



Potential Mimicry of Viral and Pancreatic β Cell Antigens Through Non-Spliced and *cis*-Spliced *Zwitter* Epitope Candidates in Type 1 Diabetes

Michele Mishto^{1,2*}, Artem Mansurkhodzhaev³, Teresa Rodriguez-Calvo^{4,5} and Juliane Liepe^{3*}

OPEN ACCESS

Edited by:

Arnaud Zaldumbide,
Leiden University Medical Center,
Netherlands

Reviewed by:

Jon D. Piganelli,
University of Pittsburgh, United States
Maki Nakayama,
University of Colorado Denver,
United States

*Correspondence:

Michele Mishto
michele.mishto@kcl.ac.uk
Juliane Liepe
jliepe@mpibpc.mpg.de

Specialty section:

This article was submitted to
Autoimmune and
Autoinflammatory Disorders,
a section of the journal
Frontiers in Immunology

Received: 20 January 2021

Accepted: 26 March 2021

Published: 15 April 2021

Citation:

Mishto M, Mansurkhodzhaev A,
Rodriguez-Calvo T and Liepe J
(2021) Potential Mimicry of
Viral and Pancreatic β Cell
Antigens Through Non-Spliced
and *cis*-Spliced *Zwitter* Epitope
Candidates in Type 1 Diabetes.
Front. Immunol. 12:656451.
doi: 10.3389/fimmu.2021.656451

¹ Centre for Inflammation Biology and Cancer Immunology (CIBC) & Peter Gorer Department of Immunobiology, King's College London, London, United Kingdom, ² Francis Crick Institute, London, United Kingdom, ³ Max-Planck-Institute for Biophysical Chemistry, Göttingen, Germany, ⁴ Institute of Diabetes Research, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich-Neuherberg, Germany, ⁵ German Center for Diabetes Research (DZD), Neuherberg, Germany

Increasing evidence suggests that post-translational peptide splicing can play a role in the immune response under pathological conditions. This seems to be particularly relevant in Type 1 Diabetes (T1D) since post-translationally spliced epitopes derived from T1D-associated antigens have been identified among those peptides bound to Human Leucocyte Antigen (HLA) class I and II complexes. Their immunogenicity has been confirmed through CD4⁺ and CD8⁺ T cell-mediated responses in T1D patients. Spliced peptides theoretically have a large sequence variability. This might increase the frequency of viral-human *zwitter* peptides, *i.e.* peptides that share a complete sequence homology irrespective of whether they originate from human or viral antigens, thereby impinging upon the discrimination between self and non-self antigens by T cells. This might increase the risk of autoimmune responses triggered by viral infections. Since enteroviruses and other viral infections have historically been associated with T1D, we investigated whether *cis*-spliced peptides derived from selected viruses might be able to trigger CD8⁺ T cell-mediated autoimmunity. We computed *in silico* viral-human non-spliced and *cis*-spliced *zwitter* epitope candidates, and prioritized peptide candidates based on: (i) their binding affinity to HLA class I complexes, (ii) human pancreatic β cell and medullary thymic epithelial cell (mTEC) antigens' mRNA expression, (iii) antigen association with T1D, and (iv) potential hotspot regions in those antigens. Neglecting potential T cell receptor (TCR) degeneracy, no viral-human *zwitter* non-spliced peptide was found to be an optimal candidate to trigger a virus-induced CD8⁺ T cell response against human pancreatic β cells. Conversely, we identified some *zwitter* peptide candidates, which may be produced by proteasome-catalyzed peptide splicing, and might increase the likelihood of pancreatic

β cells recognition by virus-specific CD8⁺ T cell clones, therefore promoting β cell destruction in the context of viral infections.

Keywords: T1D, antigen presentation, tolerance, mimicry, spliced peptides, virus

INTRODUCTION

Type 1 diabetes (T1D) is an autoimmune disease with a pivotal T cell activity. Indeed, CD4⁺ and CD8⁺ T cell-mediated responses in T1D patients play a prominent role in pancreatic β cell death, and the consequent insulin-dependent disease. CD4⁺ and CD8⁺ T cells recognize fragments (peptidic epitopes) of antigens, which are presented to T cell receptors (TCRs) $\alpha\beta$ through Human Leucocyte Antigen class I and class II (HLA-I and -II) complexes. Since autoreactive T cells selectively recognize antigenic epitopes specific for pancreatic β cells, part of the research in T1D sails in uncharted waters to discover neoantigens, epitopes and their presentation mechanisms, which can explain why CD4⁺ and CD8⁺ T cells build an autoreactive immune response in T1D. In the last decade, huge progress in mass spectrometry and bioinformatics has allowed the identification of unconventional antigenic peptides, *i.e.* peptides that could not be directly identified in the human proteome. Cryptic peptides derived from putative non-coding regions, usage of alternative open reading frames, as well as post-translational modifications emerged as a sizeable portion of the peptides that are presented by HLA-I and -II complexes to T cells (1–5). Among them, post-translationally spliced epitopes derived from T1D-associated antigens represent an attractive source of neoantigens. These peptides are produced by fusion of two non-contiguous peptide fragments of either an antigen – *i.e.* *cis*-spliced peptides – or two distinct antigens, *i.e.* *trans*-spliced

peptides (6) (Figures 1A, B). Hybrid insulin peptides (HIPs) identified by Delong and colleagues (7), belong to the latter category. Indeed, they are formed by the fusion of a splice-reactant of insulin, and another derived from other T1D-associated antigens. HIPs are presented by major histocompatibility complexes class II molecules (MHC-II) in nonobese diabetic (NOD) mice and by HLA-II (HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ, and HLA-DR) in humans (7–9). A handful of known HIPs can trigger a CD4⁺ T cell response both in NOD mice and in T1D patients (7–11). The enzymes (or biochemical reactions) catalyzing their production are not fully understood, although pioneer studies suggest that HIPs or HIPs' precursors might be produced in β cell's insulin crinosomes or professional antigen presenting cell (APC)'s lysosomes (9, 12, 13). Their identification, however, is still controversial, and the employment of different mass spectrometry data analysis strategies has led to contradictory results in HIPs' identification (12, 14).

Many more antigenic *cis*-spliced peptides have been identified in the context of HLA-I antigen processing and presentation (APP) pathway since their first appearance in the literature in 2004 (6, 15, 16). They can target *in vivo* CD8⁺ T cell responses against bacterial antigens, otherwise neglected in a mouse model of *Listeria monocytogenes* infection (17). *Cis*-spliced epitopes derived from either melanoma-associated or T1D-associated antigens are recognized by CD8⁺ T cells in peripheral blood of melanoma patients and T1D patients, respectively (18–20).

