AUTOIMMUNITY

Position β57 of I-A^{g7} controls early anti-insulin responses in NOD mice, linking an MHC susceptibility allele to type 1 diabetes onset

Louis Gioia[1]*, Marie Holt[2]*, Anne Costanzo[2]*, Siddhartha Sharma[2]*, Brian Abe[2†], Lisa Kain[2], Maki Nakayama[3], Xiaoxiao Wan[4], Andrew Su[1], Clayton Mathews[5], Yi-Guang Chen[6], Emil Unanue[4], Luc Teyton[2‡]

The class II region of the major histocompatibility complex (MHC) locus is the main contributor to the genetic susceptibility to type 1 diabetes (T1D). The loss of an aspartic acid at position 57 of diabetogenic HLA-DQβ chains supports this association; this single amino acid change influences how TCRs recognize peptides in the context of HLA-DQ8 and I-A^{g7} using a mechanism termed the P9 switch. Here, we built register-specific insulin peptide MHC tetramers to examine CD4$^+$ T cell responses to Ins$_{12-20}$ and Ins$_{13-21}$ peptides during the early prediabetic phase of disease in nonobese diabetic (NOD) mice. A single-cell analysis of anti-insulin CD4$^+$ T cells performed in 6- and 12-week-old NOD mice revealed tissue-specific gene expression signatures. TCR signaling and clonal expansion were found only in the islets of Langerhans and produced either classical T$_H$1 differentiation or an unusual T$_{reg}$ phenotype, independent of TCR usage. The early phase of the anti-insulin response was dominated by T cells specific for Ins$_{12-20}$, the register that supports a P9 switch mode of recognition. The presence of the P9 switch was demonstrated by TCR sequencing, reexpression, mutagenesis, and functional testing of TCRαβ pairs in vitro. Genetic correction of the I-Aβ57 mutation in NOD mice resulted in the disappearance of D/E residues in the CDR3β of anti-Ins$_{12-20}$ T cells. These results provide a mechanistic molecular explanation that links the characteristic MHC class II polymorphism of T1D with the recognition of islet autoantigens and disease onset.

INTRODUCTION

The association between HLA genes and autoimmune diseases was uncovered more than 40 years ago (1). Among them were type 1 diabetes (T1D) (2) and a linkage to human leukocyte antigen–DR3 (HLA-DR3) and HLA-DR4 that explained the vast majority of the genetic component of this serious disease (3, 4) in which the exclusive destruction of the β cells of the islets of Langerhans of the endocrine pancreas leads to a lifelong dependency on insulin replacement therapy. The linkage to two HLA-DR haplotypes was later redefined as an association with the HLA-DQ haplotypes that segregate with these HLA-DR genes, HLA-DQ2 and HLA-DQ8 for HLA-DR3 and HLA-DR4, respectively. Relative risk is higher for HLA-DQ2 and HLA-DQ8 homozygotes than for heterozygotes and maximal for HLA-DQ2/HLA-DQ8 heterozygotes (5). In 1987, McDevitt's group made the important observation that every HLA class II associated with T1D was carrying a distinct polymorphism at position 57 of the β chain that substituted the normal aspartic acid of all major histocompatibility complex (MHC) class II β chains at this position by a neutral residue (6). This notable observation has been confirmed since in one of the largest genetic studies of T1D (7). Structurally

speaking, the consequences of this alteration are the loss of a salt bridge with the arginine 76 of the MHC class II α chain and the appearance of a surface-exposed positively charged patch that modifies both the P9 pocket and potential T cell receptor (TCR) contacts (8, 9). We, and others, have shown that the loss of the salt bridge had no consequence on the structural integrity of the molecule (8, 10) and that it remained stable and competent for peptide binding. As expected, the modifications of surface charges at the P9 pocket affect peptide binding profoundly, and the peptide repertoire of diabetogenic MHC class II molecules is heavily biased toward the selection of peptides with acidic residues at P9 (11, 12). However, similar to most I-A and HLA-DQ molecules, diabetogenic MHC class II proteins remain very promiscuous for peptide binding as they interact mainly with the peptide backbone instead of using anchor residues (13, 14). The consequence of this mode of binding is that all MHC class II molecules without Aspβ57 can also efficiently bind peptides that do not have a negatively charged residue at the P9 position (8, 14). In this case, a large positively charged patch remains surface-exposed and potentially accessible to T cell recognition. We have evaluated this situation by immunizing nonobese diabetic (NOD) mice and HLA-DQ8 transgenic NOD mice with peptides carrying or lacking a negatively charged residue at the P9 position. In both instances, we have shown that the absence of charge at that position in the peptide resulted in the selection of TCRs that encoded either an Asp or a Glu residue at position 2 or 3 of their complementarity-determining region 3β (CDR3β) (9, 15). For one of those peptides derived from hen egg lysozyme (HEL) that has a glycine at P9, we also demonstrated biophysically and structurally that the presence of a negatively charged residue at position 2 or 3 of the CDR3β increased the affinity of the TCR for its cognate peptide-MHC (pMHC) complex by more than 30-fold (9). We called this mode of TCR

recognition the "P9 switch" and suggested that it might be important in the initiation of diseases such as celiac disease and T1D, which are tightly associated with non-Aspβ57 MHC class II molecules. However, in the absence of important reagents such as appropriate MHC tetramers and/or animal models (e.g., a celiac disease mouse model), we could not formerly test our hypothesis. The development of I-A^{g7} tetramers capable of distinguishing the two main registers of the Ins$_{9-23}$ peptide (16) was the first important step in testing the relevance of the P9 switch model in T1D because one of them, Ins$_{12-20}$, has a glycine at P9, whereas the other one, Ins$_{13-21}$, bears a glutamic acid at that position. Recognition of insulin requires its processing in intracellular compartments of antigen-presenting cells (APCs), resulting in the presentation of the Ins$_{13-21}$ peptide segment. T cells that recognize the Ins$_{13-21}$ peptide either from such processing of insulin or from exogenous peptides represent type A T cells reliant on conventional antigen presentation. In contrast, other peptide segments, such as Ins$_{12-20}$, are eliminated by the intracellular processing of insulin and are therefore not normally presented by APCs. The Ins$_{12-20}$ peptide segment can be presented when supplied as an exogenous peptide and is recognized by type B T cells. This alternate mode of presentation takes place when β cell crinosomes partially catabolize insulin molecules and release a number of insulin peptides that are then presented in islets and peripheral lymph nodes (17).

In any instance, if the P9 switch model were correct, one would expect three simultaneous findings: (i) CD4$^+$ T cells specific for pMHC with neutral P9 residues should be preferentially expanded; (ii) these T cells should be activated earlier than cells recognizing pMHC with Asp/Glu (D/E) at P9; (iii) these T cells should rely on coulombic forces for interaction. We have already shown that cells specific for a BDC2.5 mimotope with a neutral P9 position were preferentially expanded in the NOD mouse and that those cells had higher affinity for their pMHC than cells specific for BDC2.5 peptides with Asp at the P9 position (9). We now demonstrate that the nature of the amino acid at position β57 controls the selection and expansion of CD4$^+$ T cells specific for insulin and the onset of T1D in the mouse. Using single CD4$^+$ T cell analysis, we show that the tissue of origin of anti-insulin T cells could easily be identified on the basis of characteristic gene expression signatures. Hallmarks of TCR signaling and T cell expansion appeared to be localized exclusively to the islet rather than the pancreatic lymph nodes (PLNs). Although islet-infiltrating T cells were either classic T helper 1 (T$_H$1) or regulatory T cells (T$_{regs}$), this dichotomy was not associated with particular TCRαβ pairs. The early T cell response seen in the islets was dominated by anti-Ins$_{12-20}$ cells. A majority of these cells carried a D/E residue in their CDR3β and used a P9 switch for recognition. The restitution of an aspartic acid at position β57 of I-A^{g7} in vivo eliminated this population of cells. These results provide a mechanistic molecular explanation of why and how diabetogenic MHC class II molecules increase the risk of autoimmunity.

RESULTS
MHC class II tetramers specific for the two registers of Ins$_{9-23}$
The expression of functional MHC class II tetramers remains difficult especially when multiple registers can be displayed within short peptides. The length and nature of the linker when expressing tethered peptides, the on and off rates of the peptides themselves, and the potential role of flanking residues are all factors to consider when expressing MHC class II/peptide complexes. In addition, in many instances, expression does not mean proper display as many recombinant MHC class II molecules are unable to activate and/or bind T cells that are specific for the cognate peptide (18). We have successfully expressed the various registers of ovalbumin in I-A^d and I-A^b using a short seven-residue linker (19) but failed to express functional Ins$_{12-20}$ and Ins$_{13-21}$ peptides in I-A^{g7} using the same approach. The reengineering of both peptide and MHC to introduce artificial disulfide bridges, as proposed by some (20, 21), led to expression of nonfunctional molecules (i.e., no stimulation of T cell hybridomas and no T cell staining ex vivo). We then applied various modifications at anchor residues or N- and C-terminal flanking residues (P1 and P10) with the idea to limit register sliding or inadequate display; none was successful (table S1). Last, we explored the ability of various sequences such as polyglycine stretches to disrupt type II polyproline helices (22), the normal required secondary structure of peptides for binding to MHC class II molecules. The addition of three C-terminal glycines allowed the expression of Ins$_{13-21}$, whereas the addition of four glycines allowed the expression of both Ins$_{12-20}$ and Ins$_{13-21}$ (i.e., VEALYLVCG-GGGG and EALYLVCGE-GGGG, respectively). The specificity and functionality of the recombinant molecules were tested by staining and stimulating T cell clones specific for each register (Fig. 1). Accordingly, type A cells of which IIT3 is a representative (23) were stained and stimulated specifically by recombinant I-A^{g7} Ins$_{13-21}$ molecules, whereas type B cells such as 2D10 were stained and responsive to I-A^{g7} Ins$_{12-20}$ molecules.

