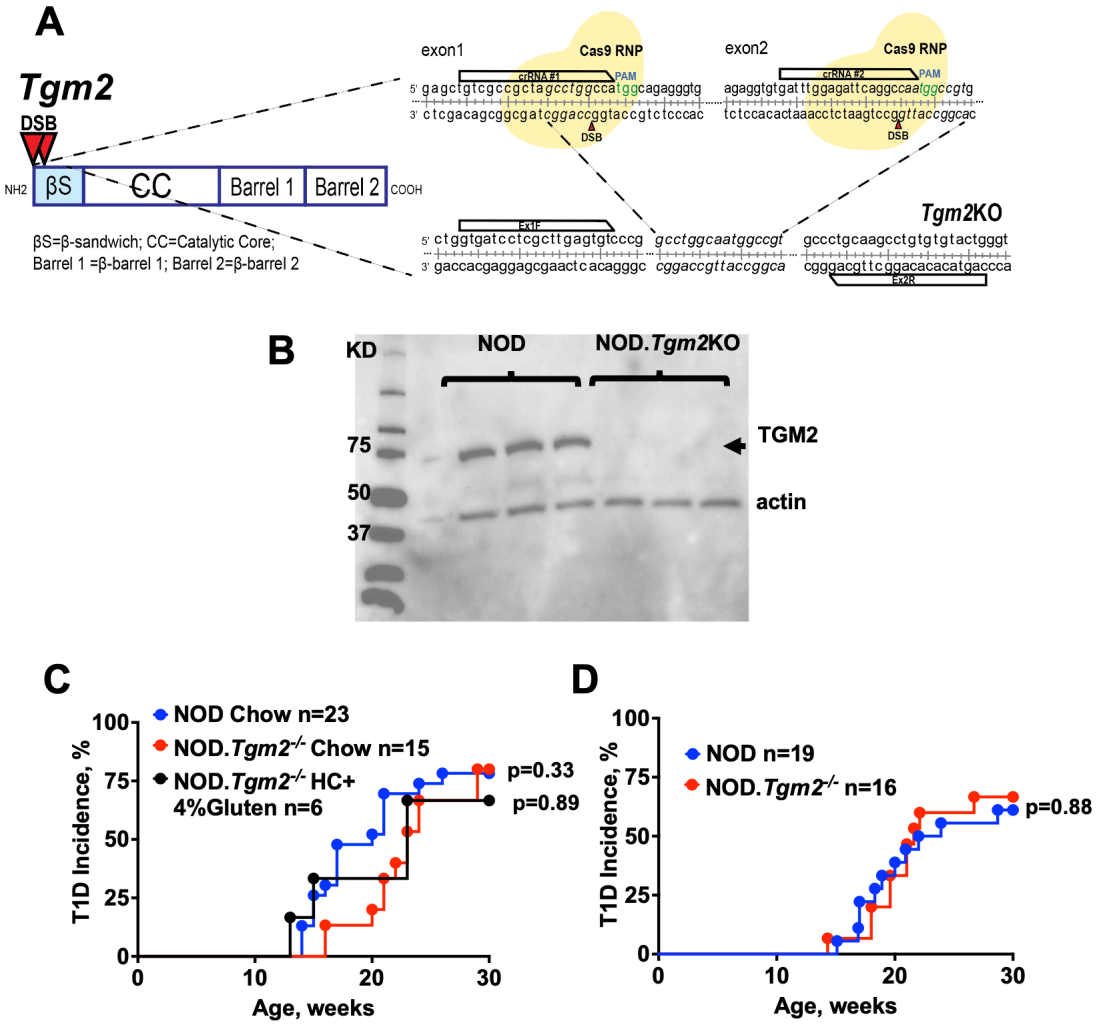

Figure 5.2: NOD mice lacking TGM2 are not resistant to T1D. (A) Targeting of *Tgm2* gene in NOD mouse. Guide RNA, PAMS and primers to detect KO allele are shown. **(B)** Western blotting with lysates of thymocytes from control NOD mice and NOD mice with deleted first+second exons of *Tgm2* gene with anti-TGM2 and anti-actin antibodies. **(C)** T1D development in TGM2-deficient mice fed either chow or HC+4%gluten diets at The University of Chicago. p value was estimated using Mantel-Cox long-rank test. n = number of animals per group. **(D)** T1D development in TGM2-deficient and TGM2-sufficient NOD mice in cohorts kept at the Jax. p value was estimated using Mantel-Cox long-rank test. n=number of animals per group.

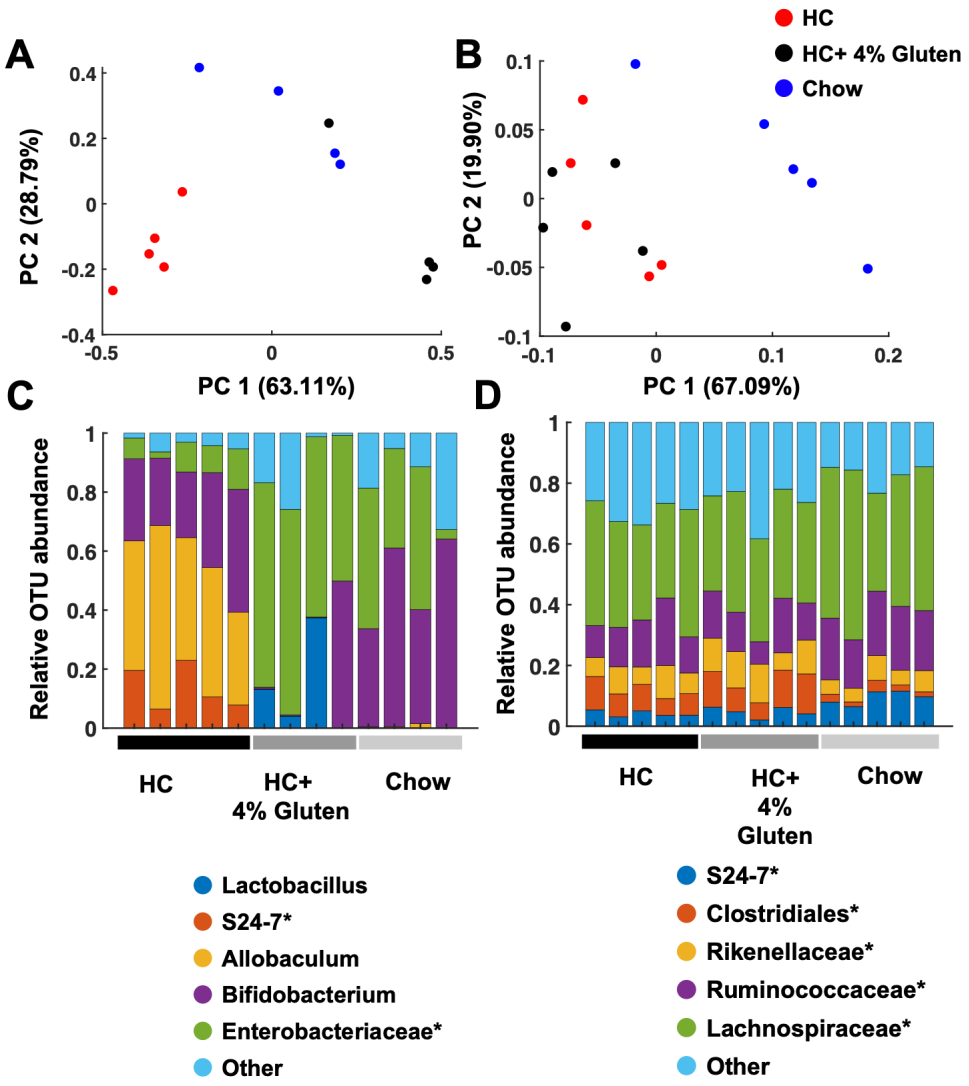

Figure 5.3: Diet induced changes to microbiota in the small intestine and cecum. **(A)** Principal Component Analysis of the small intestine (SI) microbiota of ex-GF mice simultaneously colonized with a single source of SPF NOD microbiota 8 weeks prior. **(B)** Principal Component Analysis of the cecal microbiota of ex-GF mice simultaneously colonized with a single source of SPF NOD microbiota 8 weeks prior. **(C)** Relative abundance of the top Operational Taxonomic Units (OTUs) in the SI microbiota of ex-GF mice fed indicated diets. **(D)** Relative abundance of the top Operational Taxonomic Units (OTUs) in the cecal microbiota of ex-GF mice fed indicated diets.

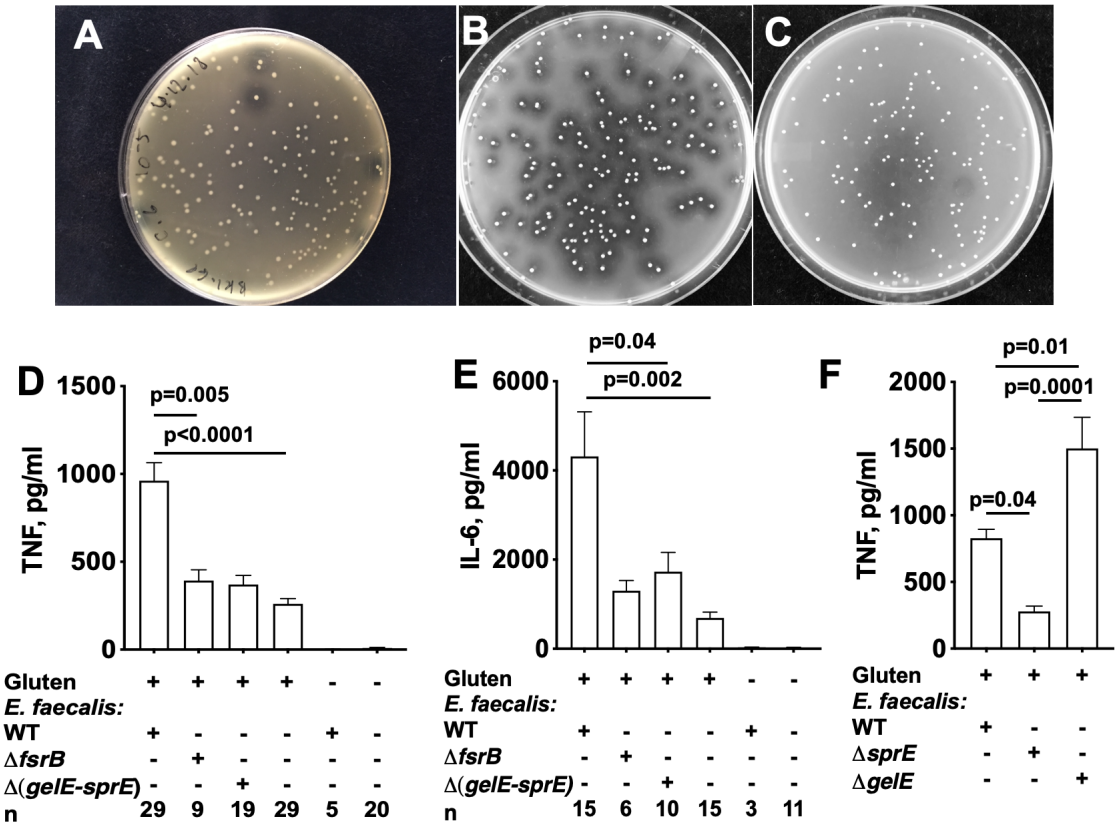

Figure 5.4: Gluten proteolysis by bacteria leads to activation of innate immunity. **(A)** Some bacteria from the small intestine of a NOD mouse plated on brain heart infusion (BHI) agar containing gliadin secrete proteases leaving transparent halos around colonies. **(B, C)** Digestion of gliadin by the wild-type *E. faecalis* strain OG1RF (B) and lack of digestion by Δ*fsrB* mutant TX5266 (C). **(D)** Production of TNF by peritoneal macrophages stimulated for 6 hrs with 10% (v/v) gluten digests by wild-type *E. faecalis* or protease-negative mutants, mean±sem. *p* values calculated using one-way ANOVA with post-hoc Tukey test. n – number of experiments per condition. **(E)** IL-6 production by macrophages after overnight stimulation with gluten digests by *E. faecalis* or protease-negative mutants, mean±sem. *p* values calculated using one-way ANOVA with post-hoc Tukey test. n – number of experiments per condition. **(F)** Comparison of TNF-eliciting activity of gluten digests performed by the wild-type *E. faecalis* strain and by *sprE*- TX5243 or *gelE*- TX5264 mutant strains. Combined results from 5 independent digestions tested independently 2 times. For each batch, TNF levels produced by macrophage incubation with gluten digest by protease-negative JRC105Δ(*gelE-sprE*) mutant was subtracted to account for digestion by intracellular proteases from dead bacteria in the culture. Mean±sem. *p* values calculated using one-way ANOVA with post-hoc Tukey test.

