A post-translational cysteine-to-serine conversion in human and mouse insulin generates a diabetogenic neoepitope

Neetu Srivastava[1,2,#], Anthony N. Vomund[1,2,#], Orion J. Peterson[1,2], Omar Abousaway[3], Tiandao Li[4], Lisa Kain[5], Pamela Stone[6], Cristina C. Clement[7], Siddhartha Sharma[5], Bo Zhang[4], Chang Liu[10], Alok V. Joglekar[11,12,13], Laura Campisi[1], Chyi-Song Hsieh[3], Laura Santambrogio[7,8,9], Luc Teyton[5], Ana Maria Arbelaez[6], Cheryl F. Lichti[1,2,*], Xiaoxiao Wan[1,2,*]

[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 and Immunotherapy Programs, 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]*Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA.*

[9]*Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA.*

[10]*Department of Pathology and Immunology, Division of Lab and Genomic Medicine, Washington University School of Medicine, St. Louis, MO, USA.*

[11]*Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.*

[12]*Center for Systems Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.*

[13]*Cancer Immunology and Immunotherapy Program, UPMC Hillman Cancer Center, Pittsburgh, PA, USA.*

[#]*Equal contribution.*

**Corresponding authors*

ABSTRACT

Type 1 diabetes (T1D) affects a genetically susceptible population that develops autoreactive T cells attacking insulin-producing pancreatic β cells. Increasingly, neoantigens are recognized as critical drivers of this autoimmune response. Here, we report a novel insulin neoepitope generated via post-translational cysteine-to-serine conversion (C>S) in human patients, which is also seen in the autoimmune-prone non-obese diabetic (NOD) mice. This modification is driven by oxidative stress within the microenvironment of pancreatic β cells and is further amplified by T1D-relevant inflammatory cytokines, enhancing neoantigen formation in both pancreatic β cells and dendritic cells. We discover that C>S-modified insulin is specifically recognized by $CD4^+$ T cells in human T1D patients and NOD mice. In humans with established T1D, HLA-DQ8-restricted, C>S-specific $CD4^+$ T cells exhibit an activated memory phenotype and lack regulatory signatures. In NOD mice, these neoepitope-specific T cells can orchestrate islet infiltration and promote diabetes progression. Collectively, these data advance a concept that microenvironment-driven and context-dependent post-translational modifications (PTMs) can generate neoantigens that contribute to organ-specific autoimmunity.

INTRODUCTION

In autoimmune diseases, the immune system mistakenly recognizes self-antigens as non-self. Neoantigens are novel forms of self-proteins that can become unique targets for autoreactive T cells, triggering autoimmune responses overlooked by immune tolerance[1,2]. Type 1 diabetes (T1D) is characterized by autoimmune destruction of insulin-producing β cells in the pancreatic islets of Langerhans, which predominantly affects individuals expressing the human leukocyte antigen class II (HLA-II) variants HLA-DQ8 and HLA-DQ2[3,4]. This genetic association forms a foundation for exploring neoantigens presented by risk-conferring HLA-II molecules. T1D involves a broad spectrum of antigens over its chronic course, with the initiation of disease driven by a restricted set of antigens, leading to subsequent epitope spreading. The non-obese diabetic (NOD) mouse, which expresses the I-A^{g7} major histocompatibility complex class II (MHC-II) molecule analogous to human HLA-DQ8[5–7], spontaneously develops diabetes, making it a suitable model for studying antigen-specific responses. Ample proof has established insulin-derived epitopes as key drivers in T1D initiation[8,9]. Identifying insulin neoantigens that are conserved between mice and humans may advance our understanding of T1D pathogenesis and uncover novel therapeutic targets.

Studies in tumor immunology have provided valuable insights into the discovery of disease-relevant neoantigens. In cancer, neoantigens often arise from single nucleotide polymorphisms (SNPs) within tumor cells, and tools such as RNA sequencing and MHC-binding predictions facilitate their discovery[10]. After immunological validation, mass spectrometry (MS)-based immunopeptidomics is used to verify the presence of these neoantigens in the immunopeptidome[11–13]. Beyond genetically encoded mutants, recent studies have revealed a new category of neoantigens, termed "substitutants", which emerge at the translational level when depletion of an amino acid causes its replacement by another one[14,15]. In contrast to cancer, most neoantigens identified in T1D to date result from unexpected post-translational modifications (PTMs) of existing β-cell proteins[16–25], which cannot be identified at the transcriptional level and require MS for their discovery[2,26–28]. Specialized searches are often needed to confirm rare modifications in the proteome or peptidome[29].

The unique physiology of insulin-producing β cells makes them particularly vulnerable to stress-induced PTMs. Insulin biosynthesis, which occurs in oxidative vesicular compartments, exposes β cells to endoplasmic reticulum (ER) and oxidative stress, creating conditions that promote the formation of neoantigens[30–32]. B cells under stress and inflammation have been shown to have elevated formation of carbonylation, deamidation, citrullination, and hybrid insulin peptides (HIPs)[16,24,33–36]. As the most abundant protein produced in β cells, insulin may therefore be subject to stress-associated modifications, creating potential neoantigens.

How neoepitopes are recognized in vivo, in relation to their native counterparts, remains poorly understood in autoimmunity. Once generated in the target tissue, a neoantigen must be properly sampled, processed, and presented by MHC-II-expressing antigen-presenting cells (APCs) in disease-relevant locations. In vivo, a given neoepitope may also need to compete with the vast self-peptidome of the APC to reach a threshold of effective presentation. T cells may specifically recognize the modification carried by the neoepitope or simply cross-react with the native version, shaping the specificity and magnitude of the consequent responses. These selective processes suggest that only limited neoepitopes may elicit meaningful pathogenic reactions. In this study, we report the identification of a conserved insulin neoepitope that meets these criteria. This neoepitope forms via post-translational cysteine-to-serine conversion (C>S) in stressed mouse and human β cells and is presented by APCs bearing I-A^{g7} or HLA-DQ8, leading to the activation of C>S-specific $CD4^{+}$ T cells exhibiting diabetogenic phenotypes in both species.

RESULTS

Identification of C>S insulin peptides in mice and humans with T1D

We previously found in mice that insulin peptides are exocytosed from islets in response to glucose stimulation[37,38]. These released peptides are transiently presented by MHC-II-expressing APCs in the blood within 30 to 120 minutes following glucose injection[38]. Specifically, glucose stimulation facilitated the identification of native insulin peptides belonging to the B-chain 9-23 region (B9-23) (Supplementary Figure 1a), a well-defined immunogenic hotspot in both NOD mice and humans with T1D[9,37,39–43]. This observation suggests that stimulating β-cell degranulation may allow for the detection of islet-derived insulin peptides in the self-peptidome of circulating APCs.

Since direct immunopeptidome analysis of islet samples from T1D patients is largely infeasible, we sought to apply the β-cell degranulation strategy used in mice to examine the HLA-II peptidome of peripheral blood mononuclear cells (PBMCs) from humans. We recruited T1D subjects with either 3- or 18-month onset (3mos, n=10; 18mos, n=10), as well as non-diabetic healthy controls (ND, n=10). Most participants carried at least one copy of the risk-conferring DR3-DQ2 or DR4-DQ8 haplotype (9 out of 10 in 3mos and 18mos; 6 out of 10 in ND) (Supplementary Table 1).

Each participant fasted, had an initial blood draw (time 0), and then underwent a mixed meal tolerance test (MMTT) to stimulate β-cell degranulation. Additional blood samples were taken at 90 and 120 minutes post-MMTT, and C-peptide levels were measured at various time points (Figure 1a). Given the limited sample material (median 5 mL blood per time point), we performed sequential isolation of pan-HLA-DQ and pan-HLA-DR peptidomes. Initial analyses on individual participants were unsatisfactory due to poor peptide recovery, so we combined samples under the same conditions for further analysis.

In the HLA-DQ and HLA-DR peptidomes of ND controls, native insulin B-chain peptides were identified at 90 and 120 minutes post-MMTT but not at baseline (time 0) (Supplementary Figure 1b). T1D subjects at both 3 and 18 months post-onset exhibited impaired C-peptide responses post-MMTT, consistent with reduced β-cell function (Figure 1b). However, several native insulin peptides were still detected in the HLA-II peptidomes of T1D subjects following MMTT (Supplementary Figure 1b), possibly reflecting release from the residual islets. We also identified insulin B-chain peptides where the cysteine 19 residue was oxidized to cysteic acid, a modification that can be introduced during sample preparation[37].

Further examination of the data revealed insulin peptides exhibiting a cysteine-to-serine conversion (C>S) at position 19 of the B-chain (e.g., SHLVEALYLV**C**GERG>SHLVEALYLV**S**GERG) in both HLA-DQ or HLA-DR peptidomes (Figure 1c-d). All these peptides with C>S were only identified post MMTT

(Figure 1d). These findings were confirmed by repeating the database search with C>S insulin sequences appended to the canonical Uniprot-Human database, which yielded additional insulin B-chain peptides with C>S. Moreover, verification of these 11 unique peptide sequences confirmed that all matched synthetic standard controls (Figure 1e; Supplementary Figure 1c). These data suggest that, besides native insulin peptides, the C>S variant can be presented by T1D-predisposing HLA-II molecules.

The identification of C>S insulin peptides in human HLA-II peptidomes prompted us to assess their presence in mice. Reanalysis of I-A^{g7} peptidomes from blood leukocytes[38] revealed a group of peptides with the same C>S conversion, which showed increased abundance following glucose challenge (Figure 1f). Also, in our previous analysis of the I-A^{g7} peptidomes of pancreatic islets and draining lymph nodes (pLNs) in NOD mice[41], we identified and verified the B9-23(C>S) peptide (Supplementary Figure 1d). Furthermore, we assessed soluble peptidomes of mesenteric lymph cannulated from 6-week-old female NOD mice with or without glucose challenge. At baseline (no glucose injection), the lymph peptidome already contained peptides from several secretory protein antigens in β cells, such as native insulin (B-chain and C-peptides), islet amyloid polypeptides (IAPP), and chromogranin A (ChgA), which further increased after glucose challenge (Figure 1g). Notably, the B9-23(C>S) peptide was identified and verified in the lymph peptidome of glucose-injected mice (Supplementary Figure 1e). Thus, insulin peptides with C>S are identified in both soluble and MHC-II-bound forms and can be presented at disease-relevant sites in NOD mice.

C>S modification in insulin peptides is recognized by register-specific $CD4^+$ T cells

To assess the impact of C>S on $CD4^+$ T-cell recognition, we first generated monoclonal $CD4^+$ T-cell hybridomas from NOD mice immunized with synthetic B9-23(C>S) peptide (SHLVEALYLV**S**GERG). We identified a panel of I-A^{g7}-restricted hybridomas, such as the S5 clone, showing robust reactivity to the B9-23(C>S) peptide, while they were largely unresponsive to the native B9-23 peptide (SHLVEALYLV**C**GERG) (Figure 2a). Conversely, the previously characterized 9B9 $CD4^+$ T-cell clone[37,38,44,45], which specifically recognizes native B9-23, showed minimal reactivity to B9-23(C>S) (Figure 2b). These results suggest that C>S alters $CD4^+$ T-cell recognition of the B9-23 peptide.

C>S is positioned within two 9-mer I-A^{g7}-binding cores, B12-20 (VEALYLV**C**G) and B13-21 (EALYLV**C**GE), which use G20 and E21 as their P9 anchor residues[46]. To evaluate how C>S affects register recognition, we synthesized peptides with B12-20(C>S) (VEALYLV**S**G) or B13-21(C>S) (EALYLV**S**GE) as putative registers, each nested within surrogate flanking residues (Supplementary Figure 2a). Both cell-based (Supplementary Figure 2b) and biochemical (Supplementary Figure 2c) binding assays showed that C>S reduced binding affinity to I-A^{g7}, with weakest binding seen for the B12-20(C>S) peptide. Despite

this, most B9-23(C>S)-responsive T-cell hybridomas preferentially recognized the B12-20(C>S), with only few reacting to B13-21(C>S) (Supplementary Figure 2d). Furthermore, such T-cell responses were completely nullified when G20 or E21 were substituted with the inhibitory residue K (Supplementary Figure 2d), indicating that G20 and E21 remain as P9 residues.

Based on these results, we produced I-A^{g7}-based tetramers incorporating either the 9-mer native or C>S versions of the B12-20 and B13-21 epitopes (Supplementary Figure 2e). Using the magnetic enrichment protocol[47], co-staining of secondary lymphoid organs (SLOs) of 8-week-old female NOD mice revealed a largely overlapping T-cell population when using the B13-21:I-A^{g7} and B13-21(C>S):I-A^{g7} tetramers, indicating cross-reactivity (Supplementary Figure 2f-g). In stark contrast, the B12-20(C>S):I-A^{g7} tetramer identified a distinct T-cell population that did not overlap with T cells labeled by the native B12-20:I-A^{g7} tetramer (Figure 2c). Moreover, the frequencies of these B12-20(C>S)-specific T cells were consistently higher than those specific for the native B12-20 epitope (Figure 2c). This indicates that the C>S change in the B12-20 register creates a unique, non-cross-reactive T-cell population. Moreover, since autoreactivity of the native B12-20 epitope has been well documented[37,41,42,48], we decided to focus on the newly identified T cells specific for the B12-20(C>S) epitope (referred to as C>S-specific hereafter) to study their role in T1D.