Ex vivo detection of insulin-reactive CD4$^+$ T cells in naïve NOD mice
After this characterization, the MHC tetramers were tested ex vivo to quantify the CD4$^+$ T cells from the thymus, spleen, and PLNs, and islet infiltration from NOD mice of different ages. The specificity of the reagents and the existence of two distinct populations of cells were confirmed by double staining (Fig. 1, A and B). In all cases, insulin-reactive cells were detected in the thymus and peripheral lymph nodes; consistently, Ins$_{12-20}$ cells were twice as numerous as Ins$_{13-21}$ cells in prediabetic mice, whereas both populations, at difference with BDC2.5 cells (24), showed no notable expansion between the thymus and PLNs (Fig. 1C). However, in the islets at 6 weeks of age, the Ins$_{12-20}$ cells showed a 200- to 400-fold expansion, which contributed to most of the infiltrating CD4$^+$ population of a majority of mice (Fig. 2, A to C). The Ins$_{13-21}$ cells, although also more numerous in the islets than in PLNs, never represented more than 20% of the infiltrating CD4$^+$ population, and a pairwise analysis of each mouse showed that Ins$_{12-20}$ cells dominated the early phase of autoimmunity (Fig. 2C). Across all 6-week-old mice (n = 11), the mean representation of Ins$_{12-20}$-specific cells in the islets was 52.05% of all CD4$^+$ T cells (SEM = 8.0%), whereas Ins$_{13-21}$-specific cells had a mean representation of 11.26% (SEM = 1.96%), establishing the highly significant difference of frequencies between the two populations (P < 0.0001). The frequencies dropped to 3.31% (SEM = 1.42%) and 1.65% (SEM = 0.82%) for Ins$_{12-20}$- and Ins$_{13-21}$-specific cells, respectively, at 12 weeks of age, with no statistical difference of representation of either population (P = 0.352). These results indicated the likely importance of the P9 switch for early pathology but not for progression. This conclusion was reinforced by examining the anti-insulin response in overtly diabetic mice and seeing the same disappearance of the differential between the two
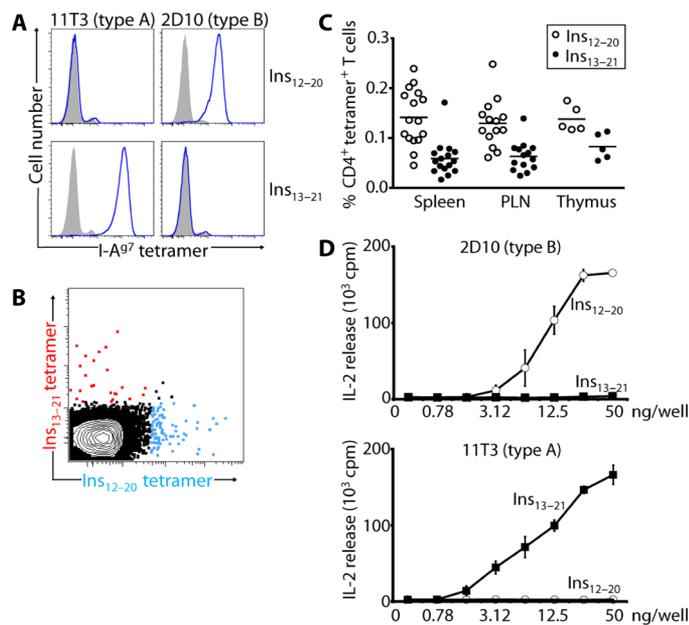
Fig. 1. Characterization of insulin peptide/I-A^{g7} tetramers and detection of register-specific CD4 cells in NOD mice. (**A**) Specific staining of type A (Ins$_{12-20}$ peptide) and type B (Ins$_{12-20}$ peptide) anti-insulin T cell hybridomas, of which 2D10 and 11T3 are representative members, respectively. (**B**) Double stain of NOD mouse splenocytes with I-A^{g7} tetramers containing Ins$_{12-20}$ peptide (blue) and Ins$_{13-21}$ peptide (red). Gated on CD3$^+$CD4$^+$ (CD8$^-$B220$^-$CD11b$^-$) population. (**C**) Enumeration of Ins$_{12-20}$- and Ins$_{13-21}$-specific CD4$^+$ T cells in the spleen, PLNs, and thymi of 8-week-old female NOD mice using register-specific tetramers. (**D**) Register-specific activation of type A and B hybridoma T cells using recombinant I-A^{g7}/peptide–coated wells. This result is representative of a minimum of five similar experiments. Each point is a biological triplicate (SDs are represented for each triplicate). cpm, counts per minute.

populations with no statistical differences in both the spleen ($P = 0.68$) and the PLNs ($P = 0.15$) (Fig. 2D).

Single-cell expression profiling of insulin-reactive CD4$^+$ T cells reveals organ residence signature and key steps of pathogenesis

Having appropriate reagents, anti-insulin register-specific T cells were sorted by cytometry as single cells in 96-well plates using a dump negative (CD8$^-$B220$^-$CD11b)$^-$CD3$^+$CD4$^+$tetramer$^+$ gate. Each cell was examined for the expression of a set of 96 genes surveying surface receptors, cytokines, signaling molecules, chemokines and chemokine receptors, transcription factors, and cytokine response genes (25). Profiles were established from individual 6- and 12-week-old mice and based on the analysis of 1580 individual cells, of which 501 were from the islets, 729 were from the PLNs, and 350 from the spleen; 967 cells were from 6-week-old mice ($n = 11$) and 613 cells were from 12-week-old mice ($n = 4$). With or without glyceraldehyde-3-phosphate dehydrogenase (GAPDH) normalization, unsupervised clustering could define six subsets. However, as illustrated in Fig. 3A, 1060 cells, besides CD3, CD4, and housekeeping genes, expressed less than 10 of the 96 chosen genes at detectable levels and needed to be removed from further statistical analysis to evaluate the differential expression between cells expressing larger numbers of genes. Accordingly, 87.2% of cells from the spleen, 66.4% from the PLNs, and 54% from the islet were removed from the unsupervised cluster analysis (Fig. 3B). In the spleen and PLNs, these numbers correlated with the recently
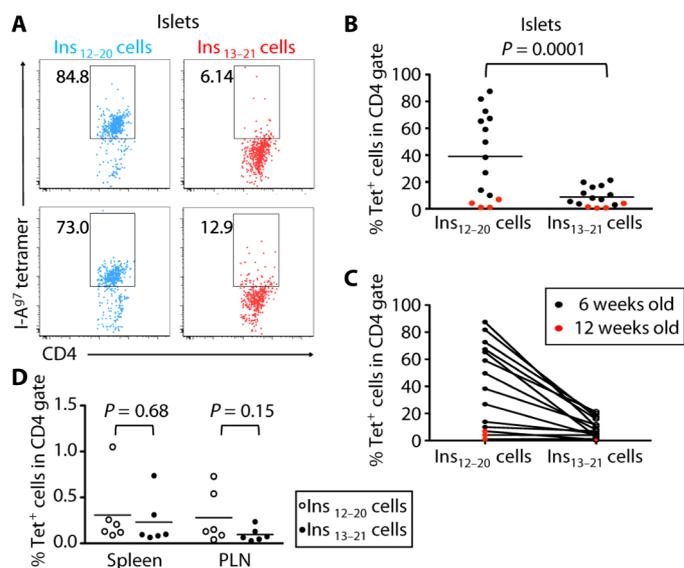
Fig. 2. Anti-insulin register-specific population dynamics in the islets during the course of disease in female NOD mice. (**A**) Two representative examples of Ins$_{12-20}$ and Ins$_{13-21}$ tetramer staining of islet-infiltrating CD4$^+$ T cells at 6 weeks of age. Numbers are percentage of total CD3$^+$CD4$^+$ cells. (**B** and **C**) Representation of the percentages of Ins$_{12-20}$ and Ins$_{13-21}$ tetramer$^+$ cells in the purified islets of eleven 6-week-old and four 12-week-old animals used in this study. Whereas (B) shows numbers and mean value for all mice, (C) shows the pairwise analysis of both tetramers in each mouse. (**D**) Percentage of Ins$_{12-20}$ and Ins$_{13-21}$ tetramer$^+$ cells in the spleen and PLNs of mice with established diabetes (20 weeks). Determinations could not be made using islet tissue because most of the islets were destroyed. P values were calculated using an unpaired two-tailed t test in Prism.

described phenotype of functional anergy defined as FR4$^+$CD73$^+$ in the CD44$^+$ population (fig. S1) (26). After this curation, the unsupervised clustering revealed four different groups (Fig. 3, C and D, and fig. S2). This cluster definition was not influenced by age (6 versus 12 weeks) or the reactivity of the cell (Ins$_{12-20}$ versus Ins$_{13-21}$). Cells from cluster 1 were represented in all three tissues and, besides CD3 and CD4, expressed genes associated with cell adhesion and sustained activation of T cells, CD86, ICAM-1, CD80, CD40 (27, 28), and two transcription factors (Zeb2 and PPARγ) involved in terminal differentiation and effector function of T cells (29). This cluster of cells was the only one that increased significantly between 6 and 12 weeks. Cluster 2 was overrepresented in the PLNs and notable by absent features: no signs of TCR- or cytokine-mediated signaling and no markers associated with the major CD4$^+$ subsets. CCR7, a chemokine receptor strongly linked to lymph node homing, was the highest expressed gene in this cluster (Table 1). Clusters 3 and 4 identified the cells resident in islets and a very small percentage of PLN cells (Fig. 3). Cluster 3 cells were prototypic T$_H$1 cells with high expression of T$_H$1 cytokines, such as interferon-γ (IFN-γ), tumor necrosis factor–α (TNF-α), and interleukin-21 (IL-21); TCR-triggered signaling molecules, such as CD3, Zap-70, and Fyn; and coactivator genes, including CD28 and inducible T-cell costimulator (ICOS). The expression of IFN-γ and TNF-α was confirmed by flow cytometry for the whole intra-islet CD4$^+$ T cell population, and very few cells were found to be double expressors (fig. S3). Cluster 4 was more intriguing, with high expression of typical T$_{reg}$-associated genes, forkhead box P3 (Foxp3), IL2Rα, and IL-10, as well as persistent expression of T-bet and, in some cells, IFN-γ and TNF-α (Fig. 3E). The presence of Foxp3$^+$CD25$^+$
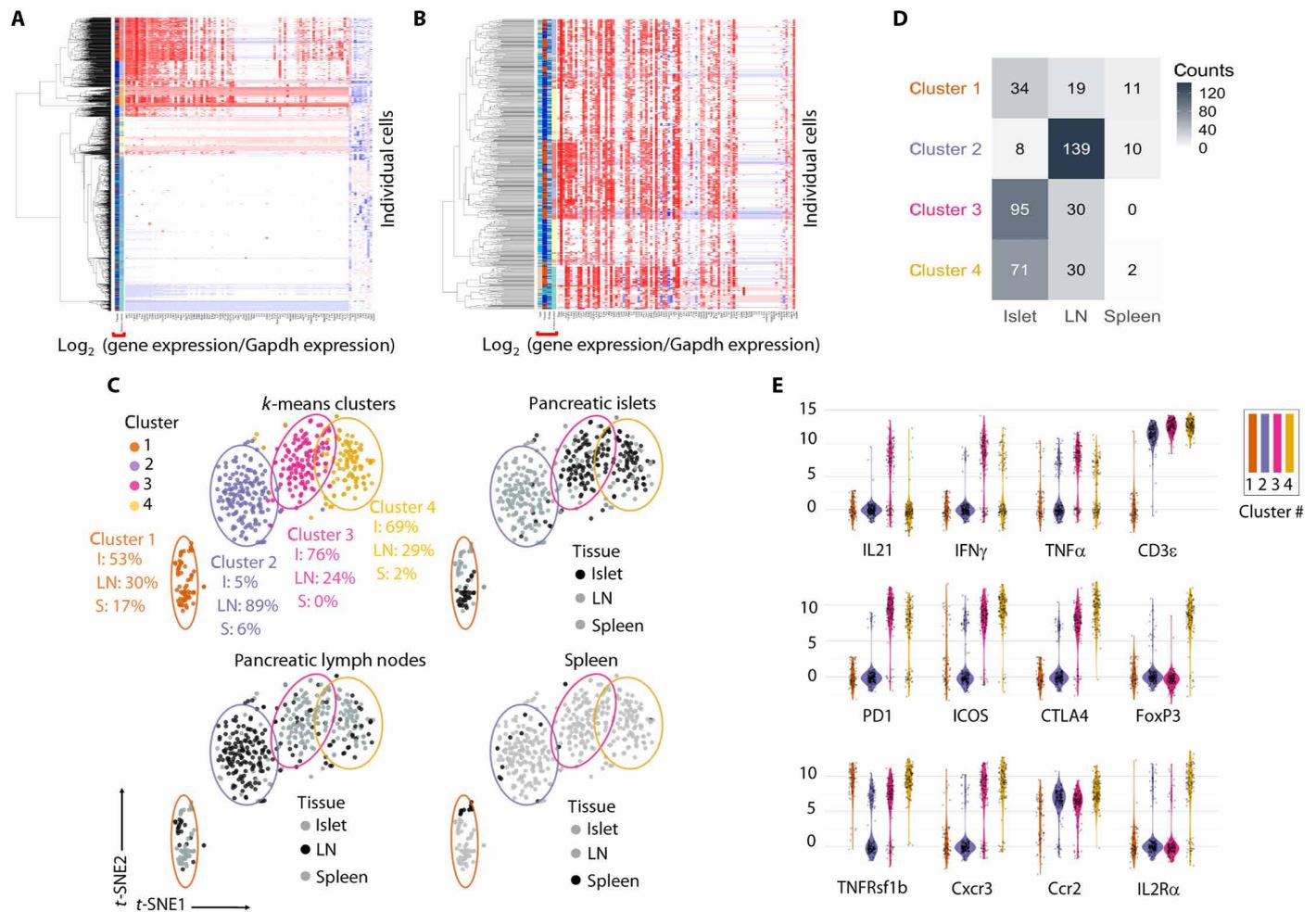
Fig. 3. Single-cell transcriptomics analysis of Ins$_{12-20}$ and Ins$_{13-21}$ tetramer$^+$ cells sorted from the spleen, PLNs, and infiltrated islets. Single tetramer$^+$ CD4 T cells were sorted on a FACSAria instrument directly in RT buffer and examined for expression of a panel of 96 genes. (**A** and **B**) Heatmaps of all cells from all organs before (A) and after (B) removal of cells expressing low levels of >80% of all genes. (**C**) t-SNE analysis of the single-cell gene expression data. The unsupervised cluster distribution (k-means) defined four separate groups of cells (top left: group 1, brown; group 2, blue; group 3, red; group 4, yellow). The distribution of each group in islets (I), PLNs (LN), and spleen (S) is indicated under each cluster. The other three panels show in black the representation of the cells from each tissue, pancreatic islets, PLNs, and spleen, within each of the four clusters, whereas cells from the other tissues are in gray. (**D**) Heatmap of the representation of each cluster in all three tissues examined and represented in (C). (**E**) Violin plot representation of the 12 statistically most differentially expressed genes that defined clusters 1 to 4.