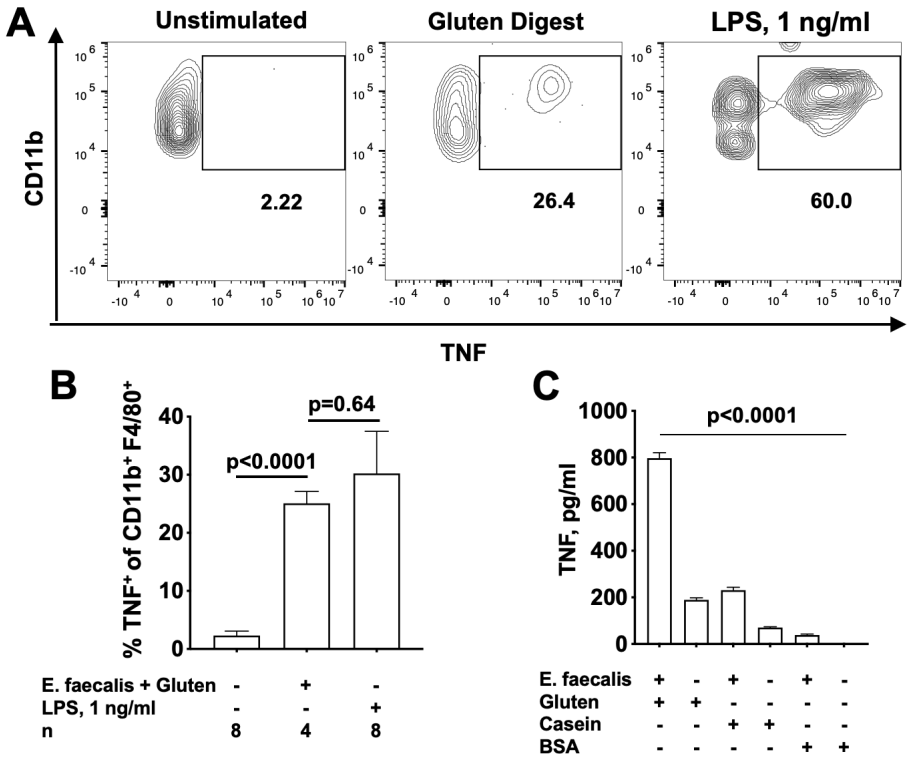

Figure 5.5: Activation of macrophages by bacterially digested proteins. (A, B) TNF expression in macrophages (CD11b+F4/80+ cells) from the pancreatic lymph nodes activated by the digest of gluten by *E. faecalis*. Treatment with 1ng/ml LPS served as positive control. FACS profiles (A) and data combined from 4 experiments (B). Mean±sem. n=number of mice. p -values calculated using Student's t test. **(C)** Activation of TNF secretion is specific to gluten. Digests of gluten, casein, or bovine serum albumin (BSA) by *E. faecalis* were used to elicit TNF secretion by peritoneal macrophages. Mean±sem. Representative of 2-3 independent experiments. p-values calculated using one-way ANOVA with post-hoc Tukey test.

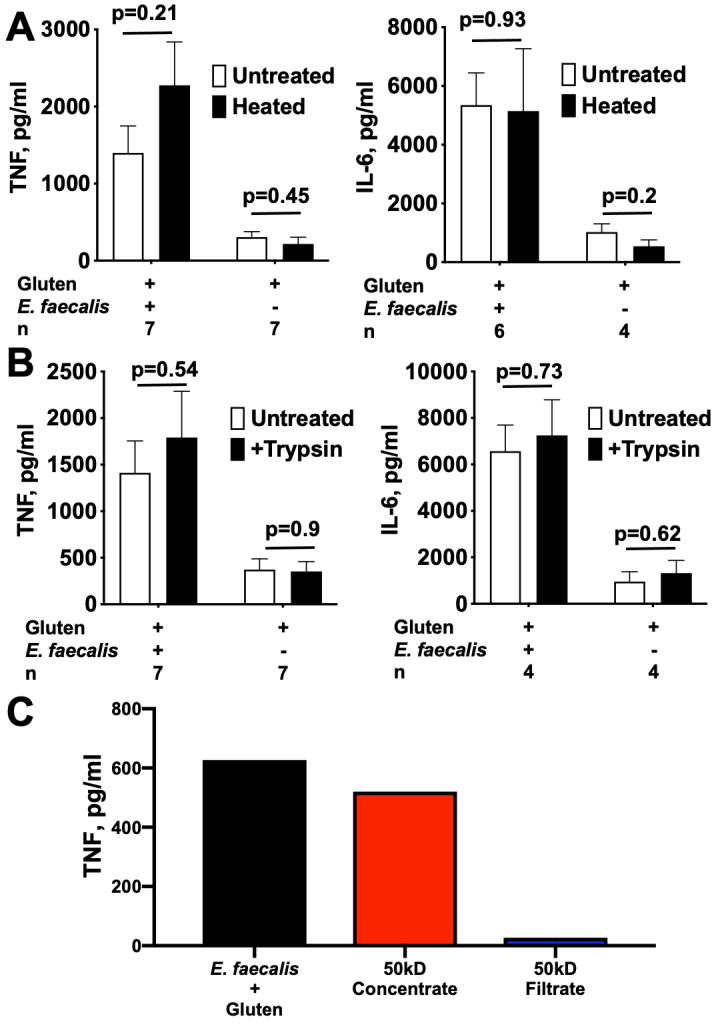

Figure 5.6: The stimulatory capacity of bacterially digested gluten is not heat or trypsin sensitive. **(A)** Gluten digests were prepared as in Figure 5.3D but heated at 85°C for 30 minutes. Mean±sem. p ¬values calculated using Student's t test. n - number of experiments per condition. **(B)** Gluten digests were prepared as in Fig. 5E and F but treated with 0.05% Trypsin overnight at 37°C. Mean±sem. p ¬values calculated using Student's t test. n – number of experiments per condition. **(C)** Gluten digested by *E. faecalis* was passed through a 50kD filter. Material retained by the filter (concetrate) was resuspended in the original volume of DMEM. Macrophages were then exposed to the original, concentrate, and filtrate. Representative of 3 independent experiments.

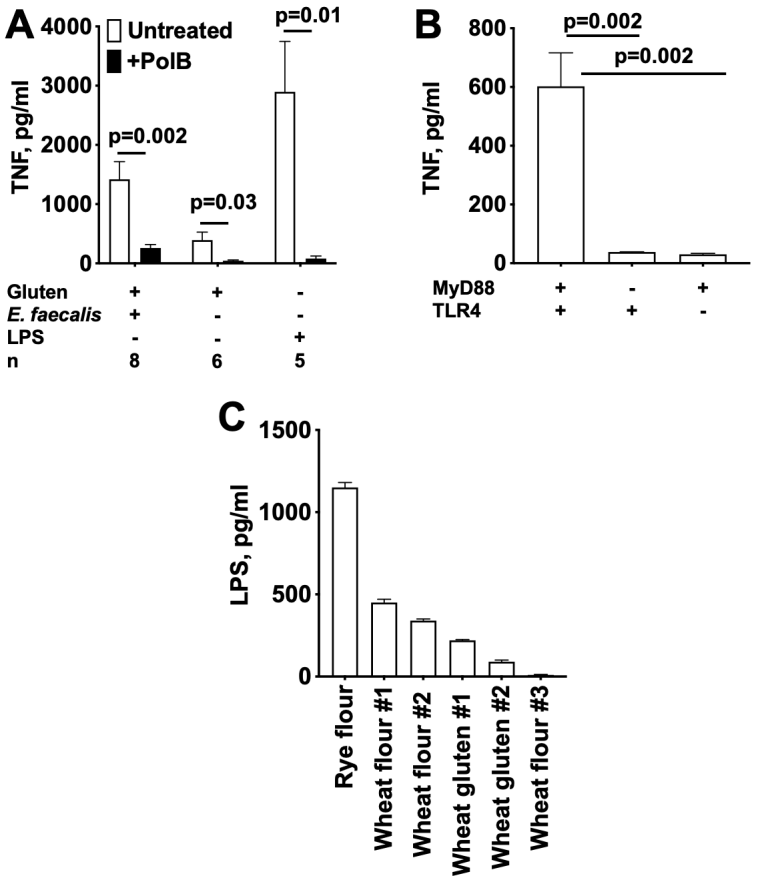

Figure 5.7: Digestion of gluten by bacterial proteases releases LPS. (A) Gluten digests were prepared as in Fig. 5.3D and either left untreated or treated with Polymyxin B beads overnight at 4°C. Polymyxin B activity was controlled by simultaneous treatment of LPS (at final concentration of 1 ng/ml). Mean±sem. p ¬values calculated using Student's t test. n – number of experiments per condition. **(B)** Stimulation of wild-type NOD and NOD.MyD88 KO or NOD.TLR4 KO macrophages with gluten digested by *E. faecalis*, mean±sem. p values calculated using one-way ANOVA with post-hoc Tukey test. Data are from a representative experiment of 4 independent experiments. **(C)** Presence of LPS detected in commercial flour and gluten samples by LAL assay. Mean±sem.

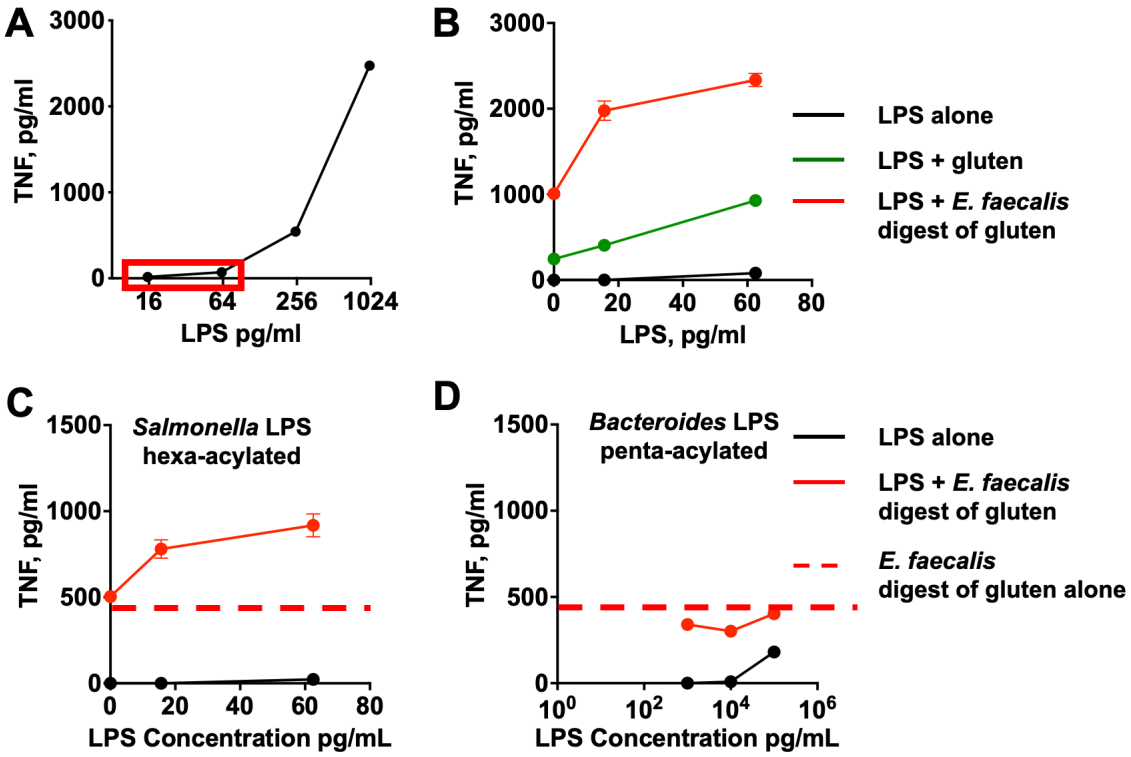

Figure 5.8: Bacterially digested gluten amplifies cytokine secretion triggered by LPS (A) Delineation of suboptimal concentrations of LPS for induction of TNF secretion by peritoneal macrophages. The red frame indicates agonist concentrations used in B and C. **(B)** Suboptimal doses of LPS (representative experiment of 6 independent experiments, black line) were mixed with undigested gluten (green lines) or *E. faecalis* digested gluten (red line) and added to macrophages for 6 hours stimulation. Mean±sem. **(C,D)** Amplification of TNF secretion by *E. faecalis'* digest of gluten in the presence of suboptimal doses of hexa-acylated LPS (C) but not of penta-acylated LPS (D). Red punctate line – the level of TNF secretion in the presence of bacterial digest of gluten without addition of LPS. Note the difference in LPS doses required to elicit macrophage responses.

