Published in final edited form as:

Nat Immunol. 2020 April ; 21(4): 455–463. doi:10.1038/s41590-020-0623-7.

The MHCII peptidome of the pancreatic islet identifies key features of autoimmune peptides

Xiaoxiao Wan[1,3], **Anthony N. Vomund**[1,3], **Orion J. Peterson**[1], **Alexander V. Chervonsky**[2], **Cheryl F. Lichti**[1,*], **Emil R. Unanue**[1,*]

[1]Department of Pathology and Immunology and the Bursky Center for Human Immunology and Immunotherapy, Washington University School of Medicine, St Louis, MO 63110, USA

[2]Department of Pathology, University of Chicago School of Medicine, Chicago, IL 60637, USA

Abstract

The nature of autoantigens that trigger autoimmune diseases has been much discussed, but direct biochemical identification is lacking for most. Addressing this question demands unbiased examination of the self-peptides displayed by a defined autoimmune major histocompatibility complex class II (MHCII) molecule. Here we examined the immunopeptidome of the pancreatic islets in non-obese diabetic (NOD) mice, which spontaneously develop autoimmune diabetes based on the I-A^{g7} variant of MHCII. The relevant peptides that induced pathogenic $CD4^+$ T cells at the initiation of diabetes derived from proinsulin. These peptides were also found in the MHCII peptidome of the pancreatic lymph nodes and spleen. The proinsulin-derived peptides followed a trajectory from their generation and exocytosis in β cells, to uptake and presentation in islets and peripheral sites. Such a pathway generated conventional epitopes but also resulted in the presentation of post-translationally modified peptides, including deamidated sequences. These analyses reveal the key features of a restricted component in the self-MHCII peptidome that caused autoreactivity.

Genetic variants of the class II major histocompatibility complex (MHCII) alleles strongly correlate with the propensity to develop tissue-specific autoimmunity. Crucial to this process is the binding of tissue-derived peptides by susceptible MHCII molecules, expressed on the surface of antigen presenting cells (APCs), for display to autoreactive $CD4^+$ T cells. The non-obese diabetic (NOD) mouse strain spontaneously develops autoimmunity targeting the β cells of pancreatic islets and mimics many of the features of human type 1 diabetes (T1D)[1]. The major susceptibility genes in NOD and human diabetes, encode the I-A^{g7} and

*Corresponding author: clichti@wustl.edu; unanue@wustl.edu.

Author contribution

E.R.U., C.F.L., X.W., and A.N.V. planned experiments, interpreted results and evaluated the data. X.W. and O.J.P. prepared the biological samples. A.N.V. isolated the MHCII molecules. C.F.L. performed the mass spectrometry experiments and analyzed the data. A.N.V. and X.W. carried out the immunological experiments. A.V.C. contributed to the germ-free mice used in the experiments, reviewed and examined the paper. X.W, C.F.L, and E.R.U. wrote the paper with inputs from all the authors.

[3]These authors contributed equally.

Competing interests

The authors declare no competing interests.

HLA-DQ8 MHCII molecules respectively, which share common structural properties[2–4]. Both molecules have a neutral substitution of the aspartic acid at position 57 of the β-chain (β57), failing to form an ion bond with an arginine at the position 76 of the alpha chain. Such change generates an anchoring site that favors binding of peptides with acidic residues at the carboxy end[5–8]. Reintroduction of aspartic acid at β57 completely abolishes development of diabetes in NOD mice[9–11].

$CD4^+$ T cells that recognize insulin peptides are the major drivers of the autoimmune process in NOD mice[12]. Early studies in NOD mice identified T cells reacting with peptides derived from the insulin B chain (InsB) and the C peptide (InsC)[13–16], both components of proinsulin. Subsequently, islets isolated from 2–3-week old NOD mice were shown to harbor insulin-reactive $CD4^+$ T cells before any other indication of autoreactivity[11,17]. In humans with T1D, insulin autoreactivity manifests through the initial appearance of insulin autoantibodies in sera; insulin-reactive T cells have been detected in inflamed islets[18–20]. Whether insulin autoreactivity is the main component in human T1D has not been established.

Many of the immunogenic insulin peptides were found in the β-cell crinophagic bodies (crinosomes), a set of degradative vesicles formed from the fusion of insulin dense core granules (DCGs) with lysosomes[21]. These insulin peptides are released from the islets in response to glucose stimulation and seed various peripheral lymphoid tissues[21]. A number of reports identified $CD4^+$ T cells directed to other β-cell-derived antigens (reviewed in ref.[22]), but their role in the development of diabetes remains unclear. As such, although a broad reactivity in autoimmune diabetes has been characterized, identification of the actual MHCII epitopes presented in the target organ and the peripheral lymphoid tissues requires critical evaluation.

We searched the MHCII peptidome in the islets, pancreatic lymph nodes (pLN) and spleens of female NOD mice and identified the main β-cell-derived peptides. We then correlated the identified peptides with their binding strength to $I\text{-}A^{g7}$ and with the presence of autoreactive $CD4^+$ T cells isolated from islets and pLNs. T cell autoreactivity mainly originated from recognition of a selected component of the tissue-specific MHCII peptidome, in which peptides derived from proinsulin were the dominant autoantigens.

Results

The islet MHCII peptidome identifies β-cell-derived peptides

We examined by mass spectrometry the MHCII peptidome of the pancreatic islets isolated from 219 female NOD mice of 8–10-week of age. The islets contained two major sets of myeloid APCs: $CD45^+CD11c^+F4/80^+$ resident macrophages and incoming $CD45^+CD11c^+F4/80^-$ dendritic cells (DCs) (Extended Data Fig. 1a). These APCs express high amounts of MHCII and present insulin peptides (reviewed in[23]). We also examined the MHCII peptidome from the pLN and spleen of 8–10-week old female NOD mice. The peptide-MHCII complexes (pMHCs) were isolated from the three sites and the nature of the bound peptides was examined by mass spectrometry (Extended Data Fig. 1b). Using Gibbs cluster analysis[24], the $I\text{-}A^{g7}$-binding motif based on all the peptides isolated from spleen was

consistent with previous studies that showed a predominance of acidic residues at the C-terminal ends of the peptide[8] (Fig. 1a). The MHCII peptidome of spleens from I-A^b NOD mice was distinctly different from that of I-A^{g7} NOD mice (Fig. 1a), indicating the specificity of the I-A^{g7} peptidome.

We isolated ~4 × 10^5 APCs from the islets of the I-A^{g7} NOD mice and detected MHCII-bound peptides from 112 proteins. The islet MHCII peptidome was enriched in β-cell-derived peptides (75 peptides from 9 different proteins), mostly derived from the β-cell secretory granules (Table 1; Supplementary Table 1). The β-cell-derived peptides were grouped in families, each having sequences of different lengths and variable flanking residues (Table 1; Supplementary Table 1). Most of the peptides had 14–17 residues, a typical length for MHCII-binding peptides, but about 10% of peptides in islets were longer, with 20 residues or more (Extended Data Fig. 1c).

InsB and InsC dominated the islet MHCII peptidome in relative abundance and in the large number of peptide variants (Table 1). Intact InsB (InsB:1–30; 31 residues) and InsC of proinsulin-1 (Ins1C:33–61; 29 residues) were detected among the MHCII-bound epitopes (Supplementary Table 1), raising the question whether these were further processed by the APCs to smaller fragments or were bound as such to MHCII. Using a standard antigen presentation assay that measures IL-2 production by $CD4^+$ T cell hybridomas during antigen-specific stimulation[17,21,25], we found that treatment of the C3g7 APCs (a B cell lymphoma cell line expressing I-A^{g7}) with chloroquine, a drug that prevents intracellular processing, did not affect the presentation of Ins1C:33–61 to a cognate $CD4^+$ T cell hybridoma (Extended Data Fig. 1d). Thus, the long MHCII-bound peptides can be directly presented without a reduction in their length. A hybrid insulin peptide (HIP) formed by the fusion of InsC with the islet amyloid polypeptide (IAPP) which had a sequence (LQTLAL-NAARD) identical to that of peptide HIP6.9[26,27], was also identified in the MHCII peptidome (Table 1). This InsC-IAPP HIP was detected at a low level in the islet and pLN MHCII peptidome and represented less than 1% of the total peptidome in each case (Table 1; Supplementary Table 1–2).

Considering the important role of peripheral presentation in the development of autoimmune diabetes[21,28,29], we examined the MHCII peptidome of the pLN (Supplementary Table 2) and spleen (Supplementary Table 3), as well as the free peptides in DCGs, crinosomes and in the secreted peptides from β cells in response to glucose challenge[21] (referred to as secretome hereafter; Supplementary Table 4). Several β-cell peptide families were identified in the MHCII peptidome of the pLN and spleen, mostly from InsB and InsC; others included IAPP, Zinc transporter 8 (ZnT8) and the hybrid InsC-IAPP, in lesser amounts (Table 1). We also examined the MHCII peptidome of $CD19^+$ B cells isolated from the spleen of 8–10-week old female NOD mice and of the whole spleen from age- and sex-matched B cell-deficient (uMT) NOD mice. These peptidomes contained similar β-cell-derived peptides (Table 1; Supplementary Table 3). Thus, the spleen MHCII peptidome was contributed by both B cells and phagocytes (macrophages and DCs). Most of the MHCII-bound sequences found in pLN and spleen were also detected as free peptides in the DCGs, crinosomes and secretome (Table 1; Supplementary Table 4), suggesting that peripheral APCs could have captured peptides released from β cells. The number and diversity of free peptides were

much higher than the MHCII-bound peptides (Table 1; Supplementary Table 4). Some peptides associated with previously reported T cell autoreactivity were not detected in any of the MHCII peptidomes. To exclude that failure to detect them was due to problems with false discovery rates (FDR) associated with immunopeptidomics experiments[30], we performed a targeted search for all the previously reported T1D autoantigens and I-A^{g7}-binding peptides[31]. No additional known T1D autoantigen sequences were identified. Thus, among the β-cell-derived peptides, those from InB and InC predominate in the MHCII peptidome of islets, pLN and spleen.

Selected β-cell-derived peptides induce autoreactivity

We carried out a screen of the representative peptides from each β-cell-derived peptide families (Table 2) for their corresponding $CD4^+$ T cell reactivity and relative MHCII binding strength. Each synthetic β-cell-derived peptides (Table 2) were added for two challenge cycles to pooled cells isolated from the islets and pLNs of 8–10-week old female NOD mice. The reactivity of the $CD4^+$ T cells was assayed by measuring the production of the cytokines IL-2 and IFN-γ by ELISPOT. We also measured the production of IL-10 and IL-4, which were invariably negative (results not shown). Stimulation of the islet and pLN $CD4^+$ T cells with an I-A^{g7}-binding hen egg-white lysozyme (HEL) peptide did not detect $CD4^+$ T cells specific for this foreign antigen (Extended Data Fig. 2a), indicating that the assay was specific and did not generate *de novo* responses. We also examined the sensitivity of the assay by spiking in preactivated HEL-reactive $CD4^+$ T cells and detected as few as 20 $CD4^+$ T cells (Extended Data Fig. 2b), indicating the high sensitivity of the assay. The relative binding strength of the peptides to MHCII was evaluated in a competitive binding assay against the reference peptide (g7-MIME) known to strongly bind to I-A^{g7} (Extended Data Fig. 2c).

