New insights into the mechanisms of proteasome-mediated peptide splicing learned from comparing splicing efficiency by different proteasome subtypes

Violette Ferrari[*,†], Vincent Stroobant[*,†], Joanna Abi Habib[*,†], Stefan Naulaerts[*,†], Benoit J. Van den Eynde[*,†,‡,§] and Nathalie Vigneron[*,†,§]

(*) Ludwig Institute for Cancer Research, Brussels, Belgium; de Duve Institute, Université Catholique de Louvain, Brussels, Belgium

(†) WELBIO (Walloon Excellence in Life Sciences and Biotechnology), Brussels, Belgium.

(‡) Ludwig Institute for Cancer Research, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

(§) These authors contributed equally to this work

Correspondence :nathalie.vigneron@bru.licr.org (NVI) Tel: +32 2 764 75 70, Fax: +32 2 764 75 90

benoit.vandeneynde@bru.licr.org (BVDE) Tel: +32 2 764 75 72, Fax: +32 2 764 75 90

Running title

Mechanisms of peptide splicing by the proteasome

Abstract

By tying peptide fragments originally distant in parental proteins, the proteasome can generate spliced peptides that are recognized by cytolytic T lymphocytes. This occurs by transpeptidation involving a peptide-acyl-enzyme intermediate and another peptide fragment present in the catalytic chamber. Four main subtypes of proteasomes exist: the standard proteasome (SP), the immunoproteasome (IP) and intermediate proteasomes β1-β2-β5i (SIP) and β1i-β2-β5i (DIP). Here, we use a TMT-quantification approach to study the production of six spliced human antigenic peptides by the four proteasome subtypes. Peptides FGF-$5_{172\text{-}176/217\text{-}220}$, tyrosinase$_{368\text{-}373/336\text{-}340}$ and gp$100_{40\text{-}42/47\text{-}52}$, are better produced by the SP than the other proteasome subtypes. The peptides SP$110_{296\text{-}301/286\text{-}289}$, gp$100_{195\text{-}202/191\text{or}192}$ and gp$100_{47\text{-}52/40\text{-}42}$ are better produced by the IP and the DIP. The current model of proteasome-catalyzed peptide splicing suggests that the production of a spliced peptide depends on the abundance of the peptide splicing partners. Surprisingly, we found that despite the fact that reciprocal peptides RTK_QLYPEW (gp$100_{40\text{-}42/47\text{-}52}$) and QLYPEW_RTK (gp$100_{47\text{-}52/40\text{-}42}$) are composed of identical splicing partners, their production varies differently according to the proteasome subtype. These differences were maintained after *in vitro* digestions involving identical amounts of the splicing fragments. Our results indicate that the amount of splicing partner is not the only factor driving peptide splicing, and suggest that peptide splicing efficiency also relies on other factors such as the affinity of the C-terminal splice reactant for the primed binding site of the catalytic subunit.

Key points

- All four proteasome subtypes (β1β2β5, β1iβ2iβ5i, β5i, β1iβ5i**)** can splice peptides
- Splicing efficiency by the four proteasome subtypes depends on the peptide sequence
- Availability of the splice reactants is a factor controlling splicing efficiency
- Affinity of the C-terminal splice reactant for the primed binding site seems key

Introduction

Cytolytic T lymphocytes (CTL) are essential players of the adaptive immune response, because they can kill virally infected cells and tumor cells. On the cell surface, CTL recognize peptides of 8-11 amino acids loaded onto MHC class I molecules (1). These peptides are mostly derived from the degradation of cellular proteins by the proteasome, a large multi-protease complex located in the cytosol. The 20S core of the proteasome is composed by four stacked heptameric rings, which constitute the catalytic chamber of the proteasome (2, 3). The two outer rings are made of seven structural alpha subunits (α1-7) while the two inner rings contain seven beta subunits (β1-7), three of which (β1, β2 and β5) are catalytically active.

Originally, antigenic peptides produced by the proteasome and recognized by CTL were believed to solely derive from linear fragments of proteins. However, some years ago, this idea was challenged by the identification of several antigenic peptides, which were produced after splicing of two non-contiguous fragments of the parental protein (4-9). Peptide splicing was shown to occur in the catalytic chamber of the proteasome through a transpeptidation reaction (4). This reaction requires the production of an acyl-enzyme intermediate, composed of a peptide fragment bound to the proteasome through the N-terminal threonine of the catalytic subunit. This acyl-enzyme intermediate is subjected to a nucleophilic attack by the N-terminal amino group of another peptide fragment present in the catalytic chamber of the enzyme, leading to the production of a spliced peptide (4). This process involves the excision of the intervening sequence and the creation of a new peptide bond between the spliced fragments. So far, six human tumor antigenic peptides, recognized by cytolytic T lymphocytes, were found to be produced by peptide splicing (4-9). Four of these peptides contain peptide fragments that are spliced in the reverse order to that in which they occur in the parental protein (6-9). In addition, one of these peptides was shown to contain two aspartate residues resulting from post-

translational deamidation of asparagine residues (7). While peptide splicing is a low efficiency process (4, 10, 11), the extreme sensitivity of CTL, which can recognize target cells presenting fewer than 10 MHC/peptide complexes (12) likely explains how such a low efficiency process is immunologically relevant and can elicit clinically meaningful CD8 T cell responses. It is for example the case for CD8 T cells isolated against a spliced tyrosinase peptide, which were used for adoptive cell transfer in a melanoma patient and were shown to induce dramatic and lasting tumor regressions (7, 13). The recent developments of peptidomic approaches have led to the detailed characterization of peptides presented by the MHC class I molecules at the cell surface, based on the mass spectrometry analysis of peptides eluted from MHC molecules (14). Using this approach, several groups have tried to estimate the proportion of spliced peptides effectively expressed at the surface of cells (15-17). Because identification of spliced peptides using this approach is based on the development of very complex algorithms, estimation of the proportion of spliced peptides varies from 0.1% to up to 30% between the investigators and therefore remains an intense matter of debate (15-18, 19 , 20, 21). Overall, this confirms that isolation of CTLs remains a crucial step for the validation of spliced peptides.

Based on their ability to cleave fluorogenic peptides and to degrade long precursor peptides, the proteasome catalytic subunits were shown to display cleavage specificities: β1 cleaves mostly after acidic residues (caspase-like activity), β2 after basic residues (trypsin-like activity) and β5 after hydrophobic residues (chymotrypsin-like activity) (22). In cells exposed to IFNγ and in some lymphoid cells, these three constitutive catalytic subunits can be replaced by their inducible counterparts β1i (LMP2), β2i (MECL1) and β5i (LMP7) to form the immunoproteasome (1). Because they contain alternative catalytic subunits, immunoproteasomes display a lower capase-like activity and higher trypsin-like and chymotrypsin-like activities than standard proteasomes (23, 24). Our group identified two

intermediate proteasome subtypes that have incorporated only one or two of the three catalytic subunits of the immunoproteasome: the single intermediate proteasome (SIP), which contains β1, β2 and β5i and the double intermediate proteasome (DIP), composed of β1i, β2 and β5i. These two intermediate proteasomes are present in normal tissues (human liver, kidney and gut), where they represent approximately 20 to 30% of the proteasome content. Intermediate proteasomes also represent half of the proteasome content of immature and mature dendritic cells (25). Although the role of intermediate proteasomes is not yet clear, their different catalytic activity leads to the production of a unique peptide repertoire, with some peptides being solely produced by one or the other of these proteasome subtypes. Identifying peptides that are efficiently processed by intermediate proteasomes SIP and DIP is of great interest for the development of immunotherapeutic cancer vaccines, as these two proteasome subtypes are shared between tumor cells and dendritic cells, which are instrumental to vaccination approaches.

Analyzing proteasome digests performed with long peptide precursors, we previously observed that both the standard proteasome and the immunoproteasome are able to splice peptides, and that some spliced peptides are better produced by the standard proteasome, while others are better produced by the immunoproteasome (7, 26). Our results suggested that this was related to differences in cleavage specificity resulting in differential production of the fragments to splice. However, thus far the ability of intermediate proteasomes to produce spliced peptides has not been investigated. In this study, we compared the four different proteasome subtypes for their ability to produce the five different spliced peptides FGF-5$_{172\text{-}176/217\text{-}220}$, sp110$_{296\text{-}301/286\text{-}289}$, Tyrosinase$_{368\text{-}373/336\text{-}340}$, gp100$_{195\text{-}202/191\text{or}192}$ and gp100$_{40\text{-}42/47\text{-}52}$ against which we have previously isolated CTL clones. We used a tandem mass tag (TMT) labeling approach to quantitatively compare, by mass spectrometry, the fragments produced during in vitro

digestion. By studying the splicing of peptide **RTK_QLYPEW** and its reverse spliced reciprocal counterpart **QLYPEW_RTK** by the four proteasome subtypes, we show that the efficiency of peptide splicing does not only depend on the abundance of the peptide splicing partners.

Materials and methods

Cell lines and CTL clones

The HEK-293-EBNA clonal cell lines expressing different proteasomes subtypes were previously described (25). These cells were cultured in IMDM medium (ThermoFisher Scientific) supplemented with 10% FBS (Sigma Life Science), GlutaMAX (2mM L-alanyl L-glutamine dipeptide, ThermoFisher Scientific),100 U/ml penicillin and 100 µg/ml streptomycin (ThermoFisher Scientific) (IMDM complete medium). 293-SIP were supplemented with hygromycin (600µg/ml) (Roche) and 293-DIP and IP were supplemented with both hygromycin (600µg/ml) and puromycin (5µg/ml) (Invivogen). WEHI 164 cl 13 cells were cultured in RPMI medium (ThermoFisher Scientific) supplemented with 5% FBS HyClone (GE Healthcare), GlutaMAX (2mM L-alanyl L-glutamine dipeptide) and 100 U/ml penicillin and 100µg/ml streptomycin (ThermoFisher Scientific). Epstein-Barr transformed B cells LG2-EBV and CAN-EBV were grown in IMDM medium supplemented with 10% FBS, GlutaMAX (2mM L-alanyl L-glutamine dipeptide), 100U/ml penicillin and 100µg/ml streptomycin. The CTL clone 14 recognize the HLA-A32-restricted spliced peptide $gp100_{40\text{-}42/47\text{-}52}$ RTK_QLYPEW (4). CTL clone C2 recognizes the spliced peptide $FGF\text{-}5_{172\text{-}176/217\text{-}220}$ NTYAS_PRFK associated with HLA-A3 molecule (5). CTL clone DRN-7 recognizes the HLA-A3-restricted spliced peptides SP110 SLPRGT_STPK, SLPRGT_STPKR and SLPRGT_STPKRR (6). CTL 888 clone 62 recognize the HLA-A24 restricted spliced peptide $Tytosinase_{368\text{-}373/336\text{-}340}$ IYMDGT_ADFSF (7). These four CTL were expanded by stimulation every 14 days with anti-CD3 OKT3 monoclonal antibody (10ng/ml) (BioXcell) and irradiated allogeneic PBLs and LG2-EBV cells (100Gy) (27). CTL clone M45-3B recognizes the HLA-A3 restricted spliced peptide $gp100_{195\text{-}202/191or192}$ RSYVPLAH_R (8). This CTL clone was propagated by stimulation every 14 days with 1µg/ml phytohemagglutinin (PHA) (Sigma Aldrich) using irradiated allogeneic PBMCs (100 Gy) and CAN-EBV cells. All CTLs were

cultured in IMDM (ThermoFisher Scientific) containing 10% human serum, rIL-2 (Proleukin sanofi) (150U/ml), and IDO inhibitor 1-methyl-L-tryptophan (200µM; Sigma Aldrich) (28).

Transient transfection and antigen presentation of 293-SP, 293-SIP, 293-DIP and 293-IP cells

Twelve hours before transfection, 293-SP, 293-SIP, 293-DIP and 293-IP were plated in 96 wells plate (Corning) (30.000 cells/wells) in 100µl of IMDM complete medium. Cells were transfected using LT transit reagent (MIR2360, Sopachem) with 50ng of pcDNA3 plasmid encoding the appropriate HLA molecule and the indicated amount of pcDNA3 plasmid encoding the parental antigenic protein. Control cells were transfected with an empty pcDNA3 and loaded or not with the appropriate antigenic peptide (1µM). Twelve hours after the transfection the relevant CTL was added (20.000 cells/wells) in 200µl of CTL medium. After an overnight co-culture the supernatants were collected to measure their TNF-β content using WEHI cells.

Capture proteasome assay (CAPA) assay

The CAPA assay was performed as previously described (29, 30). Briefly, 96-wells plate (Greiner) were coated for 2h at RT with the monoclonal antibody (mAb) MCP21 (5µg/ml) recognizing the proteasome subunit α2 (31), then, blocked for 1h with PBS containing 2% BSA (Roth). Cell pellets were lysed on ice at a cell density of 10^7 cells/ml of Tris-HCL 50mM NP40 0,1% pH 7.5 (lysis buffer). The amount of proteins in the lysate was measured using a BCA protein assay kit (ThermoFisher Scientific) and protein concentration was adjusted to 4mg/ml using lysis buffer. 50µl of cell lysate were then added in each well and the plate were incubated overnight at 4°C. The plate was washed firstly in Tris-HCL 10mM NP-40 0.1% pH 7.5 then in Tris-HCL 10mM pH 7.5. After proteasome capture, 50µl of the indicated precursor peptide

were added at a concentration of 20µg/ml in Tris-HCL 10mM. Plates were then sealed with a plastic cover and incubated at 37°C. For each time point supernatants were collected (45µl) in protein LoBind tube (eppendorf) and digestions were stopped by the addition of 2µl trifluoroacetic acid 10%. All samples were lyophilized and resuspended in 20µl of water. 5µl of each digest were pulsed in 50µl X-vivo-10 serum-free medium (Lonza) onto the indicated EBV-B cells (40.000 cells per well). After 1 hour incubation at 37°C, CTL were added (20.000 cells per well) in CTL medium containing IL2 (25U/ml). IFN-γ produced by the CTL was then measured by ELISA (29).