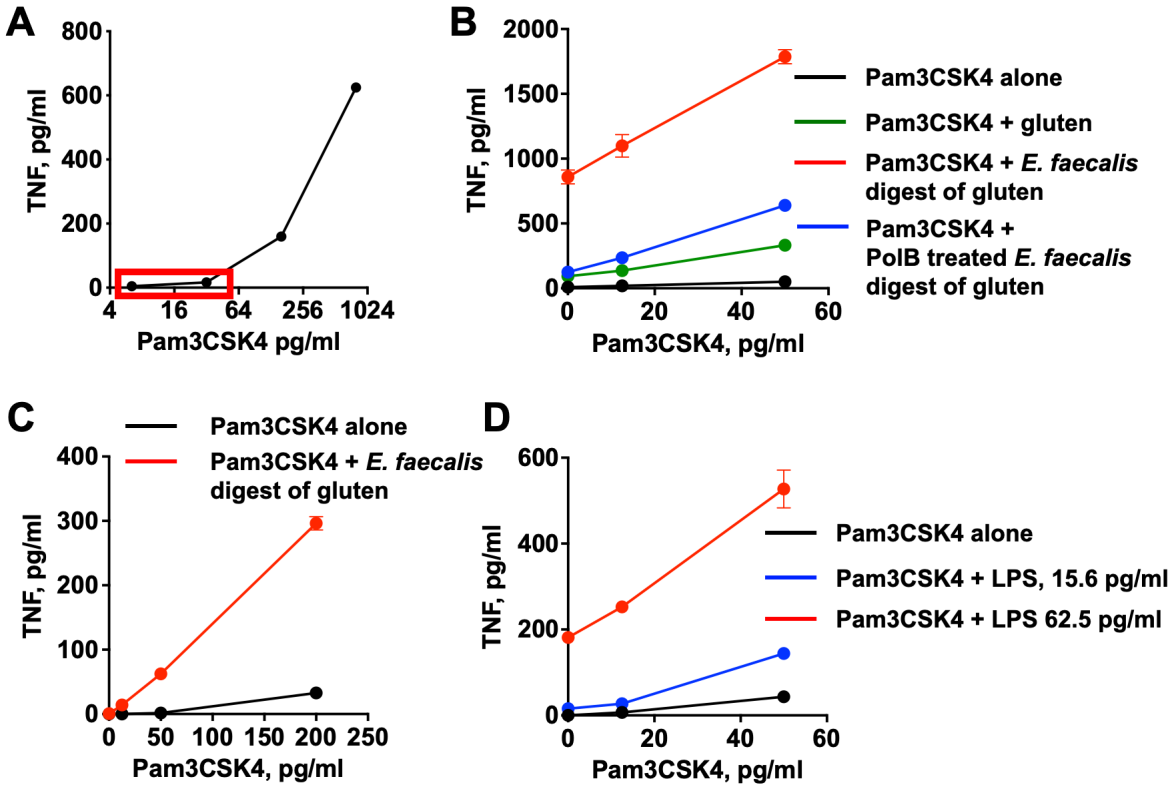

Figure 5.9: Bacterially digested gluten amplifies cytokine secretion downstream of TLR2 (A) Delineation of suboptimal concentrations of Pam3CSK4 for induction of TNF secretion by peritoneal macrophages. The red frame indicates agonist concentrations used in B, C, and D. **(B)** Suboptimal doses of Pam3CSK4 (representative experiment of 5 independent experiments, black line) were mixed with undigested gluten (green lines), *E. faecalis* digested gluten (red line), or *E. faecalis* digested gluten subsequently treated with polymyxin B (blue lines) and added to WT macrophages for 6 hours stimulation. Mean±sem. **(C)** Suboptimal doses of Pam3CSK4 (representative experiment of 4 independent experiments, black line) were mixed with undigested gluten (green lines) or *E. faecalis* digested gluten (red line) and added to TLR4KO macrophages for 6 hours stimulation. Mean±sem. **(D)** Suboptimal doses of Pam3CSK4 (representative experiment, black line) were mixed with suboptimal doses of LPS (blue and red lines, delineated in Figure 5.8A) and added to WT macrophages for 6 hours stimulation. Mean±sem.

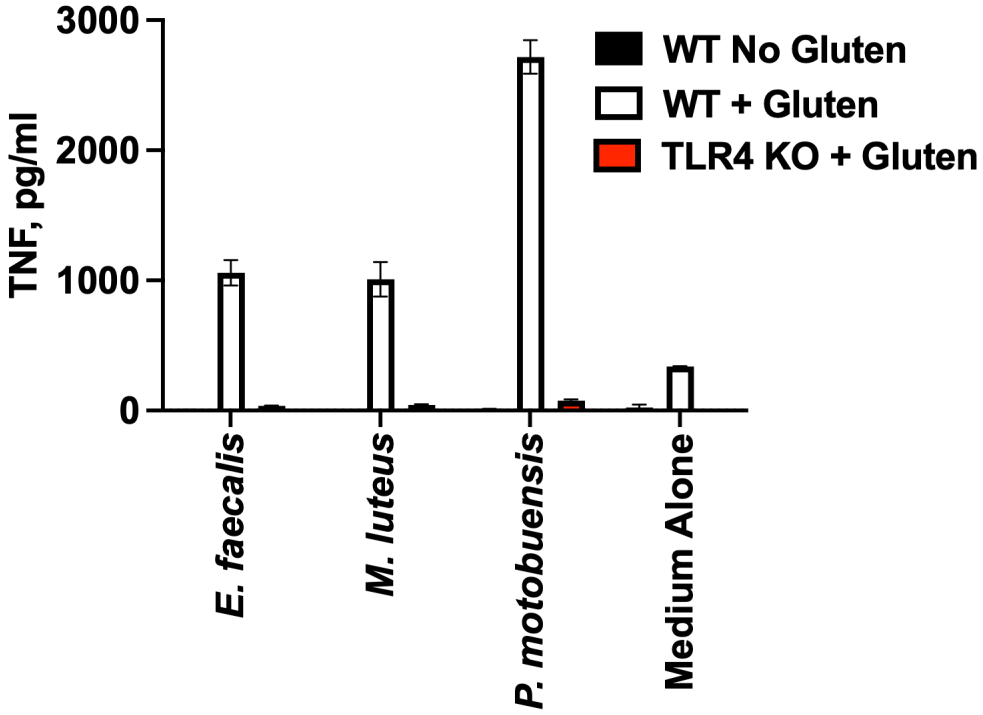

Figure 5.10: Gluten digested by SI Bacteria other than *E. faecalis* stimulates macrophages. Bacteria from the SI of NOD mice on HC+4% gluten diet were selected based on either digestion of gliadin or growths on minimal medium with added gluten. TNF secretion by NOD WT or NOD.TLR4 KO macrophages induced with digests by indicated bacteria grown in DMEM with or without gluten. Mean±sem. Representative experiment of 2 independent experiments.

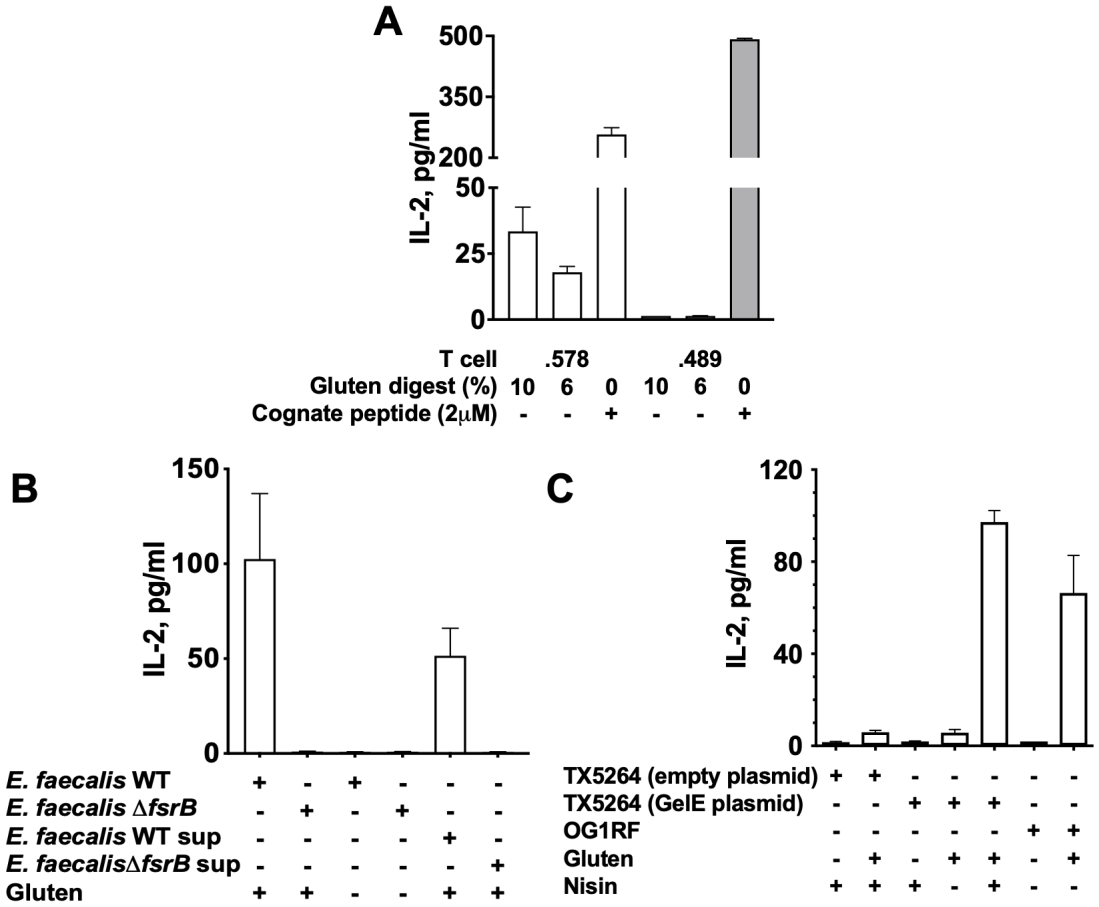

Figure 5.11: Gluten digested by bacterial proteases activates adaptive immunity **(A)** Gluten digest by *E. faecalis* activates 578_BV7_AV12 (.578) T cells specific for glia-ω2 peptide/HLA-DQ2.5 but does not activate an independent T cell TCC489.2.14 (.489) specific for gliadin alpha1 peptide/HLA-DQ8 complex. Both T cells were readily activated by their cognate peptides. Mean±sem. Representative of two experiments. **(B)** IL-2 production by the .578 T cells in the presence of HLA-DQ2.5+ antigen-presenting cells and gluten digests performed overnight by indicated bacteria or by cell-free supernatants (sup) from the same bacteria. Mean±sem from a representative experiment out of 5 independent experiments. **(C)** IL-2 production by the same T cells in the presence of APC and gluten digests produced by wild-type *E. faecalis* or *gelE*-negative mutant TX5264 transformed with an empty control plasmid or a plasmid carrying *gelE* gene under nisin-sensitive promoter. Mean±sem from one out of 2 independent experiments.

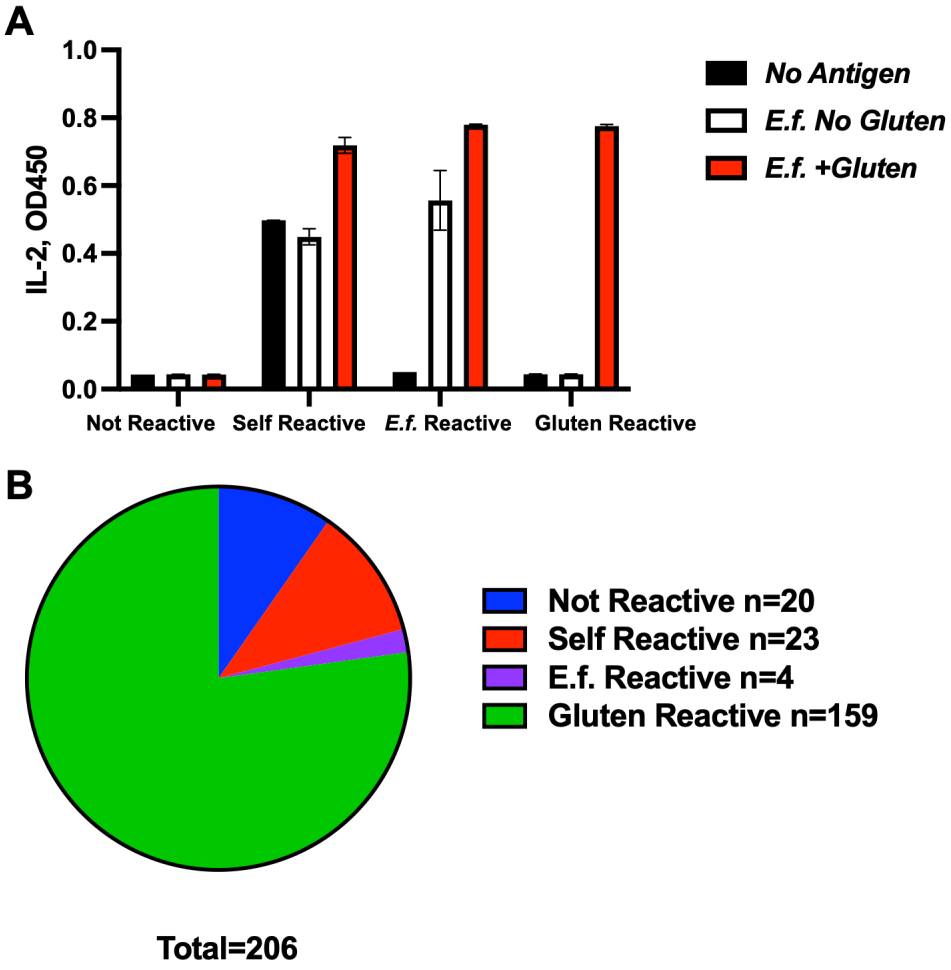

Figure 5.12: Screen of T cell hybridomas for reactivity to *E. faecalis* digested gluten (A, B) T cell hybridomas produced from mice immunized with gluten digested by *sprE*-negative mutant of *E. faecalis* TX5243 were stimulated overnight with splenocytes and either supernatant from *E. faecalis* grown in DMEM or supernatant from *E. faecalis* digested gluten. Individual hybrids were deemed gluten reactive if they produced IL-2 in response to *E/ faecalis* digested gluten only, *E.f.* reactive if they responded both to *E. faecalis* with or without gluten, and self reactive if they responded to splenocytes without additional antigen. IL-2 production by representative hybrids from each category is shown in (A). Mean±sem. A summary of reactivity of all of the tested hybrids is shown in (B).