C>S-specific $CD4^+$ T cells are primed in the periphery and invade pancreatic islets

We performed single-cell RNA sequencing (scRNA-seq) on tetramer-binding C>S-specific $CD4^+$ T cells FACS-sorted from SLOs of 8-week-old female NOD mice, with B12-20:I-A^{g7} tetramer-binding (native-specific) and tetramer-negative (polyclonal) $CD4^+$ T cells as controls. Analysis of the combined data from all three populations revealed five clusters corresponding to different T-cell activation states (Supplementary Figure 3a). Specifically, cluster 1 (mC1) represented naïve T cells, while cluster 2 (mC2) and cluster 3 (mC3) upregulated genes related to tonic TCR signaling and NFκB activation, respectively (Supplementary Figure 3b). Cluster 4 (mC4) resembled regulatory (Treg) and anergy-like T cells, cluster 5 (mC5) exhibited the strongest expression of T-cell activation genes, representing further differentiated effectors (Supplementary Figure 3b).

C>S-specific $CD4^+$ T cells were predominantly found in the mC5 cluster, distinguishing them from polyclonal and native-specific T cells, which were mainly in mC3 (Figure 2d). Compared to mC3, mC5 significantly upregulated genes associated with effector T-cell functions such as cytokine and chemokine signaling, cell migration, Th1/Th17 differentiation, and proliferation (Supplementary Figure 3c), aligning with their higher frequencies relative to native-specific T cells (Figure 2c). This feature of clonal expansion was also reflected by single-cell TCR sequencing, which showed a higher number of overlapped TCR

clonotypes in C>S-specific T cells (Supplementary Figure 3d), with several of the most expanded TCRs corresponding to the mC5 cluster (Supplementary Figure 3e).

To validate these transcriptional data, we used flow cytometry to assess several activation markers highly expressed by mC5, including CD44, CD11a (encoded by *Itgal*), and CXCR6. Most C>S-specific $CD4^+$ T cells displayed a high CD44 but low CD62L expression, in contrast to polyclonal and native-specific T cells (Figure 2e). Similarly, $CD44^+CD11a^+$ and $CD44^+CXCR6^+$ double-positive cells were also significantly higher in C>S-specific cells (Figure 2e), confirming the activation phenotype of C>S-specific $CD4^+$ T cells at the protein level.

To functionally test responses to C>S in diabetes-relevant sites, we stimulated cells pooled from islets and pLNs with equal amounts of B9-23 and B9-23(C>S) peptides and assessed T-cell responses (IFNγ production) via ELISPOT during antigen recall. As reported previously[41,44], recall with native B9-23 induced IFNγ production (Figure 2f). Notably, recall with the B9-23(C>S) peptide also triggered robust IFNγ responses (Figure 2f), indicating functional T-cell responsiveness to C>S in pLN and islets.

The observed priming of C>S-specific $CD4^+$ T cells in the periphery led us to investigate their ability to invade pancreatic islets. Using gene set enrichment analysis (GSEA), we found that the mC5 cluster closely resembled the intra-islet terminal effector $CD4^+$ T-cell subset identified in our previous scRNA-seq analysis[49], with many shared effector genes between the two datasets (Figure 2g). We then used dual B12-20(C>S):I-A^{g7} tetramer labeling in islet infiltrates from 8-week-old female NOD mice and identified an evident population of C>S-specific $CD4^+$ T cells (Figure 2h). Moreover, nearly all of these intra-islet C>S-specific $CD4^+$ T cells were activated ($CD44^+CD62L^-$), unlike the tetramer-negative polyclonal T cells (Figure 2h). These results suggest that C>S-specific $CD4^+$ T cells can enter islets and become activated in situ.

To confirm these results, we generated primary $CD4^+$ T-cell lines specific for C>S (C>S CD4-PTLs) from B9-23(C>S)-immunized NOD mice, expanded them in vitro, and adoptively transferred them to NOD.$Rag1^{-/-}$ recipients. Tetramer staining confirmed substantial and selective expansion of C>S-specific $CD4^+$ T cells prior to transfer, with minimal presence of native-specific T cells (Supplementary Figure 4a), allowing us to evaluate whether C>S-specific T cells could independently trigger autoimmune responses. To further ensure specificity, we designed a scrambled peptide of B9-23(C>S) (LELYARVGVSESHGL), which could not be recognized by the S5 T cell (Supplementary Figure 4b) but was immunogenic (Supplementary Figure 4c) and generated primary T-cell lines against this scrambled peptide (scramble CD4-PTLs) as a non-specific control.

Adoptive transfer of C>S CD4-PTLs caused extensive infiltration of $CD45^+$ leukocytes into the islets of NOD.Rag1$^{-/-}$ recipients, with clear invasion of the transferred $CD4^+$ T cells (Supplementary Figure 4d). Tetramer staining revealed evident C>S-specific $CD4^+$ T cells in the SLOs of the recipient mice, with native-specific T cells notably absent (Supplementary Figure 4e). About half of the NOD.Rag1$^{-/-}$ recipients receiving only C>S CD4-PTLs developed clinical diabetes (Supplementary Figure 4f), and all mice co-transferred with C>S CD4-PTLs and $CD8^+$ T cells (purified from the spleens of regular NOD mice) became diabetic (Supplementary Figure 4g). Conversely, none of the NOD.Rag1$^{-/-}$ mice given $CD8^+$ T cells alone, or $CD8^+$ T cells co-transferred with scramble CD4-PTLs, developed diabetes (Supplementary Figure 4g). These data suggest that C>S-specific T cells can migrate to the islets, mediate inflammation, and promote full diabetes development by assisting cytotoxic $CD8^+$ T cells.

C>S is formed in β-cell granules via redox stress

The observed pathogenicity of C>S-specific $CD4^+$ T cells prompted us to investigate the source and mechanisms of C>S formation in pancreatic islets. We hypothesized that C>S might be generated in secretory granules of pancreatic β cells, which are crucial sites for insulin biosynthesis and autoantigen generation[50,51]. Previous studies have found distinct insulin products in two sets of β-cell granules[37,44]: insulin dense-core granules (DCGs), which primarily store intact insulin, and crinosomes, a minor set of degradative vesicles generated through the fusion of excessive DCGs with lysosomes and primarily contain shorter peptide fragments.

We employed a granule-based antigen presentation (GAP) assay to assess native and C>S insulin peptides in DCGs and crinosomes (Supplementary Figure 5a). We isolated subcellular granule fractions representing DCGs and crinosomes from β cells and offered them to the I-A^{g7} expressing C3.g7 APC to allow the presentation of insulin epitopes derived from these granules. The resulting epitope-specific responses were then probed by corresponding $CD4^+$ T-cell hybridomas. We validated the identity of the crinosome and DCG fractions by assessing their content via the GAP assay and measuring lysosomal activity through acid phosphatase levels (Supplementary Figure 5b-d). To differentiate between native and C>S insulin peptides, we utilized the specific reactivity of the 9B9 and S5 T cells (Figure 2a-b). Notably, both T cells exhibited comparable reactivities to their respective peptides (Supplementary Figure 5e), allowing for quantitative comparisons.

The identification of C>S in healthy humans suggested its presence in β cells even in the absence of autoimmunity. To test this possibility, we examined granules isolated from the MIN6 β-cell line (derived from diabetes-resistant B6 mice) and primary islets from B6 mice. Both sources of β cells showed similar results. As reported previously[37,44], the native B9-23 peptide (assayed by 9B9) was mostly found in the

crinosome, with minimal presence in the DCG fraction (Supplementary Figure 5f). The crinosome fraction also activated the S5 T cell, indicating the presence of C>S peptides, but the responses were much lower, about 30~50% of those observed with 9B9 (Supplementary Figure 5f). In contrast, S5 T-cell responses to the DCG fraction were minimal (Supplementary Figure 5f). These data suggest that, under non-diabetic conditions, C>S insulin peptides are primarily localized in crinosomes at low levels.

During insulin biosynthesis, the initial disulfide bond formation occurs between cysteine 19 of the B-chain and cysteine 20 of the A-chain[52,53]. This C(B19)-C(A20) pairing is crucial for proper insulin folding in the ER. We reasoned that C>S on the C(B19) residue might indicate insulin misfolding, a phenomenon closely linked to ER stress[54,55]. To test this, we briefly exposed MIN6 cells and B6 mouse islets to the ER stress inducer tunicamycin for 2 hours, followed by extensive washing. This short treatment did not cause β-cell death or alter total protein levels in the granule samples (Supplementary Figure 5g).

The S5 T cell showed significantly stronger responses to crinosome fractions isolated from MIN6 cells and B6 mouse islets treated with tunicamycin (Supplementary Figure 5h), suggesting that ER stress can promote C>S formation. Notably, this increase in C>S formation was significantly inhibited by the antioxidants glutathione or N-acetylcysteine (NAC), a phenomenon consistently observed between the two sources of β cells (Supplementary Figure 5h), suggesting that C>S forms via redox stress. These results align with previous studies showing that ER stress can induce the production of reactive oxygen species (ROS), resulting in redox stress[56] and with evidence that cysteine can be converted to serine under redox stress conditions[57]. In contrast, although the tunicamycin treatment also elevated native B9-23 levels in crinosomes, this effect was not influenced by NAC or glutathione (Supplementary Figure 5i). Additionally, at baseline, DCGs from MIN6 cells and B6 mouse islets contained minimal native or C>S insulin peptides, while tunicamycin increased both forms (Supplementary Figure 5j-k). Glutathione and NAC also selectively inhibited the increase of C>S (Supplementary Figure 5j) but not the native (Supplementary Figure 5k) form in DCGs. Thus, unlike the native insulin peptides, C>S formation is specifically driven by redox stress.

Unbiased identification of C>S modifications in insulin under ER stress

The targeted detection of C>S using the GAP assay prompted us to investigate whether C>S could be unbiasedly identified following ER stress. Considering that C>S was mostly detected in degraded insulin peptides in crinosomes and that the proteome profile of MIN6 cells closely resembles primary β cells[58], we isolated crinosome peptidomes from equal numbers of MIN6 cells treated with vehicle or tunicamycin for 2 hours (Figure 3a). The samples were prepared in triplicate, fractionated to increase depth of coverage, and block randomized for analysis by MS to minimize confounding effects during the analytical process.

After characterizing the protein composition through the initial PEAKS DB search, we performed sequential PEAKS PTM and SPIDER searches to identify initially unspecified PTMs and single amino acid variants (SAVs), respectively. For quantification, we collated PTMs and SAVs at the peptide-spectrum match (PSM) level to perform an unbiased survey of all changes between the two conditions. Among all identified PTMs and SAVs (total 348 entries; Supplementary Table 2), C>S conversion exhibited the most significant increase following tunicamycin treatment (Figure 3b). Other cysteine modifications, such as oxidation and hydroxylation, also showed substantial increase under ER stress (Figure 3b). Dehydroalanine, which can be formed via degradation of an oxidized cysteine sulfhydryl group and is also proposed to be a precursor of serine[57], was also increased (Figure 3b). We did not observe significant changes in oxidative modifications on other residues (e.g., histidine, tryptophan, proline, methionine) (Supplementary Table 2), suggesting that the observed cysteine oxidations reflect biological changes rather than sample preparation artifacts. Collectively, these results suggest that cysteine modifications, particularly C>S, are rapidly induced in β cells responding to acute ER stress.

Tracing C>S back to all source peptides revealed that most of these changes were localized to the insulin B chain. To verify these results, we created a database in which all cysteine residues were replaced by serine and found no additional peptides with C>S. Next, we appended C>S forms 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 the modified insulin B-chain peptides, most of which exhibited C>S at the 19th position (Supplementary Figure 6a), confirming findings of the MHC-II peptidome analyses in mice and humans. Some peptides had dual C>S modifications at both the 7th and 19th positions (e.g., **C**GSHLVEALYLV**C**GERG→**S**GSHLVEALYLV**S**GERG). As the C(B7)-C(A7) interchain disulfide bond only forms under highly oxidative conditions[52], this further supports the idea that C>S is a result of insulin misfolding under oxidative stress. Quantitative analysis revealed a significant increase in the abundance of all C>S insulin peptides following tunicamycin treatment (Figure 3c; Supplementary Figure 6a).

Next, we focused on assessing all modifications on the C(B19) residue. In the vehicle control, C(B19) exhibited modifications like cysteinylation and glutathionylation, with cysteic acid formation, dehydroalanine formation, and C>S conversion less prominent (Figure 3d). However, after the two-hour tunicamycin treatment, C>S conversion emerged as the dominant modification (Figure 3d), indicating that C(B19) is already susceptible to modifications in the steady state and can be converted to serine under ER stress. Collectively, these findings suggest that in β cells undergoing ER stress, C>S modifications can occur at the post-translational level, generating insulin neoepitopes.