Digestions of synthetic peptides by purified proteasomes and CTL assay

Purification of 20S proteasome was performed as previously described (32, 33) using frozen pellets of 1×10^9 293 EBNA cells. Digestions of precursor peptides (>95% pure) were performed by incubating at 37°C in 20µl of phosphate buffer (KH2PO4 10mM pH 7,5)/ time point, 1µg of SP, SIP, DIP or IP/ time point with 5µg of RTKAWNRQLYPEWTEAQR (gp100); 7,5µg of NTYASAIHRTSTHFLPRFKQSEQP (FGF-5); 4µg of STPKRRHKKKSLPRGTASSR (sp110); 5µg of ADFSFRNTLEHNALHIYMDGTMSQV (Tyrosinase) or 5µg of RRGSRSYVPLAHSSSAFT (gp100)/ time point. Reactions were stopped at the indicated time point using 2µl trifloroacetic acid 10%. All samples were lyophilized and resuspended in 20µl of water. 5µl of each digest were pulsed in 50ul X-vivo-10 serum-free medium (Lonza) onto the indicated EBV-B cells (20.000 cells per well). After 1hour incubation at 37°C, CTL were added (15.000 cells per well) in CTL medium containing IL2 (25U/ml). After an overnight co-culture, IFN-γ produced by the CTL was measured by ELISA (29).

In figure 7, 1µg of SP SIP DIP or IP/ time point with 1µg of RTKQLYPEWTEA, RTKQLYPEWTEAQR, QLYPEWRTKA and QLYPEWRTKAW. In figure 8, an equi-molar

concentration of each of the combined peptides was used with 1µg of SP SIP DIP or IP/ time point. A no-digestion control (No proteasome) was performed for each reaction.

Peptide synthesis

All precursor peptides were synthesized on solid phase using fluorenylmethoxycarbonyl (Fmoc) chemistry. Synthetized peptides were purified by reverse-phase HPLC to >95% purity and finally characterized by mass spectrometry. The lyophilized peptides were diluted to a final concentration of 20 mg/ml in DMSO and stored at -80°C.

TMT peptide labeling

10µl of each purified proteasome digestions (one half) were lyophilized and solubilized in 10µl of water. Five µl of 200 mM HEPES buffer (pH 8.5) and 2µl of 6-plex TMT (20µg/µl in ACN) (Thermo Fisher Scientific) were added to 2µl aliquots of digestions. The solution was gently mixed, centrifuged and incubated for 90 min without agitation at room temperature. The labeled reaction was quenched by adding 1µl of 50% hydroxylamine and incubated for 30 min and were store at –80°C. Digestions with SP, SIP, DIP, IP and the no-digestion control (No proteasome) were labeled with TMT-126, TMT-127, TMT-128, TMT-129 and TMT-131 (Figure 2-5) or TMT-130 (Figure 1), respectively.

HPLC and Mass spectrometry

For TMT analysis and epitopes detection (orbitrap)

The digestions were analyzed by LC-MS/MS with an Ultimate 3000 RSLCnano system (ThermoFisher Scientific) connected in-line to an Orbitrap Fusion Lumos Tribrid mass spectrometer (ThermoFisher Scientific). Peptides was directly loaded onto a reversed-phase trap column (Acclaim PepMap100, 0.3 × 5 mm, C18, 5µm, 100 A, Thermo Scientific) at

10µl/min with aqueous solution containing 0.1% (v/v) trifluoroacetic acid and 2% ACN. After 3 min, the trap column was set on-line in backflush mode with a reversed-phase analytical column (Acclaim PepMap RSLC, 0.075 × 250 mm, C18, 2 um, 100 A, Thermo Scientific). Solvent A was HPLC grade water with 0.1% (v/v) formic acid and solvent B was HPLC grade 80% acetonitrile with 0.1% (v/v) formic acid. Separations were performed by applying a linear gradient of 4% to 45% solvent B over 40 min, followed by a linear gradient of 45% to 60% solvent B over 5 min, followed a washing step (10 min at 95% solvent B) and an equilibration step (15 min at 4% solvent B) at a constant flow rate of 300nL/min.

For TMT analysis, intact peptides were detected in the Orbitrap at a resolution of 120,000 over a scan range 280-1500 m/z with an AGC target of 4×10^5 ions and a maximum injection time of 50 ms. A data-dependent procedure of MS/MS scans was applied for the top precursor ions with 15 s dynamic exclusion. The total cycle time was set to 3 s. Precursor ions were accumulated with a 0.7 m/z isolation window and their HCD fragmentation was performed with normalized collision energy of 38. Fragment ions were transferred to the Orbitrap at a resolution of 50,000 with an AGC target of 5×10^4 ions and a maximum injection time of 50 ms.

For the targeted quantification of epitopes, the mass acquisition method combined a full scan method (same parameters as above) with a PRM method. This PRM method employed an isolation of the target precursor ions on the basis of the theoretical masses of the epitope sequences (considering the charge states +2 or +3) with a 0.7m/z window. CID fragmentation was performed with a normalized collision energy of 35, and fragment ions were detected in the Orbitrap at a resolution of 50,000 with an AGC target of 5×10^4 ions and a maximum injection time of 50 ms.

Extended spliced peptides digestion (LTQ)

The digestions were analyzed by LC-MS/MS using an Ultimate 3000 RSLC nanoHPLC system (ThermoFisher Scientific) coupled to an LTQ XL ion-trap mass spectrometer (ThermoFinnigan). Peptides were separated on a reversed-phase analytical column (50 cm µPAC C18, PharmaFluidics) with a 34 min linear gradient of 3%–52% ACN (0.1 % FA), followed by 6 min at 70% ACN (0.1 % FA), and 20 min reequilibration at 3% ACN (0.1 % FA), at a constant flow rate of 300 nL/min.

MS analysis was performed on line with the mass spectrometer equipped with a nanoelectrospray source and operated in positive mode with default parameters and active automatic gain control. Mass spectra were acquired in a mode that alternated single MS scans (m/z 300–2,000) with MS/MS scans using low- energy collision-induced dissociation with relative collision energy of 35%.

TMT-labeled peptide identification and quantification

Peptide identifications were made using SearchGui 3.3.17 (34), which was set to use the X!Tandem, MS-GF+, OMSSA and Comet search engines. Precursor Tolerance was set at 7.0 ppm and fragment tolerance at 20.0 ppm. Resulting search engine files were then harmonized and preprocessed using PeptideShaker 2.0.31(35). Project files were then exported to Reporter 0.7.20 for relative TMT quantifications (https://compomics.github.io/projects/reporter.html). Peptides displayed in the figures were identified with 100% (figures 1,2 and 5), 90% (Figure 3) and 80% confidence (Figure 4) and originate, in most cases, from a single cleavage occurring either at one of the border or inside one of the peptide splicing partner. Peptide identifications and quantifications were manually confirmed using the Thermo X-calibur 4.1.50.

Results

Processing of spliced peptide FGF-5$_{172-176/217-220}$ by the four proteasomes subtypes

Several years ago, Hanada et al. showed that a human anti-tumor CD8$^+$ T lymphocyte clone (CTL C2) isolated from a renal cell carcinoma patient recognized a peptide presented by HLA-A3 and composed of two non-contiguous fragments of the FGF-5 protein. Production of this peptide required the removal of a 40 amino-acid sequence located between fragments **NTYAS** and **PRFK**, which compose the spliced peptide **NTYAS_PRFK** recognized by the CTL (5). We subsequently showed that production of this peptide was mediated by the proteasome in a transpeptidation reaction (11). To analyze the processing of this spliced peptide by the four proteasome subtypes, we used a set of engineered HEK-293 cells, which exclusively express either standard proteasomes (293-SP), single intermediate proteasomes β5i (293-SIP), double intermediate proteasomes β1iβ5i (293-DIP) or immunoproteasomes (293-IP). These cell lines were previously obtained by successive transfections of parental HEK-293 cells, which exclusively express the SP (hereafter called 293 SP), with strong expression vectors encoding the different inducible subunits. The overexpressed subunits completely replaced the endogenous subunits in the 20S proteasome particles, and LC–MS/MS analysis of proteasomes purified from the four cell lines 293-SP, SIP, DIP and IP confirmed that they contained exclusively SP, SIP, DIP and IP, respectively (25, 33). This occurs despite continued transcription of the genes encoding the constitutive subunits, because of the preferential and cooperative incorporation of inducible subunits into nascent proteasome particles, as determined from the study of immunoproteasome assembly (36). Moreover, the lack of incorporation of constitutive subunits in the proteasomes from lines 293-SIP, 293-DIP and 293-IP appears to trigger their degradation, as a western blot of the total lysates did not show any trace of unincorporated subunits (33). These 293 cells were transiently transfected with plasmid constructs encoding the entire FGF-5 protein and HLA-A3. Transfected cells were then tested

for their ability to activate CTL C2. 293-SP stimulated CTL C2 much more efficiently than 293-SIP, 293-DIP or 293-IP, suggesting that the standard proteasome produced the spliced peptide better than the three other proteasome subtypes (Figure 1A). To study in detail the processing of the FGF-5 peptide, we performed in vitro proteasome digestions using two different methods. The CAPA (capture proteasome assay) is a technique that involves the capture of proteasomes on a microplate using the α2-specific antibody MCP21 (29, 30) . Lysates originating from the four different HEK-293 cell lines were added to wells coated with MCP21. Plates were then extensively washed to remove unbound proteins and the synthetic precursor peptide **NTYAS**AIHRTSTHFL**PRFK**QSEQP was added to the plate and digested by the captured proteasomes. Digestions were stopped at different time points and pulsed on HLA-A3^{+} cells. In agreement with the results of the transfection assay, digests performed with SP were much better recognized by CTL C2 than those obtained from SIP, DIP and IP (Figure 1B). Similar results were obtained when the precursor peptide was digested with purified 20S proteasomes (Figure 1C). These results were in line with our previous results showing that SP better produced this FGF-5 spliced peptide than IP (26). Degradation of the precursor peptide by purified proteasomes was estimated by mass spectrometry, and digests displaying similar rate of precursor degradation were analyzed by HPLC coupled to mass spectrometry (HPLC-MS). Peptide fragments were quantified using a 6-plex TMT mass tag labeling approach in order to obtain quantitative information about the relative abundance of each peptide in the digests. Using this approach, peptides contained in each of the four digests were labeled with a given TMT mass tag. TMT tags are isobaric mass tags that contain defined isotopic substitutions in their mass reporter region, which can therefore be distinguished and quantified following MS/MS fragmentation. Similar amounts of the four labeled digests, displaying similar percentages of precursor degradation (Figure 1D top), were pooled together and loaded on the mass spectrometer in the same run. Upon MS/MS fragmentation, the mass reporter tags are

cleaved and their amount precisely quantified, enabling the precise comparison of the abundance of given peptides in each digest. Unfortunately, using this approach, we were not able to detect the **NTYAS** fragment, which composes the N-terminal part of the spliced peptide **NTYAS_PRFK**. However, the complementary fragment AIHRTSTHFL**PRFK**QSEQP could be found in digests made with SP, DIP and IP and to a much lesser extent in those made with SIP. Although we could not directly estimate the efficiency at which the acyl-enzyme intermediate involving **NTYAS** was produced, these results show that SP, DIP and IP (and to a lesser extent SIP) are able to produce the cleavage between S_{176} and A_{177}, which is necessary for the production of the corresponding acyl-enzyme intermediate. Interestingly, the other splice reactant **PRFK**, which composes the C-terminus of the antigenic peptide was much more abundant in the SP digest. This is likely due to the fact that in DIP and IP digests (and to a lesser extent in the SIP digest), the **PRFK** fragment is destroyed by an internal destructive cleavage, which occurs between F_{189} and K_{190} (fragments **PRF**, NTYASAIHRTSTHFL**PRF** and **K**QSEQP) (Figure 1D). Furthermore, two other destructive cleavages between A_{174} and S_{175} (fragments **NTYA** and **S**AIHRTSTHFL**PRFK**QSEQP) and between Y_{174} and A_{175} (fragments **NTY** and **AS**AIHRTSTHFL**PRFK**QSEQP) were more intense in DIP/IP digests and in SIP/DIP/IP digests, respectively, and therefore impacted the production of the spliced peptide partner **NTYAS** (Figure 1D). Interestingly, the destructive cleavages Y_{174}/A_{175} and F_{189}/K_{190} both occurred after large hydrophobic amino acids (Y and F). The improved ability of SIP, DIP and IP to produce these cleavages could therefore result from the presence of the β5i subunit, whose S1 pocket accommodates larger amino acids than that of the β5 subunit found in SP. In conclusion, our results suggest that the increased production of spliced peptide **NTYAS_PRFK** by the SP is related to the increased ability of the SP to produce the nucleophile peptide **PRFK**, combined with destructive cleavages inside the **NTYAS** splicing partner by the other proteasome subtypes.

Processing of spliced peptide SP110$_{296\text{-}301/286\text{-}289}$ by the four proteasomes subtypes

The spliced peptide **SLPRGT_STPK** is a minor histocompatibility antigen created by a polymorphism in the *SP110* gene (6). This peptide is presented by HLA-A3 and recognized by CTL clone DRN-7 isolated from a myeloma patient after MHC-matched hematopoietic cells transplantation. This peptide is produced after excision of a 6 amino acids sequence and splicing of the fragments **SLPRGT** and **STPK**, which are assembled together in the reverse order to that of the parental protein (6). To study the processing of this peptide, we transfected HEK-293 cells with plasmids encoding SP110 and HLA-A3. Transfected cells expressing IP were much better recognized by CTL DRN-7 than those expressing DIP or SIP. Cells expressing SP did not stimulate CTL DRN-7 at all (Figure 2A). *In vitro* digestions were then performed by incubating synthetic precursor peptide **STPK**RRHKKK**SLPRGT**ASSR with immunocaptured or purified proteasomes obtained from the four HEK-293 cell lines. Digests were pulsed on HLA-A3$^+$ EBV-B cells and tested for their ability to activate CTL DRN-7. As expected, IP digests stimulated CTL DRN-7 better than the DIP and SIP digests, while the SP digests barely induced a CTL response (Figure 2B and 2C). These results were in line with the results obtained in the cellular assay and with previous analyses showing that the IP better produced this peptide than the SP (26).