Only a limited number (5 out of 21 examined) of β-cell-derived peptides correlated with the presence of autoreactive $CD4^+$ T cells in islets and pLNs (Fig. 1b). These MHCII-bound peptides were identified in both islets and the peripheral lymphoid tissues (Fig. 1b; Table 2). The major populations of reactive $CD4^+$ T cells recognized the peptide 9–23 of InsB (InsB:9–23) and several Ins1C segments, including Ins1C:33–61, Ins1C:37–61 and Ins1C:51–61(Fig. 1b), while the intact InsC of proinsulin-2 (Ins2C:33–63) induced little responses (Fig. 1b). We also detected $CD4^+$ T cells specific to the hybrid peptide InsC-IAPP (Fig. 1b). The peptides identified only in the islets generated low or negative $CD4^+$ T cell responses (Fig. 1b; Table 2), showed relatively low MHCII binding strength (Fig. 1c), and were in low abundance (Fig. 1d). $CD4^+$ T cells reactive to peptides from three known autoantigens, including chromogranin A (ChgA:358–371), ZnT8:313–326 and IAPP:78–90 were detected only in the islets and pLNs from 16- and 20-week old female NOD mice (Fig. 1e). Because these peptides were already found in the MHCII peptidome at 8–10 weeks (Table 2), “epitope spreading” may result not from the appearance of *de novo* pMHCs, but could be explained by the particular biology of the T cells, such as their diversification following the initial inflammation. $CD4^+$ T cells specific for peptides from secretogranin-1 (Scg1:438–454 and Scg1:516–535) and vitamin D-binding protein (DBP:389–405) were not detected throughout the time course (Fig. 1e). In brief, most of the $CD4^+$ T cell

autoreactivity in islets and pLN at the initial stage of diabetes is directed to peptides derived from InsB and InsC.

Diverse InsB peptides harbor a pathogenic epitope

CD4$^+$ T cells reactive to InsB:9–23[13–15] and two distinct 9 amino acid registers included in InsB:9–23, specifically InsB:12–20 and InsB:13–21, each having divergent biological properties[17,25,32], have been previously described. The InsB:12–20-specific CD4$^+$ T cells are reactive only to extracellularly-derived InsB peptides bound to I-A^{g7} and are highly diabetogenic[17,25,32]. The InsB:13–21-reactive CD4$^+$ T cells recognize APCs that process native insulin, are of a low frequency and are largely unresponsive upon antigen challenge[17,25,32]. We identified variants of InsB-derived peptides centered on the 9–23 region, including InsB:9–23, across the islets, pLN and spleen (Fig. 2a). The InsB:9–23-reactive CD4$^+$ T cells in the islets and pLN of 8–10-week old female NOD responded to InsB:12–20, but not InsB:13–21 (Fig. 2b). No T cell reactivity was detected for two other putative registers, InsB:14–22 and InsB:15–23 (Fig. 2b), indicating that the responses were highly specific to InsB:12–20.

We also identified three other InsB peptides containing the InsB:12–20 register, namely InsB:11–25, InsB:12–25 and InsB:12–26, in the spleen MHCII peptidome. The MHCII epitope predictive program TEA-DQ8[33] suggested InsB:15–23 was the preferred binding register included in InsB:11–25, followed by InsB:12–20 (Extended Data Fig. 3a). We generated 46 CD4$^+$ T cell hybridomas reactive to InsB:11–25 and found that half of these (represented by clone 21) cross-reacted with InsB:12–20 (Fig. 2c). The rest (represented by clone 58) did not recognize either InsB:12–20 or InsB:13–21 but responded to InsB:15–23 (Fig. 2c), although the recognition was highly dependent on the amino- and carboxy-flanking residues (Extended Data Fig. 3b–c). However, when the islet and pLN CD4$^+$ T cells were challenged with InsB:11–25, only CD4$^+$ T cells reactive to InsB:12–20 were identified (Fig. 2d). The islet APCs isolated from the same mice spontaneously activated clone 21, but not clone 58 (Fig. 2e). Thus, InsB:15–23-specific CD4$^+$ T cells did not directly contribute to the autoimmune process.

InsB:11–25 bound to I-A^{g7} stronger than InsB:9–23 (Fig. 2f). A brief pulse of the C3g7 APCs with InsB:11–25 resulted in sustained presentation of InsB:12–20 to the 9B9 CD4$^+$ T cell hybridoma (Fig. 2g), which has specificity to InsB:12–20[25]. In contrast, the responses to InsB:9–23 declined rapidly (Fig. 2g). These results indicated that InsB:11–25 was a better carrier of the InsB:12–20 epitope. Primary CD4$^+$ T cells activated with InsB:11–25, which contained InsB:12–20-reactive T cells, transferred diabetes into NOD *Rag1*$^{-/-}$ recipient mice (Fig. 2h). In sum, the MHCII peptidome contains a family of InsB peptides but the pathogenic CD4$^+$ T cells center on recognition of the InsB:12–20 epitope.

InsC peptides give rise to autoreactive CD4$^+$ T cells

Next we examined the CD4$^+$ T cell responses directed to Ins1C peptide segments encompassing their C-terminal region. Most of the MHCII-bound Ins1C- and Ins2C-derived peptides spanned the N-terminal region, while only a few contained the C-terminus (Fig. 3a). Production of IFN-γ and IL-2 by CD4$^+$ T cells from the draining lymph nodes 7 days

after immunization of NOD mice with the long peptide Ins1C:33–63 was only detected after restimulation with segments containing the C-terminal segment (Ins1C:33–63 and Ins1C:49–63), but not with Ins2C:33–65 (Fig. 3b). Immunization with Ins1C:41–55 or Ins1C:37–56, which lacks the C-terminus did not generate any response (Extended Data Fig. 4a). Adoptive transfer of primary $CD4^+$ T cells specific for Ins1C:49–63 into NOD *Rag1*$^{-/-}$ mice induced diabetes in the recipient mice (Fig. 3c), indicating that presentation of the Ins1C C-terminal peptides induced pathogenic T cell responses.

We identified deamidation of the four glutamine residues in the Ins1C peptides in all the MHCII peptidomes, including the glutamine to glutamic acid change at the C-terminus (Fig. 3a; Supplementary Table 1–3). Deamidation in the native Ins1C:51–61 (LQTLALEVARQ to LQTLALEVARE) was found in the spleen, while deamidation was only detected in Ins1C:53–61E (TLALEVARE) in islets (Fig. 3a; Supplementary Table 1–3). Deamidation in the C-terminus of Ins2C:53–63 (LQTLALEVAQQ) was not found across the three sites (Fig. 3a; Supplementary Table 1–3). Because deamidation and citrullination (arginine to citrulline) result in the same mass change, we compared the spectrum of the synthetic deamidated Ins1C:53–61E to the synthetic citrullinated peptide (TLALEVArQ) and found that the peptide spectrum from the islet peptidome matched deamidation, but not citrullination (Fig. 4d; Extended Data Fig. 4b).

To test whether deamidation could have been introduced during the isolation of the peptidome, we subjected the native Ins1C:53–61 (TLALEVARQ) to all the procedures for isolation of the MHCII peptidome. The deamidated version (TLALEVARE) was not detected by mass spectrometry (result not shown). Furthermore, a BLAST search with the TLALEVARE sequence indicated matches to several microbial species found in the mouse gut microbiome (results not shown). To determine whether the deamidated Ins1C:53–61E was of bacterial origin, we examined the MHCII peptidome of germ-free NOD mice. Ins1C:53–61E was detected with a spectrum identical to that found in normal NOD mice (results not shown). These observations indicated that deamidation of native InsC peptides was a biologically-induced process.

To understand how deamidation may impact autoreactivity, we compared the immunogenicity of the deamidated Ins1C:51–61E, which was identified only in the periphery (Supplementary Table 3), with native Ins1C:51–61. First, Ins1C:51–61E displayed a ~3-fold higher binding affinity to I-A^{g7} than Ins1C:51–61 (Fig. 3e), confirming that peptides with acidic residues are preferentially bound by I-A^{g7} (ref.[8]). The short Ins1C:53–61E bound poorly to I-A^{g7} (Fig. 3e), indicating the importance of flanking residues. Next, we generated a panel of $CD4^+$ T cell hybridomas reactive to Ins1C:51–61E, and found that all cross-reacted to the native Ins1C:51–61 (Fig. 3f); however, significantly lower amounts of Ins1C:51–61E were required to reach the half-maximal T cell responses (Fig. 3f). Finally, we searched the islets and pLN of 8–10-week old female NOD mice for $CD4^+$ T cells specific for Ins1C:51–61 or Ins1C:51–61E. Ins1C:51–61E generated substantially stronger responses than native Ins1C:51–61 or Ins1C:33–61 (Fig. 3g). In sum, deamidation changes in InsC did not give rise to $CD4^+$ T cells that exclusively recognized this PTM, but enhanced the immunogenicity of the unmodified peptide.

A multi-round search identifies an MHCII-bound HIP

Because hybrid peptides are not included in standard canonical protein databases, we employed a multi-round search strategy. We first searched a standard Uniprot/SwissProt mouse proteome database and then a published database of *in silico* generated HIPs[34]. If a spectrum was identified as a possible match to a HIP sequence, we then checked for possible alternatives assignments by querying PEAKS PTM[35] and SPIDER[36] to determine if other matches for a given spectrum were present. This strategy required the consideration of alternative sequences, such as those generated from non-standard modifications of the sample. Without such considerations, unknown sequences could be falsely assigned.

Five potential HIPs from the MHCII-bound peptides isolated from islets and pLN were searched for alternate assignments (Table 3). The spectra for 4 of these peptides were better assigned to native or modified InsC fragments (Table 3), indicating that the initial assignment as HIP was incorrect. In addition, the spectra of these 4 peptides did not match those of the synthetic HIP standards (Table 3; Extended Data Fig. 5). Only the InsC-IAPP HIP (LQTLAL-NAARD) was verified as a fused peptide (Table 3; Fig. 4a). We did not identify the HIP InsC-ChgA (LQTLAL-WSRMD)[27], corresponding to the epitope of the potent diabetogenic BDC2.5 $CD4^+$ T cell, possibly due to the presence of this peptide at levels below the limit of detection for the mass spectrometry method used in this study. In brief, when following strict criteria for identification of fused peptides, only one major HIP was found in the MHCII peptidome.

Crinosomes but not DCGs contain free HIPs

According to several of the proposed mechanisms for spliced peptide formation[37–39], crinosomes should be favorable for HIP formation. We reanalyzed published mass spectrometry data from crinosomes, DCGs and secretome[21]. The initial search identified 134 spectra for putative HIPs in crinosomes and 239 in the secretome, but none in DCGs (results not shown). We filtered the putative HIP assignments by either PEAKS PTM or SPIDER and then performed additional filtering by published criteria[34]. These stringent criteria identified 20 putative HIPs from crinosomes and the secretome (Supplementary Table 5). The HIPs from the two sources were similar in composition, in terms of being either cis-fused InsC-InsC or trans-fused InsC-IAPP HIPs (Supplementary Table 5). Three HIP sequences common between the crinosomes and the secretome were verified by matching their spectra to those of the corresponding synthetic peptides (Fig. 4b–d). Multiple truncated forms of a HIP formed by InsC and IAPP (EVED-TPVRSGTNPQM), which was previously identified in a proteolytic digest of islets[34], were detected in the secretome (Fig. 4b; Supplementary Table 5). The other two HIPs (Fig. 4c–d) were not previously reported. None of the HIPs from crinosomes or the secretome were identified in the MHCII peptidome (Table 3; Supplementary Table 5). These observations indicate that fused peptides are formed in crinosomes, a site of proteolytic activity and high peptide concentrations, and were secreted with insulin.

Discussion

Here we show that the initiation of diabetic autoimmunity depends on a limited number of $CD4^+$ T cells that recognize peptides from proinsulin. The immunogenic peptides were not only found in islets, but also in the peripheral lymphoid tissues. In addition to the native peptides, our analysis identified an important PTM, a deamidation in peptides derived from Ins1C, which enhanced the MHCII-binding affinity and the corresponding $CD4^+$ T cell reactivity. We also documented an MHCII-bound HIP in the islet and pLN peptidome and identified several others as free HIPs in the crinosomes, where the HIPs were formed.