cells was detected by flow cytometry at increased frequency in infiltrated islets compared with peripheral lymphoid tissues for both Ins$_{12-20}$- and Ins$_{13-21}$-specific cells (fig. S4). The overlap of highly expressed genes between clusters 3 and 4 included programmed cell death protein 1 (PD-1), cytotoxic T lymphocyte–associated protein 4 (CTLA-4), ICOS, and the TCR signaling molecules Zap70, Fyn, and Nur77 (Table 1 and Fig. 3E); we interpreted this overlap as a sign of the possible transition between these two subsets, indicative of the plasticity of T cells. In any instance, the near-equal representation of the two clusters, 45.6% of cells in cluster 3 and 34.5% in cluster 4 (Fig. 3D), in the entire anti-insulin T cell population regardless of register usage indicated that antigen alone was not the determining factor for effector function. We conclude from the single-cell gene analysis that the intra-islet events determine much of the final fate of the T cells, regardless of their specificity.

TCR sequencing of insulin-reactive CD4$^+$ T cells: Confirmation of the P9 switch model and islet pauciclonal expansion

Single-cell TCR sequencing of tetramer-sorted T cells was carried out using the same reaction used for RNA expression profiling (*25*). Briefly, after a preamplification step, a second round of polymerase chain reaction (PCR) was used to amplify and bar code each cell with a common tag on the Cα and Cβ primers. The final library was produced using a nested PCR approach with internal Cα and Cβ primers and analyzed on a MiSeq next-generation sequencer (*25*). The efficiency of sequencing of TCR chains was highly dependent on the state of activation of the cell as measured on our 96-gene set; chain pairing was successful in ~70% of cases, whereas in the remaining 30%, only one of the two TCR chains could be sequenced. In the absence of up-regulation of at least 50% of the 96 genes that we evaluated

Table 1. Differential gene expression between the four clusters defined by *t*-SNE analysis of the single-cell analysis (*P* values are indicated next to the name of each gene).

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	Cd86 8.24×10^{-35}	Ccr7 3.42×10^{-30}	Pd1 2.57×10^{-30}	Foxp3 2.01×10^{-36}
2	Zeb2 2.31×10^{-30}	Il7r 1.12×10^{-17}	Il21 2.15×10^{-23}	Ctla4 7.36×10^{-33}
3	Tnfrsf1a 8.54×10^{-27}	Ccr6 3.87×10^{-12}	Ifng 3.41×10^{-22}	Il2ra 6.24×10^{-24}
4	Icam1 1.64×10^{-26}	Ifi44 1.23×10^{-9}	Tnf 1.18×10^{-15}	Tnfrsf1b 1.17×10^{-21}
5	Tgfbr2 3.15×10^{-22}	Cxcl10 1.65×10^{-9}	Stat4 1.26×10^{-14}	Icos 2.26×10^{-21}
6	Ifngr1 4.85×10^{-22}	Ifit3 2.98×10^{-8}	Cd3e 3.25×10^{-13}	Cxcr3 1.91×10^{-19}
7	Cd80 3.56×10^{-20}	Bcl2 4.61×10^{-8}	Cd28 4.28×10^{-10}	Ifngr1 1.38×10^{-14}
8	Ccr1 3.98×10^{-20}	Il25 1.23×10^{-6}	Fyn 3.65×10^{-9}	Ccr2 1.93×10^{-13}
9	Vav1 4.47×10^{-15}	Cd3e 2.78×10^{-5}	Zap70 3.74×10^{-9}	Cd3e 2.85×10^{-12}
10	Cd40 1.18×10^{-13}		Jak2 6.08×10^{-9}	Tbx21 2.62×10^{-111}
11	Il4ra 7.38×10^{-12}		Icos 1.04×10^{-8}	Ly6e $1.76e \times 10^{-10}$
12	Pten 1.26×10^{-11}		Cxcr3 2.88×10^{-8}	Zap70 3.55×10^{-10}
13	Tnfrsf1b 2.49×10^{-11}		Ctla4 5.87×10^{-8}	Ccr5 1.72×10^{-9}
14	Pparg 9.59×10^{-11}		Bcl6 1.30×10^{-7}	Cd28 2.75×10^{-9}
15	Cd4 1.58×10^{-10}		Cd4 8.42×10^{-7}	Il10 7.28×10^{-8}
16	Hprt 7.23×10^{-10}			Fyn 7.34×10^{-8}
17	Mx1 6.89×10^{-9}			Nur77 2.06×10^{-7}
18	Socs3 7.56×10^{-9}			Ccr4 2.33×10^{-7}
19	Stat1 2.29×10^{-8}			Il.12rb 2.89×10^{-7}
20	Irf7 4.40×10^{-8}			Tnfaip3 6.97×10^{-7}
21	Ccr5 4.54×10^{-8}			Irf1 1.30×10^{-6}
22	Traf2 7.27×10^{-8}			Il27r 2.19×10^{-6}
23	Il27 3.38×10^{-7}			Pd1 1.70×10^{-5}
24	Aim2 4.03×10^{-7}			Icam1 2.90×10^{-5}
25	Gsk3b 4.84×10^{-7}			
26	Pdl1 6.35×10^{-7}			
27	Rsad2 3.89×10^{-6}			
28	Il12b 2.39×10^{-5}			

in our panel, TCR sequencing efficiency was always low (<20%). Consequently, very few sequences were recovered from the spleen, and fewer from the PLNs than from the islets. The first information that came from this analysis is that both $Ins_{12–20}$ and $Ins_{13–21}$ cells in the PLNs and islets expressed diverse T cell receptor alpha variable region (TRAV) and T cell receptor beta variable region (TRBV) segments with slightly more diversity in the islets (Fig. 4). This diversity came with some bias and an overrepresentation of TRAV6.5, TRAV6.6, TRAV5D4, TRBV2, TRBV5, and TRBV13.1 for $Ins_{12–20}$ cells in both locations, whereas TRBV1, TRBV5, and TRBV13.2 were overrepresented for $Ins_{13–21}$ cells in the islets. The profiles of $Ins_{12–20}$ V segment usage, although very similar between the islets and PLNs, did not correspond to matching TCR pairs in the two locations; similarly, no identical TCRs could be found between mice. A pairwise analysis (Fig. 4B) revealed some preferential Vα/Vβ pairing for $Ins_{12–20}$ cells, but all of these pairs corresponded to clones. As we defined the occurrence of clonal expansion by the sequencing of the same TCR pair in three or more cells in the same location, clones were only detected in the islets and never in the PLNs, and most were found

at 6 weeks of age and were specific for $Ins_{12–20}$ (Figs. 4B and 5A). When cluster identity was examined for each clone, although most of them belonged to either cluster 3 or 4, some were split between clusters (Fig. 5A). These observations strongly supported the hypothesis that antigen recognition and expansion took place in the islets and not in the lymph node and demonstrated that the functional fate, T_H1 versus T_{reg}, was not exclusively dependent on TCR usage. To support the first conclusion, we stimulated $Ins_{12–20}$ and $Ins_{13–21}$ T cell hybridomas with fresh islet cells and PLN cells in the absence of exogenous peptides. As shown in Fig. 5B, only islet cells activated Ins-specific clones, whereas PLN cells, even in much larger numbers, did not (Fig. 5B).