Four digests with approximately similar percentages of precursor degradation were labeled with TMT, pooled and analyzed by HPLC-MS/MS. Fragment **SLPRGT**, which composes the N-terminal part of the antigenic peptide was more abundant in the DIP and IP digests (Figure 2D). This fragment is produced by two different cleavages: a N-terminal cleavage between K_{296} and S_{297}, which was more intense in the DIP and IP digests (as demonstrated by the quantification of peptides **STPK**RRHKKK and **SLPRGT**ASSR) and a C-terminal cleavage occurring between T_{302} and A_{303}, which was also more intense in the DIP and IP digests as reported by

fragment **STPK**RRHKKK**SLPRGT**. Two destructive cleavages were observed at S_{297}/L_{298} and L_{298}/P_{299}. These destructive cleavages were stronger in the DIP digest, but not sufficient to decrease production of the N-terminal splice reactant **SLPRGT** by DIP when compared to SP and SIP. We detected **STPK**, **STPK**R and **STPK**RR, three fragments that could be used as splicing partners for the production of spliced peptide **SLRGTSTPK**, **SLRGTSTPK**R and **SLRGTSTPK**RR respectively. Although the natural peptide eluted from HLA-A3 molecules is **SLPRGT_STPK**, we previously observed that both peptides **SLPRGT_STPKR** and **SLPRGT_STPKRR** were also recognized by CTL DRN-7 (6). Moreover, these extended spliced peptides might be further trimmed by the proteasome to produce the final antigenic peptide (8). Unfortunately, we could not detect, after TMT labeling, any of the spliced peptides **SLPRGT_STPK**, **SLPRGT_STPKR** or **SLPRGT_STPKRR**. While **STPK** is better produced by the SIP, DIP and IP, **STPKR** and **STPKRR** were more abundant in the IP digest, suggesting that the higher propensity of IP to produce the spliced peptide recognized by CTL DRN-7 might also be at least partly linked to its increased ability to produce the corresponding extended spliced peptides (Figure 2D). Overall, the increased ability of IP to produce the SP110-spliced peptide appears related to its increased propensity to produce both splicing partners when compared to the other proteasome subtypes, particularly the SP.

Processing of spliced peptide tyrosinase$_{368-373/336-340}$ by the four proteasomes subtypes

The HLA-A24-restricted peptide **IYMDGT_ADFSF** is derived from tyrosinase and recognized by melanoma-infiltrating cytolytic T lymphocytes that induced dramatic and lasting tumor regressions after adoptive cell therapy (13). In addition to the reverse splicing reaction in the proteasome, processing of this peptide involves the deamidation of two asparagine residues (7). After transfection of the four HEK-293 cells with plasmids containing the cDNA encoding tyrosinase and HLA-A24, only 293-SP cells could stimulate the CTL (Figure 3A). In

agreement with this, only standard proteasomes incubated with the long peptide precursor **ADFSF**RNTLEHNALH**IYMDGT**MSQV could activate the CTL (Figure 3B and 3C). Digests corresponding to approximately 40% of precursor degradation were labeled with TMT and analyzed by HPLC MS/MS (Figure 3D). Peptide **IYMDGT**, which constitutes the N-terminal splicing partner could not be detected, even in digests obtained with standard proteasomes**.** However, fragment **MSQV**, which is released following cleavage of the peptide bond between T_{373} and M_{374} was more abundant in SP digests, suggesting that the SP better produced the acyl-enzyme intermediate needed for the peptide splicing reaction. Production of the other splicing partner **ADFSF** was more effective in digests from SP and IP, as compared to SIP or DIP. Several internal cleavages were observed in both splicing partners. In particular, cleavages between I_{368} and Y_{369} and between M_{370} and D_{371}, which likely decreased the production of the acyl enzyme-intermediate required for the production of the antigenic peptide, were more intense in SIP, DIP and IP than in SP digests. However, a cleavage between D_{371} and G_{372} was more intensively produced by SP and SIP, likely because these proteasomes share the β1 subunit, which is responsible for the caspase-like activity of the proteasome.

Overall, these observations suggest that spliced peptide tyrosinase$_{368\text{-}373/336\text{-}340}$ is better produced by the standard proteasome in part because the SP better produces the cleavage between T_{373} and M_{374}, which is required for the production of the acyl-enzyme intermediate involving **IYMDGT.** Again, destructive cleavages within peptide splicing partners likely also influence the peptide splicing reaction. However, our inability to detect **IYMDGT** prevents us from ruling out additional factors potentially involved.

Processing of spliced peptide gp100$_{195\text{-}202/191or192}$ by the four proteasomes subtypes

Peptide **RSYVPLAH_R** is derived from the melanosomal protein gp100 and presented by HLA-A3 (8). In contrast to the previously described spliced peptides, which are produced by

423 the association of fragments of 3 to 6 amino acids, peptide **RSYVPLAH_R** results from the
424 association of an 8-amino acid fragment with a single arginine residue. The production of this
425 spliced peptide is a two-step process involving the reverse splicing of a peptide fragment
426 containing at least three amino acids followed by the trimming of the C-terminally-extended
427 spliced peptide to produce the final antigenic peptide (8). To compare the processing of this
428 peptide, we co-transfected 293-SP, 293-SIP, 293-DIP and 293-IP cells with plasmids encoding
429 gp100 and HLA-A3. Transfected 293-IP and 293-DIP cells were recognized by CTL M45-3B,
430 while transfected 293-SIP and 293-SP cells were not or very poorly (Figure 4A). We then
431 digested synthetic precursor peptide **R**RGS**RSYVPLAH**SSSAFT with plate-captured or
432 purified proteasomes. When loaded onto HLA-A3^{+} target cells, the digests performed with IP
433 and DIP were recognized by CTL M45-3B, whereas digests produced by SIP or SP were not or
434 only slightly recognized (Figure 4B and 4C), confirming the results of the transfection assay.
435 Four digests presenting a similar level of precursor degradation were labeled with TMT as
436 above, pooled together and analyzed by HPLC-MS. In DIP and IP digests, we detected higher
437 amounts of peptide **RSYVPLAH**, which composes the N-terminal part of the antigenic peptide
438 recognized by the CTL. Production of **RSYVPLAH** is dependent on two different cleavages.
439 A first cleavage between H_{202} and S_{203}, which can be produced by all four proteasome subtypes
440 as demonstrated by the quantification of the peptide **R**RGS**RSYVPLAH** and its
441 complementary fragment SSSAFT (Figure 4D). The N-terminal cleavage between S_{194} and
442 R_{195}, which is required for the release of **RSYVPLAH**, can be detected in both the SP and IP
443 digests (fragment **RSYVPLAH**SSSAFT). The increased production of the **RSYVPLAH**
444 peptide by DIP and IP is likely due to the occurrence of a destructive cleavage between A_{201}
445 and H_{202} in the SP and SIP digests (as demonstrated by the increased production of peptides
446 **R**RGS**RSYVPLA** and **H**SSSAFT). This destructive cleavage likely relies on the activity of
447 subunit β1, which is found in both SP and SIP but not in DIP and IP. Indeed, although subunit

β1 is mainly linked to the caspase-like activity, it was suggested to display additional activities targeting small neutral amino acids (SNAAP) or branched amino acids (BRAAP) (37), activities that are enhanced by the presence of a proline residue in P3 (38). Note that the IP also makes another destructive cleavage, between Y_{197} and V_{198}, which, however, does not limit its efficiency to produce the **RSYVPLAH** splicing partner.

We previously showed that production of the spliced peptide **RSYVPLAH_R** does not occur through the splicing of a single arginine by transpeptidation (8). Rather, splicing of a longer peptide fragment leads to the production of a C-terminally extended spliced peptide that will then be further trimmed by the proteasome to produce the final spliced peptide. The arginine involved in the splicing reaction was previously shown to correspond to R_{191} or R_{192}. Unfortunately, our experimental conditions did not allow us to detect any fragment containing R_{191} or R_{192}**.**

Altogether, our results suggest that the differential processing of spliced peptide **RSYVPLAH_R** can be explained by a greater capacity of the DIP and IP to produce **RSYVPLAH**, which composes the acyl enzyme intermediate. Although, we were unable to evaluate the role of the C-terminal spliced reactant, efficiency of the splicing reaction appeared to depend on the production of the **RSYVPLAH-**enzyme intermediate.

Processing of spliced peptide gp100$_{40\text{-}42/47\text{-}52}$ by the four proteasomes

The HLA-A32-restricted spliced peptide derived from the melanoma differentiation antigen gp100 is recognized by an anti-tumor CTL clone isolated from a melanoma patient, and was the first spliced peptide shown to be produced by the proteasome in a transpeptidation reaction (4). Production of this spliced peptide **RTK_QLYPEW** involves the excision of four amino

acids and the splicing of the fragments **RTK** and **QLYPEW** by the proteasome (4). We assessed the processing of this spliced peptide by transfecting the four 293 cell lines 293-SP, 293-SIP, 293-DIP and 293-IP with plasmids encoding gp100 and HLA-A32. Transfected cells were then tested for CTL recognition. After transfection, 293-SP were the most efficient at producing the peptide (Figure 5A), while 293-SIP, DIP and IP were barely recognized. We then digested peptide **RTK**AWNR**QLYPEW**TEAQR using immunocaptured proteasomes or purified proteasomes. Digests were then pulsed onto HLA-A32^{+} cells and tested for their ability to sensitize the CTL. In line with the cell transfection assay, the SP digest efficiently stimulated CTL14, while the IP digest did not and the SIP and DIP were in-between (Figure 5B and 5C).

Digests produced with purified proteasomes and corresponding to approximately 40% of precursor degradation were labeled with TMT and analyzed by HPLC MS/MS (Figure 5D top). Fragment **RTK**, which composes the N-terminal part of the antigenic peptide, was present in all four proteasome digests, although slightly less in the IP digest, and slightly more in the SP and DIP digests(Figure 5D). The C-terminal splicing partner **QLYPEW** was found in all proteasome digests but was slightly more abundant in the SP digest. Two internal destructive cleavages were more intense either in the IP digest (Y_{49}/P_{50}) or in the SIP, DIP and IP digests (L_{48}/Y_{49}). The increased abundance of spliced peptide **RTK_QLYPEW** in the SP digest was confirmed by searching the digest for that peptide by MS/MS (Figure 6A). Increased production of spliced peptide **RTK_QLYPEW** by the SP might be partially explained by an increased abundance of the peptide **QLYPEW** in the SP digest. However, the amounts of **QLYPEW** were only modestly increased, so that this factor was unlikely to be the only explanation. Moreover, when we searched the digest for the presence of the reversed spliced peptide **QLYPEW_RTK**, which was previously described by Epstein et al (9), we surprisingly found that this peptide, which involves the same splicing partners as **RTK_QLYPEW**, was produced

498 in an opposite manner by the four proteasome subtypes, being preferentially processed by the
499 DIP and the IP (Figure 6B). This indicates that the splicing efficiency does not solely depend
500 on the abundance of the splicing partners but also relies on additional factors.
501
502 ***Differential splicing of reciprocal spliced peptides gp100$_{47-52/40-42}$ and gp100$_{40-42/47-52}$***
503 We then explored whether differential splicing of these two peptides could potentially be
504 explained by the fact that the splicing reaction might use C-terminal splice reactants whose size
505 is longer than the minimal C-terminal splicing partner. C-terminally-extended spliced peptides
506 would be preferentially produced by one proteasome subtype and would then be further cleaved
507 by that proteasome to produce the final antigenic peptide.
508
509 Three peptides starting with **QLYPEW** and extended at their C-terminus were detected in the
510 digests: **QLYPEW**TE, **QLYPEW**TEA and **QLYPEW**TEAQR (Figure 5D). Peptide
511 **QLYPEW**TEAQR was found in all digests, while slightly more in SIP. Fragments
512 **QLYPEW**TEA and **QLYPEW**TE were highly released by the standard proteasome, but not
513 by the other proteasome subtypes. In line with the presence of **QLYPEW**TEA only in the SP
514 digest (Figure 5D), the corresponding extended spliced peptide **RTK_QLYPEW**TEA was
515 solely produced by the SP (Figure 7A). In addition, just as SIP was the most efficient at
516 producing the **QLYPEW**TEAQR peptide (Figure 5D), SIP was also the most efficient at
517 producing the extended spliced peptide **RTK_QLYPEW**TEAQR (Figure 7A). So, the
518 increased production of the final antigenic peptide **RTK_QLYPEW** by SP and SIP (Figure
519 6A) might result from a better production of the spliced extended peptide fragments
520 **RTK_QLYPEW**TEAQR or **RTK_QLYPEW**TEA, which would be further trimmed by these
521 proteasomes to release the final antigenic peptide. Such phenomenon was previously shown to
522 be involved in the production of the peptide **RSYVPLAH_R** (8). To verify whether SP and

SIP are able to produce the final trimming step, we incubated the extended peptides **RTK_QLYPEW**TEA and **RTK_QLYPEW**TEAQR with SP and SIP, respectively, and we could detect the final spliced peptide **RTK_QLYPEW** in both digests (Figure 7B). Likewise, when analyzing the spliced reactants potentially involved in the production of the reverse spliced peptide **QLYPEW_RTK**, from the long precursor peptide**,** we observed that increased production of the peptide by DIP and IP might be explained by an increased production of two extended spliced reactants, **RTK**A and **RTK**AW (Figure 5D). Again, the corresponding extended spliced peptides **QLYPEW_RTK**A and **QLYPEW_RTK**AW were also more abundant in the DIP and IP digests (Figure 7C), and the final trimming could be produced by both DIP and IP (Figure 7D). However, we also observed that digestion of extended peptides **RTK_QLYPEW**TEA, **RTK_QLYPEW**TEAQR, **QLYPEW_RTK**A and **QLYPEW_RTK**AW with their respective proteasome subtypes induced major destructive cleavages (Figure 7B and 7D). These cleavages likely impair the production of the final spliced peptides from these extended precursors, suggesting that trimming of extended spliced precursors is not the principal mechanism involved and that other parameters likely regulate the splicing reaction itself.