Several features of InsB- and InsC- derived peptides contribute to their high immunogenicity. Both are from proinsulin, a major β-cell protein. The peptides were displayed as sets varying in size, including long segments. Some of the long peptides harbored flanking residues that favored a strong interaction of the 9 amino acid binding core with MHCII, resulting in enhanced $CD4^+$ T cell responses. The immunogenic peptides were presented also in peripheral lymphoid tissues, as a result of the exocytosis from crinosomes. We contend that a panoply of insulin peptides is generated in the β-cell crinosomes and flow to islet and peripheral APCs where they are bound and presented. Such a pathway is supported by our findings and by previous studies examining a diabetogenic T cell specific for InsB:12–20, the epitope examined here[17,40]. This mode of generation and presentation of peptides has two consequences[21]. First, it generated unique T cell epitopes that were not produced from the uptake and internal processing of insulin or proinsulin by the islet APCs, a process that involves editing of the internalized protein[17,25]. Second, the peptides were exported to secondary lymphoid tissues where they contributed to the reactivity of T cells[21]. Peripheral presentation is important because the absence of peripheral lymph nodes markedly dampened the development of autoimmune diabetes[28,29].

We note that the abundance of the MHCII-bound peptides did not necessarily correlate with immunogenicity. The Ins2C peptide family had the highest abundance in the islet peptidome but generated little $CD4^+$ T cell responses. Only a few Ins2C-derived peptides contained the immunogenic C-terminal segments, and none of these had deamidation that may enhance their MHCII binding affinity. Additionally, due to the expression of proinsulin-2 but not proinsulin-1 in the murine thymus, high selection pressure may be imposed upon the Ins2C-reactive T cells. In sum, the Ins2C family provides an example of a highly displayed autoantigen in the target organ that is non-immunogenic.

We documented naturally deamidated Ins1C peptides bound to MHCII. Deamidation improved binding to $I\text{-}A^{g7}$, although $CD4^+$ T cells exclusively recognizing the deamidated epitope were not detected. Thus, such a PTM will be important in improving the interaction of the $CD4^+$ T cells to the wild-type peptides, but not by giving rise to unique T cells that bypass conventional control mechanisms, an explanation frequently mentioned to explain the potential importance of peptides with PTMs. T cells have been reported to recognize deamidated sequences of peptides in T1D[41,42]. We identified deamidation in Ins1C:51–61 (LQTLALEVARQ), but not in Ins2C:53–63 (LQTLALEVAQQ), despite similar abundance. This finding underscores the influence of the amino acids proximal to the glutamine in the deamidation process and suggests a non-enzymatic process[43–45].

Although we identified the MHCII bound peptides involved at the initial stage of the autoimmune process in NOD mice, it is possible that we have missed peptides presented at a low abundance. Whether a very low amount of pMHCs in APCs can trigger $CD4^+$ T cells is an issue to contend. The finding that $CD4^+$ T cells to low-level epitopes did appear late suggests that there are events that foster their development and/or activation, i.e. the state of the islets as inflammation proceeds. We expect that additional late epitopes may involve other PTMs resulting from oxidative and stress responses. We did not find peptides derived from glutamic acid decarboxylase (GAD)[46] or islet antigen-2 (IA-2)[47]. One explanation is that some peptides are presented in tissues outside of islets. Diabetogenic T cells were reported to recognize commensal peptides that are cross-reactive with islet antigens[48,49].

Two findings may have clinical relevance. First, examining the peptidome of the secondary lymphoid tissues draining a target organ may provide relevant information on the key autoantigens. This has clinical applicability if sampling the affected organs is not feasible, as it happens in pancreatic islets. Second, considering that I-A^{g7}, HLA-DQ2 and HLADQ8 are structural homologs[2–4] and select similar peptidome[8], our analysis provides a foundation for assessing the T1D-relevant epitopes presented by them. We posit that the similar peptide families may be shared between the human and murine MHCII peptidome. HLA-DQ8-restricted T cells recognizing InsB:9–23[20] and the full-length C-peptide[50] have been identified.

Finally, we confirmed the identity of an MHCII-bound HIP and identified a few free fused peptides in crinosomes and in the secretome. Their identification in crinosomes, a structure rich in cathepsins, points to their formation via reverse proteolysis[39]. Our reported fused peptides were filtered in a very stringent manner in order to remove false positive assignments by eliminating all spectra that could be assigned to chemically or post-translationally modified forms.

Methods

Mice.

NOD/ShiLtJ (NOD), NOD.129S7(B6)-Rag1tm1Mom/J (NOD.$Rag1^{-/-}$), NOD.129S2(B6)-Ighmtm1Cgn/DoiJ (μMT), and NOD.B10Sn-H2b/J (NOD.H2b) mice were originally obtained from the Jackson Laboratory. NOD.10E11 TCR transgenic mice to HEL were generated in the laboratory. Germ-free NOD mice were provided by one of us (A.V.C). All the mice were bred and maintained under specific pathogen-free conditions in our animal facility. All experiments were approved by the Division of Comparative Medicine of Washington University School of Medicine in St. Louis (Accreditation number A3381–01).

Islet isolation.

Mice were killed and the peritoneal cavity was opened to expose the common bile duct. The duct leading into the duodenum was clamped, and under a microscope, a syringe fitted with a 30G needle was inserted into the duct close to the liver. Type XI collagenase (0.4 mg/ml; Sigma) in 5 ml Hank's BSS (no Ca^{2+}, Sigma) was injected to inflate the pancreas. The pancreas was removed and digested at 37°C for 15min. After digestion, crude islets were

washed and collected to a petri dish. Under the microscope, debris and undigested pancreas were removed and islets were handpicked.

Isolation of the MHCII peptidome.

We followed our original procedure[8], with modifications[51]. Cells isolated from islets, pLNs or spleens were suspended in lysis buffer (40mM MEGA 8, 40mM MEGA 9, 1mM PMSF, 0.2mM Iodoacetamide, 20ug/ml Leupeptin, Roche Complete Mini Protease cocktail in PBS) and rocked for 1hr at 4°C. The cell lysate was spun in a centrifuge at 20,000 xg for 30min at 4°C. In order to eliminate non-specific binding of peptides, the supernatant was first incubated with polyclonal mouse IgG (BioXcell; 1.5 mg antibody/sample) bound to Sepharose 4B at 4°C for 30min. The unbound material containing MHCII-peptide complexes was collected and was then added to a tube containing PBS-washed sepharose conjugated to the anti-I-A^{g7} antibody (AG2.42.7, 1.5 mg/sample) and incubated at 4°C overnight. The I-A^{g7}-sepharose was applied to a column and washed four times as follows: 10ml 150mM NaCl, 20mM Tris pH7.4; 10ml 400mM NaCl, 20mM Tris pH7.4; 10ml 150mM NaCl, 20mM Tris pH7.4; 10ml 20mM Tris pH8.0. Peptides were eluted with 10% acetic acid and dried with a SpeedVac. Eluted peptides were passed over detergent removal spin columns (Pierce) to remove traces of remaining detergent and were cleaned by using the C18 Spin Columns from ThermoScientific (Pierce).

Identification of effector T cells in islet/pLN.

Islets and pLNs were harvested from female NOD mice of various ages. Handpicked islets were dispersed using a non-enzymatic cell dispersion solution (Sigma) for single-cell suspensions. Total cells from the pLNs were isolated by digestion with Liberase (125 ug/ml, Roche) and DNase (50 ug/ml, Roche). Cells from islet and pLN cells were mixed together and cultured in DMEM media supplemented with 10% fetal bovine sera at a concentration of 2×10^6 cells/ml. The culture was supplemented with 20 U/ml IL-2 and 1 uM of the peptide of interest (challenge). After incubation at 37°C for 7 days, total live cells were collected from the culture using Histopaque 1119 and were subject to a second round of amplification. Accordingly, live cells were added to a fresh culture containing 20 U/ml IL-2 and 1 μM peptides as before along with irradiated NOD splenocytes (3000 RAD) as APCs (2×10^6/ml). After incubation at 37°C for 3 days, cells were harvested and were assayed for reactivity by eliciting a recall response on either IL-2 or IFN-γ coated 96-well multi-screen plates (MilliporeSigma) for ELISPOT assay. Cultured cells were plated as 2.5×10^5 /well with 10 μM antigen (recall), and the reactive cells were counted using Immunospot Software.

Competitive Binding Assay.

We used a cell assay using the C3g7 cell line expressing a high amount of I-A^{g7}. Different amounts of a known peptide (GKKVATTVHAGYG), well known to bind to I-A^{g7}, was used to compete against the binding to an HEL peptide (AMKRHGLDNYRGYSL); the HEL peptide was offered at concentration (200 pMoles in 200 μL) about 50% of its maximum binding, estimated by the biological response against a known T cell (clone 10E11). Competitor and the HEL peptides were incubated with C3g7 at 37°C for 60min. Cells then were spun down, washed, and tested against 5×10^4 10E11 T cells for 24 h. The T cell

responses were probed by the standard antigen presentation assay. IC50 values for competitor peptides were generated by fitting the data using GraphPad Prism software (GraphPad) as a four-parameter inhibitor vs response curve.

Antigen presentation assay.

The T cell responses to a given antigen were probed by IL-2 production measured by culturing supernatants in the presence of the IL-2 dependent cell line, CTLL-2[17]. The proliferation of CTLL-2 cells was measured by the incorporation of 3[H] thymidine. For experiments involving chloroquine, C3g7 APCs were treated with 100 μM chloroquine (Sigma) at 37°C for 2 h in the presence of the antigen and were washed thoroughly. The C3g7 APCs were then added directly to T cells without additional antigen. For experiments examining the rate of peptide dissociation, C3g7 cells were incubated with 1 μM peptide at 37°C for 1hr. After a thorough wash, the C3g7 cells were divided into different aliquots and were incubated at 37°C. At indicated time points (1, 2, 4, 6, 8, and 10 h), one aliquot of C3g7 was removed from the incubation, washed and cultured with the indicator T cells overnight.

Immunization and clonal selection.

8-week old male NOD mice were immunized with peptide or protein antigens (10 nmoles) emulsified in Complete Freund's adjuvant (CFA; Difco) subcutaneously in the footpads of hind legs. (The Ins1C:33–63 peptide used for immunization harbors two additional dibasic cleavage residues which do not influence the MHCII binding strength.) One week following immunization, the popliteal lymph nodes were collected, dispersed into single-cell suspensions. The T cell reactivity was probed by the IL-2 and IFN-γ ELISPOT assay using 10^6 cells upon recall with indicated antigens for 24h. For generating single-cell T cell hybridoma clones reactive to the immunized antigen, the popliteal lymph node cells were boosted with 1 μM of the same antigen for 3 days and were fused with the BW5147 fusion partner following standard protocols. The growth-positive T cell clones were screened against the corresponding antigen using the standard antigen presentation assay. The antigen-specific clones were further expanded and subcloned to generate single-cell T cell clones.

Generation of primary $CD4^+$ T cell lines and transfer.

Primary $CD4^+$ T cell lines were generated by immunization with the Ins1C:49–63 or InsB:11–25 peptide in CFA. The lines were maintained by weekly stimulation of the popliteal lymph node cells with 1 μM of either peptide in the presence of 20 U/mL of IL-2 and 2×10^6/mL of irradiated splenocytes and tested by proliferation before adoptive transfer intravenously to NOD.*Rag1*$^{-/-}$ recipients (10^6 cells/mouse). Mice were followed for blood glucose weekly and those with a level ≥ 250 mg/dL for two consecutive readings were considered diabetic.

Prediction of I-A^{g7}-binding epitopes.

