nature immunology

Article

https://doi.org/10.1038/s41590-025-02343-z

A microenvironment-driven HLA-II-associated insulin neoantigen elicits persistent memory T cell activation in diabetes

Received: 19 August 2025

Accepted: 10 October 2025

Published online: 28 November 2025

Check for updates

Neetu Srivastava[1,2,15], Anthony N. Vomund[1,2,15], Rongzhen Yu[1,2], Orion J. Peterson[1,2], Yuqing Yang[1,2], David P. Turicek[1,2], Omar Abousaway[3], Tiandao Li[4], Lisa Kain[5], Pamela Stone[6], Aisha Ansar[6], Cristina C. Clement[7], Siddhartha Sharma[5], Rima Melhem[1], Bo Zhang[4], Chang Liu[8], Alok V. Joglekar[9,10,11], Hao Hu[12], Chyi-Song Hsieh[3], Laura Campisi[1], Laura Santambrogio[7,13,14], Luc Teyton[5], Emil R. Unanue[1,2], Ana Maria Arbelaez[6], Cheryl F. Lichti[1,2] ✉ & Xiaoxiao Wan[1,2] ✉

The antigenic landscape of autoimmune diabetes reflects a failure to preserve self-tolerance, yet how novel neoantigens emerge in humans remains incompletely understood. Here we designed an immunopeptidomics-based approach to probe HLA-II-bound, islet-derived neoepitopes in patients with type 1 diabetes. We uncovered a Cys→Ser transformation, conserved between mice and humans, that reshapes autoreactivity to insulin at the single-residue level. This transformation, which we call C19S, arises from oxidative remodeling of insulin in stressed pancreatic islets and also occurs in cytokine-activated antigen-presenting cells, contributing to a feed-forward loop of neoepitope formation and presentation. Despite involving just one amino acid, C19S is recognized by HLA-DQ8-restricted, register-specific CD4$^+$ T cells that expand at diabetes onset. These neoepitope-specific CD4$^+$ T cells lack regulatory potential but acquire a poised central memory phenotype that persists throughout disease progression. These findings reveal a distinct, microenvironment-driven route of neoantigen formation that fuels sustained autoreactivity in diabetes.

Type 1 diabetes (T1D) results from the autoimmune destruction of insulin-producing β cells in the pancreatic islets of Langerhans. Despite extensive study of native β cell antigens, the mechanisms underlying the collapse of immune tolerance remain incompletely understood. Neoantigens are novel forms of self-peptides that escape immune tolerance and become targets of autoreactive T cells[1]. Several well-characterized mechanisms contribute to neoantigen formation in T1D. Post-translational modifications (PTMs) encompass a wide range of biochemical changes to β cell proteins, such as citrullination, deamination and carbonylation, which alter amino acid side chains and generate modified epitopes capable of breaking immune tolerance[2–11]. By contrast, hybrid insulin peptides (HIPs) represent a specific mechanism wherein insulin peptides fuse with fragments of other β cell proteins, forming novel junctional sequences that are absent in the native proteome[12]. These studies have both advanced our understanding of T1D pathogenesis and established β cell neoantigens as valuable biomarkers and therapeutic targets.

The diabetic autoimmune process is characterized by continuous epitope spreading, facilitating a dynamic antigenic landscape.

A full list of affiliations appears at the end of the paper. ✉e-mail: clichti@wustl.edu; wanx@wustl.edu

Whether other mechanisms lead to the formation of neoantigens and their presentation to autoreactive T cells remains unclear. In a broader context, cancer immunology studies have revealed that neoantigens can be generated through previously unrecognized mechanisms during chronic inflammation. Although somatic mutations have long been recognized as a major source of tumor-specific neoepitopes, recent studies have described a new class of neoantigens, termed substitutants, which arise when amino acid depletion leads to tRNA misincorporation, replacing one residue with another[13,14]. These findings suggest that non-mutational mechanisms can introduce single-amino acid substitutions in self-proteins and create immunogenic neoepitopes. This concept may be more relevant to autoimmune diseases, where self-tissues typically harbor few somatic mutations. Moreover, because extra-genomic neoantigens cannot be detected by transcriptomic approaches like RNA sequencing, their discovery relies on mass spectrometry (MS)-based immunopeptidomics, with specialized searches required to confirm their often rare presence in the self-proteome or peptidome[15].

Pancreatic β cells are intrinsically prone to oxidative stress and inflammation and are further exposed to several pathological signals throughout diabetes development. These conditions are thought to create a conducive tissue microenvironment that fosters the generation of diverse neoantigens[16–18]. Although studies using non-obese diabetic (NOD) mice have provided valuable insights, identifying HLA-II-presented neoepitopes derived from human islets in vivo remains a challenge. In this study, we applied a β cell degranulation strategy coupled with HLA-II immunopeptidome profiling to systematically identify islet-derived peptides presented in human peripheral blood. Unexpectedly, we detected a group of HLA-II-bound insulin B-chain peptides harboring a cysteine-to-serine (Cys→Ser) transformation at the 19th position (C19S). Unlike conventional PTMs and HIPs, C19S directly alters the sequence of post-synthetic insulin at a single residue. This distinction prompted us to investigate mechanisms of C19S generation and whether it creates a neoepitope that reshapes insulin autoreactivity in both NOD mice and humans with T1D.

Results

Identification of insulin peptides with a Cys→Ser transformation in mice and humans

To probe potential HLA-II-bound neoepitopes derived from human islets in vivo, we used a β cell degranulation strategy combined with immunopeptidome analysis in human peripheral blood mononuclear cells (PBMCs). This approach stems from our previous mouse studies, which showed that in vivo glucose stimulation triggers exocytosis of β cell-derived peptides from pancreatic islets into circulation[19,20]. These released peptides are presented by MHC-II-expressing antigen-presenting cells (APCs) in the blood within 30 to 120 min following glucose injection, providing a window for detecting rare islet-derived peptides in the self MHC-II peptidome of PBMCs[20].

To apply this strategy in humans, we recruited patients with T1D with either 3- or 18-month onset (3mos, $n = 10$; 18mos, $n = 10$), as well as non-diabetic controls (ND, $n = 10$). Most participants carried at least one copy of the risk-conferring DR3-DQ2 or DR4-DQ8 haplotypes[21,22] (Supplementary Table 1). Each participant fasted, had an initial blood draw and then underwent a mixed meal tolerance test (MMTT) to stimulate β cell degranulation (Fig. 1a). We collected blood samples at 90 and 120 min post-MMTT and monitored C-peptide levels at various time points (Fig. 1a). Individual PBMCs were pooled by disease state (ND, 3mos, 18mos) and time point (0, 90, 120 min), yielding 18 samples for sequential isolation of pan-HLA-DQ and pan-HLA-DR peptidomes.

As in prediabetic NOD mice after glucose stimulation[20], native insulin B-chain peptides were also detected in the HLA-DQ and HLA-DR peptidomes of the non-diabetic human samples after MMTT (Extended Data Fig. 1a). Despite reduced C-peptide levels in T1D cohorts (Fig. 1b and Extended Data Fig. 1b), several native insulin peptides were still detected after MMTT, possibly reflecting peptide release from residual islets.

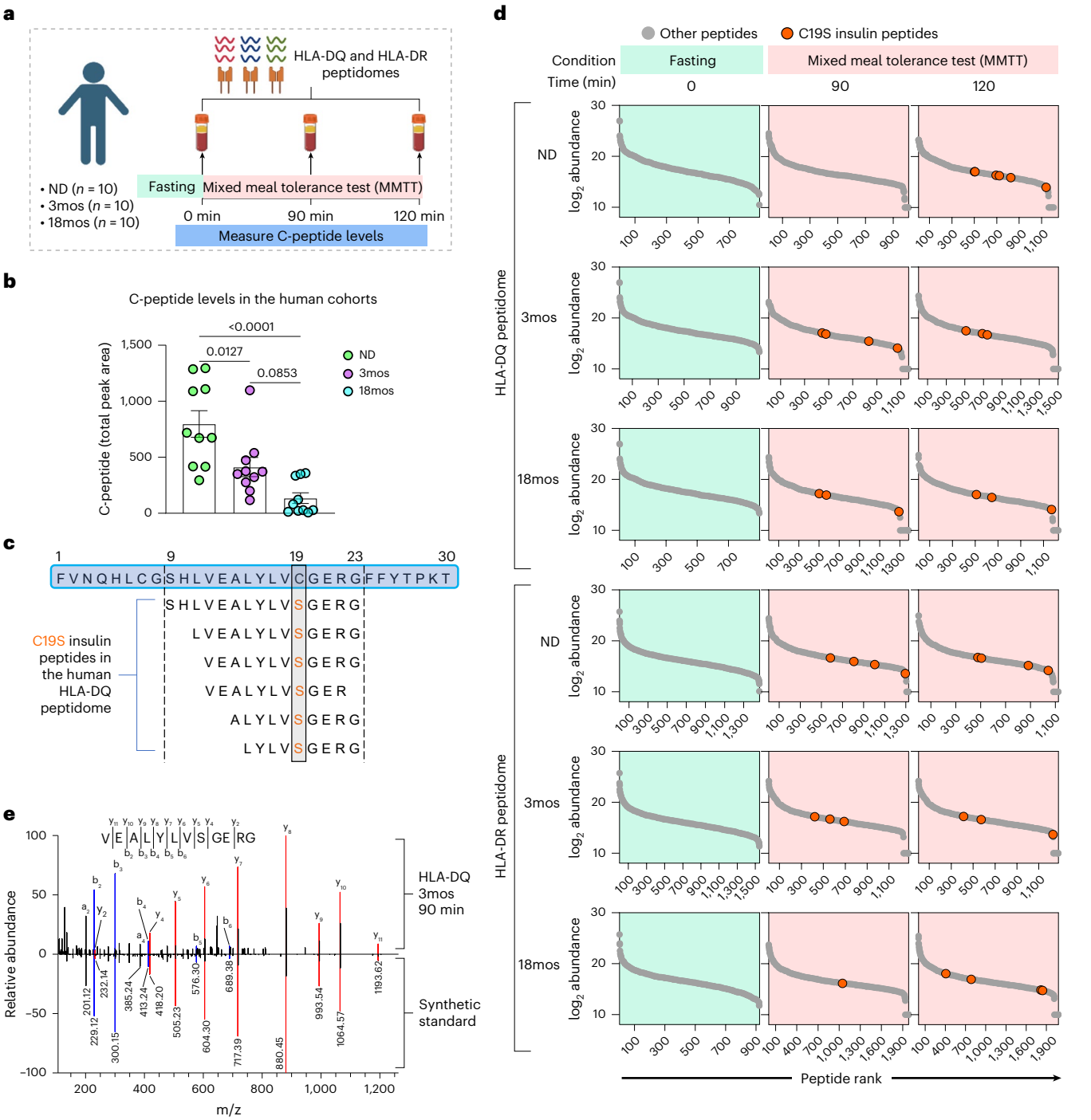
Upon further analysis, we identified a Cys→Ser transformation at position 19 of the insulin B-chain (for example, SHLVEALYLV**C**GERG → SHLVEALYLV**S**GERG), referred to as C19S. To confirm these results, we repeated the search with C19S insulin sequences appended to the canonical UniProt-Human database, which yielded additional C19S insulin peptides. All identified peptides exhibited characteristics typical of MHC-II ligands: they ranged in length from 9 to 15 residues and spanned the B-chain 9-23 segment ($InsB_{9\text{-}23}$) (Fig. 1c), a well-defined MHC-II-binding region in both NOD mice and humans with T1D[19,23–28]. In both the HLA-DQ and HLA-DR peptidomes, C19S insulin peptides were only detected after MMTT (Fig. 1d), confirming the role of MMTT in enabling their detection. For further verification, we synthesized peptides based on 11 unique C19S insulin peptide sequences identified in the HLA-II peptidomes and confirmed that each spectrum from the PBMC samples matched the synthetic standard controls (Fig. 1e and Extended Data Fig. 1c). Collectively, these data indicate that, in addition to native insulin, the C19S variant can be presented by T1D-predisposing HLA-II molecules.

These human studies prompted us to assess C19S insulin peptides in mouse MHC-II peptidomes. First, we reanalyzed the MHC-II ($I\text{-}A^{g7}$) peptidomes from blood leukocytes of prediabetic NOD mice[20] and identified a group of peptides with the same C19S transformation, which also increased in abundance following glucose challenge (Fig. 1f). Moreover, in our previous analysis of the $I\text{-}A^{g7}$ peptidomes of pancreatic islets from NOD mice[26], we identified and verified the exact $InsB_{9\text{-}23}$(C19S) peptide (SHLVEALYLV**S**GERG) (Extended Data Fig. 1d).

In addition to MHC-II-bound peptides, we assessed soluble peptidomes of mesenteric lymph cannulated from 6-week-old female NOD mice with or without glucose challenge. At baseline, the lymph peptidome already contained peptides from several secretory proteins in β cells, such as native insulin (B-chain and C-peptide), islet amyloid polypeptide (IAPP) and chromogranin-A (ChgA) (Fig. 1g). Notably, the $InsB_{9\text{-}23}$(C19S) peptide was identified and verified in the lymph peptidome of glucose-injected mice (Extended Data Fig. 1e). Thus,

Fig. 1 | Identification of β cell-derived insulin peptides with a Cys→Ser transformation in human and mouse MHC-II peptidomes. a, Schematic for analyzing human HLA-DQ and HLA-DR peptidomes of PBMCs from three cohorts of participants before and after an MMTT. ND, non-diabetic; 3mos, 3-month onset; 18mos, 18-month onset. **b**, Bar graphs (mean ± standard error of the mean (s.e.m.)) showing C-peptide levels of the three human cohorts ($n = 10$ per cohort). Each point represents the total peak area of an individual participant, calculated by summing measurements from all time points as shown in Extended Data Fig. 1b. **c**, Individual C19S insulin peptides identified in the HLA-DQ peptidome. NS, not significant. **d**, Rank plots showing each C19S insulin peptide identified in the HLA-DQ (upper) and HLA-DR (lower) peptidomes across the three human cohorts at the indicated time points of MMTT. **e**, A mirror plot showing the complete match of a representative C19S insulin peptide identified in the human HLA-DQ peptidome to its synthetic standard. **f**, Schematic (left) and dot plot (right) showing the identification and quantification of insulin B-chain peptides with C19S ($n = 8$) identified in the blood leukocyte MHC-II peptidome of NOD mice. **g**, Schematic (left) for analyzing soluble lymph peptidome. The stacked bar graph (right) shows the composition of indicated β cell-derived peptides identified in the soluble lymph peptidomes. The data are based on the cumulative abundance of all the peptides belonging to a given β cell protein. In **b**, n is the number of participants in each cohort; each dot is one participant. In **f**, n is the number of peptides; each dot is one peptide. For statistical analysis, one-way analysis of variance (ANOVA) with Tukey's multiple comparisons test was performed for **b**. Two-tailed Wilcoxon signed-rank test was performed for **f**. Illustrations in **a**, **f** and **g** were partly created with BioRender.com.

https://doi.org/10.1038/s41590-025-02343-z



C19S insulin peptides are identified in both soluble and MHC-II-bound forms at disease-relevant sites in NOD mice.

Disease-relevant signals amplify C19S generation in pancreatic β cells

Because peptidome comparisons can be confounded by reduced insulin abundance during T1D, we developed $CD4^+$ T cell hybridomas specifically recognizing the C19S insulin peptide to quantify antigen presentation. From NOD mice immunized with the synthetic $InsB_{9\text{-}23}$(C19S) peptide (SHLVEALYLV**S**GERG), we generated a panel of $CD4^+$ T cell hybridomas (Fig. 2a), including the S5 clone (Fig. 2b), which responded robustly to $InsB_{9\text{-}23}$(C19S) but not to the native $InsB_{9\text{-}23}$ peptide (SHLVEALYLV**C**GERG). Conversely, the native $InsB_{9\text{-}23}$-specific 9B9 clone[19,20,29,30] exhibited minimal reactivity to $InsB_{9\text{-}23}$(C19S) (Fig. 2b). Both T cells exhibited comparable reactivities to their respective peptides (Extended Data Fig. 2a), enabling quantitative comparisons.

Utilizing these epitope-specific T cells, we developed a granule-based antigen presentation (GAP) assay to assess changes in the generation of native versus C19S insulin by β cells under steady-state and inflammatory conditions (Fig. 2c). We focused on β cell granules, the major sites of insulin biosynthesis and autoantigen formation[31,32], and tested insulin dense-core granules (DCGs) and crinosomes (degradative DCG-lysosome fusion granules enriched for peptide fragments)[19,29]. Subcellular granule fractions representing DCGs and crinosomes were isolated from different sources of β cells and validated by granule content and lysosomal activity (acid phosphatase levels) (Extended Data Fig. 2b–d). Furthermore, C3.g7 APCs (B cell lymphoma expressing I-A^{g7}) were incubated with these granules, and presentation of C19S versus native insulin peptides was read out with the S5 and 9B9 hybridomas, respectively (Fig. 2c).

To confirm if C19S can arise in β cells without autoimmunity, as suggested by its detection in HLA-II peptidomes of non-diabetic humans, we tested granules from the MIN6 β cell line and primary islets of diabetes-resistant B6 mice. As reported previously[19,29], the native $InsB_{9\text{-}23}$ peptide (assayed by 9B9) was predominantly found in crinosomes with minimal signal in DCGs (Fig. 2d). The crinosome fraction also activated the S5 T cell, but the responses were 30% to 50% lower than those observed with 9B9 (Fig. 2d). In contrast, S5 responses to DCGs were minimal (Fig. 2d). Thus, under non-autoimmune conditions, C19S insulin peptides are primarily localized in crinosomes at low basal levels.

We next assessed whether C19S generation was associated with diabetes development. In the absence of exogenous antigen pulse, the S5 T cell responded to dispersed islet cells from 5-week-old female NOD mice, with responses further increased by 17 weeks (Fig. 2e), indicating enhanced in vivo presentation during diabetes progression. To corroborate this in disease-relevant sites, we stimulated cells pooled from islets and pancreatic lymph nodes with equal amounts of $InsB_{9\text{-}23}$ and $InsB_{9\text{-}23}$(C19S) peptides and assessed IFNγ production by ELISPOT during antigen recall. As reported previously[26,29], recall with native $InsB_{9\text{-}23}$ induced IFNγ production, whereas recall with the $InsB_{9\text{-}23}$(C19S) peptide triggered even higher responses (Fig. 2f). Thus, C19S presentation progressively increases as diabetes develops, rather than remaining static.

Because cysteine 19 of the insulin B-chain (B(C19)) is the first residue to form an interchain disulfide bond required for insulin folding[33,34], we hypothesized that C19S might reflect insulin misfolding, which is closely linked to endoplasmic reticulum (ER) stress[35,36] and downstream oxidative stress[37]. As cysteine can be converted to serine under oxidative conditions[38], we directly tested the roles of ER and oxidative stress in C19S generation in β cells. We exposed MIN6 cells and B6 mouse islets to the ER stress inducer tunicamycin for 2 h, followed by extensive washing. This treatment did not cause β cell death or alter total protein levels in the granule samples (Extended Data Fig. 2e) but markedly increased S5 T cell responses to crinosomes (Fig. 2g). The tunicamycin-induced increase in C19S was significantly inhibited by glutathione or *N*-acetylcysteine (NAC) in both MIN6 cells and mouse islets (Fig. 2g). In contrast, although tunicamycin also elevated native $InsB_{9\text{-}23}$ in crinosomes, this increase was not affected by antioxidants (Fig. 2h). At baseline, DCGs contained minimal native or C19S peptides, whereas tunicamycin increased both forms (Extended Data Fig. 2f,g). In DCGs, antioxidants also selectively inhibited the increase of C19S (Extended Data Fig. 2f) but not native insulin peptides (Extended Data Fig. 2g). Together, these results indicate that C19S generation is specifically driven by oxidative stress in the islet microenvironment.

T1D development is associated with inflammatory cytokine signaling in pancreatic islets, with IL-1β, IFNγ and TNFα known to induce ER and oxidative stress in β cells[39,40]. A pilot glucose-induced antigen transfer assay suggested that TNFα or IFNγ exposure increased C19S generation, with TNFα showing the strongest effect (Extended Data Fig. 2h). Confirming this finding, the GAP assay revealed a significant increase in S5 T cell responses to crinosome fractions from TNFα-treated MIN6 cells (Fig. 2i). This increase was inhibited by glutathione or tauroursodeoxycholic acid (TUDCA) (Fig. 2i), an ER stress inhibitor known to suppress T1D development[41]. In contrast, TNFα did not affect native $InsB_{9\text{-}23}$ levels and had minimal impact on DCGs (Fig. 2i). Thus, inflammatory cytokines, particularly TNFα, promote C19S generation by exacerbating ER and redox stress.