Another explanation for the differential production of these two spliced peptides could be that, in a given proteasome subtype, the C-terminal splicing partner interacts more readily with the primed binding site (i.e. the site accommodating the amino acids located downstream of the cleavage site) of the catalytic subunit carrying the acyl-enzyme intermediate, thereby facilitating the splicing reaction. To test this hypothesis, we performed in vitro digestions with pairs of peptides incubated with proteasomes. Because the first step of the splicing reaction is based on the production of the acyl-enzyme intermediate, we used the extended peptide fragment **RTK**AWNR, in order to produce the acyl-enzyme intermediate required for the

548 splicing of the **RTK_QLYPEW** peptide. Because most of the acyl-enzyme intermediates end
549 up being hydrolyzed while only a minority undergo peptide splicing, we postulated that the
550 amount of acyl-enzyme intermediate **RTK** bound to the proteasome catalytic subunit was
551 directly proportional to the amount of fragment **RTK** released by hydrolysis. We therefore
552 measured the amount of **RTK** released in the course of digestion (Figure 8A). Digestion times
553 were chosen so as to obtain similar releases of fragment **RTK** and consequently similar
554 amounts of acyl-enzyme intermediate **RTK** in the digests. Furthermore, at the start of the
555 digestion, we added identical amounts of the C-terminal spliced reactant **QLYPEW**, which
556 remained stable in the course of the digestion (Figure 8A). So in the conditions of this assay,
557 equal amounts of both peptide splicing partners were present in the digests, so as to rule out
558 potential influence of the amount of splicing partners available in the reaction. Production of
559 the spliced peptide was then measured by mass spectrometry and in a T-cell assay, after pulsing
560 the digests on HLA-A32^{+} cells (Figure 8A). Interestingly, we observed that in conditions where
561 the amount of acyl-enzyme intermediate **RTK** and **QLYPEW** peptide were roughly the same,
562 SP and SIP were much more efficient at producing the spliced peptide when compared to DIP
563 and IP. We then performed a similar experiment to evaluate the production of the reciprocal
564 spliced peptide **QLYPEW_RTK** (Figure 8B). We digested peptide **QLYPEW**TEAQR,
565 adjusted the digestion times so as to obtain similar amounts of **QLYPEW**-enzyme
566 intermediates, and added identical amounts of C-terminal splice reactant **RTK**. We observed
567 that, despite the presence of similar amounts of splice reactants in the four digests (Figure 8B),
568 the IP was much more efficient at producing spliced peptide **QLYPEW_RTK** when compared
569 to the other proteasome subtypes (Figure 8B). These results are in line with the differential
570 processing observed for the two spliced peptides (**RTK_QLYPEW** being better produced by
571 SP and SIP and **QLYPEW_RTK** being better processed by the IP (Figure 6)). Therefore, in
572 conditions where the amount of splicing partners are identical, splicing efficiency varies

according to the proteasome subtype. This suggests that the splicing reaction might be influenced by the ability of the nucleophile peptide to interact with the primed binding site of the catalytic subunit, where the C-terminal splice reactant likely binds during the peptide splicing reaction.

Discussion

Understanding the rules dictating the establishment of the MHC I peptide repertoire is key to the development of cancer immunotherapy based on vaccination. This can help defining the most efficient vaccination modality to use for a given target antigen. In that sense, the ability of intermediate proteasomes to process tumor antigens is valuable, as these proteasomes are present in significant amounts in both tumors and dendritic cells, which are instrumental to the vaccination approach. Therefore, such intermediate-proteasome-dependent antigens can be efficiently targeted by vaccines encoding full-length proteins, which might be more efficient than single epitope vaccines as they can trigger responses to multiple tumor antigens.

Some years ago, proteasomes were shown to produce antigenic peptides by splicing of peptide fragments originally distant in the parental protein (4-9). Here, we studied the production of six spliced peptides (4-9) by the four proteasome subtypes. We confirmed previous results obtained by our laboratory (26) showing that the standard proteasome produces better peptides FGF-$5_{172-176/217-220}$ (Figure 1), tyrosinase$_{368-373/336-340}$ (Figure 3) and gp$100_{40-42/47-52}$ (Figure 5 and 6A), while the immunoproteasome processes the SP$110_{296-301/286-289}$ better than the standard proteasome (Figure 2). The poor processing of the peptides FGF-$5_{172-176/217-220}$, gp$100_{40-42/47-52}$ and tyrosinase$_{368-373/336-340}$ by IP and intermediate proteasomes suggests that vaccine strategies using these antigens should bypass processing by dendritic cells, which mostly contain these proteasomes. In contrast, the immunoproteasome and DIP process peptides SP$110_{296-301/286-289}$ (Figure 2), gp$100_{195-202/191or192}$ (Figure 4) and gp$100_{47-52/40-42}$ (Figure 6B). This suggests that, in the frame of cancer immunotherapy, both peptides gp$100_{195-202/191or192}$ and gp$100_{47-52/40-42}$ could be targeted by vaccines based on full-length constructs, which require processing by dendritic cells. In particular, this might be important for the design of full-length vaccines based on viral

vectors or mRNA, which are known to produce $CD8^+$ T cell responses and might therefore be key to the design of efficient immunotherapeutic vaccines (39, 40).

The most unexpected observation of our study is the fact that two reciprocal spliced peptides **RTK_QLYPEW** and **QLYPEW_RTK**, which result from the splicing of the very same fragments in the reciprocal orientations, are produced by the four proteasome subtypes in a totally opposite manner: peptide **RTK_QLYPEW** is better produced by the standard proteasome while peptide **QLYPEW_RTK** is more efficiently produced by the IP and the DIP. This suggests that splicing efficiency does not solely depend on the amount of the peptide splicing partners found in the digest, but might also be ruled by other constraints. In line with this, other groups previously observed that, in vitro, splicing could also take place between splicing partners released from minor cleavages (10, 41). To investigate this further, two hypotheses were foreseen. First, we envisioned the possibility that the C-terminal splice reactant might be extended by a few amino acids at its C-terminus. This would lead to a C-terminally extended spliced peptide, which would then need to be further cleaved by the proteasome to produce the final antigenic peptide, as previously observed for peptide **RSYPLAH_R** (8). We did detect extended versions of the spliced peptides in the relevant digests and confirmed the ability of the relevant proteasome to trim those extended peptides. However, we also observed major destructive cleavages inside the extended spliced precursor peptides suggesting that this mechanism does not entirely account for the production of the spliced peptides.

Another mechanism to explain the differential processing of peptides **RTK_QLYPEW** and **QLYPEW_RTK** by the various proteasome subtypes could be that the efficiency of the splicing reaction varies according to the splicing partners and the proteasome subtype involved. Whether peptide splicing follows specific rules is a matter of intense debate (10, 41, 42). Based

on the fact that peptide splicing occurs by transpeptidation and involves the nucleophilic attack of an acyl-enzyme intermediate by a peptide present in the proteasome chamber, it is expected that splicing efficiency depends on the speed of the reaction, which is linked to the concentration of both splicing partners. Surprisingly, Mishto and others previously observed that in *in vitro* digests, peptide splicing often relies on minor proteasome cleavages (10, 41), suggesting that proteasome-catalyzed peptide splicing might obey specific rules that depend on the nature of the peptide involved in the peptide splicing reaction. Only a few studies investigated the biochemical constraints of proteasome-catalyzed peptide splicing. Berkers et al. digested pairs of peptides in which they sequentially modified one amino acid either in the acyl-enzyme intermediate or in the sequence of the C-terminal splice reactant. Splicing of the fragment was quantified by measuring the ability of the spliced peptide to bind to HLA-A2 in a fluorescence polarization assay. Although the results were limited to the HLA-A2 allele, their work suggested that splicing mostly depends on the sequence of the N-terminal reactant and on the concentration of the C-terminal reactant (42). Mishto *et al.* also suggested that a longer retention time of the N-terminal reactant at the non-primed site in the catalytic subunit might be compulsory for the peptide splicing reaction (10). More recently, another group studied the 20S standard proteasome splicing efficiency after digestion of 25 polypeptides of different lengths (41). Using a MS-based de novo sequencing approach they identified and characterized forward and reverse spliced peptides obtained after digestion with standard proteasome and determined specific splicing signatures based on the length and sequence of splice reactants.

To date, no specific rules have been defined regarding the ability of the different proteasome subtypes to splice peptides. Surprisingly, our experiments showed that when incubating peptide **RTK**AWNR and **QLYPEW** with SP, SIP, DIP or IP, in conditions in which a similar level of **RTK** hydrolysis occurred (and therefore similar amounts of **RTK** acyl-enzyme intermediate

were produced), spliced peptide **RTK_QLYPEW** was produced by SP and SIP but not by DIP or IP, even though similar amounts of the C-terminal splice reactant were available for the reaction (Figure 8). In a remarkable mirror image, when we similarly incubated peptides **QLYPEW**TEAQR and **RTK** in conditions where the amount of **QLYPEW** and **RTK** remained identical between the four proteasome subtypes, the spliced peptide **QLYPEW_RTK** was mostly produced by the IP (Figure 8). These results suggest that the ability of a C-terminal reactant to splice is influenced by the type of proteasome involved, likely through differences in the amino acid composition and/or structure of the catalytic pockets found in those proteasomes. By essence, transpeptidation involves the nucleophilic attack of an acyl-enzyme intermediate by a C-terminal peptide reactant found in the proteasome chamber. In order to perform the nucleophilic attack, the C-terminal peptide reactant has to localize in close vicinity to the acyl-enzyme intermediate. Previously, it was suggested that peptide splicing might be facilitated by the existence of a nucleophile binding site (site δ), distinct from the primed binding site, that would be located in close vicinity to the catalytic N-terminal threonine and would be occupied by the C-terminal splice reactant. However, no clear evidence for the existence of such an alternative binding site was put forward thus far, and it is likely that the C-terminal splice reactant rather binds to the primed binding site of the catalytic subunit harboring the acyl-enzyme, and is therefore adequately positioned for the nucleophilic attack. Binding of the C-terminal splice reactant to the primed binding site is likely influenced by the nature of the amino acids lining that site. In agreement with this, we previously showed that the peptide causing the nucleophilic attack on the acyl-enzyme intermediate must comprise at least 3 amino acids in order to give rise to a spliced peptide, suggesting that the C-terminal splice reactant must indeed be sufficiently long to properly bind to the primed binding site (8). The increased ability of the SP and SIP to produce the spliced peptide **RTK_QLYPEW** is likely linked to the presence of the β1 subunit found in the SP and the SIP but not in the DIP or the IP. On the other

hand, increased production of peptide **QLYPEW_RTK** by IP likely depends on the presence of the β2i subunit found in IP but not in SP, SIP or DIP. If one admits that the β1 (for **RTK_QLYPEW**) and β2i (for **QLYPEW_RTK**) subunits are the catalytic subunits where the splicing reaction takes place, it follows that, in conditions where the amount of acyl-enzyme intermediates are identical, increased splicing is likely related to an increased availability of the C-terminal splice reactant at the primed binding site of that catalytic subunit as a result of an increased affinity of the nucleophile for that site.

It is important to note however that splicing of peptide **RTK_QLYPEW** might rather require the trypsin-like activity of the β2 or the β2i subunit (and not the β1 subunit) for the production of the acyl-enzyme intermediate involving **RTK**. Likewise, production of the spliced peptide **QLYPEW_RTK** likely relies on the chymotrypsin-like activity of the β5 or the β5i subunit. Thus, the catalytic subunits that favor peptide splicing do not appear to correspond to those expected to produce the acyl-enzyme intermediate. One explanation might be that incorporation of the β1i or the β2 subunit induces a modification of the structure of the primed binding site of the β2 or β5i subunit, respectively, thereby impairing the binding of the **QLYPEW** and **RTK** peptides. Alternatively, incorporation of the β1i or the β2 subunit could partly prevent peptides **QLYPEW** and **RTK** from reaching the primed binding site of the catalytic subunit harboring the acyl-enzyme intermediate, thereby decreasing the local concentration of these peptides at the splicing site. This could happen if the C-terminal spliced reactants preferentially interacted with β1i or β2, rather than with the primed binding site of the catalytic subunit where peptide splicing takes place. Both models suggest that affinity of the C-terminal splicing partner for the primed site of the catalytic subunit harboring the acyl-enzyme might be another factor determining the efficiency of peptide splicing.

Crystallographic analysis of yeast proteasome bound to the inhibitor homobelastosin C helped defining the primed binding sites of subunits β5 and β5i, and showed striking differences in their sequence and structure. Similar differences were later observed by comparing the crystal structure of murine constitutive and immunoproteasomes (43, 44). The putative primed binding sites of β1 and β2 were modeled based on that of subunit β5. Analysis of their crystal structure in murine proteasomes revealed that the overall structure of the β2 and β2i primed sites appeared identical, while the primed site of β1i was shortened when compared to that of β1 (44). In both cases, however, the primed sites exhibited significant sequence differences that might affect their potency to interact with peptides. A better definition of the primed binding site of the catalytic subunits of the four proteasome subtypes might, in the future, help to better understand the origin of these differences in splicing efficiency.

Finally, although the role of the N-terminal splicing partner was not directly investigated here, it is another crucial parameter to take into account, since it will define which of the three catalytic subunits will be involved in the formation of the acyl-enzyme intermediate and consequently the nature of the primed site to which the C-terminal splice reactant will bind. The affinity of the N-terminal splicing partner for the non-primed site of the catalytic pocket could also play a role in the splicing reaction, but in the experiments reported here, we suppressed this effect by controlling the amount of N-terminal splicing partner produced.

