CD8+ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes

Marie Eliane Azoury, Fatoumata Samassa, Mijke Buitinga, Laura Nigi, Noemi Brusco, Aïsha Callebaut, Matthieu Giraud, Magali Irla, Ana Ines Lalanne, Alexia Carré, et al.

CD8$^+$ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes

Marie Eliane Azoury[1*], Fatoumata Samassa[1*], Mijke Buitinga[2], Laura Nigi[3], Noemi Brusco[3], Aïsha Callebaut[2], Matthieu Giraud[4], Magali Irla[5], Ana Ines Lalanne[1], Alexia Carré[1], Georgia Afonso[1], Zhicheng Zhou[1], Barbara Brandao[1], Maikel L. Colli[6], Guido Sebastiani[3], Francesco Dotta[3], Maki Nakayama[7], Decio L. Eizirik[6,8], Sylvaine You[1], Sheena Pinto[9], Mark J. Mamula[10], Yann Verdier[11], Joelle Vinh[11], Soren Buus[12], Chantal Mathieu[2], Lut Overbergh[2], Roberto Mallone [1,13]

[1]Université de Paris, Institut Cochin, CNRS, INSERM, 75014 Paris, France.
[2]KU Leuven, Laboratory of Clinical and Experimental Endocrinology, 3000 Leuven, Belgium.
[3]University of Siena, Department of Medicine, Surgery and Neuroscience, Diabetes Unit and Fondazione Umberto di Mario ONLUS, Toscana Life Sciences, 53100 Siena, Italy.
[4]Université de Nantes, INSERM UMR1064, Centre de Recherche en Transplantation et Immunologie (CRTI), 44093 Nantes, France.
[5]Aix-Marseille University, CNRS, INSERM, CIML, Centre d'Immunologie de Marseille-Luminy, Marseille, France
[6]Université Libre de Bruxelles, Center for Diabetes Research and Welbio, Medical Faculty, 1070 Brussels, Belgium.
[7]Barbara Davis Center for Childhood Diabetes, University of Colorado School of Medicine, Aurora, CO 80045, USA.
[8]Indiana Biosciences Research Institute, Indianapolis, IN 46202, USA.
[9]Deutsches Krebsforschungszentrum (DKFZ), Division of Developmental Immunology, 69120 Heidelberg, Germany.
[10]Yale University School of Medicine, New Haven, CT 06510, USA.
[11]ESPCI Paris, PSL University, Spectrométrie de Masse Biologique et Protéomique, CNRS USR3149, 75005 Paris, France.
[12]Panum Institute, Department of International Health, Immunology and Microbiology, D-2200 Copenhagen, Denmark.
[13]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, 75014 Paris, France.

*M.E.A. and F.S. contributed equally to this work.

Corresponding author: Roberto Mallone, MD PhD - INSERM U1016 Cochin Institute, G.H. Cochin-Port-Royal, Bâtiment Cassini - 123, boulevard de Port-Royal – F-75014 Paris, France. Phone: +33-1-76.53.55.83 E-mail: roberto.mallone@inserm.fr.

Word count: 4,336 **Tables count**: 1 **Figures count**: 7

Key words: autoimmunity, citrullination, cross-reactivity, lymphocytes, neo-antigens, peptidyl-arginine deiminases, post-translational modifications, thymus.

Tweet (276 characters): Autoimmune CD8+ T cells of #type 1 diabetes do not always prefer citrullinated versions of GRP78. Their preference is shaped by thymic expression of citrullinating enzymes and bacterial cross-reactivity. Native antigens are sometimes more immunogenic than citrullinated forms. @InstitutCochin, @Inserm, @Innodiagroup.

ABSTRACT

In type 1 diabetes, autoimmune β-cell destruction may be favored by neo-antigens harboring post-translational modifications such as citrullination. We studied the recognition of native and citrullinated glucose-regulated protein (GRP)78 peptides by $CD8^+$ T cells. Citrullination modulated T-cell recognition and, to a lesser extent, HLA-A2 binding. GRP78-reactive $CD8^+$ T cells circulated at similar frequencies in type 1 diabetic and healthy donors and preferentially recognized either native or citrullinated versions, without cross-reactivity. Rather, the preference for native GRP78 epitopes was associated with $CD8^+$ T cells cross-reactive with bacterial mimotopes. In the pancreas, a dominant GRP78 peptide was instead preferentially recognized when citrullinated. To further clarify these recognition patterns, we considered the possibility of citrullination in the thymus. Citrullinating peptidyl-arginine deiminase (Padi) enzymes were expressed in murine and human medullary epithelial cells (mTECs), with citrullinated proteins detected in murine mTECs. However, Padi2 and Padi4 expression was diminished in mature mTECs from NOD mice versus C57BL/6 mice. We conclude that, on one hand, the $CD8^+$ T-cell preference for native GRP78 peptides may be shaped by cross-reactivity with bacterial mimotopes. On the other hand, post-translational modifications may not invariably favor loss of tolerance because thymic citrullination, although impaired in NOD mice, may drive deletion of citrulline-reactive T cells.

INTRODUCTION

Islet-reactive $CD8^+$ T cells are held as the final mediators of β-cell destruction (1). Surprisingly, they circulate at similar frequencies in most individuals, irrespective of type 1 diabetes status (2-4). This is true for known islet antigens and for novel ones, e.g. urocortin-3 (UCN3), secretogranin-5 and proconvertase-2 (3, 4). The frequency of most of these islet epitope-reactive T cells falls in a relatively narrow range of ~1-50/10^6 $CD8^+$ T cells (0.0001-0.005%), which overlaps with that of naïve $CD8^+$ T cells recognizing viral peptides in virus seronegative donors (5, 6). In agreement with this frequency overlap, circulating islet-reactive $CD8^+$ T cells are largely naïve in both healthy and type 1 diabetic donors (2-4), suggesting a limited recruitment in the autoimmune process. The fraction engaged in this process seems instead sequestered in the pancreas, where the same T cells are enriched in type 1 diabetic patients and display a $CD45RO^+$ phenotype, compatible with prior antigen encounter (2, 3). Thus, a universal state of 'benign' islet autoimmunity exists for circulating $CD8^+$ T cells, to a much larger extent than previously appreciated (1, 7). In line with recent reports (6), we demonstrated that this benign autoimmunity is imprinted in the thymus due to a marginal deletion of autoreactive $CD8^+$ T cells, since the presence or absence of ectopic islet antigen presentation in thymic medullary epithelial cells (mTECs) does not significantly modulate this process (2, 3).

It is unknown whether this novel paradigm of benign autoimmunity applies to β-cell antigens undergoing post-translational modifications (PTMs). PTMs occurring in peripheral tissues, especially under inflammatory conditions, may favor priming of T-cell precursors that have escaped thymic deletion, due to their selection against native peptides with different physico-chemical properties (8). One post-translationally modified antigen is glucose-regulated protein (GRP)78, which is citrullinated on arginine (R) residues in inflamed mouse and human islets (9, 10) by peptidyl arginine deiminase enzymes (PADI; Padi in the mouse). GRP78 (also known as BiP and HSPA5) is an endoplasmic reticulum (ER) chaperone of the heat shock protein 70

family. It is involved in anabolic pathways, i.e. protein (re)folding and translocation of secretory proteins; and catabolic pathways, i.e. the import of polypeptides - especially hydrophobic ones – inside the ER, the ER-associated degradation pathway, peptide export outside the ER for proteasome degradation, and the regulation of the unfolded protein response (11). Citrullinated GRP78 is translocated to the plasma membrane and secreted, triggering β-cell apoptosis (9, 12). We first reported that the murine orthologue of GRP78 (Grp78) is targeted by T cells and autoantibodies (aAbs) in NOD mice (9). While interferon (IFN)-γ T-cell responses were only observed against citrullinated Grp78, aAbs were detected against both native and citrullinated Grp78, but the latter was preferentially recognized, with higher aAb titers in NOD mice than in non-autoimmune strains (9). Translating to humans, a $CD4^+$ T-cell line grown from type 1 diabetic islets responded to a $GRP78_{292-305}$ peptide only when carrying an arginine-to-citrulline (R→X) modification (13). aAbs against both native GRP78 (R-GRP78) and citrullinated GRP78 (X-GRP78) were detected at increased titers in type 1 diabetic vs. age-matched healthy individuals, but again X-GRP78 was preferentially recognized (10). $CD4^+$ T cells reactive to an HLA-DR4-binding X-$GRP78_{498-512}$ peptide displayed higher frequencies in type 1 diabetic vs. healthy donors, but these frequencies were similar to those of $CD4^+$ T cells recognizing R-$GRP78_{498-512}$. Importantly, this $GRP78_{498-512}$ epitope features a R510 residue that is citrullinated in human islets exposed *in vitro* to inflammatory cytokines (10). Collectively, these results suggest that $CD4^+$ T cells recognize preferentially, but not exclusively, X-GRP78 over its native version.

We therefore studied the recognition of R- and X-GRP78 peptides restricted for HLA-A*02:01 (HLA-A2 from hereon) by the $CD8^+$ T cells of type 1 diabetic and healthy donors. Our results show that the preferential recognition of X-GRP78 versions is not the rule, and that this recognition may be shaped by T-cell cross-reactivity with bacterial mimotopes and by PADI-mediated citrullination in the thymus and islets.

RESEARCH DESIGN AND METHODS

HLA-A2 binding measurements

Experimental binding of peptides (>90% purity, SynPeptide) to HLA-A2 was measured by flow cytometry using biotin-tagged HLA-A2 monomers (ImmunAware) assembled as described (14), captured on streptavidin-coated beads (Spherotech), and revealed by an anti-β2m monoclonal antibody (mAb; RRID:AB_626748) and an AlexaFluor (AF)488-conjugated goat anti-mouse IgG (RRID:AB_2728715).

Donors and specimens

Human studies were authorized under Ethics approval #S52697 (UZ-Leuven) and written informed consent was obtained from all participants. Heparinized blood was drawn from HLA-A*02:01$^+$ donors: 16 aAb$^+$, insulin-treated type 1 diabetic adults (median age 27 years, range 19-47; 9 females, 7 males; median disease duration 4.3 years, range 1.7-6.0), 11 healthy donors (median age 37 years, 26-60; 7 females, 4 males) and 8 type 1 diabetic adolescents (median age 16 years, range 12-17; 3 females, 5 males; median disease duration 4.2 years, range 1.0-12.0). Ficoll-purified peripheral blood mononuclear cells (PBMCs) were stored in liquid nitrogen until use. Pancreatic tissue sections were provided by the Network for Pancreatic Organ Donors with Diabetes (nPOD).

HLA-A2 multimer (MMr) staining

Our previous combinatorial HLA-A2 MMr assays (3, 4) allow to multiplex the detection of 15 different T-cell epitope specificities in the same PBMC sample. It is based on the use of unique fluorochrome pairs to label each peptide-loaded MMr, with subsequent detection of double-labeled MMr$^+$CD8$^+$ T cells with increased specificity. This format was modified so that each R-GRP78 MMr shared one fluorochrome with its corresponding X-GRP78 MMr, thus allowing

to selectively gate $CD8^+$ T cells recognizing only the R or X version (i.e. double MMr^+), or both (i.e. triple MMr^+). For example, the GRP78-261R and GRP78-261X peptides were loaded on PE/BV786- and PE/PECF594-labeled MMr pairs, respectively, thus sharing the PE fluorochrome. GRP78-261R, -261X and -261R/X MMr^+CD8^+ T cells were hence defined as $PE^+/BV786^+$, $PE^+/PECF594^+$ and $PE^+/BV786^+/PECF594^+$, respectively. MMr staining and acquisition was performed as previously detailed (3). HLA-A2-binding epitope candidates without appreciable MMr staining provided negative controls. Data were analyzed with FlowJo software as detailed in Supplementary Fig. 1-2.

In-situ HLA-A2 MMr immunohistochemistry staining was performed as described (3).

mTEC isolation

For RT-qPCR and proteomics, thymi from 6-8-week-old C57BL/6 and NOD mice were minced, gently agitated to release excess thymocytes, digested with Liberase, filtered, panned on anti-CD90.2-coated plates, stained with Fixable Viability Dye eFluor780, CD45-PE/Cy7, MHC Class II (MHC-II)-AF488, EPCAM-BV421, BP-1-APC, *Ulex europaeus* agglutinin-1 (UEA-1-biotin), streptavidin-PE and sorted on a BD Influx instrument.

For RNAseq, thymi were digested with collagenase D and collagenase/dispase, depleted with CD45 MicroBeads, stained with CD45-PerCPCy5.5, Ly51-PE, MHC-II-APC and sorted on a BD FACSAria III instrument.

These procedures are detailed in Supplementary Methods.

Reverse-transcription quantitative PCR (RT-qPCR), proteomics and RNAseq

Mouse studies were performed under authorization #P076/2020 (UZ-Leuven). Islets were isolated from 6- and 10-week-old C57BL/6 and NOD mice as described (15). Total RNA was extracted from sorted mTECs or islets, using the Single Cell RNA Purification Kit (Norgen)

and cDNA synthesized with SuperScript VILO (Invitrogen). RT-qPCR was performed on a StepOnePlus RT-PCR System (Applied Biosystems). The relative fold gene expression was calculated using the delta-delta Ct method.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) and manual verification of citrullinated residues were performed essentially as described (16). RNAseq procedures were previously reported (17).

These procedures are detailed in Supplementary Methods.

Statistics

Significance was assessed using two-tailed tests with a cutoff value of $\alpha=0.05$, as detailed for each figure and in Supplementary Methods.

Data and Resource Availability

The mTEC RNAseq datasets have been deposited in the Gene Expression Omnibus under accession numbers GSE140815, GSE140683 and GSE178456. All raw data are available from the corresponding author upon reasonable request. The study does not involve any noncommercial reagent or tool.

RESULTS

Citrullination of GRP78 peptides does not decrease HLA-A2 binding affinity

We selected 6 candidate epitopes by screening the whole GRP78 sequence for predicted HLA-A2 binders carrying R residues (**Fig. 1A**): $GRP78_{9-18}$ (GRP78-9 from hereon), $GRP78_{261-269}$ (GRP78-261), $GRP78_{298-307}$ (GRP78-298), $GRP78_{360-368}$ (GRP78-360), $GRP78_{435-443}$ (GRP78-435), and $GRP78_{502-511}$ (GRP78-502).

Since binding affinities were predicted for the native peptide sequence without citrullination, we experimentally verified binding using monomeric HLA-A2 molecules and a flow cytometry assay detecting peptide/HLA-A2/β2m complexes (4) (**Fig. 1B**). First, this assay verified the proper folding of the peptide/HLA-A2 monomers subsequently used for MMr synthesis. Second, it revealed strong (GRP78-261, -435, -298), moderate (GRP78-360, -9) and weak binders (GRP78-502) (**Fig. 1C-D**). The HLA-A2 binding of the citrullinated (X) vs. native (R) versions was similar (GRP78-261, -435), marginally increased (GRP78-298, -502) or significantly increased (GRP78-360, -9), but never decreased. The increase in HLA-A2 binding was observed for those peptides in which the R→X substitution was introduced at the C-terminal P8-P9 HLA anchor positions.

Collectively, these data show that citrullination does not negatively affect HLA-A2 binding for the candidate epitopes identified.

Blood $CD8^+$ T cells preferentially recognize either native or citrullinated peptide versions and display similar frequencies in type 1 diabetic and healthy donors

We subsequently set up combinatorial HLA-A2 MMr assays (3, 4) with two purposes. First, to define the *ex-vivo* frequency and naïve/antigen-experienced phenotype of GRP78-reactive $CD8^+$ T cells in type 1 diabetic and healthy donors. Second, to investigate whether the native and citrullinated versions of any given peptide are recognized by the same T-cell clonotypes.

For this second purpose, we modified our previous assays so that the MMr pairs loaded with each R version shared one fluorochrome with the MMr pairs loaded with its corresponding X version (**Supplementary Table 1**), thus allowing to selectively gate $CD8^+$ T cells recognizing only the R or X version, or both (**Fig. 2A-B**). A representative staining is shown in **Supplementary Fig. 1** and the analysis strategy is detailed in **Supplementary Fig. 2**.

The results from adult donors are summarized in **Fig. 2C**. Recognition of the R and X versions was rather exclusive, with very few donors recognizing both. The native R version was preferentially recognized for peptides GRP78-9, -298, -360, -435, while GRP78-261 and -502 were preferred in their X version. The R version was also preferred for those peptides whose citrullination at the C-terminus enhanced HLA-A2 binding (GRP78-9, -360 and, to a lesser extent, -298), thus excluding that this preference was simply driven by HLA-A2 binding affinity. Despite some differences in frequency, the preferred and exclusive recognition of the R or X version according to individual peptides was similar in type 1 diabetic adolescents (**Fig. 3**).

As for other islet epitopes (2-4, 18-20), the frequency of GRP78-reactive $CD8^+$ T cells was not higher in type 1 diabetic than in healthy adults (Fig. 2C), with one exception noted for GRP78-435R. All positive GRP78-reactive $CD8^+$ T-cell fractions displayed frequencies similar to those observed for the control preproinsulin $(PPI)_{6-14}$ islet epitope, confirming the reported range of ~1-50/10^6 $CD8^+$ T cells (2-4).

Collectively, these results show that, depending on the individual peptide, $CD8^+$ T cells exclusively recognize either the native or citrullinated version, with no significant cross-reactivity. Moreover, they circulate at similar frequencies in type 1 diabetic and healthy adults.