Epitopes with potential binding capacities to I-A^{g7} were predicted by a published computational program[33]. In brief, the binding strength of a given epitope to I-A^{g7} was

calculated by an algorithm that considers the structural properties of I-A^{g7} and evaluates the preference of I-A^{g7} for binding to the amino acid residues within a given 9-residue epitope at P1, P4, P6, P7, and P9. For peptides containing multiple epitopes, each epitope was given a log of odds (LOD) score that represents the accumulative preference of the five positions. LOD scores closer to positive values indicate more preferred binding to I-A^{g7}.

Mass Spectrometry.

A Dionex UltiMate 1000 system (Thermo Scientific) was coupled to an Orbitrap Fusion Lumos (Thermo Scientific) through an Easy-Spray ion source (Thermo Scientific). For the islet sample, nLC separation was carried out as follows, while other samples were analyzed as previously described[21]. Peptide samples were loaded (15 μL/min, 1 min) onto a trap column (100 μm × 2 cm, 5 μm Acclaim PepMap 100 C18, 50 °C), eluted (0.2 μL/min) onto an Easy-Spray PepMap RSLC C18 column (2 μm, 50 cm × 75 μm ID, 50 °C, Thermo Scientific) and separated with the following gradient, all % Buffer B (0.1% formic acid in ACN): 0–110 min, 2%–22%; 110–120 min, 22%–35%; 120–130 min, 35–95%; 130–150 min, isocratic at 95%; 151–153 min, 95%–2%, 153–171 min, isocratic at 2%. Spray voltage was 1900V, ion transfer tube temperature was 275°C, and RF lens was 30%. MS scans were acquired in profile mode and MS/MS scans in centroid mode, for ions with charge states 2–7, with a cycle time of 3 sec. MS spectra were recorded from 375–1500 Da at 120K resolution (at m/z 200), and HCD MS/MS was triggered above a threshold of 2.0e4, with quadrupole isolation (0.7 Da) at 30K resolution, and collision energy (CE) of 30%. Dynamic exclusion was used (60 s), and monoisotopic precursor selection was on.

MS Data Analysis.

Data files were uploaded to PEAKS × (Bioinformatics Solutions, Waterloo, ON) for processing, de novo sequencing and database searching. The sequences were searched against the UniProt Mouse database (downloaded 1/12/2019; 17,000 entries) or a custom, in-house database consisting of previously reported T1D autoantigens[31] plus all of our identified I-A^{g7} peptides. We searched with mass error tolerances of 10 ppm and 0.01 Da for parent and fragment, respectively, no enzyme specificity and oxidation (M), deamidation (NQ) and trioxidation (C) as variable modifications. FDR estimation was enabled. Peptides were filtered at a 1% FDR at the peptide-spectrum match (PSM) level, and protein filtering was disabled by setting Protein −10logP score at 0 with 0 unique peptides with a requirement for significant peptides. A multi-round search was subsequently performed in which unassigned spectra were searched against a fused peptide database[34]. Putative HIPs were manually verified by visual inspection. Sequences that passed described criteria for HIP identification were appended to the mouse proteome database and the search was repeated.

Statistics.

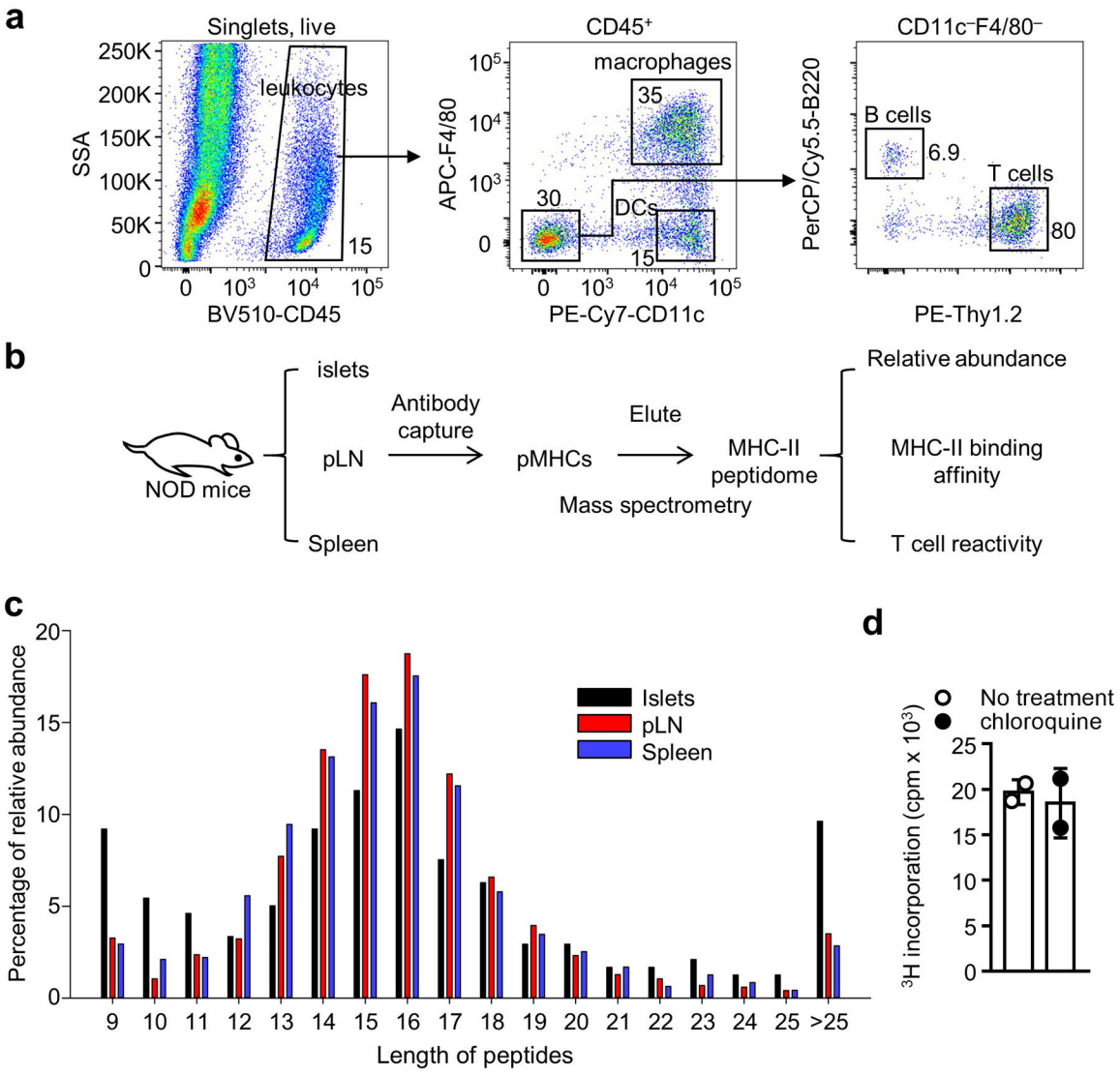
The mice were age and sex-matched and were randomized and equally distributed into experimental groups. Power analysis was used to determine the sample size in biological experiments. The sample size of mass spectrometry experiments was determined in preliminary experiments in which different sample sizes were tested to reach a satisfactory level of peptide recovery and detection. The two-tailed unpaired Student’s t-test was used to

determine the significant differences between two experimental groups containing unpaired biological replicates.

Data availability

The I-A^{g7}-binding epitope prediction program will be available upon request. The mass spectrometry data are available via ProteomeXchange with identifiers PXD015408 (MHCII peptidomes) and PXD015726 (reanalysis of crinosomes, DCGs and secretome).

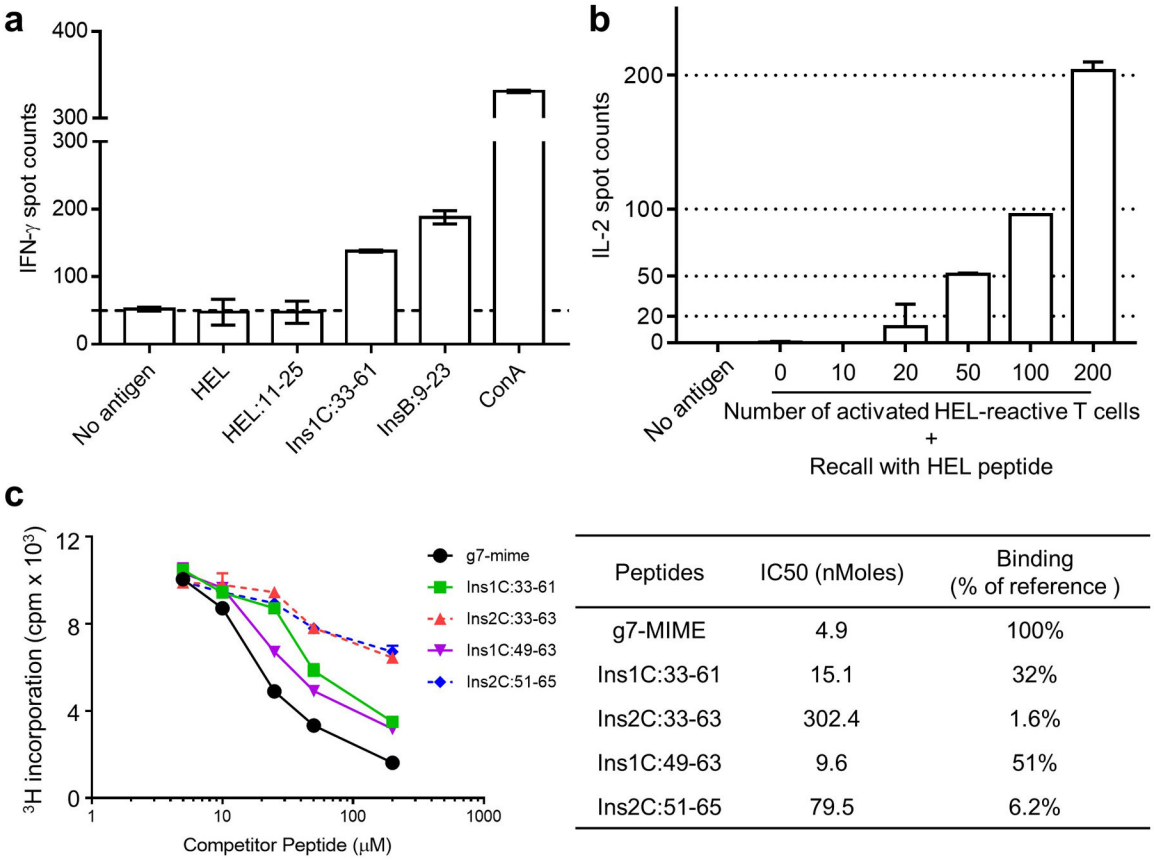
Extended Data



Extended Data Fig. 1. Assessing the MHCII peptidomes of the pancreatic islets, pancreatic lymph nodes and spleens from NOD mice.

a, Representative FACS plots showing the APC populations in the islets from 8–10-week old female NOD mice. The major APCs in the islet were the islet macrophages ($CD45^+CD11c^+F4/80^+$) and dendritic cells ($CD45^+CD11c^+F4/80^-$), along with a minor population of B cells ($CD45^+CD11c^-F4/80^-B220^+$). Data are representative of n = 4 independent experiment including n = 6 mice per experiment. **b**, Workflow for isolating the MHCII peptidome followed by mass spectrometry and immunological analyses. **c**, The lengths of all the peptides identified in the MHCII peptidomes from the indicated sites. Most of the peptides were 14–17 residues long, and a small number was longer than 25 residues. Data (mean) are from the mass spectrometry analysis of the MHCII peptidomes depicted in Supplementary Tables 1–3. **d**, Antigen presentation assay showing responses of an Ins1C-specific $CD4^+$ T cell hybridoma to C3g7 APCs treated with or without chloroquine and pulsed with the full-length 29-residue Ins1C:33–61. The assay measures IL-2 production by the $CD4^+$ T cell hybridoma during stimulation with the cognate antigen, assessed by the proliferation (^3H

incorporation) of the IL-2-depedent cell line CTLL-2 (see Methods). The Data (mean ± s.e.m.) are representative of n = 2 independent experiments with similar results.