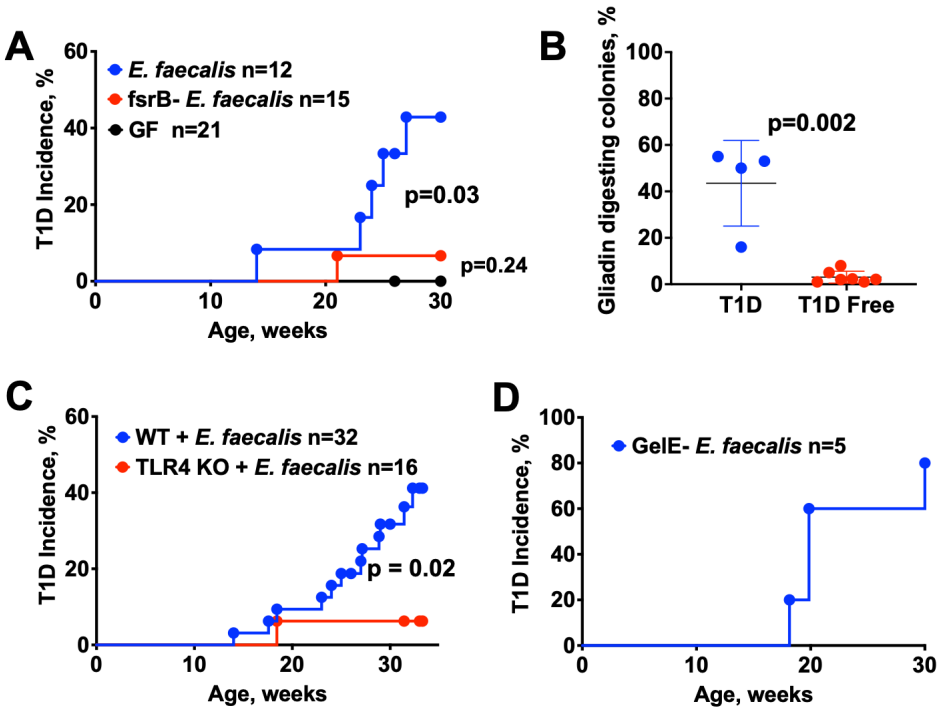

Figure 5.13: Gluten digestion by secreted microbial proteases is key to T1D promotion. **(A)** T1D incidence in GF and gnotobiotic mice colonized with either wild-type *E. faecalis* or Δ*fsrB E. faecalis* incapable of making secreted proteases and fed with HC+4%gluten diet. p values were estimated using Mantel-Cox long-rank test. **(B)** Proportion (%) of gliadin-digesting bacteria in the small intestines of ex-GF mice monocolonized with the wild-type *E. faecalis* bacteria at the time of diabetes development or at 30-week endpoint if mice stayed diabetes-free, mean±sem. p value was calculated using Student's t test. Dots-individual animals. **(C)** Comparison of T1D incidence in gnotobiotic wild-type NOD mice of NOD mice lacking TLR4 colonized with *E. faecalis* and fed with HC+4%gluten diet. p values were estimated using Mantel-Cox long-rank test. **(D)** T1D in gnotobiotic mice colonized with GelE- *E. faecalis* mutant TX5264 and fed HC+4%gluten diet.

5.4 Discussion

While HC protection of NOD mice from T1D was independent from microbial influence, gluten required the microbiota to promote T1D development. Continued protection of GF NOD mice fed HC+4%gluten also confirmed that the enhancement of diabetes seen in SPF NOD mice fed HC+4%gluten was due to the presence of gluten and not to the dilution of HC. Gluten did not rely on microbially activated TGM2. Instead, microbes were needed to proteolytically process gluten releasing biologically active products capable of activating innate and adaptive immunity. This was demonstrated using the gram-positive commensal *E. faecalis* as a model organism. *E. faecalis* could promote T1D in mice fed gluten only when it secreted two proteases, gelE and sprE. Interestingly, we found that the two proteases had two distinct functions. GelE was responsible for liberating peptides from gluten capable of stimulating TCRs reactive to those peptides. SprE was responsible for releasing LPS from gluten which stimulated MyD88 dependent innate immune signaling. Whether, either proteolytic function is sufficient to render gluten diabetogenic is currently under investigation (**Figure 5.13D**, also see Future Directions).

A prerequisite to both molecular mimicry and to bystander activation is the activation of professional APCs capable of presenting self-antigen to T cells. Much of our work therefore focused on whether bacterially digested gluten could stimulate the innate immune system and through which pathways and ligands it did so. LPS was a surprising candidate ligand (we were expecting a peptide); it is a component of the outer membrane of gram-negative bacteria and therefore not produced by wheat or by gram-positive *E. faecalis*. It has also been found to protect from rather than enhance autoimmunity in NOD mice. When LPS is injected intraperitoneally, NOD mice are completely protected from T1D [133] and NOD mice that lack the LPS receptor TLR4 show accelerated rather than reduced onset of disease [142]. One caveat here is that oral delivery of LPS to NOD mice had no effect on T1D development [134]. It is therefore possible that protective signaling pathways triggered through systemic exposure to PAMPs may not be accessible from the gut.

How then could LPS work to enhance autoimmunity? First, we found that it was present to varying degrees in different commercial flour products and gluten digested by *E. faecalis* was sensitive to the LPS inactivating compound polymyxin B. An earlier report also found that the chloroform-methanol soluble fraction of a wheat based diet was pro-diabetic implying that the diabetogenic portion of wheat was lipoidal in nature rather than proteinaceous [97]. This strongly suggested that the innate immune activating component of bacterially digested gluten was LPS. The digest's insensitivity to heat and to trypsin further suggested that its innate stimulatory component was not a protein setting it apart from gliadin derived wheat-amylase trypsin inhibitors that were reported to stimulate TLR4 [228].

Second, the products from bacterially digested gluten appeared to work in synergy with LPS to enhance stimulation of innate immune cytokine secretion. We envisioned two possibilities to explain this: First, the non-LPS products of bacterially digested gluten could have affected LPS aggregation and the binding of LPS monomers to the receptor CD14 on macrophages. It has been previously shown that LPS bound to a retrovirus [224] or to heat shock proteins [222, 223] is a significantly more potent TLR4 agonist than LPS added in solution directly to cells. Second, the ability of bacterially digested gluten to amplify cytokine secretion could be also driven by other microbial PAMPs that work synergistically with LPS. In our study and in work by others [225, 226] macrophage cytokine secretion was greatly enhanced when both TLR2 and TLR4 were simultaneously activated showing a titration curve similar to that achieved with bacterially digested gluten and LPS or Pam3CSK4.

The obvious counterargument to stressing the importance of LPS is that LPS is plentiful in the site where gluten is digested, i.e. the gut, and intestinal macrophages and DCs tend to be tolerant to stimulation by LPS. However, the gut is dominated by penta-acylated LPS [229] which had no synergistic effect with bacterially digested gluten **(Figure 5.8C, D)** and is far less stimulatory than hexa-acylated LPS. Therefore, the LPS derived from gluten may be hexa-acylated, though this point requires further investigation. In addition, macrophages from the PLNs, a site of both gut drainage and immune priming against β-

cell antigens, were not tolerant to stimulation by either hexa-acylated LPS or to bacterially digested gluten **(Figure 5.5A, B)**. Solubilized gluten could potentially drain from the gut to the PLNs to participate in the activation of APCs. While the precise mechanism still needs to be worked out, TLR4 signaling was clearly required for T1D to develop in NOD mice monocolonized with protease secreting *E. faecalis* and fed HC+4%gluten **(Figure 5.13C)**. This demonstrates a clear role for LPS in the microbe dependent promotion of T1D by gluten.

After innate immune activation, the second factor involved in both bystander or molecular mimicry are T cells that react to a foreign antigen yet contribute to the development of autoimmunity. We demonstrated here that *E. faecalis* using protease gelE could liberate peptides capable of stimulating T cells specific to gliadin. While this shows in principle that bacterial proteolytic digestion of gluten can make gluten more accessible to the adaptive immune system, it does not show that gluten reactive T cells contribute to T1D. To address this point, we have generated gluten reactive T cell hybridomas to define peptide specificities and to generate gluten reactive TCR transgenic mouse lines that will enable us to test the role of gluten reactive T cells *in vivo*. These experiments will be discussed in more detail in the next chapter. However, ongoing observations of NOD mice monocolonized with *E. faecalis* lacking the gelE protease and fed HC+4%gluten **(Figure 5.13D)** show high incidence of T1D. This suggests that peptide liberation may not be required for gluten's enhancement of T1D and that LPS liberation instead plays the critical role. We primarily used *E. faecalis* because most of the colonies we identified in the gliadin digestion assay were *E. faecalis* and because it was a species with tractable genetics. Other bacteria that formed halos on gliadin containing agar could also digest gluten to stimulate TLR4 dependent TNF secretion **(Figure 5.10)**. Proteases from other bacteria (e.g. pseudolysin from *Pseudomonas aeruginosa*) have also been reported to be capable of breaking down gluten [230]. Therefore, bacteria other than *E. faecalis* may be capable of eliciting the same phenotype, T1D, via the same biological activity, proteolytic digestion of gluten. This is important as different mouse

colonies and human populations carry different communities of microbes which may carry many different genes that encode proteases with similar functions. The relative absence of these microbes, or the presence of microbes capable of inactivating gluten digesting proteases, may explain the disparate effects of gluten on NOD mice reported in the literature.

Together, these results indicate that microbial digestion can make an otherwise innocuous dietary component (gluten) diabetogenic by modifying its ability to stimulate the innate immune system, revealing a novel role for LPS in the promotion of T1D. Whether the microbiota can modify other dietary proteins to promote autoimmunity is an open question that requires further study.

CHAPTER 6

FUTURE DIRECTIONS

Our use of specific diabetes protective and diabetes promoting diets was merely a starting point that allowed us to link previously unappreciated immunological and non-immunological host pathways to organ specific autoimmunity. Further study of these pathways will be key to understanding how environmental factors can influence autoimmunity and how to prevent these diseases from manifesting in the genetically at risk. This chapter will address the questions raised by this study and will detail the next steps that must be tackled to address them.

6.1 Which gene controls high insulin secretion in NOD mice?

Islets from NOD mice tend to secrete more insulin following glucose challenge. Importantly this trait persists independent of adaptive immunity and is therefore not caused by immune-mediated damage to the islets [188]. We showed in Chapter 3, using F1 crosses of NOD and B6g7 mice carrying the scid mutation, that high glucose-tolerance in NOD mice is a dominant trait not linked to a mitochondrial gene **(Figure3.2A, B)**. Using N2 mice (F1 x B6g7), we found that elevated glucose tolerance in NOD mice was likely controlled by a single gene **(Figure3.2B)**. Efforts are currently being undertaken in the lab to confirm that this trait is controlled by a single locus and to identify the gene. Low-glucose tolerant N2 mice were recently crossed to NOD.scid mice to produce N3 mice. If high-glucose tolerance is controlled by a single gene, 50% of N3 mice should have high-glucose tolerance and 50% should have low glucose tolerance, mirroring the N2 results. In parallel, we plan to perform a Mouse Universal Genotyping Array (MUGA), a high performance genetic quality-check platform with more than 11,000 probes [231] on genetic samples from N2 mice previously screened for glucose tolerance by IPGTT. This platform will allow us to find genetic differences in our N2 animals that correlate with high or low glucose tolerance facilitating the

identification of candidate genes. Once such genes are identified, B6 versions of the genes will be introduced into NOD mice using CRISPR-Cas9 technology to determine whether the mutations reduce glucose tolerance in NOD mice and, more importantly, whether this reduction in glucose tolerance and insulin secretion affect T1D development.

6.2 How does diet influence β-cell stress?

We demonstrated in Chapter 3 that feeding NOD mice a diet containing HC as the sole amino acid source prevented T1D in a microbe independent manner by reducing β-cell insulin secretion and associated ER stress. How exactly diet leads to reductions in insulin secretion is currently unclear and understanding these upstream pathways will elucidate which environmentally sensitive host signals influence β-cell physiology.