Inflammatory cytokines promote C>S formation in β cells and DCs

To test C>S formation in autoimmune conditions, we first used three B9-23(C>S)-specific $CD4^+$ T-cell hybridomas to examine the in vivo presentation of C>S insulin peptides in pancreatic islets from female NOD mice. All three T cells responded to dispersed islet cells in the absence of exogenous antigen (Supplementary Figure 6b), indicating natural presentation of C>S insulin peptides in islets. This spontaneous presentation was significantly higher in islets from prediabetic NOD mice with advanced autoimmunity (17-week-old) compared to those at the initial stage (5-week-old) (Supplementary Figure 6b), implicating that the autoinflammatory islet microenvironment may further enhance C>S formation.

T1D development is associated with inflammatory cytokine signaling in pancreatic islets, with IL-1β, IFNγ, and TNFα known to induce ER and redox stress in β cells[59,60]. Since inflamed islets from NOD mice may already have active cytokine signaling, which may interfere with the assessment of individual cytokine effects, we tested the influence of IL-1β, IFNγ, and TNFα on C>S formation in MIN6 cells. Initially, we conducted a pilot screen using a glucose-induced antigen transfer assay, which suggested that TNFα or IFNγ exposure increased production of C>S insulin peptides, with TNFα showing the strongest effect (Supplementary Figure 6c). Confirming these results, GAP assay also revealed a significant increase in S5 T-cell responses to crinosome fractions from TNFα-treated MIN6 cells (Supplementary Figure 6d). This increase was inhibited by glutathione or by tauroursodeoxycholic acid (TUDCA) (Supplementary Figure 6d), an ER stress inhibitor known to suppress T1D development[61]. In contrast, TNFα did not affect native B9-23 levels and had minimal impact on DCGs (Supplementary Figure 6d). Thus, inflammatory cytokines can enhance C>S production in β cells through ER and redox stress.

In addition to targeting β cells, inflammatory cytokines also act on DCs, mediating their activation and licensing. While DCs are well known for presenting diabetogenic antigens, including the native B9-23 peptide[44,62], their role in forming neoantigens remains understudied. To explore whether C>S could be generated within DCs under cytokine stimulation, we enriched DCs from NOD mice treated with FLT3 ligand and pulsed them with the native B9-23 peptide. Equal aliquots of these B9-23-pulsed DCs were then exposed to different cytokines, with or without the antioxidant NAC; the presentation levels of native and C>S forms of B9-23 were probed by the 9B9 and S5 T cells (Figure 3e). We measured the ratio of S5 to 9B9 T-cell responses (S5:9B9) to evaluate the generation of C>S from the pulsed native peptide. Stimulation with IL-1β, TNFα, and IFNγ significantly increased the S5:9B9 ratio, suggesting C>S formation from the native peptide, while the influences of IL-4, IL-6, and IL-12 were less profound (Figure 3f). Furthermore, NAC largely blocked C>S generation induced by IL-1β, TNFα, and IFNγ (Figure 3f), indicating the dependence of redox stress.

The antigen presentation assay suggested that cytokine stimulation allowed the weaker-binding C>S peptides to integrate into the broad and competitive DC self-peptidome. To investigate this, we isolated I-A^{g7} peptidomes from B9-23-pulsed DCs followed by treatment with TNFα or TNFα plus NAC and analyzed by MS (Figure 3g). While the native B9-23 peptide was identified in the DC peptidome under all conditions, the B9-23(C>S) peptide appeared in TNFα-stimulated DCs (Figure 3h). We observed a significant increase in all identified C>S insulin peptides in three independent experiments following TNFα stimulation, which was inhibited by NAC (Figure 3i). Thus, inflammatory cytokines like TNFα can trigger C>S formation from native insulin peptides in DCs, allowing MHC-II presentation of this modified neoepitope, which may serve as a mechanism for insulin epitope spreading during diabetes progression.

We then assessed C>S-specific $CD4^+$ T cells in NOD mice deficient with both TNF receptor genes (NOD.Tnfsf1a/1b$^{-/-}$), which are diabetes resistant (Supplementary Figure 7a). There was a significant reduction in the proportions of $CD44^+CD62L^-$, $CD44^+CD11a^+$, and $CD44^+CXCR6^+$ cells in tetramer-binding C>S-specific $CD4^+$ T cells in SLOs of 8-week-old female NOD.Tnfsf1a/1b$^{-/-}$ mice (Supplementary Figure 7b), suggesting that a TNF-mediated inflammatory process is required for optimal activation of C>S-specific $CD4^+$ T cells.

To further assess C>S-specific T cells in a diabetes-resistant context, we used NOD.B16A mice, which lack endogenous insulin genes but express an insulin transgene with a tyrosine-to-alanine mutation at the 16th position of the B-chain (Y16A)[9]. This mutation blocks T-cell reactivity to the native B9-23 peptide and protects against insulitis and diabetes development[9] (Supplementary Figure 7a). As expected, the native B9-23-specific 9B9 T cell showed no reactivity to the B9-23(Y16A) peptide (SHLVEAL**A**LVCGERG) (Supplementary Figure 7c). Furthermore, the Y16A mutation also nullified the recognition by the B9-23(C>S)-specific S5 T cell (Supplementary Figure 7c), indicating that NOD.B16A mice lack the necessary epitope for C>S-specific T-cell recognition. In vivo, C>S-specific $CD4^+$ T cells in NOD.B16A mice exhibited a significant reduction in both cell frequencies (Supplementary Figure 7d) and CD44 expression levels (Supplementary Figure 7e). Thus, in the absence of the C>S neoepitope, C>S-specific $CD4^+$ T cells become inactivated in the diabetes-resistant NOD.B16A mice.

C>S formation in human islets under stress and inflammation

To explore the relevance of our findings in mice to human T1D, we examined C>S formation in human islets under stress and inflammatory conditions. We first applied the GAP assay to crinosome and DCG fractions from islets of four non-diabetic donors (Supplementary Table 3) and observed similar results to those obtained from MIN6 cells and mouse islets. Tunicamycin treatment led to an increase in both native and C>S insulin peptides in the crinosome and DCG fractions (Figure 4a). However, the addition of

glutathione alongside tunicamycin selectively inhibited C>S formation, while leaving native B9-23 unchanged (Figure 4a). These results suggest that redox stress is a conserved mechanism for C>S formation in β-cell granules across both mice and humans.

To unbiasedly assess C>S formation in human islets under cytokine stimulation, we performed MS analysis of crinosome peptidomes isolated from non-diabetic human islets (Supplementary Table 3) exposed to IL-1β, TNFα, IFNγ, and a combination of all three cytokines. Peptides from common β-cell proteins, including native insulin, ChgA, and IAPP, were identified in crinosome peptidomes across all conditions in islet samples from two donors (Figure 4b). Insulin peptides with C>S were below the detection limit in the crinosome peptidomes from vehicle-treated islets, possibly due to limited sample material, but they were detected in various cytokine stimulation conditions (Figure 4b). Specifically, we identified a shared sequence of the B9-23(C>S) peptide in cytokine-stimulated islets from both donors (Figure 4c). Thus, as T1D-relevant cytokines, IL-1β, TNFα, and IFNγ are capable of promoting C>S in human β cells.

C>S specific, activated-memory CD4$^+$ T cells appear in humans with established T1D

To examine whether C>S-specific CD4$^+$ T cells are also present in humans with T1D, we generated HLA-DQ8-based tetramers incorporating the B12-20 and B12-20(C>S) epitopes (Supplementary Figure 8a). In human PBMCs, CD4$^+$ T cells specific for native insulin peptides are markedly rare and exhibit heterogenous phenotypes across different stages of T1D[43]. Given this, we decided to focus on patients with established diabetes, a clinically relevant condition, and examined a cohort of four patients expressing a single copy of HLA-DQ8 (Supplementary Figure 8b). In PBMCs from each patient, we detected a rare but distinct, non-cross-reactive population of C>S-specific CD4$^+$ T cells (without magnetic enrichment) (Figure 4d; Supplementary Figure 7c). Moreover, in each patient, both cell frequencies (Figure 4e) and absolute numbers (Figure 4f) of C>S-specific CD4$^+$ T cells were consistently higher than native-specific CD4$^+$ T cells. Thus, C>S-specific T cells are readily detectable in the blood of HLA-DQ8-positive T1D patients and, consistent with our NOD mouse observations, are more abundant than native B12-20-specific CD4$^+$ T cells.

We used a 27-marker spectral flow cytometry panel to examine the functional characteristics of C>S-specific CD4$^+$ T cells. We compared their phenotypes directly with the native-specific and polyclonal T cells in the same individual. Consistent with previous studies showing that insulin-reactive CD4$^+$ T cells in patients with established diabetes mainly exhibit a memory phenotype[43], we found that both native- and C>S-specific tetramer-binding T cells were primarily CD45RA$^-$CD95$^+$, in contrast to the tetramer-negative polyclonal population (Supplementary Figure 8d). Clustering of the three T-cell populations from all four patients (with comparable cell numbers) revealed eight subsets that fell into three major phenotypic groups: naïve (hC1, hC2), memory (hC3-hC6), and Tregs (hC7, hC8)[63–65] (Supplementary Figure 8e). Naïve cells

(hC1, hC2) expressed markers typical of inactivation ($CD45RA^+CCR7^+$), with hC2 additionally expressing CXCR5 (Supplementary Figure 8f). Memory cells were divided into central memory (hC3: $CD45RA^-CCR7^+$) and activated Th1 memory cells (hC4: $CD45RO^+CD95^+CXCR3^+KLRG1^+$), with effector memory cells (hC5, hC6) lacking both CD45RA and CCR7 (Supplementary Figure 8f). Two Treg subsets (hC7, hC8) were defined by high CD25 and low CD127 expression; hC7 represented memory Tregs ($CD45RO^+CD95^+HLA\text{-}DR^+$), while hC8 was distinguished by high expression of CD226 and PD-1 (Supplementary Figure 8f).

Two major differences emerged when comparing C>S-specific $CD4^+$ T cells with native-specific and polyclonal $CD4^+$ T cells. First, nearly 50% of C>S-specific $CD4^+$ T cells were enriched in the Th1 memory hC4 subset, whereas only ~5% of native-specific $CD4^+$ T cells fell within hC4 (Figure 4g). Confirming this, C>S-specific $CD4^+$ T cells showed higher proportions of cells expressing activation markers CD11a, CXCR3, KLRG1 and CD49d, compared to native-specific and polyclonal $CD4^+$ T cells (Supplementary Figure 8g), indicating an activated Th1 phenotype. Second, C>S-specific $CD4^+$ T cells had a substantial reduction in the Treg subsets, with only ~3% classified into hC7 and hC8[66,67] (Figure 4g). In contrast, a Treg signature was notable in native-specific $CD4^+$ T cells, as evidenced by the exclusive presence of the $CD226^+$ hC8 Treg subset (Figure 4g) and higher proportions of $CD25^+CD127lo$ cells (Supplementary Figure 8h). These phenotypic differences between C>S-specific and native-specific $CD4^+$ T cells were consistent across all four patients (Supplementary Figure 9).

Collectively, these results show that C>S-specific $CD4^+$ T cells in humans with established T1D exhibit an activated memory phenotype. This profile generally resembles their counterparts in NOD mice, suggesting that C>S-specific $CD4^+$ T cells represent a conserved autoreactive population across species, contributing to the pathogenesis of T1D.

DISCUSSION

This study reveals a novel mechanism underlying T1D pathogenesis. We demonstrate that a C>S modification at position 19 of the insulin B-chain creates a neoepitope recognized by autoreactive $CD4^+$ T cells in both NOD mice and human T1D patients. This C>S modification is closely linked to the highly oxidative microenvironment of pancreatic β cells. The high abundance of insulin and redox sensitivity of the C(B19) residue likely sets up ideal conditions for C>S formation. Consistent with this, only the most redox-sensitive cysteine in GAPDH overexpressed in human embryonic kidney cells was converted to serine during exposure to hydrogen peroxide[57]. This context-driven C>S formation is distinct from genetic mutations seen in MIDY syndrome (mutant INS gene–induced diabetes of youth), which causes permanent neonatal diabetes, a condition with a distinct pathobiology from autoimmune T1D[68–70]. Our findings align with the evidence suggesting that neoantigens with redox-associated PTMs contribute to T1D pathogenesis and that inhibiting ER and redox stress can mitigate disease progression[61,71–75].

The robust responses against to the C>S neoepitope may be attributed to several factors. First, the unique oxidative environment required for efficient C>S formation may be absent in the thymus, limiting the establishment of central tolerance. This is supported by our analysis of the thymic MHC-II peptidome in NOD mice, which detected the native B9-23 insulin peptide but not its C>S-modified form[44]. Second, inflammatory cytokines associated with T1D can further enhance C>S formation, potentially creating a feed-forward loop that amplifies the activation of C>S-specific $CD4^+$ T cells in vivo. We observed enhanced C>S formation in human islets stimulated with TNFα, IL-1β, and IFNγ, suggesting that this process may persist in the residual islets of T1D patients. Furthermore, inflammatory cytokines also promote C>S formation in dendritic cells, expanding the potential sites of neoantigen presentation. These amplification mechanisms may contribute to the shift from physiological autoreactivity to pathological autoimmunity, a process observed in many chronic autoimmune diseases but remains poorly understood[76]. Overall, the C>S modification exemplifies how a single amino acid change can profoundly impact T-cell reactivity in autoimmunity. This phenomenon is reminiscent of altered peptide ligands (APLs), where subtle alterations in peptide antigens can dramatically alter T-cell responses[77] and echoes the role of immunogenic tumor neoepitopes in driving anti-tumor immunity.