C19S represents a context-dependent single-amino-acid transformation

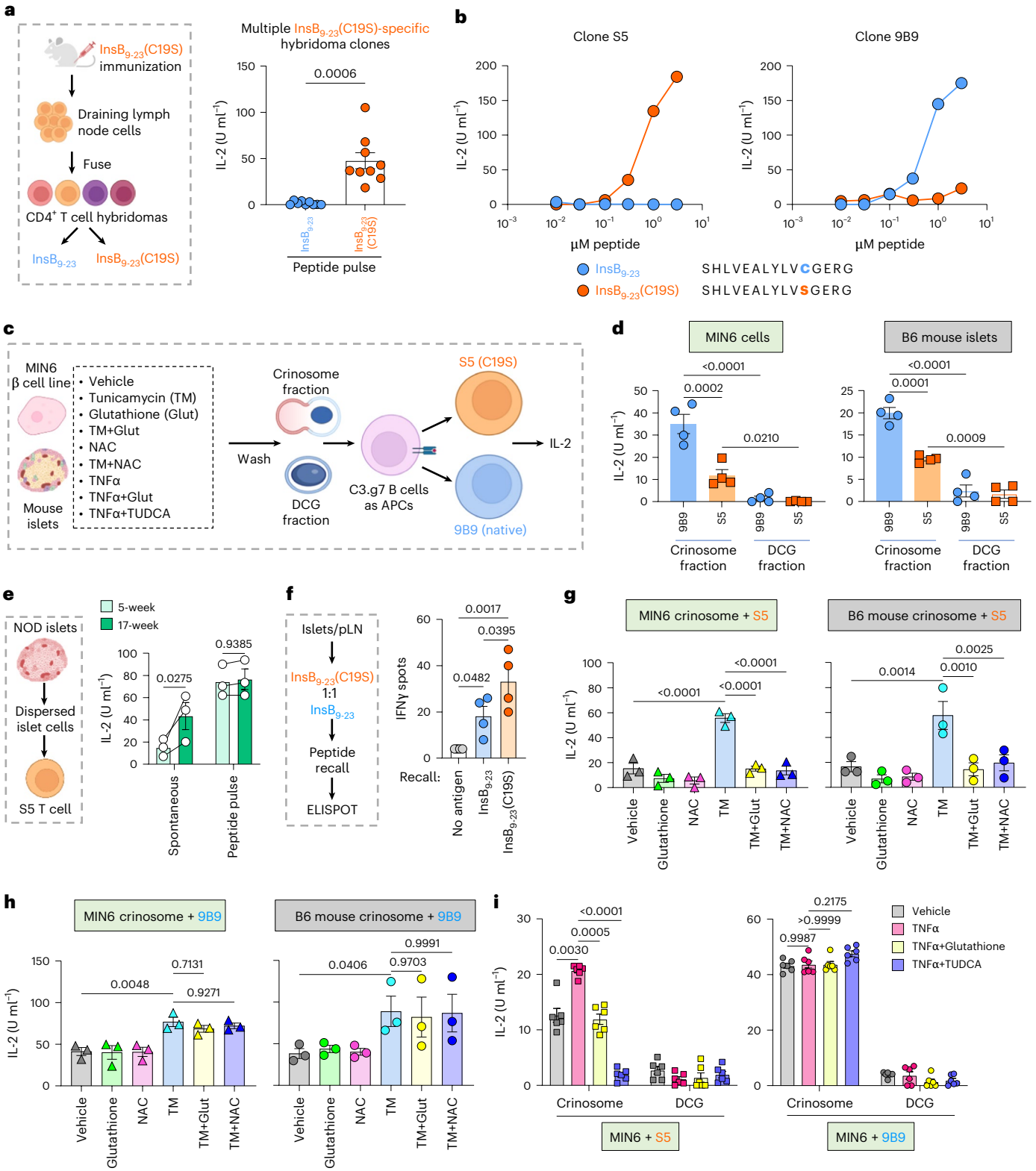
Our data from stressed islets suggest that C19S generation is microenvironment-driven and context-dependent, potentially representing a rare event. Because targeted analyses, such as the GAP assay, examine only predefined peptides, we decided to use an unbiased

Fig. 2 | Disease-relevant signals drive C19S transformation via oxidative stress in β cell granules. a, Schematic (left) for generating $CD4^+$ T cell hybridomas after immunizing NOD mice with the $InsB_{9\text{-}23}$(C19S) peptide. The bar graph (right; mean ± s.e.m.) quantifies responses of individual T cell hybridomas ($n = 9$) to 0.3 μM native $InsB_{9\text{-}23}$ or $InsB_{9\text{-}23}$(C19S) peptide. **b**, Responses of the S5 (left) and 9B9 (right) $CD4^+$ T cell hybridomas to serially diluted concentrations of the native $InsB_{9\text{-}23}$ or the $InsB_{9\text{-}23}$(C19S) peptide. **c**, Schematic of the GAP assay. **d**, Bar graphs (mean ± s.e.m.) showing responses of the 9B9 or the S5 T cell to crinosome or DCG fractions isolated from unmanipulated MIN6 cells or B6 mouse islets ($n = 4$). **e**, Schematic (left) and bar graph (right; mean ± s.e.m.) showing in vivo presentation of C19S insulin peptides in pancreatic islets from 5- and 17-week-old female NOD mice ($n = 3$). **f**, Schematic (left) and bar graph (mean ± s.e.m.) showing ELISPOT assay for T cell responses (IFNγ production) in islets and pancreatic lymph nodes during antigen recall with the $InsB_{9\text{-}23}$ or $InsB_{9\text{-}23}$(C19S) peptide ($n = 4$). **g**, Bar graphs (mean ± s.e.m.) showing responses of the S5 T cell to the crinosome fraction isolated from MIN6 cells and B6 mouse islets following indicated treatments ($n = 3$). **h**, Bar graphs (mean ± s.e.m.) showing responses of the 9B9 T cell to the crinosome fraction isolated from MIN6 cells and B6 mouse islets following indicated treatments ($n = 3$). **i**, Bar graphs (mean ± s.e.m.) showing responses of the S5 or 9B9 T cell to crinosome ($n = 6$) and DCG ($n = 5$) fractions isolated from MIN6 cells stimulated with TNFα, with or without glutathione or TUDCA. In **a**, *n* is the number of individual hybridomas; each dot is one hybridoma. In **d**, **g**, **h** and **i**, *n* is the number of biological replicates; each dot is an independent experiment. In **e** and **f**, *n* is the number of independent experiments; each dot is one sample pooled from multiple mice. For statistical analysis, a two-tailed paired *t*-test was performed for **a**. Repeated measures (RM) one-way ANOVA with Tukey's multiple comparisons test was performed for **f**. RM one-way ANOVA with Sidak's multiple comparisons tests were performed for **d**, **g** and **h**. RM two-way ANOVA with Sidak's multiple comparisons test was performed for **e**. RM two-way ANOVA with Dunnett's multiple comparisons test was performed for **i**. The data represent three (**e**, **g**, **h**), four (**d**, **f**) and five (**a**, **i**) independent experiments. Illustrations in **a**, **c** and **e** were partly created with BioRender.com.

https://doi.org/10.1038/s41590-025-02343-z

approach to examine whether C19S could be detected among the broad spectrum of PTMs and single-amino-acid variants (SAVs) and how ER stress might shape these events. Because the proteome profile of MIN6 cells closely resembles primary β cells[42], we isolated crinosome peptidomes from equal numbers of MIN6 cells treated with vehicle or tunicamycin for 2 h (Fig. 3a). The samples were prepared in triplicate, fractionated to increase depth of coverage and block randomized for MS analysis. We first profiled protein composition with PEAKS DB, then applied sequential PEAKS PTM and SPIDER searches to capture initially unspecified PTMs and all possible SAVs. For quantification, PTM/SAV identifications were collated at the peptide–spectrum match (PSM) level to enable an unbiased comparison between conditions.

Among 348 PTMs and SAVs (Supplementary Table 2), Cys→Ser exhibited the most significant increase after tunicamycin (Fig. 3b). Notably, it was identified as a SAV in the SPIDER search rather than through the PTM search. ER stress also increased several PTMs,



https://doi.org/10.1038/s41590-025-02343-z

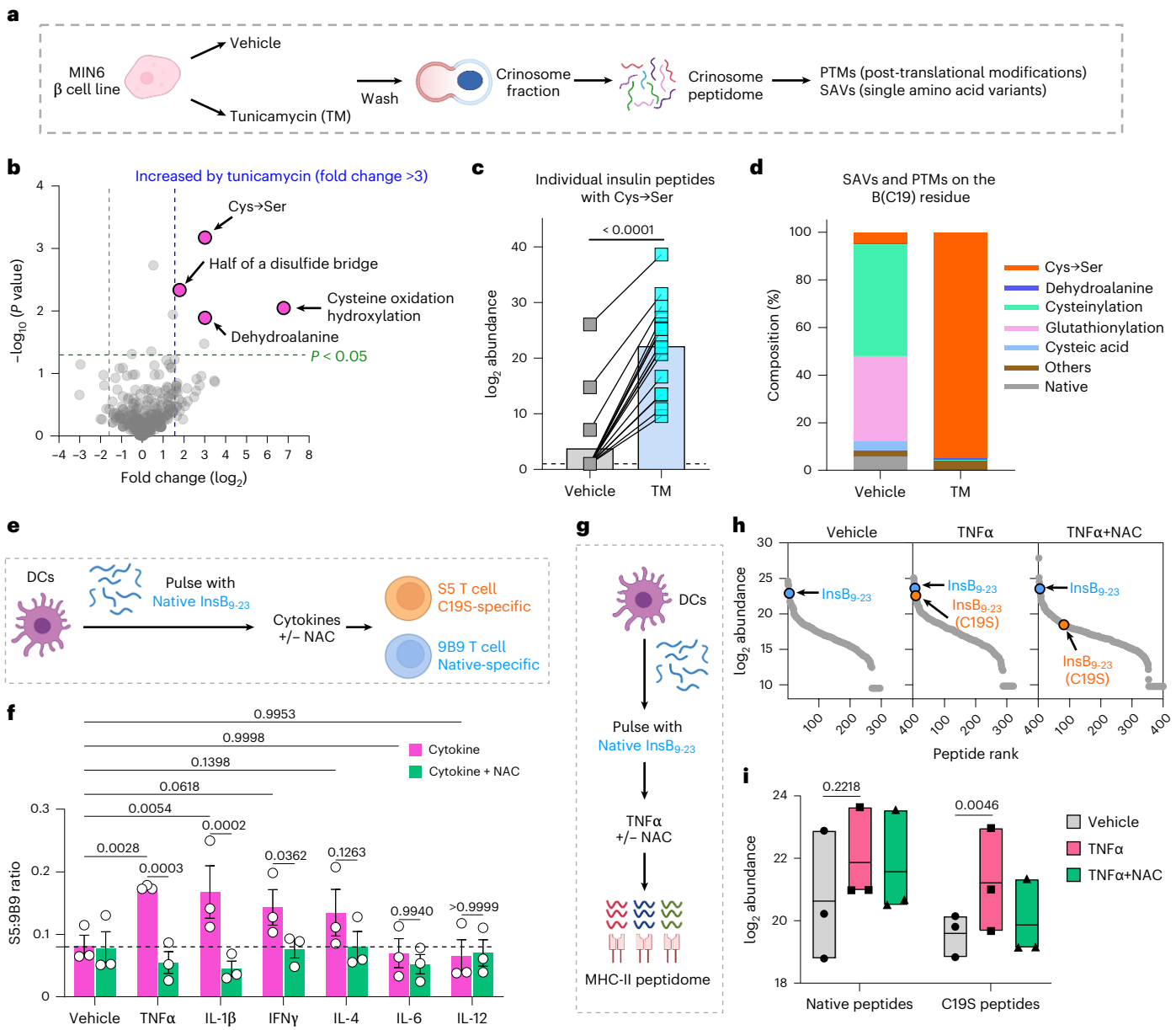


Fig. 3 | C19S represents a context-dependent single-amino-acid transformation emerging in β cells and dendritic cells. a, Schematic of the PTM/SAV search in crinosome peptidomes isolated from MIN6 cells treated with vehicle or tunicamycin for 2 h. **b**, Volcano plot showing changes in PTMs and SAVs identified in the crinosome peptidome of MIN6 cells given tunicamycin relative to vehicle treatment. See a complete list in Supplementary Table 2. **c**, Bar graph showing the abundance of individual insulin peptides with Cys→Ser transformations (each point) identified in the crinosome peptidome of MIN6 cells given vehicle or tunicamycin treatment ($n = 16$). See a complete list in Supplementary Table 3. **d**, Stacked bar graph showing the composition of indicated SAVs and PTMs on the B(C19) residue of insulin peptides identified in the crinosome peptidome of MIN6 cells given vehicle or tunicamycin treatment. **e**, Schematic of the antigen presentation assay to assess the presentation of native or C19S insulin peptides by DCs pulsed with native $InsB_{9\text{-}23}$ and stimulated with different cytokines. **f**, Bar graph (mean ± s.e.m.) showing C19S transformation from the native $InsB_{9\text{-}23}$ peptide in DCs upon cytokine stimulation, shown as the ratio of S5 to 9B9 T cell responses with or without NAC treatment ($n = 3$). **g**, Schematic of MHC-II immunopeptidome analysis of $InsB_{9\text{-}23}$-pulsed DCs stimulated with TNFα with or without NAC. **h**, Rank plots showing all identified peptides in the DC MHC-II peptidomes ranked by abundance. **i**, Box plot showing the cumulative abundance of native and C19S insulin peptides identified in the DC MHC-II immunopeptidome analyses ($n = 3$). In **c**, n is the number of peptides; each dot is one peptide. In **f** and **i**, n is the number of independent experiments; each dot is one biological replicate. Box plot in **i** shows the minimum and maximum boundaries and the mean (center line). For statistical analysis, two-tailed Wilcoxon signed-rank test was performed for **c**. RM two-way ANOVA with Sidak's multiple comparisons tests were performed for **f** and **i**. The data represent three independent experiments. Illustrations in **a**, **e** and **g** were partly created with BioRender.com.

especially cysteine oxidation and hydroxylation, but to a lesser extent than Cys→Ser (Fig. 3b). We did not observe significant changes in oxidative modifications on other residues (for example, histidine, tryptophan, proline or methionine) (Supplementary Table 2), suggesting that the observed cysteine oxidations reflect biological changes rather than sample preparation artifacts. Tracing Cys→Ser back to source peptides revealed a strong enrichment on the insulin B-chain. To verify this, we first created a database in which all cysteines were replaced by serine and found no additional peptides with Cys→Ser. We then appended Cys→Ser variants of abundant β cell proteins (insulin-1, insulin-2, IAPP, ChgA and Scg-1) to the canonical database and repeated the search. This confirmed the identification of Cys→Ser in various insulin B-chain

peptides (a total of 16) (Supplementary Table 3). All these peptides exhibited a significant increase in their abundance following tunicamycin treatment (Fig. 3c; Supplementary Table 3).

We then specifically compared Cys→Ser with other PTMs on the B(C19) residue. In the vehicle-treated control, B(C19) exhibited PTMs such as cysteinylation, glutathionylation and formation of cysteic acid and dehydroalanine, whereas Cys→Ser was only a minor component (Fig. 3d). However, after the 2-h tunicamycin treatment, Cys→Ser emerged as the dominant event (Fig. 3d), indicating that the stressed β cell microenvironment promotes this transformation over other PTMs. Thus, Cys→Ser is context dependent and preferentially occurs at the redox-sensitive B(C19) residue in stressed β cells.

C19S transformation in cytokine-activated dendritic cells

We hypothesized that effective presentation of C19S insulin peptides may also require amplification in APCs. To test this, we examined dendritic cells (DCs), which are known to migrate to inflamed islets and present insulin peptides[29,43]. DCs were enriched in vivo from NOD mice after recombinant FLT3L treatment, pulsed with native $InsB_{9\text{-}23}$ and divided into equal aliquots for cytokine stimulation with or without NAC (Fig. 3e). C19S conversion was quantified by the S5:9B9 response ratio. IL-1β, TNFα and IFNγ significantly increased the S5:9B9 ratio, whereas IL-4, IL-6 and IL-12 had weaker effects (Fig. 3f). NAC largely abrogated the increases induced by IL-1β, TNFα and IFNγ (Fig. 3f), indicating oxidative dependence.

The antigen presentation assay suggested that cytokine stimulation allowed C19S insulin peptides to enter the broad and competitive DC self-peptidome. To test this directly, we pulsed DCs with $InsB_{9\text{-}23}$, treated them with TNFα or TNFα plus NAC, and isolated I-A^{g7} peptidomes for MS analysis (Fig. 3g). Although the native $InsB_{9\text{-}23}$ peptide was identified under all conditions, the $InsB_{9\text{-}23}$(C19S) peptide appeared in TNFα-stimulated DCs and was reduced by NAC (Fig. 3h). This pattern was consistently observed across three independent DC peptidomes (Fig. 3i), demonstrating de novo C19S formation in cytokine-stimulated DCs.

C19S alters $CD4^+$ T cell recognition in a register-dependent manner

Based on the generation of $InsB_{9\text{-}23}$(C19S)-specific $CD4^+$ T cell hybridomas (Fig. 2a), we examined whether C19S alters I-A^{g7} binding and register recognition. C19S is positioned within two 9-mer I-A^{g7}-binding cores, $InsB_{12\text{-}20}$ (VEALYLV**C**G) and $InsB_{13\text{-}21}$ (EALYLV**C**GE), which use G20 and E21, respectively, as their P9 anchor residues[44]. This register shift allows native $InsB_{12\text{-}20}$ and $InsB_{13\text{-}21}$ to be recognized by two distinct $CD4^+$ T cell populations through a mechanism called P9 switch[27].

To evaluate the effects of C19S, we synthesized peptides with $InsB_{12\text{-}20}$(C19S) (VEALYLV**S**G) or $InsB_{13\text{-}21}$(C19S) (EALYLV**S**GE) as putative registers, each nested within surrogate flanks[44] (Extended Data Fig. 3a). In the $InsB_{13\text{-}21}$ register, C19S occupies P7, an I-A^{g7}-binding pocket, whereas in the $InsB_{12\text{-}20}$ register, C19S introduces a new P8 TCR contact. As a promiscuous peptide binder, I-A^{g7} has an unusual binding groove with a widened entrance to the P9 pocket, suggesting that changes in P8 could also impact binding[45,46]. By both cell-based (Extended Data Fig. 3b) and biochemical (Extended Data Fig. 3c) binding assays, all peptides with C19S exhibited a marked reduction in binding compared to their native counterparts, with $InsB_{12\text{-}20}$(C19S) being the weakest binder. Confirming this, a surrogate $InsB_{12\text{-}20}$(C19A) peptide similarly reduced binding (Extended Data Fig. 3d), corroborating recent evidence that P8 residues can modulate MHC-II binding[47,48]. Despite reduced MHC-II binding, most $InsB_{9\text{-}23}$(C19S)-specific T cell hybridomas preferentially recognized the $InsB_{12\text{-}20}$(C19S) register, with only a few recognizing $InsB_{13\text{-}21}$(C19S) (Extended Data Fig. 3e). Furthermore, substituting G20 or E21 with the inhibitory residue K completely abolished T cell recognition (Extended Data Fig. 3e), confirming these as P9 anchors in C19S-containing registers.

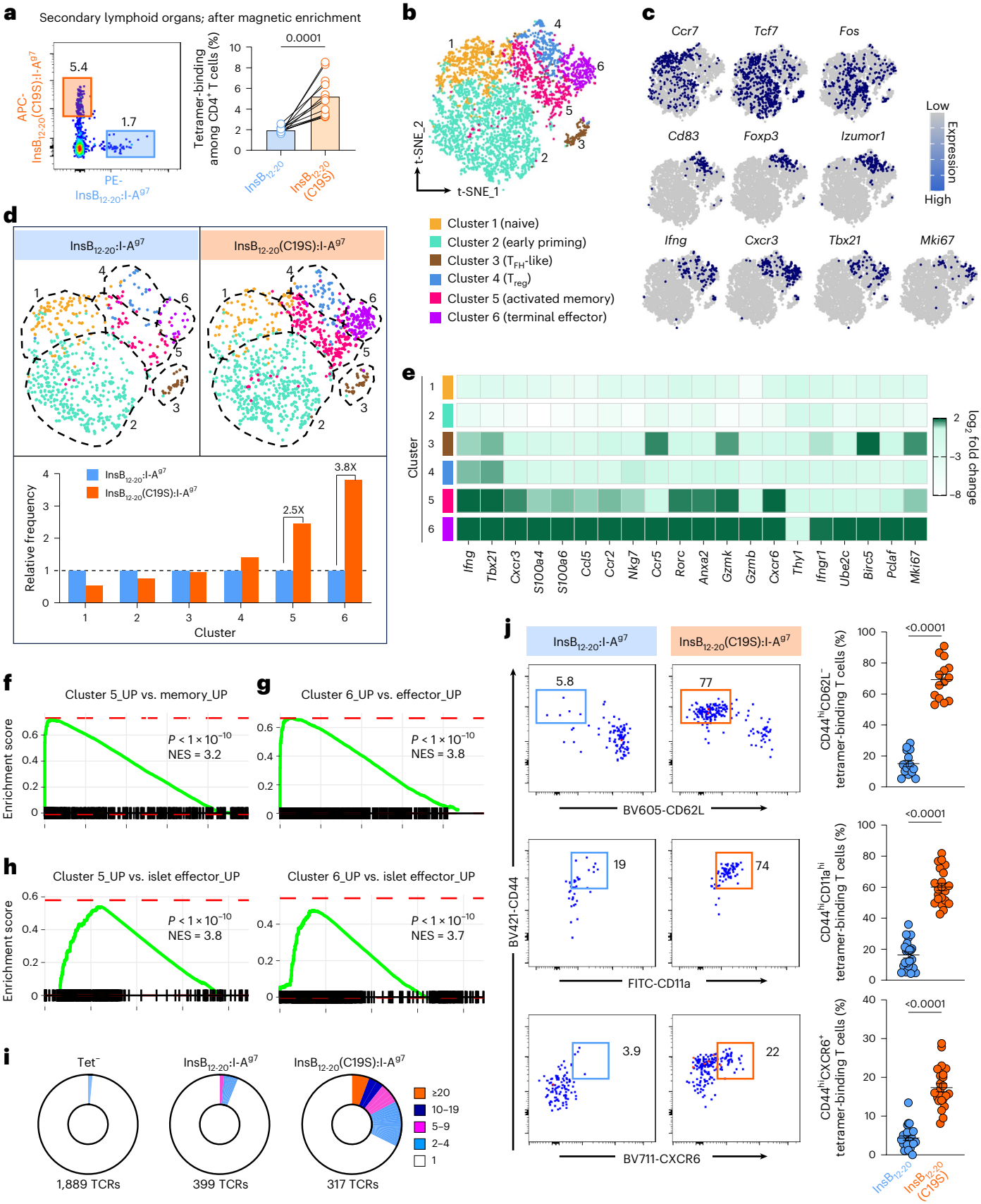
To test whether C19S alters T cell recognition, we generated I-A^{g7}-based tetramers incorporating the 9-mer $InsB_{12\text{-}20}$(C19S) and $InsB_{13\text{-}21}$(C19S) epitopes (Extended Data Fig. 3f). Using the magnetic enrichment protocol[49] (Extended Data Fig. 3g), neither tetramer stained $CD8^+$ T cells in NOD mice nor $CD4^+$ T cells from a NOD strain expressing I-A^b instead of I-A^{g7} (Extended Data Fig. 3h), confirming MHC restriction. The $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer identified a distinct $CD4^+$ T cell population in secondary lymphoid organs (SLOs) of 8-week-old female NOD mice, with minimal overlap and consistently higher frequencies than those detected by the native $InsB_{12\text{-}20}$:I-A^{g7} tetramer (Fig. 4a). Confirming this, an $InsB_{12\text{-}20}$(C19A):I-A^{g7} tetramer also identified a distinct T cell population (Extended Data Fig. 3i). In contrast, the $InsB_{13\text{-}21}$:I-A^{g7} and $InsB_{13\text{-}21}$(C19S):I-A^{g7} tetramers labeled an overlapping population, indicating cross-reactivity (Extended Data Fig. 3j). Thus, C19S in the $InsB_{12\text{-}20}$ register corresponds to a distinct $CD4^+$ T cell population in vivo.