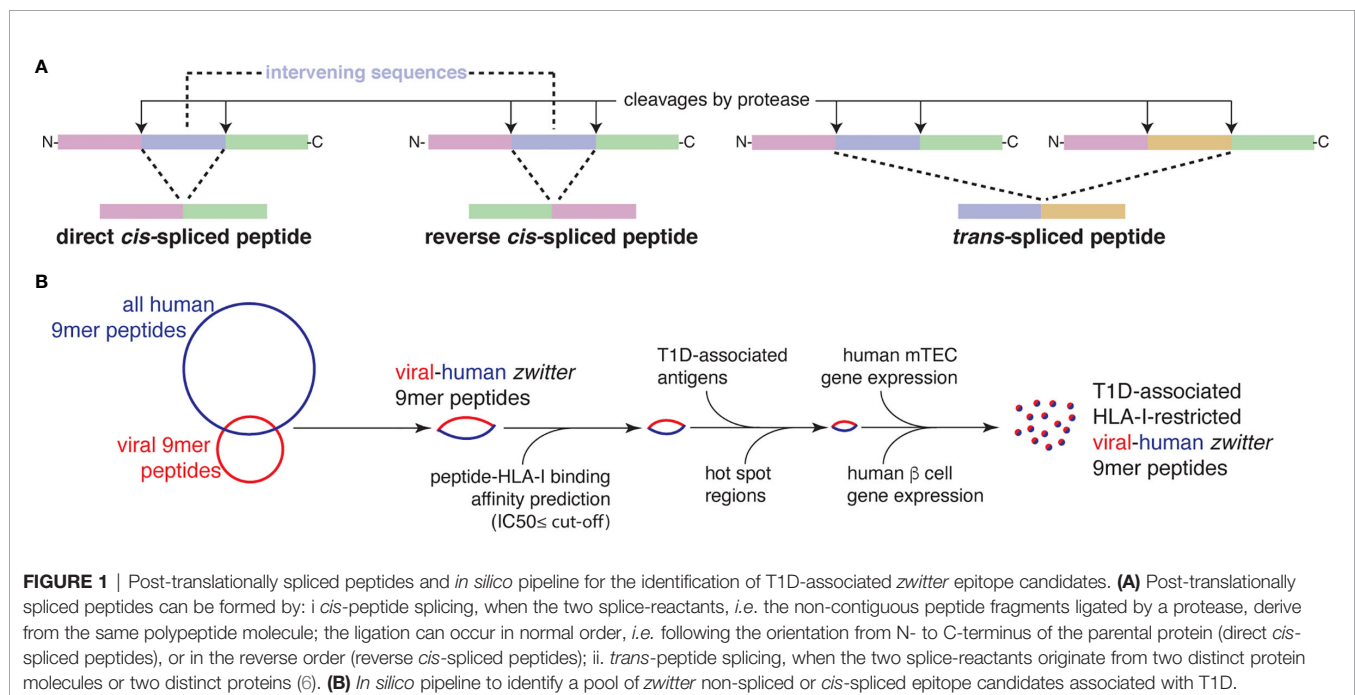


FIGURE 1 | Post-translationally spliced peptides and *in silico* pipeline for the identification of T1D-associated zwitter epitope candidates. **(A)** Post-translationally spliced peptides can be formed by: i *cis*-peptide splicing, when the two splice-reactants, *i.e.* the non-contiguous peptide fragments ligated by a protease, derive from the same polypeptide molecule; the ligation can occur in normal order, *i.e.* following the orientation from N- to C-terminus of the parental protein (direct *cis*-spliced peptides), or in the reverse order (reverse *cis*-spliced peptides); ii. *trans*-peptide splicing, when the two splice-reactants originate from two distinct protein molecules or two distinct proteins (6). **(B)** *In silico* pipeline to identify a pool of zwitter non-spliced or *cis*-spliced epitope candidates associated with T1D.

Cis-spliced as well as non-spliced peptides bound to HLA-I complexes are mainly produced by proteasomes (16, 18, 21). During *in vitro* digestions of synthetic polypeptides, proteasomes seem to produce almost as many *cis*-spliced as non-spliced peptides in terms of number of different peptide products (22). However, on average, *cis*-spliced peptides are generated in lower amounts than non-spliced peptides (23, 24), thereby suggesting that peptide splicing is catalyzed by proteasomes less efficiently than peptide hydrolysis. The identification of *cis*-spliced peptides among peptides bound to HLA-I complexes - *i.e.* HLA-I immunopeptidomes - is at least as controversial as that of HIPs; depending on the method used for *cis*-spliced peptide identification, their frequency in HLA-I immunopeptidomes was estimated to range from 1% to 34% (25).

The large variety of *cis*-spliced peptide sequences might have been problematic for models of self/non-self discrimination and tolerance of CD8⁺ T cells (26). Indeed, the vast human and virus *cis*-spliced peptide database might significantly increase the theoretical number of viral-human *zwitter* peptides. *Zwitter* is the German word for hybrid. It is etymologically derived from *zwi-*, “duplex”. For example, in chemistry, a zwitterion is an ion which possesses both positively- and negatively-charged groups. We use the term ‘*zwitter* peptides’ to describe peptides with a sequence that may be generated from viral antigens as well as from human antigens (6, 27). A large number of viral-human *zwitter cis*-spliced peptides might impinge upon CD8⁺ T cell repertoires and the recognition of viral antigens through central and peripheral tolerance. Nevertheless, according to a preliminary *in silico* study, this does not seem to be a frequent occurrence (27). This is in part due to the fact that only a tiny fraction of all theoretical *cis*-spliced peptides is *de facto* produced by proteasomes, as suggested by *in vitro* digestions of synthetic polypeptides with proteasomes measured by mass spectrometry (22).

Nonetheless, even a few viral-human *zwitter (cis*-spliced) epitopes may act as the targets of a CD8⁺ T cell autoreactive response, triggered by a viral infection. A growing number of immunological studies suggest that thymic clonal deletion prunes but does not completely eliminate autoreactive CD8⁺ T cells (28–30). Therefore, naïve CD8⁺ T cell clones specific to viral-human *zwitter* epitopes may circulate in the body and be controlled by peripheral tolerance. Central and peripheral tolerance, however, does not completely eliminate the risk that some of these naïve CD8⁺ T cell clones may recognize the cognate epitopes during a viral infection, be primed, proliferate, and then carry out an autoreactive response, for instance during a second or multiple viral infections. This might be the case of T1D, wherein CD8⁺ T cells could recognize viral-human *zwitter* epitopes which are also presented by pancreatic β cells. T1D has been historically associated with viral infections and especially with enteroviruses (EVs), such as Coxsackievirus B1 and B4 (CVB1 and CVB4; **Table S1**) (31). However, to date, the detection of circulating or infiltrating EV-reactive T cells has been challenging. CD8⁺ T cells predominate in the islets, many of which express high levels of HLA-I molecules at disease onset, likely contributing to exacerbated antigen presentation. EVs are known to infect pancreatic β cells and induce an

inflammatory response in the islets (32). Sequence similarity has been previously described through the study of canonical non-spliced epitope candidates, between the 2C non-structural protein (P2C) of CVB4 and the glutamate decarboxylase 2 (GAD2; a.k.a. GAD65) antigen, which is predominantly expressed in pancreatic β cells, although no evidence of cross-recognition by CD8⁺ T cell clones was demonstrated in that report (33). Other members of the Picornaviridae family, such as Parechoviruses (HPeV2), have also been associated to T1D (34). Similarly to EVs, a possible link was described between Rotaviruses - such as Rotavirus C (RVC) - and T1D, because of a potential molecular mimicry between the VP7 protein of a human rotavirus strain and I-A2 (35) and GAD65 (36), although in the latter case the response was limited to CD4⁺ T cells.

Viral and β cell epitopes might be generated and presented during viral infections, where β cell destruction could be triggered. Even if the infection is cleared and regulatory mechanisms are in place, additional infections could trigger further waves of β cell destruction. This might explain why T1D has been defined as a relapsing-remitting disease, where β cells may be killed only when a certain trigger (*i.e.* a viral infection) is present (37). Moreover, persistent infections are likely to be problematic, as they could perpetuate inflammation, immune activation and β cell destruction. Many herpesviruses produce lifelong infections and remain in their host in a latent state. Reactivations occur upon immune dysfunction, but might also be concomitant with newer infections. Cytomegalovirus (HCMV) is one of the most prevalent viruses of this family, infecting over half of adults in the United States by the age of 40 (<https://www.cdc.gov/cm/overview.html>). Several case reports have associated HCMV infection with T1D. Pak and colleagues (38) showed a strong correlation between HCMV genome and islet autoantibodies, while HCMV-positive cells were found in the islets of subjects with fulminant T1D (39). HCMV-specific CD8⁺ T cells have been found in the pancreas of T1D subjects at the onset of the disease (40). Conversely, several reports have found neither an association with the disease (41) nor even a delay in progression of T1D (42). It is therefore possible that more virus-specific cells are present in the islets during disease progression than previously expected. Other highly prevalent herpesviruses like Epstein-Barr virus (EBV), human herpesvirus 6A (HHV-6A) and 6B (HHV-6B) have also been associated with T1D. High circulating antibody titers against EBV have been detected in patients with T1D when compared to non-diabetic controls (43). Interestingly, HHV-6B glycoprotein B (gB) was more expressed in the islets and exocrine pancreas of donors with T1D, as compared to non-diabetic subjects (44). However, due to the persistent and latent nature of these viruses, a direct involvement in the pathogenesis of T1D is likely to be hard to prove.

If viral-human *zwitter* epitopes associated with multiple viruses existed, immune responses could be constantly triggered, even if they were of low magnitude, with potential implication for the etiopathogenesis of T1D. Therefore, we investigated *in silico* the theoretical existence of T1D-associated

viral-human *zwitter* peptides, and their potential in triggering CD8⁺ T cell-mediated autoimmunity in T1D.

MATERIALS AND METHODS

Peptide-HLA-I Binding Affinity Prediction and Immune Epitope Database (IEDB)

The study focused on non-spliced and *cis*-spliced 9 amino acid long (9mer) peptides and HLA-A*01:01, -A*02:01, -A*03:01, -A*11:01, -A23:01, -A*24:02, -B*07:02, -B*08:01, -B*15:01, -B*35:01, -B*39:06, -B*40:01, -B*44:02, -B44:03 complexes. This pool of HLA-I alleles covers over 90% of the Caucasian population. For each HLA-I allele, we computed a cut-off comparable to the threshold of a predicted IC₅₀ ≤ 500 nM of peptide-HLA-A*02:01 complex as follows: we downloaded all 9mer peptides detected through peptide elution from HLA-I complexes, and reported in the IEDB database (45). We restricted the analysis to the HLA-I alleles specified above. For each peptide-HLA-I complex, we predicted the inhibitory constant (IC₅₀) of these 9mer peptide sequences by using NetMHCpan-BA4.0 algorithm (46). IC₅₀ estimates the binding affinity of HLA-I-peptide complexes. The lower the IC₅₀, the higher the binding affinity between peptide and HLA-I complex. To have a similar IC₅₀ cut-off among HLA-I alleles, we determined the quantile of the HLA-A*02:01 for IC₅₀ = 500 nM, which resulted in 91.4%-ile of peptides present in the HLA-A*02:01-specific HLA-I immunopeptidome database of the IEDB. We then applied this quantile to the predicted IC₅₀ distributions of all other peptide-HLA-I complexes (**Figure S1**), thereby identifying the IC₅₀ cut-offs of each HLA-I allele, which corresponded to the peptide-HLA-A*02:01 IC₅₀ = 500 nM. Values are displayed in **Table S2**.