Are β-cells influenced directly by the changes in macronutrient source or is another cell type or even another organ the primary target? Distinguishing between these possibilities will not be a simple task as each possibility will take considerable high throughput screening to test and will require substantial *in vivo* validation. Direct effects on β-cells could be identified by searching for ligands present in HC and receptors in β-cells that can reduce glucose stimulated insulin secretion. A number of different bioactive products derived from milk proteins have been previously described [100] and could be tested on *in vitro* cultured islets for influences on insulin secretion. This system has the added benefit of testing whole islets so if an islet cell type other than β-cells is the primary target, it can be isolated and studied further. If any ligands are identified, biochemical techniques such as co-immunoprecipitation could then be used to identify the receptor and test for effects of the ligand *in vivo*. If the influence is on other organs, a combined transcriptomic and targeted approach can be applied. Measuring gene expression changes in multiple organs in mice fed chow or HC diets could reveal pathways strongly affected by HC that could then be knocked down to determine if they influence β-cell physiology. In addition, measuring the levels of non-β-cell factors in mice fed different diets known to influence insulin secretion could indicate which organs and

pathways connect products from HC to the islets. Potential candidates include the incretins produced by the gut, regenerating family proteins primarily produced by acinar cells and cells of the small intestine (though some members such as Reg2 can also be produced by β-cells), and serotonin produced by neurons within the islets [232–234].

In our study, mice were exposed to HC for their entire lives, from in utero until death. It is therefore possible that HC affected the development of the islets. While we identified reductions in the frequency of high-insulin secreting β-cells, reductions in insulin secretion could also be accomplished by a reduction in the number of β-cells themselves. A reduction in β-cell numbers could also theoretically reduce the activation of autoimmunity due to the decreased concentration of β-cell antigens. Thorough histological examinations of pancreata from mice on different diets could be carried out to examine β-cell frequencies within the islets and count the average number of islets present to estimate the average number of β-cells. It is also possible that rather than affecting overall cell numbers, HC exposure in utero epigenetically programs β-cells for lower insulin secretion by reducing the frequency of β-cells producing the most insulin. We observed a biphasic distribution in *Ins1* and *Ins2* transcription in β-cells from chow-fed mice that was lost in HC-fed mice **(Figure3.5A, B)** that could potentially be explained by epigenetic changes. A genome wide examination of chromatin accessibility and histone modifications coupled with transcriptomic data could reveal pathways that are induced or repressed by HC (or by the lack of other components found in chow).

6.3 How does β-cell stress provoke autoimmunity?

In our study, reduced β-cell stress and insulin secretion was associated with a reduction in diabetes incidence and the activation of autoreactive T cells in the PLNs. How exactly this β-cell stress leads to increased autoimmunity is an open question. Stressed β-cells may be more sensitive and die more easily during infiltration and attack by autoreactive T cells. A similar hypothesis is that, because stressed β-cells produce more insulin which is a potent

autoantigen [187], they are preferentially targeted by insulin-reactive T cells. In either case, β-cells producing more insulin would need to die more readily under conditions of islet infiltration. To test whether high insulin secreting β-cells are preferentially targeted by the immune system for killing requires a mosaic transgenic system that labels β-cells with high insulin secretion. This could potentially be accomplished by transgenically expressing insulin and GFP under the rat insulin promoter to ensure expression only in β-cells and targeted to the X chromosome. Random X inactivation in female mice has been previously shown to inactivate x-linked transgenes in adult tissues [235] and should theoretically silence extra insulin expression in half of the β-cells. If produced in NOD mice, ratios of GFP labeled to non-GFP labeled cells should be 1:1 in young mice; but if high insulin producing (GFP^+) cells are preferentially killed, the ratio should decrease.

Alternatively, stressed β-cells may simply accelerate T cell priming in the PLNs due to their increased levels of insulin secretion. To test whether high insulin producing β-cells more readily activate adaptive immunity, a simpler *in vitro* system can be used. β-cells expressing GFP under the mouse insulin promoter have correlated insulin expression and GFP expression [236]. Sorting GFP high and GFP low β-cells from a NOD mouse and incubating them in culture with APCs and islet reactive, transgenic T cells (G9C8, 8.3, or BDC2.5) labeled with a proliferation tracker dye, would allow you to track T cell reactivity to β-cells with different insulin expression levels. T cells should proliferate much better in wells with high insulin producing β-cells if they serve as a better source of antigen.

6.4 Which microbial proteases are required for gluten dependent promotion of T1D?

In chapter 4 we showed that replacing 1/5th of the casein in the HC diet with gluten was able to restore high T1D incidence to NOD mice. The new HC+4%gluten diet did not reverse any of the protective effects on β-cell physiology but rather enhanced autoimmunity

by activating multiple immune cell types. In chapter 5 we showed that this immune activation and elevated T1D incidence relied on the proteolytic digestion of gluten by the microbiota. The biologically active products released by proteolysis included LPS which activated innate immunity via TLR4 and peptides capable of activating gliadin reactive T cells specific to them. The LPS released from microbially-digested gluten promoted T1D development as gnotobiotic NOD mice lacking TLR4 and monocolonized with *E. faecalis* were resistant to T1D when fed HC+4%gluten diet.

The *E. faecalis* protease sprE was responsible for the release of LPS while gelE was responsible for the release of T cell activating peptides. Whether proteolytic cleavage of gluten by either protease is sufficient to promote T1D is an open question. To test this, we plan to monocolonize NOD and NOD.TLR4KO mice with *E. faecalis* mutants lacking either sprE (TX5243 [157]) or gelE (TX6254 [158]). As each protease releases different products that stimulate different arms of the immune system from gluten, these experiments will have the added benefit of telling us which arm must be activated by bacterially digested gluten to promote disease. NOD mice monocolonized with gelE- sprE+ *E. faecalis* **(Figure 5.13D)** have quite high diabetes incidence in preliminary results suggesting that the main diabetogenic function of bacterial digestion is the release of LPS. This is well supported by the protection from T1D shown by monocolonized TLR4 KO mice fed HC+4%gluten **(Figure 5.13C)**. More work needs to be done to fully confirm this conclusion.

6.5 Do gluten reactive T cells contribute to gluten's enhancement of autoimmunity?

While bacterial digestion was capable of releasing peptides capable of stimulating gliadin-reactive T cells when presented on APCs, we do not have direct experimental evidence that gluten-reactive T cells contribute to autoimmunity. We envision 3 ways that gluten-reactive T cells could participate in the autoimmune process: 1. Some gluten-reactive T cell

could cross-react to islet antigens and, following activation by gluten peptides, proceed to destroy the pancreatic islets. 2. Gluten-reactive T cells could, following activation, produce a generally inflammatory environment that lowers the activation threshold for *bona fide* autoreactive cells. 3. Gluten-reactive T cells play no role at all in the autoimmune process. Any experiments to test these hypotheses require that we have T cells with TCRs known to react to gluten, so as discussed earlier, we have generated gluten-reactive T cell hybridomas **(Figure 5.12A, B)**. The immediate goal is to identify peptide reactivities for selected hybrids in order to generate MHC II tetramers so that gluten-reactive T cells can be traced *in vivo* and to sequence TCRs for the generation of TCR-transgenic mice. The hybridomas are also being currently screened *in vitro* for cross-reactivity to islet antigens using a hybridoma generated with BDC2.5 T cells serving as a control. No cross-reactive hybrids have so far been identified.

If gluten-reactive T cells promote T1D development, they should be able to accelerate T1D progression only in the presence of gluten unless they are also cross reactive to islet antigens. To test this, transgenic mice that carry gluten reactive TCRs will be placed on the HC diet, which lacks gluten. Gluten-reactive T cells from these mice will then be mixed with splenocytes from HC-fed mice and transferred into HC-fed NOD.SCID recipients. As controls, HC-fed NOD.SCID mice will be injected with splenocytes from HC-fed NOD mice without additional gluten-reactive T cells. Half of the NOD.SCID recipients will then be switched to HC+4%gluten diet while the other half will remain on HC diet. All mice will then be monitored for diabetes development. If gluten-reactive T cells promote T1D by creating an inflammatory environment that fosters the expansion of islet-reactive T cells, then time to T1D should be shortened in recipients that received gluten-reactive T cells and were switched to HC+4%gluten diet. If gluten-reactive T cells promote T1D via cross reactivity, then time to T1D should be shortened in any recipients receiving gluten-reactive T cells. Additionally, cross-clones should proliferate in the PLNs of mice fed HC due to exposure to islet antigens. If gluten-reactive T cells do not promote T1D development, then

NOD.SCID mice receiving gluten-reactive T cells should not progress to T1D faster than NOD.SCID mice receiving splenocytes alone. It should be noted that if gluten-reactive T cells do contribute to T1D development, they probably play a secondary role to the activation of innate immunity by gluten. This intuitively makes sense as the activation of professional APCs is a prerequisite to activation of adaptive immunity.

REFERENCES

[1] Ignatavicius, D. & Workman, M. Chapter 3 assessment and care of patients with pain. medical-surgical nursing: patient-centered collaborative care. st louis (mo) (2016).

[2] Lewis, S. L., Bucher, L., Heitkemper, M. M. & Dirksen, S. R. *Clinical Companion to Medical-Surgical Nursing-E-Book* (Elsevier Health Sciences, 2014).

[3] Kawahito, S., Kitahata, H. & Oshita, S. Problems associated with glucose toxicity: role of hyperglycemia-induced oxidative stress. *World journal of gastroenterology: WJG* **15**, 4137 (2009).

[4] Atlas, D. *et al.* International diabetes federation. *IDF Diabetes Atlas, 7th edn. Brussels, Belgium: International Diabetes Federation* (2015).

[5] Sieving, P. A. Diabetic retinopathy: past, present and future. *Nova Acta Leopoldina NF* **119**, 21–28 (2014).

[6] Poretsky, L. *Principles of diabetes mellitus*, vol. 21 (Springer, 2010).

[7] Ahmed, A. M. History of diabetes mellitus. *Saudi medical journal* **23**, 373–378 (2002).

[8] Alam, S. *et al.* The history of diabetes: from olden days to discovering insulin. *Internafional Journal of Unani and Integrafive Medicine* **1**, 25–28 (2017).

[9] Von Mering, J. & Minkowski, O. Diabetes mellitus after pancreas extirpation. *Arch Exp Pathol Pharmakol* **26**, 111 (1889).

[10] Hedon, E. Greffe sous-cutanée du pancréas. *Archs Physiol Norm Pathol* **5**, 617–28 (1892).

[11] Minkowski, O. Weitere mittheilungen über den diabetes mellitus nach exstirpation des pankreas. *Berl Klin Wochenschr* **29**, 90–94 (1892).

[12] Langerhans, P. Beitrage zur mikroskopischen anatomie der bauchspeicheldruse, inaugural disseration. *Gustav Lange* (1869).

[13] Gepts, W. Islet changes in human diabetes. In *The Islets of Langerhans: Biochemistry, Physiology, and Pathology*, 321–356 (Elsevier, 1981).

[14] Opie, E. L. The relation oe diabetes mellitus to lesions of the pancreas. hyaline degeneration of the islands oe langerhans. *The Journal of experimental medicine* **5**, 527 (1901).

[15] De Mayer, J. Action de la secretion interne du pancreas sur differents organes et en particulier sur la secretion renals. *Arch di Fisiol* **7**, 96 (1909).

[16] Sharpey-Schäfer, E. A. *An introduction to the Study of the endocrine glands and internal secretions: lane medical lectures, 1913.* 18 (Stanford University Press, 1914).

[17] Vecchio, I., Tornali, C., Bragazzi, N. L. & Martini, M. The discovery of insulin: an important milestone in the history of medicine. *Frontiers in endocrinology* 613 (2018).

[18] Zuelzer, G. *et al.* Experimentelle untersuchungen über den diabetes. *Berlin Klin Wochenschr* **44**, 474–475 (1907).

[19] Scott, E. L. On the influence of intravenous injections of an extract of the pancreas on experimental pancreatic diabetes. *American Journal of Physiology-Legacy Content* **29**, 306–310 (1912).

[20] Forschibach, J. Versuche zur behandlung des diabetes mellitus mit dem zuelzerschen pankreashormon. *DMW-Deutsche Medizinische Wochenschrift* **35**, 2053–2055 (1909).

[21] Pratt, J. H. A reappraisal of researches leading to the discovery of insulin. *Journal of the History of Medicine and Allied Sciences* 281–289 (1954).

[22] Banting, F. G., Best, C. H., Collip, J. B., Campbell, W. R. & Fletcher, A. A. Pancreatic extracts in the treatment of diabetes mellitus. *Canadian Medical Association Journal* **12**, 141 (1922).

[23] FALTA, W. & BOLLER, R. Insulprer und insulinresistenter diabetes. klin. *Wchnrchr, 10* **438** (1931).

[24] Himsworth, H. P. *et al.* Diabetes mellitus. its differentiation into insolin-sensitive and insulin-insensitive types. *Lancet* **230**, 127–130 (1936).

[25] Wrenshall, G. A., Bogoch, A. & Ritchie, R. Extractable insulin of pancreas: correlation with pathological and clinical findings in diabetic and nondiabetic cases. *Diabetes* **1**, 87–107 (1952).

[26] Lister, J., Nash, J. & Ledingham, U. Constitution and insulin sensitivity in diabetes mellitus. *British medical journal* **1**, 376 (1951).

