Constitutive and Inflammatory Immunopeptidome of Pancreatic β-Cells

Nadine L. Dudek,[1] Chor Teck Tan,[1] Dhana G. Gorasia,[1] Nathan P. Croft,[1] Patricia T. Illing,[2] and Anthony W. Purcell[1]

Type 1 diabetes is characterized by the autoimmune destruction of pancreatic β-cells. Recognition of major histocompatibility complex (MHC)-bound peptides is critical for both the initiation and progression of disease. In this study, MHC peptide complexes were purified from NIT-1 β-cells, interferon-γ (IFN-γ)-treated NIT-1 cells, splenic and thymic tissue of 12-week-old NOD mice, and peptides identified by mass spectrometry. In addition to global liquid chromatography–tandem mass spectrometry analysis, the targeted approach of multiple-reaction monitoring was used to quantitate the immunodominant K^d-restricted T-cell epitope islet-specific glucose-6-phosphatase catalytic subunit-related protein $(IGRP)_{206–214}$. We identified >2,000 MHC-bound peptides; 1,100 of these presented by β-cells grown under normal conditions or after exposure to IFN-γ. These include sequences from a number of known autoantigens. Quantitation of $IGRP_{206–214}$ revealed low-level presentation by K^d (~25 complexes/cell) on NIT-1 cells after IFN-γ treatment compared with the simultaneous presentation of the endogenously processed K^d-restricted peptide Janus kinase-$1_{355–363}$ (~15,000 copies/cell). We have successfully sequenced peptides from NIT-1 β-cells under basal and inflammatory conditions. We have shown the feasibility of quantitating disease-associated peptides and provide the first direct demonstration of the disparity between presentation of a known autoantigenic epitope and a common endogenously presented peptide. ***Diabetes*** **61:3018–3025, 2012**

The cellular immune response depends upon T-cell recognition of peptides presented on the cell surface by molecules encoded by the major histocompatibility complex (MHC). Several thousand different MHC-bound peptides derived from the degradation of both intracellular and extracellular sources are displayed for scrutiny by T cells. In type 1 diabetes, the recognition of self-peptides leads to immune-mediated destruction of insulin-secreting β-cells and ultimately insulin deficiency. Presentation of peptides by class I MHC molecules on professional antigen-presenting cells and β-cells is critical for the development of disease. Nonobese diabetic (NOD) mice lacking class I MHC fail to develop diabetes, and studies aimed at specifically reducing class I expression on β-cells show an inverse correlation between the level of cell-surface expression and protection from disease (1,2). Moreover, increased expression of class I molecules on β-cells is observed in biopsies of patients with type 1 diabetes reflecting the inflammatory nature of the lesion (3–5).

Identification of peptides presented by β-cells under basal and inflammatory conditions may provide insight into the mechanisms by which β-cells become targeted by autoreactive T cells. In particular, changes in the peptides presented by β-cells under inflammatory conditions may dictate the transition from benign to destructive insulitis in NOD mice. The generation of MHC-bound peptides is governed by a number of factors including protein accessibility, half-life, and protease resistance. Cytokine-induced expression of proteasome subunits and the action of signal peptidases together with cytoplasmic and endoplasmic reticulum–associated aminopeptidases contribute to the complexity and plasticity of peptide generation. Tissue-specific differences in antigen processing or differences in the milieu of inflammatory mediators at the site of antigen presentation may therefore facilitate the development of autoimmunity.

Mass spectrometry offers a powerful approach not only to identify new targets of immunity but also to provide accurate quantitative measurement of antigen presentation (6,7). Traditionally, T-cell epitopes have been defined using peptide libraries that span a given antigen and, less frequently, by Edman sequencing or tandem mass spectrometry. However, for many diseases the vast majority of peptide epitopes remain poorly defined, particularly in heterogeneous populations such as humans (7,8). One reason for this is the great complexity of the immunopeptidome and the relatively small proportion of antigen-specific peptides contained within it. Moreover, the immunopeptidome changes in response to inflammatory stimuli are chemically diverse and contain a significant proportion of peptides of atypical or heterogeneous length or bearing posttranslational modifications.

To examine the changes in peptide presentation by β-cells under inflammatory conditions, we have established a database of peptides presented by cultured β-cells in the presence or absence of interferon-γ (IFN-γ) by sequencing MHC-bound peptides using liquid chromatography–tandem mass spectrometry (LC-MS/MS). Although sequences derived from known autoantigens could be identified among the MHC-bound peptides isolated from the surface of β-cells, these did not include previously identified T-cell epitopes. It has often been assumed that autoantigen-derived epitopes would be either unique to the tissue targeted by autoaggressive T cells or that they would be presented at relatively low levels that fail to induce tolerance. Since standard LC-MS/MS analysis did not reveal a number of known epitopes, we decided to use an approach that improves the selectivity and sensitivity of detection of known analytes. We therefore used the targeted mass spectrometry–based approach of multiple-reaction monitoring (MRM) to examine presentation of the

From the [1]Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Victoria, Australia; and the [2]Department of Immunology and Microbiology, The University of Melbourne, Victoria, Australia.

Corresponding author: Anthony W. Purcell, apurcell@unimelb.edu.au, or Nadine L. Dudek, ndudek@unimelb.edu.au.

Received 22 September 2011 and accepted 17 May 2012.

DOI: 10.2337/db11-1333

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db11-1333/-/DC1.

dominant epitope from the islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP)$_{206–214}$ presented by β-cells. This approach has for the first time allowed direct quantitation of a known diabetes T-cell epitope and demonstrated the disparity in presentation of disease- and non–disease-related peptides. Moreover, we have provided a global dataset with which to probe tissues from NOD mice and examine the fluidity of antigen presentation during the development of disease.

RESEARCH DESIGN AND METHODS

NOD mice were purchased and housed in the Bio21 Institute animal facility. Studies were carried out in accordance with accepted standards of humane animal care and were approved by the animal ethics committee, University of Melbourne. Tissues were snap-frozen in liquid nitrogen and stored at −80°C.

Cell culture. The P815 (H-2^d) mastocytoma and EL4 (H-2^b) thymoma were maintained in RPMI 1640 (Invitrogen) medium supplemented with 10% FCS, 2 mmol/L glutamine, 50 IU/mL penicillin, and 50 μg/mL streptomycin. The NIT-1 insulinoma cell line (9) was maintained in low-glucose Dulbecco's modified Eagle's medium supplemented as described above. Expanded NIT-1 cells were washed in PBS, harvested by scraping, and snap-frozen. For cytokine treatment, cells were cultured with 100 IU/mL IFN-γ (eBioscience) for 24–72 h. For peptide pulsing, 5×10^7 cells were incubated with 1 μmol/L synthetic peptide in serum-free RPMI 1640 medium for 60 min at 37°C and washed with PBS.