For the identification of peptides already determined in HLA-I immunopeptidomics or analyzed (with positive outcome) for T cell recognition, we consulted the IEDB. We downloaded and selected all HLA-I-restricted peptides for which a positive T cell assay was reported (45). The latter included experiments, for example, performed through tetramer staining, IFN-γ assays with co-culture of APCs pulsed with synthetic peptide candidates and either peripheral blood mononuclear cells (PBMCs) or CD8⁺ T cell clones as well as Cr⁵¹ cytotoxicity. For the computation of antigenic hotspot regions see below.

Estimation of Viral-Human Zwitter Peptides

Viral proteomes were obtained *via* ViralZone, and the Human proteome referred to Swiss-Prot Version 2016 excluding protein isoforms (47, 48). Only viruses with human tropism and association to T1D were included in any downstream analysis here presented (n = 8; **Table S1**). The Human proteome database contained 20,191 protein entries with a total of 11,323,862 amino acid residues.

We focused our study on 9mer peptides since they represent the majority of non-spliced and *cis*-spliced peptides in HLA-I immunopeptidomes (21, 49, 50).

We defined as viral-human *zwitter* 9mer peptide any 9mer peptide that had a sequence that could be obtained by either peptide hydrolysis or *cis*-peptide splicing, both from self-proteins and from viral proteins.

We first computed all possible non-spliced 9mer peptides from viral and human proteomes, and all normal and reverse *cis*-spliced 9mer peptides - with an intervening sequence length ≤ 25 amino acids - that could be derived from the viral and human proteomes. *Cis*-spliced peptides were computed *in silico* as previously described (49). We used an intervening sequence length restriction of 25 amino acid residues to be consistent with our previous study on HLA-I immunopeptidomes and *in cellulo* study on a tumor-associated spliced epitope (49, 51).

After, we compared all viral peptides with human peptides by aligning their sequences.

Two 9mer peptides were considered as identical, *i.e.* as viral-human *zwitter* peptides, if all of their 9 amino acid residues were exactly matching. The relative frequency of viral-human *zwitter* peptides (F_v) was calculated as:

$$F_v = 100 \frac{z_v}{p_v},$$

where z_v is the number of all unique viral-human *zwitter* peptides of a given virus v ; and p_v is the number of all possible unique 9mer peptides derived from virus v . The number of viral-human *zwitter* peptides, z , was computed for the comparison of non-spliced peptides only, of *cis*-spliced peptides only, of non-spliced viral peptides compared to *cis*-spliced human peptides, and of *cis*-spliced viral peptides compared to non-spliced human peptides. Additionally, the relative frequency of all (non-spliced and *cis*-spliced) viral-human *zwitter* peptides was computed.

Estimation of Viral-Human Zwitter Epitope Candidates Considering Antigenic Hotspots and the Potential Antigen Repertoire of Human mTECs and Pancreatic β Cells

To determine the potential hotspot regions among antigens that might be the origin of *zwitter* epitope candidates, we collected all peptide sequences present in IEDB's human HLA-I immunopeptidome database and mapped them to the reference proteome database. For each amino acid in the reference proteome database, we counted how many unique peptides of IEDB's human HLA-I immunopeptidome database contained that residue. For any given *zwitter* 9mer peptide we computed the average count over the 9 residues on its sequence, which was our hotspot score. Finally, we applied a cut-off score of 1 to define hotspot regions. Therefore, a hotspot score of 1 was computed if each residue of a given 9mer peptide was identified at least once in IEDB's human HLA-I immunopeptidome database.

To determine the potential antigen repertoire of human medullary thymic epithelial cells (mTECs) and pancreatic β cells, we extracted gene expression values from the RNA sequencing dataset of human mTECs and pancreatic β cells, published by Gonzalez-Duque and colleagues (19), for each antigen in our study. We filtered all antigens based on their

expression values, such that the expression was smaller than 0.1 RPKM in mTECs and larger than 5 RPKM in pancreatic islets.

Predicted Protein Structures

For visualization purpose, the structure of HCMV DNA primase (*UL70*) and human IA-2 (a.k.a. PTPRN) antigens was determined using iTasser (52) with default settings without inclusion or exclusion of structural templates.

Database Source and Data Availability

The human mTEC's and pancreatic β cell RNA sequencing data were provided by Gonzalez-Duque et al. (19), as well as the T1D-associated antigen list (Table S3).

The list of *zwitter* non-spliced and *cis*-spliced 9mer peptides and related information (Tables S4, S5) are accessible in the repository Mendeley dataset: <http://dx.doi.org/10.17632/z9g9knjxgw.1>

RESULTS

Estimation of Viral-Human Zwitter Epitope Candidates Potentially Associated to T1D

For a systematic estimation of the potential number of viral-human *zwitter* epitope candidates that could (i) be presented by HLA-I complexes, (ii) be involved in an autoimmune CD8⁺ T cell response in T1D patients, and (iii) be at least in part triggered by viral infection, we started from the foundations: we computed the number of 9mer peptides that might originate from human proteome, as well as those that might originate from T1D-associated viruses (Table S3). We

focused on 9mer peptides because this is the predominant length in HLA-I immunopeptidomes. We neglected TCR degeneracy (see Discussion), and therefore we focused only on peptides that might be derived from either human proteome and virus proteome with the exact same sequence of amino acids, here named viral-human *zwitter* peptides. With these restrictions, 332 non-spliced peptides were obtained that might be viral-human *zwitter* non-spliced 9mer peptides (Figure 2A and Table S4). Only HHV-6A and -6B, EBV and HCMV potentially carried these peptides. Among them, 45 were predicted to efficiently bind at least one of the selected HLA-I variants, which represents a large section of the Caucasian population (Figure 2B). Twelve viral-human *zwitter* non-spliced 9mer epitope candidates have already been eluted from HLA-I complexes and identified by mass spectrometry, and for one peptide a positive T cell assay has been described, according to the IEDB (Figures 2C, D). Six viral-human *zwitter* non-spliced 9mer peptides could be derived from a list of T1D-associated antigens proposed by Gonzalez-Dunque et al. (19), (Figure 2E). None of these latter peptides were predicted to efficiently bind the selected HLA-I alleles, although two of them were identified in HLA-I immunopeptidomes, according to the IEDB (Figures 2F, G). No viral-human *zwitter* non-spliced 9mer epitope candidates derived from T1D-associated antigens has been tested through T cell assays with a positive outcome (Figure 2H).

As expected, the scenario changed when we included *cis*-spliced peptides (Table S5). Indeed, the number of *zwitter* peptides that may be produced by *cis*-peptide splicing of either both viral and human antigens or only one or the other – herein defined as *zwitter cis*-spliced peptides - increased in each of the categories analyzed. Almost two million viral-human *zwitter cis*-spliced 9mer peptides were computed, which could be derived from the investigated viruses (Figure 2A), and more than