Distinct transcriptional heterogeneity of C19S neoepitope-reactive $CD4^+$ T cells

Because the $InsB_{12\text{-}20}$(C19S) epitope only differs from its native counterpart by a single residue, we next asked whether C19S-specific $CD4^+$ T cells show distinctive transcriptional features. We performed single-cell RNA sequencing (scRNA-seq) on $InsB_{12\text{-}20}$(C19S):I-A^{g7} and $InsB_{12\text{-}20}$:I-A^{g7} tetramer-binding $CD4^+$ T cells, with tetramer-negative (Tet^-) $CD4^+$ T cells as controls. Unsupervised clustering of the combined datasets revealed six clusters (Fig. 4b) characterized by differential expression of phenotypic T cell genes (Fig. 4c). Cluster 1 corresponded to naïve T cells, whereas cluster 2 showed features of early priming observed in the $InsB_{12\text{-}20}$-specific 8F10 T cells during weak antigen encounter[19] (Extended Data Fig. 4a) and contained about half of the $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding cells (Extended Data Fig. 4b). Cluster 3 exhibited T follicular helper (T_{FH})-like signatures[50] (Extended Data Fig. 4c) and was equally detected in both

Fig. 4 | C19S-specific $CD4^+$ T cells exhibit distinct transcriptional states marked by peripheral activation and clonal expansion. a, Representative flow cytometry plot (left) showing co-staining with $InsB_{12\text{-}20}$:I-A^{g7} and $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramers, revealing two distinct, register-specific populations. The bar graph (right) summarizes the frequencies of the tetramer-binding populations (n = 12). **b**, t-distributed stochastic neighbor embedding (t-SNE) plot showing T cell clusters merged from $InsB_{12\text{-}20}$-specific, $InsB_{12\text{-}20}$(C19S)-specific and Tet^- $CD4^+$ T cells (n = 4,787). **c**, Feature plots showing the expression of functional genes classifying $CD4^+$ T cell clusters. **d**, t-SNE plots (upper) showing cluster distribution for $InsB_{12\text{-}20}$:I-A^{g7} (*n* = 719) and $InsB_{12\text{-}20}$(C19S):I-A^{g7} (*n* = 1,256) tetramer-binding $CD4^+$ T cells. The bar graph (lower) shows the relative frequency of each cluster. **e**, Heat map showing the expression of indicated T cell activation and cycling genes among all clusters. **f**, GSEA showing enrichment of cluster 5 for memory T cell signatures from dataset GSE9650. **g**, GSEA showing enrichment of cluster 6 for effector T cell signatures from dataset GSE9650. **h**, GSEA showing enrichment of both clusters 5 and 6 for intra-islet, terminally activated $CD4^+$ T cells depicted in dataset GSE262101. **i**, Pie charts showing clonal overlap of TCRs identified in tetramer-negative, $InsB_{12\text{-}20}$:I-A^{g7} and $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding $CD4^+$ T cells. Clonal overlap was defined as TCRs with identical amino acid sequences in the full CDR3 regions of both α and β chains. **j**, Representative flow cytometry plots (left) showing indicated activation markers on $InsB_{12\text{-}20}$:I-A^{g7} and $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding $CD4^+$ T cells in SLOs of 8-week-old female NOD mice. The dot plots (right; mean ± s.e.m.) summarize the frequencies of $CD44^{hi}CD62L^-$ (*n* = 14), $CD44^{hi}CD11a^{hi}$ (*n* = 22) and $CD44^{hi}CXCR6^+$ (*n* = 23) cells in the tetramer-binding T cells. In **a** and **j**, *n* is the number of mice; each dot is one mouse. In **b** and **d**, *n* is the number of cells; each dot is one cell. For statistical analysis, two-tailed paired *t*-test was performed for **a**. One-way ANOVA with Dunnett's multiple comparisons test was performed for **j**. The data represent two (**b**–**i**) and 6-10 (**a**, **j**) independent experiments.

https://doi.org/10.1038/s41590-025-02343-z



https://doi.org/10.1038/s41590-025-02343-z

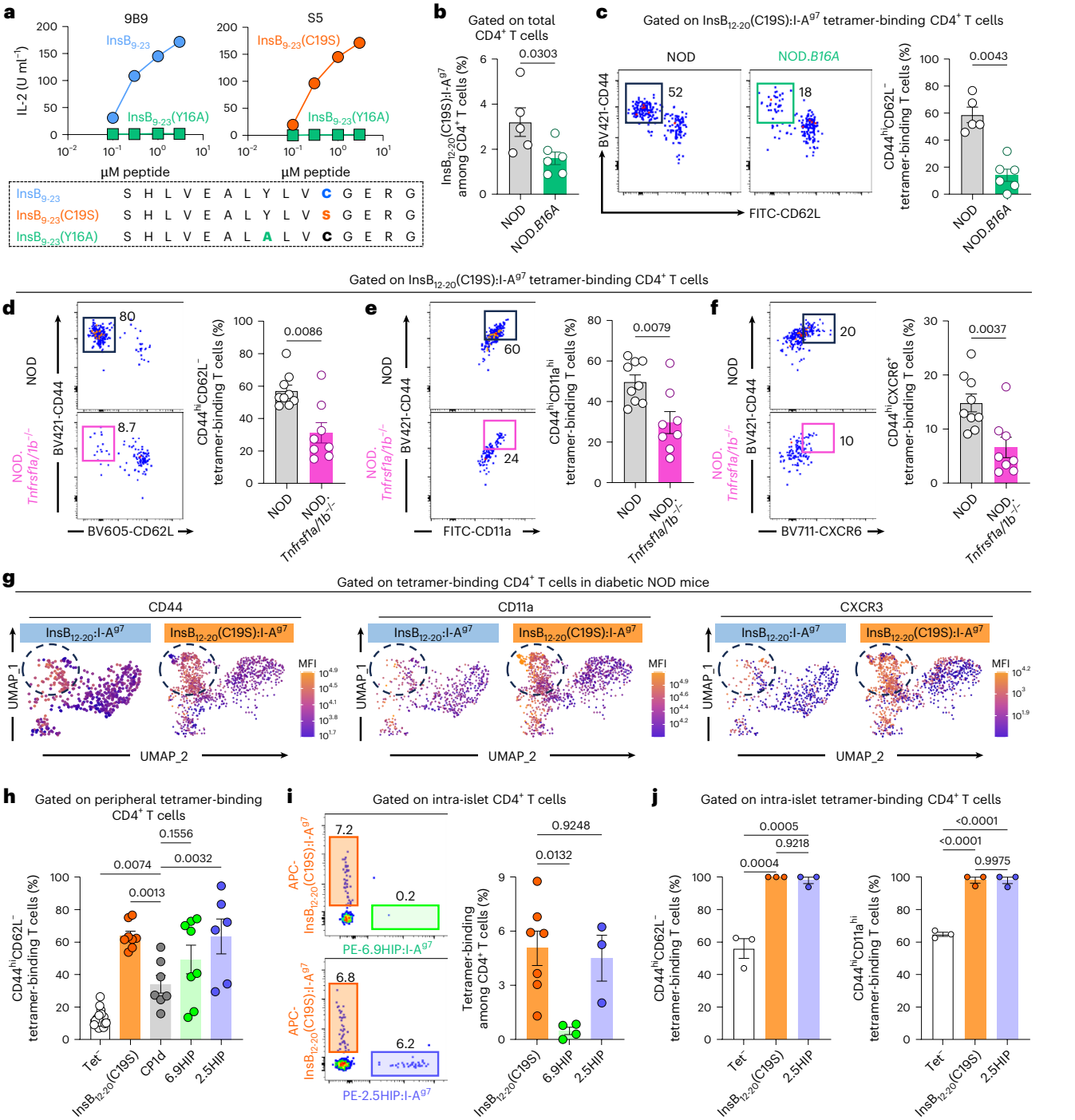


Fig. 5 | Epitope availability and inflammatory cues regulate peripheral activation, tissue infiltration and diabetogenic potential of C19S-specific CD4⁺ T cells. a, IL-2 production by 9B9 and S5 T cells in response to $InsB_{9\text{-}23}$, $InsB_{9\text{-}23}$(C19S) and $InsB_{9\text{-}23}$(Y16A) peptides. **b**, Bar graph (mean ± s.e.m.) summarizing the frequencies of $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding CD4⁺ T cells in SLOs of 8-week-old female NOD ($n = 5$) and NOD.*B16A* ($n = 6$) mice. **c**, Representative flow cytometry plot (left) and bar graph (right; mean ± s.e.m.) showing the percentage of $CD44^{hi}CD62L^{-}$ T cells within $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding CD4⁺ T cells in SLOs of 8-week-old female NOD ($n = 5$) and NOD.*B16A* ($n = 6$) mice. **d**–**f**, Representative flow cytometry plots (left) and bar graphs (right; mean ± s.e.m.) showing the expression of indicated activation markers on $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding CD4⁺ T cells in SLOs of 8-week-old female NOD ($n = 9$) and NOD.*Tnfrsf1a/1b*$^{-/-}$ ($n = 8$) mice. **g**, Feature plots illustrating the expression of CD44, CD11a and CXCR3 on $InsB_{12\text{-}20}$:I-A^{g7} versus $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding CD4⁺ T cells from SLOs of diabetic NOD mice. Shown are equal numbers of tetramer-binding T cells for either epitope merged from eight mice. **h**, Bar graph (mean ± s.e.m.) showing the frequencies of $CD44^{hi}CD62L^{-}$ cells within tetramer-negative ($n = 21$), $InsB_{12\text{-}20}$(C19S):I-A^{g7} ($n = 8$), CP1d:I-A^{g7} ($n = 7$), 6.9HIP:I-A^{g7} ($n = 8$) or 2.5HIP:I-A^{g7} ($n = 6$) tetramer-binding CD4⁺ T cells in SLOs from 8-week-old female NOD mice. **i**, Representative flow cytometry plot (left) showing co-staining of intra-islet CD4⁺ T cells with the $InsB_{12\text{-}20}$(C19S):I-A^{g7} ($n = 7$) tetramer, along with the 6.9HIP:I-A^{g7} ($n = 4$) or 2.5HIP:I-A^{g7} ($n = 3$) tetramer. The bar graph (right; mean ± s.e.m.) summarizes the frequencies of indicated tetramer-binding cells among intra-islet CD4⁺ T cells. **j**, Percentage of activated $CD44^{hi}CD62L^{-}$ and $CD44^{hi}CD11a^{hi}$ T cells among indicated intra-islet tetramer-negative ($n = 3$), $InsB_{12\text{-}20}$(C19S):I-A^{g7} ($n = 3$), or 2.5HIP:I-A^{g7} ($n = 3$) tetramer-binding CD4⁺ T cells. In **b**, **c**–**f**, **h**–**j**, n is the number of mice; each dot is one mouse. For statistical analysis, two-tailed Mann-Whitney tests were performed for **b**–**f**. One-way ANOVA with Dunnett's multiple comparisons test was performed for **h**. One-way ANOVA with Tukey's multiple comparisons tests were performed for **i** and **j**. The data in all panels represent at least three independent experiments.

https://doi.org/10.1038/s41590-025-02343-z

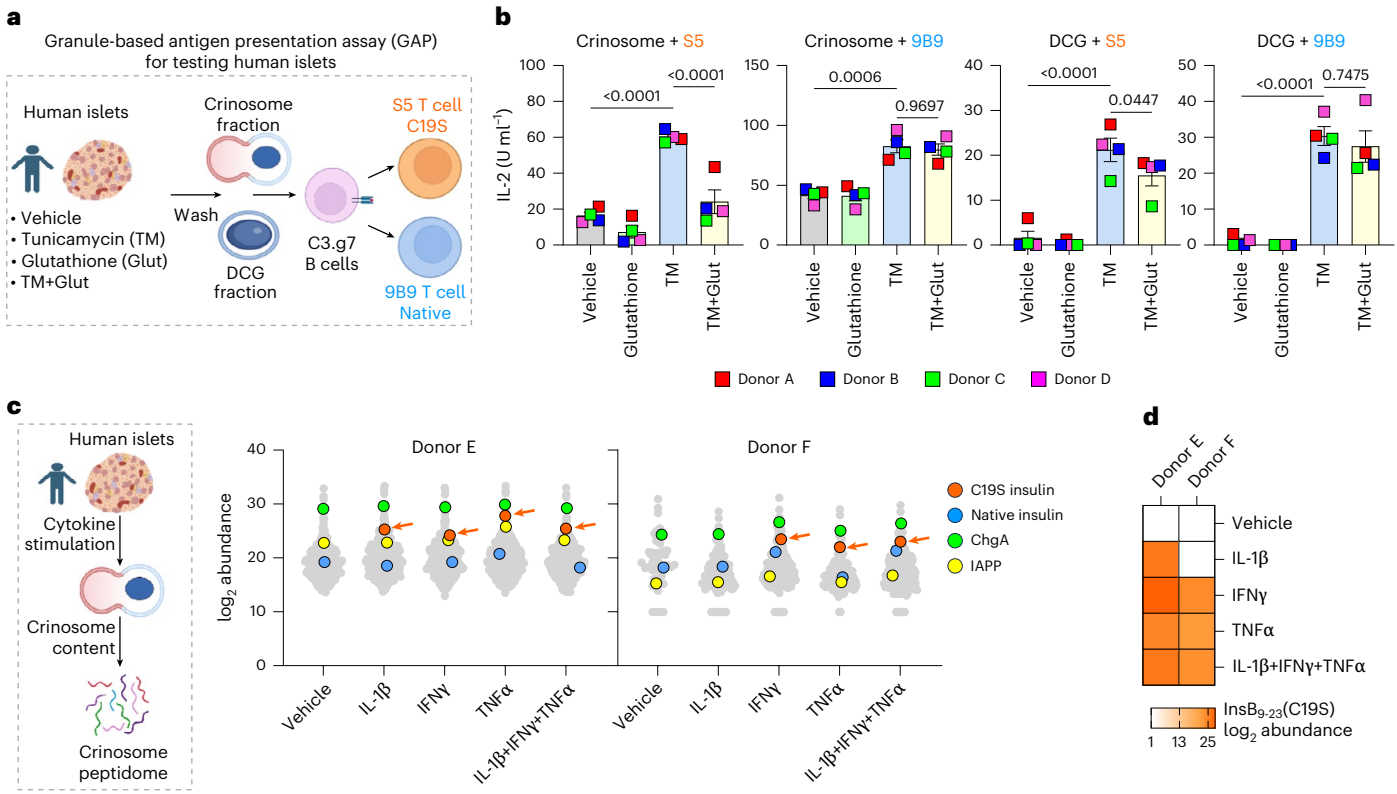


Fig. 6 | C19S transformation occurs in human islets during stress and inflammation. a, Schematic of the GAP assay using human islets as a source of β cell granules. **b**, Bar graphs (mean ± s.e.m.) showing the responses of the S5 and 9B9 T cells to crinosomes and DCGs isolated from human islets following indicated treatments ($n = 4$). **c**, Schematic (left) and dot plots (right) showing C19S insulin peptides in crinosome peptidomes of human islets from two different donors exposed to indicated cytokines. **d**, Heatmap showing the appearance of the $InsB_{9\text{-}23}$(C19S) peptide in the crinosome peptidomes of two human donors under different cytokine conditions. In **b**, *n* is the number of islet samples from different human donors; each dot is one donor. For statistical analysis, RM one-way ANOVA with Sidak's multiple comparisons test was performed for **b**. The data represent two (**c**, **d**) and four (**b**) independent experiments. Illustrations in **a** and **c** were partly created with BioRender.com.

$InsB_{12\text{-}20}$(C19S):I-A^{g7} and $InsB_{12\text{-}20}$:I-A^{g7} tetramer-binding $CD4^+$ T cells (Extended Data Fig. 4b).

Despite these similarities, $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding cells were highly enriched in clusters 5 and 6 (Fig. 4d and Extended Data Fig. 4b), which are associated with T cell activation, memory formation and terminal differentiation. Both clusters upregulated activation genes, such as *Ifng*, *Cxcr6* and *Tbx21* (Fig. 4e and Extended Data Fig. 4d). By gene set enrichment analysis (GSEA), cluster 5 was also enriched for memory T cell signatures[51] (Fig. 4f), consistent with early memory responses observed in NOD mice[52]. Cluster 6 showed the strongest expression of effector T cell genes, with further upregulation of cell cycling genes such as *Ube2c*, *Birc5*, *Pclaf* and *Mki67* (Fig. 4e), and was significantly enriched for defined effector T cell signatures[51] (Fig. 4g). Furthermore, both cluster 5 and cluster 6 were enriched for the gene profile of a previously identified intra-islet effector $CD4^+$ T cell subset associated with terminal activation[53] (Fig. 4h).

Single-cell TCR sequencing revealed a higher degree of clonal overlap (TCRs with identical CDR3α and CDR3β amino acid sequences) in $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding cells (Fig. 4i and Extended Data Fig. 4e), indicating clonal expansion. To further validate the transcriptional data, we used flow cytometry to assess several activation markers highly expressed by clusters 5 and 6, including CD44, CD11a and CXCR6. $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding cells consistently showed significantly higher proportions of $CD44^{hi}CD62L^-$, $CD44^{hi}CD11a^{hi}$ and $CD44^{hi}CXCR6^+$ subsets, as compared to $InsB_{12\text{-}20}$:I-A^{g7} tetramer-binding cells (Fig. 4j), demonstrating a more activated profile of $InsB_{12\text{-}20}$(C19S)-specific $CD4^+$ T cells at the protein level.

C19S neoepitope availability and inflammatory cues drive cognate T cell activation

The high activation state of $InsB_{12\text{-}20}$(C19S)-specific $CD4^+$ T cells led us to ask whether this profile depends on in vivo availability of the C19S neoepitope. We first sought to ablate C19S directly; however, this is impractical, as the B(C19) residue is essential for proper insulin folding[54]. While evaluating alternative models, we unexpectedly found that the mutant transgenic insulin expressed by the NOD.*B16A* mouse, which carries a tyrosine-to-alanine substitution at position 16 of the B-chain (Y16A)[23], did not support T cell recognition of C19S. As shown previously[19,23], the native $InsB_{9\text{-}23}$-specific 9B9 T cell did not respond to the $InsB_{9\text{-}23}$(Y16A) peptide (SHLVEAL**A**LVCGERG) (Fig. 5a). However, the $InsB_{9\text{-}23}$(C19S)-specific S5 T cell also failed to respond at any peptide concentration (Fig. 5a). Relative to sex/age-matched NOD controls, NOD.*B16A* mice showed significantly lower frequencies (Fig. 5b) and activation (Fig. 5c) of $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding $CD4^+$ T cells, indicating that in vivo epitope availability is required for the maintenance and activation of $InsB_{12\text{-}20}$(C19S)-specific $CD4^+$ T cells.

Given that TNFα promotes C19S generation in both β cells and DCs, we assessed NOD mice deficient in both TNF receptor genes (NOD.*Tnfrsf1a/1b*$^{-/-}$). These mice had significantly lower proportions of $CD44^{hi}CD62L^-$ (Fig. 5d), $CD44^{hi}CD11a^{hi}$ (Fig. 5e) and $CD44^{hi}CXCR6^+$ (Fig. 5f) $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding $CD4^+$ T cells, indicating that TNFα signaling supports optimal activation. To distinguish effects on peptide generation from co-stimulation, we examined overtly diabetic NOD mice, when C19S generation is likely reduced due to β cell loss, whereas inflammatory cues are prominent. Using CD44, CD11a

and CXCR3 as readouts, we observed an evident population of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells with high expression of all three activation markers, whereas T cells specific for the native epitope showed substantially lower expression (Fig. 5g). Thus, the diabetic inflammatory milieu can sustain elevated activation of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells even when antigen generation is diminished.

C19S neoepitope-reactive $CD4^+$ T cells infiltrate islets and acquire terminal activation

Both NOD.*B16A* and NOD.*Tnfrsf1a/1b*$^{-/-}$ mice in our colony were completely resistant to diabetes development (Extended Data Fig. 5a), consistent with previous studies[23,55]. The observed inactivation of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells in both strains suggested their relevance to disease pathogenesis. To test this, we enriched polyclonal $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells from $InsB_{9-23}$(C19S)-immunized NOD mice, expanded them in vitro, and adoptively transferred them to NOD.*Rag1*$^{-/-}$ recipients. Given that some $InsB_{9-23}$(C19S)-reactive T cells responded to the $InsB_{13-21}$(C19S) epitope (Extended Data Fig. 3d), we selectively expanded T cells using the weak-binding $InsB_{12-20}$(C19S) peptide to maintain register specificity. Compared to studies using similar approaches[26,56,57], this method yielded a limited number of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells for transfer (Extended Data Fig. 5b). Despite this, diabetes development was evident in NOD.*Rag1*$^{-/-}$ recipients of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells, although the incidence was incomplete (50%) (Extended Data Fig. 5c). Moreover, co-transfer of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells with splenic $CD8^+$ T cells (purified from NOD mice) resulted in 100% diabetes incidence (Extended Data Fig. 5d). In contrast, $CD8^+$ T cells alone, or $CD8^+$ T cells plus $CD4^+$ T cells enriched against a scrambled, non-cross-reactive, but immunogenic control peptide (Extended Data Fig. 5e,f), failed to induce diabetes (Extended Data Fig. 5d).

To further evaluate diabetogenic potential under physiological conditions, we assessed the peripheral activation and islet localization of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells along with $CD4^+$ T cells recognizing three well-characterized islet neoepitopes: the deamidated insulin-1 C-peptide (CP1d; GDLQTLALEVARE), the 2.5HIP formed by C-peptide and ChgA (InsC-ChgA) (LQTLAL-WSRMD) and the InsC-IAPP 6.9HIP (LQTLAL-NAARD). In SLOs of 8-week-old female NOD mice, all four tetramer-binding $CD4^+$ T cell populations exhibited significantly higher proportions of $CD44^{hi}CD62L^-$ cells compared to tetramer-negative cells (Fig. 5h). Notably, $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells were even more activated than CP1d-specific T cells (Fig. 5h), despite CP1d being a high-affinity MHC-II ligand (~45-fold stronger binding than the native C-peptide[26]). Furthermore, peripheral $InsB_{12-20}$(C19S)-specific T cell activation was similar to 2.5HIP- and 6.9HIP-reactive T cells (Fig. 5h), which are established mediators of T1D pathogenesis[12,57,58].

We next examined whether $InsB_{12-20}$(C19S)-specific T cells infiltrate pancreatic islets at early disease stages. These cells were readily detectable in the islets of 6- to 8-week-old female NOD mice and formed a population distinct from 2.5HIP- and 6.9HIP-reactive T cells (Fig. 5i). Their frequencies were consistently higher than 6.9HIP-reactive T cells (Fig. 5i), which accumulate in islets later in disease[58], and were comparable to 2.5HIP-reactive T cells that mediate early islet autoimmunity[12,57,58] (Fig. 5i). Nearly all intra-islet $InsB_{12-20}$(C19S)- and 2.5HIP-specific $CD4^+$ T cells were $CD44^{hi}CD62L^-$ and $CD44^{hi}CD11a^{hi}$ (Fig. 5j), indicating a terminal effector phenotype.

C19S transformation in stressed and inflamed human islets

We applied the GAP assay to human islets to assess C19S generation under ER and oxidative stress. Islets obtained from four non-diabetic donors (Supplementary Table 4) were hand-picked and treated with tunicamycin with or without glutathione for 2 h. The crinosome and DCG fractions were isolated and offered to C3.g7 for presentation to the 9B9 and S5 T cells (Fig. 6a). In all four donors, tunicamycin treatment consistently increased C19S in both crinosome and DCG fractions, whereas glutathione significantly inhibited this increase without altering native insulin peptide levels (Fig. 6b). Thus, as observed with MIN6 cells and mouse islets, ER and oxidative stress enhance C19S transformation in human islets.

To test whether inflammatory cytokines drive C19S in human islets, we performed MS analysis of crinosome peptidomes isolated from human islets exposed to IL-1β, TNFα, IFNγ or all three combined. Peptides from β cell proteins, including native insulin, ChgA and IAPP, were identified in all conditions (Fig. 6c). With this unbiased approach, C19S insulin peptides were below the detection limit in vehicle-treated islets, possibly due to limited sample material. However, under cytokine stimulation conditions, C19S was detected in crinosome peptidomes from two different donors (Fig. 6c), including the exact $InsB_{9-23}$(C19S) peptide (Fig. 6d). Thus, T1D-relevant cytokines amplify C19S generation in human islets.