Peptides	IC50 (nMoles)	Binding (% of reference)
g7-MIME	4.9	100%
Ins1C:33-61	15.1	32%
Ins2C:33-63	302.4	1.6%
Ins1C:49-63	9.6	51%
Ins2C:51-65	79.5	6.2%

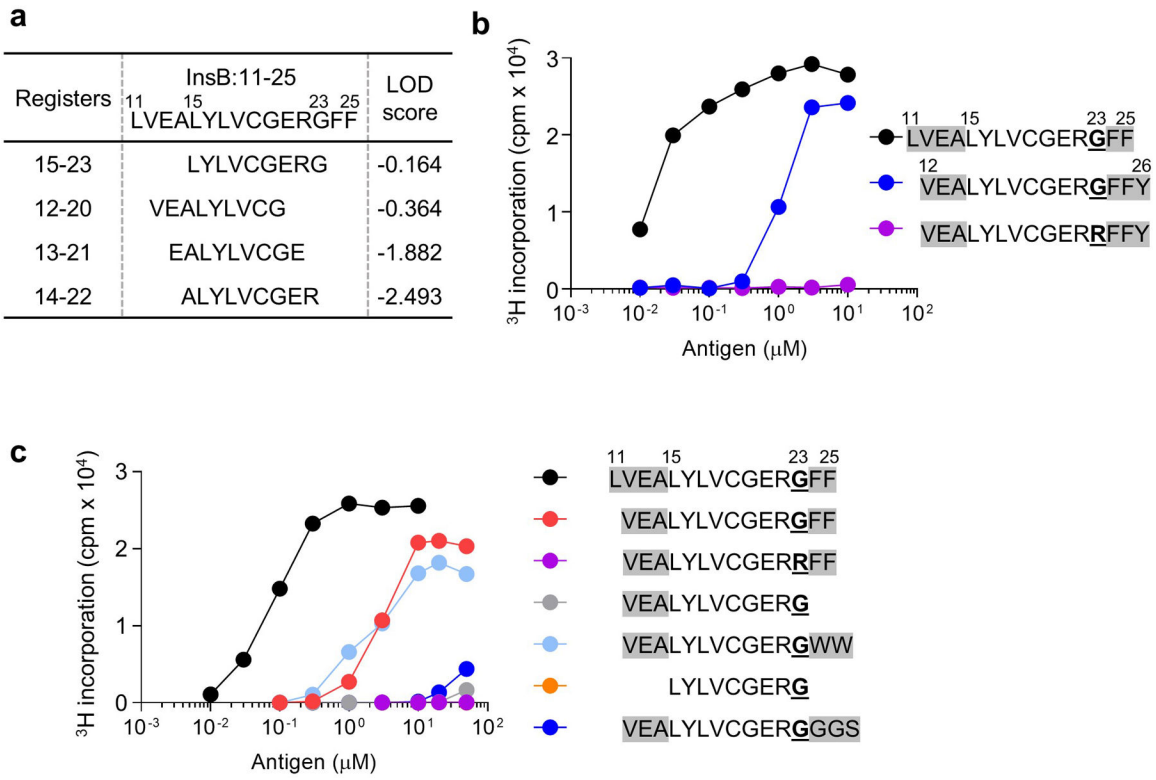
Extended Data Fig. 2. Evaluation of T cell reactivity and relative MHCII binding affinity in peptides selected from the MHCII peptidome.

a, Pooled islet and pLN cells from 8–10-week old female NOD mice were challenged with each indicated peptide for two cycles, and ELISPOT assays were conducted to read IFN-γ production in the cells upon recalling with the same peptide (see Methods). The results show positive T cell responses to two immunogenic β-cell-derived peptides, Ins1C:33–61 and InsB:9–23, indicating the presence of the effector T cells to these peptides. No IFN-γ responses were observed to the HEL protein or the I-A^{g7}-binding HEL:11–25 peptide, demonstrating that the culture assay did not generate *de novo* T cells. Data (mean ± s.e.m.) are representative of n = 3 experiments using n = 5 mice per experiment. **b**, Different numbers of pre-activated HEL-reactive T cells were spiked into the normal two-cycle culture, followed by recall with the HEL peptide and ELISPOT assay. The data depict a ~1:1 recovery ratio of the spiked HEL-reactive T cells with the IL-2 spots. The assay detected as few as ~20 pre-activated HEL-reactive T cells, indicating a high sensitivity. Data (mean ± s.e.m.) are representative of n = 2 experiments using n = 8 mice per experiment. **c**, An example of the competitive binding assay and the presentation of the data. A standard APC line (C3g7) was cultured with 1 μM HEL:11–25 peptide together with serial dilutions of a competitor peptide: the response of a specific T cell hybridoma to HEL:11–25 was probed by standard antigen presentation assay (see Methods). In every experiment, each competitor peptides were compared to a reference peptide, g7-MIME, with known strong binding to I-A^{g7}. The amounts of the peptide required to compete half-maximal T cell response to

HEL:11–25 (IC50) was estimated and compared to the reference peptide. A higher amount indicated a weaker binding affinity. The table (right) depicts the IC50 calculated from several representative peptides. The results are presented as the percent of reference after normalization to the reference g7-MIME peptide. Data are presentative of n = 4 independent experiments with similar results.

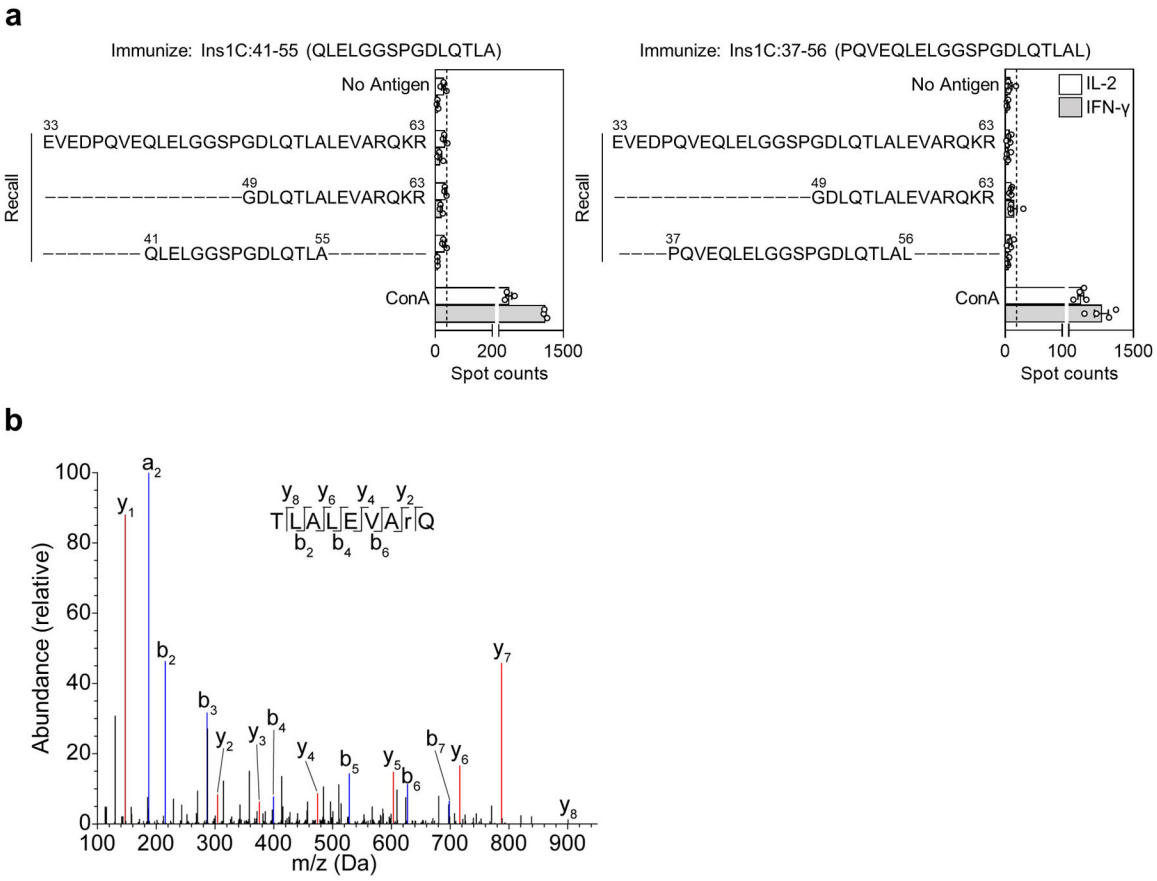
a

Registers	InsB:11-25 (11 15 23 25) LVEALYLVCGERGFF	LOD score
15-23	LYLVCGERG	-0.164
12-20	VEALYLVCG	-0.364
13-21	EALYLVCGE	-1.882
14-22	ALYLVCGER	-2.493



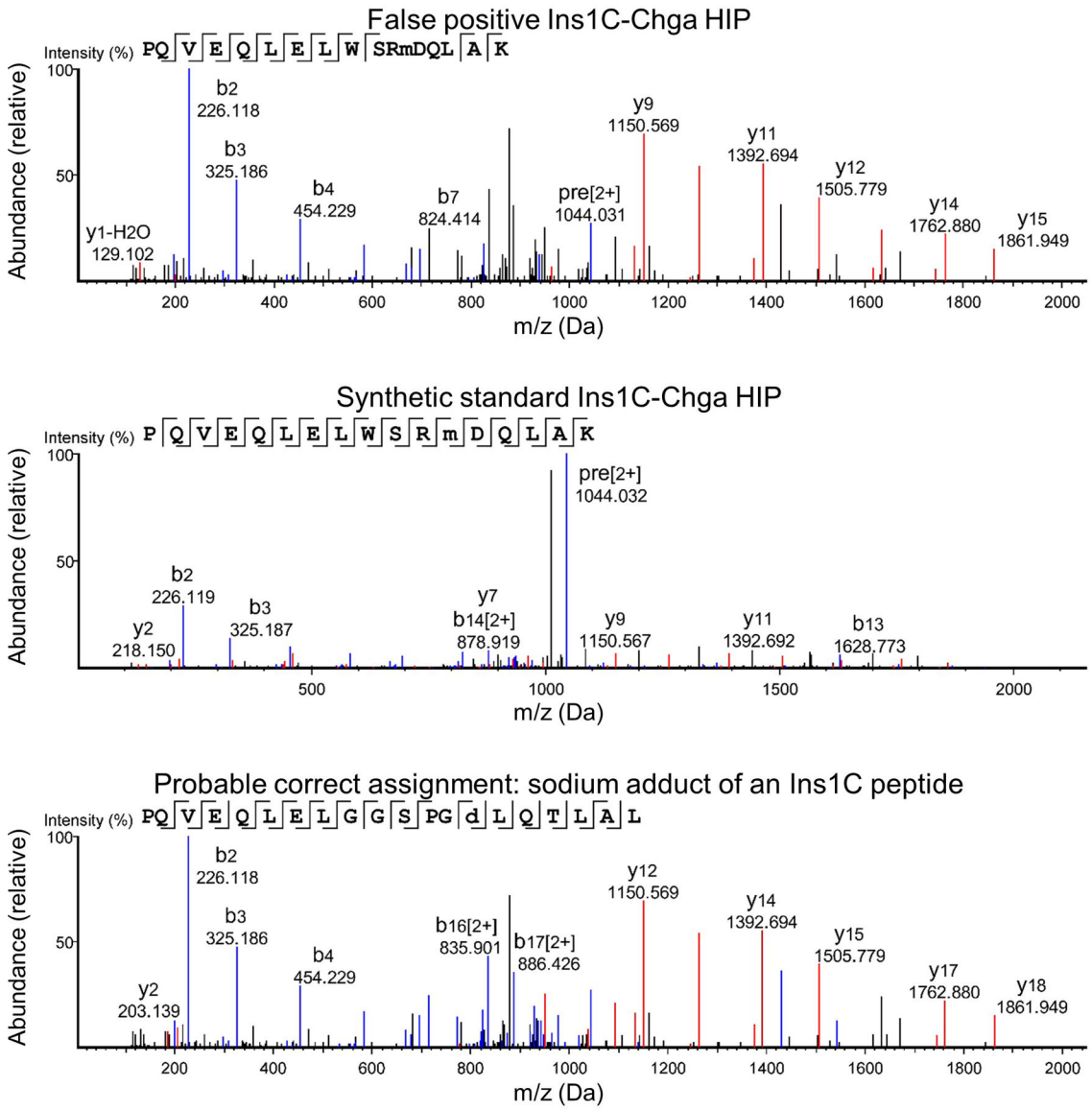
Extended Data Fig. 3. CD4$^+$ T cell recognition of the InsB:15–23 register is influenced by the nature of the flanking residues.