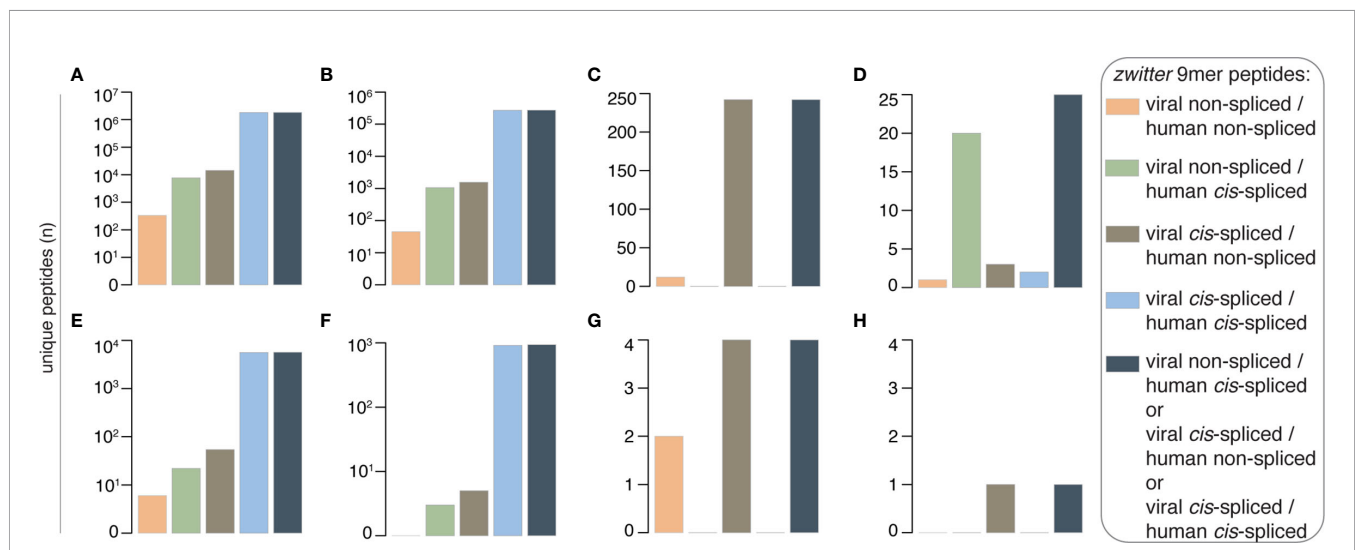


FIGURE 2 | Theoretical viral-human *zwitter* 9mer peptide frequency and potential association with T1D. (A–H) Number of theoretical viral-human 9mer (non-spliced or *cis*-spliced) (A) *zwitter* peptides, (B) *zwitter* epitope candidates predicted to efficiently bind selected HLA-I complexes, (C) *zwitter* epitope candidates described in published HLA-I immunopeptidomes, (D) *zwitter* epitope candidates that showed a positive T cell response in published studies, (E) *zwitter* epitope candidates derived from T1D-associated antigens, (F) *zwitter* epitope candidates predicted to efficiently bind selected HLA-I complexes and derived from T1D-associated antigens, (G) *zwitter* epitope candidates described in published HLA-I immunopeptidomes and derived from T1D-associated antigens, (H) *zwitter* epitope candidates that may be derived from T1D-associated antigens and showed a positive T cell response in published studies. For the identification of epitope candidates already identified in HLA-I immunopeptidomics or analyzed (with positive outcome) for T cell recognition, we consulted the IEDB database.

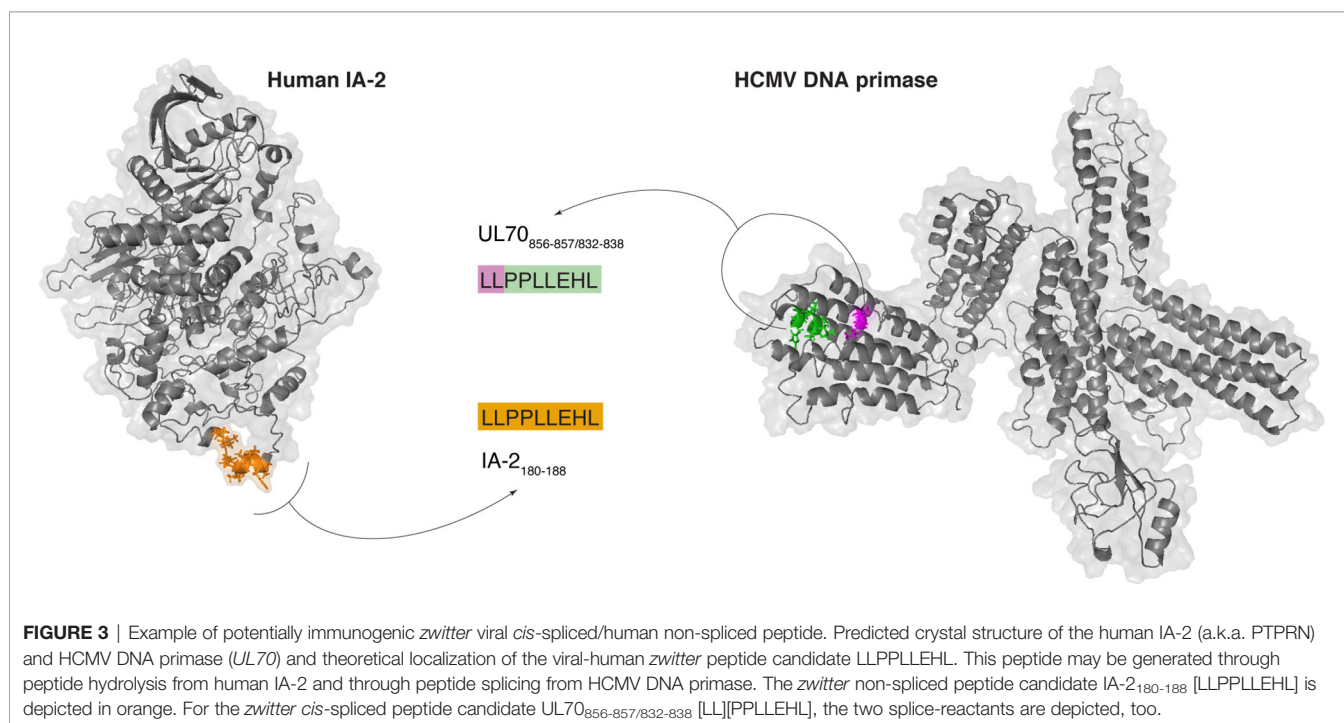
270,000 of them were predicted to efficiently bind the selected HLA-I variants (**Figure 2B**). 242 viral-human *zwitter cis*-spliced 9mer epitope candidates have already been eluted from HLA-I complexes, and identified by mass spectrometry, according to the IEDB (**Figure 2C**). However, they all belonged to the *zwitter viral cis*-spliced/human non-spliced peptide category; hence, they were all identified as human non-spliced peptides in human HLA-I immunopeptidomes (**Table S5**). For 25 viral-human *zwitter cis*-spliced 9mer epitope candidates, we identified studies showing a positive T cell assay (**Figure 2D**). Among them, 20 were viral non-spliced human *cis*-spliced epitope candidates, and the response has been detected against the viral non-spliced epitopes. The remaining five were viral *cis*-spliced epitopes and either human non-spliced or human *cis*-spliced epitope candidates in our database (**Table S5**). For them, the positive T cell response reported by other groups was either against the human non-spliced peptide, or a viral non-spliced peptide derived from a different viral strain than what was included in our database (**Table S1**).

Over 5,000 viral-human *zwitter cis*-spliced 9mer peptides could be derived from the Gonzalez-Dunque's et al. T1D-associated antigen list (**Figure 2E**). Almost a thousand of these latter peptides were also predicted to efficiently bind the selected HLA-I alleles (**Figure 2F**), and four of them were identified as non-spliced peptides in human HLA-I immunopeptidomes, according to the IEDB (**Figure 2G**). One viral-human *zwitter cis*-spliced 9mer peptide derived from T1D-associated antigens has been tested through T cell assays with a barely positive outcome (**Figure 2H**). It was the antigenic peptide [LLPPLLEHL], which may be generated through peptide hydrolysis from the human insulinoma-associated antigen 2 (IA-2; a.k.a. PTPRN) as well as, according to our computation, from the DNA primase (*UL70*)

antigen of HCMV through *cis*-peptide splicing (**Figure 3**). This antigenic peptide is presented by HLA-A*02:01 complex (predicted $IC_{50} = 45$ nM; measured $IC_{50} = 444$ nM) (53, 54). In a standard IFN- γ ELISpot assay with PBMCs, only 1 out of 11 T1D HLA-A*02:01 patients showed a CD8⁺ T cell response above the cut-off, whereas no response was detected in healthy donors against this epitope (53).

Prioritization of Viral-Human Zwitter Peptide Candidates Potentially Associated to T1D