Flow cytometry. Cells were incubated with fluorescein isothiocyanate–conjugated anti-K^d (SF1.1.10; BD Biosciences) antibody for 30 min on ice. Viable cells (determined by propidium iodide exclusion) were analyzed using a CyAn ADP flow cytometer (Beckman Coulter) in conjunction with FloJo software.

Peptides. IGRP$_{206–214}$ (VY*LKTNVFL; K^d restricted) and Janus kinase (JAK)-1$_{355–363}$ (SY*FPEITHI; K^d restricted) absolute quantitation (AQUA) peptides were synthesized to incorporate an isotopically labeled amino acid (*) using standard solid-phase fluorenylmethyloxycarbonyl chemistry and purified to >95% by reversed-phase high-performance liquid chromatography (RP-HPLC). The identity of the peptides was confirmed by mass spectrometry and the concentration of peptides after dissolution in water determined by amino acid analysis as previously described (6).

Purification of MHC peptide complexes. Cell pellets or tissues were ground in a Retsch Mixer Mill MM 400 under cryogenic conditions; resuspended in 0.5% IGEPAL, 50 mmol/L Tris (pH 8), 150 mmol/L NaCl, and protease inhibitors (Complete Protease Inhibitor Cocktail Tablet; Roche Molecular Biochemicals) at a density of 5×10^7 cells/mL; and incubated for 1 h at 4°C. Lysates were cleared by ultracentrifugation (200,000*g*) and MHC-peptide complexes immunoaffinity purified using solid-phase bound monoclonal antibodies SF1.1.10 (anti-K^d) and 28.8.6s (anti-D^{b}K^b) as previously described (10). Bound complexes were eluted with 10% acetic acid. The mixture of peptides and class I heavy-chain and β-2 microglobulin was fractionated on a 4.6-mm internal diameter × 50-mm long reversed-phase C18 HPLC column (Chromolith Speed Rod; Merck) using an ÄKTAmicro HPLC (GE Healthcare) as previously described (6). Fractions were collected, vacuum concentrated, and diluted in 0.1% formic acid.

Identification of MHC-bound peptides using LC-MS/MS. Concentrated peptide fractions were analyzed by LC-MS/MS using an Eksigent NanoUltra HPLC and an AB SCIEX 5600 TripleTOF mass spectrometer as described in the Supplementary Data online. Data were analyzed with ProteinPilot software (11) and peptide identities determined subject to strict bioinformatic criteria that included the use of a decoy database to calculate the false discovery rate. A false discovery rate cutoff of 10% was applied, and the filtered dataset was further analyzed manually to exclude redundant peptides and known contaminants as described in the Supplementary Data online. Cellular localization of source proteins was analyzed using Software Tool for Researching Annotations of Proteins (12) and pathway analysis performed though the use of Ingenuity Pathway Analysis (IPA) (Ingenuity Systems; www.ingenuity.com).

Validation and quantitation of peptide epitopes by MRM

Selective validation of global LC-MS/MS analysis by MRM. To determine the validity of the LC-MS/MS dataset generated from the TripleTOF 5600 experiments, 97 peptides from the K^d dataset and 96 peptides from the D^b dataset were selected for MRM detection. For the D^b set, two known subdominant peptides from IGRP were also included (IGRP$_{225–233}$, LRLFGIDLL, and IGRP$_{241–249}$, KWCANPDWI). MRM transitions for these peptides were designed in silico using Skyline v1.1 (University of Washington, Seattle, WA). Peptide eluates were prepared from 1×10^8 IFN-γ–treated NIT-1 cells in duplicate and analyzed using a Tempo nanoLC-1Dplus in combination with a cHiPLC-nanoflex system (Eksigent) coupled to a 5500 QTRAP mass spectrometer (AB SCIEX) using experimental conditions described in the Supplementary Data online.

Quantitation of the immunodominant islet-specific epitope IGRP$_{206–214}$ by MRM. Transitions for MRM experiments were designed after inspection of experimental tandem mass spectrometry data. Three transitions were designed per peptide and validated through a series of experiments to investigate specificity, reproducibility, and background (Supplementary Figs. 1–4). Peptide fractions were concentrated and analyzed with a Tempo nanoLC-1Dplus system coupled to a 5500 QTRAP mass spectrometer as described in the Supplementary Data online. The amount of JAK-1$_{355–363}$ and IGRP$_{206–214}$ peptide present in each HPLC fraction was quantitated by examining the area under each MRM transition peak relative to an internal standard AQUA peptide as previously described (6). The addition of the AQUA peptide occurred immediately after immunoaffinity chromatography, since this is the first practical point for addition. This provides accurate quantitation of the peptide isolated from the immunoaffinity-purified MHC complexes and allows an estimation of the number of complexes per cell (6).

RESULTS

Dataset establishment and motif analysis. For establishment of a database of peptides presented by β-cells, K^d and D^b molecules were purified from cultured NIT-1 cells grown under basal conditions (3×10^9 cells) or after 48 h IFN-γ treatment (1.5×10^9 cells). For provision of a comparative dataset, peptides were also purified from pooled thymus and spleen from 12-week-old female NOD mice (n = 10). MHC-peptide complexes were isolated from detergent lysates using K^d- or D^b-specific antibodies. Peptides were separated from class I heavy-chain and β-2 microglobulin by RP-HPLC and identified by LC-MS/MS. The complete dataset for IFN-γ–treated NIT-1 cells for both K^d (544 peptides from 484 source proteins) and D^b (339 peptides from 297 source proteins) are shown in Supplementary Tables 1 and 2, respectively, and their presence in untreated NIT-1, spleen, or thymus indicated. The sequences of a further 221 peptides (51 D^b, 170 K^d) identified in untreated NIT-1 cells (not in the thymus, spleen, or cytokine-treated cells by LC-MS/MS) are shown in Supplementary Table 3. These 221 peptides may be present in low abundance in the treated cells; however, manual inspection using extracted ion chromatograms generally did not provide definitive confirmation in the treated data. The predicted peptide-binding affinity was determined using the SYFPEITHI algorithm (13), and scores are recorded for K^d 9-mers and D^b 9-mers and 10-mers. Of those peptides sequenced from the spleen, ~80% were also detected in the thymus for both K^d and D^b repertoires. The experimental data generated from IFN-γ–treated NIT-1 cells have been deposited with Tranche (https://proteomecommons.org/tranch/) and are Minimum Information About a Proteomics Experiment compliant (14).