GRP78 peptides preferentially recognized in their native version are cross-reactive with gut bacterial mimotopes

Another recurrent feature of circulating $CD8^+$ T cells recognizing other islet epitopes is their largely naïve phenotype (2-4). Also GRP78-reactive $CD8^+$ T cells were largely (>80%) naïve in most donors, but substantial antigen-experienced fractions were observed in some. This was noteworthy for epitopes GRP78-298R, -360R and -435R preferentially recognized in their native version, which displayed >30% antigen-experienced fractions in >30% of adult donors (**Fig. 2D**). We therefore hypothesized that these R-GRP78-reactive $CD8^+$ T cells may cross-react with foreign antigens. A sequence homology search (**Table 1**) retrieved homologous bacterial sequences (mimotopes) for these 3 epitopes and for GRP78-502, which was instead predominantly recognized as citrullinated. Bacterial sequences were identical for GRP78-435R and, notably, GRP78-360R, which displayed particularly high cognate $CD8^+$ T-cell frequencies (Fig. 2C). These mimotopes mapped to proteins expressed by gut commensal bacteria (*Enterobacteriaceae*, *Proteobacteria*, *Clostridia*, *Escherichia coli*), including the well-known DnaK chaperone (21). While the GRP78-360R- and GRP78-435R-identical bacterial sequences ('homotopes') are by definition cross-reactive, we investigated whether this was also the case for the GRP78-298R and -502X mimotopes. To this end, we used a combinatorial MMr strategy similar to the one applied to define R-and X-GRP78 cross-reactivity (**Fig. 4A**). Each of 3 PBMC aliquots was stained with two MMr pairs, and each MMr pair was loaded with either the GRP78, mimotope or control GRP78-360R or Flu MP_{58-66} peptide. The two MMr pairs in each tube shared one fluorochrome with each other to allow visualization of cross-reactive T cells as triple-labeled MMr^+ events. A sizable fraction (24±4% in healthy donors, 25±14% in type 1 diabetic donors) of $CD8^+$ T cells recognizing the GRP78-298R epitope was cross-reactive with its bacterial mimotope (**Fig. 4B**, left), while this was not the case when analyzing cross-reactivity with the control GRP78-360R or Flu epitope (**Fig. 4B**, middle). As a positive control, GRP78-360R/GRP78-360R or Flu/Flu triple-MMr^+ events were correctly visualized as largely cross-reactive (**Fig. 4B**, right). This cross-reactivity was instead lower for $CD8^+$ T cells

recognizing the GRP78-502X epitope (4±3% of total in healthy donors, 13±4% in type 1 diabetic donors; **Fig. 4C**).

Collectively, these results suggest that the preferential $CD8^+$ T-cell recognition of some R-GRP78 versions is influenced by cross-reactivity with bacterial mimotopes.

Pancreas-infiltrating $CD8^+$ T cells preferentially recognize the citrullinated GRP78-9X epitope

While the preferential recognition of GRP78 native versions in peripheral blood may be favored by bacterial cross-reactivity, the situation may be different in the pancreas, where citrullination and its catalyzing enzymes PADIs are upregulated by islet inflammation (9, 10). We therefore performed *in-situ* MMr staining on pancreas sections from nPOD donors (**Supplementary Table 2**) using the prototype epitope GRP78-9R/X, whose R version was preferentially recognized in peripheral blood. A representative GRP78-9X MMr staining on pancreas sections is shown in **Fig. 5A**. As reported for other islet antigen specificities (2, 3, 22), scattered GRP78-9X MMr^+ cells were detected in the islet vicinity and in the exocrine tissue. In type 1 diabetic donors, they displayed significantly higher densities than control MMr^+ cells recognizing the melanocyte epitope $MelanA_{26-35}$ (**Fig. 5B**). However, GRP78-9X MMr^+ cells were not enriched in type 1 diabetic patients compared to non-diabetic donors, suggesting no disease specificity. The density of these GRP78-9X MMr^+ cells was lower than that of $ZnT8_{186-194}$ MMr^+ cells, and similar to that of the $UCN3_{1-9}$ MMr^+ cells. GRP78-9X MMr^+ cells were significantly more abundant than their GRP78-9R MMr^+ counterparts when pooling type 1 diabetic and aAb^+ donors, but not when pooling control non-diabetic and type 2 diabetic donors. They were instead absent in duodenal mucosa sections from type 1 diabetic cases (**Fig. 5C**), despite the presence of $CD8^+$ T cells (not shown), thus confirming pancreas specificity. Analysis of a more limited set of samples for the epitope GRP78-360R/X, which, like GRP78-9R/X, was

preferentially recognized in its R version in peripheral blood, also detected recognition of the X version in the pancreas of one type 1 diabetic donor (**Fig. 5D**).

Collectively, these data indicate that citrullinated GRP78-9X rather than its native GRP78-9R version is recognized by $CD8^+$ T cells in the pancreas, and that this pancreas-specific reactivity is not enriched in type 1 diabetic cases.

The citrullinating Padi2 enzyme is expressed in murine mTECs under the control of Aire and is diminished in NOD mice.

To summarize, we observed that blood $CD8^+$ T cells preferentially recognize the native R-GRP78 versions for several epitopes, including GRP78-9R, while the citrullinated GRP78-9X version is preferentially recognized in the pancreas. Besides the potential effect of bacterial cross-reactivity, we asked how to reconcile these findings with the expected impact of central and peripheral tolerance mechanisms in shaping these responses. It is currently held that PTM recognition should be favored, because PTMs may take place under inflammatory conditions in the periphery and not in the thymus (8). Hence, T cells recognizing citrullinated epitopes may be more likely to escape thymic deletion, and may then be preferentially primed in the inflammatory milieu of insulitis that promotes citrullination.

To address this possibility, we compared the gene expression of *Padi2*, which is the predominant islet isoform (23), in the islets and thymus between NOD and non-autoimmune C57BL/6 mice. As previously reported (9), islet *Padi2* expression was negligible in C57BL/6 mice and higher in NOD mice (**Fig. 6A**). This expression increased with age and correlated with a difference in the Padi enzymatic activity (**Fig. 6B**). In mTECs (**Fig. 6C-D**; gating strategy in **Supplementary Fig. 3A-B**), *Padi2* and *Padi4* were substantially upregulated with maturation in C57BL/6 but not in NOD mice, and lower in NOD vs. C57BL/6 mice for the mature MHC-II^{hi} mTECs driving T-cell deletion. Mature MHC-II^{hi} mTECs were also less

represented in NOD mice (**Supplementary Fig. 3C**). The Padi enzymatic activity from whole thymic tissue was not significantly different between the two strains (**Fig. 6B**), possibly reflecting the mixture of mTECs (both MHC-II^{lo} and MHC-II^{hi}) and other cells in these tissue preparations. Proteomics analysis of mTECs revealed citrullinated R residues in three proteins (**Supplementary Fig. 4**; 3 out of 920 unique proteins detected, 0.33%). *PADI* gene expression was also observed in human mTECs (**Supplementary Fig. 5**), and was dominated by the *PADI2* isoform.

We next asked whether mTECs express *Padi* isoforms under the control of the autoimmune regulator (Aire) transcription factor. Expression of all *Padi* isoforms was detected by RNAseq in MHC-II^{hi} mTECs isolated from wild-type (WT) C57BL/6 mice (**Fig. 6E**), but significantly diminished or absent in *Aire*$^{-/-}$ mice, suggesting an *Aire*-dependent/enhanced expression. As in islets, the predominant mTEC isoform was *Padi2*, which was less impacted by *Aire* knock-out in C57BL/6 mice. The enrichment of *Padi2* and *Padi4* in MHC-II^{hi} mTECs from WT vs. *Aire*$^{-/-}$ NOD mice (**Supplementary Fig. 6A**) further supported an *Aire*-dependent/enhanced expression. No modulation by *Aire* was detected for *Grp78* and the tissue transglutaminase gene *Tgm2* responsible for the PTM transglutamination. The ATAC-seq and Chip-seq profiles of *Padi* genes (**Supplementary Fig. 7**) were also in line with an Aire-dependent expression mechanism at an epigenetic level (24), i.e. closed transcription start sites, near-absence of activating chromatin modifications (H3K4Ac, H3K9Ac) and presence of repressive modifications (H3K27me3).

Further analysis of gene sets specific for modifying enzymes (**Supplementary Table 3**) revealed that, among the 22 enzymatic PTMs analyzed, citrullination was the only one under the control of Aire (**Fig. 6F**; **Supplementary Fig. 6B** for NOD mice) and enriched in mature MHC-II^{high} vs. immature MHC-II^{low} mTECs of WT mice (**Supplementary Fig. 8**). Finally,

reduced Padi2 expression in the mTECs of NOD and *Aire*$^{-/-}$ C57BL/6 mice compared to WT C57BL/6 mice was confirmed at the protein level by confocal microscopy (**Fig. 7**).

Collectively, these observations suggest that Padi gene expression and enzymatic activity may catalyze citrullination not only in the islets, but also in the thymus. In the NOD mice, the increased, age-dependent expression of *Padi2* in the islets may augment the availability of citrullinated peptides for T-cell recognition. On the other hand, the reduced *Padi2* and *Padi4* expression in the mature MHC-II^{high} mTECs of NOD mice may favor the escape of T cells reactive to citrullinated peptides.

DISCUSSION

It is currently posited that post-translationally modified epitopes may be more prone to loss of tolerance (8). This is based on the assumption that these epitopes may not be presented in the thymus, as PTMs are usually induced in a tissue-specific fashion and upregulated by inflammation. This assumption has been experimentally verified in elegant mouse models for glycosylated collagen type II (25), which is expressed in cartilages (including those from healthy donors), but not in the thymus. By analyzing GRP78 citrullination, we here document that this tenet may not always apply. Indeed, either native or citrullinated epitopes were preferentially recognized by circulating $CD8^+$ T cells. Moreover, this recognition was not promiscuous, since most T cells recognized exclusively one epitope version or the other.

These observations prompt two questions. First, how can GRP78 be recognized in its native form, despite its ubiquitous expression? We suggest that one driver for such recognition may be cross-reactivity with gut bacterial sequences, as mimotopes - and even homotopes - were detected for some GRP78 peptides preferred in their R version, but not for those preferentially recognized as citrullinated. Of note, these R-GRP78 versions were also those displaying significant fractions of antigen-experienced $CD8^+$ T cells in some individuals and, for those with an identical bacterial match (GRP78-360R and -435R), particularly high frequencies. The final outcome of this bacterial cross-recognition is likely to be tolerance, which may however be lost with gut dysbiosis, as recently described for IGRP microbial mimotopes (26, 27).

Second, why are citrullinated GRP78 peptides not universally preferred by $CD8^+$ T cells? This may reflect the presentation of citrullinated peptides in the thymus, as both murine and human mTECs robustly expressed citrullinating Padi enzymes, most notably Padi2. *Padi* expression in murine mTECs has been previously reported along with the presence of citrullination (28). The finding of Padi2 expression and Padi enzymatic activity in the murine thymus and of citrullinated proteins in purified mTECs provide direct evidence that this gene expression can

translate into PTMs. We further show that *Padi* expression is under the control of Aire in murine mTECs. The effect of Aire on promiscuous antigen expression may thus extend to citrullinated proteins. Hence, $CD8^+$ T cells recognizing native and citrullinated peptides may undergo a similar, albeit incomplete (29), thymic deletion. Interestingly, NOD mice displayed lower *Padi2* and *Padi4* expression in mature mTECs and higher expression in islets [likely in both β cells and immune infiltrates (30)], which may exert a synergistic effect on the loss of central and peripheral tolerance against citrullinated epitopes, respectively. In line with this possibility, we previously showed that spleen T cells from NOD mice, but not from C57BL/6 mice, recognize the citrullinated version of the Grp78 protein, but not its native version (9). Moreover, the lack of *Padi2*/*Padi4* upregulation between MHC-II^{lo} (and $Aire^-$) and MHC-II^{hi} ($Aire^+$) mTECs in NOD mice suggests that the Aire-driven control is impaired in this strain. The NOD thymus also harbored less MHC-II^{hi} mTECs, which may further contribute to defective central tolerance. Although inter-species differences are possible, it will be relevant to define whether an equivalent of this *Padi2* expression pattern in the mTECs and islets of NOD mice exists in type 1 diabetic patients.

The finding of a similar frequency of GRP78-reactive $CD8^+$ T cells in the blood of type 1 diabetic and healthy individuals mirrors previous observations with other islet epitopes (2-4, 18, 20). These observations are likely to reflect the sequestration of the fraction engaged in the disease in the pancreas, where the same T cells were enriched in type 1 diabetic vs. non-diabetic donors (2, 3). Given the notion that citrullination is enhanced in the inflammatory milieu of insulitis (30), we hypothesized that X-GRP78-reactive T cells may preferentially accumulate in the type 1 diabetic pancreas. However, despite the fact that the native GRP78-9R peptide recognized by circulating $CD8^+$ T cells was preferred in its citrullinated version in the pancreas, these T cells were not specific for type 1 diabetes.

This notion of tissue but not disease specificity echoes our recent report documenting that citrullination is also detectable in non-infiltrated islets of non-autoimmune C57BL/6 mice (30). It is possible that citrullination, although upregulated by inflammation (9, 10), may be, to a certain extent, physiological in the pancreas, as is the case in other tissues (31) and for other PTMs (25), e.g. iodinated thyroglobulin and glycosylated collagen. At the steady state, the outcome of this physiological citrullination may be peripheral tolerance, which may be broken under conditions of inflammation and ER stress that also upregulate GRP78 expression (11). Whether citrullination results in the presentation of citrullinated peptides by HLA Class I molecules on β cells or antigen-presenting cells remains an open question. Despite the presence of GRP78 peptides, immunopeptidomics studies on human and mouse islets (3, 32) did not retrieve citrullinated peptides. It is noteworthy that, despite the notion of protein citrullination in inflamed cartilages (10, 30, 33), direct evidence of HLA-eluted citrullinated peptides is missing even for rheumatoid arthritis. While this may reflect a technical limitation of immunopeptidomics strategies, it invites the question of whether the diabetogenic role of citrullination (30) relies on T cells recognizing X-GRP78 peptides. We documented that Padi inhibition by BB-Cl-amidine treatment protects NOD mice from diabetes and is associated with decreased citrullination of GRP78 and other proteins in the pancreas, reduced anti-X-GRP78 aAb titers, and diminished formation of neutrophil extracellular traps (NETs) (30). Interestingly, diabetes protection was observed upon treatment from 8 weeks of age, when insulitis is already present, suggesting an effect on disease amplification rather than priming. This is consistent with the higher titers of aAbs against X-GRP78 detected in long-standing type 1 diabetic patients (10). As citrullination is involved in the control of several biological processes (31), both physiological and pathological [e.g. NET formation (34)], pathogenic effects independent of antigen presentation and T-cell recognition are possible. For instance, circulating aAbs against native or citrullinated GRP78 are described in patients with cancers

(35), rheumatoid arthritis (36) or atherosclerosis (37). aAb binding to surface GRP78 activates endothelial cells (37) and monocytes/macrophages (36), upregulating inflammatory cytokines and adhesion molecules. Whether a similar process takes place in β cells in the course of type 1 diabetes deserves further investigation.

This study carries some limitations. First, we did not provide evidence of natural processing and presentation of the GRP78 epitopes studied. Most of them are in close proximity to dibasic cleavage sites, which may favor their processing by the proteasome and other proteases. For the GRP78-502 epitope, which is exclusively recognized as citrullinated, we reported direct proteomic evidence for citrullination at position 510R in cytokine-treated human islets (10) and rat INS1E β cells (9). Interestingly, this peptide overlaps with the citrullinated $GRP78_{498\text{-}512}$ epitope targeted by circulating $CD4^+$ T cells that are enriched in type 1 diabetic donors (10). Second, the limited availability of human tissues allowed us to verify the absence of GRP78-9X-reactive $CD8^+$ T cells outside the pancreas only in the duodenum.

In conclusion, our study shows that GRP78 is targeted by circulating $CD8^+$ T cells that do not always prefer citrullinated peptides. Their recognition pattern may be shaped by cross-reactivity with gut microbial mimotopes and by thymic PADI expression. These findings mitigate the current claims that citrullinated epitopes are invariably more immunogenic than their native counterparts.

Acknowledgments. We thank C. Maillard for technical assistance, the Cochin Institute CyBio and KU Leuven Flow Cytometry Core Facilities for assistance with cell analysis and sorting, and J. Perez-Hernandez (Cochin Institute) for critical reviewing of the manuscript.

Funding. This work was supported by JDRF grants 2-SRA-2016-164-Q-R (to R.M.), 2-SRA-2015-52-Q-R (to C.M. and L.O.), 2-SRA-2018-480-S-B (to M.N.) and Postdoctoral Fellowship 3-PDF-2020-942-A-N (to Z.Z.); The Leona M. and Harry B. Helmsley Charitable Trust

(Helmsley #1901-03689), the *Fondation Francophone pour la Recherche sur le Diabète* (FFRD), the EFSD/JDRF/Lilly European Programme in Type 1 Diabetes Research 2015, the *Agence Nationale de la Recherche* (ANR-19-CE15-0014-01) and the *Fondation pour la Recherche Medicale* (EQU20193007831), to R.M.; and the NIDDK (R01DK099317, R01DK032083), to M.N.. M.B. and Aï.Ca. were supported by a fellowship from the Flemish Research Foundation (12R0719N and 1189518N, respectively). F.D., D.L.E., C.M., L.O. and R.M. received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST), which receive support from the EU Horizon 2020 program, the European Federation of Pharmaceutical Industries and Associations, JDRF, and The Leona M. & Harry B. Helmsley Charitable Trust. This research was performed with the support of the Network for Pancreatic Organ donors with Diabetes (nPOD; RRID: SCR_014641), a collaborative type 1 diabetes research project sponsored by JDRF (nPOD: 5-SRA-2018-557-Q-R) and The Leona M. & Harry B. Helmsley Charitable Trust (Grant #2018PG-T1D053). Organ Procurement Organizations (OPO) partnering with nPOD to provide research resources are listed at www.jdrfnpod.org/for-partners/npod-partners. The content and views expressed are the responsibility of the authors and do not necessarily reflect an official view of nPOD.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. Conceptualization, F.S., M.G., C.M., L.O., R.M.; Methodology, M.E.A., F.S., M.B., L.N., N.B., Aï.Ca., M.G., M.I., A.I.L., Al.Ca., Z.Z., B.B., M.L.C., G.S., F.D., M.N., D.L.E., S.Y., S.P., M.J.M., Y.V., J.V., S.B., C.M., L.O., R.M.; Investigation, M.E.A., F.S., M.B., L.N., N.B., Aï.Ca., M.G., M.I., A.I.L., Al.Ca., G.A., Z.Z., B.B., M.L.C., G.S., M.N., L.O., R.M.; Data Curation, M.E.A., F.S., M.B., L.N., N.B., Aï.Ca., M.G., Al.Ca., M.L.C., G.S., M.J.M., L.O., R.M.; Writing, M.E.A., F.S., M.B., M.G., L.O., R.M.; Visualization, M.E.A., F.S., M.B., L.N., N.B., Aï.Ca., M.G., M.I., L.O., R.M.; Supervision,

L.O., R.M.; Funding Acquisition, S.Y., C.M., L.O., R.M. R.M. is the guarantor of this work and, as such, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior presentation. Parts of this study were presented at the nPOD 12th Annual Scientific Meeting, Tampa, FL, 23–26 February 2020.