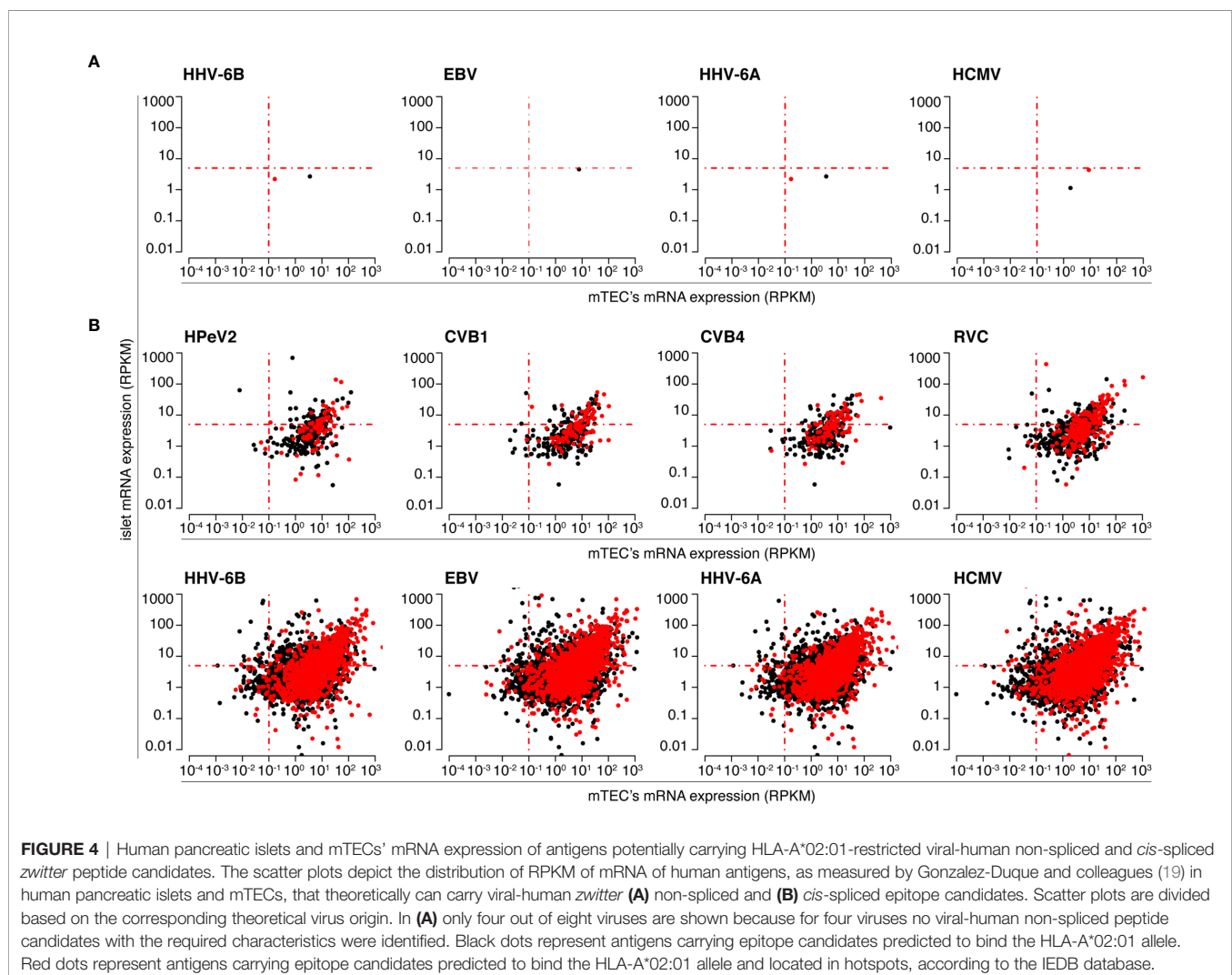
Not all non-spliced and *cis*-spliced epitope candidates that might be generated are in fact produced and presented by HLA-I complexes. APP pathway has many steps that filter peptides based on their sequence and amount (5). The number of peptide molecules in HLA-I immunopeptidomes seems to depend on the amount of the antigen, its degradation rate, and the location of the peptide within antigens. The latter is well represented by "hotspot" regions in antigens, which are overrepresented by antigenic non-spliced and *cis*-spliced peptides in HLA-I immunopeptidomes (49, 55). Furthermore, RPKMs of RNA sequencing measurements, which may be an indirect indicator of protein translation, showed a certain degree of correlation with HLA-I immunopeptidomes (55). According to the current models of thymic negative selection, we could speculate that autoreactive CD8⁺ T cell clones may be more likely to recognize self-antigenic peptides that are not (or barely) presented by mTECs or thymic dendritic cells (DCs), which would reduce the risk of being eliminated during thymic negative selection. If the same antigenic peptides were highly expressed by β cells, the risk of autoreactive CD8⁺ T cell response against these cells would be higher, although several peripheral tolerance pathways are in place to repress undesired autoimmune reactions.

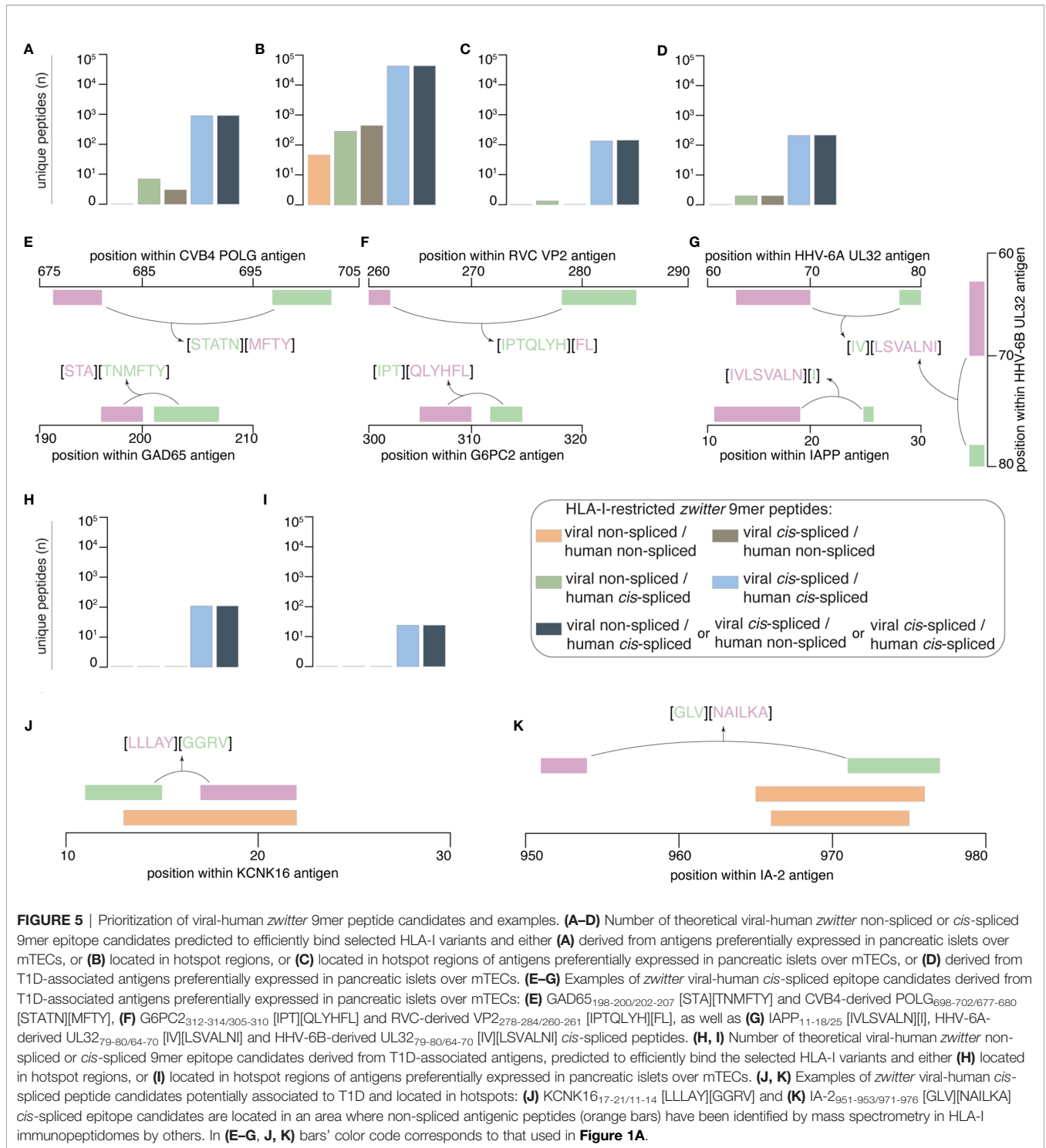


With this in mind, we prioritized viral-human *zwitter* 9mer peptide candidates (predicted to bind the selected HLA-I variants) based on: (i) RNA sequencing data of human mTECs and primary pancreatic islets (for antigen selection); (ii) localization of epitope candidates within hotspot regions of antigens; (iii) antigen association with T1D. The former and the latter databases were derived from Gonzalez-Duque and colleagues (19). The localization within antigenic hotspot regions was computed based on published HLA-I immunopeptidome databases (see Material and Methods section). The distribution of RPKM of genes coding for antigens from which viral-human *zwitter* 9mer epitope candidates may be derived - as measured in mTECs and primary pancreatic islets - is reported in **Figure 4** and **Figures S2–S14**.

When we considered a cut-off for gene expression with a RPKM larger than 5 in islets and smaller than 0.1 in mTECs, which mimicked what was proposed by Gonzalez-Duque and colleagues (19), we obtained no viral-human *zwitter* non-spliced 9mer epitope candidates predicted to efficiently bind the selected HLA-I variants (**Figure 5A**). When we considered only epitope candidates predicted to bind the HLA-I variants and located in

hotspot regions, sixteen viral-human *zwitter* non-spliced 9mer peptide candidates could be identified (**Figure 5B**). Again, the outcome is very different if *cis*-spliced peptides are included. Over 900 epitope candidates that were predicted to efficiently bind the selected HLA-I alleles, and theoretically derived from antigens preferentially expressed in pancreatic islets over mTECs – i.e. with RNA sequencing RPKM larger than 5 in islets and smaller than 0.1 in mTECs – may be viral-human *zwitter cis*-spliced 9mer epitope candidates (**Figure 5A**). Over 60,000 viral-human *zwitter* 9mer epitope candidates predicted to efficiently bind the selected HLA-I alleles and located in hotspot regions may have at least one of the two paired peptides produced by peptide splicing (**Figure 5B**). Among them, over a hundred were predicted to efficiently bind the selected HLA-I alleles, derived from antigens preferentially expressed in pancreatic islets over mTECs and located in hotspot regions. None of the viral-human *zwitter* non-spliced 9mer epitope candidates had these characteristics (**Figure 5C**). When we focused our *in silico* investigation on T1D-associated antigens (**Table S3**), although no non-spliced peptides were predicted to efficiently bind the





selected HLA-I alleles, and may be derived from antigens within the genes' RPKM cut-offs, over 200 viral-human zwitter 9mer epitope candidates with such characteristics may have at least one of the two paired peptides produced by peptide splicing (Figure 5D).