No changes were noted in the amino acid preferences within peptides purified from cells (with or without IFN-γ) or tissues. The distribution of predicted peptide-binding scores remained unaltered between the untreated and IFN-γ–treated samples for K^d (Fig. 1) and D^b (not shown). Motif analysis was therefore performed on the entire elution set obtained from NIT-1, NIT-1 plus IFN-γ, spleen, and thymus. A total of 1,146 and 910 peptide sequences for K^d and D^b, respectively, were used to refine the motifs. Peptides bound to K^d were predominantly 9 amino acids in length (80%) with a small number of 10- or 11-mer peptides (Fig. 2*A*). Consistent with the published motif (15–17), position 2 was dominated by tyrosine and, in a small number of sequences, phenylalanine. The COOH-terminal position was occupied by Ile, Leu, or Val (Fig. 2*B*). Met and Phe were also tolerated at the COOH terminus; however, this preference was more apparent in peptides of 10 or 11 amino acids. Although the dominant anchor residues

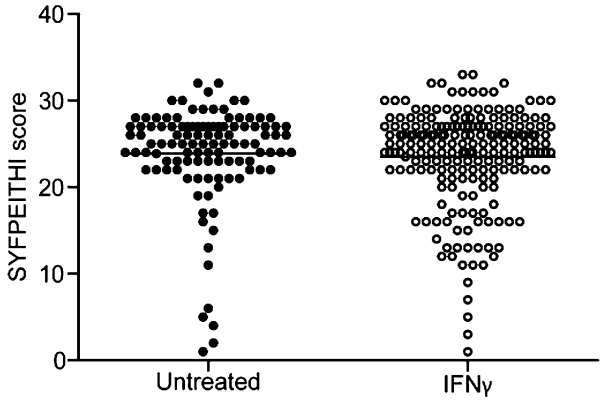


FIG. 1. SYFPEITHI binding scores for K^d 9-mer peptides eluted from NIT-1 cells (110 peptides) or NIT-1 cells treated with IFN-γ for 48 h (209 peptides).

(P2 and PΩ) remained consistent, the binding motif for peptides of 9, 10, or 11 amino acids showed differential amino acid preferences in the COOH-terminal half of the peptide (Fig. 3). In particular, a strong preference for Pro at P7 and Gly at P8 in the 11-mer peptides was noted. Pro was also present at a higher frequency (>10%) at P6 in the 11-mer dataset. Ser and Thr remained moderate to strong anchors at all peptide lengths; however, the position of these residues shifted from P7 to P8 in 10-mer peptides and to P9 in the 11-mers.

While the majority of D^b peptides were 9 amino acids in length (46%), more peptides of 10 and 11 amino acids in length (150 and 152 peptides, respectively) were isolated from D^b than K^d (Fig. 2*C*). The number of peptides longer than 9 amino acids (~53% of total sequences, 63% of these being 10- or 11-mer peptides) exceeds that previously reported for D^b (18). The dominant anchors of Asn at P5 and Leu at the COOH terminus were present in 9-, 10-, and 11-mer peptide sets (Fig. 2*D*) as previously described (18). A strong preference for Ala at P2 was seen in the 9-mers, which became dominant in the 10- and 11-mer peptides (Fig. 3). The appearance of Pro at P6 was noted in 10- and 11-mer peptides, and Pro was also a moderate P7 preference in the 10-mer data set. Gly demonstrated a change in frequency, becoming a moderate anchor at P6 in 10-mer peptides and P8 in 11-mer peptides. Of final note, the preference for glutamic acid or aspartic acid shifted from P7 in the short peptide set to P8 in the 10-mers and P9 in the 11-mer dataset.

Peptide sequences. Peptides were derived largely from ubiquitous proteins, expressed in NIT-1, spleen, and thymus (Supplementary Tables 1 and 2). As expected, the majority of sequences were derived from either nuclear or cytoplasmic proteins (Fig. 4). No change in the localization of source proteins was noted following IFN-γ treatment; however, peptides from IFN-γ–inducible proteins were found in both datasets. Of particular note, peptides from the LMP2 and LMP7 subunits of the immunoproteasome- and the proteasome-associated protein