Our data suggest that C>S modification can occur in vivo as a natural consequence of cellular processes, consistent with previous identification of similar modifications in mouse liver mitochondria and melanoma cells[57]. This implies that C>S modification may occur on redox-sensitive residues in abundant self-proteins under oxidative tissue microenvironments pertinent to other autoimmune conditions and cancers.

This study has several limitations. The availability of samples from juvenile participants restricted our ability to perform robust quantitative statistical comparisons, and future studies with larger cohorts are needed. Although we focused on HLA-DQ8, the identification of C>S insulin peptides in both HLA-DQ and HLA-DR peptidomes suggests potential binding to multiple T1D-risk HLA-II molecules, raising the possibility that C>S may have broader implications in T1D pathogenesis. Our stringent filtering criteria during peptide analysis may have inadvertently missed some modified peptides, and the use of a short-term ER stress model may explain why certain redox-associated PTMs (i.e., deamidation, citrullination, carbonylation) did not show substantial changes. We found that the loss of reactivity to the C>S neoepitope may also contribute to diabetes resistance in NOD.B16A mice. However, further dissecting its contribution relative to the native epitope is challenging. Definitive evidence may require selectively ablating C>S formation without impacting native insulin, which is technically difficult since manipulating the C(B19) residue may disturb normal insulin biosynthesis and β-cell function. Given results from the present study, we suggest that responses to the C>S neoepitope may be particularly relevant in stressed inflammatory islet microenvironment.

Recent insights into cysteine carboxyethylation-driven neoantigen formation in ankylosing spondylitis[78] and neoself-antigen presentation in systemic lupus erythematosus[79] further emphasize the diverse mechanisms by which neoantigens drive autoimmunity. Identifying conserved neoantigens may offer an integrated framework bridging mouse and human studies to better understand disease pathogenesis and therapeutic intervention.

REFERENCES

1. Mustelin, T. & Andrade, F. Autoimmunity: the neoantigen hypothesis. *Front. Immunol.* **15**, (2024).
2. Stern, L. J., Clement, C., Galluzzi, L. & Santambrogio, L. Non-mutational neoantigens in disease. *Nat Immunol* **25**, 29–40 (2024).
3. 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).
4. 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. U.S.A.* **84**, 2435–2439 (1987).
5. Corper, A. L. *et al.* A structural framework for deciphering the link between I-Ag7 and autoimmune diabetes. *Science* **288**, 505–511 (2000).
6. 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).
7. 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).
8. Wegmann, D. R., Norbury-Glaser, M. & Daniel, D. Insulin-specific T cells are a predominant component of islet infiltrates in pre-diabetic NOD mice. *Eur. J. Immunol.* **24**, 1853–1857 (1994).
9. 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).
10. Gubin, M. M. *et al.* Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* **515**, 577–581 (2014).
11. Alspach, E. *et al.* MHC-II neoantigens shape tumour immunity and response to immunotherapy. *Nature* **574**, 696–701 (2019).
12. Huber, F. *et al.* A comprehensive proteogenomic pipeline for neoantigen discovery to advance personalized cancer immunotherapy. *Nat Biotechnol* 1–13 (2024) doi:10.1038/s41587-024-02420-y.
13. Kraemer, A. I. *et al.* The immunopeptidome landscape associated with T cell infiltration, inflammation and immune editing in lung cancer. *Nat Cancer* **4**, 608–628 (2023).
14. Pataskar, A. *et al.* Tryptophan depletion results in tryptophan-to-phenylalanine substitutants. *Nature* **603**, 721–727 (2022).
15. Yang, C. *et al.* Arginine deprivation enriches lung cancer proteomes with cysteine by inducing arginine-to-cysteine substitutants. *Molecular Cell* **84**, 1904-1916.e7 (2024).
16. Yang, M.-L. *et al.* Carbonyl Posttranslational Modification Associated With Early-Onset Type 1 Diabetes Autoimmunity. *Diabetes* **71**, 1979–1993 (2022).
17. 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).
18. McGinty, J. W. *et al.* Recognition of Posttranslationally Modified GAD65 Epitopes in Subjects With Type 1 Diabetes. *Diabetes* **63**, 3033–3040 (2014).
19. Delong, T. *et al.* Diabetogenic T-Cell Clones Recognize an Altered Peptide of Chromogranin A. *Diabetes* **61**, 3239–3246 (2012).
20. Delong, T. *et al.* Islet Amyloid Polypeptide Is a Target Antigen for Diabetogenic $CD4^+$ T Cells. *Diabetes* **60**, 2325–2330 (2011).
21. Mannering, S. I. *et al.* The insulin A-chain epitope recognized by human T cells is posttranslationally modified. *Journal of Experimental Medicine* **202**, 1191–1197 (2005).

22. 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).
23. 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).
24. 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).
25. Azoury, M. E. *et al.* $CD8^+$ T Cells Variably Recognize Native Versus Citrullinated GRP78 Epitopes in Type 1 Diabetes. *Diabetes* **70**, 2879–2891 (2021).
26. McGinty, J. W., Marré, M. L., Bajzik, V., Piganelli, J. D. & James, E. A. T Cell Epitopes and Post-Translationally Modified Epitopes in Type 1 Diabetes. *Curr Diab Rep* **15**, 90 (2015).
27. Yang, M.-L., Kibbey, R. G. & Mamula, M. J. Biomarkers of autoimmunity and beta cell metabolism in type 1 diabetes. *Front. Immunol.* **13**, (2022).
28. Lichti, C. F. & Wan, ×. Using mass spectrometry to identify neoantigens in autoimmune diseases: The type 1 diabetes example. *Seminars in Immunology* **66**, 101730 (2023).
29. Clement, C. C. *et al.* Pleiotropic consequences of metabolic stress for the major histocompatibility complex class II molecule antigen processing and presentation machinery. *Immunity* **54**, 721-736.e10 (2021).
30. 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. *Journal of Diabetes Research* **2012**, 238980 (2012).
31. 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. *Molecular Metabolism* **54**, 101365 (2021).
32. Marre, M. L., James, E. A. & Piganelli, J. D. β cell ER stress and the implications for immunogenicity in type 1 diabetes. *Front. Cell Dev. Biol.* **3**, (2015).
33. Yang, M.-L. *et al.* Citrullination of glucokinase is linked to autoimmune diabetes. *Nat Commun* **13**, 1870 (2022).
34. Delong, T. *et al.* Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* **351**, 711–714 (2016).
35. Marré, M. L. *et al.* Inherent ER stress in pancreatic islet β cells causes self-recognition by autoreactive T cells in type 1 diabetes. *J Autoimmun* **72**, 33–46 (2016).
36. Wenzlau, J. M. *et al.* Mapping of a hybrid insulin peptide in the inflamed islet β-cells from NOD mice. *Frontiers in Immunology* **15**, (2024).
37. Wan, X. *et al.* Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature* **560**, 107–111 (2018).
38. Vomund, A. N. *et al.* Blood leukocytes recapitulate diabetogenic peptide-MHC-II complexes displayed in the pancreatic islets. *J Exp Med* **218**, (2021).
39. Daniel, D., Gill, R. G., Schloot, N. & Wegmann, D. Epitope specificity, cytokine production profile and diabetogenic activity of insulin-specific T cell clones isolated from NOD mice. *European Journal of Immunology* **25**, 1056–1062 (1995).
40. Michels, A. W. *et al.* Islet-Derived CD4 T Cells Targeting Proinsulin in Human Autoimmune Diabetes. *Diabetes* **66**, 722–734 (2016).
41. Wan, X. *et al.* The MHC-II peptidome of pancreatic islets identifies key features of autoimmune peptides. *Nat Immunol* **21**, 455–463 (2020).

42. 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. *Science Immunology* **4**, eaaw6329 (2019).
43. Sharma, S. *et al.* Measuring anti-islet autoimmunity in mouse and human by profiling peripheral blood antigen-specific CD4 T cells. *Science Translational Medicine* **15**, eade3614 (2023).
44. 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).
45. Srivastava, N. *et al.* Chromogranin A Deficiency Confers Protection From Autoimmune Diabetes via Multiple Mechanisms. *Diabetes* **70**, 2860–2870 (2021).
46. 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. *Journal of Experimental Medicine* **208**, 2375–2383 (2011).
47. Moon, J. J. *et al.* Tracking epitope-specific T cells. *Nat Protoc* **4**, 565–581 (2009).
48. 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).
49. Srivastava, N. *et al.* CXCL16-dependent scavenging of oxidized lipids by islet macrophages promotes differentiation of pathogenic $CD8^+$ T cells in diabetic autoimmunity. *Immunity* **0**, (2024).
50. 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).
51. Amdare, N., Purcell, A. W. & DiLorenzo, T. P. Noncontiguous T cell epitopes in autoimmune diabetes: From mice to men and back again. *Journal of Biological Chemistry* **297**, 100827 (2021).
52. Haataja, L. *et al.* Disulfide Mispairing During Proinsulin Folding in the Endoplasmic Reticulum. *Diabetes* **65**, 1050–1060 (2016).
53. Arunagiri, A. *et al.* Proinsulin misfolding is an early event in the progression to type 2 diabetes. *eLife* **8**, e44532 (2019).
54. Sun, J. *et al.* Proinsulin misfolding and endoplasmic reticulum stress during the development and progression of diabetes☆. *Molecular Aspects of Medicine* **42**, 105–118 (2015).
55. Arunagiri, A. *et al.* Misfolded proinsulin in the endoplasmic reticulum during development of beta cell failure in diabetes. *Annals of the New York Academy of Sciences* **1418**, 5–19 (2018).
56. Cao, S. S. & Kaufman, R. J. Endoplasmic Reticulum Stress and Oxidative Stress in Cell Fate Decision and Human Disease. *Antioxidants & Redox Signaling* **21**, 396–413 (2014).
57. Jeong, J. *et al.* Novel Oxidative Modifications in Redox-Active Cysteine Residues*. *Molecular & Cellular Proteomics* **10**, M110.000513 (2011).
58. Sacco, F. *et al.* Glucose-regulated and drug-perturbed phosphoproteome reveals molecular mechanisms controlling insulin secretion. *Nat Commun* **7**, 13250 (2016).
59. Delmastro, M. M. & Piganelli, J. D. Oxidative Stress and Redox Modulation Potential in Type 1 Diabetes. *Clinical and Developmental Immunology* **2011**, 1–15 (2011).
60. Oslowski, C. M. & Urano, F. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. *Methods Enzymol* **490**, 71–92 (2011).
61. Engin, F. *et al.* Restoration of the Unfolded Protein Response in Pancreatic β Cells Protects Mice Against Type 1 Diabetes. *Sci. Transl. Med.* **5**, (2013).
62. 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).
63. Farber, D. L., Yudanin, N. A. & Restifo, N. P. Human memory T cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol* **14**, 24–35 (2014).