Expansion of HLA-DQ8-restricted $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells in patients with T1D

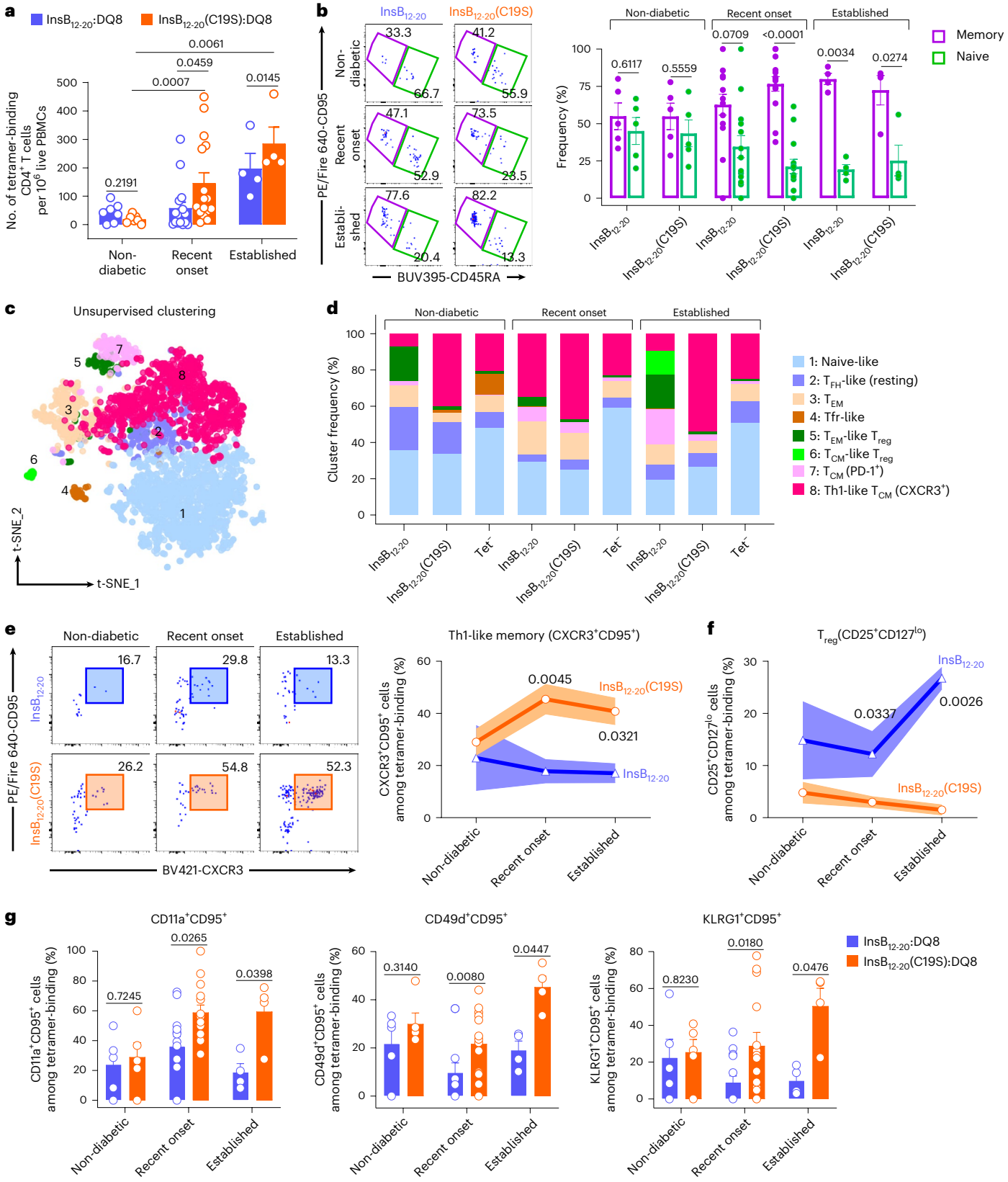
Because $InsB_{12-20}$(C19S) is conserved between mice and humans and I-A^{g7} and HLA-DQ8 share antigen presentation features[46,59,60], we tested whether HLA-DQ8-expressing individuals harbor $CD4^+$ T cells recognizing the $InsB_{12-20}$(C19S) epitope during diabetes development. Building on prior HLA-DQ8 tetramers for native $InsB_{12-20}$ and $InsB_{13-21}$(ref. 28), we generated an $InsB_{12-20}$(C19S):DQ8 tetramer (Extended Data Fig. 6a) and analyzed PBMC samples from three HLA-$DQ8^+$ cohorts: non-diabetic controls ($n = 7$), recent-onset T1D (diagnosed within 12 months; $n = 16$) and established T1D (diagnosed for > 5 years; $n = 4$) (Extended Data Fig. 6b). All participants carried at least one copy of the HLA-DQ8 haplotype (Supplementary Table 5).

Both the $InsB_{12-20}$:DQ8 and $InsB_{12-20}$(C19S):DQ8 tetramers exhibited minimal staining in $CD4^+$ T cells from non-DQ8 healthy (DQ6/DQ7) or diabetic (DQ9/DQ5) individuals (Extended Data Fig. 6c), confirming HLA-DQ8 restriction. As described previously[28], the $InsB_{12-20}$:DQ8 tetramer detected rare native insulin-reactive $CD4^+$ T cells in humans

Fig. 7 | HLA-DQ8-restricted C19S-specific $CD4^+$ T cells expand in patients with T1D and acquire persistent memory activation during disease progression. **a**, Bar graph (mean ± s.e.m.) showing the numbers of indicated DQ8 tetramer-binding T cell populations from all 27 individuals across non-diabetic ($n = 7$), recent-onset ($n = 16$) and established T1D ($n = 4$) cohorts. **b**, Representative flow cytometry plots (left) and bar graph (right; mean ± s.e.m.) showing the frequencies of naïve ($CD45RA^+CD95^-$) and memory ($CD45RA^-CD95^+$) cells within the indicated tetramer-binding $CD4^+$ T cell populations in the non-diabetic ($n = 5$), recent-onset ($n = 14$) and established T1D ($n = 4$) cohorts. **c**, t-SNE plot showing unsupervised clustering of merged tetramer-binding and Tet^- cells from all 27 participants in the three cohorts. **d**, Stacked bar graph showing the distribution of each cluster in each of the three T cell populations across non-diabetic, recent-onset and established cohorts. **e**, Representative flow cytometry plots (left) and longitudinal quantification (right; mean ± s.e.m.) showing $CXCR3^+CD95^+$ activated memory cells within $CD127^{hi}CD25^-$ conventional T cells among $InsB_{12-20}$:DQ8 or $InsB_{12-20}$(C19S):DQ8 tetramer-binding $CD4^+$ T cells. **f**, Longitudinal quantification (mean ± s.e.m.) of $CD25^+CD127^{lo}$ regulatory T cells (T_{reg} cells) within $InsB_{12-20}$:DQ8 or $InsB_{12-20}$(C19S):DQ8 tetramer-binding $CD4^+$ T cells in the three cohorts. **g**, Bar graph (mean ± s.e.m.) showing frequencies of individual activation markers (CD11a, CD49d and KLRG1) among $CD95^+$ tetramer-binding $CD4^+$ T cells in the non-diabetic ($n = 5$), recent-onset ($n = 14$) and established T1D ($n = 4$) cohorts. In **a**, **b** and **g**, n is the number of human subjects; each dot is one participant. In **b**, **e**, **f** and **g**, samples without detectable $InsB_{12-20}$:DQ8 or $InsB_{12-20}$(C19S):DQ8 tetramer-binding $CD4^+$ T cells were excluded from the analysis. For statistical analysis, two-tailed Mann-Whitney tests were performed for between-cohort comparisons in **a**. Two-tailed paired t-tests were performed for within-subject comparisons in **a**, **b** and **e**–**g**. The data (**a, b, e**–**g**) represent 14 independent experiments.

https://doi.org/10.1038/s41590-025-02343-z

(Extended Data Fig. 6c). In contrast, we identified a distinct population labeled by the $InsB_{12\text{-}20}$(C19S):DQ8 tetramer, which consistently appeared in all DQ8$^+$ samples from the non-diabetic, recent-onset and established cohorts, with little overlap with the $InsB_{12\text{-}20}$-specific CD4$^+$ T cells (Extended Data Fig. 6c). To further confirm specificity, we performed dual-color staining using the $InsB_{12\text{-}20}$(C19S):DQ8 tetramer conjugated with different fluorochromes, which again identified a discrete T cell population (Extended Data Fig. 6d). Collectively, these results demonstrate the presence of CD4$^+$ T cells specific for the $InsB_{12\text{-}20}$(C19S) epitope in DQ8$^+$ individuals.



Next, we quantified $InsB_{12-20}$:DQ8 and $InsB_{12-20}$(C19S):DQ8 tetramer-binding $CD4^+$ T cells in the three cohorts. $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells were rare in nearly all non-diabetic individuals but showed significant expansion in patients with recent-onset or established T1D (Fig. 7a and Extended Data Fig. 6e). Furthermore, in both cohorts of patients with T1D, $InsB_{12-20}$(C19S)-specific T cells were consistently more abundant than $InsB_{12-20}$-specific T cells (Fig. 7a and Extended Data Fig. 6e). Thus, $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells expand at diabetes onset and persist into the established stage, supporting a link to disease progression.

Persistent memory activation of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells in human T1D

We profiled $InsB_{12-20}$(C19S):DQ8 and $InsB_{12-20}$:DQ8 tetramer-binding $CD4^+$ T cells with a 27-marker spectral flow panel capturing activation, memory and regulatory features. Consistent with previous findings[28], native $InsB_{12-20}$-specific $CD4^+$ T cells in recent-onset patients showed a trend toward memory differentiation ($CD95^+CD45RA^-$), but this shift was not statistically significant and only became evident in established T1D (Fig. 7b). In contrast, $InsB_{12-20}$(C19S):DQ8 tetramer-binding $CD4^+$ T cells displayed a memory-dominant phenotype as early as the recent-onset stage, which persisted at similarly high levels in established T1D (Fig. 7b).

Because both $InsB_{12-20}$- and $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells exhibited a general memory phenotype, we asked whether they differed in their underlying activation states. To this end, we combined $InsB_{12-20}$:DQ8 and $InsB_{12-20}$(C19S):DQ8 tetramer-binding cells, along with tetramer-negative controls, from all 27 individuals in the non-diabetic, recent-onset and established cohorts, and performed unsupervised clustering of the merged samples. This analysis revealed 8 clusters (Fig. 7c) capturing the well-documented cellular diversity in autoreactive T cells involved in T1D[61].

As shown in Extended Data Fig. 6f, cluster 1 expressed high CD45RA and CCR7, consistent with a naïve-like phenotype. Cluster 2 lacked memory and regulatory T cell (T_{reg} cell)markers but expressed CXCR5 and low PD-1, compatible with resting T_{FH} cells[62]. Clusters 4, 5 and 6 exhibited a general T_{reg} phenotype ($CD25^+CD127^{lo}$) but differed in their differentiation status. Cluster 4 co-expressed CXCR5 and memory markers (CD45RO and CD95), consistent with circulating memory-like T follicular regulatory (T_{fr}) cells[63]. Cluster 5 showed an effector memory (T_{EM}) profile ($CD45RO^+CD95^+CCR4^+CD45RA^-CCR7^-$) and high HLA-DR, resembling T_{EM}-like T_{reg} cells capable of inducing rapid suppression[64,65]. Cluster 6, by contrast, had reduced CD25 expression, a hallmark of terminally differentiated $CD25^{low}$ Tregs that expand in autoimmunity[66]. This population ($CD45RO^+CD95^+CCR4^+CD45RA^-CCR7^+$) also resembled central memory (T_{CM})-like T_{reg} cells and expressed high levels of PD-1 and CD226, markers associated with T_{reg} persistence and regulatory function[67,68]. Clusters 3, 7 and 8 exhibited a general memory ($CD45RO^+CD95^+CD45RA^-$) signature while lacking T_{reg} markers. Cluster 3 lacked CCR7 and preferentially expressed KLRG1 and CXCR3, consistent with an activated T_{EM} profile. Clusters 7 and 8 retained CCR7, consistent with T_{CM}, and expressed activation markers CD11a, CD49d, KLRG1, CD28, CD38 and CXCR3. However, cluster 7 had higher expression of PD-1, suggesting restrained activation. In contrast, cluster 8 had higher levels of CXCR3, consistent with a Th1-like T_{CM} activation profile, a feature aligning with recent scRNA-seq studies defining T_{CM} activation in islet-reactive $CD4^+$ T cells from patients with T1D [69].

Analysis of cluster distributions revealed several distinct features of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells. Relative to other populations, these cells were more enriched in the Th1-like T_{CM} cluster 8 (Fig. 7d). In both recent-onset and established T1D, cluster 8 accounted for more than half of the $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells (Fig. 7d). Conversely, they showed minimal representation in the two Treg-associated clusters 5 and 6 across all cohorts (Fig. 7d). To validate results from unsupervised clustering, we analyzed $CXCR3^+CD95^+$ Th1-like memory cells in individual subjects. In $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells, the frequencies of $CXCR3^+CD95^+$ cells significantly increased at diabetes onset and persisted at high levels in established T1D (Fig. 7e). In contrast, $CD25^+CD127^{lo}$ Tregs showed an opposite pattern: these cells were markedly fewer overall and further declined within the $InsB_{12-20}$(C19S)-specific population across stages (Fig. 7f). Additional activation markers associated with the Th1-like T_{CM} cluster, including CD11a, KLRG1 and CD49d, were significantly higher on $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells at both recent-onset and established disease stages (Fig. 7g). Collectively, these results highlight the distinct functional profile of $InsB_{12-20}$(C19S)-specific T cells, characterized by progressive memory activation and limited regulatory potential, that emerges early and persists through T1D development.

Discussion

Our study uncovers a microenvironment-driven, extra-genomic mechanism of neoantigen formation in T1D, in which disease-relevant signals induce and amplify a Cys→Ser transformation (C19S) in insulin, leading to neoepitope presentation and $CD4^+$ T cell autoreactivity. We detected C19S insulin peptides in both human and mouse MHC-II immunopeptidomes, demonstrating that this transformation occurs in vivo across species and integrates into the broad self-peptidome. Inflammation and oxidative stress enhance C19S transformation in β cells and APCs, underscoring the role of the islet microenvironment in shaping the autoimmune peptidome. We further show that C19S-specific $CD4^+$ T cells expand in both NOD mice and $HLA\text{-}DQ8^+$ patients with T1D, displaying an activated memory phenotype linked to disease progression. Thus, C19S diversifies β cell antigenicity at the single-residue level.

C19S, as described here, is best considered a single-amino acid transformation, as it was detected by SPIDER rather than a conventional PTM search. Mechanistically, Cys→Ser also differs from PTMs that are enzyme-mediated and reversible, whereas prior work showed that oxidation can drive cysteine thiols through dehydroalanine intermediates to serine in an irreversible process[38]. A seminal study demonstrated that oxidative stress can also drive post-genetic recoding through methionine misacylation of non-cognate tRNAs, altering protein translation without involving classical PTMs[70]. Our study extends these concepts to β cells in vivo, showing that oxidative stress can generate C19S in the insulin B-chain, creating a disease-relevant neoepitope. Several features of the islet microenvironment, including the high abundance of insulin, the redox sensitivity of B(C19) residue, and the intrinsically oxidative islet milieu, converge to create a 'perfect storm' for this transformation. Our recent immunopeptidome analysis identified native $InsB_{9-23}$ as an MHC-II-bound peptide in the thymus[29]. After exhaustive searching, we did not detect the C19S variant, supporting the idea that C19S formation requires an oxidative tissue microenvironment such as pancreatic islets. Given the essential role of the B(C19) residue in insulin biosynthesis, C19S likely represents collateral damage rather than an adaptive modification. In a broader sense, our study suggests that Cys→Ser transformations may also occur in other disease contexts where ROS are elevated, such as viral infections and oxidative tumor microenvironments. Immunopeptidome analysis will be a powerful approach to indicate whether such peptides are presented by MHC-I or MHC-II molecules and whether they can act as neoantigens.

In the context of insulin autoreactivity, the previously described defective ribosomal insulin gene product in humans provides a remarkable example of how non-canonical translation events can generate highly immunogenic neoepitopes[71]. Although the defective ribosomal insulin gene product-derived epitope appears to be unique to humans, our data show that C19S transformation is a conserved process, generating identical neoepitopes in both mice and humans. In NOD mice, we found that the $InsB_{12-20}$(C19S) epitope is recognized by a distinct $CD4^+$ T cell population. Although this differs from classical altered peptide ligand recognition[72], which usually involves the same TCR responding differently to distinct peptide ligands, it illustrates the

broader principle that a single-amino-acid change in a self-peptide can profoundly alter T cell function. An intriguing observation is the inactivation of the $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells in the NOD.*B16A* mice, suggesting that the loss of reactivity to the C19S epitope may also contribute to the profound diabetes resistance phenotype in this mouse[23]. Along these lines, previous studies have indicated that C-terminal additions[73] and peptide fusions[74] of $InsB_{9-23}$ can generate strong TCR agonists. The identification of C19S further highlights that autoreactivity to the $InsB_{9-23}$ region manifests in multiple forms, and this diversity may explain why this segment remains a dominant target in autoimmune diabetes.

In humans, $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells expand at diabetes onset and exhibit a persistent T_{CM} activation phenotype. This corroborates a recent comprehensive immunopeptidome analysis of cytokine-stressed human islets, showing recognition of multiple HLA-bound peptides by patient-derived T cells[75]. From a biomarker perspective, tracking C19S-specific T_{CM} cells could offer a dynamic measure of disease activity, progression risk and therapeutic efficacy, aligning with the concept that antigen-specific $CD4^+$ T cells can be used to monitor islet autoimmunity in real time[28]. Moreover, given the self-renewal capacity and tissue-homing potential of T_{CM} cells, C19S-specific T cells may readily target newly regenerated or transplanted β cells in patients with established T1D. Developing strategies to modulate these T_{CM} cells may have therapeutic implications in islet transplantation. Overall, our study suggests that C19S-specific $CD4^+$ T_{CM} cells are primed for long-term persistence and may act as a sustained reservoir of anti-islet autoreactivity. Their persistence and reactivation capacity make them a potential target for immunotherapies aimed at disrupting pathogenic memory responses and restoring immune tolerance.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41590-025-02343-z.

References

1. Stern, L. J., Clement, C., Galluzzi, L. & Santambrogio, L. Non-mutational neoantigens in disease. *Nat. Immunol.* **25**, 29–40 (2024).
2. Yang, M.-L. et al. Carbonyl posttranslational modification associated with early-onset type 1 diabetes autoimmunity. *Diabetes* **71**, 1979–1993 (2022).
3. Peakman, M. et al. Naturally processed and presented epitopes of the islet cell autoantigen IA-2 eluted from HLA-DR4. *J. Clin. Invest.* **104**, 1449–1457 (1999).
4. McGinty, J. W. et al. Recognition of posttranslationally modified GAD65 epitopes in subjects with type 1 diabetes. *Diabetes* **63**, 3033–3040 (2014).
5. Delong, T. et al. Diabetogenic T-cell clones recognize an altered peptide of chromogranin A. *Diabetes* **61**, 3239–3246 (2012).
6. Delong, T. et al. Islet amyloid polypeptide is a target antigen for diabetogenic $CD4^+$ T cells. *Diabetes* **60**, 2325–2330 (2011).
7. Mannering, S. I. et al. The insulin A-chain epitope recognized by human T cells is posttranslationally modified. *J. Exp. Med.* **202**, 1191–1197 (2005).
8. Babon, J. A. B. 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).
9. Scotto, M. et al. Zinc transporter (ZnT)8186–194 is an immunodominant $CD8^+$ T cell epitope in HLA-$A2^+$ type 1 diabetic patients. *Diabetologia* **55**, 2026–2031 (2012).
10. Marre, M. L. 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).
11. Azoury, M. E. et al. $CD8^+$ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes. *Diabetes* **70**, 2879–2891 (2021).
12. Delong, T. et al. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* **351**, 711–714 (2016).
13. Pataskar, A. et al. Tryptophan depletion results in tryptophan-to-phenylalanine substitutants. *Nature* **603**, 721–727 (2022).
14. Yang, C. et al. Arginine deprivation enriches lung cancer proteomes with cysteine by inducing arginine-to-cysteine substitutants. *Mol. Cell* **84**, 1904–1916 (2024).
15. Lichti, C. F. & Wan, X. Using mass spectrometry to identify neoantigens in autoimmune diseases: the type 1 diabetes example. *Semin. Immunol.* **66**, 101730 (2023).
16. Zhong, J., Rao, X., Xu, J.-F., Yang, P. & Wang, C.-Y. The role of endoplasmic reticulum stress in autoimmune-mediated beta-cell destruction in type 1 diabetes. *J. Diabetes Res.* **2012**, 238980 (2012).
17. Sahin, G. S., Lee, H. & Engin, F. An accomplice more than a mere victim: the impact of β-cell ER stress on type 1 diabetes pathogenesis. *Mol. Metab.* **54**, 101365 (2021).
18. Marre, M. L. β cell ER stress and the implications for immunogenicity in type 1 diabetes. *Front. Cell Dev. Biol.* **3**, 67 (2015).
19. Wan, X. et al. Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature* **560**, 107–111 (2018).
20. Vomund, A. N., et al. Blood leukocytes recapitulate diabetogenic peptide-MHC-II complexes displayed in the pancreatic islets. *J. Exp. Med.* **218**, e20202530 (2021).
21. Todd, J. A., Bell, J. I. & McDevitt, H. O. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* **329**, 599–604 (1987).
22. Acha-Orbea, H. & McDevitt, H. O. The first external domain of the nonobese diabetic mouse class II I-A beta chain is unique. *Proc. Natl Acad. Sci. USA* **84**, 2435–2439 (1987).
23. 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).
24. Daniel, D., Gill, R. G., Schloot, N. & Wegmann, D. Epitope specificity, cytokine production profile and diabetogenic activity of insulin-specific T cell clones isolated from NOD mice. *Eur. J. Immunol.* **25**, 1056–1062 (1995).
25. Michels, A. W. et al. Islet-derived CD4 T cells targeting proinsulin in human autoimmune diabetes. *Diabetes* **66**, 722–734 (2016).
26. Wan, X. et al. The MHC-II peptidome of pancreatic islets identifies key features of autoimmune peptides. *Nat. Immunol.* **21**, 455–463 (2020).
27. Gioia, L. et al. Position β57 of I-Ag7 controls early anti-insulin responses in NOD mice, linking an MHC susceptibility allele to type 1 diabetes onset. *Sci. Immunol.* **4**, eaaw6329 (2019).
28. Sharma, S. et al. Measuring anti-islet autoimmunity in mouse and human by profiling peripheral blood antigen-specific CD4 T cells. *Sci. Transl. Med.* **15**, eade3614 (2023).
29. Hu, H. et al. Crinophagic granules in pancreatic β cells contribute to mouse autoimmune diabetes by diversifying pathogenic epitope repertoire. *Nat. Commun.* **15**, 8318 (2024).
30. Srivastava, N. et al. Chromogranin A deficiency confers protection from autoimmune diabetes via multiple mechanisms. *Diabetes* **70**, 2860–2870 (2021).