[27] Roitt, I., Doniach, D., Campbell, P. & Hudson, R. V. Auto-antibodies in hashimoto's disease (lymphadenoid goitre). *The Lancet* **268**, 820–821 (1956).

[28] Anderson, J., Goudie, R., Gray, K. & Timbury, G. Auto-antibodies in addison's disease. *The Lancet* **269**, 1123–1124 (1957).

[29] Witebsky, E., Rose, N. R., Terplan, K., Paine, J. R. & Egan, R. W. Chronic thyroiditis and autoimmunization. *Journal of the American Medical Association* **164**, 1439–1447 (1957).

[30] VOV MEYENBURG, H. Uber" insulitis" bei diabetes. *Schweiz Med Wochenschr* **21**, 554–557 (1940).

[31] LeCompte, P. M. Insulitis in early juvenile diabetes. *AMA archives of pathology* **66**, 450–457 (1958).

[32] Renold, A. E., Soeldner, S. & Steinke, J. Immunological studies with homologous and heterologous pancreatic insulin in the cow. In *Ciba Foundation Symposium-Aetiology of Diabetes Mellitus and its Complications (Colloquia on Endocrinology)*, vol. 15, 122–139 (Wiley Online Library, 1964).

[33] Gepts, W. Pathologic anatomy of the pancreas in juvenile diabetes mellitus. *Diabetes* **14**, 619–633 (1965).

[34] Nerup, J. *et al.* Hl-a antigens and diabetes mellitus. *The Lancet* **304**, 864–866 (1974).

[35] Singal, D. & Blajchman, M. Histocompatibility (hl-a) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. *Diabetes* **22**, 429–432 (1973).

[36] Bottazzo, G., Florin-Christensen, A. & Doniach, D. Islet-cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies. *The Lancet* **304**, 1279–1283 (1974).

[37] MacCuish, A., Irvine, W., Barnes, E. & Duncan, L. Antibodies to pancreatic islet cells in insulin-dependent diabetics with coexistent autoimmune disease. *The Lancet* **304**, 1529–1531 (1974).

[38] Lendrum, R. *et al.* Islet-cell antibodies in diabetes mellitus. *The Lancet* **308**, 1273–1276 (1976).

[39] Nakhooda, A., Like, A., Chappel, C., Murray, F. & Marliss, E. The spontaneously diabetic wistar rat: metabolic and morphologic studies. *Diabetes* **26**, 100–112 (1977).

[40] Makino, S. *et al.* Breeding of a non-obese, diabetic strain of mice. *Experimental Animals* **29**, 1–13 (1980).

[41] Knip, M. Natural course of preclinical type 1 diabetes. *Hormone Research in Paediatrics* **57**, 6–11 (2002).

[42] Pozzilli, P., Signore, A., Williams, A. J. & Beales, P. E. Nod mouse colonies around the world-recent facts and figures. *Immunology today* **14**, 193–196 (1993).

[43] Charre, S. *et al.* Abnormalities in dendritic cell and macrophage accumulation in the pancreas of nonobese diabetic (nod) mice during the early neonatal period. *Histology and histopathology* (2002).

[44] Turley, S., Poirot, L., Hattori, M., Benoist, C. & Mathis, D. Physiological β cell death triggers priming of self-reactive t cells by dendritic cells in a type-1 diabetes model. *The Journal of experimental medicine* **198**, 1527–1537 (2003).

[45] Zakharov, P. N., Hu, H., Wan, X. & Unanue, E. R. Single-cell rna sequencing of murine islets shows high cellular complexity at all stages of autoimmune diabetes. *Journal of Experimental Medicine* **217** (2020).

[46] Füchtenbusch, M., Larger, E., Thebault, K. & Boitard, C. Transfer of diabetes from prediabetic nod mice to nod-scid/scid mice: association with pancreatic insulin content. *Hormone and metabolic research* **37**, 63–67 (2005).

[47] Bendelac, A., Carnaud, C., Boitard, C. & Bach, J. Syngeneic transfer of autoimmune diabetes from diabetic nod mice to healthy neonates. requirement for both l3t4+ and lyt-2+ t cells. *The Journal of experimental medicine* **166**, 823–832 (1987).

[48] Christianson, S. W., Shultz, L. D. & Leiter, E. H. Adoptive transfer of diabetes into immunodeficient nod-scid/scid mice: relative contributions of cd4+ and cd8+ t-cells from diabetic versus prediabetic nod. non-thy-1a donors. *Diabetes* **42**, 44–55 (1993).

[49] Bluestone, J. A., Herold, K. & Eisenbarth, G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* **464**, 1293–1300 (2010).

[50] Wong, F. S. *et al.* Activation of insulin-reactive cd8 t-cells for development of autoimmune diabetes. *Diabetes* **58**, 1156–1164 (2009).

[51] Katz, J. D., Wang, B., Haskins, K., Benoist, C. & Mathis, D. Following a diabetogenic t cell from genesis through pathogenesis. *Cell* **74**, 1089–1100 (1993).

[52] Nagata, M., Santamaria, P., Kawamura, T., Utsugi, T. & Yoon, J.-W. Evidence for the role of cd8+ cytotoxic t cells in the destruction of pancreatic beta-cells in nonobese diabetic mice. *The Journal of Immunology* **152**, 2042–2050 (1994).

[53] Akashi, T. *et al.* Direct evidence for the contribution of b cells to the progression of insulitis and the development of diabetes in non-obese diabetic mice. *International immunology* **9**, 1159–1164 (1997).

[54] Serreze, D. V. *et al.* B lymphocytes are essential for the initiation of t cell-mediated autoimmune diabetes: analysis of a new" speed congenic" stock of nod. ig mu null mice. *The Journal of experimental medicine* **184**, 2049–2053 (1996).

[55] Wong, F. S. *et al.* Investigation of the role of b-cells in type 1 diabetes in the nod mouse. *Diabetes* **53**, 2581–2587 (2004).

[56] Noorchashm, H. *et al.* I-ag7-mediated antigen presentation by b lymphocytes is critical in overcoming a checkpoint in t cell tolerance to islet β cells of nonobese diabetic mice. *The Journal of Immunology* **163**, 743–750 (1999).

[57] Batista, F. D. & Neuberger, M. S. Affinity dependence of the b cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* **8**, 751–759 (1998).

[58] Ogata, H. *et al.* The toll-like receptor protein rp105 regulates lipopolysaccharide signaling in b cells. *The Journal of experimental medicine* **192**, 23–30 (2000).

[59] Victora, G. D. *et al.* Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* **143**, 592–605 (2010).

[60] Varanasi, V., Avanesyan, L., Schumann, D. M. & Chervonsky, A. V. Cytotoxic mechanisms employed by mouse t cells to destroy pancreatic β-cells. *Diabetes* **61**, 2862–2870 (2012).

[61] Rui, J. *et al.* β cells that resist immunological attack develop during progression of autoimmune diabetes in nod mice. *Cell metabolism* **25**, 727–738 (2017).

[62] Mrena, S. *et al.* Models for predicting type 1 diabetes in siblings of affected children. *Diabetes care* **29**, 662–667 (2006).

[63] Barnett, A., Eff, C., Leslie, R. & Pyke, D. Diabetes in identical twins. *Diabetologia* **20**, 87–93 (1981).

[64] Cudworth, A. & Woodrow, J. Hl-a antigens and diabetes mellitus. *The Lancet* **304**, 1153 (1974).

[65] Yui, M. *et al.* Production of congenic mouse strains carrying nod-derived diabetogenic genetic intervals: an approach for the genetic dissection of complex traits. *Mammalian Genome* **7**, 331–334 (1996).

[66] Redondo, M. J., Steck, A. K. & Pugliese, A. Genetics of type 1 diabetes. *Pediatric diabetes* **19**, 346–353 (2018).

[67] Pugliese, A. *et al.* The insulin gene is transcribed in the human thymus and transcription levels correlate with allelic variation at the ins vntr-iddm2 susceptibility locus for type 1 diabetes. *Nature genetics* **15**, 293–297 (1997).

[68] Dooley, J. *et al.* Genetic predisposition for beta cell fragility underlies type 1 and type 2 diabetes. *Nature genetics* **48**, 519–527 (2016).

[69] Sollid, L. M. & Jabri, B. Triggers and drivers of autoimmunity: lessons from coeliac disease. *Nature Reviews Immunology* **13**, 294–302 (2013).

[70] Craig, M. E. *et al.* Prevalence of celiac disease in 52,721 youth with type 1 diabetes: international comparison across three continents. *Diabetes care* **40**, 1034–1040 (2017).

[71] Kahaly, G. J., Frommer, L. & Schuppan, D. Celiac disease and endocrine autoimmunity–the genetic link. *Autoimmunity reviews* **17**, 1169–1175 (2018).

[72] Norris, J. M., Johnson, R. K. & Stene, L. C. Type 1 diabetes—early life origins and changing epidemiology. *The lancet Diabetes & endocrinology* **8**, 226–238 (2020).

[73] Atkinson, M. A., Eisenbarth, G. S. & Michels, A. W. Type 1 diabetes. *The Lancet* **383**, 69–82 (2014).

[74] Kondrashova, A. *et al.* A six-fold gradient in the incidence of type 1 diabetes at the eastern border of finland. *Annals of medicine* **37**, 67–72 (2005).

[75] Bach, J.-F. & Chatenoud, L. The hygiene hypothesis: an explanation for the increased frequency of insulin-dependent diabetes. *Cold Spring Harbor perspectives in medicine* **2**, a007799 (2012).

[76] Airaghi, L. & Tedeschi, A. Negative association between occurrence of type 1 diabetes and tuberculosis incidence at population level. *Acta Diabetologica* **43**, 43–45 (2006).

[77] Lefebvre, D. E., Powell, K. L., Strom, A. & Scott, F. W. Dietary proteins as environmental modifiers of type 1 diabetes mellitus. *Annual review of nutrition* **26**, 175–202 (2006).

[78] Honda, K. & Littman, D. R. The microbiome in infectious disease and inflammation. *Annual review of immunology* **30**, 759 (2012).

[79] Chervonsky, A. V. Microbiota and autoimmunity. *Cold Spring Harbor perspectives in biology* **5**, a007294 (2013).

[80] Vatanen, T. *et al.* The human gut microbiome in early-onset type 1 diabetes from the teddy study. *Nature* **562**, 589–594 (2018).

[81] David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).

[82] Faith, J. J., McNulty, N. P., Rey, F. E. & Gordon, J. I. Predicting a human gut microbiota's response to diet in gnotobiotic mice. *Science* **333**, 101–104 (2011).

[83] Giulietti, A. *et al.* Vitamin d deficiency in early life accelerates type 1 diabetes in non-obese diabetic mice. *Diabetologia* **47**, 451–462 (2004).

[84] Zella, J. B., McCary, L. C. & DeLuca, H. F. Oral administration of 1, 25-dihydroxyvitamin d3 completely protects nod mice from insulin-dependent diabetes mellitus. *Archives of biochemistry and biophysics* **417**, 77–80 (2003).

[85] Schmid, S. *et al.* Delayed exposure to wheat and barley proteins reduces diabetes incidence in non-obese diabetic mice. *Clinical Immunology* **111**, 108–118 (2004).

[86] Gysemans, C. *et al.* Unaltered diabetes presentation in nod mice lacking the vitamin d receptor. *Diabetes* **57**, 269–275 (2008).

[87] Hyppönen, E., Läärä, E., Reunanen, A., Järvelin, M.-R. & Virtanen, S. M. Intake of vitamin d and risk of type 1 diabetes: a birth-cohort study. *The Lancet* **358**, 1500–1503 (2001).

[88] Lamb, M. M. *et al.* Dietary glycemic index, development of islet autoimmunity, and subsequent progression to type 1 diabetes in young children. *The Journal of Clinical Endocrinology & Metabolism* **93**, 3936–3942 (2008).

[89] Lamb, M. M. *et al.* Sugar intake is associated with progression from islet autoimmunity to type 1 diabetes: the diabetes autoimmunity study in the young. *Diabetologia* **58**, 2027–2034 (2015).

[90] Chen, K. *et al.* Specific inulin-type fructan fibers protect against autoimmune diabetes by modulating gut immunity, barrier function, and microbiota homeostasis. *Molecular nutrition & food research* **61**, 1601006 (2017).

[91] Stillie, R., Bell, R. C. & Field, C. J. Diabetes-prone biobreeding rats do not have a normal immune response when weaned to a diet containing fermentable fibre. *British journal of nutrition* **93**, 645–653 (2005).

[92] Hakola, L. *et al.* Association of cereal, gluten, and dietary fiber intake with islet autoimmunity and type 1 diabetes. *JAMA pediatrics* **173**, 953–960 (2019).

[93] Thorburn, A. N., Macia, L. & Mackay, C. R. Diet, metabolites, and "western-lifestyle" inflammatory diseases. *Immunity* **40**, 833–842 (2014).

[94] Mariño, E. *et al.* Gut microbial metabolites limit the frequency of autoimmune t cells and protect against type 1 diabetes. *Nature immunology* **18**, 552–562 (2017).

[95] Hoorfar, J., Buschard, K. & Brogren, C. Impact of dietary protein and fat source on the development of insulin-dependent diabetes in the bb rat. *Diabetes Research (Edinburgh, Scotland)* **20**, 33–41 (1992).