64. Höllbacher, B. *et al.* Transcriptomic Profiling of Human Effector and Regulatory T Cell Subsets Identifies Predictive Population Signatures. *Immunohorizons* **4**, 585–596 (2020).
65. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* **10**, 4706 (2019).
66. 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).
67. Fuhrman, C. A. *et al.* Divergent Phenotypes of Human Regulatory T Cells Expressing the Receptors TIGIT and CD226. *J Immunol* **195**, 145–155 (2015).
68. Liu, M. *et al.* Proinsulin misfolding and diabetes: mutant *INS* gene-induced diabetes of youth. *Trends in Endocrinology & Metabolism* **21**, 652–659 (2010).
69. Støy, J. *et al.* In celebration of a century with insulin – Update of insulin gene mutations in diabetes. *Molecular Metabolism* **52**, 101280 (2021).
70. Ataie-Ashtiani, S. & Forbes, B. A Review of the Biosynthesis and Structural Implications of Insulin Gene Mutations Linked to Human Disease. *Cells* **12**, 1008 (2023).
71. Hayward, A. R., Shriber, M. & Sokol, R. Vitamin E supplementation reduces the incidence of diabetes but not insulitis in NOD mice. *J Lab Clin Med* **119**, 503–507 (1992).
72. Tse, H. M. *et al.* NADPH oxidase deficiency regulates Th lineage commitment and modulates autoimmunity. *J Immunol* **185**, 5247–5258 (2010).
73. Li, S. *et al.* Prevention of Autoimmune Diabetes in NOD Mice by Dimethyl Fumarate. *Antioxidants (Basel)* **10**, 193 (2021).
74. Lee, H. *et al.* Beta Cell Dedifferentiation Induced by IRE1α Deletion Prevents Type 1 Diabetes. *Cell Metab* **31**, 822-836.e5 (2020).
75. Argaev Frenkel, L., Rozenfeld, H., Rozenberg, K., Sampson, S. R. & Rosenzweig, T. N-Acetyl-l-Cysteine Supplement in Early Life or Adulthood Reduces Progression of Diabetes in Nonobese Diabetic Mice. *Curr Dev Nutr* **3**, nzy097 (2019).
76. Clement, C. C. *et al.* PDIA3 epitope-driven immune autoreactivity contributes to hepatic damage in type 2 diabetes. *Science Immunology* **7**, eabl3795 (2022).
77. 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).
78. Zhai, Y. *et al.* Cysteine carboxyethylation generates neoantigens to induce HLA-restricted autoimmunity. *Science* **379**, eabg2482 (2023).
79. Mori, S. *et al.* Neoself-antigens are the primary target for autoreactive T cells in human lupus. *Cell* **187**, 6071-6087.e20 (2024).
80. Kleverov, M. *et al.* Phantasus, a web application for visual and interactive gene expression analysis. *eLife* **13**, e85722 (2024).
81. Nanaware, P. P. *et al.* Role of the afferent lymph as an immunological conduit to analyze tissue antigenic and inflammatory load. *Cell Reports* **43**, 114311 (2024).
82. Van Gassen, S. *et al.* FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636–645 (2015).
83. Quintelier, K. *et al.* Analyzing high-dimensional cytometry data using FlowSOM. *Nat Protoc* **16**, 3775–3801 (2021).
84. 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 extend our gratitude to Kodi S. Ravichandran and Jonathan Kipnis for their insightful guidance throughout this study and for their critical review of the manuscript. We thank Jennifer Ponce, Christelle Schatz, and Jinsheng Yu for their expertise in performing scRNA-seq analysis, as well as Dorjan Brinja and Erica 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. This work is dedicated to the memory of Emil R. Unanue, who inspired our exploration of post-translationally modified insulin peptides generated in crinophagic granules. This work is 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 Institute of Health (R01AI62591 and R01DK134437 to X.W.).

Author Contribution

N.S. designed and conducted in vivo immunoassays and T-cell biology experiments. A.N.V. carried out antigen presentation experiments and isolated immunopeptidomes. O.J.P. prepared pancreatic islets and assisted in multiple experimental procedures. O.A. contributed to the in vivo evaluation of T-cell phenotypes. 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^{g7}- and DQ8-based tetramers. C.C.C. and L.S. designed and supplied samples for the lymph peptidome experiments. A.V.J. provided critical advice on TCR analysis. S.S. and L.C. offered crucial advice on the human spectral flow cytometry analysis. A.M.A. organized human subject recruitment and conducted clinical studies. C.F.L. performed mass spectrometry 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.

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 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-CD3e (145-2C11), anti-CD4 (RM4-5), anti-CD8a (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), all from Biolegend. Anti-CD11a (M17/4) was purchased from eBioscience. 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), all from Biolegend. Anti-CD4 (SK3), anti-CD45RA (HI100), anti-KLRG1 (13F12F2), and anti-CD226 (11A8.7.4) were purchased from eBioscience. 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 Bioscience.

Human PBMC immunopeptidome analysis. All human protocols have been reviewed and approved by the Institutional Review Boards of Washington University School of Medicine in St. Louis. After fasting for at least 8 hours, 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 (maximum 360 mL). Additional blood draws were performed at 90 and 120 minutes post MMTT. C-peptide levels were measured 10 minutes before and immediately before MMTT, as well as 15, 30, 60, 90, and 120 minutes post MMTT. All subjects had their glucose measured upon arrival at the study site; ND controls with glucose levels > 126 mg/dL and T1D patients with glucose levels > 250 mg/dL did not continue the study. T1D subjects also avoided doses of rapid-acting insulin 8 hours before testing unless needed to correct significant hyperglycemia.

For each participant (ND, 3mos, 18mos, n=10 per group), three blood samples (time 0, 90, 120 minutes) 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.

For sequential isolation of the HLA-DQ and HLA-DR peptidomes, the frozen cell pellets from each time point of each patient were thawed on the isolation day, pooled based on their conditions (ND, 3mos, 18mos), and suspended in fresh lysis buffer (40 mM MEGA 8 (MilliporeSigma, Cat #: O3129), 40 mM MEGA 9 (MilliporeSigma, Cat #: N1138), 1 mM phenylmethylsulfonyl fluoride (MilliporeSigma, Cat #: P7626), 0.2 mM iodoacetamide (MilliporeSigma, Cat #: I6125), 20 µg/mL leupeptin (MilliporeSigma, Cat #: L2884), and Roche complete Mini Protease cocktail (Roche Diagnostics, Cat #: 11836153001) in PBS). There were a total of 9 samples, each containing 35 to 51 million PBMCs. The lysate suspensions were rocked for 2 hours at 4 °C and then centrifuged at 20,000xg for 30 minutes at 4 °C. To eliminate non-specific binding of peptides, the supernatant was first incubated with polyclonal mouse IgG (Leinco, Cat #: N229; 1.5 mg antibody per sample) bound to Sepharose 4B (MilliporeSigma, Cat #: C9142) at 4 °C for 30 minutes. The unbound fraction containing peptide-MHC-II complexes was collected and added to a tube containing PBS-washed Sepharose conjugated to the anti-pan-HLA-DQ antibody (Leinco, Cat # H262, 3.0 mg per sample). This mixture was incubated at 4 °C overnight. The HLA-DQ Sepharose was applied to a column and the flow-through portion was collected and incubated with the anti-pan-HLA-DR antibody (Leinco, Cat # H261, 3.0 mg per sample) conjugated with Sepharose at 4 °C overnight. Either HLA-DQ or HLA-DR Sepharose was washed four times as follows: 10 mL 150 mM NaCl and 20 mM Tris (pH 7.4), 10 mL 400 mM NaCl and 20 mM Tris (pH 7.4), 10 mL 150 mM NaCl and 20 mM Tris (pH 7.4), and 10 mL 20 mM Tris (pH 8.0). Peptides were eluted with 10% acetic acid (ThermoFisher, Cat #: A38SI-212) and dried using a SpeedVac. Eluted peptides were resuspended and passed through detergent removal spin columns (Pierce, Cat #: 87777) to remove traces of remaining detergent, and then cleaned using C18 Spin Columns (Pierce, Cat #: 89870).

Generation of CD4$^+$ T-cell hybridomas. NOD mice (5-8-week-old) were immunized with the B9-23(C>S) 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 B9-23(C>S) peptide for 3 days and fused with the BW5147 fusion partner following standard protocols. The growth-positive T cell clones were screened against the B9-23(C>S) peptide using an antigen presentation assay. The antigen-responsive clones were further expanded and subcloned to generate monoclonal T-cell hybridomas.

Competition binding assay. For the cell-based binding assay, C3.g7 cells (10^4) were mixed with competitor peptide and incubated for 30 minutes at 37°C. The HEL:11-25 peptide (1 µM) was added to the C3.g7/competitor peptide mixture and incubated for an additional 60 minutes at 37°C. Cells were

centrifuged and washed. A $CD4^+$ T cell hybridoma specific for the HEL:11-25 peptide (clone 10E11) was added to the washed and loaded C3.g7 cells and incubated overnight at 37°C. The supernatants were then collected, and the production of IL-2 was measured by ELISA. IC50 values for competitor peptides were generated by fitting the data using GraphPad Prism software as a four-parameter inhibitor vs. response curve. For the biochemical binding assay, peptides were diluted in PBS and then added to the reaction mix to provide a range between 5-1000 pmoles peptide/sample. For each sample, 0.1 U thrombin, 1 µg of soluble g7-clip, and 0.125 pmoles of 125I-labeled g7-MIME peptide were added to the diluted unlabeled peptides. The reaction was allowed to sit at room temperature overnight. BioGel P6 columns (BioRad, Cat # 732-6299) were used to separate unbound, labeled peptide, and the fraction of displaced labeled peptide was quantified with a gamma counter (Perkin Elmer, Wallac 1272 CliniGamma). Binding curves were determined by plotting the amount of displaced labeled peptide vs the concentration of unlabeled competitor. The binding data is reported as the concentration of unlabeled peptide required to reduce g7-MIME binding by 50%.

Isolation of pancreatic islets. To isolate pancreatic islets, the peritoneal cavity was opened to expose the common bile duct. Using a dissection microscope, the bile duct leading to the duodenum was clamped. A solution of type XI collagenase (0.4 mg/mL; 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 then carefully removed and digested at 37°C for 12–14 minutes. 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 minutes at 37°C.

ELISPOT. For evaluating T cell responses by immunization, mice were immunized with the indicated peptides (10 nmol) emulsified with CFA at the footpad. On day 7, the popliteal lymph node cells were restimulated with the indicated peptides (10 µM) for 24 hours, and IFNγ production was assessed by ELISPOT. To evaluate T cell responses in islets and pancreatic lymph nodes without immunization, total islet and pLN cells were harvested from 8-week-old female NOD mice, pooled, and cultured with a mixture of equal amounts of B9-23 and B9-23(C>S) peptide (1 µM), in the presence of IL-2 (20 U/mL) for 7 days. Total live cells were collected from the culture using Histopaque 1119. The live cells were added to a culture containing 20 U/mL IL-2 and 1 µM peptide as before, along with 2×10^6 irradiated NOD splenocytes (3000 RAD) per mL of culture, and incubated for 3 days at 37°C. In all experiments, cells were harvested and assayed for reactivity by eliciting a recall response on IFNγ-coated 96-well multi-screen plates (Merck

Millipore, Cat #: S2EM004M99) for ELISPOT. The reactive cells were counted using Immunospot Software (C.T.L.).

Generation of primary T-cell lines and adoptive transfer. NOD mice were immunized with either B9-23(C>S) or scrambled peptide emulsified with CFA. Primary T cell lines were generated by culturing popliteal lymph node cells with 1 μM of either peptide in the presence of 20 U/mL of IL-2 and 2×10^6/mL of irradiated splenocytes. After one week of culture, total live cells were collected using Histopaque-1119. The cells underwent four cycles of expansion with B12-20(C>S) or scrambled control peptide for four weeks, each week supplemented with 0.1 μM peptide, 20 U/mL of IL-2 and 2×10^6/mL of irradiated splenocytes. For $CD8^+$ T cell co-transfer, polyclonal $CD8^+$ T cells were purified using anti-mouse CD8 isolation magnetic beads (Miltenyi Biotec). A total of 1×10^6 $CD4^+$ PTLs or $CD8^+$ T cells were transferred intravenously into NOD.Rag1$^{-/-}$ recipients. Urine glucose levels in mice were checked weekly (AimStrip US-G; Germaine Laboratories) and those with a level ⩾250 mg/dL for two consecutive readings were considered diabetic.

I-A^{g7} tetramer staining using magnetic enrichment. The generation of I-A^{g7}-based monomers has been previously described[42]. 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 hour 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 for 1 hour 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 of anti-APC and anti-PE microbeads (Miltenyi Biotec, Cat #: 130-090-855 and 130-048-801) in 100 μl of MACS buffer at 4°C for 20 minutes. After washing, the cells were resuspended in 1 mL of MACS buffer and applied to LS columns (Miltenyi Biotec, Cat #: 130-042-401) for enriching tetramer-bound T cells. Both 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 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).

Cell sorting and library preparation for single-cell RNA sequencing analysis. For scRNA-seq experiments, spleen and lymph node cells were pooled from 8-week-old female NOD mice (n=6-10). After red blood cell lysis, cells were washed with MACS buffer, and the $CD19^+$ 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 then stained with fluorochrome-conjugated tetramers followed by magnetic enrichment using anti-PE and anti-APC beads. Surface staining was performed as described above and tetramer-positive and polyclonal cells were sorted using the BD FACS Aria II sorter (BD Biosciences).

cDNA was prepared after the GEM generation and barcoding, followed by the GEM-RT reaction and bead cleanup steps. Purified cDNA was amplified for 11–16 cycles before being cleaned up using SPRI Select beads. Samples were then run on a Bioanalyzer to determine the cDNA concentration. V(D)J target enrichment (TCR) was performed on the full-length cDNA. 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) with Feature Barcoding technology for Cell Surface Protein and Immune Receptor Mapping, with appropriate modifications to the PCR cycles based on the calculated cDNA concentration. For sample preparation on the 10× Genomics platform, the following kits were used: Chromium Next GEM Single Cell 5′ Kit v2, 16 rxns (Cat #: PN-1000263), Chromium Next GEMChip 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).The concentration of each library was accurately determined through qPCR using the KAPA library Quantification Kit according to the manufacturer's protocol (KAPA Biosystems/Roche) to achieve the desired cluster counts for the Illumina NovaSeq6000 instrument. Normalized libraries were sequenced on a NovaSeq6000 S4 Flow Cell using the XP workflow and a 151×10×10×151 sequencing recipe according to the manufacturer's protocol. A median sequencing depth of 50,000 reads/cell was targeted for each Gene Expression library, and 5000 reads/cell for each V(D)J and Feature library.