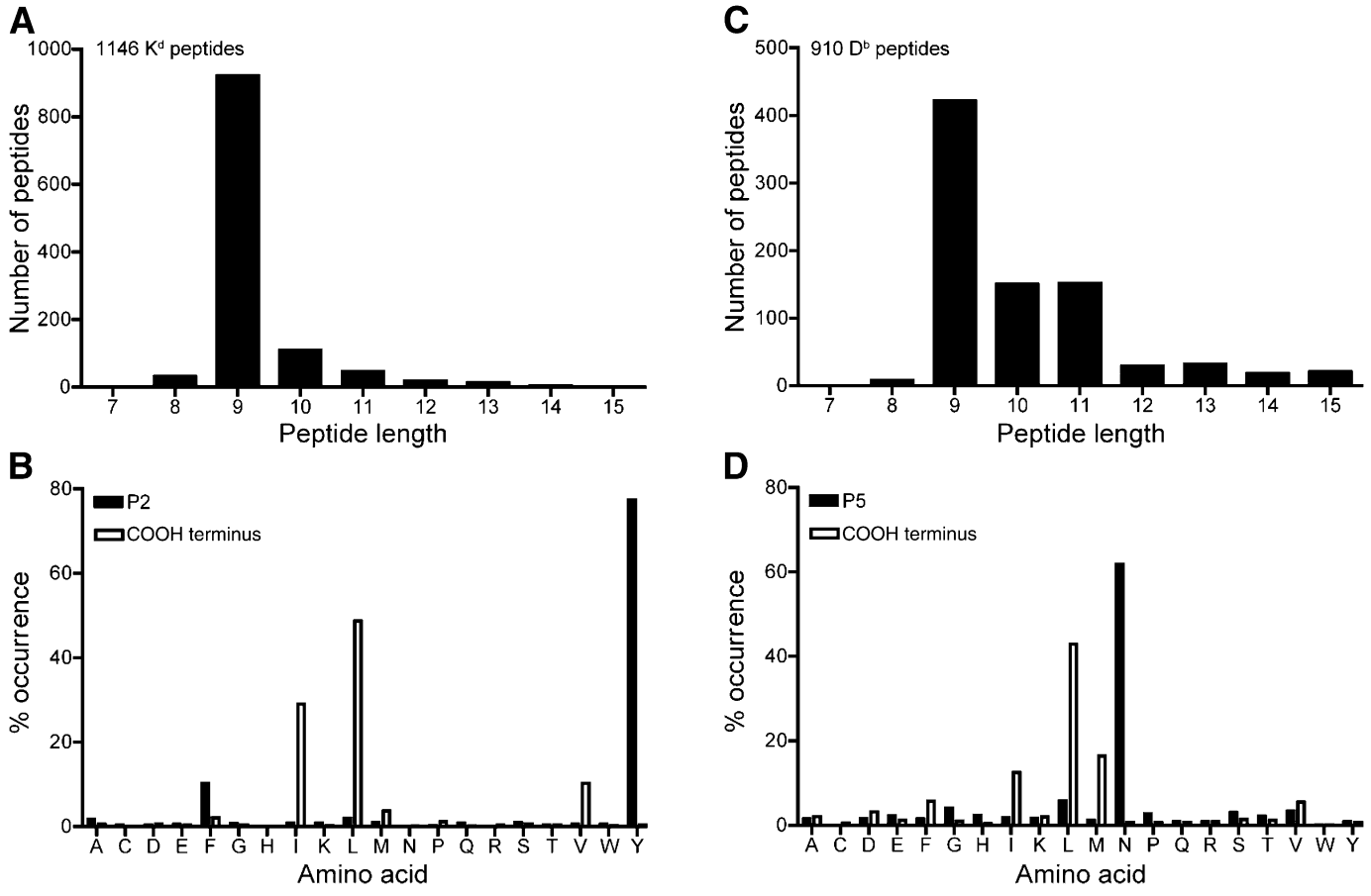


FIG. 2. Length distribution and anchor preference for K^d (*A* and *B*) and D^b (*C* and *D*). Number of peptide sequences used for each analysis is shown in the *top panels*. Anchor preference is expressed as the frequency of occurrence of each amino acid at P2 and CΩ for K^d and at P5 and CΩ for D^b.

K^d 9 mers (921 total)	P1	P2	P3	P4	P5	P6	P7	P8	P9
Dominant (>40%)		Y							L
Strong (>20%)	S		L				T		I
Moderate (>10%)	K	F	Q	P E	L S T V	L K I	S N	E	V
Increased (5-10%)	G A T Y V I		I V F N S K	A S K G N Q D V	A G I N M	R V A Q G F	E Q H	S T L A Y Q F G H N D	

D^b 9 mers (422 total)	P1	P2	P3	P4	P5	P6	P7	P8	P9
Dominant (>40%)					N				L
Strong (>20%)	S A	A	L V						M
Moderate (>10%)	V G	S G	I P	V E T I L		L V I	D E T	T Y H E F	I
Increased (5-10%)	T I	Q M T L		Q R S A	L	K A P T G Q F	V Q S H	V S L A	V F

K^d 10 mers (108 total)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Dominant (>40%)		Y								L
Strong (>20%)	S		L					S		
Moderate (>10%)	K		N	P E Q	L I V M	L G	G S T	T	S L G	I
Increased (5-10%)	T G A V F I R Y	A S	Q I S V F K Y	R K S A L G	N P A Q S T	V T K F I R Y A N Q S	P Q L A R V	E D I N Q G P	V D E A N Y	M V F

D^b 10 mers (150 total)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Dominant (>40%)		A			N					L
Strong (>20%)	S A	S	I P							
Moderate (>10%)		G	V L	R E Q V		F G P	P	E S	H F	I M
Increased (5-10%)	V G T	V		L S A K N		A S L H Y T	K L V A S F E I G T	V P D Y H Q L	T Y A E K N L G S V	

K^d 11 mers (47 total)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Dominant (>40%)		Y									L
Strong (>20%)	S		L				P	G	S		I
Moderate (>10%)	Y T		N	P E A Q	V L M N A	P V L K	S G D		T	G S T	
Increased (5-10%)	V K G	Q	F G H I K	S N	I G S	S A E	T V	K L T H A	A L V	E Q Y A L N	F

D^b 11 mers (152 total)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Dominant (>40%)		A			N						L
Strong (>20%)	S A	S	P								
Moderate (>10%)	V		V I L	T V S L		L P S		G	S D	T V H	M I
Increased (5-10%)	G T	G Q M	M F	Q E A R I	V S G	K D I G H T	T P L G Q F E A S I V	K A S Q L D E P R I H F	A N H L P R Q K	L Y S P E F	F

FIG. 3. K^d and D^b motif analysis. Amino acid preferences found in naturally processed K^d and D^b peptides eluted from NIT-1, IFN-γ–treated NIT-1, spleen, and thymus. Motifs are shown for 9-, 10-, and 11-mer peptides, with the total number of sequences used in each analysis indicated. Amino acids found at each position were classed as dominant (>40% occurrence), strong (>20%), moderate (>10%), and increased (5–10%).

ECM29 were found in IFN-γ–treated cells. This protein binds to the 26S proteasome and has been proposed to couple proteasomes to secretory compartments (19,20). IPA of source proteins within the larger K^d dataset showed a change in the top canonical pathways from caveolar-mediated endocytosis and insulin receptor signaling to ephrin receptor and AMP-activated protein kinase signaling. Phosphatidylinositol 3-kinase/AKT signaling remained a top