REFERENCES

1. Carré A, Richardson SJ, Larger E, Mallone R: Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset? Diabetologia 2021;64:15-25

2. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, Kuranda K, Nigi L, Eugster A, Osterbye T, Maugein A, McLaren JE, Ladell K, Larger E, Beressi JP, Lissina A, Appay V, Davidson HW, Buus S, Price DA, Kuhn M, Bonifacio E, Battaglia M, Caillat-Zucman S, Dotta F, Scharfmann R, Kyewski B, Mallone R, ImMaDiab Study G: Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 2018;3:eaao4013

3. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, Lalanne AI, Sebastiani G, Carre A, Pinto S, Culina S, Corcos N, Bugliani M, Marchetti P, Armanet M, Diedisheim M, Kyewski B, Steinmetz LM, Buus S, You S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Dotta F, Scharfmann R, Eizirik DL, Verdier Y, Vinh J, Mallone R: Conventional and Neo-antigenic Peptides Presented by beta Cells Are Targeted by Circulating Naive CD8+ T Cells in Type 1 Diabetic and Healthy Donors. Cell Metab 2018;28:946-960 e946

4. Azoury ME, Tarayrah M, Afonso G, Pais A, Colli ML, Maillard C, Lavaud C, Alexandre-Heymann L, Gonzalez-Duque S, Verdier Y, Vinh J, Pinto S, Buus S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Eizirik DL, You S, Mallone R: Peptides Derived From Insulin Granule Proteins are Targeted by CD8(+) T Cells Across MHC Class I Restrictions in Humans and NOD Mice. Diabetes 2020;69:2678-2690

5. Alanio C, Lemaitre F, Law HK, Hasan M, Albert ML: Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. Blood 2010;115:3718-3725

6. Yu W, Jiang N, Ebert PJ, Kidd BA, Muller S, Lund PJ, Juang J, Adachi K, Tse T, Birnbaum ME, Newell EW, Wilson DM, Grotenbreg GM, Valitutti S, Quake SR, Davis MM: Clonal deletion prunes but does not eliminate self-specific alphabeta CD8(+) T lymphocytes. Immunity 2015;42:929-941

7. Mallone R, Eizirik DL: Presumption of innocence for beta cells: why are they vulnerable autoimmune targets in type 1 diabetes? Diabetologia 2020;63:1999-2006

8. James EA, Pietropaolo M, Mamula MJ: Immune Recognition of beta-Cells: Neoepitopes as Key Players in the Loss of Tolerance. Diabetes 2018;67:1035-1042

9. Rondas D, Crevecoeur I, D'Hertog W, Bomfim Ferreira G, Staes A, Garg AD, Eizirik DL, Agostinis P, Gevaert K, Overbergh L, Mathieu C: Citrullinated glucose-regulated protein 78 is an autoantigen in type 1 diabetes. Diabetes 2015;64:573-586

10. Buitinga M, Callebaut A, Marques Camara Sodre F, Crevecoeur I, Blahnik-Fagan G, Yang ML, Bugliani M, Arribas-Layton D, Marre M, Cook DP, Waelkens E, Mallone R, Piganelli JD, Marchetti P, Mamula MJ, Derua R, James EA, Mathieu C, Overbergh L: Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. Diabetes 2018;67:2337-2348

11. Wang J, Lee J, Liem D, Ping P: HSPA5 Gene encoding Hsp70 chaperone BiP in the endoplasmic reticulum. Gene 2017;618:14-23

12. Vig S, Buitinga M, Rondas D, Crèvecoeur I, van Zandvoort M, Waelkens E, Eizirik DL, Gysemans C, Baatsen P, Mathieu C, Overbergh L: Cytokine-induced translocation of GRP78 to the plasma membrane triggers a pro-apoptotic feedback loop in pancreatic beta cells. Cell Death Dis 2019;10:309

13. Babon JA, DeNicola ME, Blodgett DM, Crevecoeur I, Buttrick TS, Maehr R, Bottino R, Naji A, Kaddis J, Elyaman W, James EA, Haliyur R, Brissova M, Overbergh L, Mathieu C, Delong T, Haskins K, Pugliese A, Campbell-Thompson M, Mathews C, Atkinson MA, Powers

AC, Harlan DM, Kent SC: Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. Nat Med 2016;22:1482-1487
14. Leisner C, Loeth N, Lamberth K, Justesen S, Sylvester-Hvid C, Schmidt EG, Claesson M, Buus S, Stryhn A: One-pot, mix-and-read peptide-MHC tetramers. PLoS One 2008;3:e1678
15. Casteels K, Waer M, Laureys J, Valckx D, Depovere J, Bouillon R, Mathieu C: Prevention of autoimmune destruction of syngeneic islet grafts in spontaneously diabetic nonobese diabetic mice by a combination of a vitamin D3 analog and cyclosporine. Transplantation 1998;65:1225-1232
16. Callebaut A, Derua R, Vig S, Delong T, Mathieu C, Overbergh L: Identification of Deamidated Peptides in Cytokine-Exposed MIN6 Cells through LC-MS/MS Using a Shortened Digestion Time and Inspection of MS2 Spectra. J Proteome Res 2021;20:1405-1414
17. Guyon C, Jmari N, Padonou F, Li YC, Ucar O, Fujikado N, Coulpier F, Blanchet C, Root DE, Giraud M: Aire-dependent genes undergo Clp1-mediated 3'UTR shortening associated with higher transcript stability in the thymus. eLife 2020;9
18. Wiedeman AE, Muir VS, Rosasco MG, DeBerg HA, Presnell S, Haas B, Dufort MJ, Speake C, Greenbaum CJ, Serti E, Nepom GT, Blahnik G, Kus AM, James EA, Linsley PS, Long SA: Autoreactive CD8+ T cell exhaustion distinguishes subjects with slow type 1 diabetes progression. J Clin Invest 2019;130:480-490
19. Yeo L, Pujol-Autonell I, Baptista R, Eichmann M, Kronenberg-Versteeg D, Heck S, Dolton G, Sewell AK, Harkonen T, Mikk ML, Toppari J, Veijola R, Knip M, Ilonen J, Peakman M: Circulating beta cell-specific CD8(+) T cells restricted by high-risk HLA class I molecules show antigen experience in children with and at risk of type 1 diabetes. Clin Exp Immunol 2020;199:263-277
20. Skowera A, Ladell K, McLaren JE, Dolton G, Matthews KK, Gostick E, Kronenberg-Versteeg D, Eichmann M, Knight RR, Heck S, Powrie J, Bingley PJ, Dayan CM, Miles JJ, Sewell AK, Price DA, Peakman M: Beta-cell-specific CD8 T cell phenotype in type 1 diabetes reflects chronic autoantigen exposure. Diabetes 2015;64:916-925
21. Genest O, Wickner S, Doyle SM: Hsp90 and Hsp70 chaperones: Collaborators in protein remodeling. J Biol Chem 2019;294:2109-2120
22. Bender C, Rodriguez-Calvo T, Amirian N, Coppieters KT, von Herrath MG: The healthy exocrine pancreas contains preproinsulin-specific CD8 T cells that attack islets in type 1 diabetes. Sci Adv 2020;6
23. Crevecoeur I, Gudmundsdottir V, Vig S, Marques Camara Sodre F, D'Hertog W, Fierro AC, Van Lommel L, Gysemans C, Marchal K, Waelkens E, Schuit F, Brunak S, Overbergh L, Mathieu C: Early differences in islets from prediabetic NOD mice: combined microarray and proteomic analysis. Diabetologia 2017;60:475-489
24. Handel AE, Shikama-Dorn N, Zhanybekova S, Maio S, Graedel AN, Zuklys S, Ponting CP, Hollander GA: Comprehensively Profiling the Chromatin Architecture of Tissue Restricted Antigen Expression in Thymic Epithelial Cells Over Development. Front Immunol 2018;9:2120
25. Raposo B, Merky P, Lundqvist C, Yamada H, Urbonaviciute V, Niaudet C, Viljanen J, Kihlberg J, Kyewski B, Ekwall O, Holmdahl R, Bäcklund J: T cells specific for post-translational modifications escape intrathymic tolerance induction. Nat Commun 2018;9:353
26. Hebbandi Nanjundappa R, Ronchi F, Wang J, Clemente-Casares X, Yamanouchi J, Sokke Umeshappa C, Yang Y, Blanco J, Bassolas-Molina H, Salas A, Khan H, Slattery RM, Wyss M, Mooser C, Macpherson AJ, Sycuro LK, Serra P, McKay DM, McCoy KD, Santamaria P: A Gut Microbial Mimic that Hijacks Diabetogenic Autoreactivity to Suppress Colitis. Cell 2017;171:655-667 e617

27. Tai N, Peng J, Liu F, Gulden E, Hu Y, Zhang X, Chen L, Wong FS, Wen L: Microbial antigen mimics activate diabetogenic CD8 T cells in NOD mice. J Exp Med 2016;213:2129-2146
28. Engelmann R, Biemelt A, Cordshagen A, Johl A, Kuthning D, Muller-Hilke B: The Prerequisites for Central Tolerance Induction against Citrullinated Proteins in the Mouse. PLoS One 2016;11:e0158773
29. Davis MM: Not-So-Negative Selection. Immunity 2015;43:833-835
30. Sodré MCF, Bissenova S, Bruggeman Y, Tilvawala R, Cook DP, Berthault C, Mondal S, Callebaut A, You S, Scharfmann R, Mallone R, Thompson PR, Mathieu C, Buitinga M, Overbergh L: Peptidylarginine deiminase inhibition prevents diabetes development in NOD mice. Diabetes 2021;70:516-528
31. Alghamdi M, Alasmari D, Assiri A, Mattar E, Aljaddawi AA, Alattas SG, Redwan EM: An Overview of the Intrinsic Role of Citrullination in Autoimmune Disorders. J Immunol Res 2019;2019:7592851
32. Wan X, Vomund AN, Peterson OJ, Chervonsky AV, Lichti CF, Unanue ER: The MHC-II peptidome of pancreatic islets identifies key features of autoimmune peptides. Nat Immunol 2020;21:455-463
33. Tilvawala R, Nguyen SH, Maurais AJ, Nemmara VV, Nagar M, Salinger AJ, Nagpal S, Weerapana E, Thompson PR: The Rheumatoid Arthritis-Associated Citrullinome. Cell Chem Biol 2018;25:691-704.e696
34. Papayannopoulos V: Neutrophil extracellular traps in immunity and disease. Nat Rev Immunol 2018;18:134-147
35. Mintz PJ, Kim J, Do KA, Wang X, Zinner RG, Cristofanilli M, Arap MA, Hong WK, Troncoso P, Logothetis CJ, Pasqualini R, Arap W: Fingerprinting the circulating repertoire of antibodies from cancer patients. Nat Biotechnol 2003;21:57-63
36. Lu MC, Lai NS, Yu HC, Huang HB, Hsieh SC, Yu CL: Anti-citrullinated protein antibodies bind surface-expressed citrullinated Grp78 on monocyte/macrophages and stimulate tumor necrosis factor alpha production. Arthritis Rheum 2010;62:1213-1223
37. Crane ED, Al-Hashimi AA, Chen J, Lynn EG, Won KD, Lhotak S, Naeim M, Platko K, Lebeau P, Byun JH, Shayegan B, Krepinsky JC, Rayner KJ, Marchio S, Pasqualini R, Arap W, Austin RC: Anti-GRP78 autoantibodies induce endothelial cell activation and accelerate the development of atherosclerotic lesions. JCI insight 2018;3:e99363

TABLES

GRP78-298R (R version preferred) *Enterobacteriaceae bacterium* acetyl-CoA C-acyltransferase UniProt A0A3C2BKS9	**ALSSQHQARI** *ALSSQH-ARI* ALSSQHKARI
GRP78-360R (R version preferred) *Proteobacteria* DnaK chaperone UniProt A0A2G6J6V3	**VLVGGSTRI** *VLVGGSTRI* VLVGGSTRI
GRP78-435R (R version preferred) *Clostridium species* restriction endonuclease subunit S UniProt A0A3R6J648, R7CB01, A0A373M041	**KLIPRNTVV** *KLIPRNTVV* KLIPRNTVV
GRP78-502X (X version preferred) *Escherichia coli* DnaK chaperone UniProt P0A6Y8	**EIDVNGILXV** *+ID+-GIL-V* DIDADGILHV

Table 1. GRP78 mimotopes from human intestinal commensal bacteria. Homologous bacterial sequences were searched using BLAST. The alignment with GRP78 epitope sequences (in bold) is shown in italics, where – and + indicate non-conservative and conservative substitutions, respectively. The human GRP78 sequence is from UniProt P11021.

FIGURE LEGENDS

Figure 1. HLA-A2 binding of native (R) and citrullinated (X) GRP78 epitopes. (A) Epitope candidates (shown in bold within the GRP78 sequence, UniProt P11021) were selected *in silico* based on a NetMHC 4.0 predicted $K_D \leq 300$ nM or a SYFPEITHI score ≥20 (www.cbs.dtu.dk/services/NetMHC; www.syfpeithi.de). For length variants of the same peptide, the one with the best binding affinity was retained. R residues comprised within the selected candidates and dibasic cleavage motifs are underlined. The GRP78 signal peptide is indicated in italics. The experimental HLA-A2 binding affinity measured for the R and X versions in the experiments shown in subsequent panels is reported in arbitrary units (AU). **(B)** Schematic view and set-up of the assay. Biotinylated monomeric HLA-A2 molecules are incubated with peptide and β2m. After capturing on streptavidin-coated beads, an anti-β2m Ab is added, followed by an AF488-labeled secondary Ab. The bead-associated fluorescence is therefore detected only if the test peptide supports the folding of the HLA-A2/β2m complex. Histograms (gated on single $AF488^+$ beads) depict representative flow cytometry staining obtained with increasing concentrations of HLA-A2 molecules folded with the non-binding $CHGA_{382\text{-}390}$ (HPVGEADYF) and the binding Flu $MP_{58\text{-}66}$ (GILGFVFTL) peptides (negative and positive control, respectively). The 2.5 nM concentration was retained for subsequent experiments based on the best separation of the positive and negative peaks. **(C)** Representative staining of the tested GRP78 peptides in their native (R, blue) or citrullinated versions (X, red). The R/X position within the peptide sequence, either outside (P1, P5) or inside (P8, P9) the C-terminal region, is indicated in each histogram. **(D)** Relative median fluorescence intensity (MFI) AU values for peptide-HLA-A2 complexes at 2.5 nM, normalized to the $ChgA_{382\text{-}390}$ negative control peptide. Data is expressed as median±SD of triplicate measurements and the R/X position is indicated for each peptide. $^{**}p \leq 0.003$ by Student t test.

Figure 2. Blood CD8$^+$ T cells recognize either R- or X-GRP78 peptide versions and display similar frequencies in type 1 diabetic and healthy adults. (A) Schematic view of the combinatorial analysis applied to GRP78 MMr$^+$CD8$^+$ T cells. Each peptide is loaded onto HLA-A2 MMrs coupled with two different fluorochromes. Six fluorochromes are used to obtain 15 unique pairs. Consequently, each peptide-reactive cell population is stained with a unique pair of fluorochromes. Native (R) and citrullinated (X) versions of the same GRP78 peptide are stained with fluorochrome pairs that share one fluorochrome (green). By gating on double-MMr$^+$ events and excluding all other fluorochromes, cells recognizing only the R version or X version are visualized. By gating on triple-MMr$^+$ events, cells recognizing the R only, X only or both versions are visualized. Subsequent staining with CD45RA and CCR7 allows to characterize the naïve/memory phenotype of each population. **(B)** Representative examples for the GRP78-261 (top) and -360 peptide (bottom). This strategy is detailed in Supplementary Fig. 1-2. **(C)** Frequencies of MMr$^+$CD8$^+$ cells reactive exclusively to the GRP78-R (blue symbols) or GRP78-X peptides (red symbols), or to both (green symbols) in HLA-A2$^+$ type 1 diabetic (T1D, circles; n=16) and healthy adults (triangles; n=11). PPI_{6-14} and Flu-MP_{58-66} peptides were included as controls. At least 0.13x10^6 CD8$^+$ T cells were counted for each donor (median 5.4x10^6, range 0.13-280x10^6). **(D)** Percent antigen-experienced cells out of total MMr$^+$ cells for the GRP78 peptides depicted in (C). Data points with <5 MMr$^+$ cells were excluded (median 13 MMr$^+$ cells, range 5-524 for GRP78 peptides). Bars show median values. NA, not available (i.e. <5 MMr$^+$ cells counted). *p<0.05, **p<0.01, ***p<0.001 by Mann-Whitney U test.

Figure 3. Blood CD8$^+$ T cells recognize either R- or X-GRP78 peptide versions in type 1 diabetic adolescents. (A) Frequencies of MMr$^+$CD8$^+$ cells reactive exclusively to the GRP78-R (blue symbols) or GRP78-X peptides (red symbols), or to both (green symbols) in HLA-A2$^+$ type 1 diabetic (T1D) adolescents (n=8). PPI_{6-14} and Flu-MP_{58-66} peptides were included as

controls. At least $0.39x10^6$ $CD8^+$ T cells were counted for each donor (median $1.1x10^6$, range $0.39\text{-}2.4x10^6$). **(B)** Percent antigen-experienced cells out of total MMr^+ cells for the GRP78 peptides depicted in (A). Data points with <5 MMr^+ cells were excluded (median 15 MMr^+ cells, range 5-115 for GRP78 peptides). Bars show median values. NA, not available (i.e. <5 MMr^+ cells counted). *p<0.05, **p<0.01, ***p<0.001 by Mann-Whitney U test.