Single-cell RNA sequencing 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 were 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 employed to quantify each feature in each cell. The expression levels of cell surface proteins were measured via an antibody and antigen-multimer staining assay using TotalSeq-C. 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: the log2 count of target $\geqslant 10$, while the other log2 counts $\leqslant 6$. 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 at an FDR < 0.05. GSEA was performed using Phantasus[80] (https://artyomovlab.wustl.edu/phantasus/).

β-cell granule isolation. Human islets from de-identified, non-diabetic donors were purchased from Prodo Laboratory. 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 then washed and resuspended in 1.0 mL PBS. Cell lysis was accomplished by passing islet cells through a Cell Homogenizer (Isobiotec) 5 times using the 10 μm clearance ball bearing to shear the cells. Cell lysate was spun for 10 min at 1,000 g, 4°C to pellet cell debris. The supernatant was transferred to a new tube followed by a repeat of the 1000 g spin. Pellets from both spins were discarded, and the supernatants were 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, and the supernatant was transferred to a clean tube. Pellets from both 5,000 g spins were combined into one tube (as the crinosome fraction) and suspended in media. The supernatant was spun for 20 min at 15,000 g, and the pellet was retained. The supernatant was transferred to a new tube and was spun for 30 min at 25,000 g, 4°C. The pellet (as the DCG fraction) was suspended in media. Acid phosphatase activity in the granule contents was assessed using an ELISA-based acid phosphatase assay kit (MilliporeSigma, Cat #: CS0740). Total protein levels in the granule contents were measured by a Micro BCA protein assay kit (Pierce, Cat #: 23225).

Granule-based antigen presentation (GAP) assay. Subcellular fractions isolated by 5,000 g (5k), 15,000 (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/well) in a 96-well culture plate. After incubation for 2 hours, the $CD4^+$ T cell hybridomas (5×10^4/well) were added. After an overnight incubation, T-cell responses were assessed by measuring IL-2 production in the culture supernatants using ELISA.

β-cell and DC treatment. For experiments involving reducing agents, 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 hours. For experiments involving ER stressors, islet cell suspensions were treated for 2 hours with either 1 μM thapsigargin (Sigma, Cat# T9033) or 1 μg/mL tunicamycin (Sigma, Cat# T7765). In experiments involving reducers and ER stressors, islet cell suspensions were treated with 100 μM NAC, glutathione, or TUDCA for 16 hours followed by addition of thapsigargin or tunicamycin for 2 hours. In all experiments, cells were washed extensively before granule isolation and presentation.

For treating human islets with cytokines, whole human islets received from Prodo Laboratories were cultured in CMRL 1066 media with 10% FBS, 1% Na Pyruvate, 1% Glutamine, 1% NEAA, and Pen/Strep. Islet cells were dispersed, collected by centrifugation, washed with media, and counted. Cells were split into 5 equal aliquots (~10^6 each) treated with vehicle, 50 ng/mL human TNFα (PeproTech, Cat #300-01A), 50 ng/mL human IL-1β (PeproTech, Cat #200-01B), 50 ng/mL human IFNγ (PeproTech, Cat #300-02), or all three cytokines combined. Cells were incubated for 16 hours at 37°C, washed twice with PBS, and resuspended for granule isolation.

For mouse DC 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). In microfuge tubes, $CD11c^+$ DCs (~$1.5x10^6$ cells) were pulsed with 25 μM B9-23 peptide for 1 hour, and then were left untreated or treated with the following mouse cytokines (100 ng/mL): TNFα (PeproTech, Cat # 315-01A), IL-1β (PeproTech, Cat # 211-11B), IFNγ (PeproTech, Cat # 315-05), IL-4 (PeproTech, Cat # 214-14), IL-6 (PeproTech, Cat # 216-16), and IL-12 (PeproTech, Cat # 210-12) for 16 hours. DCs were collected, washed with PBS, and used for antigen presentation assay.

Soluble crinosome and lymph peptidomes and DC MHC-II peptidome. 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[81]. 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 (ThermoFisher, Cat #: A118-500)/95% acetonitrile (Burdick & Jackson, #015-4), and then dried with a Speed-Vac.

For mouse DC MHC-II immunopeptidome analysis, mouse splenic $CD11c^+$ DCs were pulsed with 25 μM B9-23 peptide for 1 hour, followed by three different treatments: overnight culture without further additions,

addition of 100 ng/mL TNFα for overnight culture, or 100 μM N-acetyl-cysteine (NAC) for 1 hour followed by TNFα for overnight culture in the presence of peptide/NAC. Cells were then collected, washed, and lysed for peptidome processing. The Mouse DC MHC-II peptidomes were isolated by a procedure similar to the human PBMC peptidomes, except that the anti-I-A^{g7} antibody conjugated to Sepharose was used.

Mass Spectrometry. A Dionex UltiMate 3000 system (Thermo Scientific) was coupled to an Orbitrap Fusion Lumos (Thermo Scientific) through an EASY-Spray ion source (Thermo Scientific). Peptide samples were reconstituted in 2% acetonitrile (ACN)/0.1% formic acid and loaded (15 μL/minutes for 3 minutes) onto a trap column (100 μm × 2 cm, 5 μm Acclaim PepMap 100 C18, 50 °C), then eluted (200 nL/minutes) onto an EASY-Spray PepMap RSLC C18 column (2 μm, 50 cm × 75 μm ID, 50 °C, Thermo Scientific) and separated with the following gradient, all % Buffer B (0.1% formic acid in ACN): 0–110 minutes, 2–22%; 110–120 minutes, 22–35%; 120–130 minutes, 35–95%; 130–150 minutes, isocratic at 95%; 150–151 minutes, 95–2%, 151–171 minutes, isocratic at 2%. Spray voltage was 1,600-1,800 V, ion transfer tube temperature was 275 °C, and RF lens was 30%. Mass spectrometry scans were acquired in profile mode and MS/MS scans in centroid mode, for ions with charge states 2–5, with a cycle time of 1.5 s. For HCD, mass spectra were recorded from 375–1500 Da at 120K resolution (at m/z = 200), and MS/MS was triggered above a threshold of 2.0×10^4, with quadrupole isolation (0.7 Da) at 15K resolution, and collision energy of 30%. Dynamic exclusion was used (60 s).

Mass spectrometry data analysis. Processing and database searching were performed with PEAKS Studio Xpro (version 10.6, build 20201221, Bioinformatics Solutions Inc.). For crinosomes from MIN6 cells, files were searched against a Uniprot-Mouse database (Jan. 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 C>S 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 C>S in MIN6 cells and for crinosomes from mouse islets, database searching was performed with a Uniprot-Mouse database appended with C>S forms of β-cell proteins insulin 1, insulin 2, chromogranin-A, islet amyloid polypeptide, and secretogranin using the same parameters as above. For crinosomes from human islets,

searches were performed against a Uniprot-Human database appended with C>S forms of β-cell proteins insulin, chromogranin-A, islet amyloid polypeptide, 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 by peptide sequence and 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 (B12–20 or B12–20(C>S)), were produced in-house[43]. 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 for 1 hour 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/100 μL for 1 hour at room temperature. After incubation, the cells were washed twice with MACS buffer, then incubated with 20 μL of anti-APC and anti-PE microbeads in 100 μL of MACS buffer at 4°C for 20 minutes. The cells were then washed twice with FACS buffer (1% BSA in 1X PBS) and used for flow cytometry analysis.

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.

High-dimensional analysis of spectral flow cytometry data was performed using DownSample V3, FlowSOM, and ClusterExplorer software in FlowJo version 10.9[82–84]. Native- and C>S-specific tetramer-binding T cells were gated as live, singlet, $CD19^{-}CD11b^{-}CD3^{+}CD4^{+}$ cells. Polyclonal $CD4^{+}$ T cells (live, singlet, $CD19^{-}CD14^{-}CD3^{+}CD4^{+}$tetramer-negative) were downsampled to match the number of tetramer-positive $CD4^{+}$ T cells. The FCS files for Native 12-20^{+}, C>S^{+}, and polyclonal $CD4^{+}$ T cells from individual donors were concatenated. The combined file was uploaded into FlowJo, and Uniform Manifold Approximation and Projection (UMAP) was performed on a total of 1,179 cells. The following markers were used to generate UMAP plots: CD45RO, CD45RA, CCR, CD38, CD11a, CD95, CD27, CD25, CD49d, CCR4, KLRG1, CCR6, CXCR3, HLA-DR, PD1, CD127, CD28, CD226, and CXCR5. UMAP 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.

Figure 1

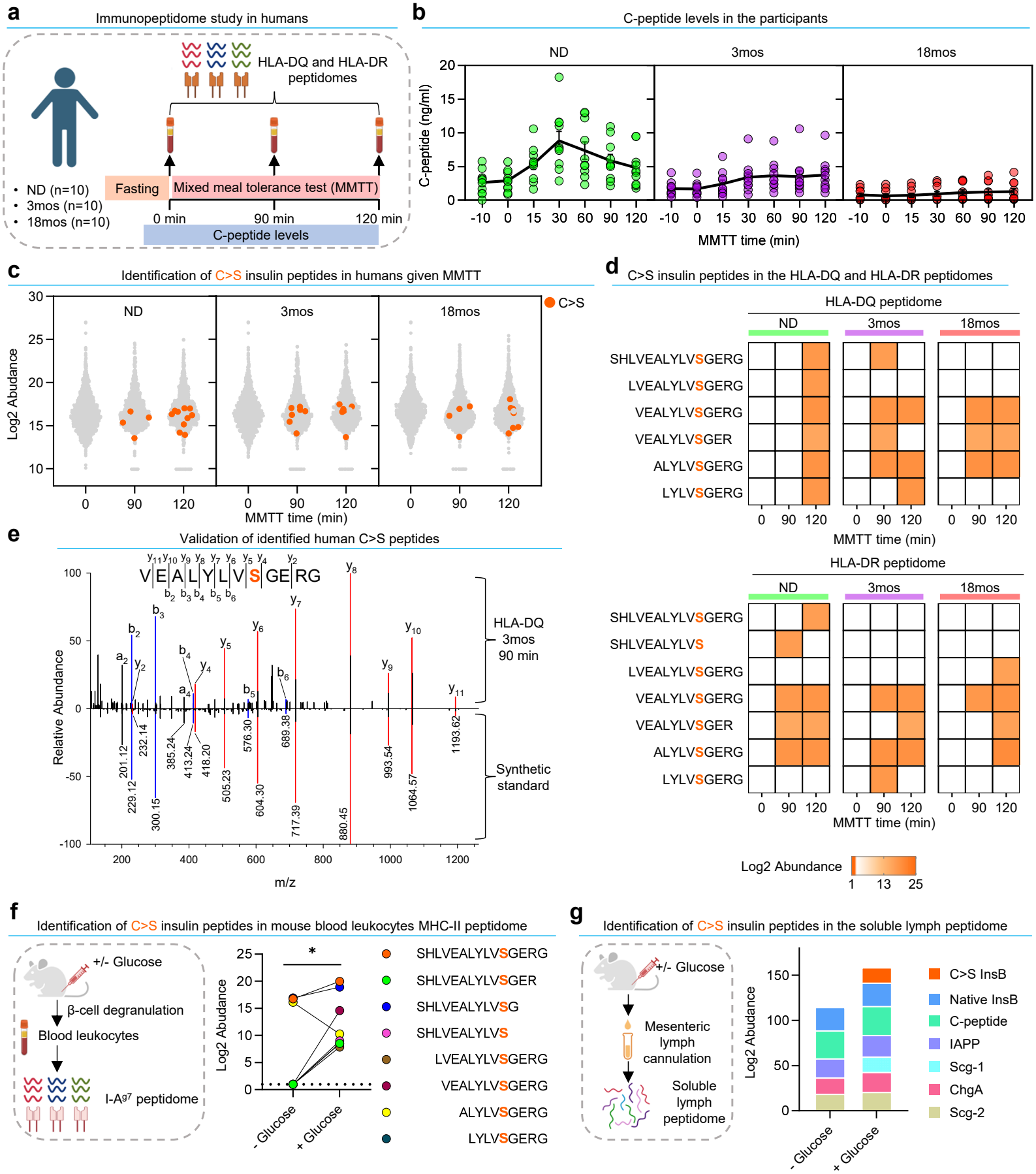
a
Immunopeptidome study in humans
HLA-DQ and HLA-DR peptidomes
ND (n=10)
3mos (n=10)
18mos (n=10)
Fasting
Mixed meal tolerance test (MMTT)
0 min
90 min
120 min
C-peptide levels
b
C-peptide levels in the participants
ND
3mos
18mos
C-peptide (ng/ml)
MMTT time (min)
c
Identification of C>S insulin peptides in humans given MMTT
C>S
Log2 Abudance
MMTT time (min)
d
C>S insulin peptides in the HLA-DQ and HLA-DR peptidomes
HLA-DQ peptidome
HLA-DR peptidome
SHLVEALYLVSGERG
LVEALYLVSGERG
VEALYLVSGERG
VEALYLVSGER
ALYLVSGERG
LYLVSGERG
SHLVEALYLVS
Log2 Abundance
e
Validation of identified human C>S peptides
VEALYLVSGERG
HLA-DQ 3mos 90 min
Synthetic standard
Relative Abundance
m/z
f
Identification of C>S insulin peptides in mouse blood leukocytes MHC-II peptidome
+/- Glucose
β-cell degranulation
Blood leukocytes
I-Ag7 peptide
Log2 Abudance
- Glucose
+ Glucose
SHLVEALYLVSGERG
SHLVEALYLVSGER
SHLVEALYLVSG
SHLVEALYLVS
LVEALYLVSGERG
VEALYLVSGERG
ALYLVSGERG
LYLVSGERG
g
Identification of C>S insulin peptides in the soluble lymph peptidome
+/- Glucose
Mesenteric lymph cannulation
Soluble lymph peptidome
Log2 Abudance
C>S InsB
Native InsB
C-peptide
IAPP
Scg-1
ChgA
Scg-2

Figure 1. Identification of insulin peptides harboring a cysteine-to-serine conversion in mice and humans.