From the latter, we mention the example of the zwitter cis-spliced epitope candidate, which may be derived from the human

GAD65 antigen as [STA][TNMFTY] and from the CVB4 Genome Polyprotein POLG as [STATN][MFTY] (Figure 5E). This zwitter cis-spliced 9mer epitope candidate was predicted to bind HLA-A*01:01, -A*11:01 and -B*35:01 with IC₅₀ < 100 nM (Table S5). In this pool of epitope candidates, we also have a peptide that may be derived from RVC, and specifically from the Inner Capsid Protein VP2 as VP2_{278-284/260-261} [IPTQLYH][FL].

The human counterpart would be the peptide [IPT][QLYHFL] derived from Glucose-6-phosphatase 2 (G6PC_{2312-314/305-310}), which was predicted to bind HLA-B*07:02 and -B*35:01 with IC₅₀ < 200 nM (**Table S5, Figure 5F**). This pool of *zwitter cis*-spliced 9mer epitope candidates also contained many peptides that may be derived from HHV-6A and -6B. In some cases, the viral *zwitter cis*-spliced peptide may be derived from both HHV-6A and -6B, as the peptide [IV][LSVALNI], *i.e.* from HHV-6A Packaging protein UL32 and HHV-6B UL32 homolog (**Figure 5G**). The human counterpart would be the *cis*-spliced peptide [IVLSVALN][I] derived from Islet amyloid polypeptide (IAPP), which is predicted to efficiently bind the HLA-A*02:01 complex (**Table S5**).

When we considered peptides derived from hotspot regions of T1D-associated antigens and predicted to efficiently bind the selected HLA-I alleles, no non-spliced and over 100 *cis*-spliced epitope candidates were identified (**Figure 5H**). Among them, twenty-five may be derived from antigens whose genes were highly expressed in pancreatic islets and marginally expressed in mTECs (**Figure 5I**). The viral epitope candidates may be derived from either EBV or HCMV antigens (**Table S5**). The two T1D-associated antigens are IA-2 and Potassium channel subfamily K member 16 (KCNK16). For the latter, we mention the example of the human KCNK16_{17-21/11-14} *cis*-spliced peptide candidate [LLLAY][GGRV], which may also be annotated as EBV Glycoprotein 42 BZLF_{21-25/34-37}. If these two *cis*-spliced epitope candidates were actually generated by proteasomes, it would be through the ligation of the same splice-reactants [LLLAY] and [GGRV]. The KCNK16_{17-21/11-14} *cis*-spliced peptide was predicted to efficiently bind HLA-A*02:01 complex (predicted IC₅₀ = 93 nM; **Table S5**). It partially overlapped with the HLA-A*03:01-restricted non-spliced peptide KCNK16₁₃₋₂₁ [RVLPLLAY] (**Figure 5J**), which was identified in HLA-I immunopeptidomes of human ECN90 pancreatic β cell line upon IFN- γ stimulation (19).

The other half of the viral-human *zwitter cis*-spliced epitope candidates that are included in this final list may be derived from IA-2 protein, which was largely expressed in human pancreatic islets whereas it was barely expressed in human mTECs (**Figure 4**). Among them, we briefly describe the *zwitter* IA-2_{951-953/971-976} *cis*-spliced peptide candidate [GLV][NAILKA], which may also be annotated as EBV Major DNA-binding protein DBP_{853-855/835-840}. Also, in this case, the peptide splicing reaction would be between the same two splice-reactants [GLV] and [NAILKA]. The *zwitter cis*-spliced peptide IA-2_{951-953/971-976} was predicted to efficiently bind HLA-A*02:01 complex (**Table S5**). The C-terminal splice-reactant of this *cis*-spliced peptide may be derived from an IA-2 area where the non-spliced epitope candidates IA-2₉₆₆₋₉₇₄ [VAEEVNAIL] and IA-2₉₆₅₋₉₇₅ [AVAEVNAILK] were also found (**Figure 5K**). The latter non-spliced epitope candidates have been predicted to bind other HLA-I alleles, and were identified in HLA-I immunopeptidomes by mass spectrometry (19, 56).

IA-2 and KCNK16 antigens are not overrepresented in the IEDB's HLA-I immunopeptidome database (data not shown), therefore their predominance in this latter group of optimal viral-

human *zwitter cis*-spliced 9mer epitope candidates may be due to a partial sequence homology between their sequence and the viral antigen sequences. This is true for IA-2 antigen, which has one of the largest pools of theoretical viral-human *zwitter cis*-spliced 9mer peptide candidates among the T1D-associated antigens (**Table S5**).

DISCUSSION

Our study is the first attempt to evaluate the potential role of antigenic *cis*-spliced peptides in a CD8⁺ T cell-mediated autoimmune response triggered by viral infections. Due to the theoretically extremely large pool of *cis*-spliced peptide sequences, and the limited knowledge of proteasome-catalyzed peptide hydrolysis and peptide splicing dynamics (6, 22, 23, 57–59), any *in silico* analysis of *zwitter cis*-spliced epitope candidates faces further hurdles and a higher degree of complexity when compared to analysis of canonical peptides. Therefore, in this study we neglected multiple layers of complexity of the CD8⁺ T cell response and HLA-I-restricted APP pathway, and focused only on *zwitter* 9mer peptides that share a complete homology between viral- and human-derived peptide candidates. This analysis provided a first estimation of the scale of the pool of viral-human *zwitter cis*-spliced epitope candidates potentially associated with T1D.

Where we disregarded APP pathway and antigen expression, this theoretical pool varied from a few hundred non-spliced peptides to millions of *cis*-spliced peptides. When we considered that, in the context of the CD8⁺ T cell cytotoxicity against pancreatic β cells, immunogenic epitopes are supposed to efficiently bind HLA-I clefts and to be derived from antigens (preferentially from antigen hotspots) that are expressed in pancreatic β cells and, ideally, barely expressed in mTECs, this initial figure seems to decrease. No viral-human *zwitter* 9mer non-spliced epitope candidates, and a hundred *cis*-spliced epitope candidates were left. On the one hand, this figure can further shrink if we considered that not all possible non-spliced and *cis*-spliced epitope candidates are actually produced by proteasomes. Based on *in vitro* digestion experiments with synthetic polypeptides and purified proteasomes, measured by mass spectrometry, we estimated that around one fourth of non-spliced and less than 0.4% of *cis*-spliced peptides that might have been produced were in fact produced by proteasomes in those conditions, and many of them in such a small amount that make them most likely immunologically irrelevant (22, 27). On the other hand, CD8⁺ TCR $\alpha\beta$ are prone to a certain degree of degeneracy of their specificity. This might lead to cross-recognition of multiple antigenic peptides, thereby increasing the immunological overlap between self and non-self antigens. The immunological relevance of CD8⁺ TCR $\alpha\beta$ cross-reactivity is still a matter of debate (60–63), although seminal studies on potential cross-reactivity of T1D-relevant CD8⁺ T cell clones for T1D-associated human antigens and pathogens have already been published (28, 64, 65). In future analyses of *cis*-spliced epitope candidates, including TCR degeneracy would represent a computational challenge, although it might significantly

increase the number of viral-human *zwitter* epitope candidates potentially associated to T1D. With today's limited knowledge of TCR degeneracy, we can also speculate that its introduction in the analysis would also increase the number of false viral-human *zwitter* epitope candidates. Additionally, future experimental analyses should consider the islet microenvironment, where proinflammatory molecules could promote immune activation and antigen presentation to APCs, further refining the pool of epitope candidates that could effectively be presented.

Regarding the *zwitter* epitope candidates derived from T1D-associated viruses and pancreatic β cell antigens described in this study, we found interesting examples potentially derived from CVB4, HHV-6A and -6B as well as RVC. Of course, the *in silico* identification of *zwitter* epitope candidates derived from T1D-associated viruses and pancreatic β cell antigens, even if confirmed *in vitro* and *in cellulo*, would not represent the only key to understanding T1D pathogenesis. It is likely that the strong genetic component of this disease plays a crucial role. We think, however, that in some T1D patients, an autoimmune response could be triggered by viral infection, which in turn might target viral-human *zwitter cis*-spliced epitopes. Although the exact mechanisms by which this occurs are currently unknown, it is possible that type I interferons, secreted in response to viral infections, play an important role. Exposure of human β cells to IFN- α leads to changes in chromatin accessibility, mRNA and protein expression, and the subsequent activation of pathways involved in protein modification, degradation and ER stress (66). IFN- α is capable of shaping the islet microenvironment by inducing the upregulation of several RNA-binding proteins with direct effects in immune cells and the potential to induce extensive changes in alternative splicing, activation and differentiation (66). In addition, the hyper-expression of HLA-I and other anti-viral response markers is associated with islet immune infiltration, which suggests that inflammatory and anti-viral responses play a crucial role in creating an islet microenvironment that potentially attracts APCs and favors antigen presentation (67). The mechanisms leading to epitope formation and presentation, as well as its possible outcomes in terms of T cell activation and cytotoxicity remain elusive, and thus will need to be elucidated in future.