Figure 4. GRP78 peptides preferentially recognized in their native R version are cross-reactive with gut bacterial mimotopes. (A) Schematic view of the combinatorial analysis applied. Following a strategy similar to the one detailed in Figure 2A, three PBMC aliquots were stained with two MMr pairs sharing one fluorochrome (green). By gating on double or triple MMr^+ events, cells recognizing only one peptide or both are visualized, respectively. **(B-C)** Representative dot plots and cumulative results from healthy and type 1 diabetic donors (pie charts) obtained for GRP78-298R (B), GRP78-502X (C) and their respective peptide mimotopes. The first column shows the results obtained on the first T-cell aliquot (test sample) by crossing MMrs loaded with GRP78 and bacterial mimotope peptides, with the T-cell fraction recognizing both highlighted in green. The second column displays the cross-reactivity between GRP78 peptide and GRP78-360R or Flu $MP_{58\text{-}66}$ epitope (negative control; second PBMC aliquot). The third column displays the cross-reactivity between GRP78-360R or Flu $MP_{58\text{-}66}$ peptides charged on both MMr pairs (positive control; third PBMC aliquot). Dot plots and pie charts display the percent MMr^+ cells binding either MMr (blue or red) or both (green). In dot plots, the MMr^- population is displayed in black to visualize the position of MMr^+ events relative to it. **p=0.016 and *p=0.031 by Wilcoxon signed rank test for the comparison of cross-reactive (green) fractions and of mimotope-reactive (red) fractions, respectively, between GRP78-298R and -502X (pooled healthy and type 1 diabetic donors).

Figure 5. Pancreas-infiltrating CD8$^+$ T cells preferentially recognize the citrullinated GRP78-9X epitope. Pancreas sections from nPOD cases (Supplementary Table 2) were immunohistochemically stained *in situ* with MMrs loaded with the indicated GRP78 peptides, with a negative control melanocyte MelanA$_{26–35}$ peptide, and with positive control UCN3$_{1–9}$ and ZnT8$_{186–194}$ islet epitopes (low and high reactivity, respectively). **(A)** Representative staining with GRP78-9X MMrs (scale bar 100 μm) and higher magnification of the dotted area (scale bar 33 μm). **(B-D)** Number of MMr$^+$ cells/mm^2 section area of pancreas (B, D) and duodenal mucosa (C). Each point represents an individual case; bars indicate median values. *p<0.05; **p ≤0.01 by Mann-Whitney U test.

Figure 6. Gene expression of Padi citrullinating enzymes in murine mTECs and islets. (A, C, D) RT-qPCR expression (mean+SD) of *Padi2* in islets from 6-week-old and 10-week-old mice **(A)** and of *Padi2* **(C)** and *Padi4* **(D)** in mTECs from 6-8-week-old C57BL/6 and NOD mice ; *p≤0.03, **p≤0.005, ***p<0.001 by one-way ANOVA. **(B)** Padi enzymatic activity (mean+SD) in whole pancreata and thymi of 6-week-old C57BL/6 and NOD mice; ***p<0.001 by Student t test. **(E)** Expression (median+range) of *Padi* isoforms in WT and *Aire*$^{-/-}$ C57BL/6 mice. Murine mature MHC-II^{high} mTECs were prepared from 4-6-week-old mice (n=3/each) and analyzed by RNAseq; *p≤0.03, **p≤0.003, ***p<0.001 by Welch t test. **(F)** Enrichment of PTM enzyme gene sets in MHC-II^{high} mTECs from WT *vs Aire*$^{-/-}$ mice. Red lines indicate the 0.01 P_{FWER} cut-off value for statistical significance (for citrullination, P_{FWER}=0.0016, corresponding to a WT normalized enrichment score of 2.4).

Figure 7. Padi2 protein expression in murine mTECs. Confocal microscopy of frozen thymic tissue section from 6-8-week-old C57BL/6 WT and *Aire*$^{-/-}$ mice and NOD mice (scale bar 20 μm). mTECs are visualized by keratin-14 staining.

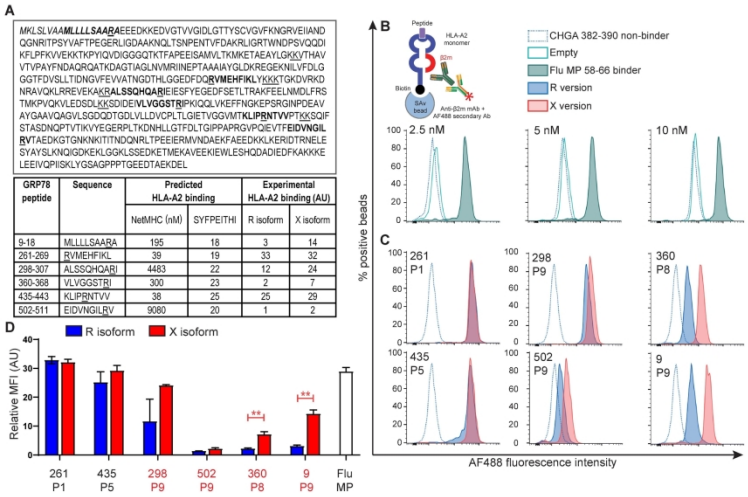


Figure 1. HLA-A2 binding of native (R) and citrullinated (X) GRP78 epitopes. (A) Epitope candidates (shown in bold within the GRP78 sequence, UniProt P11021) were selected in silico based on a NetMHC 4.0 predicted KD≤300 nM or a SYFPEITHI score ≥20 (www.cbs.dtu.dk/services/NetMHC; www.syfpeithi.de). For length variants of the same peptide, the one with the best binding affinity was retained. R residues comprised within the selected candidates and dibasic cleavage motifs are underlined. The GRP78 signal peptide is indicated in italics. The experimental HLA-A2 binding affinity measured for the R and X versions in the experiments shown in subsequent panels is reported in arbitrary units (AU). **(B)** Schematic view and set-up of the assay. Biotinylated monomeric HLA-A2 molecules are incubated with peptide and β2m. After capturing on streptavidin-coated beads, an anti-β2m Ab is added, followed by an AF488-labeled secondary Ab. The bead-associated fluorescence is therefore detected only if the test peptide supports the folding of the HLA-A2/β2m complex. Histograms (gated on single AF488+ beads) depict representative flow cytometry staining obtained with increasing concentrations of HLA-A2 molecules folded with the non-binding CHGA382-390 (HPVGEADYF) and the binding Flu MP58-66 (GILGFVFTL) peptides (negative and positive control, respectively). The 2.5 nM concentration was retained for subsequent experiments based on the best

separation of the positive and negative peaks. **(C)** Representative staining of the tested GRP78 peptides in their native (R, blue) or citrullinated versions (X, red). The R/X position within the peptide sequence, either outside (P1, P5) or inside (P8, P9) the C-terminal region, is indicated in each histogram. **(D)** Relative median fluorescence intensity (MFI) AU values for peptide-HLA-A2 complexes at 2.5 nM, normalized to the ChgA382-390 negative control peptide. Data is expressed as median±SD of triplicate measurements and the R/X position is indicated for each peptide. **p≤0.003 by Student's t test.

209x296mm (300 x 300 DPI)

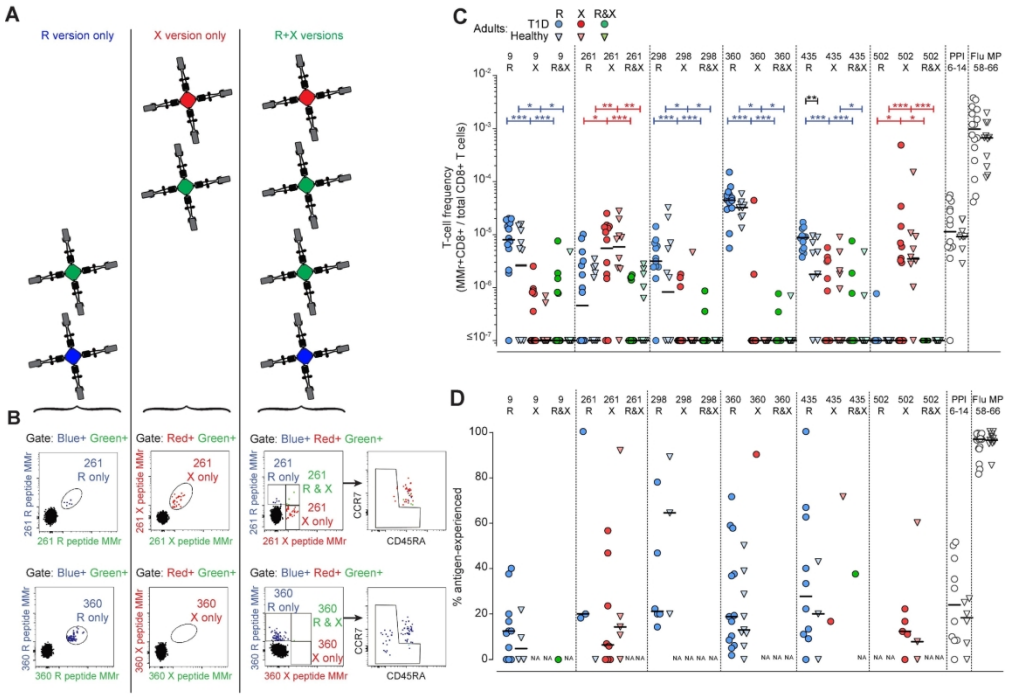


Figure 2. Blood CD8+ T cells recognize either R- or X-GRP78 peptide versions and display similar frequencies in type 1 diabetic and healthy adults. (A) Schematic view of the combinatorial analysis applied to GRP78 MMr+CD8+ T cells. Each peptide is loaded onto HLA-A2 MMrs coupled with two different fluorochromes. Six fluorochromes are used to obtain 15 unique pairs. Consequently, each peptide-reactive cell population is stained with a unique pair of fluorochromes. Native (R) and citrullinated (X) versions of the same GRP78 peptide are stained with fluorochrome pairs that share one fluorochrome (green). By gating on double-MMr+ events and excluding all other fluorochromes, cells recognizing only the R version or X version are visualized. By gating on triple-MMr+ events, cells recognizing the R only, X only or both versions are visualized. Subsequent staining with CD45RA and CCR7 allows to characterize the naïve/memory phenotype of each population. **(B)** Representative examples for the GRP78-261 (top) and -360 peptide (bottom). This strategy is detailed in Supplementary Fig. 1-2. **(C)** Frequencies of MMr+CD8+ cells reactive exclusively to the GRP78-R (blue symbols) or GRP78-X peptides (red symbols), or to both (green symbols) in HLA-A2+ type 1 diabetic (T1D, circles; n=16) and healthy adults (triangles; n=11). PPI6-14 and Flu-MP58-66 peptides were included as controls. At least 0.13×10^6 CD8+ T cells were counted for each donor (median 5.4×10^6, range $0.13\text{-}280 \times 10^6$). **(D)** Percent antigen-experienced cells out of total MMr+ cells for the GRP78 peptides depicted in (C). Data points with <5 MMr+ cells were excluded (median 13 MMr+ cells, range 5-524 for GRP78 peptides). Bars show median values. NA, not available (i.e. <5 MMr+ cells counted). *p<0.05, **p<0.01, ***p<0.001 by Mann-Whitney U test.

183x148mm (300 x 300 DPI)

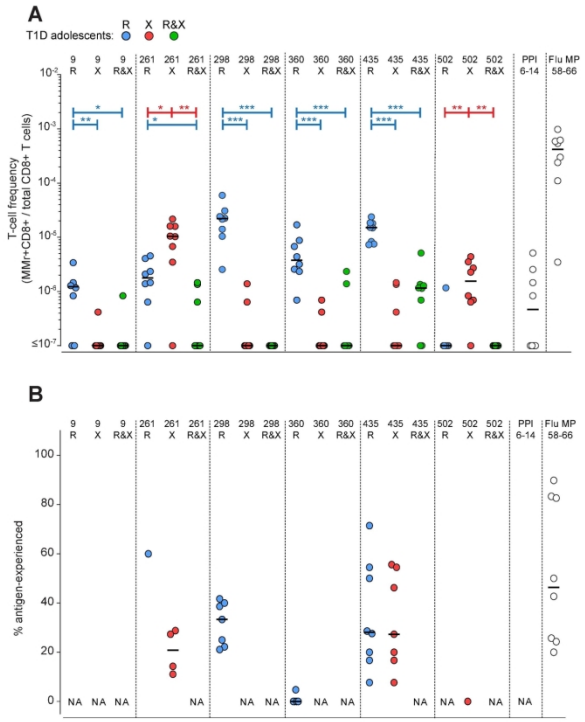


Figure 3. Blood CD8+ T cells recognize either R- or X-GRP78 peptide versions in type 1 diabetic adolescents. (A) Frequencies of MMr+CD8+ cells reactive exclusively to the GRP78-R (blue symbols) or GRP78-X peptides (red symbols), or to both (green symbols) in HLA-A2+ type 1 diabetic (T1D) adolescents (n=8). PPI6-14 and Flu-MP58-66 peptides were included as controls. At least $0.39x10^6$ CD8+ T cells were counted for each donor (median $1.1x10^6$, range $0.39\text{-}2.4x10^6$). **(B)** Percent antigen-experienced cells out of total MMr+ cells for the GRP78 peptides depicted in (A). Data points with <5 MMr+ cells were excluded (median 15 MMr+ cells, range 5-115 for GRP78 peptides). Bars show median values. NA, not available (i.e. <5 MMr+ cells counted). *p<0.05, **p<0.01, ***p<0.001 by Mann-Whitney U test.

183x148mm (300 x 300 DPI)

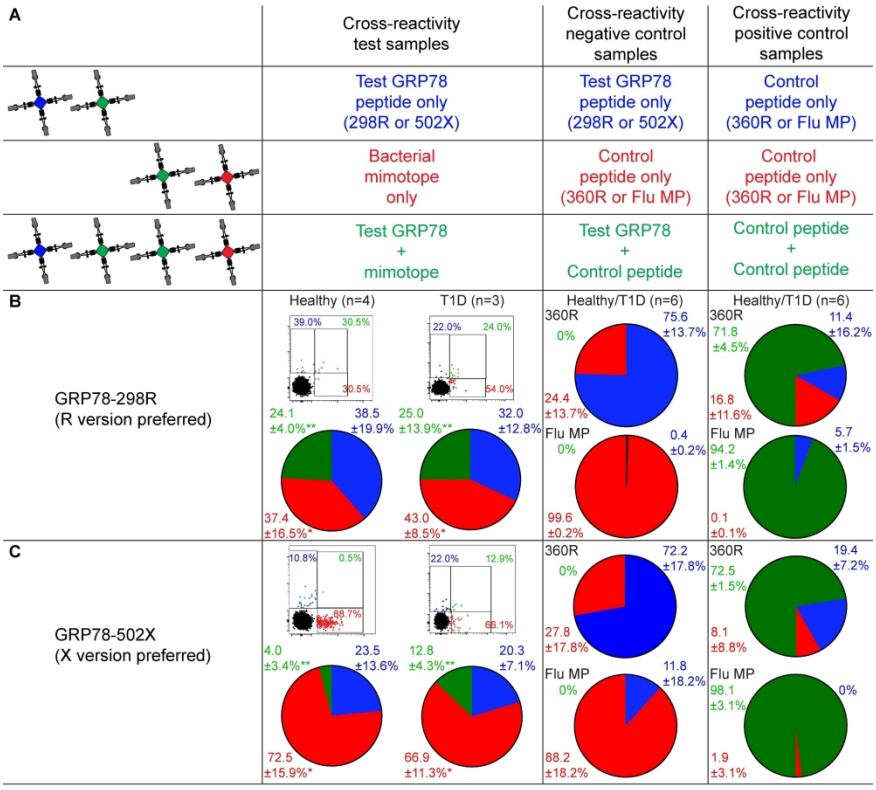


Figure 4. GRP78 peptides preferentially recognized in their native R version are cross-reactive with gut bacterial mimotopes. (A) Schematic view of the combinatorial analysis applied. Following a strategy similar to the one detailed in Figure 2A, three PBMC aliquots were stained with two MMr pairs sharing one fluorochrome (green). By gating on double or triple MMr+ events, cells recognizing only one peptide or both are visualized, respectively. **(B-C)** Representative dot plots and cumulative results from healthy and type 1 diabetic donors (pie charts) obtained for GRP78-298R (B), GRP78-502X (C) and their respective peptide mimotopes. The first column shows the results obtained on the first T-cell aliquot (test sample) by crossing MMrs loaded with GRP78 and bacterial mimotope peptides, with the T-cell fraction recognizing both highlighted in green. The second column displays the cross-reactivity between GRP78 peptide and GRP78-360R or Flu MP58-66 epitope (negative control; second PBMC aliquot). The third column displays the cross-reactivity between GRP78-360R or Flu MP58-66 peptides charged on both MMr pairs (positive control; third PBMC aliquot). Dot plots and pie charts display the percent MMr+ cells binding either MMr (blue or red) or both (green). In dot plots, the MMr^- population is displayed in black to visualize the

position of MMr+ events relative to it. **p=0.016 and *p=0.031 by Wilcoxon signed rank test for the comparison of cross-reactive (green) fractions and of mimotope-reactive (red) fractions, respectively, between GRP78-298R and -502X (pooled healthy and type 1 diabetic donors).