[96] Beales, P. *et al.* A multi-centre, blinded international trial of the effect of a1 and a2 β-casein variants on diabetes incidence in two rodent models of spontaneous type i diabetes. *Diabetologia* **45**, 1240–1246 (2002).

[97] Coleman, D. L., Kuzava, J. E. & Leiter, E. H. Effect of diet on incidence of diabetes in nonobese diabetic mice. *Diabetes* **39**, 432–436 (1990).

[98] Elliott, R., Reddy, S., Bibby, N. & Kida, K. Dietary prevention of diabetes in the non-obese diabetic mouse. *Diabetologia* **31**, 62–64 (1988).

[99] Patrick, C. *et al.* Promotion of autoimmune diabetes by cereal diet in the presence or absence of microbes associated with gut immune activation, regulatory imbalance, and altered cathelicidin antimicrobial peptide. *Diabetes* **62**, 2036–2047 (2013).

[100] Meisel, H. Biochemical properties of peptides encrypted in bovine milk proteins. *Current medicinal chemistry* **12**, 1905–1919 (2005).

[101] Hira, T., Hara, H. & Aoyama, Y. Stimulative effect of a casein hydrolysate on exocrine pancreatic secretion that is independent of luminal trypsin inhibitory activity in rats. *Bioscience, biotechnology, and biochemistry* **63**, 1192–1196 (1999).

[102] Claustre, J. *et al.* Effects of peptides derived from dietary proteins on mucus secretion in rat jejunum. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **283**, G521–G528 (2002).

[103] Courtois, P. *et al.* Gut permeability and intestinal mucins, invertase, and peroxidase in control and diabetes-prone bb rats fed either a protective or a diabetogenic diet. *Digestive diseases and sciences* **50**, 266–275 (2005).

[104] MacMurray, A. J. *et al.* Lymphopenia in the bb rat model of type 1 diabetes is due to a mutation in a novel immune-associated nucleotide (ian)-related gene. *Genome research* **12**, 1029–1039 (2002).

[105] Funda, D. P., Kaas, A., Bock, T., Tlaskalová-Hogenová, H. & Buschard, K. Gluten-free diet prevents diabetes in nod mice. *Diabetes/metabolism research and reviews* **15**, 323–327 (1999).

[106] Maurano, F. *et al.* Small intestinal enteropathy in non-obese diabetic mice fed a diet containing wheat. *Diabetologia* **48**, 931–937 (2005).

[107] Funda, D. P., Kaas, A., Tlaskalová-Hogenová, H. & Buschard, K. Gluten-free but also gluten-enriched (gluten+) diet prevent diabetes in nod mice; the gluten enigma in type 1 diabetes. *Diabetes/metabolism research and reviews* **24**, 59–63 (2008).

[108] Lund-Blix, N. A. *et al.* Maternal and child gluten intake and association with type 1 diabetes: The norwegian mother and child cohort study. *PLoS medicine* **17**, e1003032 (2020).

[109] Lund-Blix, N. A. *et al.* Gluten intake and risk of islet autoimmunity and progression to type 1 diabetes in children at increased risk of the disease: the diabetes autoimmunity study in the young (daisy). *Diabetes Care* **42**, 789–796 (2019).

[110] Elliott, R. & Martin, J. Dietary protein: a trigger of insulin-dependent diabetes in the bb rat? *Diabetologia* **26**, 297–299 (1984).

[111] Virtanen, S. M. *et al.* Infant feeding in finnish children< 7 yr of age with newly diagnosed iddm. *Diabetes care* **14**, 415–417 (1991).

[112] Scott, F. W. Food-induced type 1 diabetes in the bb rat. *Diabetes Metabolism Reviews-Clinical and Scientific Advances in Diabetes Research* **12**, 341–360 (1996).

[113] Knip, M. *et al.* Effect of hydrolyzed infant formula vs conventional formula on risk of type 1 diabetes: the trigr randomized clinical trial. *Jama* **319**, 38–48 (2018).

[114] Knip, M. & Simell, O. Environmental triggers of type 1 diabetes. *Cold Spring Harbor perspectives in medicine* **2**, a007690 (2012).

[115] Couper, J. Environmental triggers of type 1 diabetes. *Journal of paediatrics and child health* **37**, 218–220 (2001).

[116] Barbeau, W. E. What is the key environmental trigger in type 1 diabetes—is it viruses, or wheat gluten, or both? *Autoimmunity Reviews* **12**, 295–299 (2012).

[117] Adams, S. F. The seasonal variation in the onset of acute diabetes: the age and sex factors in 1,000 diabetic patients. *Archives of Internal Medicine* **37**, 861–864 (1926).

[118] Coppieters, K. T., Boettler, T. & von Herrath, M. Virus infections in type 1 diabetes. *Cold Spring Harbor perspectives in medicine* **2**, a007682 (2012).

[119] Gale, E. Congenital rubella: citation virus or viral cause of type 1 diabetes? (2008).

[120] Honeyman, M. How robust is the evidence for viruses in the induction of type 1 diabetes? *Current opinion in immunology* **17**, 616–623 (2005).

[121] Gamble, D., Kinsley, M., FitzGerald, M., Bolton, R. & Taylor, K. Viral antibodies in diabetes mellitus. *Br Med J* **3**, 627–630 (1969).

[122] Clements, G. B., Galbraith, D. & Taylor, K. Coxsackie b virus infection and onset of childhood diabetes. *Lancet (London, England)* **346**, 221–223 (1995).

[123] Vehik, K. *et al.* Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. *Nature medicine* **25**, 1865–1872 (2019).

[124] Coleman, T., Taylor, K. & Gamble, D. The development of diabetes following coxsackie b virus infection in mice. *Diabetologia* **10**, 755–759 (1974).

[125] Dippe, S., Miller, M., Bennett, P., Maynard, J. & Berquist, K. Lack of causal association between coxsackie b4 virus infection and diabetes. *The Lancet* **305**, 1314–1317 (1975).

[126] Serreze, D. V., Ottendorfer, E. W., Ellis, T. M., Gauntt, C. J. & Atkinson, M. A. Acceleration of type 1 diabetes by a coxsackievirus infection requires a preexisting critical mass of autoreactive t-cells in pancreatic islets. *Diabetes* **49**, 708–711 (2000).

[127] Tracy, S. *et al.* Toward testing the hypothesis that group b coxsackieviruses (cvb) trigger insulin-dependent diabetes: inoculating nonobese diabetic mice with cvb markedly lowers diabetes incidence. *Journal of virology* **76**, 12097–12111 (2002).

[128] Filippi, C. M., Estes, E. A., Oldham, J. E., von Herrath, M. G. *et al.* Immunoregulatory mechanisms triggered by viral infections protect from type 1 diabetes in mice. *The Journal of clinical investigation* **119**, 1515–1523 (2009).

[129] Bach, J.-F. The hygiene hypothesis in autoimmunity: the role of pathogens and commensals. *Nature Reviews Immunology* **18**, 105–120 (2018).

[130] Bach, J.-F. The effect of infections on susceptibility to autoimmune and allergic diseases. *New England journal of medicine* **347**, 911–920 (2002).

[131] Alyanakian, M.-A. *et al.* Transforming growth factor-β and natural killer t-cells are involved in the protective effect of a bacterial extract on type 1 diabetes. *Diabetes* **55**, 179–185 (2006).

[132] Mori, Y., Kodaka, T., Kato, T., Kanagawa, E. M. & Kanagawa, O. Critical role of ifn-γ in cfa-mediated protection of nod mice from diabetes development. *International immunology* **21**, 1291–1299 (2009).

[133] Aumeunier, A. *et al.* Systemic toll-like receptor stimulation suppresses experimental allergic asthma and autoimmune diabetes in nod mice. *PloS one* **5**, e11484 (2010).

[134] Kihl, P. *et al.* Oral lps dosing induces local immunological changes in the pancreatic lymph nodes in mice. *Journal of Diabetes Research* **2019** (2019).

[135] Rowland, I. *et al.* Gut microbiota functions: metabolism of nutrients and other food components. *European journal of nutrition* **57**, 1–24 (2018).

[136] Gordon, H. A. & Pesti, L. The gnotobiotic animal as a tool in the study of host microbial relationships. *Bacteriological reviews* **35**, 390–429 (1971).

[137] Stecher, B. & Hardt, W.-D. The role of microbiota in infectious disease. *Trends in microbiology* **16**, 107–114 (2008).

[138] Shin, S. C. *et al.* Drosophila microbiome modulates host developmental and metabolic homeostasis via insulin signaling. *Science* **334**, 670–674 (2011).

[139] Ivanov, I. I. *et al.* Induction of intestinal th17 cells by segmented filamentous bacteria. *Cell* **139**, 485–498 (2009).

[140] Atarashi, K. *et al.* Induction of colonic regulatory t cells by indigenous clostridium species. *Science* **331**, 337–341 (2011).

[141] Wen, L. *et al.* Innate immunity and intestinal microbiota in the development of type 1 diabetes. *Nature* **455**, 1109–1113 (2008).

[142] Burrows, M. P., Volchkov, P., Kobayashi, K. S. & Chervonsky, A. V. Microbiota regulates type 1 diabetes through toll-like receptors. *Proceedings of the National Academy of Sciences* **112**, 9973–9977 (2015).

[143] Gilbert, J. A. *et al.* Current understanding of the human microbiome. *Nature medicine* **24**, 392–400 (2018).

[144] Stewart, C. J. *et al.* Temporal development of the gut microbiome in early childhood from the teddy study. *Nature* **562**, 583–588 (2018).

[145] Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science* **341**, 1237439 (2013).

[146] Khan, A. A. *et al.* Polymorphic immune mechanisms regulate commensal repertoire. *Cell reports* **29**, 541–550 (2019).

[147] Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789–799 (2014).

[148] Carmody, R. N. *et al.* Diet dominates host genotype in shaping the murine gut microbiota. *Cell host & microbe* **17**, 72–84 (2015).

[149] Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *nature* **444**, 1027–1031 (2006).

[150] Ang, Q. Y. *et al.* Ketogenic diets alter the gut microbiome resulting in decreased intestinal th17 cells. *Cell* **181**, 1263–1275 (2020).

[151] Frazier, K. *et al.* High-fat diet disrupts reg3γ and gut microbial rhythms promoting metabolic dysfunction. *Cell Host & Microbe* (2022).

[152] Schroeder, B. O. & Bäckhed, F. Signals from the gut microbiota to distant organs in physiology and disease. *Nature medicine* **22**, 1079–1089 (2016).

[153] Kane, M. *et al.* Successful transmission of a retrovirus depends on the commensal microbiota. *Science* **334**, 245–249 (2011).

[154] Kasarda, D. D. Can an increase in celiac disease be attributed to an increase in the gluten content of wheat as a consequence of wheat breeding? *Journal of agricultural and food chemistry* **61**, 1155–1159 (2013).

[155] Murray, B. *et al.* Generation of restriction map of enterococcus faecalis og1 and investigation of growth requirements and regions encoding biosynthetic function. *Journal of bacteriology* **175**, 5216–5223 (1993).

[156] Qin, X., Singh, K. V., Weinstock, G. M. & Murray, B. E. Characterization of fsr, a regulator controlling expression of gelatinase and serine protease in enterococcus faecalis og1rf. *Journal of bacteriology* **183**, 3372–3382 (2001).

[157] Sifri, C. D. *et al.* Virulence effect of enterococcus faecalis protease genes and the quorum-sensing locus fsr in caenorhabditis elegans and mice. *Infection and immunity* **70**, 5647–5650 (2002).

[158] Singh, K. V., Qin, X., Weinstock, G. M. & Murray, B. E. Generation and testing of mutants of enterococcus faecalis in a mouse peritonitis model. *The Journal of infectious diseases* **178**, 1416–1420 (1998).

[159] Kristich, C. J., Chandler, J. R. & Dunny, G. M. Development of a host-genotype-independent counterselectable marker and a high-frequency conjugative delivery system and their use in genetic analysis of enterococcus faecalis. *Plasmid* **57**, 131–144 (2007).

[160] Bryan, E. M., Bae, T., Kleerebezem, M. & Dunny, G. M. Improved vectors for nisin-controlled expression in gram-positive bacteria. *Plasmid* **44**, 183–190 (2000).

[161] Waters, C. M., Antiporta, M. H., Murray, B. E. & Dunny, G. M. Role of the enterococcus faecalis gele protease in determination of cellular chain length, supernatant pheromone levels, and degradation of fibrin and misfolded surface proteins. *Journal of bacteriology* **185**, 3613–3623 (2003).

[162] Lieberman, S. M. *et al.* Identification of the β cell antigen targeted by a prevalent population of pathogenic cd8+ t cells in autoimmune diabetes. *Proceedings of the National Academy of Sciences* **100**, 8384–8388 (2003).

[163] Takaki, T. *et al.* Requirement for both h-2db and h-2kd for the induction of diabetes by the promiscuous cd8+ t cell clonotype ai4. *The Journal of Immunology* **173**, 2530–2541 (2004).

[164] Wong, F. S. *et al.* Identification of an mhc class i-restricted autoantigen in type 1 diabetes by screening an organ-specific cdna library. *Nature medicine* **5**, 1026–1031 (1999).