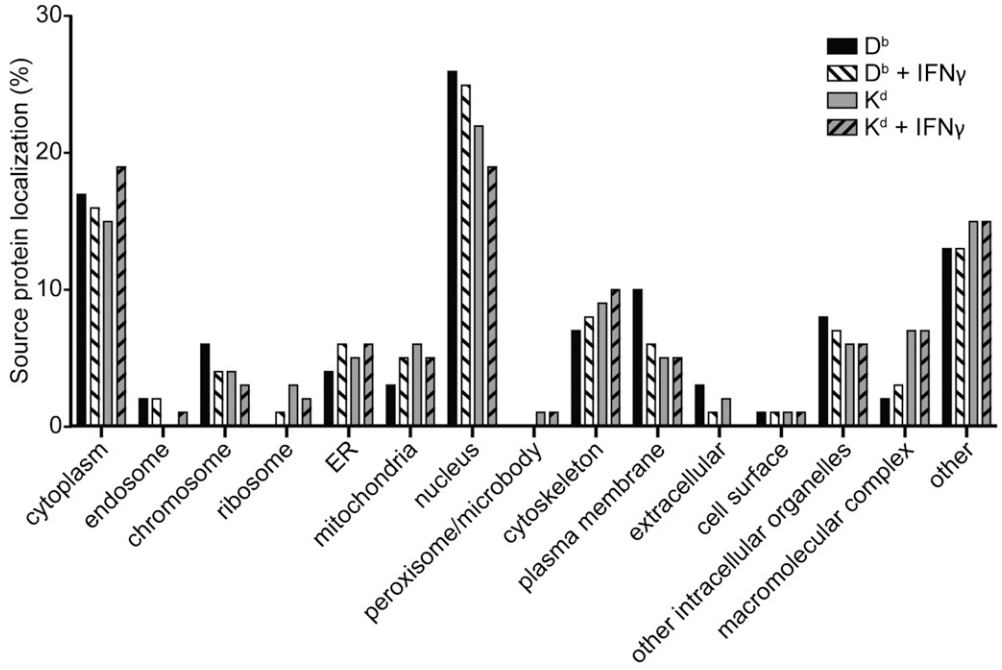


FIG. 4. Source protein localization of naturally processed K^d- and D^b-restricted peptides in the presence or absence of IFN-γ treatment. Data were analyzed using Software Tool for Researching Annotations of Proteins. ER, endoplasmic reticulum.

canonical pathway in both IFNγ–treated and untreated datasets.

Consistent with previous reports, peptides from neuroendocrine or neuronal origin (e.g., α-internexin, neurofilament medium polypeptide, reticulon 4, γ-aminobutyric acid, and glutamate receptors) were identified in both datasets. A number of peptide sequences generated from proteins involved in secretory processes were also identified in untreated and treated NIT-1 cells. Within the IFN-γ–treated dataset, a single peptide derived from the known autoantigen IGRP (D^b-restricted SGVLIIHHL) and a series of overlapping peptides from the potential autoantigen neuropeptide Y (D^b and K^d restricted) were identified (Supplementary Tables 1 and 2). In addition to developing diabetes, NOD mice develop a lupus-like syndrome. Two peptides from proteins associated with Sjögren Syndrome and systemic lupus erythematosus were identified within the D^b elution set, namely Sjögren Syndrome nuclear autoantigen 1 homolog (AALQNYNNEL) and calreticulin (EEESPGQAKDEL). A number of post-translationally modified peptides were also identified, displaying pyroglutamate formation, deamidation, NH_2-terminal acetylation, and glutathionylation. Such modifications occurred in both untreated and treated datasets.

The robustness of peptide identification was examined using a cohort of peptides identified in the discovery experiments in β-cell lines or in NOD mouse primary tissue. Duplicate samples of NIT-1 cells were solubilized and H-$2K^d$ and D^b peptides isolated. With use of in silico–generated MRMs, 77/97 H-$2K^d$– and 78/96 H-$2D^b$–bound peptides were confirmed (representing a minimal 79% validation rate [Supplementary Table 4]). It should be noted that less material was used for the MRM analysis compared with the global LC-MS/MS discovery experiments (15-fold less). The high confirmation rate without the requirement for synthetic peptides provides a high-throughput means of validating initial discovery datasets.

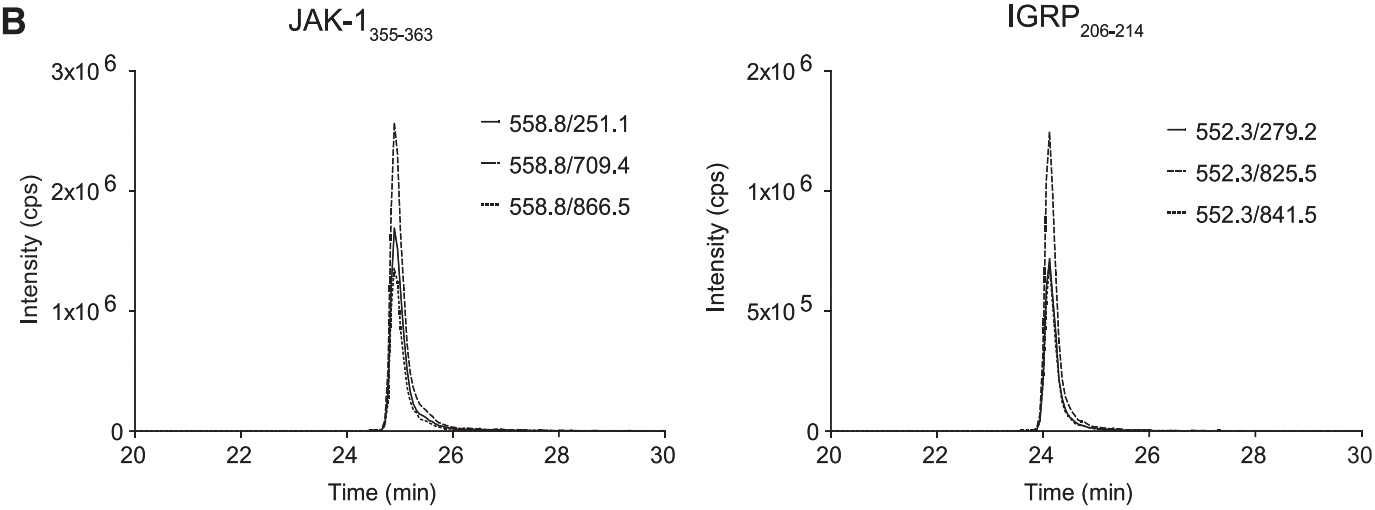
Quantitation of $IGRP_{206–214}$. In these global LC-MS/MS experiments, we did not detect any of the immunodominant peptides previously reported in NOD mice. To determine whether this was due to the sensitivity of the approach, we used MRM for the specific detection of the immunodominant K^d-restricted $IGRP_{206–214}$ peptide. MRM transitions were established for $IGRP_{206–214}$ and the endogenously presented peptide JAK-$1_{355–363}$ (Fig. 5*A*). JAK-$1_{355–363}$ (derived from the interferon signaling molecule JAK-1) was selected, as it was readily identified in NIT-1, IFN-γ–treated NIT-1, thymus, and spleen samples and is a high-affinity K^d ligand. For confirmation of the specificity of these peptides for K^d, P815 (K^dD^d) and EL4 (K^bD^b) cells were incubated with a mix of $IGRP_{206–214}$ and JAK-$1_{355–363}$ AQUA peptides, class I complexes purified, and isolated peptides subjected to MRM analysis. Both $IGRP_{206–214}$ and JAK-$1_{355–363}$ were detected in the K^d eluate of peptide-pulsed P815 cells, as shown in Fig. 5*B*. Although P815 cells endogenously present JAK-$1_{355–363}$, the use of the AQUA version of the peptide to pulse these cells allows a distinction to be made between endogenous presented peptide and exogenous loading of the mass-shifted AQUA peptide by liquid chromatography–mass spectrometry. Neither peptide was detected in fractions affinity purified from peptide-pulsed EL4 cells (K^b or D^b) or the D^d fractions of P815 cells (Fig. 5*C*).