215x297mm (300 x 300 DPI)

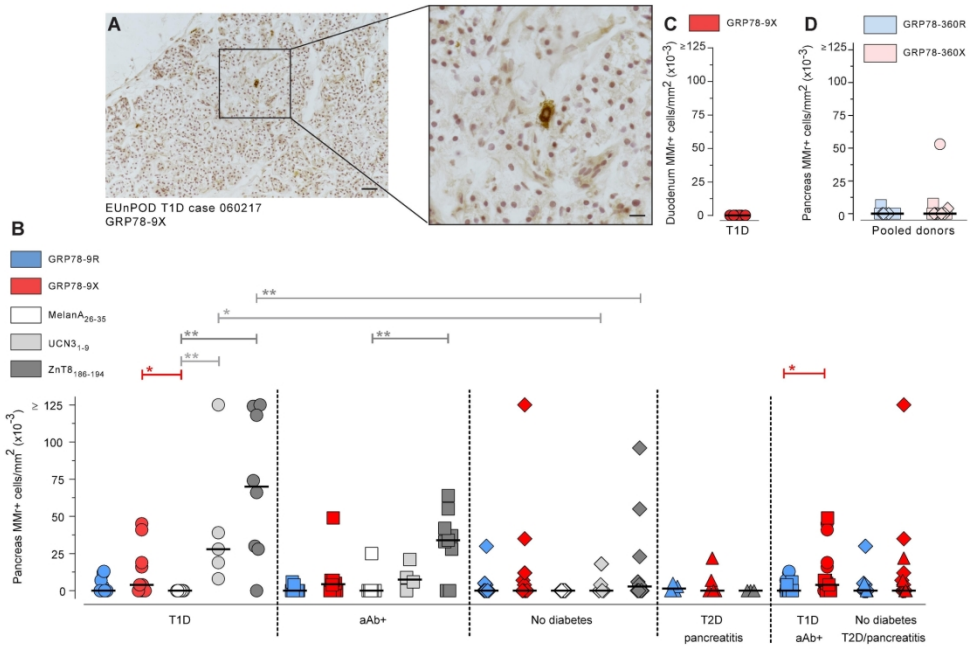


Figure 5. Pancreas-infiltrating CD8+ T cells preferentially recognize the citrullinated GRP78-9X epitope. Pancreas sections from nPOD cases (Supplementary Table 2) were immunohistochemically stained in situ with MMrs loaded with the indicated GRP78 peptides, with a negative control melanocyte MelanA26–35 peptide, and with positive control UCN31–9 and ZnT8186–194 islet epitopes (low and high reactivity, respectively). **(A)** Representative staining with GRP78-9X MMrs (scale bar 100 μm) and higher magnification of the dotted area (scale bar 33 μm). **(B-D)** Number of MMr+ cells/mm^2 section area of pancreas (B, D) and duodenal mucosa (C). Each point represents an individual case; bars indicate median values. *p<0.05; **p ≤0.01 by Mann-Whitney U test.

183x148mm (300 x 300 DPI)

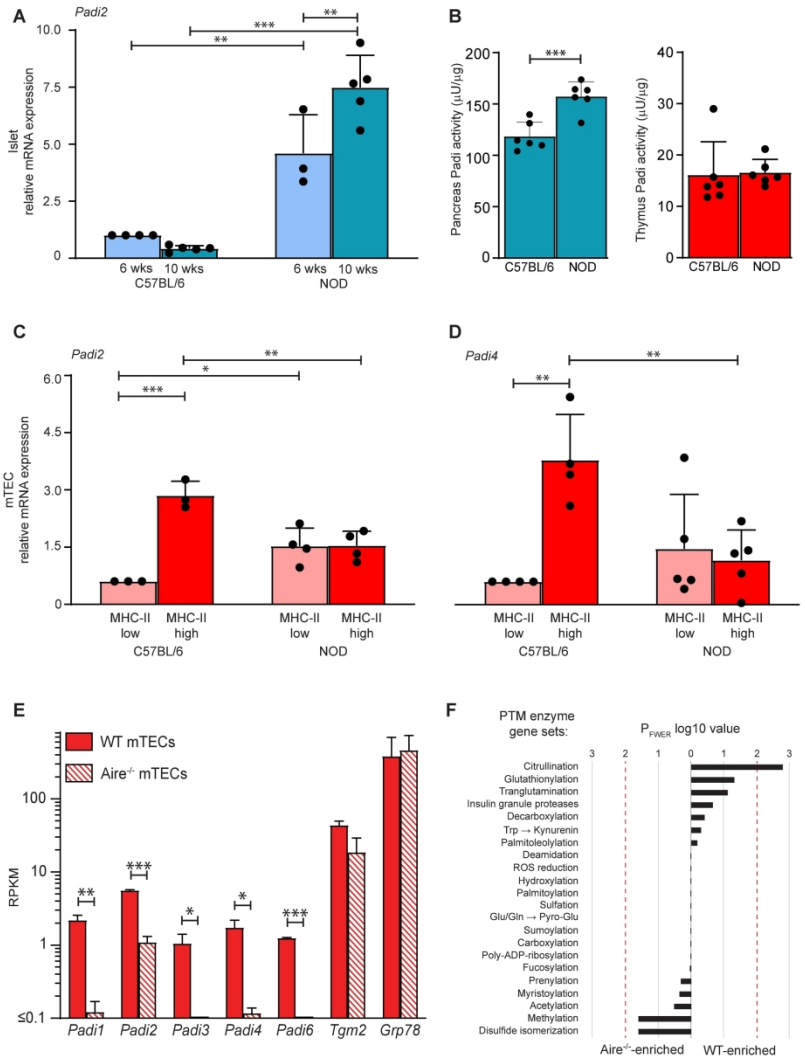


Figure 6. Gene expression of Padi citrullinating enzymes in murine mTECs and islets. (A, C, D) RT-qPCR expression (mean+SD) of Padi2 in islets from 6-week-old and 10-week-old mice (A) and of Padi2 (C) and Padi4 (D) in mTECs from 6-8-week-old C57BL/6 and NOD mice ; *p≤0.03, **p≤0.005, ***p<0.001 by one-way ANOVA. **(B)** Padi enzymatic activity (mean+SD) in whole pancreata and thymi of 6-week-old C57BL/6 and NOD mice; ***p<0.001 by Student t test. **(E)** Expression (median+range) of Padi isoforms in WT and Aire-/- C57BL/6 mice. Murine mature MHC-II^{high} mTECs were prepared from 4-6-week-old mice (n=3/each) and analyzed by RNAseq; *p≤0.03, **p≤0.003, ***p<0.001 by Welch t test. **(F)** Enrichment of PTM enzyme gene sets in MHC-II^{high} mTECs from WT vs Aire-/- mice. Red lines indicate the 0.01 P_{FWER} cut-off value for statistical significance (for citrullination, P_{FWER}=0.0016, corresponding to a WT normalized enrichment score of 2.4).

209x296mm (300 x 300 DPI)

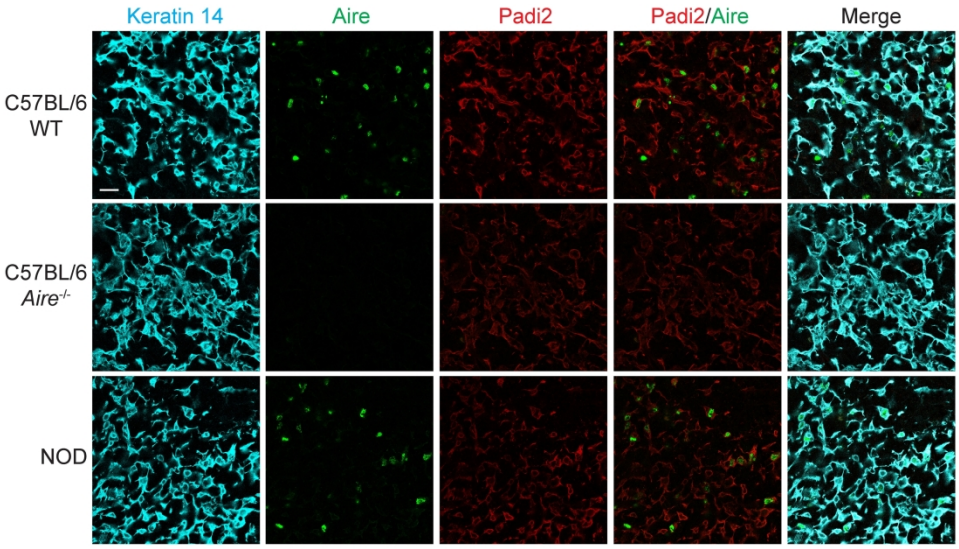


Figure 7. Padi2 protein expression in murine mTECs. Confocal microscopy of frozen thymic tissue section from 6-8-week-old C57BL/6 WT and Aire-/- mice and NOD mice (scale bar 20 μm). mTECs are visualized by keratin-14 staining.

296x209mm (300 x 300 DPI)

SUPPLEMENTARY MATERIAL

CD8+ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes

Marie Eliane Azoury[1*], Fatoumata Samassa[1*], Mijke Buitinga[2], Laura Nigi[3], Noemi Brusco[3], Aïsha Callebaut[2], Matthieu Giraud[4], Magali Irla[5], Ana Ines Lalanne[1], Alexia Carré[1], Georgia Afonso[1], Zhicheng Zhou[1], Barbara Brandao[1], Maikel L. Colli[6], Guido Sebastiani[3], Francesco Dotta[3], Maki Nakayama[7], Decio L. Eizirik[6,8], Sylvaine You[1], Sheena Pinto[9], Mark J. Mamula[10], Yann Verdier[11], Joelle Vinh[11], Soren Buus[12], Chantal Mathieu[2], Lut Overbergh[2] Roberto Mallone [1,13]

[1]Université de Paris, Institut Cochin, CNRS, INSERM, 75014 Paris, France.
[2]KU Leuven, Laboratory of Clinical and Experimental Endocrinology, 3000 Leuven, Belgium.
[3]University of Siena, Department of Medicine, Surgery and Neuroscience, Diabetes Unit and Fondazione Umberto di Mario ONLUS, Toscana Life Sciences, 53100 Siena, Italy.
[4]Université de Nantes, INSERM UMR1064, Centre de Recherche en Transplantation et Immunologie (CRTI), 44093 Nantes, France.
[5]Aix-Marseille University, CNRS, INSERM, CIML, Centre d'Immunologie de Marseille-Luminy, Marseille, France[6]Université Libre de Bruxelles, Center for Diabetes Research and Welbio, Medical Faculty, 1070 Brussels, Belgium.
[7]Barbara Davis Center for Childhood Diabetes, University of Colorado School of Medicine, Aurora, CO 80045, USA.
[8]Indiana Biosciences Research Institute, Indianapolis, IN 46202, USA.
[9]Deutsches Krebsforschungszentrum (DKFZ), Division of Developmental Immunology, 69120 Heidelberg, Germany.
[10]Yale University School of Medicine, New Haven, CT 06510, USA.
[11]ESPCI Paris, PSL University, Spectrométrie de Masse Biologique et Protéomique, CNRS USR3149, 75005 Paris, France.
[12]Panum Institute, Department of International Health, Immunology and Microbiology, D-2200 Copenhagen, Denmark.
[13]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, 75014 Paris, France.

*M.E.A. and F.S. contributed equally to this work.

SUPPLEMENTARY METHODS

mTEC isolation

For RT-qPCR and proteomics, thymic lobes were harvested from 6-8-week-old C57BL/6 mice and NOD mice, cleaned of connective tissue and fat, cut into fragments, washed in RPMI and gently agitated thrice for 30 s to release excess thymocytes. At each wash, the fragments were allowed to settle for 10 min on ice, followed by removal of the thymocyte-rich supernatant. Fragments were subsequently digested for 20 min at 37°C in ultra-low attachment 6-well plates (Corning) using 5 ml of ice-cold Liberase TH (2.5 U/mL; Roche) supplemented with 100 U/ml DNase I (Applichem). Supernatants were then removed and collected in ice-cold PBS supplemented with 0.5% BSA and 2mM EDTA. This enzymatic digestion was repeated twice for 15 min with 2.5 ml Liberase. Syringe needles (18G and 25G) were used to break up tissue after the second and third digestion, respectively. Cells were filtered through a 100-µm mesh before enrichment for TECs. To this end, $2x10^8$ cells/ml were incubated in PBS/2% FCS with 2.6 µg/ml anti-CD90.2 mAb (clone 30-H12, Sigma) on a rotating shaker for 30 min at 4°C, washed and plated ($1x10^7$ cells/ml) on panning plates previously coated overnight with 10 µg/ml goat anti-rat Ab (Sigma) in 50 mM Tris buffer pH 9.5. After 30 min at room temperature, cells were collected, washed and incubated with the Fixable Viability Dye eFluor780 (eBioscience). Cells ($40x10^6$/ml) were then stained for CD45-PE/Cy7 (RRID:AB_2734986), MHC Class II (MHC-II)-AF488 (RRID:AB_493138 and RRID:AB_493147 for C57BL/6 and NOD mice, respectively), EPCAM-BV421 (RRID:AB_2563983), BP-1-APC (RRID:AB_2762698), *Ulex europaeus* agglutinin-1 (UEA-1-biotin; Vector Laboratories), and streptavidin-PE (eBioscience). Cells ($40x10^6$/ml) were sorted with a BD Influx (100-µm nozzle, 18 psi) equipped with 488 nm (200 mW), 640 nm (120 mW), 405 nm (100 mW) and 561 nm (150 mW) lasers. Sorted cells were collected in 1.5-ml tubes in PBS/2% FCS.

For RNAseq, mouse thymi were extracted from a pool of 4 C57BL/6 mice, cut into small pieces, agitated into RPMI medium to release thymocytes, and digested with collagenase D (1 mg/ml; Roche) and DNase I (1 mg/ml; Sigma) for 30 min at 37 °C. The remaining fragments were digested with collagenase/dispase (2mg/ml; Roche) and DNase I (2 mg/ml) at 37 °C to obtain a cell suspension. After filtration through a 70-µm mesh, cells were resuspended in PBS containing 1% fetal bovine serum (FBS) and 5 mM EDTA. The remaining thymocytes were removed by magnetic depletion with CD45 MicroBeads (Miltenyi). Cells were then stained with CD45-PerCPCy5.5 (1:50; RRID:AB_893340), Ly51-PE (1:800; RRID:AB_313365), and MHC-II I-A/I-E-APC (1:1,200; RRID:AB_469455). MHC-II$^{high/low}$ mTECs

(CD45$^-$Ly51$^-$MHC-II$^{high/low}$) were sorted into TRIzol (ThermoFisher) on a BD FACSAria III instrument.

RT-qPCR and measurement of Padi enzymatic activity

Total RNA was extracted from sorted mTECs or 50 islets, using the Single Cell RNA Purification Kit (Norgen) and cDNA synthesized with SuperScript VILO (Invitrogen). The RT-qPCR forward (F) and reverse (R) primers were: *Padi2*-F: 5'-CCGCCGGGTATGAAATAGTCC-3', *Padi2*-R: 5'-CGCCGGTGTACTTGACCAC-3'; housekeeping *Act*-F: 5'-AGAGGGAAATCGTGCGTGAC-3', *Act*-R: 5'-CAATAGTGATGACCTGGCCGT-3'; housekeeping *Rpl27*-F: 5'-GTCGAGATGGGCAAGTTCAT-3', *Rpl27*-R: 5'-TTCTTCACGATGACGGCTTT-3'; and predesigned *Hprt* TaqMan primers (ThermoFisher). RT-qPCR was performed using 4 pmol primers, 0.2 µl cDNA and 5 µl Fast SYBR Green Master Mix (Applied Biosystems) on a StepOnePlus RT-PCR System (Applied Biosystems). The relative fold gene expression was calculated using the delta-delta Ct method.

Padi activity was measured with the Ab-based assay for Padi activity (ABAP; ModiQuest Research).

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

mTECs were washed thrice with PBS at 400g for 5 min and lysed in a buffer composed of 7 M urea, 2 M thiourea, 4% CHAPS, 40 mM Tris base, 1% dithiothreitol and Complete protease inhibitors (Roche). Debris were removed by centrifugation for 10 min at 13,000 rpm at 4°C. Samples were dialyzed and reduction, alkylation, protein precipitation, digestion and desalting performed as described (1). The purified peptides were vacuum-dried and dissolved in mobile phase A, containing 2% acetonitrile and 0.1% formic acid. An estimated final amount of 0.5 µg/µl was injected. The sample was separated by reversed-phase chromatography using a micropillar array column (µPAC^{TM} C18, 200 cm, PharmaFluidics). A linear gradient of mobile phase B (0.1% formic acid in 98% acetonitrile) from 1% to 40% in 80 min was followed by a steep increase to 100% mobile phase B in 5 min. After 5 min at 100% mobile phase B, a steep decrease to 1% mobile phase B was achieved in 5 min and 1% mobile phase B was maintained for 35 min with a flow rate of 750 nl per minute. LC-MS/MS was performed on a Q-Exactive Plus equipped with a nanospray ion source (ThermoFisher). Full-scan spectrum (350 to 1850 m/z, resolution 70,000, automatic gain control 3E6, maximum injection time 100 ms) was followed by high-energy collision-induced dissociation (HCD) tandem mass spectra

with a run time of 120 min. Peptide ions were selected for fragmentation by tandem MS as the 20 most intense peaks of a full-scan mass spectrum. HCD scans were acquired in the Orbitrap (resolution 17,500, automatic gain control 1E5, maximum injection time 80 ms). Peptides were identified by MASCOT (Matrix Science) using SwissProt (Homo sapiens, 169,779 entries) as a database via Proteome Discoverer 2.2, incorporating Percolator for peptide validation. Oxidation (M), deamidation (N/Q), and deamidation (R) (referring to citrullination), were included as variable modifications, carbamidomethylation (C) as a fixed modification. Two missed cleavages were allowed, peptide tolerance was set at 5 ppm and MS/MS tolerance at 20 mmu. MS/MS spectra were checked manually for the presence of citrullinated residues as described (2).

RNAseq and gene set enrichment analysis (GSEA)

Total RNA was extracted following the TRIzol protocol, adding GlycoBlue (ThermoFisher) as an RNA carrier. This RNA was used to generate poly-A-selected transcriptome libraries using the non-directional TruSeq V3 RNA Sample Prep Kit (Illumina) following the manufacturer's protocol. Sequencing was carried out on an Illumina HiSeq 2000 machine and was paired-end (2x100 bp) for MHC-II^{high} mTECs and single-end (50 bp) for MHC-II^{low} mTECs. These datasets have been deposited under GEO: (GSE submitted). RNAseq data from published (3) MHC-II^{high} and MHC-II^{low} mTEC (GSE140815) and from *Aire*$^{-/-}$ MHC-II^{high} mTEC (GSE140683) biological replicates were obtained with the same procedure.

Low-quality sequencing reads were removed using the Illumina CASAVA 1.8 pipeline and homogenized to 50-bp reads by trimming and merging the paired-end 100-bp reads. These reads were then aligned to the mm9 mouse reference genome using Bowtie (4). Read numbers were generated using the intersectBed with parameters -f 1.0 and the coverageBed programs of the BEDtool distribution (5), with a GTF annotation file that was generated from the UCSC Table Browser, in choosing mm9 mouse genome/Genes and Gene Predictions/RefSeq Genes/refGene. Reads per kilobase per million mapped reads (RPKM) values were then computed for each sample according to the transcript length of each gene and the global number of reads that map to the mm9 genome.