a, Predicted I-A^{g7}-binding registers included in the InsB:11–25 peptide. A preferred binding register is indicated by a log of odds (LOD) score (see Methods). **b, c**, An InsB:11–25-specific T cell clone (clone 58) was examined for its responses to C3g7 APCs pulsed with peptides containing the InsB:15–23 binding core with varied flanking residues (grey shaded). **b**, Clone 58 also reacted with the InsB:12–26 peptide, an MHCII-bound sequence identified in the spleen peptidome. Mutation of the G23 into R23, an inhibitory residue, nullified the response, suggesting that the G23 was the P9 anchoring residue. Clone 58 is unreactive to InsB:12–20 or InsB:13–21. **c**, Comparison of T cell (clone 58) recognition between InsB:11–25 and peptides with varied residues flanking the InsB:15–23 register. The data showed the importance of the flanking residues in T cell recognition. The InsB:15–23 segment without any flanking residues did not induce any responses. Reducing, removing, or mutating the flanking residues at either the amino or carboxy end compromised the recognition to different extents. We note the role of having hydrophobic residues an the carboxy flank: changing the native FF residues into hydrophobic WW preserved the responses. Data (**b,c**) are representative of n = 2 independent experiments with similar results.



Extended Data Fig. 4. Analysis of the immunogenic segment in InsC peptides.

a, ELISPOT assay showing IFN-γ and IL-2 production by $CD4^+$ T cells from 8-week old male NOD mice immunized with two Ins1C peptides lacking the complete C-terminus upon recalling with indicated relevant peptides. Responses to either Ins1C:41–55 (left) or Ins1C:37–56 (right) were indistinguishable from background (no antigen); the positive control (ConA) generated strong responses. Data (mean ± s.e.m.) summarize results from n = 3 independent experiments from n = 6 mice. **b**, The mass spectrometry spectrum of the synthetic citrullinated Ins1C:53–61 peptide (TLALEVArQ; the lowercase r indicates the presence of citrulline). This spectrum was distinct from the deamidated Ins1C:53–61E peptide found in the islet MHCII peptidome, confirming that the biologically induced PTM was deamidation but not citrullination.



Extended Data Fig. 5. An example of a false positive MHCII-bound HIP.

Mass spectrum of a putative Ins1C-ChgA (PQVEQLEL-WSRMDQLAK) peptide in the islet MHCII peptidome (upper) that failed to match the synthetic standard peptide (middle). Further analysis suggested a probable correct match of the putative Ins1C-ChgA HIP to an Ins1C peptide fragment with sodium adduct (lower).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Members of the Unanue laboratory provided much advice on many aspects of this project. We thank K. Frederick, for maintaining the animal colony; P. Zakharov, for assistance in prediction of the MHCII epitopes; and M. Gross for advice on mass spectrometry. This study is supported by grants from the National Institute of Health

(DK120340, DK058177, and AI114551), the Juvenile Diabetes Research Foundation and the Kilo Diabetes and Vascular Research Foundation.

References

1. Hattori M et al. The NOD mouse: recessive diabetogenic gene in the major histocompatibility complex. Science 231, 733–735 (1986). [PubMed: 3003909]
2. Todd JA, Bell JI & McDevitt HO HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 329, 599–604 (1987). [PubMed: 3309680]
3. Acha-Orbea H & McDevitt HO The first external domain of the nonobese diabetic mouse class II I-A beta chain is unique. Proc. Natl. Acad. Sci. U.S.A 84, 2435–2439 (1987). [PubMed: 2882518]
4. Miyazaki T et al. Direct evidence for the contribution of the unique I-ANOD to the development of insulitis in non-obese diabetic mice. Nature 345, 722–724 (1990). [PubMed: 2113614]
5. Corper AL et al. A structural framework for deciphering the link between I-Ag7 and autoimmune diabetes. Science 288, 505–511 (2000). [PubMed: 10775108]
6. Latek RR et al. Structural basis of peptide binding and presentation by the type I diabetes-associated MHC class II molecule of NOD mice. Immunity 12, 699–710 (2000). [PubMed: 10894169]
7. Lee KH, Wucherpfennig KW & Wiley DC Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. Nat. Immunol 2, 501–507 (2001). [PubMed: 11376336]
8. Suri A, Walters JJ, Gross ML & Unanue ER Natural peptides selected by diabetogenic DQ8 and murine I-A(g7) molecules show common sequence specificity. J. Clin. Invest 115, 2268–2276 (2005). [PubMed: 16075062]
9. Lund T et al. Prevention of insulin-dependent diabetes mellitus in non-obese diabetic mice by transgenes encoding modified I-A beta-chain or normal I-E alpha-chain. Nature 345, 727–729 (1990). [PubMed: 2163026]
10. Singer SM et al. Prevention of diabetes in NOD mice by a mutated I-Ab transgene. Diabetes 47, 1570–1577 (1998). [PubMed: 9753294]
11. Gioia L et al. Position β57 of I-A ^{g7} controls early anti-insulin responses in NOD mice, linking an MHC susceptibility allele to type 1 diabetes onset. Science Immunology 4, eaaw6329 (2019). [PubMed: 31471352]
12. 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). [PubMed: 15889095]
13. Daniel D, Gill RG, Schloot N & Wegmann D Epitope specificity, cytokine production profile and diabetogenic activity of insulin-specific T cell clones isolated from NOD mice. Eur. J. Immunol 25, 1056–1062 (1995). [PubMed: 7537670]
14. Wegmann DR, Norbury-Glaser M & Daniel D Insulin-specific T cells are a predominant component of islet infiltrates in pre-diabetic NOD mice. Eur. J. Immunol 24, 1853–1857 (1994). [PubMed: 8056042]
15. Halbout P, Briand J-P, Bécourt C, Muller S & Boitard C T cell response to preproinsulin I and II in the nonobese diabetic mouse. J. Immunol 169, 2436–2443 (2002). [PubMed: 12193712]
16. Levisetti MG, Lewis DM, Suri A & Unanue ER Weak proinsulin peptide-major histocompatibility complexes are targeted in autoimmune diabetes in mice. Diabetes 57, 1852–1860 (2008). [PubMed: 18398138]
17. Mohan JF et al. Unique autoreactive T cells recognize insulin peptides generated within the islets of Langerhans in autoimmune diabetes. Nat. Immunol 11, 350–354 (2010). [PubMed: 20190756]
18. Ziegler AG et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. JAMA 309, 2473–2479 (2013). [PubMed: 23780460]
19. Babon JAB et al. Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. Nat. Med 22, 1482–1487 (2016). [PubMed: 27798614]
20. Michels AW et al. Islet-Derived CD4 T Cells Targeting Proinsulin in Human Autoimmune Diabetes. Diabetes 66, 722–734 (2017). [PubMed: 27920090]
21. Wan X et al. Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. Nature 560, 107–111 (2018). [PubMed: 30022165]

22. Unanue ER Antigen presentation in the autoimmune diabetes of the NOD mouse. Annu. Rev. Immunol 32, 579–608 (2014). [PubMed: 24499272]
23. Unanue ER, Ferris ST & Carrero JA The role of islet antigen presenting cells and the presentation of insulin in the initiation of autoimmune diabetes in the NOD mouse. Immunol. Rev 272, 183–201 (2016). [PubMed: 27319351]
24. Andreatta M, Alvarez B & Nielsen M GibbsCluster: unsupervised clustering and alignment of peptide sequences. Nucleic Acids Res 45, W458–W463 (2017). [PubMed: 28407089]
25. Mohan JF, Petzold SJ & Unanue ER Register shifting of an insulin peptide-MHC complex allows diabetogenic T cells to escape thymic deletion. J. Exp. Med 208, 2375–2383 (2011). [PubMed: 22065673]
26. Wiles TA et al. An insulin-IAPP hybrid peptide is an endogenous antigen for CD4 T cells in the non-obese diabetic mouse. J. Autoimmun 78, 11–18 (2017). [PubMed: 27802879]
27. Delong T et al. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. Science 351, 711–714 (2016). [PubMed: 26912858]
28. Gagnerault M-C, Luan JJ, Lotton C & Lepault F Pancreatic lymph nodes are required for priming of beta cell reactive T cells in NOD mice. J. Exp. Med 196, 369–377 (2002). [PubMed: 12163565]
29. Levisetti MG, Suri A, Frederick K & Unanue ER Absence of lymph nodes in NOD mice treated with lymphotoxin-beta receptor immunoglobulin protects from diabetes. Diabetes 53, 3115–3119 (2004). [PubMed: 15561941]
30. Faridi P, Purcell AW & Croft NP In Immunopeptidomics We Need a Sniper Instead of a Shotgun. Proteomics 18, e1700464 (2018). [PubMed: 29377634]
31. Babad J, Geliebter A & DiLorenzo TP T-cell autoantigens in the non-obese diabetic mouse model of autoimmune diabetes. Immunology 131, 459–465 (2010). [PubMed: 21039471]
32. Mohan JF, Calderon B, Anderson MS & Unanue ER Pathogenic $CD4^+$ T cells recognizing an unstable peptide of insulin are directly recruited into islets bypassing local lymph nodes. J. Exp. Med 210, 2403–2414 (2013). [PubMed: 24127484]
33. Chang KY & Unanue ER Prediction of HLA-DQ8beta cell peptidome using a computational program and its relationship to autoreactive T cells. Int. Immunol 21, 705–713 (2009). [PubMed: 19461125]
34. Aaron Wiles T et al. Identification of Hybrid Insulin Peptides (HIPs) in Mouse and Human Islets by Mass Spectrometry. J. Proteome Res. 18, 814–825 (2019). [PubMed: 30585061]
35. Han X, He L, Xin L, Shan B & Ma B PeaksPTM: Mass spectrometry-based identification of peptides with unspecified modifications. J. Proteome Res 10, 2930–2936 (2011). [PubMed: 21609001]
36. Han Y, Ma B & Zhang K SPIDER: software for protein identification from sequence tags with de novo sequencing error. J Bioinform Comput Biol 3, 697–716 (2005). [PubMed: 16108090]
37. Berkers CR, de Jong A, Ovaa H & Rodenko B Transpeptidation and reverse proteolysis and their consequences for immunity. Int. J. Biochem. Cell Biol 41, 66–71 (2009). [PubMed: 18817889]
38. Ebstein F et al. Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes. Sci Rep 6, 24032 (2016). [PubMed: 27049119]
39. Mishto M & Liepe J Post-Translational Peptide Splicing and T Cell Responses. Trends Immunol. 38, 904–915 (2017). [PubMed: 28830734]
40. Mohan JF & Unanue ER Unconventional recognition of peptides by T cells and the implications for autoimmunity. Nat. Rev. Immunol 12, 721–728 (2012). [PubMed: 22955843]
41. van Lummel M et al. Posttranslational modification of HLA-DQ binding islet autoantigens in type 1 diabetes. Diabetes 63, 237–247 (2014). [PubMed: 24089515]
42. Marre ML et al. Modifying Enzymes Are Elicited by ER Stress, Generating Epitopes That Are Selectively Recognized by CD4+ T Cells in Patients With Type 1 Diabetes. Diabetes 67, 1356–1368 (2018). [PubMed: 29654212]
43. Wright HT Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. Crit. Rev. Biochem. Mol. Biol 26, 1–52 (1991). [PubMed: 1678690]
44. Robinson NE et al. Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides. J. Pept. Res 63, 426–436 (2004). [PubMed: 15140160]