Altogether our results suggest that efficiency of a splicing reaction depends on several parameters among which (1) the amount of peptide fragments available for the splicing reaction including extended fragments that may be subsequently trimmed, (2) internal destructive cleavages that could destroy these fragments or the spliced peptide itself and (3) the affinity of

724 the C-terminal splice reactant for the primed binding site of the catalytic subunit. It is likely
725 the conjunction of these parameters that modulates the production of spliced peptides.
726

727 Acknowledgments

728 We thank Rui Cheng, Thérèse Aerts and Giuseppe Attardo for excellent technical assistance
729 and Isabelle Grisse for editorial assistance.
730

731 References

732 1. Rock, K. L., and A. L. Goldberg. 1999. Degradation of cell proteins and the
733 generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* 17: 739-779.
734 2. Löwe, J., D. Stock, B. Jap, P. Zwickl, W. Baumeister, and R. Huber. 1995. Crystal
735 structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution.
736 *Science* 268: 533-539.
737 3. Unno, M., T. Mizushima, Y. Morimoto, Y. Tomisugi, K. Tanaka, N. Yasuoka, and T.
738 Tsukihara. 2002. The structure of the mammalian 20S proteasome at 2.75 A
739 resolution. *Structure* 10: 609-618.
740 4. Vigneron, N., V. Stroobant, J. Chapiro, A. Ooms, G. Degiovanni, S. Morel, P. van der
741 Bruggen, T. Boon, and B. J. Van den Eynde. 2004. An antigenic peptide produced by
742 peptide splicing in the proteasome. *Science* 304: 587-590.
743 5. Hanada, K., J. W. Yewdell, and J. C. Yang. 2004. Immune recognition of a human
744 renal cancer antigen through post-translational protein splicing. *Nature* 427: 252-256.
745 6. Warren, E. H., N. J. Vigneron, M. A. Gavin, P. G. Coulie, V. Stroobant, A. Dalet, S.
746 S. Tykodi, S. M. Xuereb, J. K. Mito, S. R. Riddell, and B. J. Van den Eynde. 2006. An
747 antigen produced by splicing of noncontiguous peptides in the reverse order. *Science*
748 313: 1444-1447.
749 7. Dalet, A., P. F. Robbins, V. Stroobant, N. Vigneron, Y. F. Li, M. El-Gamil, K.
750 Hanada, J. C. Yang, S. A. Rosenberg, and B. J. Van den Eynde. 2011. An antigenic
751 peptide produced by reverse splicing and double asparagine deamidation. *Proc. Natl.*
752 *Acad. Sci. U. S. A/* 108: E323-3331.
753 8. Michaux, A., P. Larrieu, V. Stroobant, J. F. Fonteneau, F. Jotereau, B. J. Van den
754 Eynde, A. Moreau-Aubry, and N. Vigneron. 2014. A spliced antigenic peptide
755 comprising a single spliced amino acid is produced in the proteasome by reverse
756 splicing of a longer peptide fragment followed by trimming. *J. Immunol.* 192: 1962-
757 1971.
758 9. Ebstein, F., K. Textoris-Taube, C. Keller, R. Golnik, N. Vigneron, B. J. Van den
759 Eynde, B. Schuler-Thurner, D. Schadendorf, F. K. Lorenz, W. Uckert, S. Urban, A.
760 Lehmann, N. Albrecht-Koepke, K. Janek, P. Henklein, A. Niewienda, P. M. Kloetzel,
761 and M. Mishto. 2016. Proteasomes generate spliced epitopes by two different
762 mechanisms and as efficiently as non-spliced epitopes. *Sci. Rep.* 6: 24032.
763 10. Mishto, M., A. Goede, K. T. Taube, C. Keller, K. Janek, P. Henklein, A. Niewienda,
764 A. Kloss, S. Gohlke, B. Dahlmann, C. Enenkel, and P. M. Kloetzel. 2012. Driving

765 forces of proteasome-catalyzed peptide splicing in yeast and humans. *Mol. Cell*
766 *Proteomics* 11: 1008-1023.
767 11. Dalet, A., N. Vigneron, V. Stroobant, K. Hanada, and B. J. Van den Eynde. 2010.
768 Splicing of distant peptide fragments occurs in the proteasome by transpeptidation and
769 produces the spliced antigenic peptide derived from fibroblast growth factor-5. *J.*
770 *Immunol.* 184: 3016-3024.
771 12. Sykulev, Y., M. Joo, I. Vturina, T. J. Tsomides, and H. N. Eisen. 1996. Evidence that
772 a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response.
773 *Immunity* 4: 565-571.
774 13. Robbins, P. F., M. el-Gamil, Y. Kawakami, E. Stevens, J. R. Yannelli, and S. A.
775 Rosenberg. 1994. Recognition of tyrosinase by tumor-infiltrating lymphocytes from a
776 patient responding to immunotherapy. *Cancer Res.* 54: 3124-3126.
777 14. Chong, C., F. Marino, H. Pak, J. Racle, R. T. Daniel, M. Müller, D. Gfeller, G.
778 Coukos, and M. Bassani-Sternberg. 2018. High-throughput and Sensitive
779 Immunopeptidomics Platform Reveals Profound Interferonγ-Mediated Remodeling of
780 the Human Leukocyte Antigen (HLA) Ligandome. *Mol. Cell. Proteomics* 17: 533-
781 548.
782 15. Liepe, J., F. Marino, J. Sidney, A. Jeko, D. E. Bunting, A. Sette, P. M. Kloetzel, M. P.
783 Stumpf, A. J. Heck, and M. Mishto. 2016. A large fraction of HLA class I ligands are
784 proteasome-generated spliced peptides. *Science* 354: 354-358.
785 16. Faridi, P., C. Li, S. H. Ramarathinam, J. P. Vivian, P. T. Illing, N. A. Mifsud, R.
786 Ayala, J. Song, L. J. Gearing, P. J. Hertzog, N. Ternette, J. Rossjohn, N. P. Croft, and
787 A. W. Purcell. 2018. A subset of HLA-I peptides are not genomically templated:
788 Evidence for cis- and trans-spliced peptide ligands. *Sci. Immunol.* 3: eaar3947.
789 17. Mylonas, R., I. Beer, C. Iseli, C. Chong, H. S. Pak, D. Gfeller, G. Coukos, I. Xenarios,
790 M. Muller, and M. Bassani-Sternberg. 2018. Estimating the Contribution of
791 Proteasomal Spliced Peptides to the HLA-I Ligandome. *Mol. Cell Proteomics* 17:
792 2347-2357.
793 18. Liepe, J., J. Sidney, F. K. M. Lorenz, A. Sette, and M. Mishto. 2019. Mapping the
794 MHC Class I-Spliced Immunopeptidome of Cancer Cells. *Cancer Immunol. Res.* 7:
795 62-76.
796 19. Paes, W., G. Leonov, T. Partridge, T. Chikata, H. Murakoshi, A. Frangou, S.
797 Brackenridge, A. Nicastri, A. G. Smith, G. H. Learn, Y. Li, R. Parker, S. Oka, P.
798 Pellegrino, I. Williams, B. F. Haynes, A. J. McMichael, G. M. Shaw, B. H. Hahn, M.
799 Takiguchi, N. Ternette, and P. Borrow. 2019. Contribution of proteasome-catalyzed
800 peptide cis-splicing to viral targeting by CD8(+) T cells in HIV-1 infection. *Proc.*
801 *Natl. Acad. Sci. U. S. A.* 116: 24748-24759.
802 20. Lichti, C. F., N. Vigneron, K. R. Clauser, B. J. Van den Eynde, and M. Bassani-
803 Sternberg. 2022. Navigating Critical Challenges Associated with
804 Immunopeptidomics-Based Detection of Proteasomal Spliced Peptide Candidates.
805 *Cancer Immunol. Res.*
806 21. Erhard, F., L. Dolken, B. Schilling, and A. Schlosser. 2020. Identification of the
807 Cryptic HLA-I Immunopeptidome. *Cancer Immunol. Res.* 8: 1018-1026.
808 22. Dick, T. P., A. K. Nussbaum, M. Deeg, W. Heinemeyer, M. Groll, M. Schirle, W.
809 Keilholz, S. Stevanovic, D. H. Wolf, R. Huber, H. G. Rammensee, and H. Schild.
810 1998. Contribution of proteasomal beta-subunits to the cleavage of peptide substrates
811 analyzed with yeast mutants. *J. Biol. Chem.* 273: 25637-25646.
812 23. Gaczynska, M., K. L. Rock, and A. L. Goldberg. 1993. Gamma-interferon and
813 expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature* 365:
814 264-267.

24. Gaczynska, M., K. L. Rock, T. Spies, and A. L. Goldberg. 1994. Peptidase activities of proteasomes are differentially regulated by the major histocompatibility complex-encoded genes for LMP2 and LMP7. *Proc. Natl. Acad. Sci. U. S. A.* 91: 9213-9217.
25. Guillaume, B., J. Chapiro, V. Stroobant, D. Colau, B. Van Holle, G. Parvizi, M. P. Bousquet-Dubouch, I. Theate, N. Parmentier, and B. J. Van den Eynde. 2010. Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proc. Natl. Acad. Sci. U. S. A.* 107: 18599-18604.
26. Dalet, A., V. Stroobant, N. Vigneron, and B. J. Van den Eynde. 2011. Differences in the production of spliced antigenic peptides by the standard proteasome and the immunoproteasome. *Eur. J. Immunol.* 41: 39-46.
27. Murata, M., E. H. Warren, and S. R. Riddell. 2003. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. *J. Exp. Med.* 197: 1279-1289.
28. Uyttenhove, C., L. Pilotte, I. Theate, V. Stroobant, D. Colau, N. Parmentier, T. Boon, and B. J. Van den Eynde. 2003. Evidence for a tumoral immune resistance mechanism based on tryptophan degradation by indoleamine 2,3-dioxygenase. *Nat. Med.* 9: 1269-1274.
29. Vigneron, N., J. Abi Habib, and B. J. Van den Eynde. 2015. The capture proteasome assay: A method to measure proteasome activity in vitro. *Anal. Biochem.* 482: 7-15.
30. Vigneron, N., J. Abi Habib, and B. J. Van den Eynde. 2015. The capture proteasome assay (CAPA) to evaluate subtype-specific proteasome inhibitors. *Data Brief* 4: 146-151.
31. Hendil, K. B., and W. Uerkvitz. 1991. The human multicatalytic proteinase: affinity purification using a monoclonal antibody. *J. Biochem. Biophys. Methods* 22: 159-165.
32. Schmidtke, G., S. Emch, M. Groettrup, and H. G. Holzhutter. 2000. Evidence for the existence of a non-catalytic modifier site of peptide hydrolysis by the 20 S proteasome. *J Biol Chem* 275: 22056-22063.
33. Abi Habib, J., E. De Plaen, V. Stroobant, D. Zivkovic, M. P. Bousquet, B. Guillaume, K. Wahni, J. Messens, A. Busse, N. Vigneron, and B. J. Van den Eynde. 2020. Efficiency of the four proteasome subtypes to degrade ubiquitinated or oxidized proteins. *Sci. Rep.* 10: 15765.
34. Barsnes, H., and M. Vaudel. 2018. SearchGUI: A Highly Adaptable Common Interface for Proteomics Search and de Novo Engines. *J. Proteome Res.* 17: 2552-2555.
35. Vaudel, M., J. M. Burkhart, R. P. Zahedi, E. Oveland, F. S. Berven, A. Sickmann, L. Martens, and H. Barsnes. 2015. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat. Biotechnol.* 33: 22-24.
36. Griffin, T. A., D. Nandi, M. Cruz, H. J. Fehling, L. V. Kaer, J. J. Monaco, and R. A. Colbert. 1998. Immunoproteasome assembly: cooperative incorporation of interferon gamma (IFN-gamma)-inducible subunits. *J. Exp. Med.* 187: 97-104.
37. Valmori, D., U. Gileadi, C. Servis, P. R. Dunbar, J. C. Cerottini, P. Romero, V. Cerundolo, and F. Lévy. 1999. Modulation of proteasomal activity required for the generation of a cytotoxic T lymphocyte-defined peptide derived from the tumor antigen MAGE-3. *J. Exp. Med.* 189: 895-906.
38. Cardozo, C., A. Vinitsky, C. Michaud, and M. Orlowski. 1994. Evidence that the nature of amino acid residues in the P3 position directs substrates to distinct catalytic sites of the pituitary multicatalytic proteinase complex (proteasome). *Biochemistry* 33: 6483-6489.
39. McAuliffe, J., H. F. Chan, L. Noblecourt, R. A. Ramirez-Valdez, V. Pereira-Almeida, Y. Zhou, E. Pollock, F. Cappuccini, I. Redchenko, A. V. Hill, C. S. K. Leung, and B.

J. Van den Eynde. 2021. Heterologous prime-boost vaccination targeting MAGE-type antigens promotes tumor T-cell infiltration and improves checkpoint blockade therapy. *J. Immunother. Cancer* 9: e003218.

40. Sahin, U., P. Oehm, E. Derhovanessian, R. A. Jabulowsky, M. Vormehr, M. Gold, D. Maurus, D. Schwarck-Kokarakis, A. N. Kuhn, T. Omokoko, L. M. Kranz, M. Diken, S. Kreiter, H. Haas, S. Attig, R. Rae, K. Cuk, A. Kemmer-Bruck, A. Breitkreuz, C. Tolliver, J. Caspar, J. Quinkhardt, L. Hebich, M. Stein, A. Hohberger, I. Vogler, I. Liebig, S. Renken, J. Sikorski, M. Leierer, V. Muller, H. Mitzel-Rink, M. Miederer, C. Huber, S. Grabbe, J. Utikal, A. Pinter, R. Kaufmann, J. C. Hassel, C. Loquai, and O. Tureci. 2020. An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. *Nature* 585: 107-112.

41. Paes, W., G. Leonov, T. Partridge, A. Nicastri, N. Ternette, and P. Borrow. 2020. Elucidation of the Signatures of Proteasome-Catalyzed Peptide Splicing. *Front. Immunol.* 11: 563800.

42. Berkers, C. R., A. de Jong, K. G. Schuurman, C. Linnemann, H. D. Meiring, L. Janssen, J. J. Neefjes, T. N. Schumacher, B. Rodenko, and H. Ovaa. 2015. Definition of Proteasomal Peptide Splicing Rules for High-Efficiency Spliced Peptide Presentation by MHC Class I Molecules. *J. Immunol.* 195: 4085-4095.

43. Groll, M., O. V. Larionov, R. Huber, and A. de Meijere. 2006. Inhibitor-binding mode of homobelactosin C to proteasomes: new insights into class I MHC ligand generation. *Proc. Natl. Acad. Sci. U. S. A.* 103: 4576-4579.

44. Huber, E. M., M. Basler, R. Schwab, W. Heinemeyer, C. J. Kirk, M. Groettrup, and M. Groll. 2012. Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell* 148: 727-738.

Footnote

VF is supported by a fellowship from the FNRS-FRIA (Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture, grant No. 1.E091.14). JAH is supported by a fellowship from the Fonds National de la Recherche Scientifique (FNRS) (TELEVIE grant No. 7455115F). SN is supported by an EOS grant (FNRS/FWO) (EOS O000518F). This work received support from the FNRS, Belgium (Grants EOS O000518F and PDR T.0091.18); the Walloon excellence in life sciences and biotechnology (WELBIO-CR-2010-11R) and the Fonds Joseph Maisin, Belgium.