- **a.** Schematic for analyzing human HLA-DQ and HLA-DR peptidomes of PBMCs from three cohorts of participants before and after a mixed meal tolerance test (MMTT).
- **b.** C-peptide levels of the three human cohorts at indicated time points before and after MMTT. Each point represents an individual participant.
- **c.** Identification of C>S insulin peptides in the human PBMC peptidomes. Grey dots represent all peptides combined from HLA-DQ and HLA-DR peptidomes at the indicated MMTT time points, with C>S (orange) insulin peptides indicated.
- **d.** Heatmaps showing each C>S insulin peptide identified in the HLA-DQ (upper) and HLA-DR (lower) peptidomes across the three human cohorts at indicated time points of MMTT.
- **e.** A mirror plot showing the complete match of a representative C>S insulin peptide identified in the human HLA-DQ peptidome to its synthetic standard.
- **f.** Schematic (left) and quantification (right) of insulin B-chain peptides with C>S (each point) identified in the blood leukocyte MHC-II peptidome of NOD mice, showing increased abundance after glucose challenge. *$P < 0.05$; Wilcoxon signed-rank test.
- **g.** Schematic (left) for analyzing mouse lymph peptidome and composition (right) of indicated β-cell-derived peptides identified in the mouse lymph peptidomes. The data are based on the cumulative abundance of all the peptides belonging to a given β-cell protein.

Figure 2

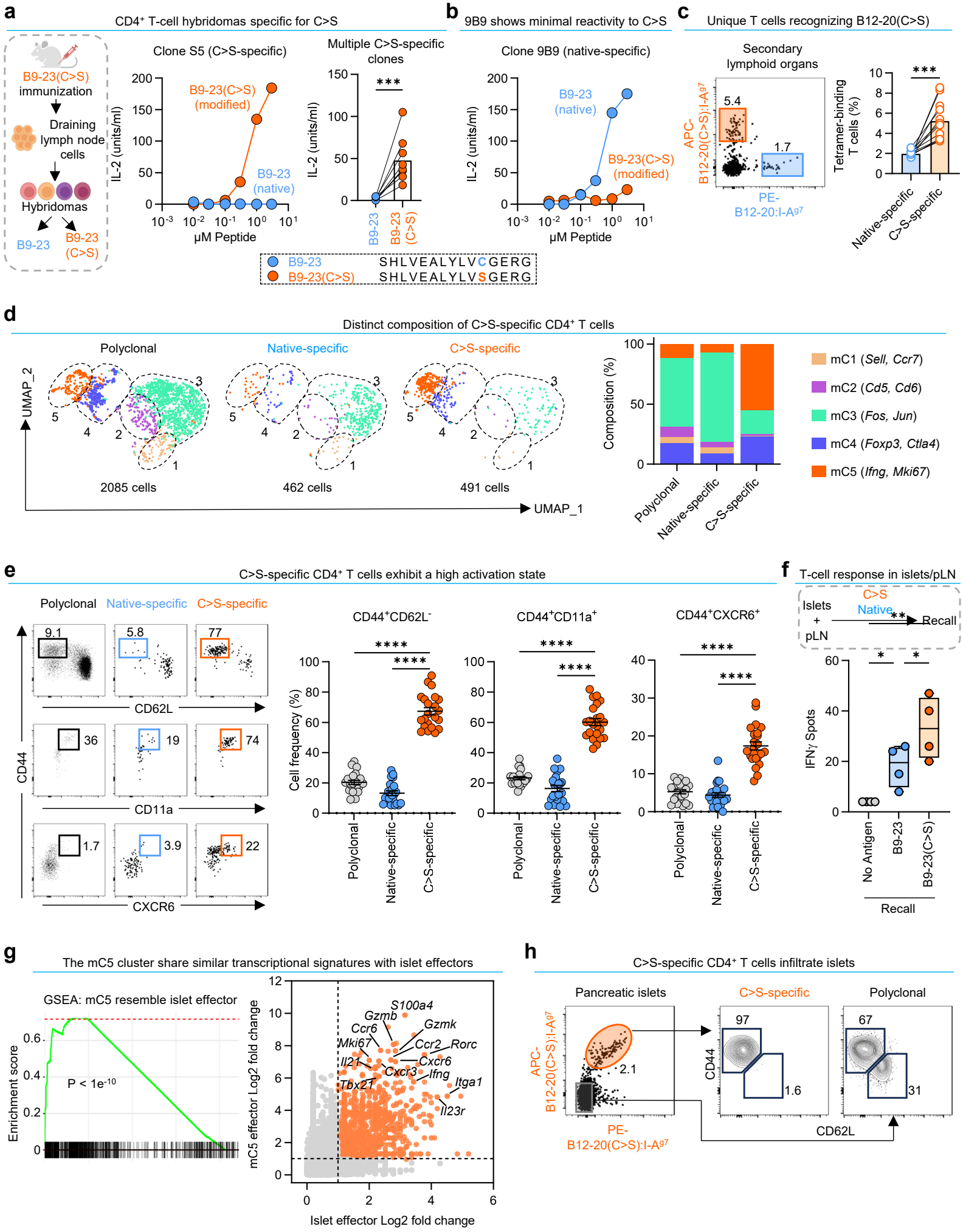
a
CD4+ T-cell hybridomas specific for C>S
B9-23(C>S) immunization
Draining lymph node cells
Hybridomas
B9-23
B9-23 (C>S)
Clone S5 (C>S-specific)
B9-23(C>S) (modified)
B9-23 (native)
IL-2 (units/ml)
μM Peptide
Multiple C>S-specific clones
***
B9-23
B9-23 (C>S)
b
9B9 shows minimal reactivity to C>S
Clone 9B9 (native-specific)
B9-23 (native)
B9-23(C>S) (modified)
B9-23 SHLVEALYLVCGERG
B9-23(C>S) SHLVEALYLVSGERG
c
Unique T cells recognizing B12-20(C>S)
Secondary lymphoid organs
5.4
1.7
APC-B12-20(C>S):I-Ag7
PE-B12-20:I-Ag7
Tetramer-binding T cells (%)
Native-specific
C>S-specific
d
Distinct composition of C>S-specific CD4+ T cells
Polyclonal
Native-specific
C>S-specific
UMAP_2
UMAP_1
2085 cells
462 cells
491 cells
Composition (%)
mC1 (Sell, Ccr7)
mC2 (Cd5, Cd6)
mC3 (Fos, Jun)
mC4 (Foxp3, Ctla4)
mC5 (Ifng, Mki67)
e
C>S-specific CD4+ T cells exhibit a high activation state
CD62L
CD11a
CXCR6
CD44
CD44+CD62L-
CD44+CD11a+
CD44+CXCR6+
Cell frequency (%)
****
f
T-cell response in islets/pLN
Islets + pLN
C>S
Native
Recall
IFNγ Spots
No Antigen
B9-23
B9-23(C>S)
Recall
g
The mC5 cluster share similar transcriptional signatures with islet effectors
GSEA: mC5 resemble islet effector
Enrichment score
P < 1e-10
mC5 effector Log2 fold change
Islet effector Log2 fold change
S100a4
Gzmb
Gzmk
Ccr6
Mki67
Ccr2
Rorc
Il21
Cxcr6
Cxcr3
Ifng
Tbx21
Itga1
Il23r
h
C>S-specific CD4+ T cells infiltrate islets
Pancreatic islets
C>S-specific
Polyclonal
APC-B12-20(C>S):I-Ag7
PE-B12-20(C>S):I-Ag7
2.1
97
1.6
67
31
CD44
CD62L

Figure 2. C>S is recognized by register-specific CD4$^+$ T cells that invade pancreatic islets.

- **a.** Reactivity of CD4$^+$ T-cell hybridomas specifically recognizing B9-23(C>S) but not native B9-23. Data show responses of the S5 T cell (left) to serially diluted peptide concentrations, with quantification (right) of multiple individual T-cell hybridomas (each point) responding to 0.3 μM peptide.
- **b.** Reactivity of the 9B9 T cell to the native B9-23 but not the B9-23(C>S) peptide.
- **c.** Flow cytometry analysis (left) showing co-staining with B12-20:I-A^{g7} and B12-20(C>S):I-A^{g7} tetramers revealing two distinct register-specific T-cell populations. The quantification (right) summarizes the frequencies of the tetramer-binding populations in individual mice (each point) from six independent experiments. The experiments used magnetic enrichment in SLOs of 8-week-old female NOD mice; gating strategy shown in Supplementary Figure 2f. ***P < 0.001; two-tailed paired t test.
- **d.** UMAP plots (left) showing the distribution of indicated T-cell clusters in polyclonal, native-specific, and C>S-specific CD4$^+$ T cells. The bar chart (right) shows the relative frequency of each cluster.
- **e.** Flow cytometry analysis (left) of indicated T-cell activation markers in polyclonal, native-specific, and C>S-specific CD4$^+$ T cells in SLOs of 8-week-old female NOD mice. The quantification (mean ± SEM) summarizes cell frequencies in individual mice (each point) from 7-10 independent experiments. ****P < 0.0001; one-way ANOVA with Dunnett's multiple comparisons test.
- **f.** ELISPOT assay for T-cell responses (IFNγ production) in islets and pLNs during antigen recall with the B9-23 or B9-23(C>S) peptide. Data (mean ± SEM) summarize results from two independent experiments; each point represent a biological replicate including 3-4 mice. *P<0.05; **P<0.01; repeated measures one-way ANOVA with Tukey's multiple comparisons test.
- **g.** GSEA (left) showing significant transcriptional profile enrichment of mC5 with islet terminal effector CD4$^+$ T cells (GEO: GSE262101). The correlation plot (right) depicts highly expressed effector T-cell genes shared between the two datasets.
- **h.** Flow cytometry analysis showing C>S-specific CD4$^+$ T cells in pancreatic islets of 8-week-old female NOD mice using dual B12-20(C>S):I-A^{g7} tetramer labeling. The tetramer-binding and the tetramer-negative polyclonal CD4$^+$ T cells (gating: live, singlets, CD45$^+$CD4$^+$CD8$^-$; without magnetic enrichment) were assessed for CD44 and CD62L expression. Data are representative of four independent experiments with similar results.