We next used this approach to detect peptides presented by NIT-1 cells. NIT-1 cells were treated with IFN-γ and MHC-peptide complexes isolated at 0, 24, 48, and 72 h posttreatment. The increase in K^d expression following cytokine treatment is shown in Fig. 6*A*. Both JAK-$1_{355–363}$ and $IGRP_{206–214}$ were detected from NIT-1 cells treated with IFN-γ over 24–72 h (Fig. 6*B*). A striking difference was observed between the number of MHC/peptide complexes for the two peptides isolated from the same sample. JAK-$1_{355–363}$ was quantified at ~15,000 copies per cell after 72 h of IFN-γ treatment; however, $IGRP_{206–214}$ did not exceed >25 copies per cell. In the absence of IFN-γ treatment, JAK-$1_{355–363}$ but not $IGRP_{206–214}$ could be detected. To determine whether the inability to detect $IGRP_{206–214}$ in untreated cells was due to the low level of class I

A

Peptides	Sequence	Transition 1	Transition 2	Transition 3
JAK-$1_{355-363}$	SYFPEITHI	553.8/251.1	553.8/709.4	553.8/856.5
JAK-$1_{355-363}$	SY*FPEITHI	558.8/251.1	558.8/709.4	558.8/866.5
$IGRP_{206-214}$	VYLKTNVFL	548.8/279.2	548.8/818.5	548.8/834.5
$IGRP_{206-214}$	VY*LKTNVFL	552.3/279.2	552.3/825.5	552.3/841.5

B



C

Cell Line	Class I allele	JAK-$1_{355-363}$	$IGRP_{206-214}$
P815	D^b	Not detected	Not detected
EL4	K^b/D^b	Not detected	Not detected

FIG. 5. MRM detection of $IGRP_{206–214}$ and JAK-$1_{355–363}$. *A*: LC-MS/MS acquisition was performed on each peptide to determine the best collision energy and to obtain the full fragment ion spectrum; the three highest-intensity peaks were selected to be built into MRM transitions. *B*: MHC-peptide complexes were immunoaffinity purified from 5×10^7 P815 cells pulsed with 1 μmol/L each of $IGRP_{206–214}$ and JAK-$1_{355–363}$ AQUA peptides. Peptide fractions were analyzed on the mass spectrometer in MRM scanning mode. Values show triplicate biological replicates, ±SEM, for K^d eluate. *C*: Absence of MRM signal for JAK-$1_{355–363}$ or $IGRP_{206–214}$ on D^d purified from P815 cells or D^b or K^b from EL-4 cells. cps, counts per second.

expression, a second larger sample of untreated cells was prepared (3×10^9 cells). JAK-$1_{355–363}$ was detected at ~2,000 copies per cell (Fig. 6*C*), consistent with the 0-h time point in the IFN-γ dataset. $IGRP_{206–214}$ was also detected in this sample at ~1 copy/cell.

DISCUSSION

We have established a database of MHC class I peptides endogenously processed and presented by the NIT-1 β-cell line and in tissues of NOD mice. This is the largest dataset of H-2K^d and D^b peptides reported, allowing refinement of the binding motifs for these molecules. Because these peptide sequences are derived from high confidence mass spectrometry identifications of naturally processed and presented peptide ligands, they do not share the inherent bias in amino acid motif definition seen in pool Edman sequencing (21) because of the dominance of highly abundant peptides. In general, the peptide motifs observed in this study were in good agreement with published motifs (15–18) and were further developed to include amino acids present with dominant, strong, and moderate frequencies. Moreover, discrete motifs were also apparent for peptides longer than the canonical 8–9 amino acids. Such long peptides have typically been difficult or impossible to predict because of the lack of a significant number of bona fide longer ligands. In particular, bioinformatics tools have not performed well for D^b in comparison with other alleles (18), attributed to the low number of long natural ligands used to train predictive algorithms. In our study, we identified >400 naturally processed peptides of 10 or 11 amino acids in length, with the selection of such long peptides more frequent in the D^b compared with the K^d repertoire. For both class I molecules, Gly and Pro became more common in peptides of 10 and 11 amino acids, reflecting a requirement for flexibility and kinking or bulging in the peptide backbone of longer peptides (22). For longer peptides bound to H-2D^b, Ala became a dominant motif at P2. Identification of these long, naturally processed peptides and refinement of the current motifs

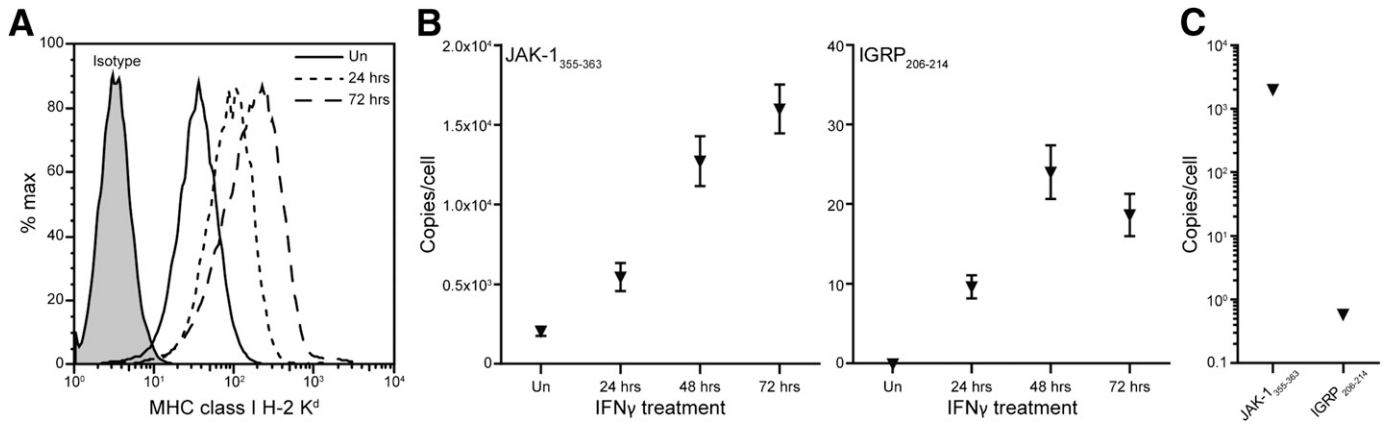


FIG. 6. Detection and quantitation of IGRP$_{206–214}$ and JAK-1$_{355–363}$ presented by NIT-1 β-cells. *A*: Flow cytometric analysis of K^d expression on NIT-1 cells treated with IFN-γ. *B*: NIT-1 cells (5×10^7) were treated with IFN-γ for 0–72 h and K^d peptide complexes immunoaffinity purified at each time point. Samples were spiked with AQUA peptides prior to HPLC separation and fractions subjected to liquid chromatography–MRM analysis. The amount of peptide is shown as copies/cell; values are triplicate biological replicates ±SEM. *C*: K^d MHC-peptide complexes were purified from a single pellet of 3×10^9 untreated NIT-1 cells for peptide quantitation. max, maximum; Un, untreated.

should enhance the success of bioinformatic predictions, providing new opportunities to predict autoantigen-derived epitopes in the NOD mouse.