Figure Legends

Figure 1. Impact of the four proteasome subtypes on the processing of spliced peptide FGF-5$_{172\text{-}176/217\text{-}220}$ NTYAS_PRFK. (**A**) Presentation of spliced peptide **NTYAS_PRFK** by 293-EBNA cells expressing different proteasome subtypes. Cells were transfected with HLA-A3 and the indicated amount of plasmid encoding the FGF-5. As control, cells were pulsed with 1µM **NTYAS_PRFK** peptide. FGF-5-specific CTL clone C2 was added and TNF production was measured after 16h co-culture. Error bars show standard deviations of duplicate samples. One representative out of three experiments is shown. (**B,C**) Detection of peptide **NTYAS_PRFK** in digests obtained by incubating precursor peptide **NTYAS**AIHRTSTHFL**PRFK**QSEQP with immunocaptured proteasomes (CAPA assay) (**B**) or purified 20S proteasomes (**C**). Digests were loaded on HLA-A3+ CAN-EBV cells and the FGF-5 specific CTL C2 was added. IFN-γ was measured after an overnight co-culture. Error bars show standard deviations of duplicates. One of three representative experiments is shown. In (C), recognition by the CTL of digests displaying about 50% of precursor degradation (quantified in D) is shown. (**D**) Quantitative analysis of the fragments present in the 210-min SP digest, the 90-min SIP and DIP digests and the 120-min IP digest, representing about 50%

of precursor degradation. Digests were differentially labeled with TMT, pooled together and the relative abundance of the precursor peptide and the fragments that are relevant for the production of the antigenic peptide was evaluated by HPLC-MS/MS.

Figure 2. Impact of the four proteasome subtypes on the processing of spliced peptide $SP110_{296\text{-}301/286\text{-}289}$ SLPRGT_STPK. (**A**) Presentation of spliced peptide **SLPRGT_STPK** by 293-EBNA cells expressing different proteasome subtypes. Cells were transfected with HLA-A3 and the indicated amount of plasmid encoding SP110. As control, cells were pulsed with 1µM **SLPRGT_STPK** peptide. SP110-specific CTL DRN-7 was added and TNF production was measured after 16h co-culture. Error bars show standard deviations of duplicate samples. One representative out of three experiments is shown. (**B,C**) Detection of peptide **SLPRGT_STPK** in digests obtained by incubating precursor peptide **STPK**RRHKKK**SLPRGT**ASSR with immunocaptured proteasomes (CAPA assay) (**B**) or purified 20S proteasomes (**C**). Digests were loaded on HLA-$A3^+$ CAN-EBV cells and the SP110-specific CTL DRN-7 was added. IFN-γ was measured after an overnight co-culture. Error bars show standard deviations of duplicates. One of three representative experiments is shown. In (C), recognition by the CTL of digests displaying about 50% of precursor degradation (quantified in D) is shown. (**D**) Quantitative analysis of the fragments present in the 480-min SP, SIP and IP digests and the 300-min DIP digest, corresponding to about 50% of the precursor degradation. Digests were differentially labeled with TMT, pooled together and the relative abundance of the precursor peptide and the fragments that are relevant for the production of the antigenic peptide was evaluated by HPLC-MS/MS.

Figure 3. Impact of the four proteasome subtypes on the processing of spliced peptide tyrosinase$_{368\text{-}373/336\text{-}340}$ IYMDGT_ADFSF. (**A**) Presentation of spliced peptide **IYMDGT_ADFSF** by 293-EBNA cells expressing different proteasome subtypes. Cells were transfected with HLA-A24 and the indicated amount of plasmid encoding tyrosinase. As control, cells were pulsed with 1µM **IYMDGT_ADFSF** peptide. Tyrosinase-specific CTL 888 cl62 was added and TNF production was measured after 16h co-culture. Error bars show standard deviations of duplicate samples. One representative out of three experiments is shown. (**B,C**) Detection of peptide **IYMDGT_ADFSF** in digests obtained by incubating precursor peptide **ADFSF**RNTLEHNALH**IYMDGT**MSQV with immunocaptured proteasomes (CAPA assay) (**B**) or purified 20S proteasomes (**C**). Digests were loaded on HLA-A24^{+} LG2-EBV cells and tyrosinase-specific CTL 888 cl62 was added. IFN-γ was measured after an overnight co-culture. Error bars show standard deviations of duplicates. One of three representative experiments is shown. In (C), recognition by the CTL of digests displaying about 40% of precursor degradation (quantified in D) is shown. (**D**) Quantitative analysis of the fragments present in the 60-min SP and IP digests and the 90-min SIP and DIP digests, corresponding to about 40% of the precursor degradation. Digests were differentially labeled with TMT, pooled together and the relative abundance of the precursor peptide and the fragments that are relevant for the production of the antigenic peptide was evaluated by HPLC-MS/MS.

Figure 4. Impact of the four proteasome subtypes on the processing of spliced peptide gp100 R$_{195\text{-}202/191or192}$ RSYVPLAH_. (**A**) Presentation of spliced peptide **RSYVPLAH_R** by 293-EBNA cells expressing different proteasome subtypes. Cells were transfected with HLA-A3 and the indicated amount of plasmid encoding gp100. As control, cells were pulsed with 1µM **RSYVPLAH_R** peptide. Gp100-specific CTL clone M45-3B was added and TNF production was measured after 16h co-culture. Error bars show standard deviations of duplicate

samples. One representative out of three experiments is shown. (**B,C**) Detection of peptide **RSYVPLAH_R** in digests obtained by incubating precursor peptide **R**RGS**RSYVPLAH**SSSAFT with immunocaptured proteasomes (CAPA assay) (**B**) or purified 20S proteasomes (**C**). Digests were loaded on HLA-A3^{+} CAN-EBV cells and the gp100-specific CTL M45-3B was added. IFN-γ was measured after an overnight co-culture. Error bars show standard deviations of duplicates. One of three representative experiments is shown. In (C), recognition by the CTL of digests displaying about 50% of precursor degradation (quantified in D) is shown. (**D**) Quantitative analysis of the fragments present in the 240-min SP and SIP digests and the 30-min DIP and IP digests, corresponding to about 50% of the precursor degradation. Digests were differentially labeled with TMT, pooled together and the relative abundance of the precursor peptide and the fragments that are relevant for the production of the antigenic peptide was evaluated by HPLC-MS/MS.

Figure 5. Impact of the four proteasome subtypes on the processing of spliced peptide gp100$_{40\text{-}42/47\text{-}52}$ RTK_QLYPEW. (**A**) Presentation of spliced peptide **RTK_QLYPEW** by 293-EBNA cells expressing different proteasome subtypes. Cells were transfected with HLA-A32 and the indicated amount of plasmid encoding gp100. As control, cells were pulsed with 1μM **RTK_QLYPEW** peptide. Gp100-specific CTL clone 14 was added and TNF production was measured after 16h co-culture. Error bars show standard deviations of duplicate samples. One representative out of three experiments is shown. (**B,C**) Detection of peptide **RTK_QLYPEW** in digests obtained by incubating precursor peptide **RTK**AWNR**QLYPEW**TEAQR with immunocaptured proteasomes (CAPA assay) (**B**) or purified 20S proteasomes (**C**). Digests were loaded on HLA-A32^{+} LG2-EBV cells and gp100-specific CTL 14 was added. IFN-γ was measured after an overnight co-culture. Error bars show standard deviations of duplicates. One of three representative experiments is shown. In (C),

recognition by the CTL of digests displaying about 40% of precursor degradation (quantified in D) is shown. (**D**) Quantitative analysis of the fragments present in the 120-min SP, 90-min SIP and DIP digests and the 190-min IP digest, corresponding to about 40% of the precursor degradation. Digests were differentially labeled with TMT, pooled together and the relative abundance of the precursor peptide and the fragments that are relevant for the production of the antigenic peptide was evaluated by HPLC-MS/MS.

Figure 6. MS/MS detection of the reciprocal spliced peptides RTK_QLYPEW and QLYPEW_RTK in proteasome digests. MS/MS detection of RTK_QLYPEW (A) or QLYPEW_RTK (B) in the digests of figure 5D.

Figure 7. Production of RTK_QLYPEW and QLYPEW_RTK from extended spliced peptide precursors. (**A**) MS/MS detection of **RTK_QLYPEW**TEA and **RTK_QLYPEW**TEAQR in the digests of figure 5D. (**B**) MS detection of the indicated peptides in digests of precursor peptides **RTKQLYPEW**TEA and **RTKQLYPEW**TEAQR by SP and SIP respectively (60-min digests). (**C**) MS/MS detection of **QLYPEWRTK**A and **QLYPEWRTK**AW in the digests of figure 5D. (**D**) MS detection of the peptides **QLYPEW_RTK** and **QLYPEW** in digests of precursor peptides **QLYPEWRTK**A and **QLYPEWRTK**AW by DIP and IP (60-min digests).

Figure 8. Splicing efficiency of RTK_QLYPEW and QLYPEW_RTK. (**A**) Quantification of **RTK** and **QLYPEW** in SP, SIP, DIP and IP digests (90-min SP and DIP, 30-min SIP and 210-min IP) performed with peptide pair **RTK**AWNR + **QLYPEW**. Digestion times were chosen so that release of fragment RTK was similar between the four digests. Digests were differentially labeled with TMT, pooled together and the relative abundance of **RTK** (top panel), **QLYPEW** (mid-high panel) and **RTK_QLYPEW** (mid-low panel) was evaluated by

HPLC-MS/MS. The digests were also loaded on HLA-A32^{+} LG2-EBV cells and CTL 14 was added. IFN-γ was measured by ELISA after overnight co-culture (bottom panel). (**B**) Quantification of **QLYPEW** and **RTK** in SP, SIP, DIP and IP digests (240-min SP, 60-min SIP, 360-min DIP and 210-min IP) performed with peptide pair **QLYPEW**TEAQR + **RTK**. Digestion times were chosen so that release of fragment QLYPEW was similar between the four digests. Digests were differentially labeled with TMT, pooled together and the relative abundance of **QLYPEW** (top panel), **RTK** (middle panel) and **QLYPEW_RTK** (bottom panel) was evaluated by HPLC-MS/MS.

Figure 1

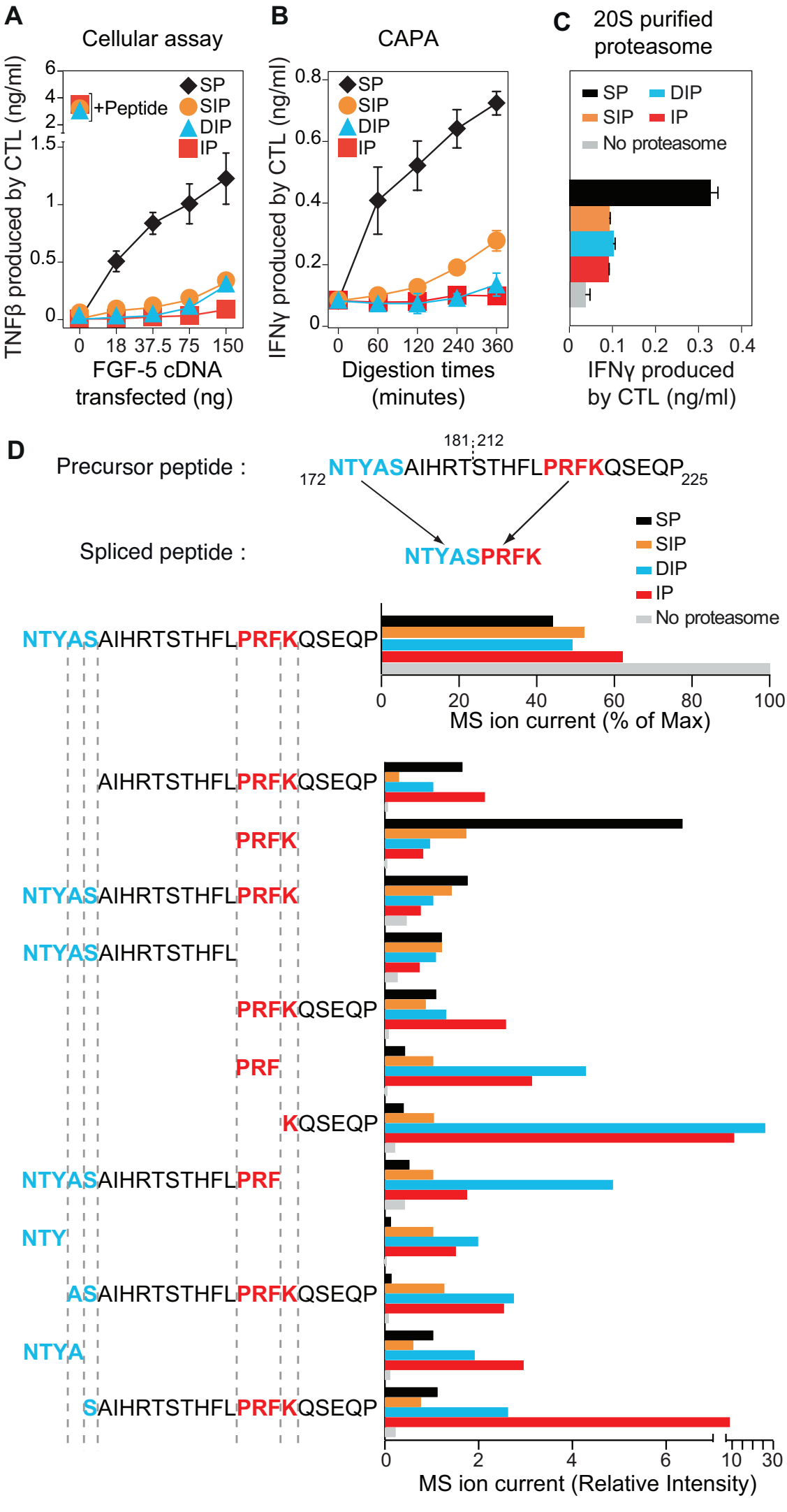

A
Cellular assay
TNFβ produced by CTL (ng/ml)
+Peptide
SP
SIP
DIP
IP
FGF-5 cDNA transfected (ng)
B
CAPA
IFNγ produced by CTL (ng/ml)
Digestion times (minutes)
C
20S purified proteasome
No proteasome
IFNγ produced by CTL (ng/ml)
D
Precursor peptide :
Spliced peptide :
NTYASPRFK
MS ion current (% of Max)
MS ion current (Relative Intensity)