31. Groegler, J., Callebaut, A., James, E. A. & Delong, T. The insulin secretory granule is a hotspot for autoantigen formation in type 1 diabetes. *Diabetologia* **67**, 1507–1516 (2024).
32. Amdare, N., Purcell, A. W. & DiLorenzo, T. P. Noncontiguous T cell epitopes in autoimmune diabetes: from mice to men and back again. *J. Biol. Chem.* **297**, 100827 (2021).
33. Haataja, L. et al. Disulfide mispairing during proinsulin folding in the endoplasmic reticulum. *Diabetes* **65**, 1050–1060 (2016).
34. Arunagiri, A. et al. Proinsulin misfolding is an early event in the progression to type 2 diabetes. *Elife* **8**, e44532 (2019).
35. Sun, J. et al. Proinsulin misfolding and endoplasmic reticulum stress during the development and progression of diabetes. *Mol. Asp. Med.* **42**, 105–118 (2015).
36. Arunagiri, A. et al. Misfolded proinsulin in the endoplasmic reticulum during development of beta cell failure in diabetes. *Ann. N. Y. Acad. Sci.* **1418**, 5–19 (2018).
37. Cao, S. S. & Kaufman, R. J. Endoplasmic reticulum stress and oxidative stress in cell fate decision and human disease. *Antioxid. Redox Signal.* **21**, 396–413 (2014).
38. Jeong, J. et al. Novel oxidative modifications in redox-active cysteine residues. *Mol. Cell. Proteom.* **10**, M110.000513 (2011).
39. Delmastro, M. M. & Piganelli, J. D. Oxidative stress and redox modulation potential in type 1 diabetes. *Clin. Developmental Immunol.* **2011**, 1–15 (2011).
40. Oslowski, C. M. & Urano, F. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. *Methods Enzymol.* **490**, 71–92 (2011).
41. Engin, F. et al. Restoration of the unfolded protein response in pancreatic β cells protects mice against type 1 diabetes. *Sci. Transl. Med.* **5**, 211ra156 (2013).
42. Sacco, F. et al. Glucose-regulated and drug-perturbed phosphoproteome reveals molecular mechanisms controlling insulin secretion. *Nat. Commun.* **7**, 13250 (2016).
43. Ferris, S. T. et al. A minor subset of *Batf3*-dependent antigen-presenting cells in islets of Langerhans is essential for the development of autoimmune diabetes. *Immunity* **41**, 657–669 (2014).
44. Mohan, J. F., Petzold, S. J. & Unanue, E. R. Register shifting of an insulin peptide–MHC complex allows diabetogenic T cells to escape thymic deletion. *J. Exp. Med.* **208**, 2375–2383 (2011).
45. Stratmann, T. et al. The I-Ag7 MHC class II molecule linked to murine diabetes is a promiscuous peptide Binder. *J. Immunol.* **165**, 3214–3225 (2000).
46. Latek, R. R. 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).
47. Rappazzo, C. G., Huisman, B. D. & Birnbaum, M. E. Repertoire-scale determination of class II MHC peptide binding via yeast display improves antigen prediction. *Nat. Commun.* **11**, 4414 (2020).
48. Collesano, L., Łuksza, M. & Lässig, M. Energy landscapes of peptide-MHC binding. *PLoS Comput. Biol.* **20**, e1012380 (2024).
49. Moon, J. J. et al. Tracking epitope-specific T cells. *Nat. Protoc.* **4**, 565–581 (2009).
50. Podestà, M. A. et al. Stepwise differentiation of follicular helper T cells reveals distinct developmental and functional states. *Nat. Commun.* **14**, 7712 (2023).
51. Wherry, E. J. et al. Molecular signature of $CD8^+$ T cell exhaustion during chronic viral infection. *Immunity* **27**, 670–684 (2007).
52. Zakharov, P. N., et al. Single-cell RNA sequencing of murine islets shows high cellular complexity at all stages of autoimmune diabetes. *J. Exp. Med.* **217**, e20192362 (2020).
53. Srivastava, N. et al. CXCL16-dependent scavenging of oxidized lipids by islet macrophages promotes differentiation of pathogenic $CD8^+$ T cells in diabetic autoimmunity. *Immunity* **57**, 1629–1647 (2021).
54. Støy, J. et al. In celebration of a century with insulin: update of insulin gene mutations in diabetes. *Mol. Metab.* **52**, 101280 (2021).
55. Kägi, D. et al. TNF receptor 1-dependent β cell toxicity as an effector pathway in autoimmune diabetes. *J. Immunol.* **162**, 4598–4605 (1999).
56. Mohan, J. F. et al. Unique autoreactive T cells recognize insulin peptides generated within the islets of Langerhans in autoimmune diabetes. *Nat. Immunol.* **11**, 350–354 (2010).
57. Mitchell, J. S. et al. $CD4^+$ T cells reactive to a hybrid peptide from insulin-chromogranin A adopt a distinct effector fate and are pathogenic in autoimmune diabetes. *Immunity* **57**, 2399–2415 (2024).
58. Baker, R. L. et al. CD4 T cells reactive to hybrid insulin peptides are indicators of disease activity in the NOD mouse. *Diabetes* **67**, 1836–1846 (2018).
59. Corper, A. L. et al. A structural framework for deciphering the link between I-Ag7 and autoimmune diabetes. *Science* **288**, 505–511 (2000).
60. Lee, K. H., Wucherpfennig, K. W. & Wiley, D. C. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. *Nat. Immunol.* **2**, 501–507 (2001).
61. Pugliese, A. Autoreactive T cells in type 1 diabetes. *J. Clin. Invest.* **127**, 2881–2891 (2017).
62. Morita, R. et al. Human Blood $CXCR5^+CD4^+$ T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity* **34**, 108–121 (2011).
63. Sage, P. T., Alvarez, D., Godec, J., Andrian, U. H. von & Sharpe, A. H. Circulating T follicular regulatory and helper cells have memory-like properties. *J. Clin. Invest.* **124**, 5191–5204 (2014).
64. Wendering, D. J. et al. Effector memory–type regulatory T cells display phenotypic and functional instability. *Sci. Adv.* **10**, eadn3470 (2024).
65. Ashley, C. W. & Baecher-Allan, C. Cutting Edge: Responder T cells regulate human DR^+ effector regulatory T cell activity via granzyme B1. *J. Immunol.* **183**, 4843–4847 (2009).
66. Ferreira, R. C. et al. Cells with Treg-specific *FOXP3* demethylation but low CD25 are prevalent in autoimmunity. *J. Autoimmun.* **84**, 75–86 (2017).
67. Thirawatananond, P. et al. Treg-specific CD226 deletion reduces diabetes incidence in NOD mice by improving regulatory T-cell stability. *Diabetes* **72**, 1629–1640 (2023).
68. Fuhrman, C. A. et al. Divergent phenotypes of human regulatory T cells expressing the receptors TIGIT and CD226. *J. Immunol.* **195**, 145–155 (2015).
69. Balmas, E. et al. Islet-autoreactive $CD4^+$ T cells are linked with response to alefacept in type 1 diabetes. *JCI Insight* **8**, e167881 (2023).
70. Netzer, N. et al. Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* **462**, 522–526 (2009).
71. Kracht, M. J. L. et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat. Med.* **23**, 501–507 (2017).
72. Evavold, B. D. & Allen, P. M. Separation of IL-4 production from Th cell proliferation by an altered T cell receptor ligand. *Science* **252**, 1308–1310 (1991).
73. Wang, Y. et al. How C-terminal additions to insulin B-chain fragments create superagonists for T cells in mouse and human type 1 diabetes. *Sci. Immunol.* **4**, eaav7517 (2019).

74. Wenzlau, J. M. et al. Insulin B-chain hybrid peptides are agonists for T cells reactive to insulin B:9-23 in autoimmune diabetes. *Front Immunol.* **13**, 926650 (2022).
75. Nanaware, P. P. et al. The antigen presentation landscape of cytokine-stressed human pancreatic islets. *Cell Rep.* **44**, 115927 (2025).

[1]Department of Pathology and Immunology, Division of Immunobiology, Washington University School of Medicine, St. Louis, MO, USA. [2]The Andrew M. and Jane M. Bursky Center for Human Immunology, Washington University School of Medicine, St. Louis, MO, USA. [3]Department of Internal Medicine, Division of Rheumatology, Washington University School of Medicine, St. Louis, MO, USA. [4]Department of Developmental Biology, Center of Regenerative Medicine, Washington University School of Medicine, St. Louis, MO, USA. [5]Department of Immunology and Microbiology, Scripps Research Institute, La Jolla, CA, USA. [6]Department of Pediatrics, Division of Endocrinology, Diabetes, and Metabolism, Washington University School of Medicine, St. Louis, MO, USA. [7]Department of Radiation Oncology, Weill Cornell Medicine, New York, NY, USA. [8]Department of Pathology and Immunology, Division of Laboratory and Genomic Medicine, Washington University School of Medicine, St. Louis, MO, USA. [9]Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [10]Center for Systems Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [11]Cancer Immunology and Immunotherapy Program, UPMC Hillman Cancer Center, Pittsburgh, PA, USA. [12]Department of Neurology, Hope Center for Neurological Disorders, Knight Alzheimer's Disease Research Center, Washington University School of Medicine, St. Louis, MO, USA. [13]Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA. [14]Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. [15]These authors contributed equally: Neetu Srivastava, Anthony N. Vomund. ✉e-mail: clichti@wustl.edu; wanx@wustl.edu

Methods

Mice

NOD/ShiLtJ (NOD), NOD.129S7(B6)-*Rag1*^{tm1Mom}/J (NOD.*Rag1*$^{-/-}$), NOD.Cg-Tg (Ins2*Y16A) 1EllIns1t^{m1Jja}Ins2^{tm1Jja}/GseJ (NOD.*B16A*), NOD.*Tnfrsf1a/1b*$^{-/-}$ and C57BL/6 mice were originally obtained from The Jackson Laboratory. All mice were bred, maintained and used in experiments in our specific-pathogen-free animal facility in accordance with the Division of Comparative Medicine at Washington University School of Medicine (Association for Assessment and Accreditation of Laboratory Animal Care accreditation no. A3381-01; Protocol No. 23-0429).

Antibodies

The following antibodies were used for mouse studies: anti-CD45 (30-F11), anti-CD11c (N418), anti-CD11b (M1/70), anti-CD3ε (145-2C11), anti-CD4 (RM4-5), anti-CD8α (53-6.7), anti-CD44 (IM7), anti-CD90.2 (53-2.1), anti-B220 (RA3-6B2), anti-CXCR6 (SA051D1), anti-CD62L (MEL-14), anti-CD69 (H1.2F3) and anti-CD19 (6D5) were all purchased from BioLegend. Anti-CD11a (M17/4) was purchased from eBioscience (Thermo Fisher Scientific). The anti-I-A^{g7} antibody (AG2.42.7) was generated in our laboratory.

The following antibodies were used for human studies: anti-CD19 (HIB19), anti-CD14 (HCD14), anti-CD3 (SK7), anti-CD8 (SK1), anti-CCR7 (G043H7), anti-CXCR3 (G025H7), anti-CD25 (BC96), anti-HLA-DR (L243), anti-CD28 (CD28.2), anti-CD95 (DX2), anti-PD-1 (EH12.2H7), anti-CD45RO (UCHL1), anti-CD38 (S17015A) and anti-CXCR5 (J252D4) were all purchased from BioLegend. Anti-CD4 (SK3), anti-CD45RA (HI100), anti-KLRG1 (13F12F2) and anti-CD226 (11A8.7.4) were purchased from eBioscience (Thermo Fisher Scientific). Anti-CD127 (HIL-7R-M21), anti-CD27 (M-T271), anti-CCR6 (11A9), anti-CCR4 (1G1), anti-CD11a (G25.2) and anti-CD49d (9F10) antibodies were purchased from BD Biosciences.

Human subject recruitment

All human participants were recruited through St. Louis Children's Hospital at Washington University School of Medicine under protocols approved by the Institutional Review Board at Washington University in St. Louis. Written informed consent was obtained from all participants or their legal guardians. For the HLA-II immunopeptidome and MMTT study, 30 participants were enrolled over a 4-year period, with 10 participants per cohort: (1) non-diabetic controls (n = 10), (2) patients with T1D at 3-month onset (n = 10) and (3) patients with T1D at 18-month onset (n = 10). Some non-diabetic participants were siblings of individuals with T1D. For the human T cell biology studies, a separate recruitment was conducted over a three-year period and included three cohorts: (1) non-diabetic controls, (2) patients with recent-onset T1D (diagnosed within 12 months) and (3) individuals with established T1D (diagnosed for >5 years). More than 65 individuals were screened for HLA-DQ haplotypes, and PBMC samples from participants expressing at least one copy of the HLA-DQ8 allele were used for the study. Because only two DQ8$^+$ individuals were identified among recruited non-diabetic controls, an additional five DQ8$^+$ samples were obtained from Precision for Medicine. In total, 27 samples were included in the T cell biology experiments: non-diabetic (n = 7), recent-onset (n = 16) and established (n = 4).

Human HLA typing

High-resolution HLA class I and class II typing for participants in the immunopeptidome study was performed by HistoGenetics using ~1 ml whole blood. For the T cell biology study, HLA-DQ typing was conducted at the HLA Laboratory at Washington University School of Medicine using ~0.5 ml of whole blood.

Human study sample size

The immunopeptidomics component of this study was exploratory in nature, and sample size was guided by practical and technical considerations. Due to logistical constraints associated with the MMTT and repeated blood draws, approximately 3 to 5 ml blood was collected per participant per time point. By pooling samples from 10 individuals per cohort, we estimated a yield of ~3–12 × 10^6 HLA-II$^+$ cells per group. Based on our experience, this yield fell within a workable range enabling detection of representative β cell-derived peptides in the HLA-II peptidome. For the human T cell biology study, sample size was first guided by our prior publication[28] and the mouse studies in the present study. We performed a power analysis using a two-sided t-test (α = 0.05, power = 0.80). Assuming a threefold difference in activation marker frequency and a standard deviation of 15% to 20%, the resulting effect size (Cohen's d = 1.5–2.0) indicated that 5 to 8 subjects per group would be sufficient to detect this difference with 80% power. Our final cohort included 7 non-diabetic controls, 16 patients with recent-onset T1D and 4 individuals with established T1D, and it yielded statistically significant results.

Human subject MMTT and sample collection

After fasting for at least 8 h, each participant had a baseline blood drawn (time 0) and was given an MMTT by drinking Boost High Protein Nutritional Energy Drink (Mead-Johnson) at 6 ml kg^{-1} (maximum 360 ml). Additional blood draws were performed at 90 and 120 min post-MMTT. C-peptide levels were measured 10 min before and immediately before MMTT, as well as 15, 30, 60, 90 and 120 min post-MMTT. All subjects had their glucose measured upon arrival at the study site; ND controls with glucose levels >126 mg dl^{-1} and T1D patients with glucose levels >250 mg dl^{-1} did not continue the study. Subjects with T1D also avoided doses of rapid-acting insulin 8 h before testing unless needed to correct hyperglycemia. For each participant, three blood samples (time 0, 90 and 120 min) were collected in BD P800 Blood Collection System tubes (BD Biosciences, Cat#366421) containing a cocktail of protease inhibitors. An aliquot of whole blood (1 ml) was used for HLA typing by HistoGenetics. The rest was used to isolate PBMCs (buffy coats), which were then frozen as cell pellets at −80 °C.

Human PBMC immunopeptidome analysis

Frozen PBMC pellets were thawed on the isolation day, pooled by cohort (ND, 3mos, 18mos), and lysed in PBS containing 40 mM MEGA 8 (MilliporeSigma, Cat#O3129), 40 mM MEGA 9 (MilliporeSigma, Cat#N1138), 1 mM PMSF (MilliporeSigma, Cat#P7626), 0.2 mM iodoacetamide (MilliporeSigma, Cat#I6125), 20 µg ml^{-1} leupeptin (MilliporeSigma, Cat#L2884), and Roche Complete Mini Protease cocktail (Roche Diagnostics, Cat#11836153001). Eighteen samples (35–51 million PBMCs each) were rocked 2 h at 4 °C, clarified at 20,000 g for 30 min at 4 °C, precleared with mouse IgG–Sepharose (Leinco, Cat#N229; Sepharose 4B, MilliporeSigma, Cat#C9142) for 30 min at 4 °C, then incubated with anti-pan-HLA-DQ–Sepharose (Leinco, Cat#H262, 3.0 mg per sample) overnight at 4 °C. Flow-through was incubated with anti-pan-HLA-DR–Sepharose (Leinco, Cat#H261, 3.0 mg per sample) overnight at 4 °C. Resins were washed sequentially (10 ml each) as in the original protocol. Peptides were eluted with 10% acetic acid (Thermo Fisher Scientific, Cat#A38SI-212), dried (SpeedVac), passed through detergent removal spin columns (Pierce, Cat#87777), cleaned on C18 Spin Columns (Pierce, Cat#89870) and dried.

Generation of CD4$^+$ T cell hybridomas

NOD mice (5–8 weeks old) were immunized with the InsB$_{9\text{-}23}$(C19S) peptide (10 nmol) emulsified in Complete Freund's Adjuvant (CFA; Difco) subcutaneously in the footpads of the hind legs. One week following immunization, the popliteal lymph nodes were collected and dispersed into single-cell suspensions. The cells were then boosted with 1 µM of the InsB$_{9\text{-}23}$(C19S) peptide for 3 days and fused with the BW5147 fusion partner following standard protocols. The growth-positive T cell clones were screened against the InsB$_{9\text{-}23}$(C19S) peptide using an antigen presentation assay. The antigen-responsive clones were further expanded and subcloned to generate monoclonal T cell hybridomas.

Competition binding assay

All synthetic peptides were purchased from Peptide 2.0. For the cell-based assay, C3.g7 cells (10^4) were mixed with competitor peptide and incubated 30 min at 37 °C; HEL_{11-25} (1 µM) was then added for 60 min at 37 °C. Cells were centrifuged, washed, and co-cultured overnight at 37 °C with HEL_{11-25}–specific $CD4^+$ hybridoma 10E11. Supernatants were collected and IL-2 measured by ELISA. IC_{50} values were fit in GraphPad Prism using a four-parameter inhibitor versus response curve. For the biochemical assay, peptides in PBS (5–1,000 pmol/sample) were mixed with 0.1 U thrombin, 1 µg soluble g7-clip and 0.125 pmol ^{125}I-labeled g7-MIME and incubated overnight at room temperature. BioGel P6 columns (Bio-Rad, Cat#732-6299) separated unbound labeled peptide; displaced label was quantified on a gamma counter (PerkinElmer, Wallac 1272 CliniGamma). Binding curves plotted displaced label vs. unlabeled competitor concentration; the concentration reducing g7-MIME binding by 50% was reported.

Isolation of pancreatic islets

The peritoneal cavity was opened to expose the common bile duct to isolate pancreatic islets. The bile duct leading to the duodenum was clamped using a dissection microscope. A solution of type XI collagenase (0.4 mg ml^{-1}; Sigma-Aldrich) in isolation buffer (composed of 1x HBSS, 10 mM HEPES, and 1 mM MgCl2; pH 7.4) was injected through the common bile duct to perfuse the pancreas. The inflated pancreas was carefully removed and digested at 37 °C for 12-14 min. Crude islets were collected and repeatedly washed with wash buffer (1x HBSS, 10 mM HEPES, 1 mM MgCl2, 1 mM CaCl2; pH 7.4). Under a microscope, pure islets were hand-picked, excluding any acinar tissue. The purified islets were then dispersed into a single-cell suspension using a non-enzymatic cell dissociation solution (Sigma-Aldrich) for 3 min at 37 °C.

ELISPOT

For immunization assays, mice were immunized with the indicated peptides (10 nmol) emulsified with CFA at the footpad. On day 7, popliteal lymph node cells were restimulated with the indicated peptides (10 µM) for 24 h, and IFNγ production was measured by ELISPOT. For assays without immunization, total islet and pancreatic lymph node cells from 8-week-old female NOD mice were pooled and cultured with equal amounts of $InsB_{9-23}$ and $InsB_{9-23}$(C19S) (1 µM) in the presence of IL-2 (20 U ml^{-1}) for 7 days. Total live cells were collected using Histopaque-1119, then recultured for 3 days at 37 °C with IL-2 (20 U ml^{-1}), peptide (1 µM) as above, and 2×10^6 irradiated NOD splenocytes (3000 RAD) per milliliter. In all experiments, recall responses were assayed on IFNγ-coated 96-well multi-screen plates (Merck Millipore, Cat#S2EM004M99) for ELISPOT, and reactive cells were counted using Immunospot Software (C.T.L.).

T cell expansion and adoptive transfer

NOD mice were immunized with InsB9-23(C19S) or scrambled peptide emulsified with CFA. Popliteal lymph node cells were then cultured with 1 µM of the respective peptide in 20 U ml^{-1} IL-2 and 2×10^6 ml^{-1} irradiated splenocytes to generate primary T cell lines. After 1 week, total live cells were collected using Histopaque-1119 and expanded for four weekly cycles with the nested $InsB_{12-20}$(C19S) or scrambled control peptide (0.1 µM each week) in 20 U ml^{-1} IL-2 and 2×10^6 ml^{-1} irradiated splenocytes. Because $InsB_{12-20}$(C19S) binds weakly, only 5–8% of $CD4^+$ T cells were tetramer-positive, and these tetramer-binding cells were prone to cell death after FACS sorting. For transfer, 1×10^6 $CD4^+$ T cells (containing ~5×10^4 $InsB_{12-20}$(C19S)-specific cells) were injected intravenously per NOD.*Rag1*$^{-/-}$ recipient. For $CD8^+$ T cell co-transfer, polyclonal $CD8^+$ T cells were purified with anti-mouse CD8 isolation magnetic beads (Miltenyi Biotec) and 1×10^6 $CD8^+$ T cells were injected intravenously into NOD.*Rag1*$^{-/-}$ recipients. Urine glucose was measured weekly (AimStrip US-G; Germaine Laboratories); mice with ≥250 mg dl^{-1} for two consecutive readings were scored diabetic.

I-A^{g7} tetramer staining using magnetic enrichment

The generation of I-A^{g7}-based monomers has been previously described[27]. Biotinylated monomers were tetramerized by incubating them with APC- or PE-labeled streptavidin (Agilent Technologies, Cat#PJ27S and PJRS301-1) at a 5:1 molar ratio of biotinylated molecules to labeled streptavidin for 1 h at room temperature. Single-cell suspensions from the spleen and lymph nodes (inguinal, axillary and pancreatic lymph nodes) were washed with PBS and incubated with the APC- or PE-labeled tetramers at a final concentration of 10 µg ml^{-1} for 1 h at room temperature. Following incubation, the cells were washed twice with MACS buffer (2 mM EDTA, 0.5% BSA in PBS) and then incubated with 20 µl anti-APC and anti-PE microbeads (Miltenyi Biotec, Cat#130-090-855 and 130-048-801) in 100 µl MACS buffer at 4 °C for 20 min. After washing, the cells were resuspended in 1 ml MACS buffer and applied to LS columns (Miltenyi Biotec, Cat#130-042-401) for enriching tetramer-bound T cells. The flow-through (negative fraction) and column-bound cells (positive fraction) were collected for subsequent flow cytometry analysis.

Flow cytometry analysis

Single-cell suspensions from spleen and lymph nodes (inguinal, axillary, and pancreatic lymph nodes) were prepared by mechanically disrupting the tissue using a syringe plunger through a 70-µm filter. Red blood cells were lysed, and the cell suspension was washed with PBS, and then stained with tetramers. Following tetramer staining, cells were washed with FACS buffer (1% BSA in 1xPBS) and incubated with FcR blocking antibody (2.4G2) for 10 min at 4 °C. Surface staining was performed with fluorescently labeled antibodies (1:200 v/v) by incubating at 4 °C for 30 min. Cells were washed twice with FACS buffer. Samples were examined using BD FACSCanto II, BD FACSymphony (BD Biosciences), or Cytek Aurora (Cytek Biosciences; 5-laser configuration). Data were analyzed using FlowJo 10.10.0 software (TreeStar).

Cell sorting and library preparation for scRNA-seq analysis

For scRNA-seq, spleen and lymph node cells from 8-week-old female NOD mice ($n = 6–10$) were pooled. After red blood cell lysis, cells were washed with MACS buffer and $CD19^+$ B cells and $CD11b^+$ myeloid cells were depleted using CD19 beads (Miltenyi Biotec, Cat#130-050-301) and CD11b beads (Miltenyi Biotec, Cat#130-097-142). Cells were stained with fluorochrome-conjugated tetramers, magnetically enriched using anti-PE and anti-APC beads, and surface-stained as described above. Tetramer-positive and polyclonal cells were sorted on a BD FACS Aria II (BD Biosciences). cDNA was generated after GEM generation/barcoding, followed by the GEM-RT reaction and bead cleanup. Purified cDNA was amplified for 11–16 cycles and cleaned with SPRI Select beads, then run on a Bioanalyzer to determine concentration. V(D)J target enrichment (TCR) was performed on full-length cDNA. Gene Expression, Enriched TCR and Feature libraries were prepared per the 10× Genomics Chromium Single Cell 5′ Reagent Kits User Guide (v2 Chemistry Dual Index) with Feature Barcoding for Cell Surface Protein and Immune Receptor Mapping, with PCR cycles adjusted based on cDNA concentration. Kits used: Chromium Next GEM Single Cell 5′ Kit v2, 16 rxns (Cat#PN-1000263), Chromium Next GEM Chip K Single Cell Kit, 48 rxns (Cat#PN-1000286), Chromium Single Cell Human TCR Amplification Kit (Cat#PN-1000252), Dual Index Kit TT Set A, 96 rxns (Cat#PN-1000215), 5′ Feature Barcode Kit, 16 rxns (Cat#PN-1000256) and Dual Index Kit TN Set A, 96 rxns (Cat#PN-1000250). Library concentrations were quantified by qPCR using the KAPA Library Quantification Kit (KAPA Biosystems/Roche) to achieve the desired cluster counts on the Illumina NovaSeq 6000.

scRNA-seq data analysis

Libraries were processed using Cell Ranger (v7.1.0). Low-quality barcodes and UMIs were filtered and mapped to the mouse genome (mm10). Both datasets were further aggregated using the Cell Ranger

aggr pipeline (v7.1.0). The Cell Ranger aggr pipeline automatically equalizes the average read depth between samples. The gene expression from both datasets was filtered, normalized and clustered, and the resulting Cloupe file was created and imported into the Loupe Browser for further analyses and visualization. To further define the unique molecular features of the cell population in the study, the feature barcode approach was used to quantify each feature in each cell. The Cell Ranger (v7.1.0) pipeline outputs the feature barcode counts for each cell barcode. Specific antibody detection and filtering were performed as follows: log2 count of the target ≥10 and other $\log_2$ counts ≤6. PCA was performed to reduce the dimensionality of the dataset to its most important features, and the principal components were visualized by UMAP plots. The final analysis excluded cells with any expression of *Cd19*, *Lyz2*, *Adgre1*, *Flt3*, *Ncr1*, *Cd8b1*, *Cd79a*, *Cd79b*, *Tcrg-C1*, *Xcr1*, *Sirpa*, *Cd68*, *Fcgr1*, *Ly6g*, *Ly6c1* and *H2-Ab1* transcripts. The differentially expressed genes between clusters or libraries were identified using the default algorithms. Bonferroni-adjusted *P* values were used to determine significance. GSEA was performed using Phantasus[76] (https://artyomovlab.wustl.edu/phantasus/).