Examination of the CDR3 regions of TCRα and TCRβ isolated from islets and PLNs revealed the same length (~14.0 ± 1.0 amino acids) and an increase in frequency of negatively charged residues at position +3 of CDR3β of $Ins_{12–20}$ T cells, suggesting the existence of a P9 switch (Fig. 6A). In most cases, this D/E residue was encoded by a non-germline sequence, indicating the role of the antigen recognition in the selection and expansion of these populations (table S2). To confirm the functionality of this switch, three TCRαβ pairs isolated

A

B
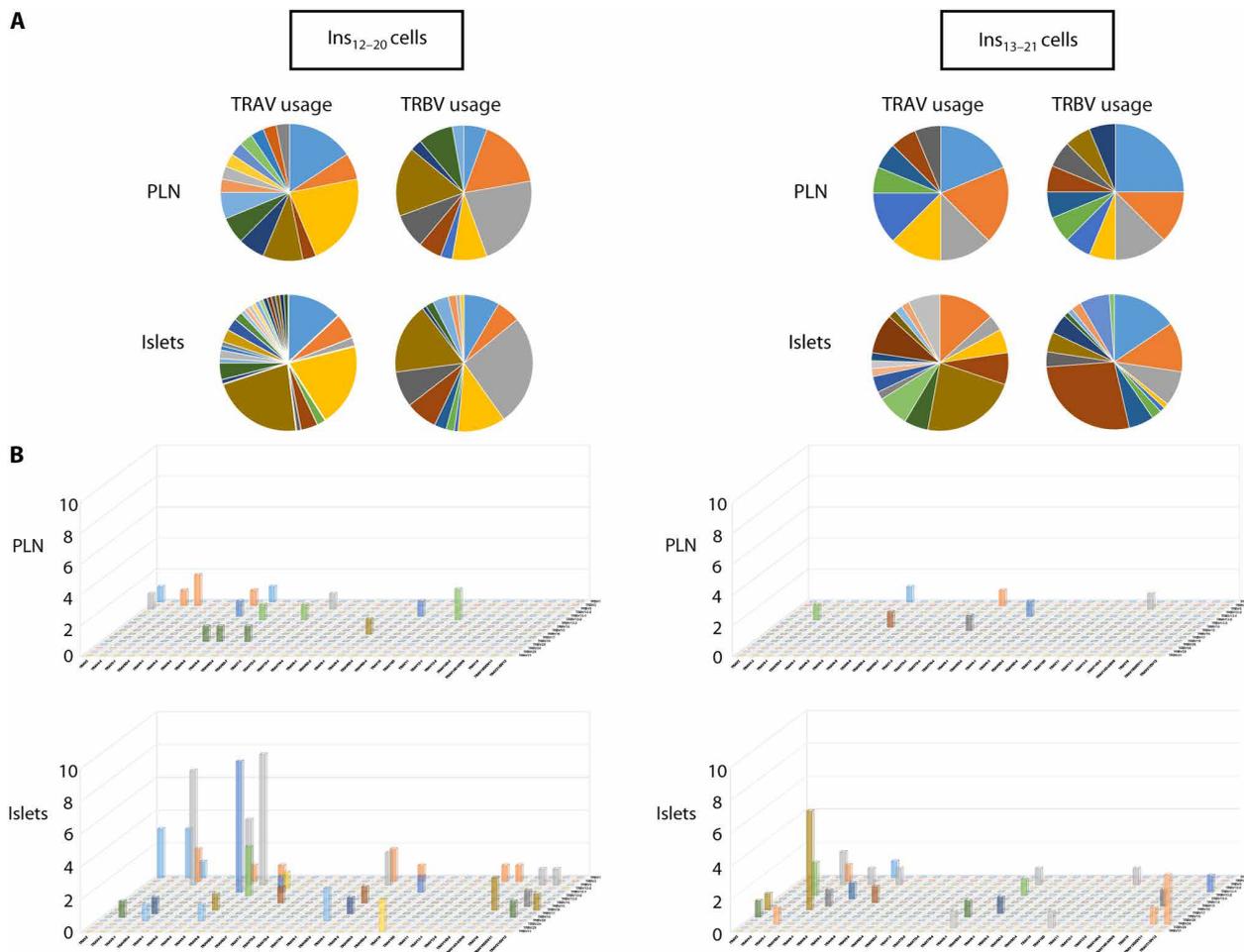
Fig. 4. Single-cell TCRαβ paired sequencing in PLNs and islets. (**A**) TRAV and TRBV usage in PLNs and islets for Ins_{12-20} and Ins_{13-21} tetramer$^+$ cells. (**B**) Clonal distribution of TCRαβ pairs in PLNs and islets for Ins_{12-20} and Ins_{13-21} tetramer$^+$ cells. The *x* axis lists the various TRAV segments, the *z* axis lists the TRBV segments (see Materials and Methods for list), and the *y* axis indicates the number of cells for each TCRαβ pair. Numbers of sequences: islets Ins_{12-20}, 107 cells; PLNs Ins_{12-20}, 36 cells; islets Ins_{13-21}, 84 cells; PLNs Ins_{13-21}, 16 cells.

from islet cells and retaining a negative charge in CDR3β were cloned and re-expressed in BW5147 TCR-negative cells using the lentiviral vector pMIG2 (*30*), and used for functional testing on plastic-bound pMHC molecules. Beyond confirming specificity for Ins_{12-20}, the titration of ligand demonstrated that the functional avidity of the three TCR pairs varied substantially with stimulatory index 50 (SI_{50}) corresponding to between 60 and 200 ng of ligand per well (Fig. 6B). Exposure of these clones to an I-A^{g7} Ins_{12-20} complex mutated at position 9 of the peptide for a glutamic acid (Ins_{12-20} P9E) resulted in the complete absence of stimulation (Fig. 6B). The same result was obtained for 2D10 and 8F10 T cells, two T cells isolated from immunized mice, and whose SI_{50} values are 5 and 50 ng of ligand per well, respectively (Fig. 6C). The reciprocal experiment was performed by mutating the glutamic acid residue 114 in the CDR3β of Cl2 TCR and testing both wild-type (WT) and mutant TCRs on the I-A^{g7} Ins_{12-20} complex (Fig. 6D); the mutant Cl2 T cells were unresponsive. Last, the same experiment was performed using recombinant 2D10 TCR and its CDR3β$_{116}$D➔G mutant using plasmon resonance for affinity measurements (Fig. 6E). Both WT and mutant TCRs were captured at the surface of a lipid bilayer by their C-terminal histidine tails using a 1,2-dioleoyl-sn-glycero-3-[(N-(5-amino-1-

carboxypentyl)imidodiacetic acid)succinyl] (DOGS) lipid and tested for binding to I-A^{g7} Ins_{12-20}. Whereas the WT 2D10 TCR bound with measurable affinity ($K_{ass} = 1.9 \times 10^4$ M s^{-1}, $K_{diss} = 0.01$ s^{-1}, and $K_d = 4.44 \times 10^{-7}$), the point mutation resulted in an almost complete loss of measurable binding (association rate was not quantifiable), confirming the functionality of the switch in this TCR.

Genetic demonstration of the role of the switch in insulin recognition

The importance of mutations at positions β56 and β57 of I-A^{g7} in the onset of insulitis and diabetes was tested 25 years ago by transgenesis in the NOD mouse using Aβ^k and Aβ^d transgenes (*31–34*). The more focused targeting of position I-A^{g7}β57S➔D was also tested by transgenesis and resulted in a marked reduction in diabetes but a persistence of peri-insulitis (*35, 36*). These transgenic mouse models were very imperfect, mainly because they all still carried a normal I-A^{g7} molecule. We showed that the simple dilution of a WT IA^{g7} onto a NOD-$Ab^{-/-}$ background (NOD-$Ab^{-/-}$ × NOD F1 mouse) was sufficient to prevent diabetes, establishing that the quantity of MHC and/or antigen displayed by MHC on APCs controlled disease onset (fig. S5). Because of these limitations, none of the transgenic experiments
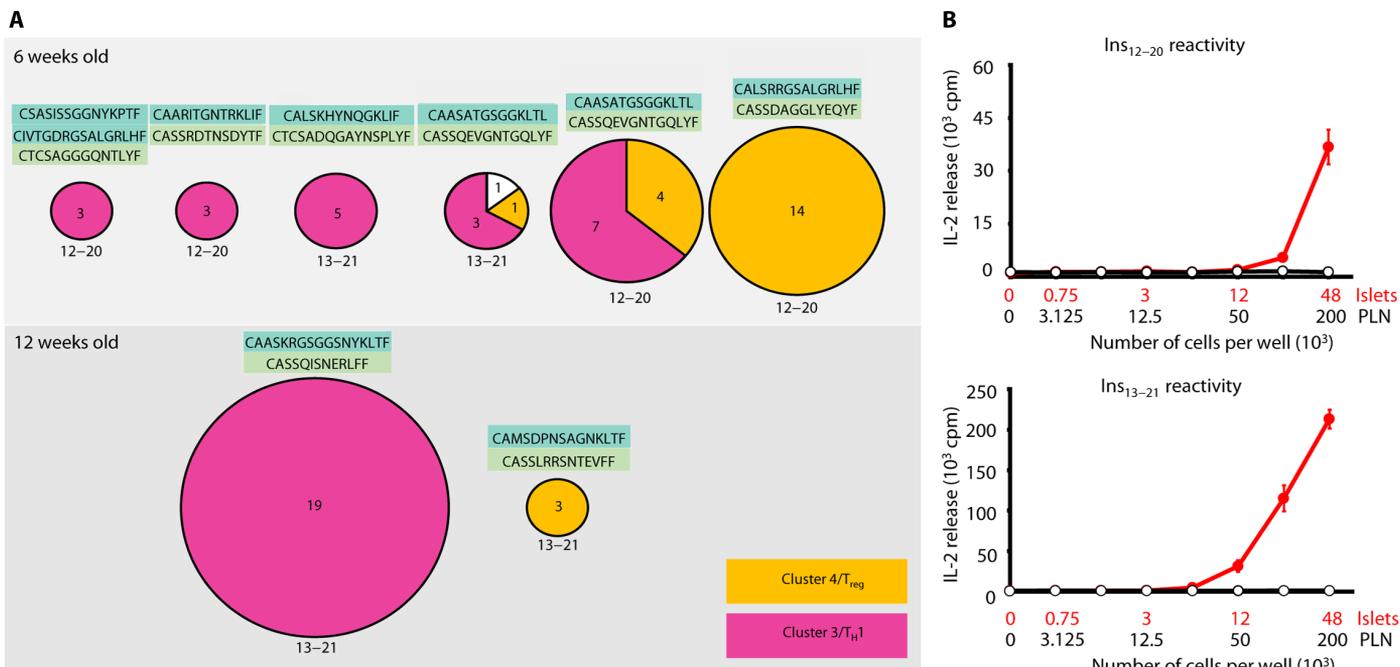
Fig. 5. Analysis of the clonal expansion of anti-insulin CD4[+] cells in the islets. (**A**) Clone CDR3α and CDR3β sequences are shown next to each circle. Each circle size is proportional to the clone size. Antigen specificity for Ins_{12-20} versus Ins_{13-21} is shown underneath each circle, whereas the cluster affiliation is represented in color (pink for cluster 3 and yellow for cluster 4). (**B**) The islet clonal expansion correlated with the ability of fresh islet cells to stimulate Ins_{12-20}- and Ins_{13-21}-specific T cell hybridomas (red line) in the absence of peptide addition, whereas PLN cells could not. Islet and PLN cell numbers are indicated. This experiment is representative of five similar experiments. Each dilution of cells was tested in triplicates for PLNs and duplicates for islets, and SDs are shown.
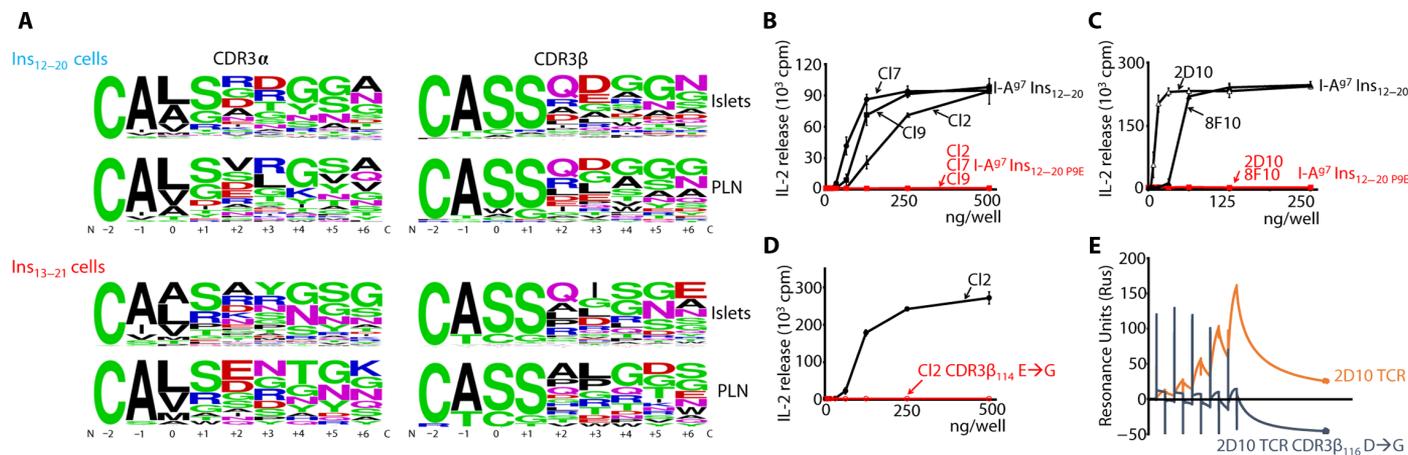
Fig. 6. Ins_{12-20}-specific CD4[+] T cells use a P9 switch for activation. (**A**) Analysis of CDR3α and CDR3β sequences of Ins_{12-20}- and Ins_{13-21}-specific cells in both PLNs and islets. Schematic representation of the percentage of each amino acid at each position of the N terminus of each CDR3 normalized against the frequency of each residue at each position. Numbering of the CDR3 residues follows the nomenclature of our previous publications (*9*, *15*). (**B** and **C**) Usage of the switch by Ins_{12-20} T cell clones. Clones 2, 7, and 9 isolated from islets (B) and 2D10 and 8F10 isolated after immunization (C) were tested for IL-2 release after stimulation on coated recombinant peptide/I-A[g7] molecules retaining either the native peptide (black) or mutant P9E version (red). (**D**) The effects of the mutation 114E→G in the CDR3β of TCR are shown for clone 2 using recombinant I-A[g7] Ins_{12-20}–coated wells. WT hybridoma is in black and mutant is in red. (**E**) Direct SPR measurements of the loss of affinity of recombinant 2D10 TCR after mutation D→G at position 116 of the CDR3β WT (orange trace) and mutant (gray trace) TCRs were captured on lipid bilayers using DOGS lipids on an L1 chip, and twofold dilutions of recombinant I-A[g7] Ins_{12-20} strep-tag were injected onto the surface using an irrelevant TCR in the flow cell used for subtraction. All functional experiments are representative of a minimum of three independent repeats. SPR measurements were performed twice.