REFERENCES

- Granados DP, Laumont CM, Thibault P, Perreault C. The nature of self for T cells—a systems-level perspective. *Curr Opin Immunol* (2015) 34:1–8. doi: 10.1016/j.coi.2014.10.012
- Granados DP, Sriranganadane D, Daouda T, Zieger A, Laumont CM, Caron-Lizotte O, et al. Impact of genomic polymorphisms on the repertoire of human MHC class I-associated peptides. *Nat Commun* (2014) 5:3600. doi: 10.1038/ncomms4600
- Laumont CM, Daouda T, Laverdure JP, Bonneil E, Caron-Lizotte O, Hardy MP, et al. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun* (2016) 7:10238. doi: 10.1038/ncomms10238
- Laumont CM, Vincent K, Hesnard L, Audemard E, Bonneil E, Laverdure JP, et al. Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med* (2018) 10:eaa5516. doi: 10.1126/scitranslmed.aau5516
- Mishto M, Liepe J. Post-Translational Peptide Splicing and T Cell Responses. *Trends Immunol* (2017) 38:904–15. doi: 10.1016/j.it.2017.07.011

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JL and MM designed the study. JL and AM carried out the data analysis and MM the data mining. MM and TR-C critically revised the immunological implication of the analysis. JL, TR-C, and MM wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The study was in part supported by: (i) MPI-BPC collaboration agreement 2020, Cancer Research UK [C67500; A29686] and National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London and/or the NIHR Clinical Research Facility to MM; (ii) Helmholtz Zentrum Munich junior group funding to TR-C; (iii) European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 945528) to JL.

ACKNOWLEDGMENTS

We thank M. Pandey for proofreading the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.656451/full#supplementary-material>

- Liepe J, Ovaas H, Mishto M. Why do proteases mess up with antigen presentation by re-shuffling antigen sequences? *Curr Opin Immunol* (2018) 52:81–6. doi: 10.1016/j.coi.2018.04.016
- Delong T, Wiles TA, Baker RL, Bradley B, Barbour G, Reisdorph R, et al. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* (2016) 351:711–4. doi: 10.1126/science.aad2791
- Arribas-Layton D, Guyer P, Delong T, Dang M, Chow IT, Speake C, et al. Hybrid Insulin Peptides Are Recognized by Human T Cells in the Context of DRB1*04:01. *Diabetes* (2020) 69:1492–502. doi: 10.2337/db19-0620
- Wang Y, Sosinowski T, Novikov A, Crawford F, White J, Jin N, et al. How C-terminal additions to insulin B-chain fragments create superagonists for T cells in mouse and human type 1 diabetes. *Sci Immunol* (2019) 4:eaav7517. doi: 10.1126/sciimmunol.aav7517
- Babon JA, DeNicola ME, Blodgett DM, Crevecoeur I, Buttrick TS, Maehr R, et al. Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. *Nat Med* (2016) 22:1482–7. doi: 10.1038/nm.4203

11. Baker RL, Rihanek M, Hohenstein AC, Nakayama M, Michels A, Gottlieb PA, et al. Hybrid Insulin Peptides Are Autoantigens in Type 1 Diabetes. *Diabetes* (2019) 68:1830–40. doi: 10.2337/db19-0128
12. 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) 21:455–63. doi: 10.1038/s41590-020-0623-7
13. Reed B, Crawford F, Hill RC, Jin N, White J, Krovi SH, et al. Lysosomal cathepsin creates chimeric epitopes for diabetogenic CD4 T cells via transpeptidation. *J Exp Med* (2021) 218:e20192135. doi: 10.1084/jem.20192135
14. Wiles TA, Powell R, Michel R, Beard KS, Hohenstein A, Bradley B, et al. Identification of Hybrid Insulin Peptides (HIPs) in Mouse and Human Islets by Mass Spectrometry. *J Proteome Res* (2019) 18:814–25. doi: 10.1021/acs.jproteome.8b00875
15. Hanada K, Yewdell JW, Yang JC. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* (2004) 427:252–6. doi: 10.1038/nature02240
16. Vigneron N, Stroobant V, Chapiro J, Ooms A, Degiovanni G, Morel S, et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science* (2004) 304:587–90. doi: 10.1126/science.1095522
17. Platteel ACM, Liepe J, Textoris-Taube K, Keller C, Henklein P, Schalkwijk HH, et al. Multi-level Strategy for Identifying Proteasome-Catalyzed Spliced Epitopes Targeted by CD8+ T Cells during Bacterial Infection. *Cell Rep* (2017) 20:1242–53. doi: 10.1016/j.celrep.2017.07.026
18. Ebstein F, Textoris-Taube K, Keller C, Golnik R, Vigneron N, Van den Eynde BJ, et al. Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes. *Sci Rep* (2016) 6:24032. doi: 10.1038/srep24032
19. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, et al. 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(6):946–60. doi: 10.1016/j.cmet.2018.07.007
20. Faridi P, Woods K, Ostrouska S, Deceneux C, Aranha R, Duscharla D, et al. Spliced Peptides and Cytokine-Driven Changes in the Immunopeptidome of Melanoma. *Cancer Immunol Res* (2020) 8(10):1322–34. doi: 10.1101/623223
21. Liepe J, Marino F, Sidney J, Jeko A, Bunting DE, Sette A, et al. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* (2016) 354:354–8. doi: 10.1126/science.aaf4384
22. Specht G, Roetschke HP, Mansurkhodzhaev A, Henklein P, Textoris-Taube K, Urlaub H, et al. Large database for the analysis and prediction of spliced and non-spliced peptide generation by proteasomes. *Sci Data* (2020) 7:146. doi: 10.1038/s41597-020-0487-6
23. Mishto M, Goede A, Taube KT, Keller C, Janek K, Henklein P, et al. Driving Forces of Proteasome-catalyzed Peptide Splicing in Yeast and Humans. *Mol Cell Proteomics* (2012) 11:1008–23. doi: 10.1074/mcp.M112.020164
24. Mishto M, Mansurkhodzhaev A, Ying G, Bitra A, Cordfunke RA, Henze S, et al. An in silico-in vitro Pipeline Identifying an HLA-A(*02:01)+ KRAS G12V(+) Spliced Epitope Candidate for a Broad Tumor-Immune Response in Cancer Patients. *Front Immunol* (2019) 10:2572. doi: 10.3389/fimmu.2019.02572
25. Mishto M. What we see, what we do not see and what we do not want to see in HLA class I Immunopeptidomes. *Proteomics* (2020) 20:e2000112. doi: 10.1002/pmic.202000112
26. Grignolio A, Mishto M, Faria AM, Garagnani P, Franceschi C, Tieri P. Towards a liquid self: how time, geography, and life experiences reshape the biological identity. *Front Immunol* (2014) 5:153. doi: 10.3389/fimmu.2014.00153
27. Mansurkhodzhaev A, Barbosa CRR, Mishto M, Liepe J. Proteasome-generated spliced peptides and their potential role in CD8+ T cell tolerance. *Front Immunol (in press)* (2021) 12:614276. doi: 10.3389/fimmu.2021.614276
28. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, et al. 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:eaa04013. doi: 10.1126/sciimmunol.aao4013
29. Yu W, Jiang N, Ebert PJ, Kidd BA, Muller S, Lund PJ, et al. Clonal Deletion Prunes but Does Not Eliminate Self-Specific alpha-beta CD8(+) T Lymphocytes. *Immunity* (2015) 42:929–41. doi: 10.1016/j.immuni.2015.05.001
30. ElTanbouly MA, Noelle RJ. Rethinking peripheral T cell tolerance: checkpoints across a T cell's journey. *Nat Rev Immunol* (2020) 21:257–67. doi: 10.1038/s41577-020-00454-2
31. Rodriguez-Calvo T. Enterovirus infection and type 1 diabetes: unraveling the crime scene. *Clin Exp Immunol* (2019) 195:15–24. doi: 10.1111/cei.13223
32. Dotta F, Censini S, van Halteren AG, Marselli L, Masini M, Dionisi S, et al. Coxsackie B4 virus infection of beta cells and natural killer cell insulinitis in recent-onset type 1 diabetic patients. *Proc Natl Acad Sci USA* (2007) 104:5115–20. doi: 10.1073/pnas.0700442104
33. Varela-Calvino R, Skowera A, Arif S, Peakman M. Identification of a naturally processed cytotoxic CD8 T-cell epitope of coxsackievirus B4, presented by HLA-A2.1 and located in the PEVKEK region of the P2C nonstructural protein. *J Virol* (2004) 78:13399–408. doi: 10.1128/JVI.78.24.13399-13408.2004
34. Kolehmainen P, Koskineemi M, Oikarinen S, Veijola R, Simell O, Ilonen J, et al. Human parechovirus and the risk of type 1 diabetes. *J Med Virol* (2013) 85:1619–23. doi: 10.1002/jmv.23659
35. Honeyman MC, Stone NL, Harrison LC. T-cell epitopes in type 1 diabetes autoantigen tyrosine phosphatase IA-2: potential for mimicry with rotavirus and other environmental agents. *Mol Med* (1998) 4:231–9. doi: 10.1007/BF03401920
36. Jones DB, Crosby I. Proliferative lymphocyte responses to virus antigens homologous to GAD65 in IDDM. *Diabetologia* (1996) 39:1318–24. doi: 10.1007/s001250050576
37. von Herrath M, Sanda S, Herold K. Type 1 diabetes as a relapsing-remitting disease? *Nat Rev Immunol* (2007) 7:988–94. doi: 10.1038/nri2192
38. Pak CY, Eun HM, McArthur RG, Yoon JW. Association of cytomegalovirus infection with autoimmune type 1 diabetes. *Lancet* (1988) 2:1–4. doi: 10.1016/S0140-6736(88)92941-8
39. Yoneda S, Imagawa A, Fukui K, Uno S, Kozawa J, Sakai M, et al. A Histological Study of Fulminant Type 1 Diabetes Mellitus Related to Human Cytomegalovirus Reactivation. *J Clin Endocrinol Metab* (2017) 102:2394–400. doi: 10.1210/jc.2016-4029
40. Rodriguez-Calvo T, Krogvold L, Amirian N, Dahl-Jorgensen K, von Herrath M. One in Ten CD8(+) Cells in the Pancreas of Living Individuals With Recent-Onset Type 1 Diabetes Recognizes the Preproinsulin Epitope PPI15-24. *Diabetes* (2021) 70:752–8. doi: 10.2337/db20-0908
41. Aarnisalo J, Veijola R, Vainionpaa R, Simell O, Knip M, Ilonen J. Cytomegalovirus infection in early infancy: risk of induction and progression of autoimmunity associated with type 1 diabetes. *Diabetologia* 51 (2008) 51:769–72. doi: 10.1007/s00125-008-0945-8
42. Ekman I, Vuorinen T, Knip M, Veijola R, Toppari J, Hyoty H, et al. Early childhood CMV infection may decelerate the progression to clinical type 1 diabetes. *Pediatr Diabetes* (2019) 20:73–7. doi: 10.1111/pedi.12788
43. Bian X, Wallstrom G, Davis A, Wang J, Park J, Throop A, et al. Immunoproteomic Profiling of Antiviral Antibodies in New-Onset Type 1 Diabetes Using Protein Arrays. *Diabetes* (2016) 65:285–96. doi: 10.2337/db15-0179
44. Sabouri S, Benkahla MA, Kiosses WB, Rodriguez-Calvo T, Zapardiel-Gonzalo J, Castillo E, et al. Human herpesvirus-6 is present at higher levels in the pancreatic tissues of donors with type 1 diabetes. *J Autoimmun* (2020) 107:102378. doi: 10.1016/j.jaut.2019.102378
45. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* (2019) 47:D339–43. doi: 10.1093/nar/gky1006
46. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J Immunol* (2017) 199:3360–8. doi: 10.4049/jimmunol.1700893
47. Hulo C, Castro E, Masson P, Bougueleret L, Bairoch A, Xenarios I, et al. ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res* (2011) 39:D576–82. doi: 10.1093/nar/gkq901
48. Bairoch A, Apweiler R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. *Nucleic Acids Res* (1999) 27:49–54. doi: 10.1093/nar/27.1.49
49. Liepe J, Sidney J, Lorenz FKM, Sette A, Mishto M. Mapping the MHC Class I-Spliced Immunopeptidome of Cancer Cells. *Cancer Immunol Res* (2019) 7:62–76. doi: 10.1158/2326-6066.CIR-18-0424
50. Faridi P, Li C, Ramarathinam SH, Vivian JP, Illing PT, Mifsud NA, et al. A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. *Sci Immunol* (2018) 3:eaar3947. doi: 10.1126/sciimmunol.aar3947