β cell granule isolation

Human islets from de-identified, non-diabetic donors were purchased from Prodo Laboratories; primary mouse islets were isolated as described above. Both human and mouse islets were hand-picked and dispersed using a non-enzymatic dispersion solution (Sigma-Aldrich, Cat#C5914) for 3 min at 37 °C. Islet cell suspensions (MIN6 cells, dispersed human and mouse islet cells) were washed and resuspended in 1.0 ml PBS. Cells were lysed by passing them through a Cell Homogenizer (Isobiotec) 5 times using the 10 µm clearance ball bearing. Lysates were spun for 10 min at 1,000 *g*, 4 °C to pellet debris; the supernatant was transferred to a new tube and the 1,000 *g* spin was repeated. Pellets from both spins were discarded. Supernatants were then spun for 10 min at 5,000 *g*, 4 °C; the supernatant was transferred to a new tube and the pellet was retained. The 5,000 *g* spin was repeated, the supernatant was transferred to a clean tube, and the pellets from both 5,000 *g* spins were combined (crinosome fraction) and suspended in medium. The supernatant was centrifuged for 20 min at 15,000 *g*, and the pellet was retained. The remaining supernatant was transferred to a new tube and centrifuged for 30 min at 25,000 *g*, 4 °C; the pellet (DCG fraction) was suspended in medium. Acid phosphatase activity in granule contents was assessed using an ELISA-based acid phosphatase assay kit (MilliporeSigma, Cat#CS0740). Total protein levels in granule contents were measured using a Micro BCA protein assay kit (Pierce, Cat#23225).

GAP assay

Subcellular fractions isolated by 5,000 *g* (5k), 15,000 *g* (15k) or 25,000 *g* (25k) spins (from MIN6 cells, B6 mouse islets and human islets) were offered to the C3.g7 APC line (B cell lymphoma expressing I-A^{g7}; 5×10^4 per well) in a 96-well culture plate. After incubation for 2 h, the $CD4^+$ T cell hybridomas (5×10^4 per well) were added. After an overnight incubation, T cell responses were assessed by measuring IL-2 production in the culture supernatants using ELISA.

Human and mouse islet treatment

For reducing-agent experiments, islet cell suspensions (MIN6 cells, dispersed human and mouse islet cells) were treated with 100 µM N-acetyl cysteine (NAC), glutathione (Sigma, Cat#G4376), or TUDCA (Millipore, Cat#580549) for 16 h. For ER stressor experiments, suspensions were treated with 1 µM thapsigargin (Sigma, Cat#T9033) or 1 µg ml^{-1} tunicamycin (Sigma, Cat#T7765) for 2 h. For combined treatments, cells were first incubated with 100 µM NAC, glutathione, or TUDCA for 16 h, followed by thapsigargin or tunicamycin for 2 h. In all cases, cells were washed extensively before granule isolation and presentation. For human islet cytokine treatment, whole islets from Prodo Laboratories were cultured in CMRL 1066 medium supplemented with 10% FBS, 1% Na pyruvate, 1% glutamine, 1% NEAA and Pen/Strep. Islets were dispersed, centrifuged, washed, counted and divided into five equal aliquots (~10^6 each). Cells were treated with vehicle, 50 ng ml^{-1} human TNFα (PeproTech, Cat#300-01A), 50 ng ml^{-1} human IL-1β (PeproTech, Cat#200-01B), 50 ng ml^{-1} human IFNγ (PeproTech, Cat#300-02), or all three cytokines combined, and incubated for 16 h at 37 °C. Cells were then washed twice with PBS and resuspended for granule isolation.

Dendritic cell cytokine stimulation and antigen presentation

Splenic $CD11c^+$ cells were purified from NOD mice given recombinant Flt-3L-Ig (BioXcell, Cat#BE0342) using Miltenyi Beads (Miltenyi, Cat#130-125-835). CD11c+ DCs (~1.5×10^6) were pulsed with 25 µM InsB9-23 peptide for 1 h, then either left untreated or treated for 16 h with the following mouse cytokines (100 ng ml^{-1}): TNFα (PeproTech, Cat#315-01 A), IL-1β (PeproTech, Cat#211-11B), IFNγ (PeproTech, Cat#315-05), IL-4 (PeproTech, Cat#214-14), IL-6 (PeproTech, Cat#216-16) or IL-12 (PeproTech, Cat#210-12). DCs were then washed with PBS and used for antigen presentation assays.

Soluble crinosome and lymph peptidomes

Crinosome subcellular fractions were frozen at -80 °C and thawed at 37 °C for five cycles to release the contents of granules. Lymph samples were collected by cannulation from 6-week-old female NOD mice with or without glucose injection as described previously[77]. For both crinosome and lymph samples, a complete protease inhibitor cocktail was added, and they were concentrated by speed vac to a volume of ~100 µl. C18 Spin Columns (Pierce, Cat#89870) were used to clean up the released proteins/peptides according to the manufacturer's instructions. Samples were eluted in 0.1% formic acid (Thermo Fisher Scientific, Cat#A118-500)/95% acetonitrile (Burdick & Jackson, #015-4), and then dried with a SpeedVac.

MS

A Dionex UltiMate 3000 system (Thermo Fisher Scientific) was coupled to an Orbitrap Fusion Lumos through an EASY-Spray ion source. Peptide samples were reconstituted in 2% acetonitrile (ACN)/0.1% formic acid and loaded (15 µl min^{-1} for 3 min) onto a trap column (100 µm × 2 cm, 5 µm Acclaim PepMap 100 C18, 50 °C), then eluted (200 nL min^{-1}) onto an EASY-Spray PepMap RSLC C18 column (2 µm, 50 cm × 75 µm ID, 50 °C, Thermo Fisher Scientific). Separation used 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%; 150-151 min, 95-2%; 151-171 min, isocratic at 2%. Spray voltage was 1,600-1,800 V, ion transfer tube temperature was 275 °C and RF lens was 30%. MS scans were acquired in profile mode from 375-1500 Da at 120 K resolution (at m/z = 200) with a total cycle time of 1.5 s and dynamic exclusion of 60 s; MS/MS scans were acquired in centroid mode for ions with charge states 2-5 above a threshold of 2.0×10^4 using quadrupole isolation (0.7 Da) and HCD collision energy of 30%, with HCD MS/MS spectra recorded at 15 K resolution.

MS data analysis

Processing and database searching were performed with PEAKS Studio Xpro (version 10.6, build 20201221, Bioinformatics Solutions). Files were searched against a UniProt-Mouse database (January 2022; 22,101 entries) with no enzyme specificity and carbamidomethylation (C), oxidation (M) and deamidation (NQ) as variable modifications. Parent mass tolerance was 10 ppm (20 ppm for MIN6 experiment), and fragment ion tolerance was 0.02 Da. The Common Repository for Adventitious Proteins database (www.thegpm.org/crap/) was used for contaminant identification, and FDR estimation was enabled. Subsequent PEAKS PTM and SPIDER searches were used to identify PTMs and SAVs, respectively. SPIDER search results, which contained both PTMs and SAVs, were filtered at a 1% FDR at the peptide level and exported for

further analysis. To further verify Cys→Ser transformation specificity, the MIN6 data files were searched against the UniProt-Mouse database in which all cysteine residues were converted to serine using a custom Python script. Tolerances were set as above, with methionine oxidation and deamidation as variable modifications. For verification of Cys→Ser transformations in MIN6 cells and for crinosomes from mouse islets, database searching was performed with a UniProt-Mouse database appended with Cys→Ser variants of β cell proteins insulin-1, insulin-2, ChgA, IAPP and secretogranins using the same parameters as above. For crinosomes from human islets, searches were performed against a UniProt-Human database appended with Cys→Ser variants of β cell proteins insulin, ChgA, IAPP and secretogranins.

Quantification of PTMs and SAVs

The PSM results from the initial SPIDER search were used for this analysis. A working list was obtained by compiling a list of all PTMs and SAVs using the peptide sequence and the source data file. Using the COUNTIFS function in Excel, totals were obtained for each PTM or SAV by data file. Results for fractions were then summed to obtain totals by sample, and these numbers were exported to GraphPad Prism for statistical analysis. For peptide-level quantification, peak areas from PEAKS were used. Each area was log2 transformed; a value of 2 was used in the event of missing values for peak area.

HLA-DQ8 tetramer preparation and staining of human PBMCs.

Biotinylated monomers of HLA-DQ8 (HLA-DQA1*0301 and HLA-DQB1*0302), complexed with the insulin B-chain peptide ($InsB_{12-20}$ or $InsB_{12-20}$(C19S)), were produced in-house[28]. The biotinylated monomers were purified through a two-step process: first by Ni-NTA affinity chromatography, followed by gel filtration using a Sephacryl-300 HR column (Cytiva). Aliquots were flash-frozen and stored at −80 °C until use. The biotinylated monomers were tetramerized by incubating with streptavidin-PE or streptavidin-APC at a 5:1 molar ratio of biotinylated monomers to labeled streptavidin overnight at room temperature. PBMCs were washed twice with PBS and incubated with the APC- or PE-labeled tetramers at a final concentration of 2.5 µg per 100 µl for 1 h at room temperature. After incubation, the cells were washed twice with MACS buffer, then incubated with 20 µl anti-APC and anti-PE microbeads in 100 µl MACS buffer at 4 °C for 20 min. The cells were washed twice with FACS buffer (1% BSA in 1x PBS) and used for flow cytometry analysis.

Human PBMC spectral flow cytometry

Cryopreserved PBMCs were thawed and washed twice with PBS (cell viability >90%). Following incubation with PE- and APC-labeled tetramers, cells were blocked with anti-human FcR (human TruStain FcX, BioLegend) for 10 min at 4 °C. To maintain positive/negative resolution in a multi-color panel, sequential staining was performed as follows: PBMCs were first stained with anti-CXCR3 and anti-CCR4 for 5 min, then with anti-KLRG1 and anti-CD25 for 5 min at 37 °C, after which the surface antibody cocktail was added directly and incubated for 30 min at room temperature. Cells were then stained with Live/Dead Blue (1:1,000) for 25 min at 4 °C, washed twice, and acquired on a Cytek Aurora (Cytek Biosciences; 5-laser configuration). Spectral unmixing was performed in SpectroFlo (Cytek Biosciences). Traditional analysis of unmixed data used FlowJo v10.10.

High-dimensional analysis of spectral flow cytometry data

Data analysis was performed using DownSample V3, FlowSOM and ClusterExplorer software in FlowJo version 10.10 (refs. 78–80). $InsB_{12-20}$- and $InsB_{12-20}$(C19S)-specific tetramer-binding and tetramer-negative T cells were gated as live, singlet, $CD19^-CD14^-CD3^+CD4^+$ cells. $CD4^+$ tetramer-negative T cells were downsampled to match the number of tetramer-positive $CD4^+$ T cells. The FCS files for the three $CD4^+$ T cell populations from all 27 participants were concatenated. The combined file was uploaded into FlowJo, and t-SNE was performed. The following markers were used to generate t-SNE plots: CD45RO, CD45RA, CCR7, CD38, CD11a, CD95, CD27, CD25, CD49d, CCR4, KLRG1, CCR6, CXCR3, HLA-DR, PD-1, CD127, CD28, CD226 and CXCR5. t-SNE data were subjected to unsupervised clustering using FlowSOM to generate hierarchical consensus clusters. These clusters were further analyzed, and heat maps were generated using ClusterExplorer software and GraphPad Prism. Data visualization and analysis were performed using R (R version 4.4.3). FCS files were imported and processed using the Bioconductor package flowCore (v2.18.0). All plots were generated using the ggplot2 package (v3.5.2).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Previously published data are accessible via ProteomeXchange with identifiers PXD024400 (mouse blood leukocytes) and PXD015408 (mouse islet and pancreatic lymph node). Original MS data, peptide-spectrum matches and databases are available via the MassIVE repository with the identifier MSV000096258 (ProteomeXchange identifier PXD057371). A complete list of PTM/SAV search in the crinosome peptidome from stressed MIN6 cells is shown in Supplementary Table 2. scRNA-seq data (GEO: GSE309696) are available via the National Center for Biotechnology Information Gene Expression Omnibus database. Source data are provided with this paper.

References

76. Kleverov, M. et al. Phantasus, a web application for visual and interactive gene expression analysis. *Elife* **13**, e85722 (2024).
77. Nanaware, P. P. et al. Role of the afferent lymph as an immunological conduit to analyze tissue antigenic and inflammatory load. *Cell Rep.* **43**, 114311 (2024).
78. Van Gassen, S. et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytom. A* **87**, 636–645 (2015).
79. Quintelier, K. et al. Analyzing high-dimensional cytometry data using FlowSOM. *Nat. Protoc.* **16**, 3775–3801 (2021).
80. Ujas, T. A., Obregon-Perko, V. & Stowe, A. M. A guide on analyzing flow cytometry data using clustering methods and nonlinear dimensionality reduction (tSNE or UMAP). *Methods Mol. Biol.* **2616**, 231–249 (2023.

Acknowledgements

We thank K. S. Ravichandran and J. Kipnis for their insightful guidance throughout this study and for their critical review of the manuscript. We are grateful to J. Ponce, C. Schatz and J. Yu for their expertise in performing scRNA-seq analysis, as well as D. Brinja and E. Lantelme for their assistance with cell sorting. We acknowledge the dedicated efforts of the team at St. Louis Children's Hospital in recruiting human subjects for this study and express our sincere appreciation to all the participants and their families. We thank T. Turicek and L. Turicek for their assistance in preparing the manuscript. This work was supported by the Diabetes Research Center at Washington University (P30 DK020579 to X.W.), the Juvenile Diabetes Research Foundation (5-CDA-2022-1175-A-N to X.W.) and the National Institutes of Health (R01AI162591 and R01DK134437 to X.W.).

Author contributions

N.S. designed and conducted in vivo immunoassays and T cell biology experiments. A.N.V. carried out antigen presentation experiments and isolated immunopeptidomes. R.Y. performed T cell biology experiments and spectral flow cytometry analysis. O.J.P. prepared pancreatic islets and assisted in multiple experimental procedures.

Y.Y., D.P.T., O.A. and H.H. assisted with various T cell biology experiments and manuscript writing. T.L. and B.Z. provided advice on the design of the scRNA-seq analysis and performed data analysis. L.K. and L.T. developed and produced I-A^{97}- and DQ8-based tetramers. C.C.C. and L.S. designed and supplied samples for the lymph peptidome experiments. C.L. analyzed HLA typing data. A.V.J. provided critical advice on TCR analysis. S.S., R.M. and L.C. offered crucial advice on the human spectral flow cytometry analysis. P.S. and A.A. recruited human subjects. E.R.U. inspired us to search for non-canonical peptides in crinophagic granules and instructed us on applying the GAP assay to human islets. A.M.A. supervised human subject recruitment and conducted clinical studies. C.F.L. performed MS experiments and analyzed data. X.W. analyzed data and designed the study. A.V.J., L.C., C.-S.H., L.S., L.T. and A.M.A. critically evaluated data analyses and the manuscript. N.S., C.F.L. and X.W. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

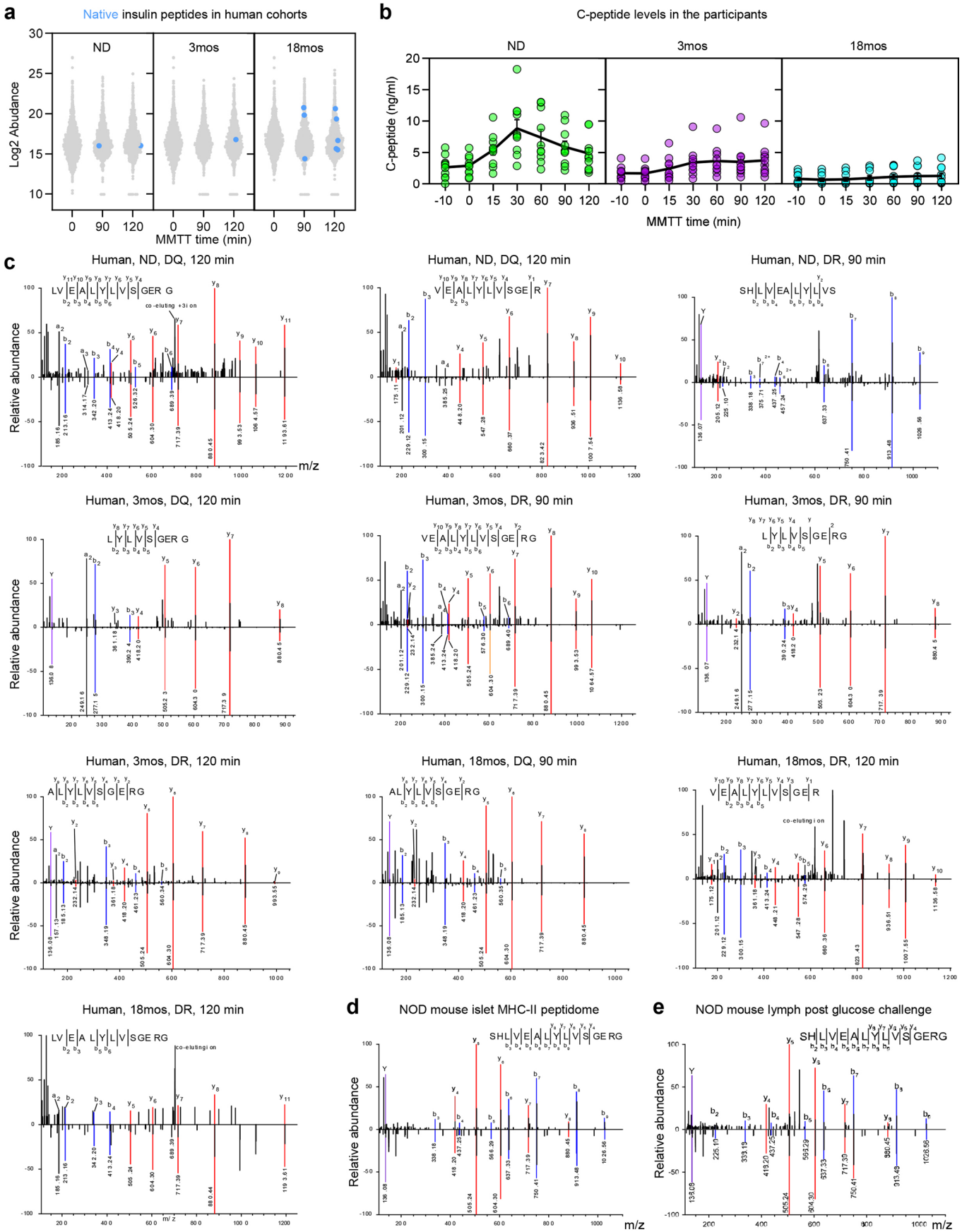
Extended data is available for this paper at https://doi.org/10.1038/s41590-025-02343-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41590-025-02343-z.

Correspondence and requests for materials should be addressed to Cheryl F. Lichti or Xiaoxiao Wan.