could address the mechanisms by which position β57 altered T cell responses and progression toward autoimmunity. To access this information directly, the serine β57 of I-A[g7] was mutated to an aspartic acid using zinc-finger nuclease (ZFN) technology directly in NOD mice (*37*). Before examination, the new mouse line was backcrossed twice into the WT NOD background to eliminate any potential off-target effect of ZFN cleavage and repair. Neither males nor females of this mutant mouse line developed diabetes, and no

periinsulitis could be observed in mice, even after 20 weeks. Mutant I-A^{g7}S$_{57}$➔D tetramers were recombinantly expressed for both insulin peptides and used to quantify anti-insulin CD4$^+$ T cells in the mutant mouse. As shown in Fig. 7, the numbers of Ins$_{12-20}$ and Ins$_{13-21}$ T cells were comparable in the thymus, spleen, and PLNs of 8-week-old female mice; cell numbers could not be obtained from islets because so few CD4$^+$ T cells could be isolated. Although tetramer$^+$ cells were in almost identical numbers in the spleen of the mutant mouse as in WT mice, numbers in PLNs were three times lower in the β57 mutant animal (Figs. 1 and 7). The TCRs of the rare insulin-specific cells isolated from the PLNs were sequenced and compared with WT PLN TCR sequences (Fig. 7B). The disappearance of D/E residues in CDR3β of Ins$_{12-20}$-specific T cells from mutant animals as compared with WT mice was highly significant (13.3% versus 28.2%; $P < 0.001$), confirming the direct role of serine β57 in the selection of this particular TCR repertoire. The influence of β57 was also shown functionally by comparing the presentation of HEL and insulin by APCs from WT or mutant mice. As shown in Fig. 7C, the presentation of HEL was diminished when the antigen was presented to the anti-HEL TCR 21.30 by mutant as compared with presenting cells from regular NOD mice, confirming our previous studies (9), although the recognition of the Ins$_{9-23}$ peptide was completely abrogated for all Ins$_{12-20}$-specific T cells that we tested and of which clone 9 is a representative. The presentation of Ins$_{9-23}$ peptide to the prototypic anti-Ins$_{13-21}$ 11T3 T cell hybridoma was nearly identical between NOD and β57 mutant mice (fig. S6).

DISCUSSION

The development of reagents that could selectively detect the two main adjacent registers of the Ins$_{9-23}$ peptide allowed us to dissect some fundamental aspects of the biology of I-A^{g7} molecules and their impact in the pathogeny of T1D. Beyond showing the importance of peptide registers in the biology of I-A and HLA-DQ molecules (15, 19, 38), these new tetramers confirmed the existence of two adjacent registers in the B$_{9-23}$ segment of insulin (16, 39).

The interrogation of the biology of autoreactive T cells against a nominal antigen using single-cell analysis and a set of selected T cell–related genes revealed some important aspects of autoimmunity. Foremost, the fact that a majority of anti-insulin T cells were expressing very low levels of any of the 96 genes that we selected as indicators of T cell functions, in the spleen, lymph nodes, and even islets, seems to correlate well with the strong state of peripheral anergy toward insulin that is observed in most mouse genetic backgrounds, including the NOD mouse. This "anergy" was also correlated phenotypically with the expression of CD73 and FR4 in the CD44$^+$ population (26). Functionally, this overall concept of "high degree of tolerance toward insulin" is confirmed by the outcome of immunization of mice with B$_{9-23}$ peptide or insulin; we and others have shown that insulin administration rarely produces measurable CD4$^+$ T cell expansion, regardless of the adjuvant used, although it induces cytokine secretion measurable by enzyme-linked immunospot (23, 40). Recent studies indicate that there is a continuous and low level of presentation of insulin peptides in all peripheral lymph nodes including the pancreatic nodes. The peptides are derived from crinosomes released into the circulation after a glucose challenge (17). This low-level and persistent presentation in peripheral nodes induced in the 8F10 T cells that recognize Ins$_{12-20}$ represents weak activation but not anergy. Ablation of lymph nodes results in complete absence of diabetes and T cell priming (41). This response at the periphery will likely vary among T cells depending on their TCR specificities and affinities. Regardless, the islet presentation will profoundly influence the response based on its display of a number of regulatory molecules such as PD-1. Therefore, it seems logical that tolerance to insulin will have a higher chance of being broken in the environment of the islet where all forms of insulin and its by-products and abortive products are presented. Studies in TCR transgenic mice have supported this hypothesis (39), and our gene expression and T cell repertoire analyses of insulin-specific cells strongly supported the same conclusion. What allows the breakage of tolerance in the islets? Is it linked to antigen processing and presentation? The effects of the dilution of I-A^{g7} in MHC knockout NOD mice demonstrate that this is likely the case. However, besides the existence of a threshold for the quantities of particular pMHC complexes needed to activate autoreactive cells, the
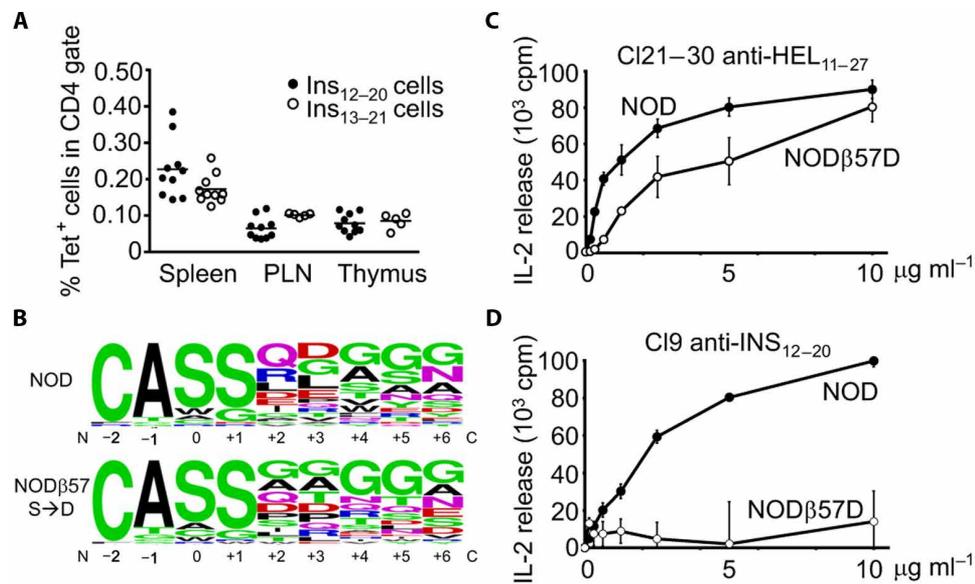
Fig. 7. The in vivo substitution of an aspartic acid for serine β57 of I-A^{g7} modifies the anti-insulin T cell repertoire and eliminates the switch in the anti-Ins$_{12-20}$ T cell population. (**A**) I-A^{g7}57S➔D Ins$_{12-20}$ and Ins$_{13-21}$ tetramers were used to enumerate register specific cells in the spleen and PLNs of NODβ57S➔D mutant animals. Numbers are expressed as percentages of CD3$^+$CD4$^+$ cells. (**B**) The TCR repertoire of Ins$_{12-20}$-specific cells isolated from the PLNs of NOD and NODβ57S➔D animals was examined. The sequences of the respective CDR3β are presented. (**C**) The usage of the switch for TCR carrying a D/E residue in CDR3β was tested for HEL using T cell clones 21 to 30, and the long Ins$_{9-23}$ peptide using clone 9 using either NOD (black symbols) or NODβ57S➔D (empty symbols) splenocytes as APCs. The same experiment was repeated three times with similar results.

classical model of T cell activation tells us that effective costimulation as well as a milieu providing cytokines and growth factors (42) will be necessary. In an inflammatory context, it appears that the islet-resident macrophage is the most likely cell responsible for self-antigen presentation. Not only do they capture immunogenic peptides from the crinophagic granules upon insulin stimulation (17), they also express a very efficient antigen processing and presentation machinery (43), and their elimination results in the prevention of disease (44). Of course, it cannot be completely ruled out that in the context of the islet, β cells themselves contribute to T cell stimulation because they have been shown to express MHC class I and class II upon TNF-α/IFN-γ stimulation (45), two cytokines that are found in the lesions. A scenario involving antigen presentation by two different cell types might explain one of the most unexpected findings of our study: the ability of T cell clones to differentiate either as T_H1 or as T_{regs} (Fig. 5A). For instance, the engagement of PD-1 by the PD-1 ligand expressed by β cells might favor T_{reg} over T_H1 differentiation (46). In addition, the ability for a single T cell to receive signals from two presenting cells might also explain the persistence of T_H1-associated genes such as T-box transcription factor X21 (Tbx21) and IFN-γ in some T_{regs} (47, 48). It is also important to contrast findings in the islets and the PLNs, a location where we never found clonal expansion and in which T_{regs} did not express Tbx21 in addition to Foxp3, yielding only classical T_{regs}. These differences suggest that T_{regs} received additional differentiation signals in the islets.