51. Dalet A, Vigneron N, Stroobant V, Hanada K, Van den Eynde BJ. Splicing of distant Peptide fragments occurs in the proteasome by transpeptidation and produces the spliced antigenic peptide derived from fibroblast growth factor-5. *J Immunol* (2010) 184:3016–24. doi: 10.4049/jimmunol.0901277
52. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* (2015) 12:7–8. doi: 10.1038/nmeth.3213
53. Ouyang Q, Standifer NE, Qin H, Gottlieb P, Verchere CB, Nepom GT, et al. Recognition of HLA class I-restricted beta-cell epitopes in type 1 diabetes. *Diabetes* (2006) 55:3068–74. doi: 10.2337/db06-0065
54. Sidney J, Vela JL, Friedrich D, Kolla R, von Herrath M, Wesley JD, et al. Low HLA binding of diabetes-associated CD8+ T-cell epitopes is increased by post translational modifications. *BMC Immunol* (2018) 19:12. doi: 10.1186/s12865-018-0250-3
55. Pearson H, Daouda T, Granados DP, Durette C, Bonneil E, Courcelles M, et al. MHC class I-associated peptides derive from selective regions of the human genome. *J Clin Invest* (2016) 126:4690–701. doi: 10.1172/JCI88590
56. Sarkizova S, Klaeger S, Le PM, Li LW, Oliveira G, Keshishian H, et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol* (2020) 38:199–209. doi: 10.1038/s41587-019-0322-9
57. Liepe J, Holzthutter HG, Bellavista E, Kloetzel PM, Stumpf MP, Mishto M. Quantitative time-resolved analysis reveals intricate, differential regulation of standard- and immuno-proteasomes. *Elife* (2015) 4:e07545. doi: 10.7554/eLife.07545
58. Berkers CR, de Jong A, Schuurman KG, Linnemann C, Meiring HD, Janssen L, et al. Definition of Proteasomal Peptide Splicing Rules for High-Efficiency Spliced Peptide Presentation by MHC Class I Molecules. *J Immunol* (2015) 195:4085–95. doi: 10.4049/jimmunol.1402455
59. Paes W, Leonov G, Partridge T, Nicastrì A, Ternette N, Borrow P. Elucidation of the Signatures of Proteasome-Catalyzed Peptide Splicing. *Front Immunol* (2020) 11:563800. doi: 10.3389/fimmu.2020.563800
60. Ishizuka J, Grebe K, Shenderov E, Peters B, Chen Q, Peng Y, et al. Quantitating T cell cross-reactivity for unrelated peptide antigens. *J Immunol* (2009) 183:4337–45. doi: 10.4049/jimmunol.0901607
61. Rossjohn J, Gras S, Miles JJ, Turner SJ, Godfrey DI, McCluskey J. T cell antigen receptor recognition of antigen-presenting molecules. *Annu Rev Immunol* (2015) 33:169–200. doi: 10.1146/annurev-immunol-032414-112334
62. Assmus LM, Guan J, Wu T, Farenc C, Sng XYX, Zareie P, et al. Overlapping Peptides Elicit Distinct CD8(+) T Cell Responses following Influenza A Virus Infection. *J Immunol* (2020) 205(7):1731–42. doi: 10.4049/jimmunol.2000689
63. Szomolay B, Liu J, Brown PE, Miles JJ, Clement M, Llewellyn-Lacey S, et al. Identification of human viral protein-derived ligands recognized by individual MHCI-restricted T-cell receptors. *Immunol Cell Biol* (2016) 94:573–82. doi: 10.1038/icb.2016.12
64. Cole DK, Bulek AM, Dolton G, Schauenberg AJ, Szomolay B, Rittase W, et al. Hotspot autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. *J Clin Invest* (2016) 126:2191–204. doi: 10.1172/JCI85679
65. Whalley T, Dolton G, Brown PE, Wall A, Wooldridge L, van den Berg H, et al. GPU-Accelerated Discovery of Pathogen-Derived Molecular Mimics of a T-Cell Insulin Epitope. *Front Immunol* (2020) 11:296. doi: 10.3389/fimmu.2020.00296
66. Colli ML, Ramos-Rodriguez M, Nakayasu ES, Alvelos MI, Lopes M, Hill JLE, et al. An integrated multi-omics approach identifies the landscape of interferon-alpha-mediated responses of human pancreatic beta cells. *Nat Commun* (2020) 11:2584. doi: 10.1038/s41467-020-16327-0
67. Apaolaza PS, Balcacean D, Zapardiel-Gonzalo J, Nelson G, Lenchik N, Akhbari P, et al. Islet expression of type I interferon response sensors is associated with immune infiltration and viral infection in type 1 diabetes. *Sci Adv* (2021) 7:eabd6527. doi: 10.1126/sciadv.abd6527

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Copyright © 2021 Mishto, Mansurkhodzhaev, Rodriguez-Calvo and Liepe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.