Figure 2

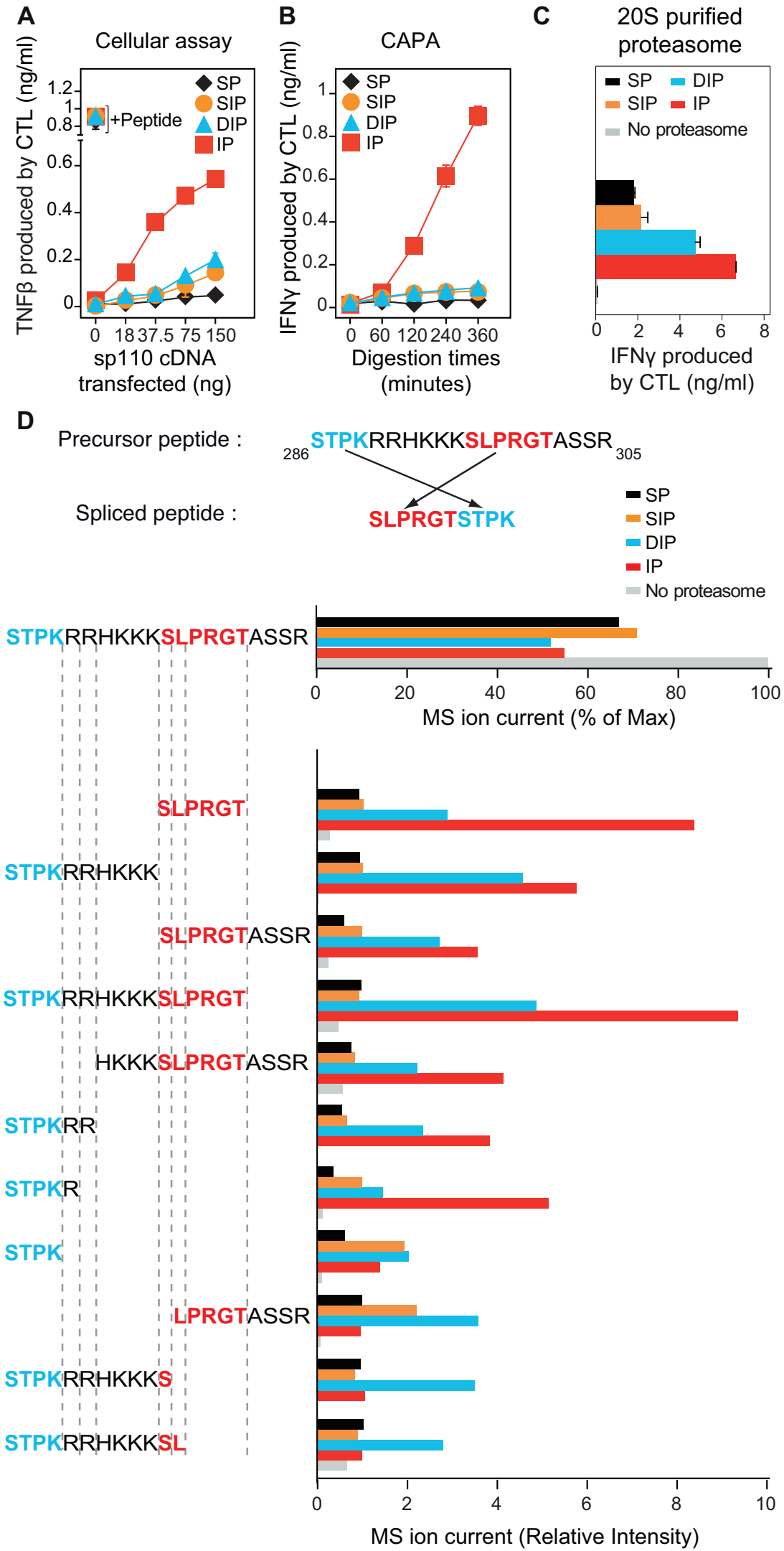

A
Cellular assay
TNFβ produced by CTL (ng/ml)
+Peptide
SP
SIP
DIP
IP
sp110 cDNA transfected (ng)
B
CAPA
IFNγ produced by CTL (ng/ml)
Digestion times (minutes)
C
20S purified proteasome
No proteasome
IFNγ produced by CTL (ng/ml)
D
Precursor peptide :
STPKRRHKKKSLPRGTASSR
Spliced peptide :
SLPRGTSTPK
MS ion current (% of Max)
SLPRGT
STPKRRHKKK
SLPRGTASSR
STPKRRHKKKSLPRGT
HKKKSLPRGTASSR
STPKRR
STPKR
STPK
LPRGTASSR
STPKRRHKKKS
STPKRRHKKKSL
MS ion current (Relative Intensity)

Figure 3

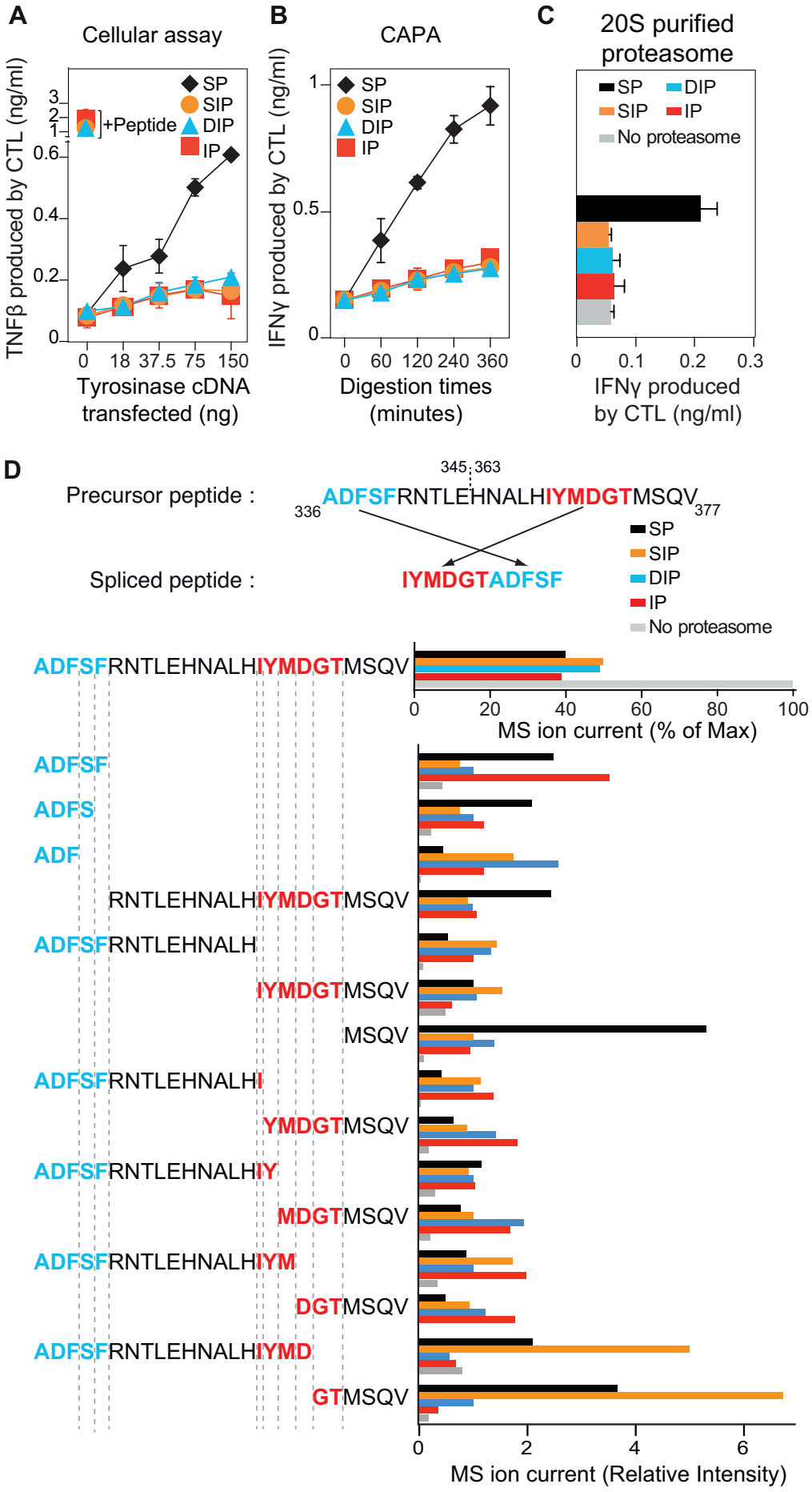
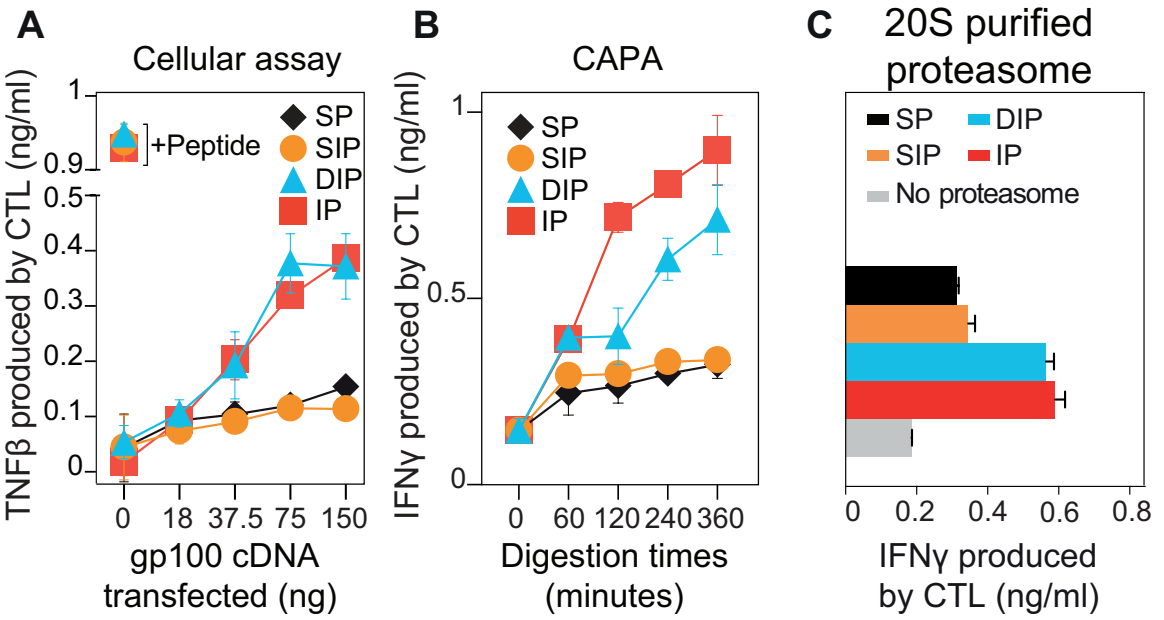

A
Cellular assay
TNFβ produced by CTL (ng/ml)
SP
SIP
DIP
IP
+Peptide
Tyrosinase cDNA transfected (ng)
B
CAPA
IFNγ produced by CTL (ng/ml)
Digestion times (minutes)
C
20S purified proteasome
No proteasome
IFNγ produced by CTL (ng/ml)
D
Precursor peptide :
ADFSFRNTLEHNALHIYMDGTMSQV
Spliced peptide :
IYMDGTADFSF
MS ion current (% of Max)
MS ion current (Relative Intensity)

Figure 4


A
Cellular assay
TNFβ produced by CTL (ng/ml)
+Peptide
SP
SIP
DIP
IP
gp100 cDNA transfected (ng)
B
CAPA
IFNγ produced by CTL (ng/ml)
Digestion times (minutes)
C
20S purified proteasome
No proteasome
IFNγ produced by CTL (ng/ml)

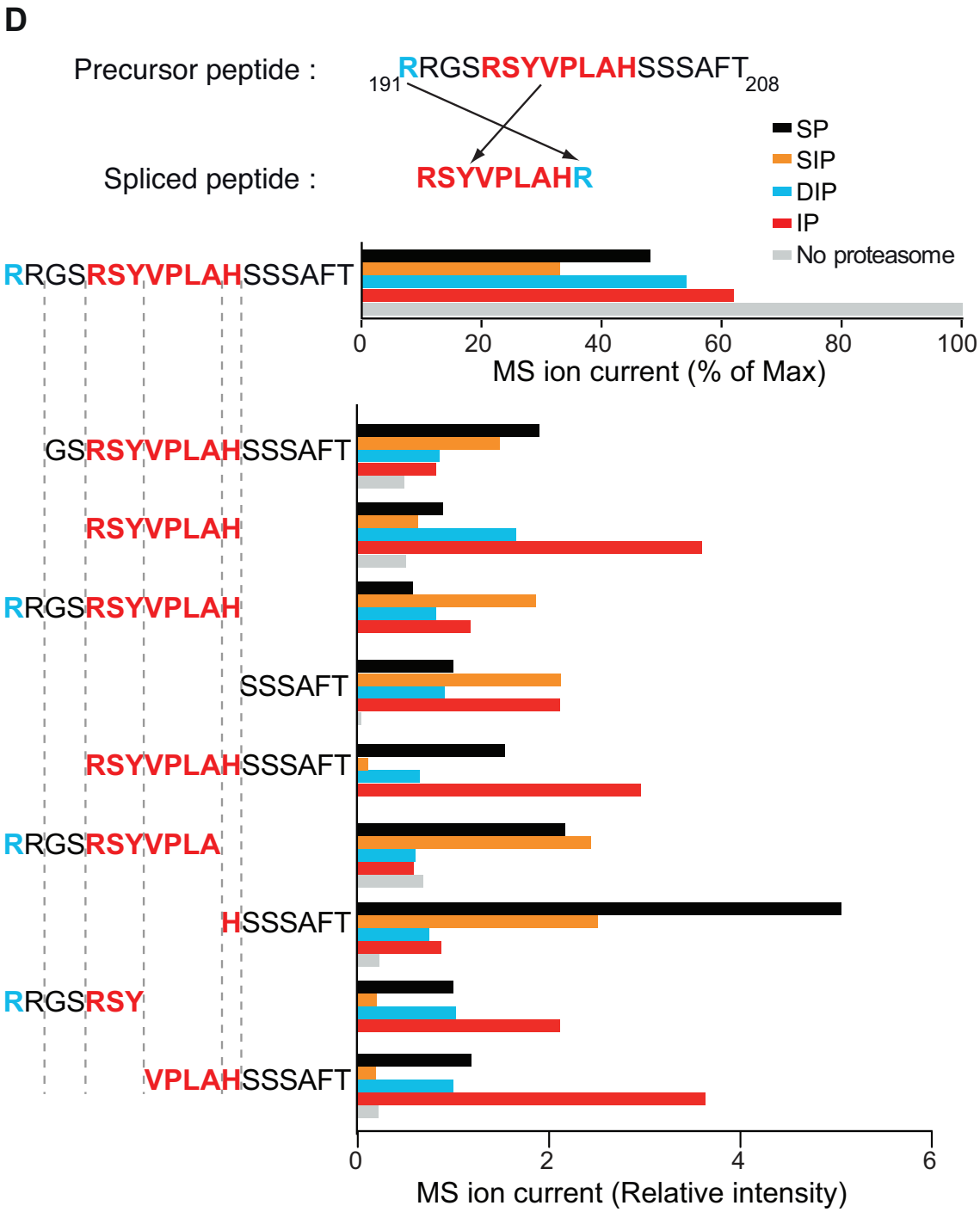

D
Precursor peptide :
191RRGSRSYVPLAHSSSAFT208
Spliced peptide :
RSYVPLAHR
No proteasome
MS ion current (% of Max)
RRGSRSYVPLAHSSSAFT
GSRSYVPLAHSSSAFT
RSYVPLAH
RRGSRSYVPLAH
SSSAFT
RSYVPLAHSSSAFT
RRGSRSYVPLA
HSSSAFT
RRGSRSY
VPLAHSSSAFT
MS ion current (Relative intensity)