The second most important conclusion of our study is the demonstration of the critical role that the P9 switch mode of recognition plays in the pathogeny of T1D, as illustrated with the Ins_{12-20} peptide. Building on the prior finding that the P5 position of the Ins_{12-20} peptide has been shown to be essential for diabetes development (49), we show here that the P9 switch is essential for the function of Ins_{12-20}-specific T cells and that at the population level, a D/E residue in CDR3β is selected and necessary for T cell recognition. Examination of infiltrating $CD4^+$ T cells at 6 weeks of age when insulitis is minimal also demonstrated the preferential expansion of Ins_{12-20}-specific T cells at the early preclinical phase of disease, indicating the proliferative advantage that the P9 switch provides to T cells. It is quite remarkable to see that more than 50% of the $CD4^+$ islet infiltrate is made of these cells at 6 weeks. It will be of upmost interest to determine whether the remaining cells also use a switch for recognition and expansion and whether peptides are provided by other fragments of insulin. Particular attention will be given to the C-peptide, a major source of antigen in humans (50) and mice (51, 52); almost all humans C-peptide–derived epitopes (50), as well as the immunodominant murine C-peptide determinant (51), have no negative charge at P9. In humans, the recognition of these peptides by TCRs fulfilling the canons of a P9 switch mode of T cell recognition with D/E amino acids in their CDR3β has already been reported (53). In mice, similar observations were made for noninsulin epitopes devoid of D/E P9 residues (9).

Last, the kinetics of the P9 switch–driven anti-insulin response in mice with its early burst and rapid disappearance have important translational implications. If they also exist in humans, it is only in at-risk patients that these cells would be detectable; their capture from peripheral blood could constitute the earliest possible marker of islet autoimmunity and change the timing of therapeutic intervention.

Our observations regarding the potential role of the P9 switch mode of T cell recognition in the onset of diabetes in the NOD mouse were challenged by the examination of the mutant β57S➔D

mouse. Our prediction was that the disappearance of the positively charged surface at P9 would eliminate T cells capable of using the P9 switch. As a result, the Ins_{12-20} peptide should select a normal T cell repertoire without D/E residues in the N-terminal segment of the CDR3β. Although this expectation was proven correct by direct sequencing of the anti-Ins_{12-20} TCRs, the enumeration of anti-insulin cells showed their marked reduction in PLNs as compared with WT NOD mice and their near-complete elimination in the islet. This latest point established the direct role of P9 switch cells in the initiation of infiltration and the inability of cells without switch to substitute for them. Therefore, the binding advantage that we have documented for P9 switch cells appears necessary and sufficient to break tolerance in the context of islet antigen presentation. This simple relationship between position β57 of diabetogenic MHC molecules and the selection of a unique T cell repertoire appears to be the molecular mechanism that links MHC and T1D. Further studies will attempt to confirm this conclusion in humans. This task might prove challenging because although the switch plays an essential role in initiation, it appears not necessary for continuation of disease. However, the examination of at-risk and just-diagnosed patients should allow us to overcome this difficulty. The initial analysis of anti–C-peptide $CD4^+$ T cells that showed the presence of D/E residues in CDR3β makes us optimistic (53). It is likely that antibodies or small molecules capable of interfering with the position β57 of HLA-DQ8, if applied early enough in the development of the autoimmune process, would be capable of preventing progression and onset of disease.

MATERIALS AND METHODS
Study design
The study was intended to examine the very early phase of T1D in mice and dissect the kinetics and characteristics of the anti-insulin response. We took advantage of new tetramer reagents we developed to follow anti-insulin $CD4^+$ T cells in mice and performed single-cell analysis. The deep examination of the T cells specific for the Ins_{12-20} fragment was aimed at testing the relevance of the P9 switch model of T cell recognition in disease onset and progression.

Mice
NOD/LtJ mice were purchased from the Jackson laboratory and housed under pathogen-free conditions. NOD/LtJ mice in which MHC class II genes were deleted were also from the Jackson laboratory [NODCaj.129S2(B6)-*H2-Ab1*^{tm1Doi}/LwnJ] but were bred in-house for the past decade and genotyped in the laboratory. The mutant NOD/LtJ mouse strain in which position β57 of I-A^{g7} was mutated from S➔D by ZFN technology (NOD/ShiLtJ-*H2-Ab1*^{em2Ygch}) was produced by C.M. and Y.-G.C. (37) and was bred and genotyped in our facility. Care and handling of mice followed Institutional Animal Care and Use Committee rules. Glycemia of all animals was measured weekly from a drop of blood collected from the tip of the tail.

Cells and transfection
BW5147 (TCRα$^-$β$^-$) cells and T cell hybridomas were maintained in Dulbecco's modified Eagle's medium–10% fetal bovine serum (FBS) medium. For reexpression of TCR in BW5147 after sequencing, Vα and Vβ segments were synthesized and cloned into pMIG II (Addgene) and spinfected with polybrene for 2 hours at room temperature at $1000g$ (30). After recovery, cells expressing TCR were sorted by staining with the anti–TCRβ chain antibody H57-597 (BD Biosciences).

Construction, expression, and purification of I-A^{g7}/peptide tetramers and recombinant TCR molecules

The generation of I-A^{g7}/peptide complexes has previously been reported in detail (54). The peptide sequences of I-A^{g7} HEL$_{11-27}$, I-A^{g7} Ins$_{12-20}$, and I-A^{g7} Ins$_{13-21}$ molecules are AMKRHGLDNYRGYSLGN, VEALYLVCG, and EALYLVCGE, respectively. The biotinylation sequence #85 from Schatz (55) follows the acidic zipper on the α chain. Biotinylation was performed in vivo by cotransfecting an inducible BirA plasmid and adding 50 μM biotin at the time of induction. After purification by ion exchange and size exclusion, biotinylated molecules were tetramerized at 4°C overnight using phycoerythrin (PE)–labeled streptavidin (BioSource International) or PE–cyanine 7– or APC-labeled streptavidin (eBioscience) using a 5:1 molar ratio of biotinylated molecules to labeled streptavidin. For surface plasmon resonance (SPR) measurements on lipid bilayers, the same pMHC molecules were produced with a C-terminal strep-tag II sequence and purified on Strep-Tactin columns (IBA Lifesciences).

TCR complementary DNA for the α and β chains of insulin-reactive T cell hybridomas and primary cells were subcloned into the fly expression vector pRMHa3 or pMT/Bip/V5-His (Invitrogen) and sequenced. The final constructs code for the VαCα and the VβCβ chains, respectively, followed by a linker sequence (SSADL), a thrombin site (LVPRGS), a leucine zipper (acidic for the α chain and basic for the β chain), and a hexahistidine tag. Vectors were cotransfected into SC2 cells along with a vector encoding a puromycin-resistance gene, and stable cell lines were established. Soluble TCRs were purified from culture supernatants, as previously described (56).

Peptides

Synthetic peptides HEL$_{11-27}$ (AMKRHGLDNYRGYSLGN), Ins$_{9-23}$ (SHLVEALYLVCGERG), and forms of Ins$_{12-20}$ (TEGVEALYLVC-GGGS) and Ins$_{13-21}$ (TEGEALYLVCGEGGS) in which the N and C termini were extended by TEG- and GGS- sequences, respectively (39), were synthesized by AnaSpec Inc. and were ≥95% pure.

Site-directed mutagenesis

PCR site-directed mutagenesis was carried out using the QuikChange II PfuUltra High-Fidelity DNA Polymerase (Agilent) and verified by DNA sequencing.

Islet isolation and cell preparation

Islets were isolated using a modified protocol derived from Li et al. (57). Briefly, pancreata were isolated and incubated for 60 min at 37°C after injection of collagenase P through the common bile duct. After several washes, islets were isolated on a metal strainer and hand-picked under a dissection microscope. An average of 200 to 300 islets were obtained per mouse. For making single-cell suspension, islets were digested with trypsin or dispase at 37°C and washed extensively before being counted.

Cell preparation, cell staining, and flow cytometry analysis

Non-islet single-cell suspensions were prepared by mechanical disruption of the corresponding organs in Hanks' balanced salt solution buffer, and erythrocytes were removed by osmotic lysis. Cells were washed with fluorescence-activated cell sorting (FACS) buffer [phosphate-buffered saline (PBS) containing 3% FBS and 0.05% NaN$_3$] and incubated with avidin (0.5 mg/ml; Sigma-Aldrich) and Fc block (BD Biosciences) in FACS buffer for 1 hour at room temperature. Cells were then washed once and stained with PE-labeled

or PE-Cyanine 7–labeled MHC/peptide tetramers at a final concentration of 10 μg/ml in FACS buffer for 1 hour at room temperature. For costaining of surface markers, fluorescein isothiocyanate–anti-CD4 (clone RM4-4, BD Biosciences), allophycocyanin–anti-B220 (clone RA3-6B2, eBioscience), and allophycocyanin–anti-CD8 (clone 53-6.7, eBioscience) were used. Exclusion of dead cells was performed by propidium iodide labeling (Invitrogen). Flow cytometry was performed using a FACS LSR-II instrument (Becton Dickinson) or a MACSQuant analyzer (Miltenyi Biotec), and data were analyzed using FlowJo software (TreeStar Inc.). Staining of T cell hybridomas with MHC tetramers was carried out in the same way.