The enrichment of sets of genes driving PTMs between wild-type (WT) and *Aire*$^{-/-}$ MHC-II^{hi} mTECs and between MHC-II^{high} and MHC-II^{low} mTECs was analyzed with the GSEA software (6). To this end, the sequencing read numbers corresponding to each group of biological replicates in each comparison were normalized using DESeq (7) and used to run GSEA with a “classic” scoring scheme and n=10,000 permutations. The conservative family-wise error rate

statistics (P_{FWER}) was used to determine whether a set of genes was significantly over-represented in a group of mTECs versus another.

Microarray data from MHC-II^{high} mTECs isolated from 3 WT and 3 *Aire*$^{-/-}$ NOD mice were obtained from the NCBI GEO database (GSE12073) (8). CEL raw data files were processed with GenePattern (9) and the ExpressionFileCreator module for normalization using the RMA method (10) with background correction. Normalized data were plotted as fold change expression. False discovery rate *p* values were computed using the NCBI GEO2R online tool.

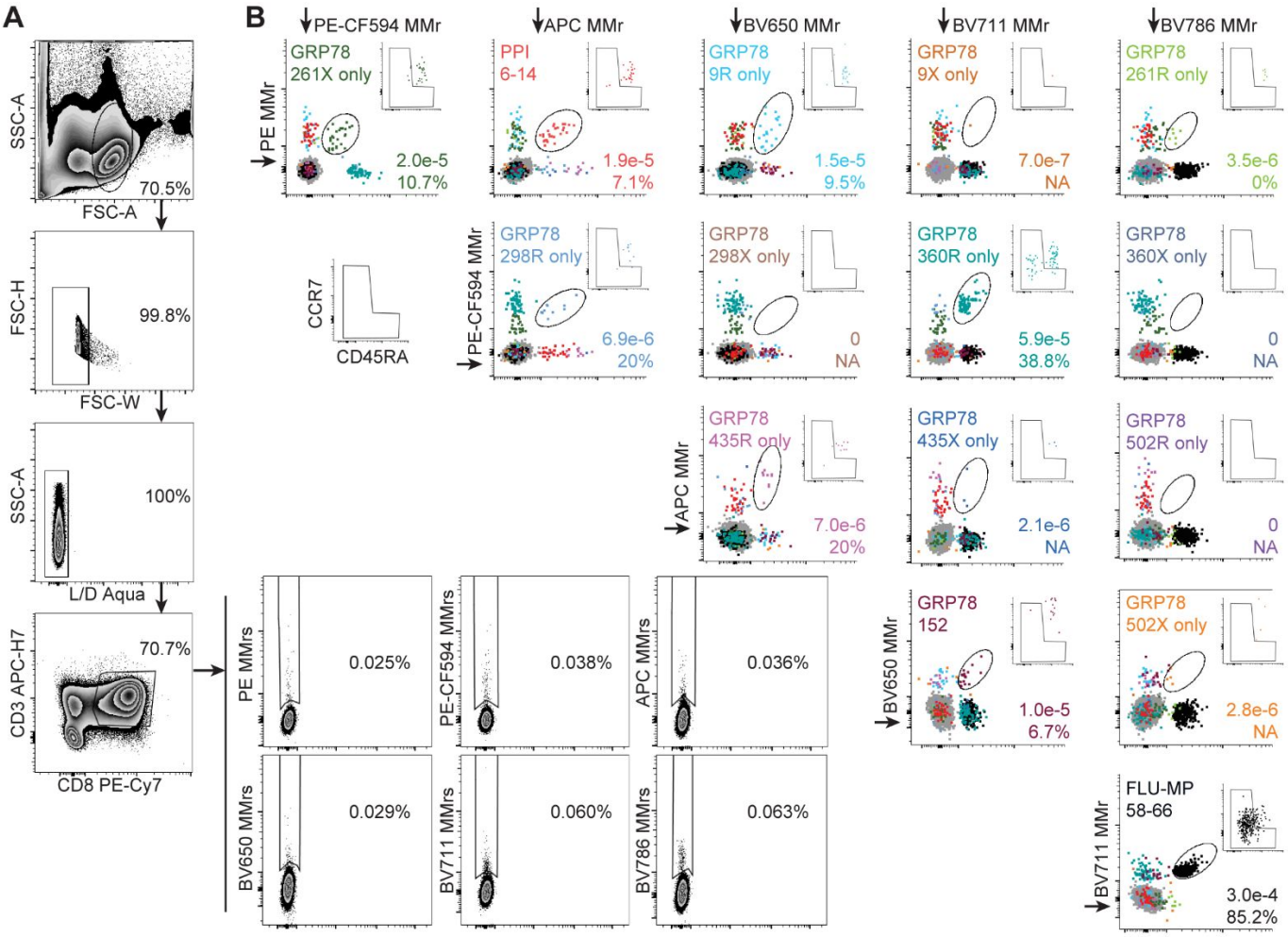
Enrichment of PTM enzyme gene sets between WT and *Aire*$^{-/-}$ NOD MHC-II^{high} mTECs was analyzed using GSEA on normalized microarray expression data using a "classic" scoring scheme and nominal *p* values.

Confocal microscopy

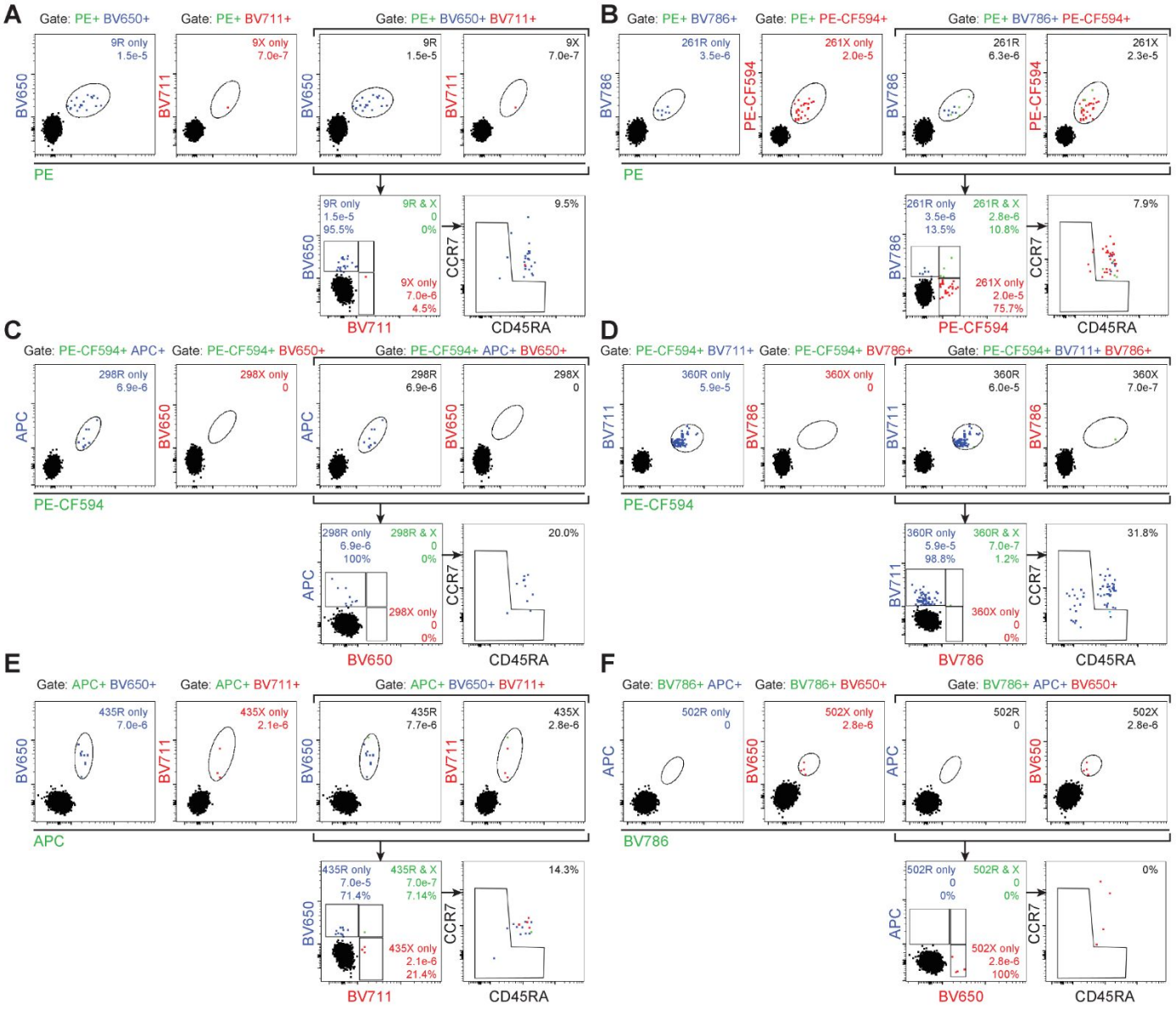
Frozen thymic tissue sections from 6-8-week-old WT and *Aire*$^{-/-}$ C57BL/6 mice and from NOD mice were fixed with 2% paraformaldehyde (Sigma) and incubated for 10 min in a saturation buffer composed of 0.1 M Tris pH 7.4, 5% mouse serum, 2% bovine serum albumin (Axday), 0.01% Triton X-100. Sections were then stained as previously described (11), with primary antibodies: rabbit anti-mouse keratin-14 (RRID:AB_2565048), mouse anti-PADI2 (RRID: AB_2881762), AF488-labeled rat anti-mouse Aire (RRID:AB_10854132); and secondary antibodies: Cy5-labeled goat anti-rabbit IgG (RRID: AB_2534032) and Cy3-labeled goat anti-mouse IgG (RRID:AB_893530). Sections were mounted with Mowiol (Calbiochem) and acquired on a LSM 780 confocal microscope (Carl Zeiss Microscopy).

Statistics

Significance was assessed with a cutoff value of $\alpha=0.05$ using two-tailed Student t test or Mann-Whitney U test according to normal or non-normal distribution, respectively. Variables displaying unequal variances were analyzed with the Welch t test. Paired datasets were analyzed with the Wilcoxon signed rank test.



Supplementary Figure 1. Gating strategy for the combinatorial analysis of GRP78 MMr$^+$CD8$^+$ T cells in T1D and healthy donors. (A) Frozen-thawed PBMCs from healthy donor C59 were magnetically depleted of CD8$^-$ cells before staining, acquisition and analysis. Cells were sequentially gated on small lymphocytes, singlets, live cells (Live/Dead Aqua$^-$), CD3$^+$CD8$^+$ T cells, and total PE$^+$, PE-CF594$^+$, APC$^+$, BV650$^+$, BV711$^+$, and BV786$^+$ MMr$^+$ T cells. Using Boolean operators, these latter gates allowed selective visualization of each double-MMr$^+$ population by including only those events positive for the corresponding fluorochrome pair (see Research Design and Methods). **(B)** The final readout obtained for the 15 peptides analyzed: 12 native/citrullinated GRP78 peptides, a GRP78$_{152\text{-}161}$ (GRP78-152) peptide without R residues, and PPI$_{6\text{-}14}$ and Flu-MP$_{58\text{-}66}$ controls). Events corresponding to each epitope-reactive population are overlaid in different colors within each plot, with MMr$^-$ events overlaid in light gray. The small dot plots on the right of each panel depict CD45RA (x axis) and CCR7 (y axis) expression in the corresponding MMr$^+$ fraction. Numbers in each panel indicate the MMr$^+$CD8$^+$ T-cell frequency out of total CD8$^+$ T cells and the percent naive (CD45RA$^+$CCR7$^+$) fraction among MMr$^+$ cells. This first gating strategy only allows to visualize CD8$^+$ T cells reactive to either the R or X isoform of each GRP78 peptides, and these T cells are therefore designated as 'R only' or 'X only', respectively. This strategy was therefore further modified to differentially visualize CD8$^+$ T cells reactive to R only, X only or both peptides (see Supplemental Figure 2).

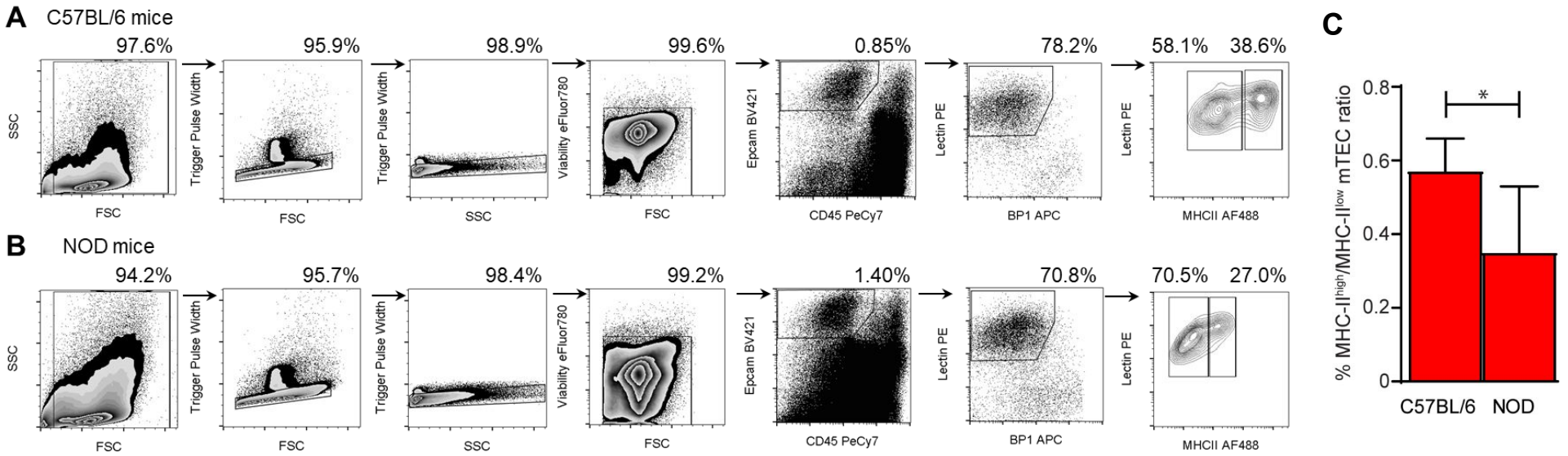


Supplementary Figure 2. Modified combinatorial gating strategy to analyze the cross-reactivity between R-GRP78 and X-GRP78 peptide isoforms. A representative example is shown for healthy donor C59. Each panel shows one of the peptides studied: GRP78-9 **(A)**, -261 **(B)**, -298 **(C)**, -360 **(D)**, -435 **(E)**, and -502 **(F)**.

First, $CD8^+$ T cells recognizing only the R-GRP78 (blue, 'R only') or X-GRP78 isoform (red, 'X only') were gated on double MMr^+ events, excluding the other four fluorochromes (first two dot plots of each panel, which are the same of Supplemental Fig. 1; MMr^- events overlaid in black). For example, in panel B '261R only' MMr^+ cells (PE^+BV786^+) were visualized by gating on $PE^+BV786^+PE\text{-}CF594^-APC^-BV650^-BV711^-$ events; and '261X only' MMr^+ cells ($PE^+PE\text{-}CF594^+$) were visualized by gating on $PE^+PE\text{-}CF594^+APC^-BV650^-BV711^-BV786^-$ events.

Second, $CD8^+$ T cells recognizing R-GRP78 or X-GRP78 peptides (i.e. including cross-reactive T cells) were gated on triple MMr^+ events, i.e. by excluding three fluorochromes but not the non-shared fluorochrome coding for the MMr loaded with the other peptide isoform (third and fourth dot plots of each panel). For example, in panel B '261R' and '261X' MMr^+ cells (i.e. cross-reactive or not; PE^+BV786^+ and $PE^+PE\text{-}$

CF594^{+}, respectively) were visualized by gating on PE^{+}BV786^{+}APC^{-}BV650^{-}BV711^{-} (i.e. not excluding PE-CF594 as before) and PE^{+}PE-CF594^{+}APC^{-}BV650^{-}BV711^{-} (i.e. not excluding BV786), respectively.
Third, these two dot plots are combined and visualized by plotting them for the two fluorochromes not shared by the R and X isoform (first dot plot in the bottom line of each panel). In this final dot plot, single-positive events correspond to the 'R only' and 'X only' fractions (blue and red, respectively), while double-positive events correspond to the 'R and X' fraction (i.e. cross-reactive; green). Events negative for all MMr fluorochromes (PE^{-}PE-CF594^{-}APC^{-}BV650^{-}BV711^{-}BV786^{-}) are overlaid in black in each dot plot to set the double- or triple-MMr^{+} gates. The bottom right dot plots of each panel depict CD45RA and CCR7 expression in the corresponding MMr^{+} fractions using the same blue, red and green color code for 'R only', 'X only' and 'R and X' MMr^{+} events, respectively.

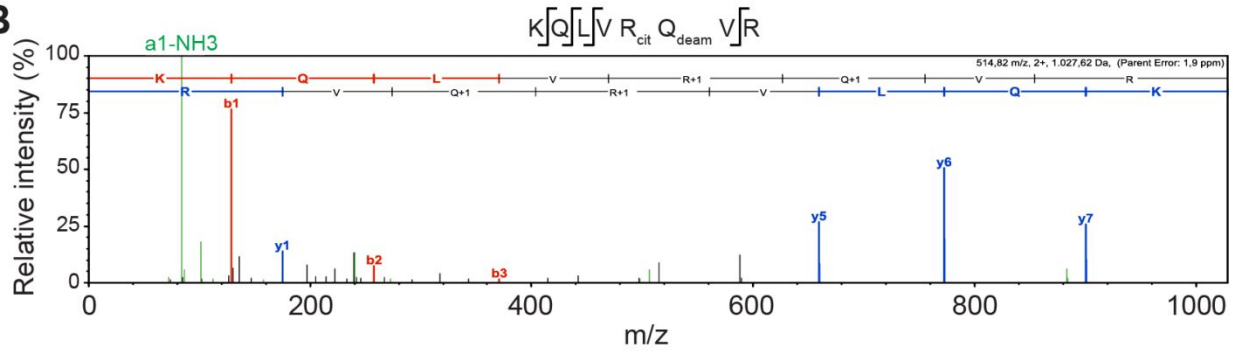


Supplementary Figure 3. Gating strategy and distribution of MHC-II^{low} and MHC-II^{high} mTECs in C57BL/6 and NOD mice. (A) mTECs from C57BL/6 mice. **(B)** mTECs from NOD mice. mTECs were gated as CD45^{-} EPCAM^{+} BP1$^{neg/low}$ Lectin (UEA-1)^{high} and MHC-II^{low} or MHC-II^{high}. **(C)** Ratio (median ± range) of MHC-II^{high}/MHC-II^{low} mTECs in 6-8-week-old C57BL/6 (n=4) and NOD mice (n=5); *p=0.032 by Mann-Whitney U test.