Figure 3

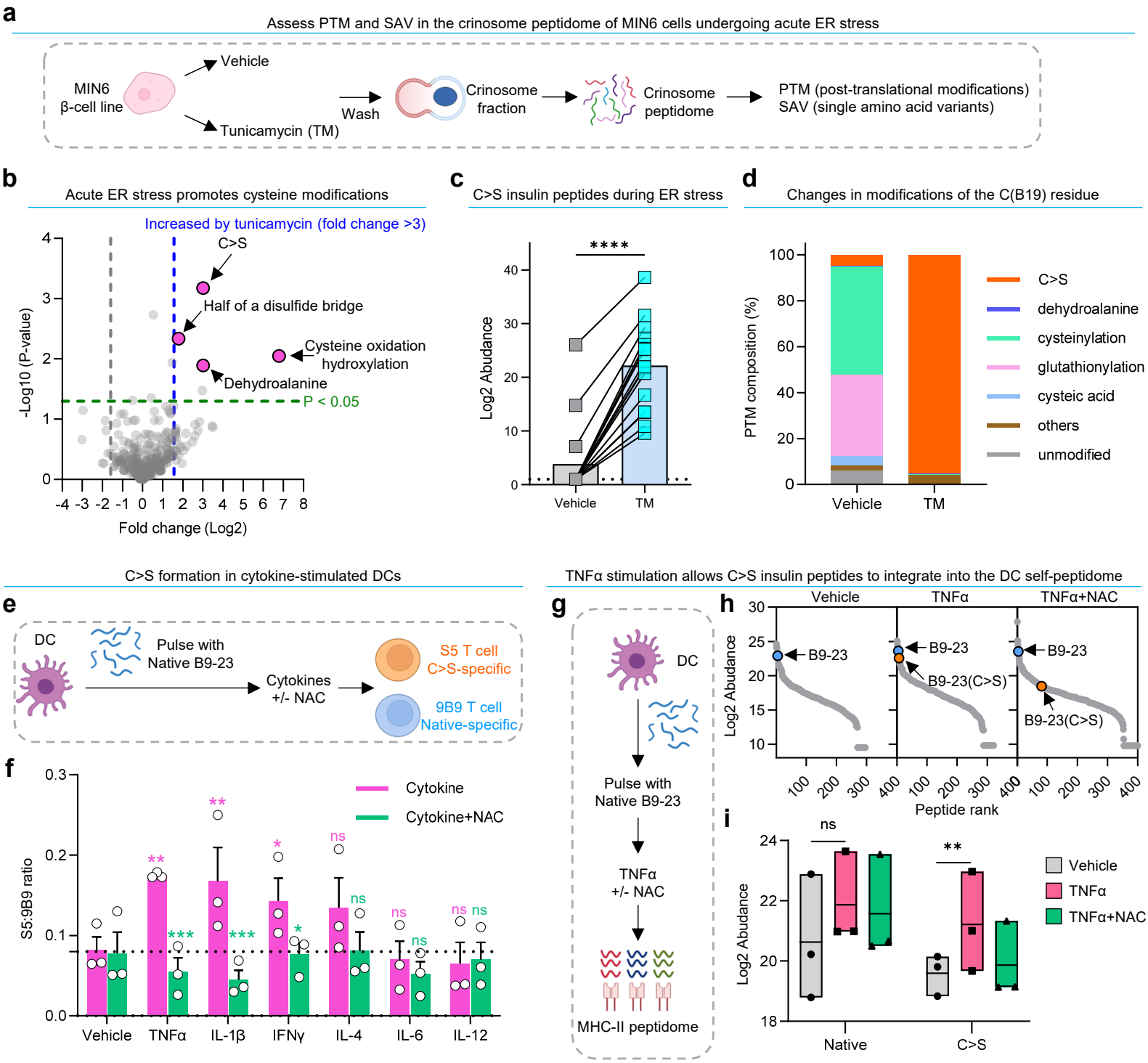
a
Assess PTM and SAV in the crinosome peptidome of MIN6 cells undergoing acute ER stress
MIN6
β-cell line
Vehicle
Tunicamycin (TM)
Wash
Crinosome
fraction
Crinosome
peptidome
PTM (post-translational modifications)
SAV (single amino acid variants)
b
Acute ER stress promotes cysteine modifications
Increased by tunicamycin (fold change >3)
C>S
Half of a disulfide bridge
Cysteine oxidation
hydroxylation
Dehydroalanine
P < 0.05
-Log10 (P-value)
Fold change (Log2)
c
C>S insulin peptides during ER stress
****
Log2 Abudance
Vehicle
TM
d
Changes in modifications of the C(B19) residue
PTM composition (%)
C>S
dehydroalanine
cysteinylation
glutathionylation
cysteic acid
others
unmodified
Vehicle
TM
C>S formation in cytokine-stimulated DCs
e
DC
Pulse with
Native B9-23
Cytokines
+/- NAC
S5 T cell
C>S-specific
9B9 T cell
Native-specific
f
Cytokine
Cytokine+NAC
S5:9B9 ratio
Vehicle
TNFα
IL-1β
IFNγ
IL-4
IL-6
IL-12
TNFα stimulation allows C>S insulin peptides to integrate into the DC self-peptidome
g
DC
Pulse with
Native B9-23
TNFα
+/- NAC
MHC-II peptidome
h
Vehicle
TNFα
TNFα+NAC
B9-23
B9-23(C>S)
Log2 Abudance
Peptide rank
i
ns
**
Log2 Abudance
Vehicle
TNFα
TNFα+NAC
Native
C>S

Figure 3. Redox stress promotes C>S formation in both β cells and DCs.

- **a.** Schematic of the PTM and SAV search in crinosome peptidomes isolated from MIN6 β cells treated with vehicle or tunicamycin for 2 hours.
- **b.** A volcano plot showing changes in PTMs and SAVs identified in the crinosome peptidome of MIN6 cells given tunicamycin relative to vehicle treatment. Several cysteine modifications with significant increase ($P < 0.05$, fold change > 3) after treatment with tunicamycin were indicated. See a complete list in Supplementary Table 2.
- **c.** Abundance of individual insulin peptides with C>S (each point) identified in the crinosome peptidome of MIN6 cells given vehicle or tunicamycin treatment. See a complete list in Supplementary Figure 6a. ****$P < 0.0001$; Wilcoxon signed-rank test
- **d.** Composition of indicated modifications on the C(B19) residue of insulin peptides identified in the crinosome peptidome MIN6 cells given vehicle or tunicamycin treatment.
- **e.** Schematic of the antigen presentation assay to assess the presentation of native or C>S insulin peptides by DCs pulsed with native B9-23 and stimulated with different cytokines.
- **f.** C>S conversion from the native B9-23 peptide in DCs upon cytokine stimulation, shown as the ratio of S5 to 9B9 T-cell responses with or without NAC treatment. Data (mean ± SEM) summarize results from three independent experiments (each point). ns, not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; repeated measures two-way ANOVA with Šídák's multiple comparisons test. Pink stars: cytokine vs. vehicle; green stars: cytokine vs. cytokine $^{+}$ NAC.
- **g.** Schematic (left) of MHC-II immunopeptidome analysis of B9-23-pulsed DCs stimulated with TNFα with or without NAC.
- **h.** Rank plots showing all identified peptides in the DC MHC-II peptidomes (grey dots), ranked by abundance. Native B9-23 and modified B9-23(C>S) peptides are denoted.
- **i.** Abundance of native and C>S insulin peptides identified in the DC MHC-II immunopeptidome analyses. Each point represents accumulative abundance of all identified native or C>S insulin peptides from one of the three independent experiments. ns, not significant; **$P < 0.01$; repeated measures two-way ANOVA with Tukey's multiple comparisons test.

Figure 4

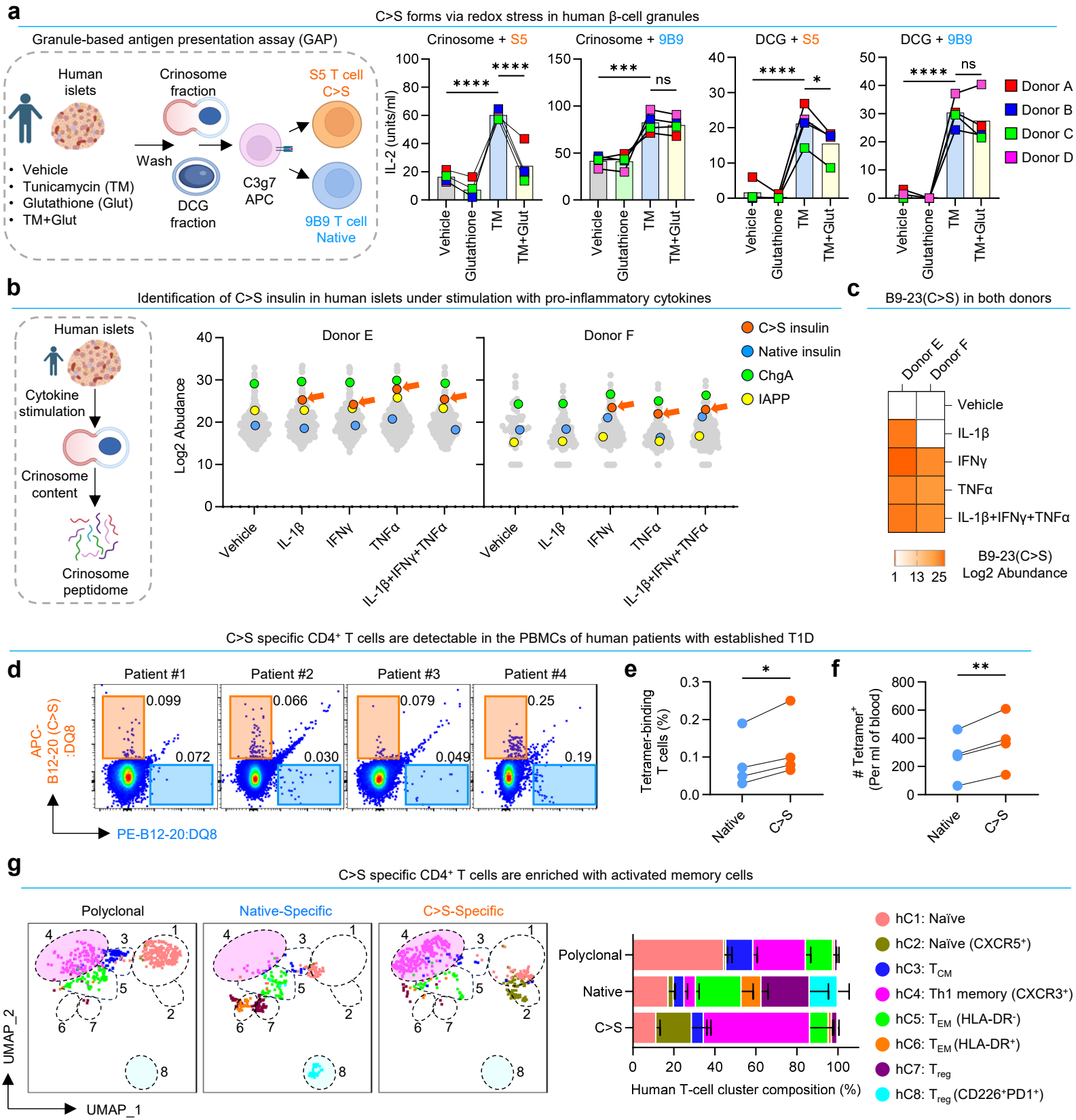
a
C>S forms via redox stress in human β-cell granules
Granule-based antigen presentation assay (GAP)
Human islets
Crinosome fraction
S5 T cell C>S
Wash
C3g7 APC
DCG fraction
9B9 T cell Native
Vehicle
Tunicamycin (TM)
Glutathione (Glut)
TM+Glut
Crinosome + S5
Crinosome + 9B9
DCG + S5
DCG + 9B9
IL-2 (units/ml)
Donor A
Donor B
Donor C
Donor D
b
Identification of C>S insulin in human islets under stimulation with pro-inflammatory cytokines
Human islets
Cytokine stimulation
Crinosome content
Crinosome peptidome
Donor E
Donor F
Log2 Abudance
C>S insulin
Native insulin
ChgA
IAPP
Vehicle
IL-1β
IFNγ
TNFα
IL-1β+IFNγ+TNFα
c
B9-23(C>S) in both donors
B9-23(C>S) Log2 Abundance
C>S specific CD4+ T cells are detectable in the PBMCs of human patients with established T1D
d
Patient #1
Patient #2
Patient #3
Patient #4
APC-B12-20 (C>S):DQ8
PE-B12-20:DQ8
e
Tetramer-binding T cells (%)
Native
C>S
f
Tetramer+ (Per ml of blood)
g
C>S specific CD4+ T cells are enriched with activated memory cells
Polyclonal
Native-Specific
C>S-Specific
UMAP_2
UMAP_1
Human T-cell cluster composition (%)
hC1: Naïve
hC2: Naïve (CXCR5+)
hC3: TCM
hC4: Th1 memory (CXCR3+)
hC5: TEM (HLA-DR-)
hC6: TEM (HLA-DR+)
hC7: Treg
hC8: Treg (CD226+PD1+)

Figure 4. C>S neoepitope generation in stressed human islets and C>S-specific CD4$^+$ T cells in patients with established T1D.

- **a.** Schematic (left) and quantification (right) of GAP assay showing formation of C>S in crinosomes and DCGs isolated from human islets following indicated treatments. Data (mean ± SEM) summarize results from four donors (each point) examined in four independent experiments. ns, not significant; *P < 0.05; ***P < 0.001; ****P < 0.0001; repeated measures one-way ANOVA with Šídák's multiple comparisons test.
- **b.** Schematic (left) and identification (right) of C>S insulin peptides in crinosome peptidomes of human islets exposed to indicated cytokines. Grey dots represent source proteins of all identified peptides, with native (blue) or C>S (orange) insulin, ChgA (green), and IAPP (yellow) indicated.
- **c.** A heatmap showing the B9-23(C>S) peptide identified in the human crinosome peptidomes.
- **d.** Spectral flow cytometry analysis showing HLA-DQ8 tetramer-binding CD4$^+$ T cells recognizing the B12-20(C>S) or B12-20 epitope in PBMCs of four individual patients with established T1D. The tetramer binding T cells were gated as Supplementary Figure 8c without magnetic enrichment.
- **e.** Quantification of the frequencies of B12-20(C>S):DQ8 and B12-20:DQ8 tetramer-binding CD4$^+$ T cells in individual patients (each point). *P < 0.05; two-tailed paired t test.
- **f.** Quantification of the cell numbers of B12-20(C>S):DQ8 and B12-20:DQ8 tetramer-binding CD4$^+$ T cells in individual patients (each point). **P < 0.01; two-tailed paired t test.
- **g.** UMAP plots (left) and quantification (right) showing the composition of the eight human T-cell clusters within polyclonal, native-specific, and C>S-specific T-cell populations.