CD4$^+$ T cell responses

For plate-bound activation assays, pMHCs were coated overnight in PBS at 4°C at various concentrations in flat-bottom 96-well plates and washed twice in PBS before plating 4×10^4 T cell hybridomas per well. After 24 hours of incubation at 37°C, supernatants were harvested and assayed for IL-2 content using a radioactive IL-2–dependent natural killer cell line bioassay. For assays with live APCs, T cell hybridomas were incubated with 1×10^5 APCs per well and increasing concentration of peptides. All assays were performed using biological triplicates and performed in at least three independent experiments.

Single-cell analysis

For single-cell analysis, T cells were sorted with pMHC tetramers on a FACSAria III machine (BD Biosciences) directly into reverse transcriptase (RT) buffer (10 μl per well). Protocols for preamplification and amplification for gene expression analysis and TCR sequencing have been detailed in recent publications (25). Gene expression analysis was performed on a Biomark unit with 96 × 96 Dynamic Array IFCs, whereas sequencing was performed on an Access Array instrument using 48 × 48 Library Prep IFCs (Fluidigm). Gene expression data quality control was performed using the Fluidigm software package and manual inspection of each melting curve to remove cells in which primer dimers were interfering with signal. After curation, data were analyzed with a homemade R package (see below). DNA sequencing was performed on a MiSeq instrument (Illumina). After stitching, TCR sequences were confronted to the IMGT database to determine TRAV and TRBV assignment (58). The following list of TRAV (x axis) and TRBV (y axis) segments appears in Fig. 4: TRAV2, TRAV4-2, TRAV5-1, TRAV5D-4, TRAV6-1, TRAV6-2, TRAV6-3, TRAV6-5, TRAV6-6, TRAV6D-4, TRAV6D-7, TRAV7-2, TRAV7D-2, TRAV7D-4, TRAV7N-4, TRAV8-1, TRAV8D-2, TRAV9-1, TRAV9-3, TRAV9D-3, TRAV9D-4, TRAV10, TRAV10D, TRAV11, TRAV12-1, TRAV12-2, TRAV14D-2, TRAV14D-3/DV8, TRAV16, TRAV16D/DV11, TRAV21/DV12, and TRBV1 TRBV2 TRBV5 TRBV12-2 TRBV13-1 TRBV13-2 TRBV13-3 TRBV14 TRBV15 TRBV16 TRBV17 TRBV19 TRBV20 TRBV24 TRBV29 TRBV31.3.

Surface plasmon resonance

A Biacore T200 instrument was used to determine interactions between purified MHC/peptide complexes and TCR molecules. TCR was immobilized on lipid bilayers coated on L1 chips by capture of their histidine tag. Lipid bilayers were produced by fusion of liposomes on the hydrophobic surface. Phosphatidylcholine (PC), cholesterol (C), dioleoyl phosphatidylcholine (DOPC), L-α-PE, L-α-phosphatidylglycerol (PG), and DOGS-NTA were all from Avanti Polar Lipids, and liposome composition was 60:20:5:5:10 (PC:C:DOPC:PE:PG:DOGS-NTA) weight ratio. Liposomes were

produced by extrusion at 1 mg/ml in PBS on 400-nm polycarbonate disks (*59*). Strep-tag–modified pMHC molecules were injected over the surface after stabilization of baseline in filtered and degassed PBS at a flow rate of 20 to 30 μl/min. To avoid regeneration of the surface between injections, kinetics studies were performed using the single cycle mode using an irrelevant TCR in the subtraction flow cell. K_d values, as well as on and off rates, were obtained by nonlinear curve fitting of subtracted curves using the 1:1 Langmuir binding model using the BIAevaluation program (version 3.0.2).

Statistics

Fisher's exact test and Student's *t* test (two-tailed) were used for statistical analyses.

R analysis methods

Single-cell expression data, in the form of quantitative PCR (qPCR) C_t values from the Fluidigm Biomark HD system, were analyzed in the R programming environment according to recommendations from the Fluidigm SINGuLAR Analysis Toolset documentation. Log2 expression values were calculated using the default limit of detection C_t value of 24, and negative log2 expression values were set to 0. Log2 expression values were then normalized per individual cell by dividing each cell's gene expression values by its Gapdh expression value. These normalized values were median-centered for further analysis. The expression of 96 genes was quantified in 1746 individual cells. Cells were removed from the analysis if no expression of any gene was detected in the isolate (111 cells removed) or if *Gapdh* expression was markedly low (55 cells removed). Cells with relatively low total expression were filtered by clustering the cells and removing clusters with low cumulative normalized log2 expression values (1131 cells removed). Differential gene expression was calculated using the nonparametric Kruskal-Wallis test by ranks, due to the multimodal distributions of most genes. *P* values were corrected by the Benjamini-Hochberg procedure.

The "Rtsne" R package was used for *t*-distributed stochastic neighbor embedding (t-SNE) dimensionality reduction. Violin, table, and *t*-SNE plots were created using the "ggplot2" R package (*60*). Heatmaps are based on the heatmap.2 function from the "gplots" R package. The columns of the heatmaps are ordered from lowest to highest differential expression *P* values, and the rows are ordered by hierarchical clustering based on Euclidean distance.

SUPPLEMENTARY MATERIALS

immunology.sciencemag.org/cgi/content/full/4/38/eaaw6329/DC1

Fig. S1. Examination of the "anergic state" of total and Ins-tetramer+CD4+ T cells in the spleen and PLNs based on FR4+/CD73+ surface labeling of the CD44− and CD44++ populations.

Fig. S2. Violin plot representation of the top most up-regulated and differentially expressed genes in clusters 1 to 4.

Fig. S3. Direct IFN-γ and TNF-α staining in islet-infiltrating CD4+ T cells in three 9-week-old NOD mice.

Fig. S4. Direct detection of CD25+Foxp3+ cells in the spleen, PLNs, and islets of 10-week-old female NOD mice.

Fig. S5. Diabetes incidence in NOD and heterozygote NOD × MHC class II−/− female mice.

Fig. S6. Response of the Ins13–21-specific T cell hybridoma 11T3 to the Ins9–23 peptide presented by NOD and NODβ57D splenocytes.

Table S1. Summary of the various peptide constructs that were used to produce recombinant I-A97/peptide complexes in fly cells.

Table S2. Comparison of all CDR3β analyzed for the present study to their genomic segments.

Table S3. Excel spreadsheet with raw data.

REFERENCES AND NOTES

1. H. O. McDevitt, W. F. Bodmer, HL-A, immune-response genes, and disease. *Lancet* **1**, 1269–1275 (1974).

2. J. Nerup, P. Platz, O. O. Andersen, M. Christy, J. Lyngsoe, J. E. Poulsen, L. P. Ryder, L. S. Nielsen, M. Thomsen, A. Svejgaard, HL-A antigens and diabetes mellitus. *Lancet* **2**, 864–866 (1974).

3. C. Polychronakos, Q. Li, Understanding type 1 diabetes through genetics: Advances and prospects. *Nat. Rev. Genet.* **12**, 781–792 (2011).

4. J. A. Todd, Etiology of type 1 diabetes. *Immunity* **32**, 457–467 (2010).

5. T. A. Aly, A. Ide, M. M. Jahromi, J. M. Barker, M. S. Fernando, S. R. Babu, L. Yu, D. Miao, H. A. Erlich, P. R. Fain, K. J. Barriga, J. M. Norris, M. J. Rewers, G. S. Eisenbarth, Extreme genetic risk for type 1A diabetes. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14074–14079 (2006).

6. J. A. Todd, J. I. Bell, H. O. McDevitt, HLA-DQβ gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* **329**, 599–604 (1987).

7. X. Hu, A. J. Deutsch, T. L. Lenz, S. Onengut-Gumuscu, B. Han, W. M. Chen, J. M. Howson, J. A. Todd, P. I. W. de Bakker, S. S. Rich, S. Raychaudhuri, Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat. Genet.* **47**, 898–905 (2015).

8. A. L. Corper, T. Stratmann, V. Apostolopoulos, C. A. Scott, K. C. Garcia, A. S. Kang, I. A. Wilson, L. Teyton, A structural framework for deciphering the link between I-Ag7 and autoimmune diabetes. *Science* **288**, 505–511 (2000).

9. K. Yoshida, A. L. Corper, R. Herro, B. Jabri, I. A. Wilson, L. Teyton, The diabetogenic mouse MHC class II molecule I-Ag7 is endowed with a switch that modulates TCR affinity. *J. Clin. Invest.* **120**, 1578–1590 (2010).

10. R. R. Latek, A. Suri, S. J. Petzold, C. A. Nelson, O. Kanagawa, E. R. Unanue, D. H. Fremont, 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).

11. A. Suri, J. J. Walters, M. L. Gross, E. R. Unanue, Natural peptides selected by diabetogenic DQ8 and murine I-A(g7) molecules show common sequence specificity. *J. Clin. Invest.* **115**, 2268–2276 (2005).

12. M. van Lummel, P. A. van Veelen, A. Zaldumbide, A. de Ru, G. M. C. Janssen, A. K. Moustakas, G. K. Papadopoulos, J. W. Drijfhout, B. O. Roep, F. Koning, Type 1 diabetes-associated HLA-DQ8 transdimer accommodates a unique peptide repertoire. *J. Biol. Chem.* **287**, 9514–9524 (2012).

13. C. A. Scott, P. A. Peterson, L. Teyton, I. A. Wilson, Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. *Immunity* **8**, 319–329 (1998).

14. T. Stratmann, V. Apostolopoulos, V. Mallet-Designe, A. L. Corper, C. A. Scott, I. A. Wilson, A. S. Kang, L. Teyton, The I-Ag7 MHC class II molecule linked to murine diabetes is a promiscuous peptide binder. *J. Immunol.* **165**, 3214–3225 (2000).

15. Z. Hovhannisyan, A. Weiss, A. Martin, M. Wiesner, S. Tollefsen, K. Yoshida, C. Ciszewski, S. A. Curran, J. A. Murray, C. S. David, L. M. Sollid, F. Koning, L. Teyton, B. Jabri, The role of HLA-DQ8 β57 polymorphism in the anti-gluten T-cell response in coeliac disease. *Nature* **456**, 534–538 (2008).

16. M. G. Levisetti, A. Suri, S. J. Petzold, E. R. Unanue, The insulin-specific T cells of nonobese diabetic mice recognize a weak MHC-binding segment in more than one form. *J. Immunol.* **178**, 6051–6057 (2007).

17. X. Wan, B. H. Zinselmeyer, P. N. Zakharov, A. N. Vomund, R. Taniguchi, L. Santambrogio, M. S. Anderson, C. F. Lichti, E. R. Unanue, Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature* **560**, 107–111 (2018).

18. V. I. Mallet-Designe, T. Stratmann, D. Homann, F. Carbone, M. B. A. Oldstone, L. Teyton, Detection of low-avidity CD4+ T cells using recombinant artificial APC: Following the antiovalbumin immune response. *J. Immunol.* **170**, 123–131 (2003).

19. E. Landais, P. A. Romagnoli, A. L. Corper, J. Shires, J. D. Altman, I. A. Wilson, K. C. Garcia, L. Teyton, New design of MHC class II tetramers to accommodate fundamental principles of antigen presentation. *J. Immunol.* **183**, 7949–7957 (2009).