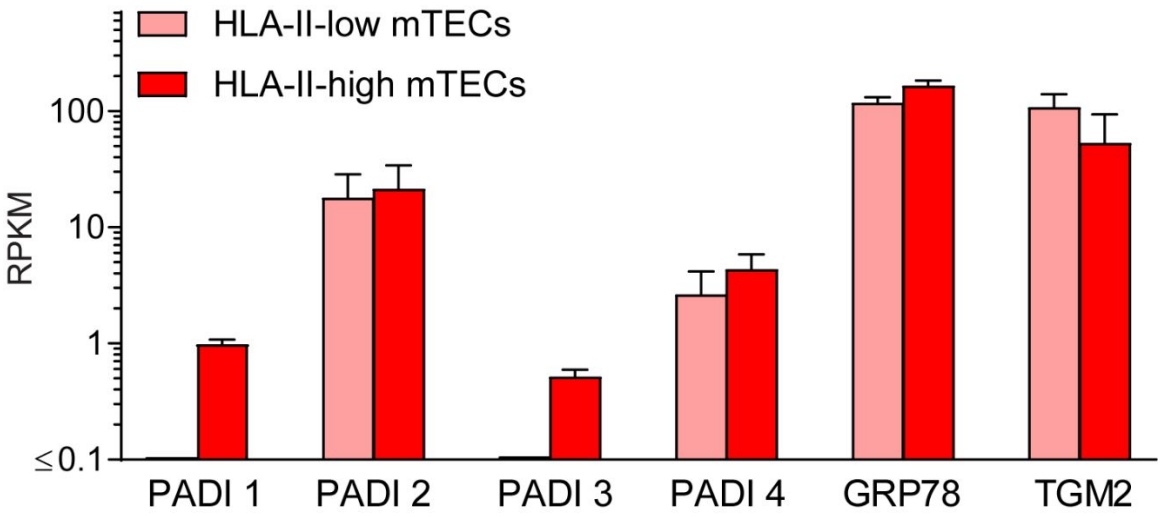
A

Name	Sequence	Modifications
EF-hand domain-containing protein 1	NGYAVVRRPTMGIGGDR	1xCitrullinated [R7]; 1xOxidation [M11]
Hepatocyte growth factor-regulated tyrosine kinase substrate	AAEEAERQR	1xCitrullinated [R7]; 1xDeamidated [Q8]
Myosin-9	KQLVRQVR	1xCitrullinated [R5]1xDeamidated [Q6]

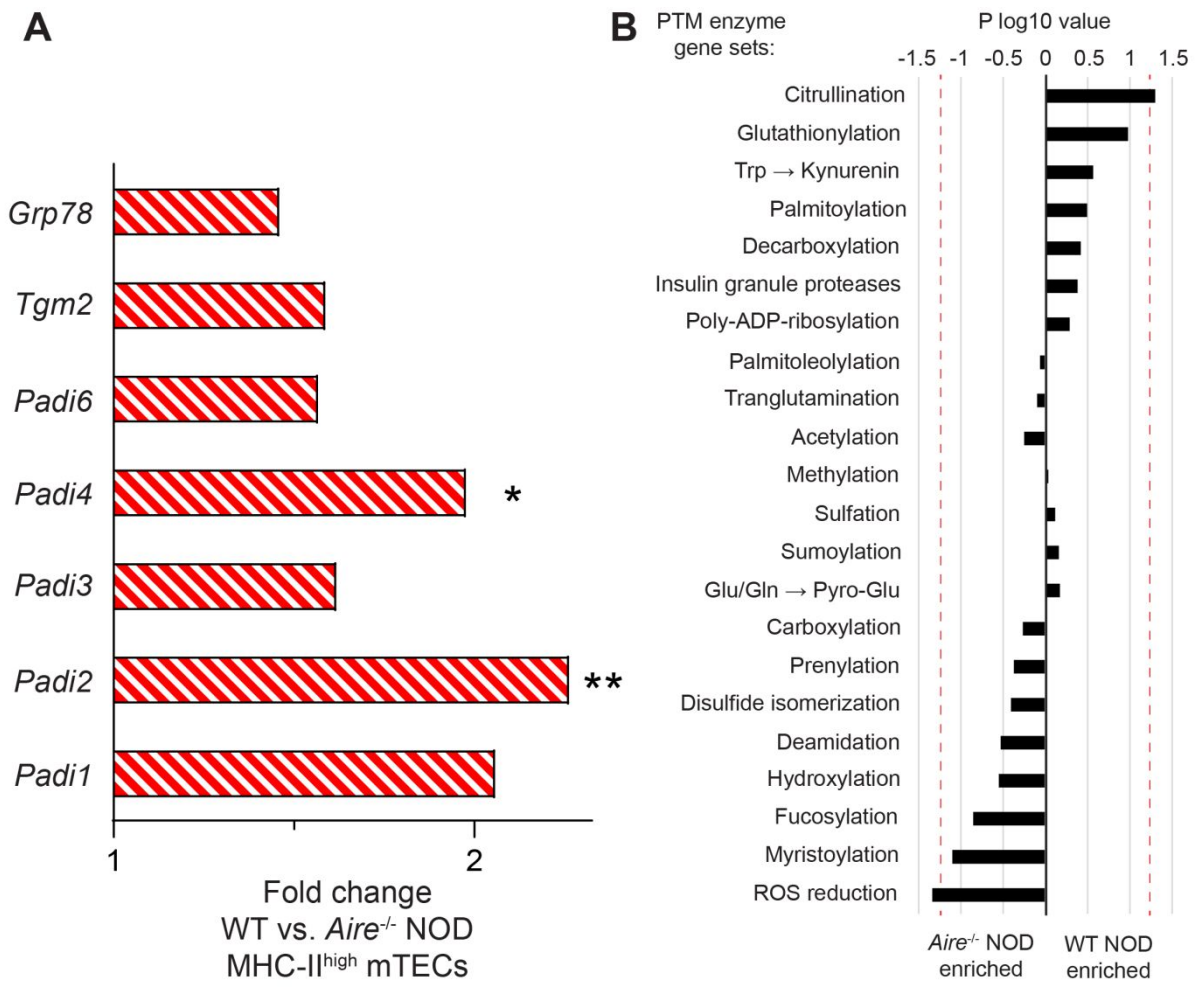
B



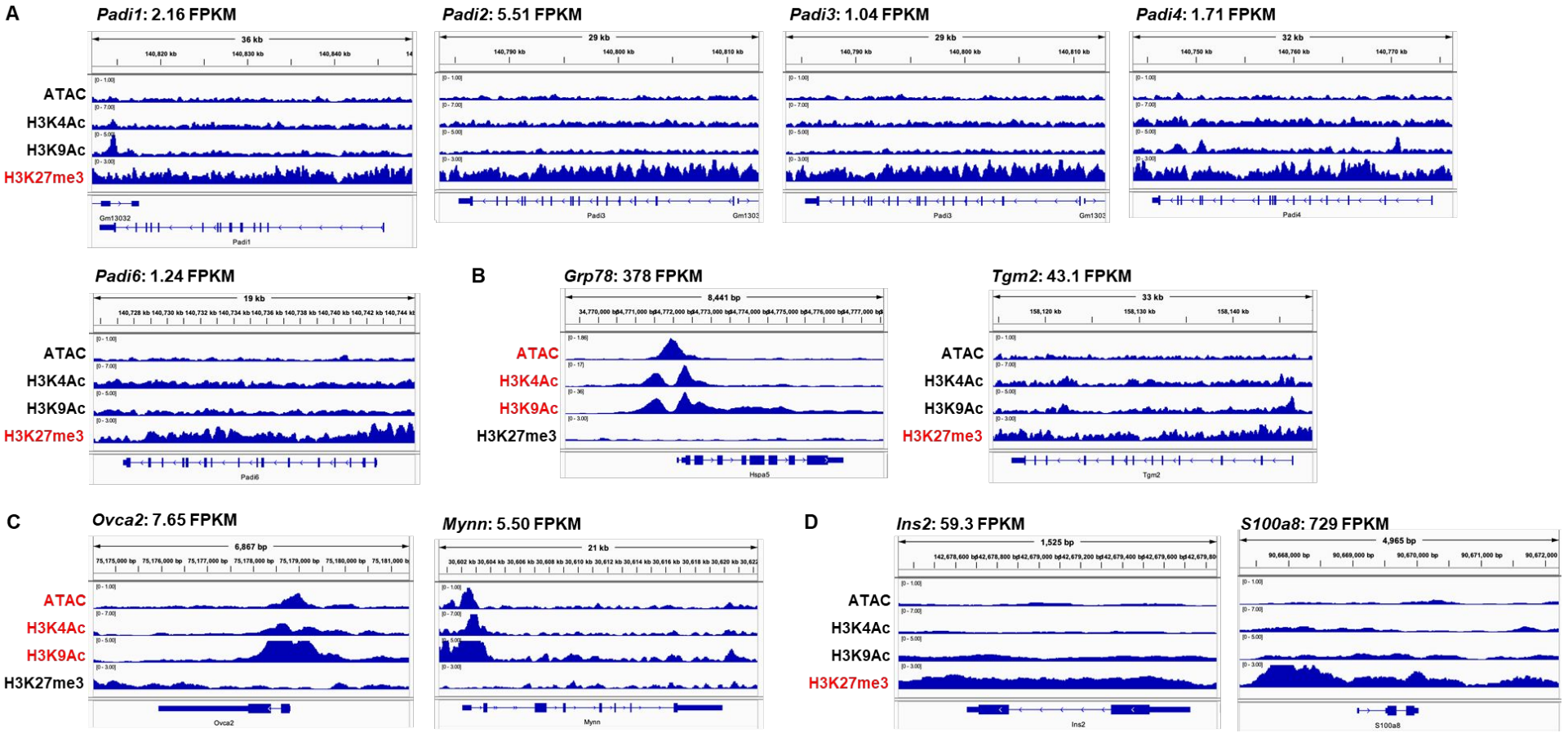
Supplementary Figure 4. Citrullinated peptides identified in mTECs of 10-week-old C57BL/6 mice. (A) List of the citrullinated peptides identified. **(B)** Spectrum of the R5-citrullinated myosin-9 peptide KQLVRQVR.



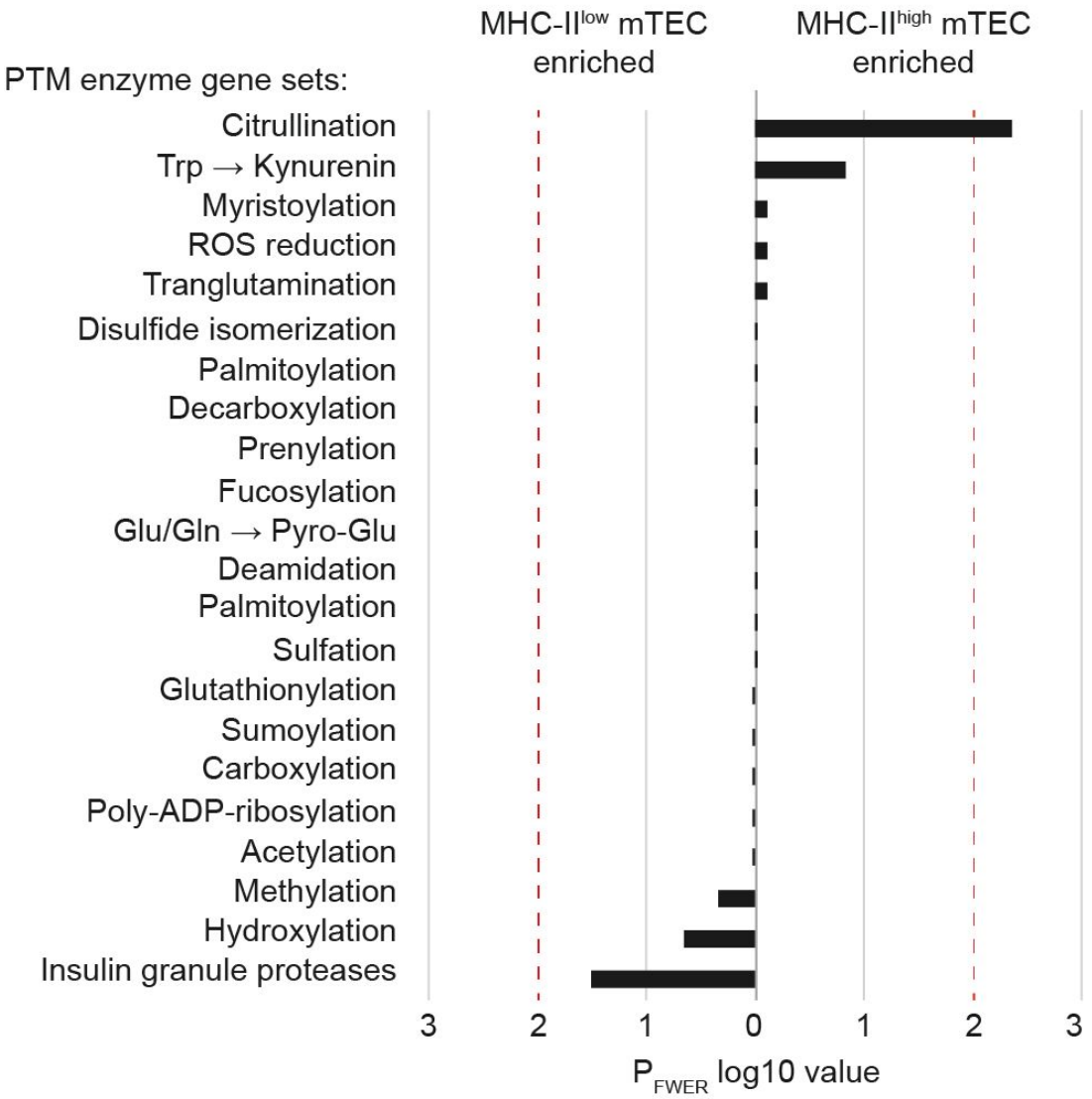
Supplementary Figure 5. Gene expression of PADI citrullinating enzymes in human immature HLA Class II (HLA-II)^{low} and mature HLA-II^{high} mTECs (n=3 donors). Data was extracted from our published RNAseq datasets (12).



Supplementary Figure 6. Enrichment of *Padi* isoform genes and PTM enzyme gene sets in mature MHC-II^{high} mTECs from WT vs. *Aire*$^{-/-}$ NOD mice (n=3/each). (A) Fold change expression of *Padi* isoform genes in MHC-II^{high} mTECs from 4-week-old WT vs. *Aire*$^{-/-}$ NOD mice analyzed by DNA microarray. **p*=0.04, ***p*=0.01 by false discovery rate method (GEO2R). **(B)** Enrichment of PTM enzyme gene sets in MHC-II^{high} mTECs from WT vs *Aire*$^{-/-}$ NOD mice. Red lines indicate the 0.05 *p* cut-off value for statistical significance [for citrullination, *p*=0.05, corresponding to a WT normalized enrichment score (NES) of 1.5]. Data is from GEO GSE12073 (8).



Supplementary Figure 7. ATAC-Seq and Chip-seq profiles of *Padi* genes in MHC-II^{high} mTECs from C57BL/6 mice. ATAC-seq and Chip-seq datasets (GSE114713) from Handel et al. (13) were analyzed. **(A)** *Padi* genes display three hallmarks of Aire-dependent expression: a closed chromatin ATAC profile (first row), near-absence of the activating chromatin modifications H3K4Ac and H3K9Ac (second and third row) and presence of the repressive modification H3K27me3 throughout the gene locus (fourth raw). **(B)** Conversely, the *Grp78* gene displays an open chromatin ATAC profile, presence of the activating chromatin modifications H3K4Ac and H3K9Ac at transcriptional start sites and absence of the repressive modification H3K27me3, in line with an Aire-independent expression. The *Tgm2* gene displays an intermediate profile. **(C-D)** Profiles of reference Aire-independent (C) and Aire-dependent genes (D) are shown as controls.



Supplementary Figure 8. Enrichment of PTM enzyme gene sets in immature MHC-II^{low} *vs* mature MHC-II^{high} mTECs from WT C57BL/6 mice (n=3). Red lines indicate the 0.01 P_{FWER} cut-off value for statistical significance (for citrullination, P_{FWER}=0.0045, corresponding to a MHC-II^{high} mTEC normalized enrichment score of 2.3).