Treatment of NIT-1 cells with IFN-γ did not significantly change the length of peptides presented by either K^d or D^b or the distribution of predicted binding affinities. This suggests that cytokine treatment does not bias toward high-affinity ligands. Recent studies examining the peptide repertoire of a number of human alleles have documented the presence of NH_2-terminally extended peptides within the endogenous peptide pool (23). Consistent with this observation, we have also identified a small number of nested NH_2-terminally extended peptides. Trimming of the NH_2 terminus by the aminopeptidase ERAAP (endoplasmic reticulum aminopeptidase associated with antigen processing) is required for most class I peptides (24). The presence of NH_2-terminally extended peptides is of interest given the reported immunogenicity of such sequences produced in the absence of ERAAP through formation of novel MHC-peptide complexes (25). Although the precise COOH terminus of MHC class I peptides is thought to be generated primarily by the proteasome in the cytoplasm, we also found peptides containing COOH-terminal extensions within the datasets for both K^d and D^b. The identification of these peptides suggests that additional trimming of the COOH terminus may occur after translocation into the endoplasmic reticulum and MHC complex formation.

The majority of peptides sequenced were common to NIT-1, spleen, and thymus. Many of the β-cell–specific peptides were of neuronal origin, consistent with previous data on elution of class II peptides in NIT-1 cells overexpressing the class II transactivator (26). We identified a number of sequences from known autoantigens such as insulin, chromogranin A, and islet amyloid polypeptide. We also identified peptides from other secretory proteins including neuropeptide Y, secretogranin, and proSAAS. These included nested sets of peptides ranging in length from 9 to 23 amino acids, and a number of these were bound to both K^d and D^b (e.g., secretogranin-1, SGKEVKGEEKGENQNSKFEVRLL; secretogranin2, YLNQEQAEQGREHL; neuropeptide Y, RYYSALRHYINLITRQRYG; and chromogranin A, LEGEDDPDRSM). The presence of unusually long peptides bound to class I molecules isolated from β-cells highlights a potential nonconventional antigen-presentation role in NOD mice. Whether these represent cleavage intermediates that are further processed following cytokine treatment and whether peptides derived from these sequences are recognized by autoreactive T cells warrants further investigation given the number of autoantigenic epitopes already identified from secretory granules. Recently, islet amyloid polypeptide was identified as the target of the BDC T-cell clone 5.2.9 (27). The CD4$^+$ restricted epitope (KCNTATCATQRLANFLVRSS) was mapped using synthetic peptides spanning the entire molecule. We have identified a peptide that overlaps the COOH terminus of this sequence in untreated NIT-1 cells, bound by K^d (LVRSSNNLGP [Supplementary Table 3]). The propensity of certain immunogenic regions of targeted proteins to yield a number of class I and class II restricted epitopes is well documented. We also sequenced a D^b-restricted peptide derived from IGRP in IFN-γ–treated NIT-1 cells (IGRP$_{7–15}$ SGVLIIHHL). This peptide has a moderate binding score of 17 using the SYFPEITHI algorithm. The immunogenicity of this epitope has previously been tested using synthetic peptides (28). IGRP$_{7–15}$ failed to elicit a response from islet-infiltrating T cells isolated from NOD mice of varying ages. Although our data demonstrate that IGRP$_{7–15}$ is naturally processed and presented by NIT-1 β-cells, it is likely that the levels of presentation are sufficient to induce tolerance.

The ability to detect a tolerogenic peptide from IGRP by LC-MS/MS and not the dominant IGRP$_{206–214}$ epitope led us to address the issue of epitope abundance using the sensitive approach of MRM. Using MRM, we were able to detect both the dominant IGRP$_{206–214}$ peptide and the subdominant epitope IGRP$_{225–233}$ (Supplementary Table 4). Moreover, this targeted method allows quantitation of peptide epitopes through the inclusion of an isotopically labeled standard. We found that IGRP$_{206–214}$ was presented at very low levels on NIT-1 cells grown under basal conditions (~1 copy/cell). Conversely, JAK-1$_{355–363}$ was presented on the same cells at ~2,000 copies/cell. Treatment of cells with IFN-γ increased the level of presentation of both IGRP$_{206–214}$ and JAK-1$_{355–363}$. Although the total number of IGRP$_{206–214}$ peptide MHC complexes remained extremely low (25 copies/cell) compared with JAK-1$_{355–363}$ (15,000 copies/cell), the fold increase in presentation of the two epitopes was greater for IGRP$_{206–214}$, and did not

merely mirror the fold increase in class I expression (fivefold increase in K^d at 72 h). This underscores the complexities of epitope generation and for the first time highlights the disparity between presentation levels of a known autoantigen and a normal endogenous ligand.

We have used a discovery-based approach to characterize the repertoire of peptides presented by pancreatic β-cells in the presence and absence of an inflammatory signal. Global LC-MS/MS analysis has successfully identified peptides from known autoantigens; however, many of these may represent sequences that are tolerogenic and, indeed, a large number of these peptides were present in the thymus. The quantitation of $IGRP_{206–214}$ provides the first demonstration of low-level presentation of an immunodominant autoantigenic determinant in absolute terms. Moreover, the disparity in levels of presentation of $IGRP_{7–15}$ and the immunodominant $IGRP_{206–214}$ highlights differences in epitope liberation even within the same antigen. Such quantitative information can be used to guide discovery-based approaches; finding the threshold at which autoantigenic peptides are likely to be present will allow sequencing of peptides that may normally be masked by more abundant species. We propose that the work flows of LC-MS/MS and MRM analysis can be used in tandem to mine deeper into the immunopeptidome and that comparative analysis of datasets sequenced from cell lines and tissues in the presence and absence of inflammation can be used to guide selection of potential autoantigens. This may be particularly relevant to granule-associated proteins, the presentation of which clearly changed upon cytokine stimulation in this study. One of the most important applications of MRM analysis of MHC-peptide epitopes will lie in the ability to examine hundreds of peptides in a single analysis. This will not only allow changes in the presentation of known autoantigenic peptides to be determined but will also enable changes in the presentation of peptides generated from a single antigen to be assessed. This will have important implications for understanding epitope liberation and the contribution of epitope abundance to immunodominance hierarchies.