20. F. Crawford, B. Stadinski, N. Jin, A. Michels, M. Nakayama, P. Pratt, P. Marrack, G. Eisenbarth, J. W. Kappler, Specificity and detection of insulin-reactive CD4+ T cells in type 1 diabetes in the nonobese diabetic (NOD) mouse. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16729–16734 (2011).

21. B. D. Stadinski, L. Zhang, F. Crawford, P. Marrack, G. S. Eisenbarth, J. W. Kappler, Diabetogenic T cells recognize insulin bound to IAg7 in an unexpected, weakly binding register. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10978–10983 (2010).

22. Y. Mansiaux, A. P. Joseph, J. C. Gelly, A. G. de Brevern, Assignment of PolyProline II conformation and analysis of sequence–structure relationship. *PLOS ONE* **6**, e18401 (2011).

23. J. F. Mohan, S. J. Petzold, E. R. Unanue, Register shifting of an insulin peptide-MHC complex allows diabetogenic T cells to escape thymic deletion. *J. Exp. Med.* **208**, 2375–2383 (2011).

24. T. Stratmann, N. Martin-Orozco, V. Mallet-Designe, L. Poirot, D. McGavern, G. Losyev, C. M. Dobbs, M. B. Oldstone, K. Yoshida, H. Kikutani, D. Mathis, C. Benoist, K. Haskins, L. Teyton, Susceptible MHC alleles, not background genes, select an autoimmune T cell reactivity. *J. Clin. Invest.* **112**, 902–914 (2003).

25. M. Holt, A. Costanzo, L. Gioia, B. Abe, A. I. Su, L. Teyton, Gene profiling and T cell receptor sequencing from antigen-specific CD4 T cells. *Methods Mol. Biol.* **1712**, 217–238 (2018).

26. L. A. Kalekar, S. E. Schmiel, S. L. Nandiwada, W. Y. Lam, L. O. Barsness, N. Zhang, G. L. Stritesky, D. Malhotra, K. E. Pauken, J. L. Linehan, M. G. O'Sullivan, B. T. Fife, K. A. Hogquist, M. K. Jenkins, D. L. Mueller, CD4+ T cell anergy prevents autoimmunity and generates regulatory T cell precursors. *Nat. Immunol.* **17**, 304–314 (2016).

27. S. Mukherjee, P. K. Maiti, D. Nandi, Role of CD80, CD86, and CTLA4 on mouse CD4+ T lymphocytes in enhancing cell-cycle progression and survival after activation with PMA and ionomycin. *J. Leukoc. Biol.* **72**, 921–931 (2002).

28. J. R. Podojil, S. D. Miller, Cross-linking of CD80 on CD4+ T cells activates a calcium-dependent signaling pathway. *J. Immunol.* **182**, 766–773 (2009).

29. C. X. Dominguez, R. A. Amezquita, T. Guan, H. D. Marshall, N. S. Joshi, S. H. Kleinstein, S. M. Kaech, The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *J. Exp. Med.* **212**, 2041–2056 (2015).

30. J. Holst, K. M. Vignali, A. R. Burton, D. A. Vignali, Rapid analysis of T-cell selection in vivo using T cell-receptor retrogenic mice. *Nat. Methods* **3**, 191–197 (2006).

31. T. Lund, L. O'Reilly, P. Hutchings, O. Kanagawa, E. Simpson, R. Gravely, P. Chandler, J. Dyson, J. K. Picard, A. Edwards, D. Kioussis, A. Cooke, Prevention of insulin-dependent diabetes mellitus in non-obese diabetic mice by transgenes encoding modified I-A β-chain or normal I-E α-chain. *Nature* **345**, 727–729 (1990).

32. T. Miyazaki, M. Uno, M. Uehira, H. Kikutani, T. Kishimoto, M. Kimoto, H. Nishimoto, J. I. Miyazaki, K. I. Yamamura, 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).

33. R. M. Slattery, L. Kjer-Nielsen, J. Allison, B. Charlton, T. E. Mandel, J. F. A. P. Miller, Prevention of diabetes in non-obese diabetic I-Ak transgenic mice. *Nature* **345**, 724–726 (1990).

34. S. M. Singer, R. Tisch, X. D. Yang, H. O. McDevitt, An Abd transgene prevents diabetes in nonobese diabetic mice by inducing regulatory T cells. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9566–9570 (1993).

35. R. Quartey-Papafio, T. Lund, P. Chandler, J. Picard, P. Ozegbe, S. Day, P. R. Hutchings, L. O'Reilly, D. Kioussis, E. Simpson, Aspartate at position 57 of nonobese diabetic I-Ag7 beta-chain diminishes the spontaneous incidence of insulin-dependent diabetes mellitus. *J. Immunol.* **154**, 5567–5575 (1995).

36. S. M. Singer, R. Tisch, X. D. Yang, H. K. Sytwu, R. Liblau, H. O. McDevitt, Prevention of diabetes in NOD mice by a mutated I-Ab transgene. *Diabetes* **47**, 1570–1577 (1998).

37. Y. G. Chen, C. E. Mathews, J. P. Driver, The role of NOD mice in type 1 diabetes research: Lessons from the past and recommendations for the future. *Front. Endocrinol.* **9**, 51 (2018).

38. J. M. Robertson, P. E. Jensen, B. D. Evavold, DO11.10 and OT-II T cells recognize a C-terminal ovalbumin 323-339 epitope. *J. Immunol.* **164**, 4706–4712 (2000).

39. J. F. Mohan, B. Calderon, M. S. Anderson, E. R. Unanue, 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).

40. G. Fousteri, A. Dave, A. Bot, T. Juntti, S. Omid, M. von Herrath, Subcutaneous insulin B:9-23/IFA immunisation induces Tregs that control late-stage prediabetes in NOD mice through IL-10 and IFNγ. *Diabetologia* **53**, 1958–1970 (2010).

41. M. G. Levisetti, A. Suri, K. Frederick, E. R. Unanue, Absence of lymph nodes in NOD mice treated with lymphotoxin-β receptor immunoglobulin protects from diabetes. *Diabetes* **53**, 3115–3119 (2004).

42. H. J. Fischer, C. Sie, E. Schumann, A. K. Witte, R. Dressel, J. van den Brandt, H. M. Reichardt, The insulin receptor plays a critical role in T cell function and adaptive immunity. *J. Immunol.* **198**, 1910–1920 (2017).

43. S. T. Ferris, P. N. Zakharov, X. Wan, B. Calderon, M. N. Artyomov, E. R. Unanue, J. A. Carrero, The islet-resident macrophage is in an inflammatory state and senses microbial products in blood. *J. Exp. Med.* **214**, 2369–2385 (2017).

44. J. A. Carrero, D. P. McCarthy, S. T. Ferris, X. Wan, H. Hu, B. H. Zinselmeyer, A. N. Vomund, E. R. Unanue, Resident macrophages of pancreatic islets have a seminal role in the initiation of autoimmune diabetes of NOD mice. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10418–E10427 (2017).

45. R. Pujol-Borrell, I. Todd, M. Doshi, G. F. Bottazzo, R. Sutton, D. Gray, G. R. Adolf, M. Feldmann, HLA class II induction in human islet cells by interferon-γ plus tumour necrosis factor or lymphotoxin. *Nature* **326**, 304–306 (1987).

46. L. M. Francisco, V. H. Salinas, K. E. Brown, V. K. Vanguri, G. J. Freeman, V. K. Kuchroo, A. H. Sharpe, PD-L1 regulates the development, maintenance, and function of induced regulatory T cells. *J. Exp. Med.* **206**, 3015–3029 (2009).

47. T. G. Tan, D. Mathis, C. Benoist, Singular role for T-BET+CXCR3+ regulatory T cells in protection from autoimmune diabetes. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14103–14108 (2016).

48. M. L. Sprouse, M. A. Scavuzzo, S. Blum, I. Shevchenko, T. Lee, G. Makedonas, M. Borowiak, M. L. Bettini, M. Bettini, High self-reactivity drives T-bet and potentiates Treg function in tissue-specific autoimmunity. *JCI Insight* **3**, 97322 (2018).

49. M. Nakayama, N. Abiru, H. Moriyama, N. Babaya, E. Liu, D. Miao, L. Yu, D. R. Wegmann, J. C. Hutton, J. F. Elliott, G. S. Eisenbarth, Prime role for an insulin epitope in the development of type 1 diabetes in NOD mice. *Nature* **435**, 220–223 (2005).

50. M. So, C. M. Elso, E. Tresoldi, M. Pakusch, V. Pathiraja, J. M. Wentworth, L. C. Harrison, B. Krishnamurthy, H. E. Thomas, C. Rodda, F. J. Cameron, J. McMahon, T. W. H. Kay, S. I. Mannering, Proinsulin C-peptide is an autoantigen in people with type 1 diabetes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10732–10737 (2018).

51. M. G. Levisetti, D. M. Lewis, A. Suri, E. R. Unanue, Weak proinsulin peptide-major histocompatibility complexes are targeted in autoimmune diabetes in mice. *Diabetes* **57**, 1852–1860 (2008).

52. P. Halbout, J. P. Briand, S. Bécourt, S. Muller, C. Boitard, T cell response to preproinsulin I and II in the nonobese diabetic mouse. *J. Immunol.* **169**, 2436–2443 (2002).

53. V. Pathiraja, J. P. Kuehlich, P. D. Campbell, B. Krishnamurthy, T. Loudovaris, P. T. Coates, T. C. Brodnicki, P. J. O'Connell, K. Kedzierska, C. Rodda, P. Bergman, E. Hill, A. W. Purcell, N. L. Dudek, H. E. Thomas, T. W. Kay, S. I. Mannering, Proinsulin-specific, HLA-DQ8, and HLA-DQ8-transdimer–restricted CD4+ T cells infiltrate islets in type 1 diabetes. *Diabetes* **64**, 172–182 (2015).

54. C. A. Scott, K. C. Garcia, F. R. Carbone, I. A. Wilson, L. Teyton, Role of chain pairing for the production of functional soluble IA major histocompatibility complex class II molecules. *J. Exp. Med.* **183**, 2087–2095 (1996).

55. P. J. Schatz, Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli. Biotechnology* **11**, 1138–1143 (1993).

56. K. C. Garcia, M. D. Tallquist, L. R. Pease, A. Brunmark, C. A. Scott, M. Degano, E. A. Stura, P. A. Peterson, I. A. Wilson, L. Teyton, Alphabeta T cell receptor interactions with syngeneic and allogeneic ligands: Affinity measurements and crystallization. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13838–13843 (1997).

57. D. S. Li, Y. H. Yuan, H. J. Tu, Q. L. Liang, L. J. Dai, A protocol for islet isolation from mouse pancreas. *Nat. Protoc.* **4**, 1649–1652 (2009).

58. M. P. Lefranc, IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.* **31**, 307–310 (2003).

59. H. Celia, E. Wilson-Kubalek, R. A. Milligan, L. Teyton, Structure and function of a membrane-bound murine MHC class I molecule. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5634–5639 (1999).

60. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009).