	PE	PE-CF594	APC	BV650	BV711	BV786
PE		261X	PPI6-14	9R	9X	261R
PE-CF594			298R	298X	360R	360X
APC				435R	435X	502R
BV650					152	502X
BV711						Flu MP58-66
BV786						

Supplementary Table 1. Combinatorial HLA-A2 MMr panel used in Figure 2. See Supplementary Figures 1-2 for details.

	nPOD case	RRID	Sex	Age (yrs)	T1D T2D (yrs)	Positive aAbs	C-peptide (ng/ml)	Pancreas MMr+ cells						
								GRP78 9R	GRP78 9X	GRP78 360R	GRP78 360X	ZnT8 186-194	UCN3 1-9	MelA 26-35
T1D (n=10)	6070*	SAMN15879127	F	23	7	IA-2/mIAA	<0.05	0	16		53	74	39	0
	6113*	SAMN15879170	F	13	2	mIAA	<0.05	0	0		0			
	6161*	SAMN15879217	F	19	7	IA-2/mIAA	<0.05	12	45			124		0
	6211*	SAMN15879267	F	24	4	GAD/IA-2/ZnT8/mIAA	<0.05	3	4		0	30		0
	6212	SAMN15879268	M	20	5	mIAA	<0.05	0	0		0	0		
	6237	SAMN15879293	F	18	12	GAD/mIAA	<0.05	7	0		0	267	8	0
	6242	SAMN15879298	M	39	19	IA-2/mIAA	<0.05	0	0	0		66		0
	6258*	SAMN15879313	F	39	37	mIAA	<0.05	0	4	0		118	19	0
	6325*	SAMN15879379	F	20	6	GAD/IA-2	0.14	13	19	0		28	28	0
	060217	NA	F	39	21	GAD	NA	0	41	0			142	
aAb+ (n=9)	6080	SAMN15879137	F	69	NA	GAD/mIAA	1.84	0	0		0	55		25
	6101	SAMN15879158	M	65	NA	GAD	26.18	0	0		0	0		
	6123	SAMN15879180	F	23	NA	GAD	2.01	0			0	0		
	6151	SAMN15879207	M	30	NA	GAD	5.49	0	5		0	28	9	0
	6154	SAMN15879210	F	49	NA	GAD	<0.05	0	0	0		64	21	0
	6171	SAMN15879227	F	4	NA	GAD	8.95	0	7		8	37		0
	6347	SAMN15879401	M	9	NA	mIAA	3.26	6	49	7		33		0
	6388	SAMN15879441	F	25	NA	GAD/Miaa	1.38	0	7	0		34	0	0
	6397	SAMN15879450	F	21	NA	GAD	12.8	4	4	0		42	6	0
No diabetes (n=11)	6103	SAMN15879160	M	2	NA	—	0.98	30	0			55	0	0
	6174	SAMN15879230	M	21	NA	—	3.00	4	0		4			
	6179	SAMN15879235	F	20	NA	—	2.74	0	7		0	96	0	0
	6182	SAMN15879238	M	3	NA	—	2.28	0	4		0	23	4	0
	6227	SAMN15879283	F	17	NA	—	2.75	5	12		0	3	18	0
	6234	SAMN15879290	F	20	NA	—	6.89	0	0	0		6	0	0
	6271	SAMN15879325	M	17	NA	—	11.47	0	0	0		0		
	6287	SAMN15879341	F	57	NA	—	4.75	0	0	0		4		0
	6289	SAMN15879343	M	19	NA	—	8.05	0	0	0		0		
	6357	SAMN15879410	M	5	NA	—	8.82		128			0		
	6366	SAMN15879419	F	21	NA	—	0.41	0	35		0	0		
T2D/pancreatitis (n (n=4+1)	6028	SAMN15879085	M	33	17	—	22.40		0			0		
	6059	SAMN15879116	F	19	0.3	—	10.68	0	0			0		
	6273	SAMN15879327	F	45	2	—	3.17	3	7					
	6275	SAMN15879329	M	48	2	—	3.46	0	0			0		
	6439	SAMN15879492	M	27	NA	—	5.33	5	22					

Supplementary Table 2. nPOD cases analyzed by *in-situ* tissue MMr staining. The clinical characteristics of each case are reported along with the counts (x10^{-3}) of MMr^{+} cells/mm^2 pancreas section area for each of the indicated peptides. Positive sections are marked in red. Case #6287 (presenting a circumscribed neuroendocrine tumor in the pancreatic pan-body region; pan-tail region analyzed here) was classified as a non-diabetic control. T1D case #060217 is from the EUnPOD biobank. Control duodenum sections were also tested for T1D cases marked with an asterisk. NA, not applicable or not available; mIAA, micro-insulin aAbs. Results for epitopes other than GRP78 are from (12, 14).

PTM	Mouse gene	Mouse UniProt #	Human Uniprot #
Acetylation	Kat2a	Q9JHD2	Q92830
	Naa10	Q9QY36	P41227
	Naa11	Q3UX61	Q9BSU3
	Naa15	Q80UM3	Q9BXJ9
	Naa16	Q9DBB4	Q6N069
	Naa20	P61600	P61599
	Naa25	Q8BWZ3	Q14CX7
	Naa30	Q8CES0	Q147X3
	Naa35	Q6PHQ8	Q5VZE5
	Naa38	Q9D2U5	Q9BRA0
	Naa50	Q6PGB6	Q9GZZ1
	Naa60	Q9DBU2	Q9H7X0
	Nat8	Q9JIY7	Q9UHE5
	Nat9	Q3UG98	Q9BTE0
	Sirt1	Q923E4	Q96EB6
	Sirt2	Q8VDQ8	Q8IXJ6
	Sirt7	Q8BKJ9	Q9NRC8
Carboxylation	Ggcx	Q9QYC7	A7YA96
Cisteinylation (non-enzymatic)	n/a		
Citrullination	Padi1	Q9Z185	Q9ULC6
	Padi2	Q08642	Q9Y2J8
	Padi3	Q9Z184	Q9ULW8
	Padi4	Q9Z183	Q9UM07
	Padi6	Q8K3V4	Q6TGC4
Deamidation	Ntan1	Q64311	Q96AB6
	Ntaq1	Q80WB5	Q96HA8
Decarboxylation	Gad1	P48318	Q99259
	Gad2	P48320	Q05329
	Gadl1	Q80WP8	Q6ZQY3
Fucosylation	Fut1	O09160	P19526
	Fut10 (Sec1)	Q5F2L2	Q6P4F1
	Fut11	Q8BHC9	Q495W5
	Fut2 (Sec2)	Q9JL27	Q10981
	Fut3	n/a	P21217
	Fut4	Q11127	P22083
	Fut7	Q11131	Q11130
	Fut8	Q9WTS2	Q9BYC5
	Fut9	O88819	Q9Y231
	Pofut1	Q91ZW2	Q9H488
	Pofut2	Q8VHI3	Q9Y2G5
Glutathionylation	Ggt1*	Q60928	P19440
	Ggt3p*	n/a	A6NGU5
	Ggt5*	Q9Z2A9	P36269
	Ggt6	Q6PDE7	Q6P531
	Ggt7	Q99JP7	Q9UJ14
	Gsta1	P13745	P08263
	Gsta2	P10648	P09210
	Gsta3	P30115	Q16772
	Gsta4	P24472	O15217
	Gsta5	n/a	Q7RTV2
	Gstk1	Q9DCM2	Q9Y2Q3
	Gstm1	P10649	P09488
	Gstm2	P15626	P28161
	Gstm3	P19639	P21266
	Gstm4	Q8R5I6	Q03013
	Gstm5	P48774	P46439
	Gstm6	O35660	n/a
	Gstm7	Q80W21	n/a
	Gsto1	O09131	P78417
	Gsto2	Q8K2Q2	Q9H4Y5
	Gstp1	P19157	P09211
	Gstp2	P46425	n/a
	Gstp3	Q8VC73	n/a
	Gstt1	Q64471	P30711
Glutathionylation	Gstt2	Q61133	P0CG29
	Gstt2b	n/a	P0CG30
	Gstt3	Q99L20	n/a
	Gstt4	Q9D4P7	A0A1W2PR19
	Lancl1	O89112	O43813
	Mgst1	Q91VS7	P10620
	Mgst2	A2RST1	Q99735
	Mgst3	Q9CPU4	O14880
Hydroxylation	Asph	Q8BSY0	Q12797
	Egln1	Q91YE3	Q9GZT9
	Egln2	Q91YE2	Q96KS0
	Egln3	Q91UZ4	Q9H6Z9
	Jmjd6	Q9ERI5	Q6NYC1
	Kdm8	Q9CXT6	Q8N371
Malonylation (non-enzymatic)	n/a		
Methylation	Antkmt	Q501J2	Q9BQD7
	Atpsckmt	Q9D1Z3	Q6P4H8
	Camkmt	Q3U2J5	Q7Z624
	Ehmt2	Q9Z148	Q96KQ7
	Icmt	Q9EQK7	O60725
	Kmt5a	Q2YDW7	Q9NQR1
	Lcmt1	A2RTH5	Q9UIC8
	Lcmt2	Q8BYR1	O60294
	Mettl11b	B2RXM4	Q5VVY1
	Mettl18	Q9CZ09	O95568
	Mettl21a	Q9CQL0	Q8WXB1
	Mettl21c	Q8BLU2	Q5VZV1
	Mettl21e	Q8CDZ2	n/a
	Mettl22	Q8R1C6	Q9BUU2
	N6amt1	Q6SKR2	Q9Y5N5
	Ndufaf7	Q9CWG8	Q7L592
	Ntmt1 (Mettl11a)	Q8R2U4	Q9BV86
	Pcmt1*	P23506	P22061
	Pcmtd1*	P59913	Q96MG8
	Prmt1	Q9JIF0	Q99873
	Prmt2	Q9R144	P55345
	Prmt3	Q922H1	O60678
	Prmt4 (Carm1)	Q9WVG6	Q86X55
	Prmt5 (Anm5)	Q8CIG8	O14744
	Prmt6 (Anm6)	Q6NZB1	Q96LA8
	Prmt7 (Anm7)	Q922X9	Q9NVM4
	Prmt8 (Anm8)	Q6PAK3	Q9NR22
	Prmt9 (Anm9)	Q3U3W5	Q6P2P2
	Setd2	E9Q5F9	Q9BYW2
	Setd4	P58467	Q9NVD3
	Setd7	Q8VHL1	Q8WTS6

continued

Supplementary Table 3. Mouse genes coding for enzymes involved in PTMs. This list was used for the analysis of gene sets presented in Fig. 5G and Supplemental Fig. 5. PTMs involving very large sets of enzymes were excluded, namely glycosylation, phosphorylation, oxidation and ubiquitination. This list does not include enzymes described to selectively modify specific proteins or histones. Asterisks indicate enzymes indirectly involved in PTMs.

PTM	Mouse gene	Mouse UniProt #	Human Uniprot #
Myristoylation	Nmt1	O70310	P30419
	Nmt2	O70311	O60551
Nitration (non-enzymatic)	n/a		
Nitrosylation (non-enzymatic)	n/a		
Palmitoylation	Zdhhc1	Q8R0N9	Q8WTX9
	Zdhhc2	P59267	Q9UIJ5
	Zdhhc3	Q8R173	Q9NYG2
	Zdhhc4	Q9D6H5	Q9NPG8
	Zdhhc5	Q8VDZ4	Q9C0B5
	Zdhhc6	Q9CPV7	Q9H6R6
	Zdhhc7	Q91WU6	Q9NXF8
	Zdhhc8	Q5Y5T5	Q9ULC8
	Zdhhc9	P59268	Q9Y397
	Zdhhc11	Q14AK4	Q9H8X9
	Zdhhc11b	n/a	P0C7U3
	Zdhhc12	Q8VC90	Q96GR4
	Zdhhc13	Q9CWU2	Q8IUH4
	Zdhhc14	Q8BQQ1	Q8IZN3
	Zdhhc15	Q8BGJ0	Q96MV8
	Zdhhc16	Q969W1	Q969W1
	Zdhhc17	Q80TN5	Q8IUH5
	Zdhhc18	Q5Y5T2	Q9NUE0
	Zdhhc19	Q810M5	Q8WVZ1
	Zdhhc20	Q5Y5T1	Q5W0Z9
	Zdhhc21	Q9D270	Q8IVQ6
	Zdhhc22	A0PK84	Q8N966
	Zdhhc23	Q5Y5T3	Q8IYP9
	Zdhhc24	Q6IR37	Q6UX98
Palmitoleoylation	Porcn	Q9JJJ7	Q9H237
	Notum	Q8R116	Q6P988
Poly-ADP-ribosylation	Parp1	P11103	P09874
	Parp2	O88554	Q9UGN5
	Parp3	Q3ULW8	Q9Y6F1
	Parp4	E9PYK3	Q9UKK3
	Parp6	Q6P6P7	Q2NL67
	Parp8	Q3UD82	Q8N3A8
	Parp9	Q8CAS9	Q8IXQ6
	Parp10	Q8CIE4	Q53GL7
	Parp11	Q8CFF0	Q9NR21
	Parp12	Q8BZ20	Q9H0J9
	Parp14	Q2EMV9	Q460N5
	PARP15	n/a	Q460N3
	Parp16	Q7TMM8	Q8N5Y8
	Tiparp (Parp7)	Q8C1B2	Q7Z3E1
	Tnks (Parp5a)	Q6PFX9	O95271
	Tnks2 (Parp5b)	Q3UES3	Q9H2K2
Prenylation	Fnta	Q61239	P49354
	Fntb	Q8K2I1	P49356
	Pggt1b	A0A494BAX1	P53609
	Rce1	P57791	Q9Y256

PTM	Mouse gene	Mouse UniProt #	Human Uniprot #
Glu/Gln → Pyro-Glu	Qpct	Q9CYK2	Q16769
	Qpctl	Q8BH73	Q9NXS2
Disulfide isomerization	Creld1	Q91XD7	Q96HD1
	Creld2	Q9CYA0	Q6UXH1
	Ero1a*	Q8R180	Q96HE7
	Ero1b*	Q8R2E9	Q86YB8
	P4hb (Pdia1)	P09103	P07237
	Pdia2	D3Z6P0	Q13087
	Pdia3	P27773	P30101
	Pdia4	P08003	P13667
	Pdia5	Q921X9	Q14554
	Pdia6	Q922R8	Q15084
	Qsox1	Q8BND5	O00391
	Qsox2	Q3TMX7	Q6ZRP7
	Tmx3	Q8BXZ1	Q96JJ7
ROS reduction	Cat (Cata)*	P24270	P04040
	Cybb*	Q61093	P04839
	Mpo*	P11247	P05164
	Nos1*	Q9Z0J4	C9J5P6
	Nos2*	P29477	P35228
	Sod1*	P08228	P00441
	Sod2*	P09671	P04179
	Sod3*	O09164	P08924
Sulfation	Tpst1	O70281	O60507
	Tpst2	O88856	O60704
Sulfone (non-enzymatic)	n/a		
Sumoylation	Sae1	Q9R1T2	Q9UBE0
	Uba2 (Sae2)	Q9Z1F9	Q9UBT2
Transglutamination	Tgm1	Q9JLF6	P22735
	Tgm2	P21981	P21980
	Tgm3	Q08189	Q08188
	Tgm4	Q8BZH1	P49221
	Tgm5	Q9D7I9	O43548
	Tgm6 (Tgm3l)	Q8BM11	O95932
	Tgm7	A2ART8	Q96PF1
Trp → Kynurenine	Afmid	Q8K4H1	Q63HM1
	Ido1	P28776	P14902
	Ido2	Q8R0V5	Q6ZQW0
	Tdo2 (T23o)	P48776	P48775
Insulin granule proteases	Cpe	Q00493	P16870
	Cpn1	Q9JJN5	P15169
	Cpn2	Q9DBB9	P22792
	Ctsb	P10605	P07858
	Ctsd	P18242	P07339
	Ctsf	Q9R013	Q9UBX1
	Ctsl	P06797	P07711
	Dpp7	Q9ET22	Q9UHL4
	Furin	P23188	P09958
	Pcsk1	P63239	P29120
	Pcsk2	P21661	P16519
	Tpp1	O89023	O14773

SUPPLEMENTARY REFERENCES

1. Buitinga M, Callebaut A, Marques Camara Sodre F, Crevecoeur I, Blahnik-Fagan G, Yang ML, Bugliani M, Arribas-Layton D, Marre M, Cook DP, Waelkens E, Mallone R, Piganelli JD, Marchetti P, Mamula MJ, Derua R, James EA, Mathieu C, Overbergh L: Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. Diabetes 2018;67:2337-2348
2. Callebaut A, Derua R, Vig S, Delong T, Mathieu C, Overbergh L: Identification of Deamidated Peptides in Cytokine-Exposed MIN6 Cells through LC-MS/MS Using a Shortened Digestion Time and Inspection of MS2 Spectra. J Proteome Res 2021;20:1405-1414
3. Guyon C, Jmari N, Padonou F, Li YC, Ucar O, Fujikado N, Coulpier F, Blanchet C, Root DE, Giraud M: Aire-dependent genes undergo Clp1-mediated 3'UTR shortening associated with higher transcript stability in the thymus. eLife 2020;9
4. Langmead B, Trapnell C, Pop M, Salzberg SL: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10:R25
5. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841-842
6. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545-15550
7. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biol 2010;11:R106
8. Guerau-de-Arellano M, Martinic M, Benoist C, Mathis D: Neonatal tolerance revisited: a perinatal window for Aire control of autoimmunity. J Exp Med 2009;206:1245-1252
9. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP: GenePattern 2.0. Nat Genet 2006;38:500-501
10. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-264
11. Serge A, Bailly AL, Aurrand-Lions M, Imhof BA, Irla M: For3D: Full organ reconstruction in 3D, an automatized tool for deciphering the complexity of lymphoid organs. J Immunol Methods 2015;424:32-42
12. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, Lalanne AI, Sebastiani G, Carre A, Pinto S, Culina S, Corcos N, Bugliani M, Marchetti P, Armanet M, Diedisheim M, Kyewski B, Steinmetz LM, Buus S, You S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Dotta F, Scharfmann R, Eizirik DL, Verdier Y, Vinh J, Mallone R: Conventional and Neo-antigenic Peptides Presented by beta Cells Are Targeted by Circulating Naive CD8+ T Cells in Type 1 Diabetic and Healthy Donors. Cell Metab 2018;28:946-960 e946
13. Handel AE, Shikama-Dorn N, Zhanybekova S, Maio S, Graedel AN, Zuklys S, Ponting CP, Hollander GA: Comprehensively Profiling the Chromatin Architecture of Tissue Restricted Antigen Expression in Thymic Epithelial Cells Over Development. Front Immunol 2018;9:2120
14. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, Kuranda K, Nigi L, Eugster A, Osterbye T, Maugein A, McLaren JE, Ladell K, Larger E, Beressi JP, Lissina A, Appay V, Davidson HW, Buus S, Price DA, Kuhn M, Bonifacio E, Battaglia M, Caillat-Zucman S, Dotta F, Scharfmann R, Kyewski B, Mallone R, ImMaDiab Study G: Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 2018;3:eaao4013