Peer review information *Nature Immunology* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: S, Houston in collaboration with the *Nature Immunology* team Peer reviewer reports are available.

https://doi.org/10.1038/s41590-025-02343-z

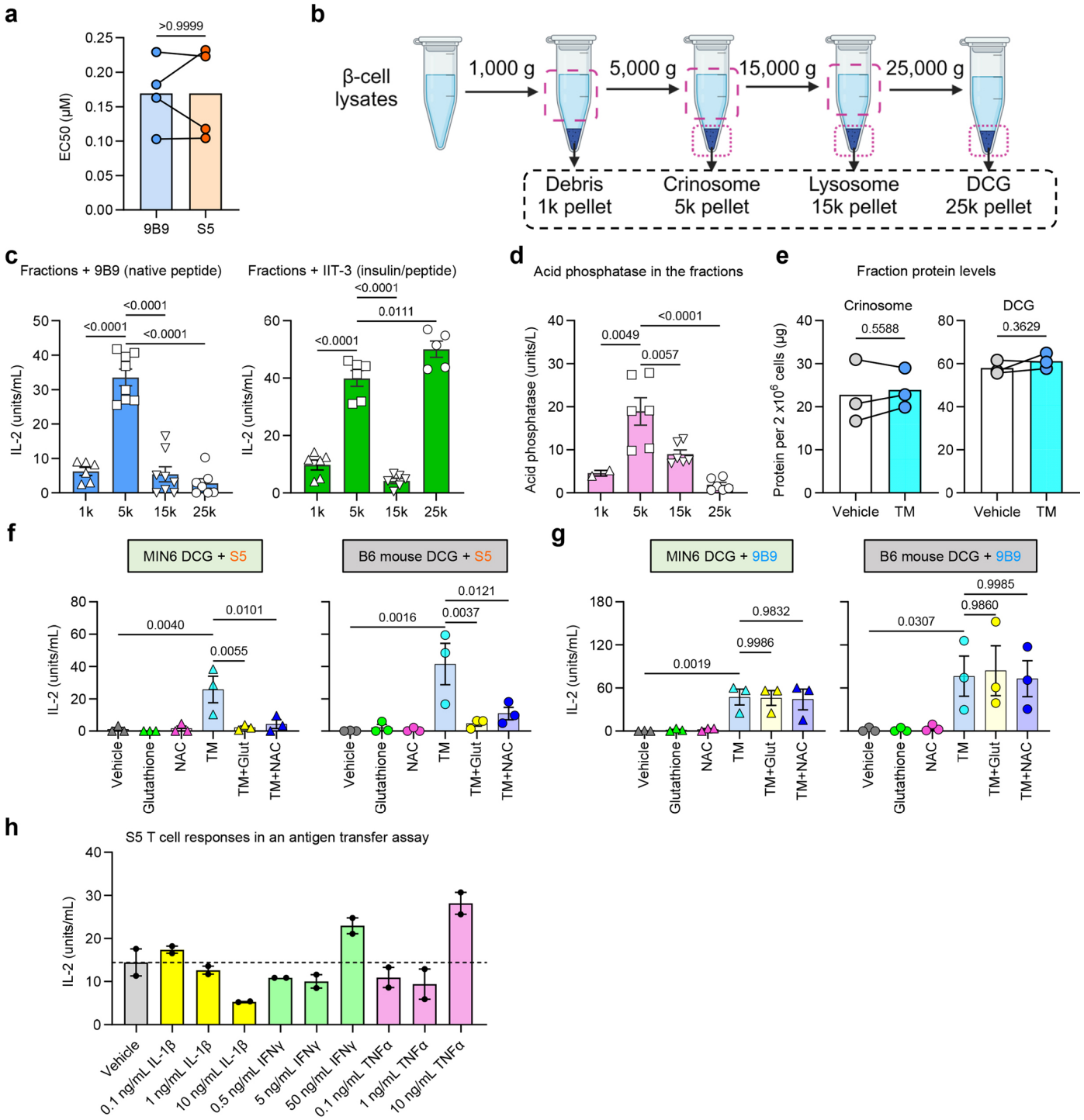


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41590-025-02343-z

Extended Data Fig. 1 | Identification and verification of C19S insulin peptides in humans and mice. a, Dot plots showing the abundance of native insulin B-chain peptides identified in the blood leukocyte MHC-II peptidome of NOD mice, showing increased abundance after glucose challenge. **b**, Dot plots (mean ± s.e.m.) showing C-peptide levels of the three human cohorts at the indicated time points before and after MMTT (n = 10 per cohort). **c**, Mirror plots showing the spectra of indicated C19S insulin B-chain peptides identified in the human HLA-DQ and HLA-DR PBMC peptidomes, all matching the synthetic standards. **d**, Mirror plot showing the spectra of the $InsB_{9\text{-}23}$(C19S) peptide identified in the MHC-II peptidome of pancreatic islets, completely matching the synthetic standard. **e**, Mirror plot showing the spectrum of the $InsB_{9\text{-}23}$(C19S) peptide identified in the mouse lymph peptidome (after glucose challenge), with a complete match to its synthetic standard. In mirror plots (**c**, **d**, **e**), the spectra of the peptides identified in biological samples are shown in the upper panel, with the corresponding synthetic standards in the lower panel.

https://doi.org/10.1038/s41590-025-02343-z

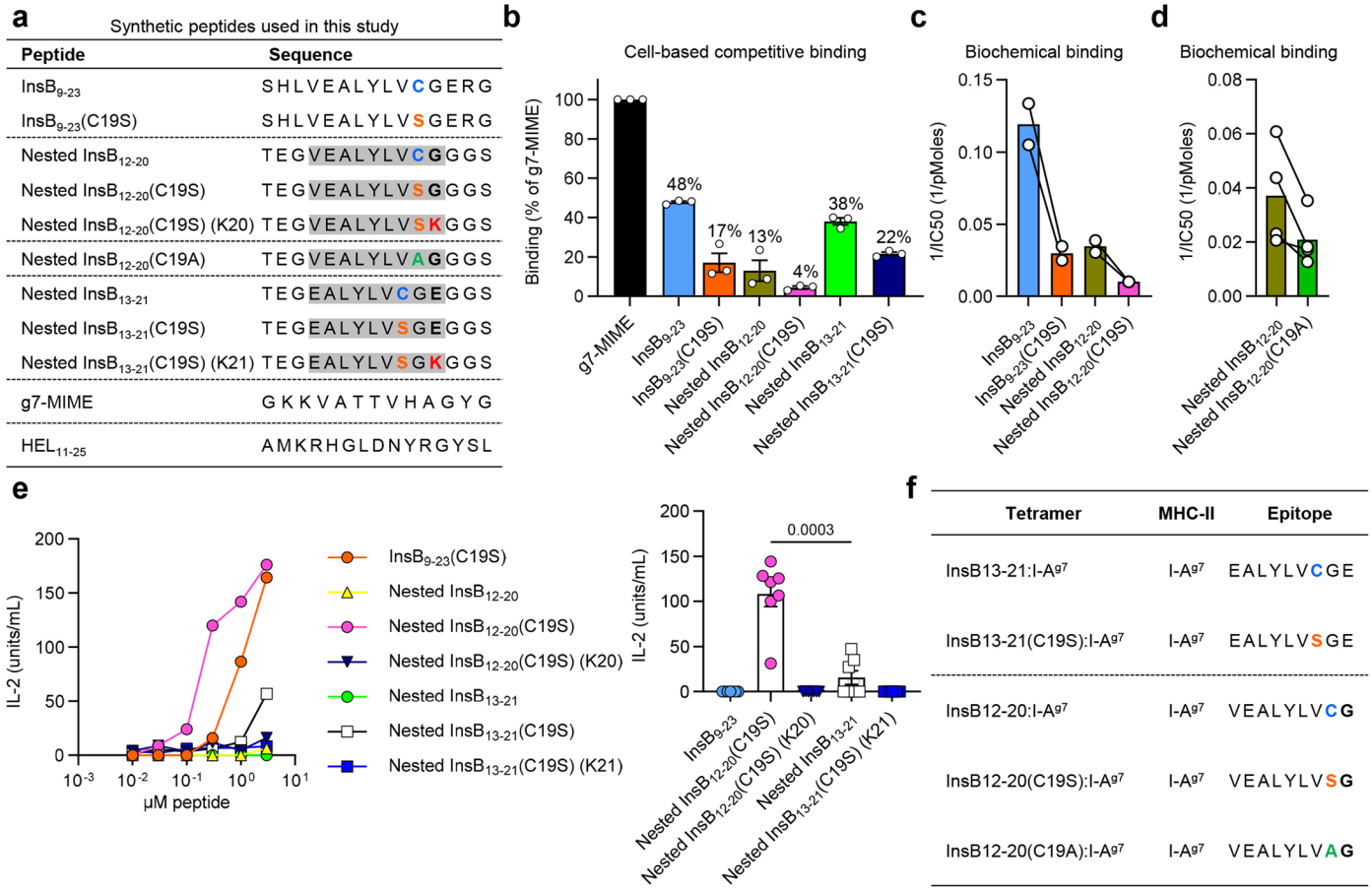


Extended Data Fig. 2 | Analysis of C19S transformation in β cells under stress and inflammation. a, Bar graph (mean ± s.e.m.) showing the peptide concentrations required to reach half-maximal responses (EC50) for 9B9 and S5 T cells to the $InsB_{9\text{-}23}$ and $InsB_{9\text{-}23}$(C19S) peptides, respectively (n = 4). **b**, Schematic of the isolation and verification of individual subcellular fractions. **c**, Bar graphs (mean ± s.e.m.) showing the responses of the 9B9 (n = 8) and IIT-3 (n = 6) T cells to C3.g7 APCs offered with each indicated granule fraction. 9B9 recognizes the $InsB_{12\text{-}20}$ epitope, presented from free $InsB_{9\text{-}23}$ but not intact insulin; IIT-3 recognizes $InsB_{13\text{-}21}$, presented from both free peptide and intact insulin. This distinguishes fractions enriched for degraded peptide (crinosomes) versus intact insulin (DCGs). **d**, Bar graph (mean ± s.e.m.) showing acid phosphatase levels in the indicated granule fractions from MIN6 cells (n = 2-6). **e**, Bar graphs (mean ± s.e.m.) showing total protein levels in the crinosome and DCG fractions from MIN6 cells (n = 3). **f**,**g**, Bar graphs (mean ± s.e.m.) showing responses of the S5 (**f**) and 9B9 (**g**) T cells to the DCG fraction isolated from MIN6 cells and B6 mouse islets following the indicated treatments (n = 3). **h**, Bar graph (mean ± s.e.m.) showing C19S generation in MIN6 cells stimulated with inflammatory cytokines, quantified by antigen-transfer assay. After 16-hour cytokine exposure, equal numbers of treated MIN6 cells were co-cultured with C3.g7 APCs under high glucose to facilitate peptide release and acquisition; S5 T cells then probed C19S formation. In **a**, **c**, **d**, **e**, **f**, **g**, and **h**, n is the number of biological replicates; each dot is one replicate. For statistical analysis, two-tailed paired t tests were performed for **a** and **e**. One-way ANOVA with Dunnett's multiple comparisons tests were performed for **c** and **d**. RM one-way ANOVA with Sidak's multiple comparisons tests were performed for **f** and **g**. The data represent two (**h**), three (**e**, **f**, **g**), four (**a**), and six (**c**, **d**) independent experiments. The illustration in b was partly created with BioRender.com.

https://doi.org/10.1038/s41590-025-02343-z

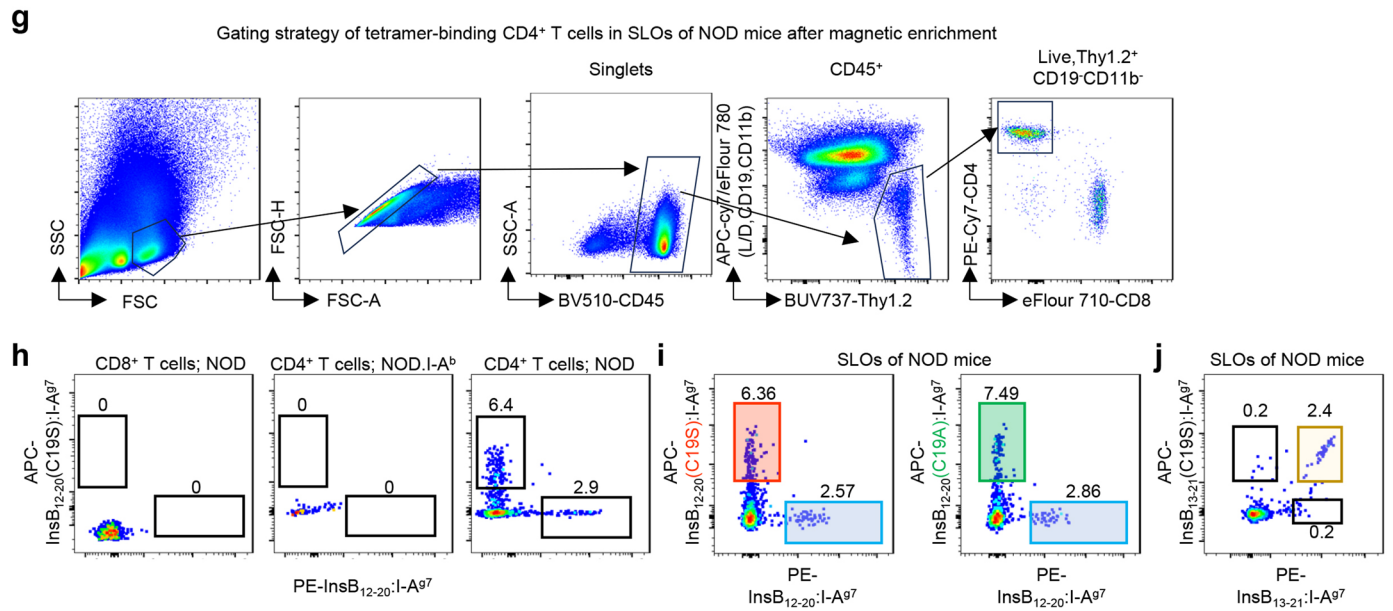
a Synthetic peptides used in this study

Peptide	Sequence
$InsB_{9-23}$	SHLVEALYLVCGERG
$InsB_{9-23}$(C19S)	SHLVEALYLVSGERG
Nested $InsB_{12-20}$	TEGVEALYLVCGGGS
Nested $InsB_{12-20}$(C19S)	TEGVEALYLVSGGGS
Nested $InsB_{12-20}$(C19S) (K20)	TEGVEALYLVSKGGS
Nested $InsB_{12-20}$(C19A)	TEGVEALYLVAGGGS
Nested $InsB_{13-21}$	TEGEALYLVCGEGGS
Nested $InsB_{13-21}$(C19S)	TEGEALYLVSGEGGS
Nested $InsB_{13-21}$(C19S) (K21)	TEGEALYLVSGKGGS
g7-MIME	GKKVATTVHAGYG
HEL_{11-25}	AMKRHGLDNYRGYSL



f

Tetramer	MHC-II	Epitope
InsB13-21:I-A^{g7}	I-A^{g7}	EALYLVCGE
InsB13-21(C19S):I-A^{g7}	I-A^{g7}	EALYLVSGE
InsB12-20:I-A^{g7}	I-A^{g7}	VEALYLVCG
InsB12-20(C19S):I-A^{g7}	I-A^{g7}	VEALYLVSG
InsB12-20(C19A):I-A^{g7}	I-A^{g7}	VEALYLVAG

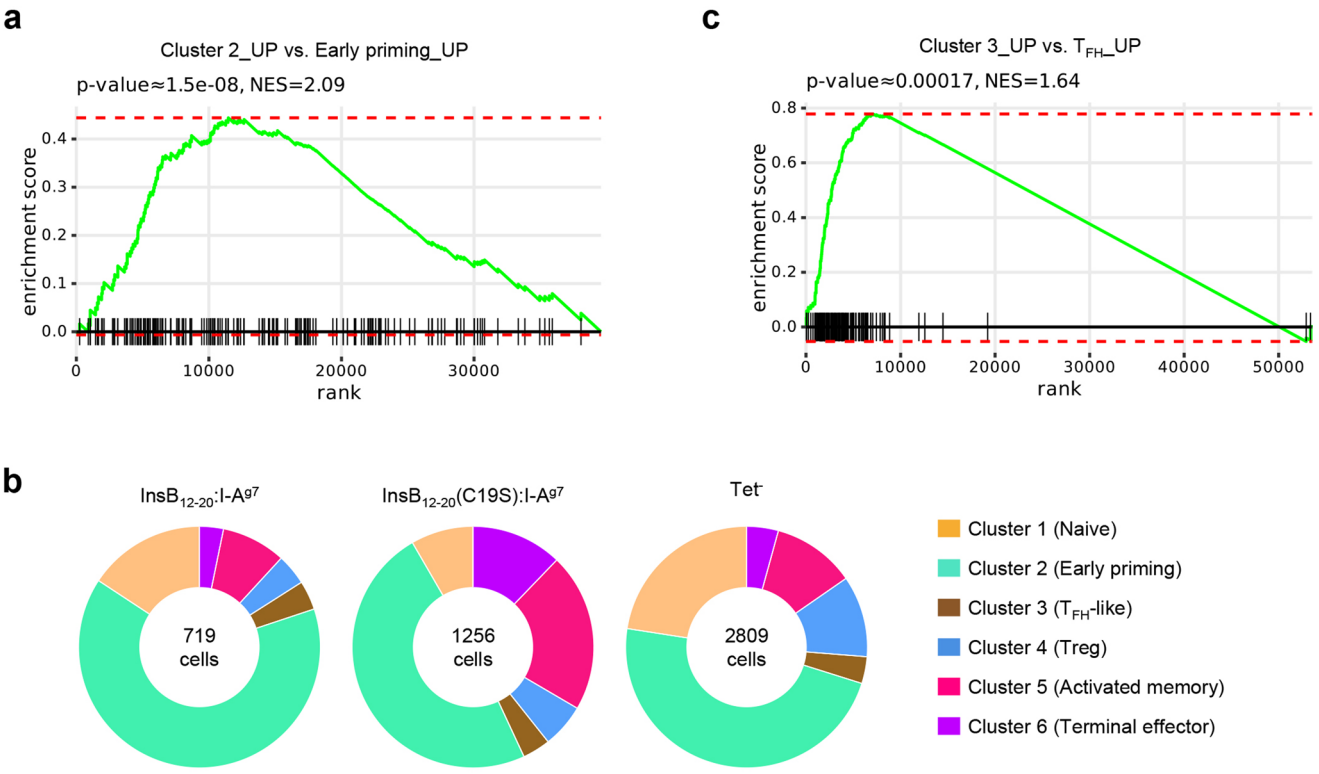


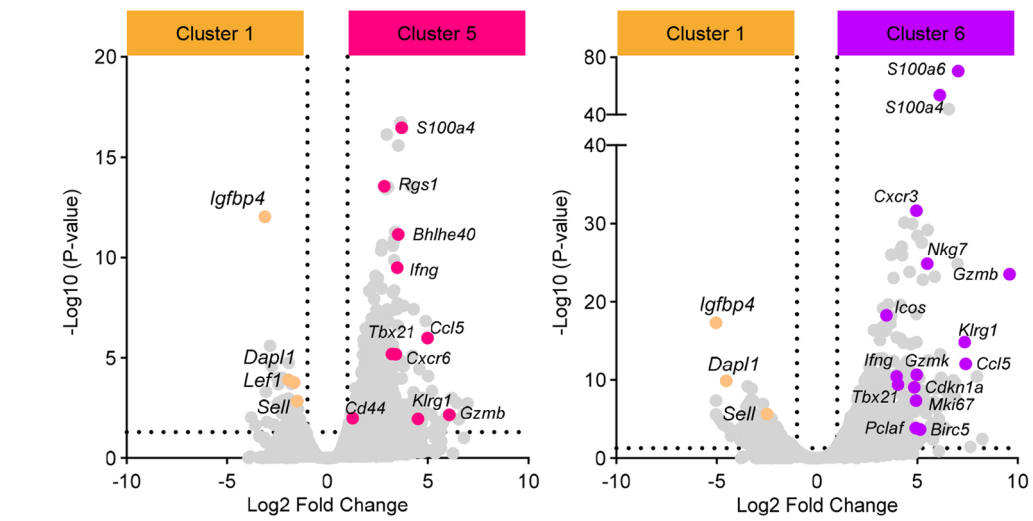
Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41590-025-02343-z

Extended Data Fig. 3 | C19S diminishes MHC-II binding but is recognized by register-specific CD4$^+$ T cells. a, Table showing sequences of synthetic peptides used in experiments (**b**, **c**, **d**, **e**). **b**, Bar graph (mean ± s.e.m.) showing relative binding affinity of indicated peptides to I-A^{g7} measured by a cell-based binding assay, which evaluates the effects of indicated competitor peptides in inhibiting the binding of the hen egg lysozyme 11-25 (HEL_{11-25}) peptide to C3.g7 cells (see Methods) (n = 3). **c**, Bar graph showing relative binding strengths of indicated peptides to soluble I-A^{g7} molecules measured by a biochemical competitive binding assay (n = 2). **d**, Bar graph showing I-A^{g7} binding of the $InsB_{12-20}$ and $InsB_{12-20}$(C19A) peptides measured by a biochemical competitive binding assay (n = 2). **e**, Responses of a representative $InsB_{9-23}$(C19S)-reactive CD4$^+$ T cell hybridoma (left) to C3.g7 cells pulsed with indicated synthetic peptides. The bar graph (right; mean ± s.e.m.) summarizes results from multiple hybridomas responding to 0.3 μM peptides (n = 6). **f**, Table summarizing I-A^{g7}-based tetramers containing native and C19S insulin peptide registers used in mouse T cell studies. **g**, Representative flow cytometry plots showing the gating strategy for tetramer-binding CD4$^+$ T cells in SLOs of NOD mice. **h**, Representative flow cytometry plots showing co-staining with $InsB_{12-20}$:I-A^{g7} and $InsB_{12-20}$(C19S):I-A^{g7} tetramers in CD4$^+$ or CD8$^+$ T cells from regular NOD mice as well as in CD4$^+$ T cells from NOD.I-A^b mice. **i**, Representative flow cytometry plots showing tetramer staining of CD4$^+$ T cells from SLOs of the same 10-week-old female NOD mice, comparing the $InsB_{12-20}$(C19S):I-A^{g7} and $InsB_{12-20}$(C19A):I-A^{g7} tetramers to the native $InsB_{12-20}$:I-A^{g7} tetramer. **j**, Representative flow cytometry plot showing co-staining with $InsB_{13-21}$:I-A^{g7} and $InsB_{13-21}$(C19S):I-A^{g7} tetramers identifying a largely overlapping CD4$^+$ T cell population. In **b**, **c**, **d**, and **e**, n is the number of independent experiments; each dot is one experiment. For statistical analysis, two-tailed paired t test was performed for **e**. The data represent two (**c**), three (**a**, **g**, **h**, **i**, **j**), and four (**d**) independent experiments.

https://doi.org/10.1038/s41590-025-02343-z





e CDR3 sequences of the three most expanded clones of $InsB_{12-20}$(C19S)-specific $CD4^+$ T cells

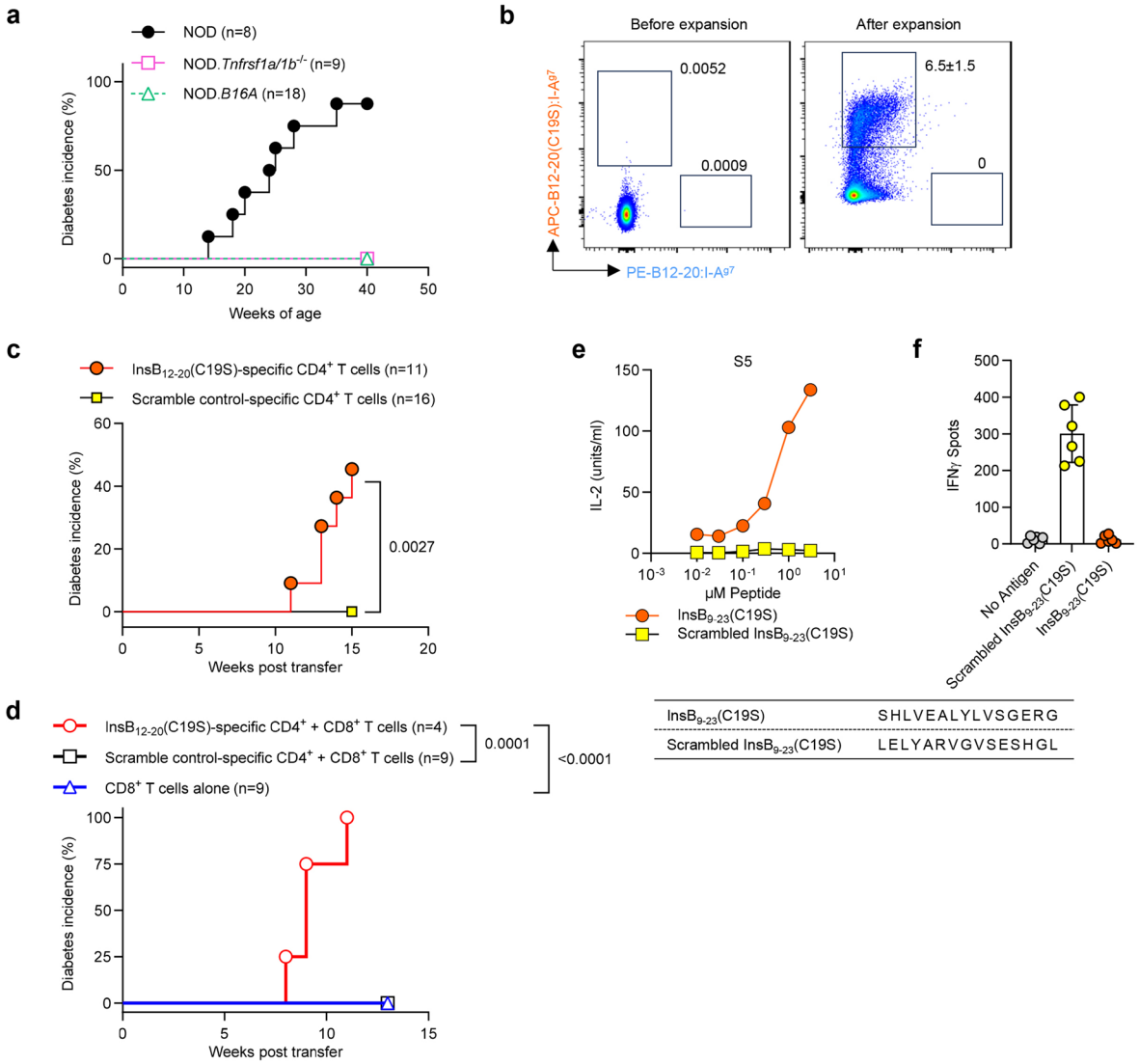
	CDR3α	CDR3β
Clonotype 1	CAASGAGTGSNRLTF	CASSQETGGQAPLF
Clonotype 2	CAASNTGKLTF	CASSQENTYTGQLYF
Clonotype 3	CAASASGGSNYKLTF	CASSQDRGGGNTLYF

Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41590-025-02343-z

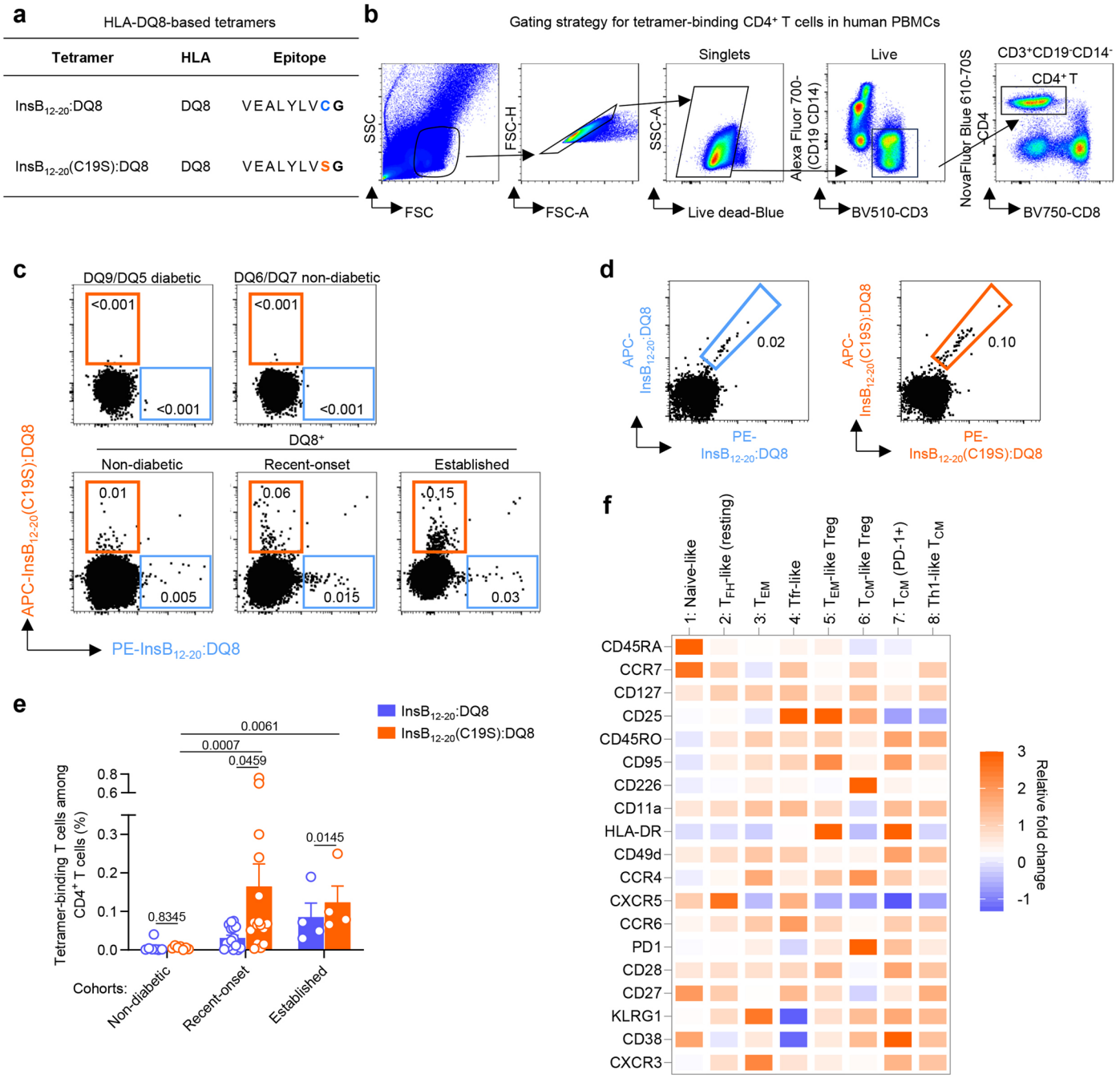
Extended Data Fig. 4 | scRNA-seq analysis of $InsB_{12\text{-}20}$- and $InsB_{12\text{-}20}$(C19S)-specific $CD4^+$ T cells in NOD mice. a, GSEA showing enrichment of cluster 2 for early priming signatures from dataset GSE114824. **b**, Pie charts depicting the distribution of indicated clusters among $InsB_{12\text{-}20}$-specific, $InsB_{12\text{-}20}$(C19S)-specific, and Tet^- $CD4^+$ T cells. **c**, GSEA showing enrichment of cluster 3 for T_{FH} cell signatures from dataset GSE225724. **d**, Volcano plots showing the high expression of effector T cell genes in clusters 5 and 6 relative to the naïve T cell cluster 1. **e**, CDR3α and CDR3β amino acid sequences of the three most expanded clonotypes among $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding $CD4^+$ T cells.

https://doi.org/10.1038/s41590-025-02343-z



Extended Data Fig. 5 | C19S neoepitope-specific CD4⁺T cells mediate islet inflammation and coordinate with CD8⁺T cells to drive diabetes development. **a**, Diabetes incidence of female WT NOD, NOD.*Tnfrsf1a/1b*$^{-/-}$, and NOD.*B16A* mice in our colony. **b**, Representative flow cytometry plots showing $InsB_{12\text{-}20}$:I-A^{g7} and $InsB_{12\text{-}20}$(C19S):I-A^{g7} tetramer-binding $CD4^+$ T cells before and after enrichment with the nested $InsB_{12\text{-}20}$(C19S) peptide. The data show a mild level of enrichment (see Methods). **c**, Diabetes incidence of NOD.*Rag1*$^{-/-}$ mice transferred with $CD4^+$ T cells enriched with the $InsB_{12\text{-}20}$(C19S) or the scrambled control peptide. **d**, Diabetes incidence of NOD.*Rag1*$^{-/-}$ recipients of indicated combinations of $CD4^+$ and $CD8^+$ T cells. **e**, Reactivity of the S5 T cell to the $InsB_{9\text{-}23}$(C19S) peptide and its scrambled version. **f**, ELISPOT assay (mean ± s.e.m.) showing T cell responses (IFNγ production) in draining lymph node cells from NOD mice immunized with the scrambled control peptide upon antigen recall with the indicated peptides (n = 6). In **f**, n is the number of mice; each dot is one mouse. For statistical analysis, log-rank (Mantel-Cox) tests were performed for **c** and **d**. The data represent three (**e**, **f**) and four (**b**) independent experiments.

https://doi.org/10.1038/s41590-025-02343-z



Extended Data Fig. 6 | Analysis of HLA-DQ8-restricted, InsB$_{12\text{-}20}$- and InsB$_{12\text{-}20}$(C19S)-specific CD4$^+$ T cells in T1D patients. a, Table summarizing the HLA-DQ8-based tetramers used in this study. **b**, Representative flow cytometry plots showing the gating strategy for tetramer-binding CD4$^+$ T cells in human PBMCs without magnetic enrichment by spectral flow cytometry analysis. **c**, Representative flow cytometry plots showing co-staining of indicated human PBMC samples with the PE-conjugated InsB$_{12\text{-}20}$:DQ8 and APC-conjugated InsB$_{12\text{-}20}$(C19S):DQ8 tetramers. Flow cytometry plots of three individuals of the non-diabetic, recent-onset, and established T1D cohorts were merged and shown. **d**, Representative flow cytometry plots showing dual-color tetramer staining of PBMCs from recent-onset T1D patients using InsB$_{12\text{-}20}$:DQ8 and InsB$_{12\text{-}20}$(C19S):DQ8 tetramers simultaneously conjugated with PE and APC. Flow cytometry plots from three individual recent-onset patients were merged and shown. **e**, Bar graph (mean ± s.e.m.) showing the frequencies of indicated DQ8 tetramer-binding T cells among CD4$^+$ T cells from all 27 individuals across non-diabetic (n = 7), recent-onset (n = 16), and established T1D (n = 4) cohorts. **f**, Heatmap showing the expression of indicated functional markers by each human T cell cluster. In **e**, n is the number of human subjects; each dot is one participant. For statistical analysis in e, two-tailed Mann-Whitney tests were performed for between-cohort comparisons and two-tailed paired t tests were performed for within-subject comparisons. The data (**e**) represent 14 independent experiments.

nature portfolio

Corresponding author(s): Xiaoxiao Wan

Last updated by author(s): Sep 28, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The mass spectrometry data were collected using XCalibur v 4.0 (ThermoFisher Scientific). Cells were sorted using Aria II (BD bioscience) or analyzed using FACS CantoII (BD bioscience), FACSymphony A5 . Flow cytometry data were collected using the FACS DIVA 9.0 software (BD Bioscience. Spectral flow cytometry data was collected on Cytek Aurora cytometer (Cytek Biosciences) and analyzed with using SpectroFlo software (Cytek Biosciences). The T cell reativity data were collected using the Cell counting software v 5.0 (ImmunoSpot).10x Genomics Library preparation for 5'ScRNA seq: Gene Expression, Enriched TCR, and Feature libraries were prepared according to the recommended protocols from the 10× Genomics Chromium Single Cell 5' Reagent Kits User Guide (v2 Chemistry Dual Index). Normalized libraries were sequenced on a NovaSeq6000 S4 Flow Cell using the XP workflow and a 50x10x10x151 sequencing recipe according to manufacturer protocol.
Data analysis	The mass spectrometry data were analyzed using PEAKS Studio Xpro (version 10.6, build 20201221, Bioinformatics Solutions Inc.). PEAKS PTM and SPIDER searches were used to identify PTMs and SAVs, respectively. The flow cytometry data were analyzed using the Flowjo 10.9.0 software (Tree Star Software).Graphing and statistical analysis was done using Graphpad Prism (version 10).Sc RNA sequencing: Libraries were counted using Cell ranger (v7.1.0). Gene expression data was normalized data and clustered in Loupe Browser for analysis.GSEA was performed using Phantasus80 (https://artyomovlab.wustl.edu/phantasus/).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data generated and supporting the findings of this study are available within the paper. Previously published data are accessible via ProteomeXchange with identifiers PXD024400 (mouse blood leukocytes) and PXD015408 (mouse islet and pancreatic lymph node). Original mass spectrometry data, peptide-spectrum matches, and databases are available via the MassIVE repository with the identifier MSV000096258 (ProteomeXchange identifier PXD057371). A complete list of PTM/SAV search in the crinosome peptidome from stressed MIN6 cells is shown in Supplementary Table 2. Single-cell RNA sequencing (GEO: GSE309696) are available via the National Center for Biotechnology Information Gene Expression Omnibus database.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	The study included both male and female participants. Sex and gender were not used as biological variables in the analysis. In vitro assays involving human islets were conducted using de-identified primary human islets obtained from deceased organ donors through Prodo Laboratories (Aliso Viejo, CA), with no sex-specific data available.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity data were not used in the analysis. Participant recruitment was based on disease status (non-diabetic vs. T1D), HLA-DQ8 genotype, and age.
Population characteristics	The study included both juvenile and adult participants. Cohorts included non-diabetic controls and individuals with type 1 diabetes (T1D) at either 3-month or 18-month post-diagnosis. A separate cohort of HLA-DQ8$^+$ participants was used for immunophenotyping studies.
Recruitment	Healthy adult and children subjects were recruited from Washington University's Volunteers for Health database, by word of mouth, and by targeted mailings. T1D subjects were recruited from the surrounding community by word-of-mouth, Washington University's Volunteers for Health and the WU Adult and Pediatric Endocrinology and Diabetes Clinics. Subjects were identified via their contact with Volunteers for Health office or upon contacting the study coordinator in response to the recruitment materials. After initial contact, each potential volunteer was interviewed via phone to determine eligibility for the study. If eligible, the volunteer was mailed a copy of the consent document to review and scheduled for a study visit. At the study visit they were consented by a physician-investigator (the PI or an M.D. postdoctoral fellow in endocrinology, diabetes and metabolism) or study coordinator. A detailed description of the study was given by the physician-investigator or study coordinator. After sufficient time to answer any questions, if the subject decided to participate, the consent document was signed at the time of the study visit. In addition, eligible minors age 7-12 were also provided with verbal assent and eligible minors age 13-17 were provided with written assent approved by the WU HRPO. For the HLA-II immunopeptidome and MMTT study, 30 participants were enrolled over a four-year period, with 10 participants per cohort: (1) non-diabetic controls (n=10); (2) T1D patients at 3-month onset (n=10); and (3) T1D patients at 18-month onset (n=10). Some non-diabetic participants were siblings of individuals with T1D. For the human T cell biology studies, a separate recruitment was conducted over a two-year period and included three cohorts: (1) non-diabetic controls, (2) recent-onset T1D patients (diagnosed within 12 months), and (3) individuals with established T1D (diagnosed for >5 years). More than 45 individuals were screened for HLA-DQ haplotypes, and PBMC samples from participants expressing at least one copy of the HLA-DQ8 allele were used for the study. Because only two DQ8$^+$ individuals were identified among recruited non-diabetic controls, an additional five DQ8$^+$ samples were obtained from Precision for Medicine, Inc. In total, 19 samples were included in the T cell biology experiments: non-diabetic (n=7), recent-onset (n=8), and established (n=4).
Ethics oversight	All human studies were approved by the Institutional Review Boards of Washington University School of Medicine in St. Louis (IRB Protocol Numbers: 201807004; 202404062).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The immunopeptidomics component of this study was exploratory in nature, and sample size was guided by practical and technical considerations. Due to logistical constraints associated with the MMTT and repeated blood draws, approximately 3–5 mL of blood was collected per participant per time point. By pooling samples from 10 individuals per cohort, we estimated a yield of ~3–12 × 10^6 HLA-II$^+$ cells per group. Based on our experience, this yield fell within a workable range enabling detection of representative β-cell-derived peptides in the HLA-II peptidome. For the human T cell biology study, sample size was first guided by our prior publication28 and the mouse studies in the present study. Based on our experimental design comparing activation phenotypes between C19S- and native insulin-specific CD4$^+$ T cells, we performed a power analysis using a two-sided t-test (α = 0.05, power = 0.80). Assuming a threefold difference in activation marker frequency and a standard deviation of 15–20%, the resulting effect size (Cohen's d = 1.5–2.0) indicated that 5–8 subjects per group would be sufficient to detect this difference with 80% power. Our final cohort included 7 non-diabetic controls, 8 recent-onset patients, and 4 individuals with established T1D, and yielded statistically significant results. For mouse flow cytometry analysis, individual mouse was used and multiple independent experiments with at least 3 mice in each group were performed. For single cell analysis, the number of the mice was determined by pilot experiments identifying the number of the cells can be isolated from one mouse.
Data exclusions	For analysis of memory status, Treg frequency, and activation marker expression of InsB12–20:DQ8 or InsB12–20(C19S):DQ8 tetramer-binding CD4$^+$ T cells in Fig. 7b, e, f, and g, human participants without detectable tetramer-binding CD4$^+$ T cells were excluded from the analysis. No exclusions were applied for quantifying the numbers and frequencies of the tetramer-binding T cells in all human subjects. No data were excluded for any mouse studies.
Replication	All data were reproduced independently as stated in the figure legends
Randomization	Flow Cytometry, sex and age matched mice were used and the mice were assigned randomly to experimental groups. Mass spectrometry samples were block randomized whenever possible. For human studies, participants were assigned to cohorts based on clinical diagnosis and HLA-DQ8 genotype, without randomization.
Blinding	Investigators were blinded to participants' age, ethnicity, and race during experimental design, data collection, and analysis. Blinding to disease state was not feasible due to the nature of the experimental design and the necessity of sample stratification by clinical diagnosis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

The following antibodies were used for mouse studies: All antibodies were used at 1:200 dilution.
anti-CD45 (30-F11) BV50 103138 Biolegend
anti-CD11c (N418) APCcy7 117324 Biolegend
anti-CD11b (M1/70) APCcy7 101226 Biolegend
anti-CD3e (145-2C11) FITC 100306 Biolegend
anti-CD4 (RM4-5) PECy7 100528 Biolegend
anti-CD8a (53-6.7) PerCP-eFluor 710 46-0081-82 eBioiscience
anti-CD44 (IM7) BV421 103040 Biolegend
anti-CD90.2 (53-2.1)BUV737 741701 BD Bioscience
anti-CD90.2 (53-2.1)FITC 140304 eBioscience
anti-B220 (RA3-6B2) PerCy5.5 103236 Biolegend
anti-CXCR6 (SA051D1) BV711 151111 Biolegend
anti-CD62L (MEL-14) FITC 104406 Biolegend
anti-CD62L (MEL-14) BV605 104438 Biolegend
anti-CD69 (H1.2F3) BV785 104543 Biolegend
anti-CD19 (6D5) APCcy7 115530 Biolegend

anti-CXCR3 (CXCR3-173) BV650 126531 Biolegend
Anti-CD11a (M17/4) FITC 11-0111-82 eBioscience
The anti-I-Ag7 antibody (AG2.42.7) was generated in our laboratory.
The following antibodies were used for human studies:
anti-CD19 (HIB19), AF700 302225 (1:200)Biolegend
anti-CD14 (HCD14) AF700 325614 (1:200) Biolegend
anti-CD3 (SK7) BV510 344827 (1/100) Biolegend
anti-CD8 (SK1) BV750 344756 (1/200) Biolegend
anti-CCR7 (G043H7) AF488 353205 (1/50) Biolegend
anti-CXCR3 (G025H7)BV421 302635 (1/100) Biolegend
anti-CD25 (BC96) BV711 302636 (1/100) Biolegend
anti-HLA-DR (L243) PECF594 307653 (1/100) Biolegend
anti-CD28 (CD28.2) BV785 302949 (1/100) Biolegend
anti-CD95 (DX2) PE Fire 640 305657 (1/100) Biolegend
anti-PD-1 (EH12.2H7)BV605 329923 (1/50) Biolegend
anti-CD45RO (UCHL1) PerCP 304221 (1/50) Biolegend
anti-CD38 (S17015A) APC Fire 810 397123 (1/100) Biolegend
anti-CXCR5 (J252D4) Spark NIR 685 356947 (1/50) Biolegend
Anti-CD4 (SK3) NovaFluor Blue 610-70s H001T02B06-A (1/100) eBioscience
anti-CD45RA (HI100)BUV395 363-0458-41 (1/100) eBioscience
anti-KLRG1 (13F12F2) BUV737 367-9488-41 (1/100) eBioscience
anti-CD226 (11A8.7.4) PECy7 25-2269-42 (1/100) eBioscience
Anti-CD127 (HIL-7R-M21)BV650 563225 (1/100) BD Biosciences
anti-CD27 (M-T271) APC-H7 560223 (1/50) BD Biosciences
anti-CCR6 (11A9) BUV805 749361 (1/50) BD Biosciences
anti-CCR4 (1G1) BUV615 613000 (1/50) BD Biosciences
anti-CD11a (G25.2) BUV496 750271 (1/100) BD Biosciences
anti-CD49d (9F10) BUV563 749455 (1/100) BD Biosciences.

Validation

The reactivity of anti-I-Ag7 antibody was confirmed by flow cytometry experiments showing that 95%-100% of the C3g7 APCs were stained positive by the Pacific Blue-conjugated AG2.42.7 antibody, whereas less than 1% of the cells of a I-Ag7 negative control cells line were stained positive with the antibody. Other antibodies are commercially available and validation materials are available on the appropriate websites. Tetramers were validated by the publication using the same tetramers before.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

The CD4+ T cell hybridomas and the C3.g7 B cell lymphoma line (used as APCs) were generated in the laboratory.The CD4+ T cell hybridomas were generated from primary CD4+ T cells from male NOD mice. C3.g7 cells were generated by transducing I-Ag7 to the M12.C3 B cell lymphoma line.

Authentication

The reactivity of each CD4+ T cell hybridoma to their specific antigens were tested and confirmed by their ability to produce IL-2 upon stimulation with antigen presenting cells with cognate antigens.

Mycoplasma contamination

All the cell lines were confirmed as mycoplasma free.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified lines were used in the study.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

NOD/ShiLtJ (NOD), NOD.129S7(B6)-Rag1tm1Mom/J (NOD.Rag1-/-), NOD.Cg-Tg (Ins2*Y16A) 1EllIns1tm1JjaIns2tm1Jja/GseJ (NOD.B16A), NOD.Tnfrsf1a/1b-/-, and C57BL/6 mice were originally obtained from The Jackson Laboratory.Mice were housed in a specific pathogen–free (SPF) facility at the Washington University School of Medicine Division of Comparative Medicine (DCM). Animals were maintained in individually ventilated cages with autoclaved bedding, under controlled temperature (20–24 °C), humidity (40–60%), and a 12-h light/dark cycle. Sterile irradiated chow and acidified/autoclaved water were provided ad libitum, and all cages were supplied with nesting material for enrichment. All procedures were conducted in accordance with institutional guidelines and protocols approved by the Washington University Institutional Animal Care and Use Committee (IACUC)

Wild animals

This study did not include wild animals.

Reporting on sex

The sex of the mice used for the experiments was stated in the main text of the manuscript as well as in the figure legends

Field-collected samples

This study did not include samples collected from the field.

Ethics oversight

All experiments were approved by the Division of Comparative Medicine of Washington University School of Medicine in St. Louis (Accreditation number A3381-01; Protocol No. 23-0429).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Single-cell suspensions from spleen and lymph nodes (inguinal, axillary, and pancreatic lymph node) were prepared by mechanically disrupting the tissue using a syringe plunger through a 70-µm filter. Red blood cells were lysed, and the cell suspension was washed with PBS, and then stained with tetramers. Following tetramer staining, cells were washed with FACS buffer (1% BSA in 1xPBS) and incubated with FcR blocking antibody (2.4G2) for 10 minutes at 4°C. Surface staining was performed with fluorescently labeled antibodies (1:200 v/v) by incubating at 4°C for 30 minutes. Cells were washed twice with FACS buffer. Samples were examined using BD FACS Canto II Cell Analyzer or BD FACS Symphony Cell Analyzer (BD Biosciences). Data was analyzed using FlowJo 10.9.0 software (TreeStar). For spectral flow analysis of human PBMCs. Cryopreserved PBMCs were thawed and washed twice with FACS buffer. Following incubation of PBMC with PE and APC labeled tetramers, cells were stained with an anti-human FcR blocking antibody (human TruStain FcX blocking solution, Biolegend) for 10 minutes at 4°C. To maintain the resolution of positive and negative signals in a multi-color panel, we performed sequential staining of the following markers. PBMCs were first stained with anti-CXCR3 and anti-CCR4 antibodies for 5 minutes, then with anti-KLRG1 and anti-CD25 for 5 minutes at 37°C. The surface antibody cocktail was added directly and incubated for an additional 30 minutes at room temperature. Cells were stained with Live/Dead Blue viability dye (1:1000) for 25 minutes at 4°C. Cells were washed twice, and the flow cytometry data were acquired using Cytek Aurora cytometer (Cytek Biosciences) equipped with 5 lasers, followed by spectral unmixing using SpectroFlo software (Cytek Biosciences). Traditional flow cytometry of unmixed data was analyzed using FlowJo v10.9.
Instrument	Samples were examined using BD FACS Canto II Cell Analyzer or BD FACS Symphony Cell Analyzer (BD Biosciences). For spectral cytometry samples weree xamined using Cytek Aurora cytometer (Cytek Biosciences) equipped with 5 lasers.
Software	Data were collected using FACS DIVA v 8.0.1 (BD Biosciences) andSpectroFlo software (Cytek Biosciences). Data was analyzed using FlowJo 10.9.0 software (TreeStar).
Cell population abundance	Cells were incubated with viability dyes to ensure assessment of live cells. The post sort was performed to confirm a purity above 99%.
Gating strategy	Gating strategies are provided in the supplementary figs, where relevant. Brifely, For mouse tetramer: Cells were gated on singlets, live cells CD45+CD11b- CD19- Thy1.2+ CD4+CD8-Tetramer+ For human tetramer:Cells were gated on singlets, live cells CD19-CD14-CD3+CD4+CD8-Tetramer+

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.