45. Vader LW et al. Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. J. Exp. Med 195, 643–649 (2002). [PubMed: 11877487]

46. Tisch R et al. Immune response to glutamic acid decarboxylase correlates with insulitis in non-obese diabetic mice. Nature 366, 72–75 (1993). [PubMed: 8232539]

47. Kelemen K, Wegmann DR & Hutton JC T-cell epitope analysis on the autoantigen phogrin (IA-2beta) in the nonobese diabetic mouse. Diabetes 50, 1729–1734 (2001). [PubMed: 11473031]

48. Tai N et al. Microbial antigen mimics activate diabetogenic CD8 T cells in NOD mice. J. Exp. Med 213, 2129–2146 (2016). [PubMed: 27621416]

49. Hebbandi Nanjundappa R et al. A Gut Microbial Mimic that Hijacks Diabetogenic Autoreactivity to Suppress Colitis. Cell 171, 655–667.e17 (2017). [PubMed: 29053971]

50. So M et al. Proinsulin C-peptide is an autoantigen in people with type 1 diabetes. Proc. Natl. Acad. Sci. U.S.A 115, 10732–10737 (2018). [PubMed: 30275329]

51. Sofron A, Ritz D, Neri D & Fugmann T High-resolution analysis of the murine MHC class II immunopeptidome. Eur. J. Immunol 46, 319–328 (2016). [PubMed: 26495903]

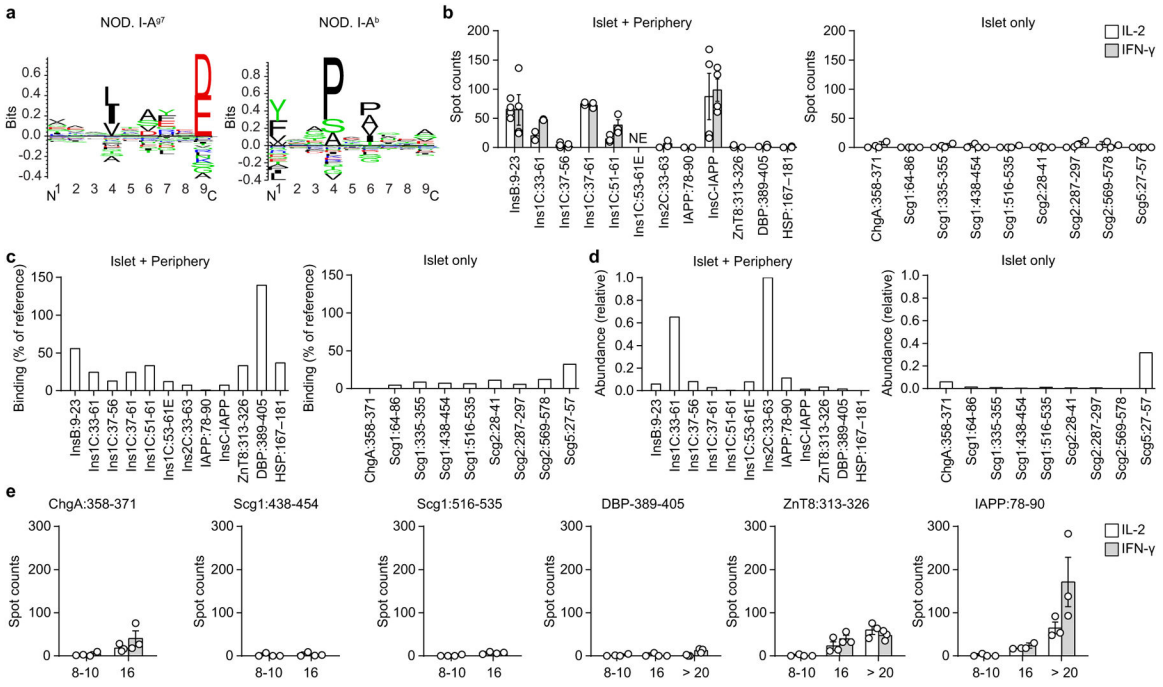


Figure 1. Analysis of the β-cell-derived peptides identified in the MHCII peptidome.

a, Epitope mapping by Gibbs cluster analysis of all the MHCII-bound peptides from the spleen of 8–10-week old NOD.I-A^{g7} (left) and NOD.I-A^{b} (right) mice. **b-d**, Individual β-cell-derived peptides identified in the MHCII peptidomes of the pancreatic islets and the peripheral lymphoid tissues (Islet+periphery; left) or in the islets alone (Islet only; right) were examined for their corresponding T cell autoreactivity (**b**), relative MHCII binding affinity (**c**) and relative abundance (**d**). **b**, ELISPOT assay showing IL-2 and IFN-γ production by CD4^{+} T cells from islets and pLN in 8–10-week old female NOD mice upon challenging and recalling with the indicated peptides. NE, not examined. Data (mean ± s.e.m.) are from n = 5 mice per independent experiment (each data point). **c**, Competitive binding assay showing the relative MHCII binding affinity of the indicated peptides. Data (mean) are from n = 2 independent experiments with similar results. **d**, Abundance (normalized peak area) of the indicated peptides relative to the most abundant peptide (Ins2C:33–63). Data (mean) are from the mass spectrometry analysis of the MHCII peptidome of the pancreatic islets isolated from n = 219 female NOD mice of 8–10-week of age. **e**, ELISPOT assay showing IL-2 and IFN-γ production by CD4^{+} T cells from islets and pLN in female NOD mice of different ages upon challenging and recalling with the indicated peptides. Data (mean ± s.e.m.) are from n = 5 mice per independent experiment (each data point).

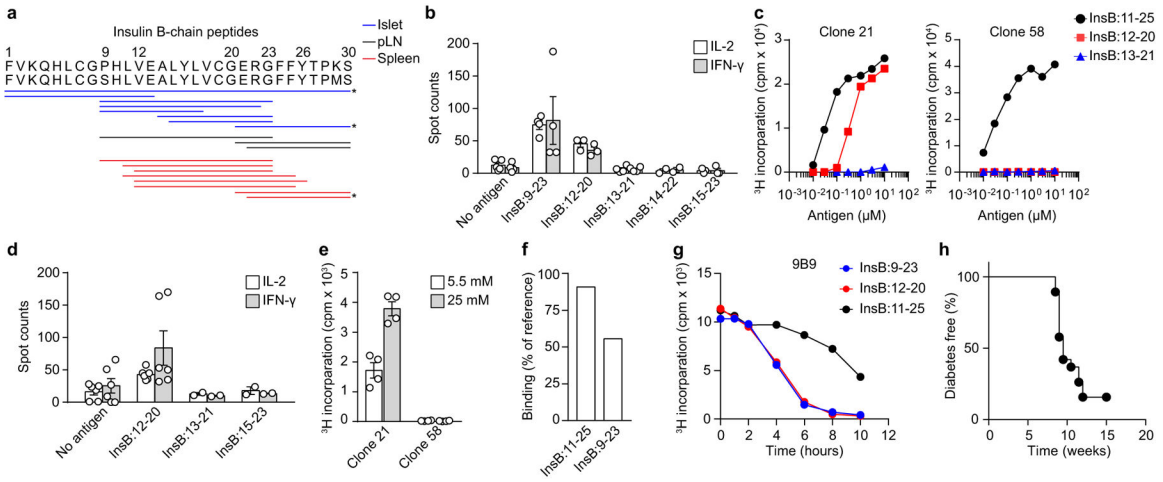


Figure 2. Diverse InsB-derived peptides induce diabetogenic CD4$^+$ T cells.

a, Summary of the InsB-derived peptides identified in the MHCII peptidome at the indicated sites. The peptides are aligned to the sequence of insulin-1 (top) or insulin-2 (bottom) B-chain. The asterisks denote insulin-1 B-chain peptides. **b**, ELISPOT assay showing IL-2 and IFN-γ production by CD4$^+$ T cells from islets and pLN in 8–10-week old female NOD mice upon challenging with InsB:9–23 and recalling with the indicated epitopes. Data (mean ± s.e.m.) are from n = 5 mice per independent experiment (each data point). **c**, Antigen presentation assay showing responses of two representative InsB:11–25-reactive CD4$^+$ T cell hybridomas, clone 21 (left) and clone 58 (right), to C3g7 APCs pulsed with the indicated epitopes. Data (mean) are representative n = 3 independent experiments with similar results. **d**, ELISPOT assay showing IL-2 and IFN-γ production by CD4$^+$ T cells from islets and pLN in 8–10-week old female NOD mice upon challenging with InsB:11–25 and recalling with indicated epitopes. Data (mean ± s.e.m.) are from n = 5 mice per independent experiment (each data point). **e**, Antigen presentation assay showing responses of the clone 21 and clone 58 CD4$^+$ T cell hybridomas to dispersed islet cells in cultures with indicated glucose concentrations. Data (mean ± s.e.m.) are from n = 4 independent experiments. **f**, Competitive binding assay showing the relative MHCII binding affinity of InsB:9–23 versus InsB:11–25. Data (mean) are from n = 2 independent experiments with similar results. **g**, Antigen presentation assay showing the responses of the InsB:12–20-specific CD4$^+$ T cell hybridoma (9B9) to C3g7 APCs briefly pulsed with the indicated peptides followed by extensive washes at the indicated time points. Data are representative of n = 3 independent experiments with similar results. **h**, Diabetes incidence of NOD.*Rag1*$^{-/-}$ mice transferred with primary CD4$^+$ T cell lines to InsB:11–25. Data are from n = 2 independent experiments from n = 19 mice.