Figure 5

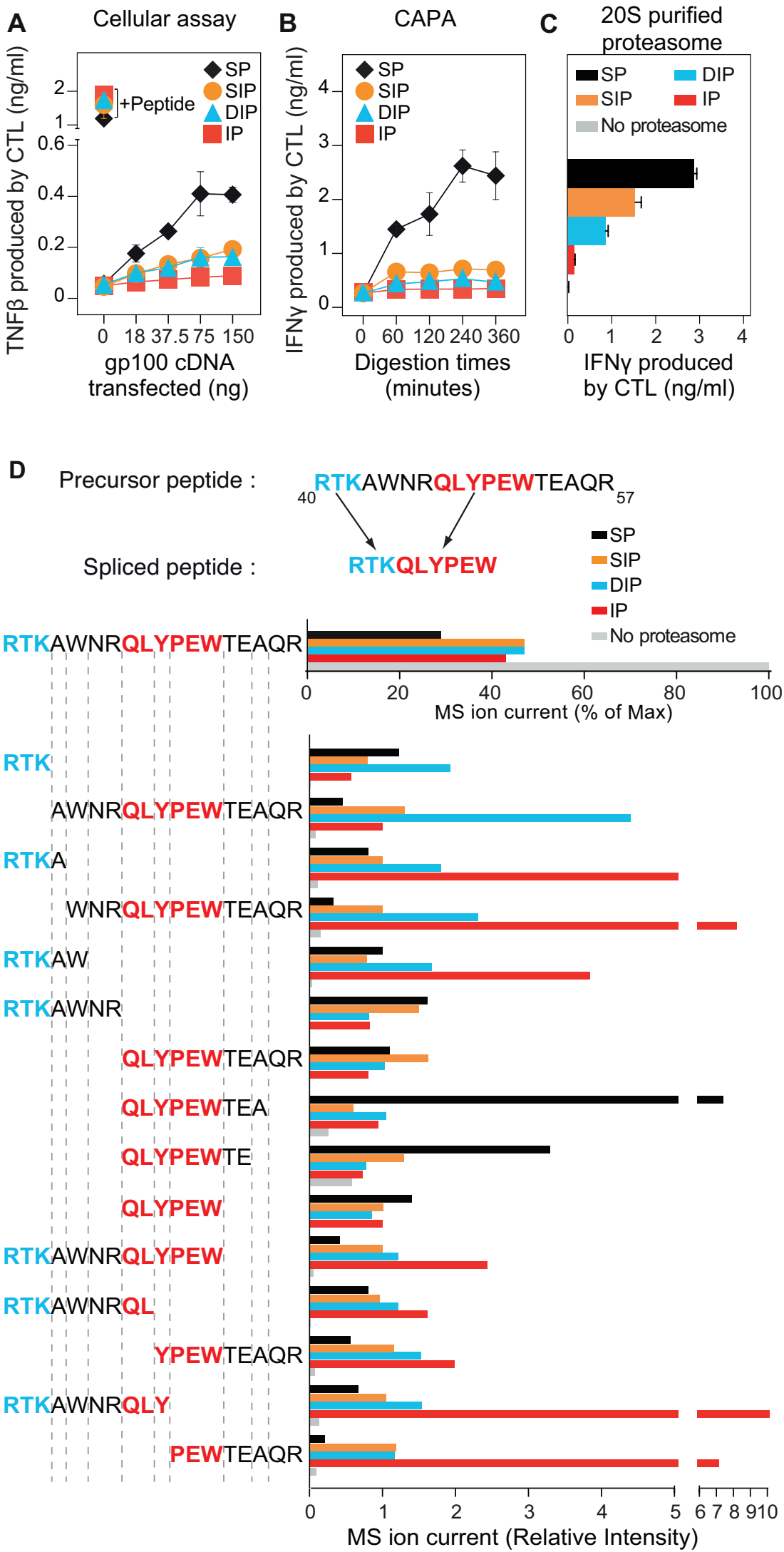
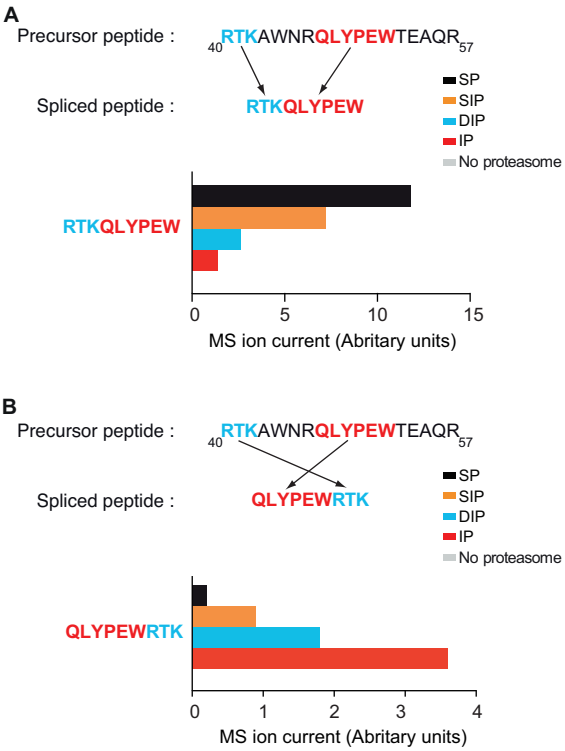

A
Cellular assay
TNFβ produced by CTL (ng/ml)
+Peptide
SP
SIP
DIP
IP
gp100 cDNA transfected (ng)
B
CAPA
IFNγ produced by CTL (ng/ml)
Digestion times (minutes)
C
20S purified proteasome
No proteasome
IFNγ produced by CTL (ng/ml)
D
Precursor peptide :
RTKAWNRQLYPEWTEAQR
Spliced peptide :
RTKQLYPEW
MS ion current (% of Max)
MS ion current (Relative Intensity)

Figure 6


A
Precursor peptide :
40 RTKAWNRQLYPEWTEAQR 57
Spliced peptide :
RTKQLYPEW
SP
SIP
DIP
IP
No proteasome
RTKQLYPEW
0
5
10
15
MS ion current (Abritary units)

B
Precursor peptide :
40 RTKAWNRQLYPEWTEAQR 57
Spliced peptide :
QLYPEWRTK
SP
SIP
DIP
IP
No proteasome
QLYPEWRTK
0
1
2
3
4
MS ion current (Abritary units)

Figure 7

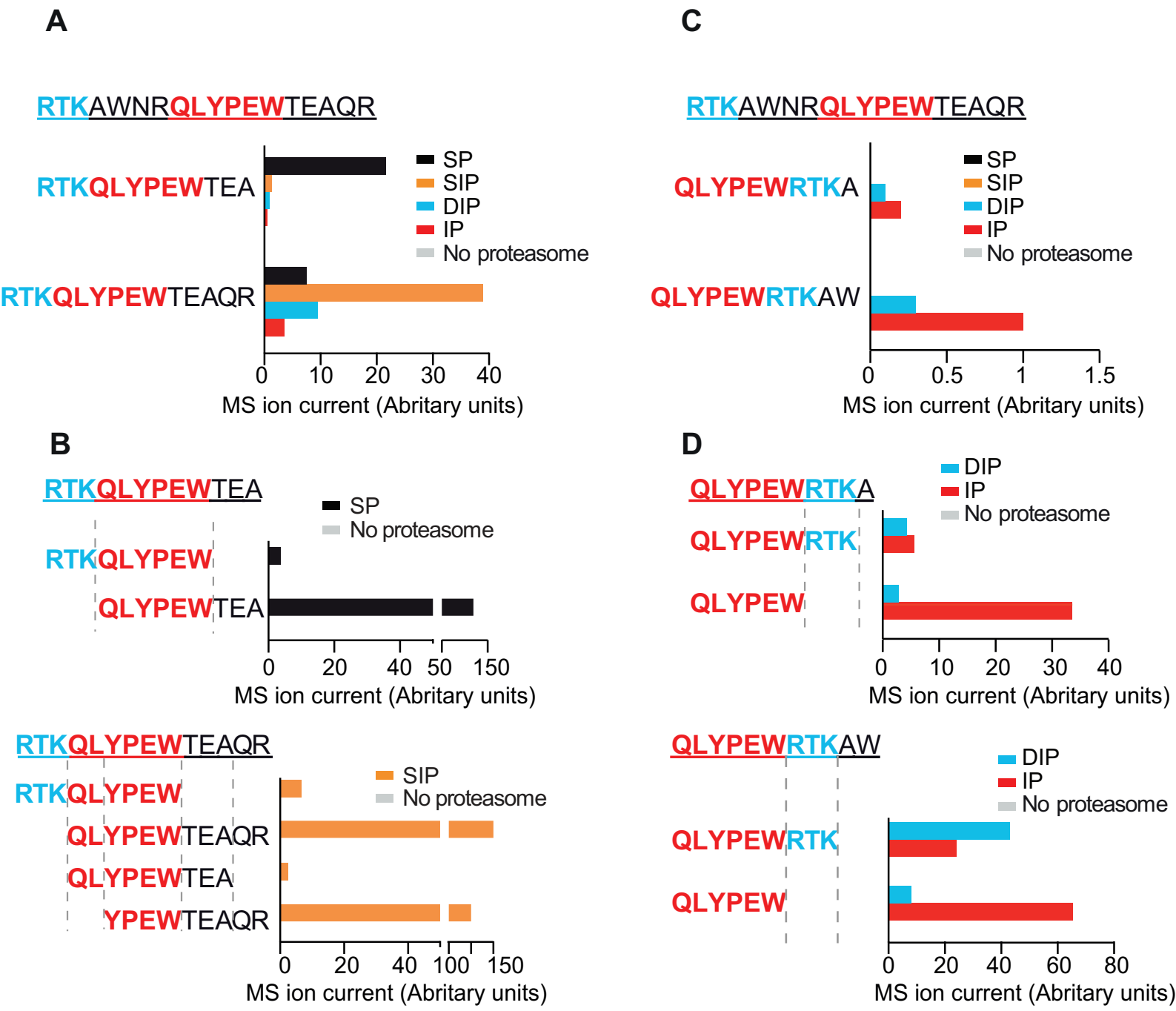
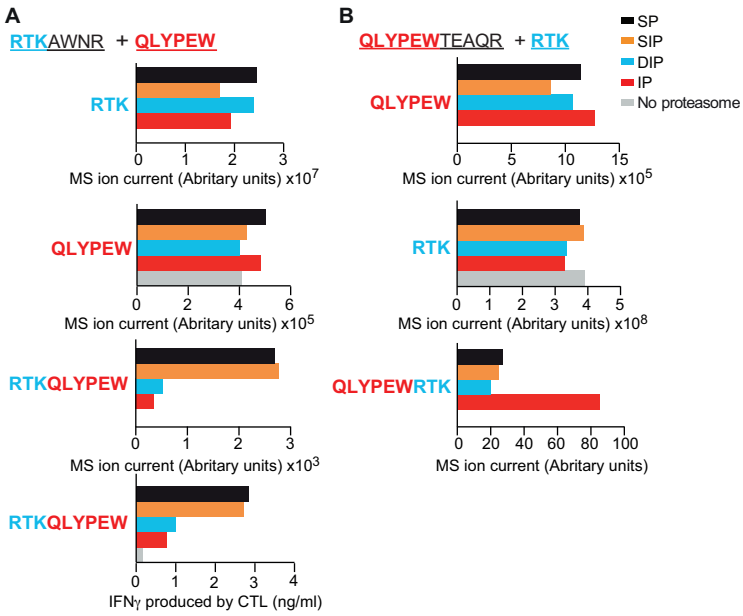

A
RTKAWNRQLYPEWTEAQR
RTKQLYPEWTEA
RTKQLYPEWTEAQR
SP
SIP
DIP
IP
No proteasome
0 10 20 30 40
MS ion current (Abritary units)
C
RTKAWNRQLYPEWTEAQR
QLYPEWRTKA
QLYPEWRTKAW
SP
SIP
DIP
IP
No proteasome
0 0.5 1 1.5
MS ion current (Abritary units)
B
RTKQLYPEWTEA
RTKQLYPEW
QLYPEWTEA
SP
No proteasome
0 20 40 50 150
MS ion current (Abritary units)
RTKQLYPEWTEAQR
RTKQLYPEW
QLYPEWTEAQR
QLYPEWTEA
YPEWTEAQR
SIP
No proteasome
0 20 40 100 150
MS ion current (Abritary units)
D
QLYPEWRTKA
QLYPEWRTK
QLYPEW
DIP
IP
No proteasome
0 10 20 30 40
MS ion current (Abritary units)
QLYPEWRTKAW
QLYPEWRTK
QLYPEW
DIP
IP
No proteasome
0 20 40 60 80
MS ion current (Abritary units)

Figure 8


A
RTKAWNR + QLYPEW
RTK
0 1 2 3
MS ion current (Abritary units) x10^7
QLYPEW
0 2 4 6
MS ion current (Abritary units) x10^5
RTKQLYPEW
0 1 2 3
MS ion current (Abritary units) x10^3
RTKQLYPEW
0 1 2 3 4
IFNγ produced by CTL (ng/ml)
B
QLYPEWTEAQR + RTK
SP
SIP
DIP
IP
No proteasome
QLYPEW
0 5 10 15
MS ion current (Abritary units) x10^5
RTK
0 1 2 3 4 5
MS ion current (Abritary units) x10^8
QLYPEWRTK
0 20 40 60 80 100
MS ion current (Abritary units)