[165] Zheng, G. X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nature communications* **8**, 1–12 (2017).

[166] Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).

[167] Wolf, F. A., Angerer, P. & Theis, F. J. Scanpy: large-scale single-cell gene expression data analysis. *Genome biology* **19**, 1–5 (2018).

[168] Szklarczyk, D. *et al.* The string database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic acids research* **39**, D561–D568 (2010).

[169] Mootha, V. K. *et al.* Pgc-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics* **34**, 267–273 (2003).

[170] Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545–15550 (2005).

[171] Sturm, G. *et al.* Scirpy: a scanpy extension for analyzing single-cell t-cell receptor-sequencing data. *Bioinformatics* **36**, 4817–4818 (2020).

[172] Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences* **89**, 10915–10919 (1992).

[173] McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction (2018). URL `https://arxiv.org/abs/1802.03426`.

[174] Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the illumina hiseq and miseq platforms. *The ISME journal* **6**, 1621–1624 (2012).

[175] Walters, W. *et al.* Improved bacterial 16s rrna gene (v4 and v4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *Msystems* **1**, e00009–15 (2016).

[176] Caporaso, J. G. *et al.* Qiime allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335–336 (2010).

[177] Caporaso, J. G. *et al.* Pynast: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266–267 (2010).

[178] Edgar, R. C. Search and clustering orders of magnitude faster than blast. *Bioinformatics* **26**, 2460–2461 (2010).

[179] McDonald, D. *et al.* An improved greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME journal* **6**, 610–618 (2012).

[180] Zamakhchari, M. *et al.* Identification of rothia bacteria as gluten-degrading natural colonizers of the upper gastro-intestinal tract. *PloS one* **6**, e24455 (2011).

[181] Berger, M., Sarantopoulos, C., Ongchangco, D., Sry, J. & Cesario, T. Rapid isolation of gluten-digesting bacteria from human stool and saliva by using gliadin-containing plates. *Experimental Biology and Medicine* **240**, 917–924 (2015).

[182] Qiao, S.-W. *et al.* Posttranslational modification of gluten shapes tcr usage in celiac disease. *The Journal of Immunology* **187**, 3064–3071 (2011).

[183] Sollid, L. M., Qiao, S.-W., Anderson, R. P., Gianfrani, C. & Koning, F. Nomenclature and listing of celiac disease relevant gluten t-cell epitopes restricted by hla-dq molecules. *Immunogenetics* **64**, 455–460 (2012).

[184] Krautkramer, K. A. *et al.* Diet-microbiota interactions mediate global epigenetic programming in multiple host tissues. *Molecular cell* **64**, 982–992 (2016).

[185] Haskins, K., Portas, M., Bradley, B., Wegmann, D. & Lafferty, K. T-lymphocyte clone specific for pancreatic islet antigen. *Diabetes* **37**, 1444–1448 (1988).

[186] Daniel, D., Gill, R. G., Schloot, N. & Wegmann, D. Epitope specificity, cytokine production profile and diabetogenic activity of insulin-specific t cell clones isolated from nod mice. *European journal of immunology* **25**, 1056–1062 (1995).

[187] Nakayama, M. *et al.* Prime role for an insulin epitope in the development of type 1 diabetes in nod mice. *Nature* **435**, 220–223 (2005).

[188] Chaparro, R. J. *et al.* Nonobese diabetic mice express aspects of both type 1 and type 2 diabetes. *Proceedings of the National Academy of Sciences* **103**, 12475–12480 (2006).

[189] Toye, A. *et al.* A genetic and physiological study of impaired glucose homeostasis control in c57bl/6j mice. *Diabetologia* **48**, 675–686 (2005).

[190] Montgomery, M. K. *et al.* Mouse strain-dependent variation in obesity and glucose homeostasis in response to high-fat feeding. *Diabetologia* **56**, 1129–1139 (2013).

[191] Christensen, A. A. & Gannon, M. The beta cell in type 2 diabetes. *Current diabetes reports* **19**, 1–8 (2019).

[192] Back, S. H. & Kaufman, R. J. Endoplasmic reticulum stress and type 2 diabetes. *Annual review of biochemistry* **81**, 767–793 (2012).

[193] Bottazzo, G. F. Death of a beta cell: homicide or suicide? *Diabetic medicine* **3**, 119–130 (1986).

[194] Freiesleben De Blasio, B., Bak, P., Pociot, F., Karlsen, A. E. & Nerup, J. Onset of type 1 diabetes: a dynamical instability. *Diabetes* **48**, 1677–1685 (1999).

[195] Yurkovetskiy, L. A., Pickard, J. M. & Chervonsky, A. V. Microbiota and autoimmunity: exploring new avenues. *Cell host & microbe* **17**, 548–552 (2015).

[196] Berridge, M. J. The endoplasmic reticulum: a multifunctional signaling organelle. *Cell calcium* **32**, 235–249 (2002).

[197] Eizirik, D. L., Cardozo, A. K. & Cnop, M. The role for endoplasmic reticulum stress in diabetes mellitus. *Endocrine reviews* **29**, 42–61 (2008).

[198] Rasschaert, J. *et al.* Global profiling of double stranded rna-and ifn-γ-induced genes in rat pancreatic beta cells. *Diabetologia* **46**, 1641–1657 (2003).

[199] Kohanim, Y. K., Tendler, A., Mayo, A., Friedman, N. & Alon, U. Endocrine autoimmune disease as a fragility of immune surveillance against hypersecreting mutants. *Immunity* **52**, 872–884 (2020).

[200] Group, T. S. Study design of the trial to reduce iddm in the genetically at risk (trigr). *Pediatric diabetes* **8**, 117–137 (2007).

[201] Knip, M. *et al.* Dietary intervention in infancy and later signs of beta-cell autoimmunity. *New England Journal of Medicine* **363**, 1900–1908 (2010).

[202] Urade, R., Sato, N. & Sugiyama, M. Gliadins from wheat grain: an overview, from primary structure to nanostructures of aggregates. *Biophysical reviews* **10**, 435–443 (2018).

[203] Shiferaw, B. *et al.* Crops that feed the world 10. past successes and future challenges to the role played by wheat in global food security. *Food Security* **5**, 291–317 (2013).

[204] Sildorf, S. M., Fredheim, S., Svensson, J. & Buschard, K. Remission without insulin therapy on gluten-free diet in a 6-year old boy with type 1 diabetes mellitus. *Case Reports* **2012**, bcr0220125878 (2012).

[205] Smyth, D. J. *et al.* Shared and distinct genetic variants in type 1 diabetes and celiac disease. *New England Journal of Medicine* **359**, 2767–2777 (2008).

[206] Steiner, D. J., Kim, A., Miller, K. & Hara, M. Pancreatic islet plasticity: interspecies comparison of islet architecture and composition. *Islets* **2**, 135–145 (2010).

[207] Miragaia, R. J. *et al.* Single-cell transcriptomics of regulatory t cells reveals trajectories of tissue adaptation. *Immunity* **50**, 493–504 (2019).

[208] Spence, A. *et al.* Revealing the specificity of regulatory t cells in murine autoimmune diabetes. *Proceedings of the National Academy of Sciences* **115**, 5265–5270 (2018).

[209] Delacher, M. *et al.* Precursors for nonlymphoid-tissue treg cells reside in secondary lymphoid organs and are programmed by the transcription factor batf. *Immunity* **52**, 295–312 (2020).

[210] Santamaria, P. *et al.* Beta-cell-cytotoxic cd8+ t cells from nonobese diabetic mice use highly homologous t cell receptor alpha-chain cdr3 sequences. *The Journal of Immunology* **154**, 2494–2503 (1995).

[211] Verdaguer, J. *et al.* Spontaneous autoimmune diabetes in monoclonal t cell nonobese diabetic mice. *The Journal of experimental medicine* **186**, 1663–1676 (1997).

[212] Graser, R. T. *et al.* Identification of a cd8 t cell that can independently mediate autoimmune diabetes development in the complete absence of cd4 t cell helper functions. *The Journal of Immunology* **164**, 3913–3918 (2000).

[213] Amrani, A. *et al.* Perforin-independent β-cell destruction by diabetogenic cd8+ t lymphocytes in transgenic nonobese diabetic mice. *The Journal of clinical investigation* **103**, 1201–1209 (1999).

[214] Lieberman, S. M. *et al.* Individual nonobese diabetic mice exhibit unique patterns of cd8+ t cell reactivity to three islet antigens, including the newly identified widely expressed dystrophia myotonica kinase. *The Journal of Immunology* **173**, 6727–6734 (2004).

[215] DiLorenzo, T. P. *et al.* Major histocompatibility complex class i-restricted t cells are required for all but the end stages of diabetes development in nonobese diabetic mice and use a prevalent t cell receptor α chain gene rearrangement. *Proceedings of the National Academy of Sciences* **95**, 12538–12543 (1998).

[216] Warshauer, J. T. *et al.* A human mutation in stat3 promotes type 1 diabetes through a defect in cd8+ t cell tolerance. *Journal of Experimental Medicine* **218**, e20210759 (2021).

[217] Utsugi, T. *et al.* Major histocompatibility complex class i–restricted infiltration and destruction of pancreatic islets by nod mouse-derived β-cell cytotoxic cd8+ t-cell clones in vivo. *Diabetes* **45**, 1121–1131 (1996).

[218] Malkiel, S., Liao, L., Cunningham, M. W. & Diamond, B. T-cell-dependent antibody response to the dominant epitope of streptococcal polysaccharide, n-acetylglucosamine, is cross-reactive with cardiac myosin. *Infection and immunity* **68**, 5803–5808 (2000).

[219] Lang, K. S. *et al.* Toll-like receptor engagement converts t-cell autoreactivity into overt autoimmune disease. *Nature medicine* **11**, 138–145 (2005).

[220] Bouziat, R. *et al.* Reovirus infection triggers inflammatory responses to dietary antigens and development of celiac disease. *Science* **356**, 44–50 (2017).

[221] Kawalec, M., Potempa, J., Moon, J. L., Travis, J. & Murray, B. E. Molecular diversity of a putative virulence factor: purification and characterization of isoforms of an extracellular serine glutamyl endopeptidase of enterococcus faecalis with different enzymatic activities. *Journal of bacteriology* **187**, 266–275 (2005).

[222] Gao, B. & Tsan, M.-F. Recombinant human heat shock protein 60 does not induce the release of tumor necrosis factor α from murine macrophages. *Journal of Biological Chemistry* **278**, 22523–22529 (2003).

[223] Habich, C. *et al.* Heat shock protein 60: specific binding of lipopolysaccharide. *The Journal of Immunology* **174**, 1298–1305 (2005).

[224] Wilks, J. *et al.* Mammalian lipopolysaccharide receptors incorporated into the retroviral envelope augment virus transmission. *Cell host & microbe* **18**, 456–462 (2015).

[225] Sato, S. *et al.* Synergy and cross-tolerance between toll-like receptor (tlr) 2-and tlr4-mediated signaling pathways. *The Journal of Immunology* **165**, 7096–7101 (2000).

[226] Beutler, E., Gelbart, T. & West, C. Synergy between tlr2 and tlr4: a safety mechanism. *Blood Cells, Molecules, and Diseases* **27**, 728–730 (2001).

[227] Dahal-Koirala, S. *et al.* Tcr sequencing of single cells reactive to dq2. 5-glia-α2 and dq2. 5-glia-ω2 reveals clonal expansion and epitope-specific v-gene usage. *Mucosal immunology* **9**, 587–596 (2016).

[228] Junker, Y. *et al.* Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *Journal of Experimental Medicine* **209**, 2395–2408 (2012).

[229] Vatanen, T. *et al.* Variation in microbiome lps immunogenicity contributes to autoimmunity in humans. *Cell* **165**, 842–853 (2016).

[230] Wei, G. *et al.* Identification of pseudolysin (lasb) as an aciduric gluten-degrading enzyme with high therapeutic potential for celiac disease. *The American journal of gastroenterology* **110**, 899 (2015).

[231] Sigmon, J. S. *et al.* Content and performance of the minimuga genotyping array: a new tool to improve rigor and reproducibility in mouse research. *Genetics* **216**, 905–930 (2020).

[232] Drucker, D. J. The biology of incretin hormones. *Cell metabolism* **3**, 153–165 (2006).

[233] Parikh, A., Stephan, A.-F. & Tzanakakis, E. S. Regenerating proteins and their expression, regulation, and signaling. *Biomolecular concepts* **3**, 57–70 (2012).

[234] Ohara-Imaizumi, M. *et al.* Serotonin regulates glucose-stimulated insulin secretion from pancreatic β cells during pregnancy. *Proceedings of the National Academy of Sciences* **110**, 19420–19425 (2013).

[235] Dandolo, L., Stewart, C. L., Mattei, M.-G. & Avner, P. R. Inactivation of an x-linked transgene in murine extraembryonic and adult tissues. *Development* **118**, 641–649 (1993).

[236] Katsuta, H. *et al.* Subpopulations of gfp-marked mouse pancreatic β-cells differ in size, granularity, and insulin secretion. *Endocrinology* **153**, 5180–5187 (2012).