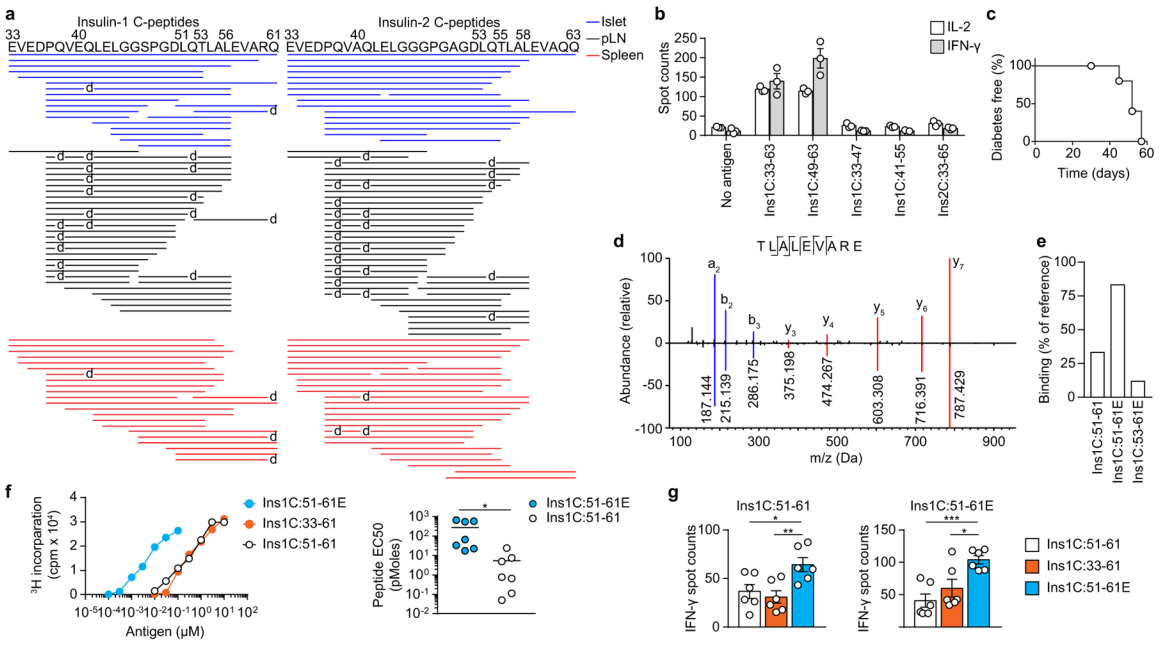


Figure 3. Pathogenic CD4$^+$ T cells recognize native and deamidated epitopes included in Ins1C-derived peptides.

a, Summary of the peptide sequences derived from Ins1C (left) or Ins2C (right) identified in the MHCII peptidome at the indicated sites. The lower case d denotes the site of deamidation. **b**, ELISPOT assay showing IL-2 and IFN-γ production by CD4$^+$ T cells from 8-week old male NOD mice immunized with a long Ins1C:33–63 peptide and recalling with indicated segments. Data (mean ± s.e.m.) are from n = 6 mice from n = 3 independent experiments (each data point). **c**, Diabetes incidence of NOD.*Rag1*$^{-/-}$ mice transferred with a primary CD4$^+$ T cell line specific to Ins1C:49–63. Data are from n = 2 independent experiments from n = 10 mice. **d**, A mirror plot showing the match of the deamidated Ins1C:53–61E peptide identified in the islet MHCII peptidome (upper) with the synthetic version (lower). **e**, Competitive binding assay showing the relative MHCII binding affinity of the native Ins1C:51–61 versus the deamidated Ins1C:51–61E and Ins1C:53–61E. Data (mean) are from n = 2 independent experiments with similar results. **f**, Antigen presentation assay showing responses of a representative Ins1C:51–61E-reactive CD4$^+$ T cell clone to C3g7 APCs pulsed with indicated peptides (left). The dot plot (right) shows the amounts of the peptides required for reaching the half-maximal responses (Peptide EC50). Data are from n = 7 clones (each data point) tested in n = 4 independent experiments. P values: *P = 0.0358. **g**, ELISPOT assay showing IL-2 and IFN-γ production by CD4$^+$ T cells from islets and pLN in 8–10-week old female NOD mice upon challenging with Ins1C:51–61 (left) or Ins1C:51–61E (right) and recalling with indicated peptides. Data (mean ± s.e.m.) are from n = 25 mice from n = 5 independent experiments (each data point). P values for Ins1C:51–61: *P = 0.0121; **P = 0.0062. P values for Ins1C:51–61E: *P = 0.0156; ***P = 0.0004. P values (**f**, **g**) were calculated using an unpaired two-tailed Student's t-test.

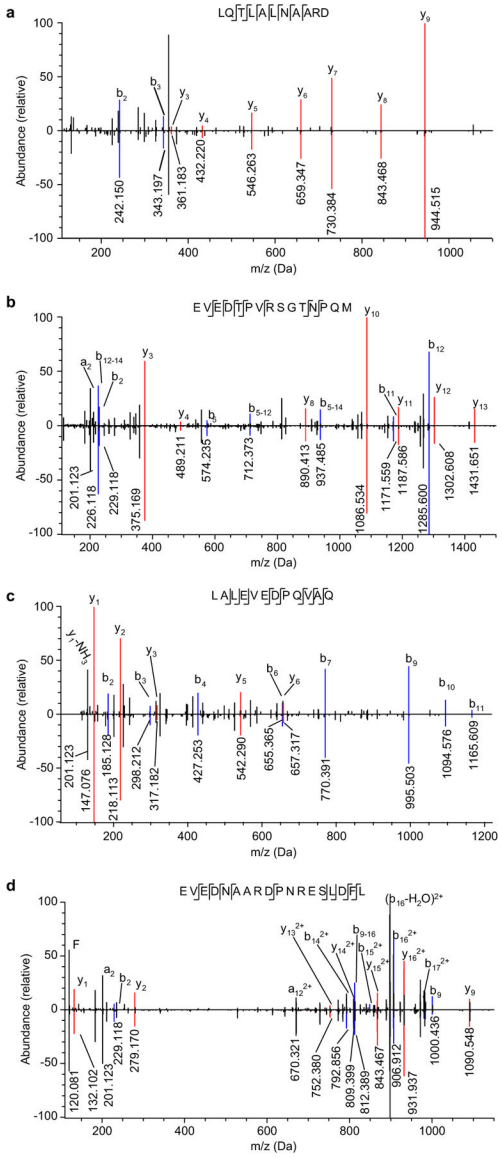


Figure 4. Validation of the potential HIPs identified in the MHCII peptidomes and the free HIPs shared between the crinosomes and the secretome.

a, A mirror plot showing the match of an InsC-IAPP HIP (LQTLAL-NAARD; upper) identified in the islet MHCII peptidome with the synthetic standard (lower). An identical spectrum of this HIP was identified in the MHCII peptidome of the pLN. **b-d**, Three potential free HIPs were identified both in the crinosomes and the secretome. **b**, A mirror plot showing the match of an InsC-IAPP HIP (EVED-TPVRSGTNPQM; upper) to the synthetic standard (lower). **c**, A mirror plot showing the match of an InsC-Ins2C HIP (LAL-EVEDPQVAQ; upper) to the synthetic standard (lower). **d**, A mirror plot showing the match of an InsC-IAPP HIP (EVED-NAARDPNRESLDFL; upper) to the synthetic standard (lower). The spectra in **b-d** are from the secretome; identical spectra were observed in the crinosomes.

Table 1.
β-cell-derived peptide families identified in the MHCII peptidomes.

The β-cell-derived peptide families are identified in the MHCII peptidome (blue) of the islets, pancreatic lymph nodes (pLN), spleens and $CD19^+$ splenic B cells isolated from 8–10-week old female NOD mice, as well as from the spleens of 8–10-week old female B cell-deficient (μMT) NOD mice. The peptide families are ranked according to their relative abundance (accumulative peak area of all the peptides in a given family normalized to the most abundant family) in the islet MHCII peptidome. The number of the individual peptide sequences belonging to each family is shown. Data are from n = 219 mice from n = 1 experiment (islets); n = 257 mice from n = 2 independent experiments (pLN); n = 16 mice from n = 4 independent experiments (spleen); n = 4 mice from n= 1 experiment (B cells); and n = 4 mice from n = 1 experiment (μMT). The free peptides (yellow) are identified in the insulin dense core granules (DCGs) and crinosomes, as well as in β-cell secretion upon glucose challenge (secretome). Data are from reanalysis of previously published datasets (ref.[21]).

Description	Protein coverage	Relative abundance (Islets)	MHC-II-bound peptides					Free peptides		
			Islets	pLN	Spleen			DCG	Crinosome	Secretome
					Whole	B cells	μMT			
Insulin-2 C-peptide	Ins2C:33–63	1.00	16	38	23	31	41	129	160	227
Insulin-1 C-peptide	Ins1C:33–61	0.82	13 #	30	16	7	20 #	108	133	182
Neuroendocrine protein 7B2	Scg5:27–57	0.55	8					4	13	31
Islet amyloid polypeptide	IAPP:78–93	0.120	13 #	3 #			7 #	23	32	36
Insulin-2 B-chain	Ins2B:1–23	0.030	7 # *	4 # *	3 (1) #	5 (2) #	7 (3) #	3	14 (13)	28 (22)
Chromogranin-A	Chga:358–371	0.027	2					27	32	34
Insulin-1 B-chain	Ins1B:1–30	0.017	2		2 (0) #	3(0) #	5 (1) #	2	6 (5)	18 (12)
Zinc transporter 8	ZnT8:313–322	0.011	2	1		1	1			
Chromogranin-A	Chga:393–402	0.006	2					6	11	12
Chromogranin-A	Chga:19–29	0.005	2 #						1	
Vitamin D-binding protein	DBP:389–405	0.005	1	11	5					
Hybrid InsC-IAPP	InsC-IAPP	0.004	1	1						
Secretogranin-1	Scg1:64–86	0.004	1					1	4	11
Secretogranin-1	Scg1:516–535	0.003	1					1	11	16
Secretogranin-1	Scg1:333–355	0.003	1						2	3
Secretogranin-2	Scg2:28–41	0.002	1					1	9	9
Secretogranin-2	Scg2:184–216	0.001	1					7	19	40
60 kD heat shock protein	HSP:167–181	0.001	2 #	12	13	10	2			
Secretogranin-2	Scg2:569–610	<0.001	1					10	28	57

* One or more peptides contained a C19S substitution

One or more peptides that did not pass 1% FDR cutoff but were manually verified

$0_{\text{The number of unique sequences specifically belonging to the insulin-1 or insulin-2 B-chain family}}$

Table 2:
Representative β-cell-derived peptides identified in the MHC-II peptidomes.

Shown are representative peptide sequences from each peptide families depicted in Table 1. The peptides are identified in the MHCII peptidome of both islets and peripheral lymphoid tissues (Islet + Periphery; blue) or only in the islets (Islet only; yellow).

Sites	Peptide	Sequence
Islet + Periphery	InsB:9–23	SHLVEALYLVCGERG
	Ins1C:33–61	EVEDPQVEQLELGGSPGDLQTLALEVARQ
	Ins1C:37–56	PQVEQLELGGSPGDLQTLAL
	Ins1C:37–61	PQVEQLELGGSPGDLQTLALEVARQ
	Ins1C:51–61	LQTLALEVARQ
	Ins1C:53–61E	TLALEVARE
	Ins2C:33–63	EVEDPQVAQLELGGGPGAGDLQTLALEVAQQ
	IAPP:78–90	NAARDPNRESLDF
	InsC-IAPP	LQTLALNAARD
	ZnT8:313–326	ILSVHVATAASQDS
	DBP:389–405	SPLLKRQLTSFIEKGQE
	HSP:167–181	EIAQVATISANGDKD
Islet only	ChgA:358–371	WSRMDQLAKELTAEKR
	Scg1:64–86	SGKEVKGEEKGENQNSKFEVRLL
	Scg1:335–355	SHHLAHYRASEEEPEYGEESRSY
	Scg1:438–454	LLDEGHYPVRESPIDTA
	Scg1:516–535	LGALFNPYFDPLQWKNSDFE
	Scg2:28–41	ASFQRNQLLQKEPD
	Scg2:287–297	SGQLGLPDEEN
	Scg2:569–578	IPVGSLKNED
	Scg5:27–57	YSPRTPDRVSETDIQRLLHGVMEQLGIARPR

Table 3.

Putative MHCII-bound HIP sequences obtained from the initial database search.

One HIP sequence was verified by matching to a synthetic standard (blue). For four peptides the spectra of the synthetic standards did not match the biological samples (yellow), a probable correct assignment is given.

Source	Putative HIP sequence	Probable correct assignment	Match to synthetic standard
Islets, pLN	LQTLAL-NAARD	NA	Yes
Islets, pLN	PQVEQLELGGSPGDLQ-APV	PQVEQLELGGSPGDLQT(-H_2O)LA	no
Islets	PQVEQLEL-WSRMDQLAK	PQVEQLELGGSPGD(+Na)LQTLAL	no
pLN	PQVEQLELGGSP-APVDN	PQVEQLELGGSPGDLQt(-H_2O)	no
pLN	PQVEQLEL-GNP	PQVEQLELGGGP	no