ACKNOWLEDGMENTS

The authors acknowledge funding from Diabetes Australia Research Trust and infrastructure funding from the Australian Research Council. A.W.P. is supported by a National Health and Medical Research Council Australia Senior Research Fellowship.

No potential conflicts of interest relevant to this study were reported.

N.L.D. performed the majority of experiments, designed the study, and wrote the manuscript. C.T.T. designed and executed the MRM experiments. D.G.G., N.P.C., and P.T.I. provided technical help and expert guidance in mass spectrometry. A.W.P. designed the study, supervised experimental work, and wrote the manuscript. N.L.D. and A.W.P. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Yamanouchi J, Verdaguer J, Han B, Amrani A, Serra P, Santamaria P. Cross-priming of diabetogenic T cells dissociated from CTL-induced shedding of beta cell autoantigens. J Immunol 2003;171:6900–6909
2. Kay TW, Parker JL, Stephens LA, Thomas HE, Allison J. RIP-beta 2-microglobulin transgene expression restores insulitis, but not diabetes, in beta 2-microglobulin null nonobese diabetic mice. J Immunol 1996;157:3688–3693
3. Itoh N, Hanafusa T, Miyazaki A, et al. Mononuclear cell infiltration and its relation to the expression of major histocompatibility complex antigens and adhesion molecules in pancreas biopsy specimens from newly diagnosed insulin-dependent diabetes mellitus patients. J Clin Invest 1993;92:2313–2322
4. Imagawa A, Hanafusa T, Itoh N, et al. Immunological abnormalities in islets at diagnosis paralleled further deterioration of glycaemic control in patients with recent-onset Type I (insulin-dependent) diabetes mellitus. Diabetologia 1999;42:574–578
5. Gianani R, Campbell-Thompson M, Sarkar SA, et al. Dimorphic histopathology of long-standing childhood-onset diabetes. Diabetologia 2010;53:690–698
6. Tan CT, Croft NP, Dudek NL, Williamson NA, Purcell AW. Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. Proteomics 2011;11:2336–2340
7. Dudek NL, Perlmutter P, Aguilar MI, Croft NP, Purcell AW. Epitope discovery and their use in peptide based vaccines. Curr Pharm Des 2010;16:3149–3157
8. Alleva DG, Crowe PD, Jin L, et al. A disease-associated cellular immune response in type 1 diabetics to an immunodominant epitope of insulin. J Clin Invest 2001;107:173–180
9. Hamaguchi K, Gaskins HR, Leiter EH. NIT-1, a pancreatic beta-cell line established from a transgenic NOD/Lt mouse. Diabetes 1991;40:842–849
10. Purcell AW, Gorman JJ. The use of post-source decay in matrix-assisted laser desorption/ionisation mass spectrometry to delineate T cell determinants. J Immunol Methods 2001;249:17–31
11. Shilov IV, Seymour SL, Patel AA, et al. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. Mol Cell Proteomics 2007;6:1638–1655
12. Bhatia VN, Perlman DH, Costello CE, McComb ME. Software tool for researching annotations of proteins: open-source protein annotation software with data visualization. Anal Chem 2009;81:9819–9823
13. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999;50:213–219
14. Taylor CF, Paton NW, Lilley KS, et al. The minimum information about a proteomics experiment (MIAPE). Nat Biotechnol 2007;25:887–893
15. Mitaksov V, Fremont DH. Structural definition of the H-2Kd peptide-binding motif. J Biol Chem 2006;281:10618–10625
16. Romero P, Corradin G, Luescher IF, Maryanski JL. H-2Kd-restricted antigenic peptides share a simple binding motif. J Exp Med 1991;174:603–612
17. Quesnel A, Casrouge A, Kourilsky P, Abastado JP, Trudelle Y. Use of synthetic peptide libraries for the H-2Kd binding motif identification. Pept Res 1995;8:44–51
18. Delgado JC, Escobar H, Crockett DK, Reyes-Vargas E, Jensen PE. Identification of naturally processed ligands in the C57BL/6 mouse using large-scale mass spectrometric peptide sequencing and bioinformatics prediction. Immunogenetics 2009;61:241–246
19. Gorbea C, Goellner GM, Teter K, Holmes RK, Rechsteiner M. Characterization of mammalian Ecm29, a 26 S proteasome-associated protein that localizes to the nucleus and membrane vesicles. J Biol Chem 2004;279:54849–54861
20. Gorbea C, Pratt G, Ustrell V, et al. A protein interaction network for Ecm29 links the 26 S proteasome to molecular motors and endosomal components. J Biol Chem 2010;285:31616–31633
21. Purcell AW, McCluskey J, Rossjohn J. More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov 2007;6:404–414
22. Tynan FE, Reid HH, Kjer-Nielsen L, et al. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. Nat Immunol 2007;8:268–276
23. Escobar H, Crockett DK, Reyes-Vargas E, et al. Large scale mass spectrometric profiling of peptides eluted from HLA molecules reveals N-terminal-extended peptide motifs. J Immunol 2008;181:4874–4882
24. Cascio P, Hilton C, Kisselev AF, Rock KL, Goldberg AL. 26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide. EMBO J 2001;20:2357–2366
25. Hammer GE, Gonzalez F, James E, Nolla H, Shastri N. In the absence of aminopeptidase ERAAP, MHC class I molecules present many unstable and highly immunogenic peptides. Nat Immunol 2007;8:101–108
26. Suri A, Walters JJ, Rohrs HW, Gross ML, Unanue ER. First signature of islet beta-cell-derived naturally processed peptides selected by diabetogenic class II MHC molecules. J Immunol 2008;180:3849–3856
27. Delong T, Baker RL, Reisdorph N, et al. Islet amyloid polypeptide is a target antigen for diabetogenic CD4+ T Cells. Diabetes 2011;60:2325–2330
28. Han B, Serra P, Amrani A, et al. Prevention of diabetes by manipulation of anti-IGRP autoimmunity: high efficiency of a low-affinity peptide. Nat Med 2005;11:645–652