Production of spliced peptides by the proteasome

Nathalie Vigneron[1,2], Vincent Stroobant[1,2], Violette Ferrari[1,2], Joanna Abi Habib[1,2], Benoit J. Van den Eynde[1,2,3]*

[1]Ludwig Institute for Cancer Research, Brussels, Belgium
[2]de Duve Institute, *Université catholique de Louvain*, Brussels, Belgium
[3]WELBIO (Walloon Excellence in Life Sciences and Biotechnology), Brussels, Belgium

Running title: Peptide splicing by the proteasome

*Author to whom correspondence should be addressed: Dr. Van den Eynde, Ludwig Institute for Cancer Research, de Duve Institute, *Université catholique de Louvain*, Avenue Hippocrate 75 B1.74.03, B-1200 Brussels, Belgium; Tel.: +32-2-7647580; Fax: +32-2-7647590; E-Mail: benoit.vandeneynde@bru.licr.org.

Keywords: proteasome, $CD8^+$ cytolytic T lymphocytes, peptide splicing, transpeptidation, antigenic peptides

Abstract

$CD8^+$ cytolytic T lymphocytes are essential players of anti-tumor immune responses. On tumors, they recognize peptides of about 8-to-10 amino acids that generally result from the degradation of cellular proteins by the proteasome. Until a decade ago, these peptides were thought to solely correspond to linear fragments of proteins that were liberated after the hydrolysis of the peptide bonds located at their extremities. However, several examples of peptides containing two fragments originally distant in the protein sequence challenged this concept and demonstrated that proteasome could also splice peptides together by creating a new peptide bond between two distant fragments. Unexpectedly, peptide splicing emerges as an essential way to increase the peptide repertoire diversity as these spliced peptides were shown to represent up to 25% of the peptides presented on a cell by MHC class I. Here, we review the different steps that led to the discovery of peptide splicing by the proteasome as well as the lightening offered by the recent progresses of mass spectrometry and bioinformatics in the analysis of the spliced peptide repertoire.

Introduction

The immune system continually monitors the emergence of abnormal cells, such as virally infected cells or cancer cells. This process, called immune surveillance, is partly achieved by $CD8^+$ cytolytic T lymphocytes (CTL), a specific subset of T cells, which recognize peptides of 8 to 10 amino acids that are presented at the cell surface by major histocompatibility complexes (MHC) class I molecules. Peptides bound to MHC class I generally arise from the degradation of cellular proteins, so that the peptide repertoire reflects the assortment of proteins expressed by a cell. In healthy cells, those peptides are derived from self-proteins and are therefore not recognized by the immune system. In contrast, tumor cells and cells infected by viruses can be recognized by circulating T cells, because they express peptides derived from viral, tumoral or "mutated" proteins, against which negative selection in the thymus did not take place.

Peptides recognized by CTLs are mainly produced by the proteasome, a large protease, which degrades the pool of cellular proteins [1] (Figure 1). Peptides resulting from proteasome degradation are then transported into the lumen of the endoplasmic reticulum (ER) by a dedicated transporter called transporter associated with antigen processing (TAP) [2]. Once in the ER, peptides can be further trimmed by ER-associated aminopeptidases (ERAP1 and ERAP2)[3-5] and loaded on MHC class I, with the help of the peptide loading complex, which is composed of TAP, Tapasin, ERp57 and calreticulin [6]. Stable MHC/peptide complexes then exit the ER and migrate through the secretory pathway to reach the cell surface, where MHC/peptide complexes are then displayed for CTL recognition (Figure 1). For many years, peptides recognized by CTLs were thought to correspond exclusively to linear fragments of proteins. This notion was confronted to the identification of CTL-recognized peptides that were composed of two fragments originally distant in the parental protein (Figure 2A) [7-9]. Splicing of these peptide fragments involves the creation of a new peptide bond between the two separated fragments, a process that was shown to take place in the proteasome in the course of

protein degradation [8, 9]. Here, we review the recent advances unraveling the importance of spliced peptides in the MHC class I repertoire.

The production of antigenic peptides by the proteasome

The proteasome is a large protease complex, which is responsible for the bulk of protein degradation in cells. The proteasome core, called 20S proteasome, is composed of four stacked heptameric rings that delimit a catalytic chamber inside which proteins are degraded. The two outer rings are made of α-subunits (α_{1-7}), while the two inner rings are made of β-subunits (β_{1-7}), three of which (β1, β2 and β5) initiate the hydrolysis of peptide bonds inside proteins [10-12]. Entrance into the catalytic chamber of proteasome is drastically regulated to avoid detrimental protein degradation. The first level of regulation resides in the α-rings, which form a very narrow pore that is further obstructed by the N-terminal tails of the α-subunits. As a consequence, free 20S proteasome has a low activity, at least toward folded protein substrates [13]. Exceptions to this rule are oxidant-damaged proteins that appear to be mainly degraded by the 20S proteasome [14, 15]. Accessory proteins, such as Hsp90 [16] or arginine-rich histones [17] might facilitate access of these proteins into the catalytic chamber. Hydrophobic patches, which are often exposed following protein oxidation, could also facilitate the entry of oxidized proteins into the catalytic chamber [14], in line with the observation that hydrophobic peptides promote the opening of the α-ring of the 20S proteasome [18].

Opening of the α-ring is orchestrated by the proteasome regulatory particles, whose HbYX motif binds to pockets located between the α subunits, promoting the rearrangement of the N-terminal tails and opening of the gate [19]. Four types of regulators exist: the PA700 (19S regulator), PA28αβ (11S regulator αβ), PA28γ (11S regulator γ) and PA200 (reviewed in [20]). Another particle PI31 was shown to bind the proteasome [21], but its effective role in proteasome function is still a matter of debate [22, 23]. The PA700 regulator associate to one or both sides of the 20S proteasome to form the 26S proteasome [21]. PA700 enables the

proteasome to bind, deubiquitylate, unfold and translocate proteins tagged with polyubiquitin chains inside the catalytic chamber. Interestingly, the nature of the regulator associated to the 20S proteasome can influence the proteasome ability to release some specific antigenic peptides. Thus, binding of PA28αβ to the proteasome can either favour or impair the production of some specific antigenic peptides [24-26].

Peptide bond hydrolysis is instigated by the N-terminal threonine of the catalytic subunits [10, 11, 27] (Figure 2B). More precisely, the hydroxyl group of the N-terminal threonine produces a nucleophilic attack on the carbonyl group of the peptide bond, leading to the formation of an acyl-enzyme intermediate where the N-terminal peptide fragment remains attached to the proteasome by an ester link. A water molecule located in the catalytic chamber then hydrolyzes the acyl-enzyme intermediate, and the peptide fragment released then diffuses in the cytosol.

Identification of spliced peptides

Anti-tumor T cells recognizing spliced peptides

A few years ago, Hanada *et al.* made the surprising observation that a CTL clone recognizing a renal cell carcinoma was composed of two peptide fragments originally distant in the FGF-5 protein [7] (Figure 2A and Figure 3). Production of this peptide required the excision of a 40-amino acid sequence and the splicing of the peptide fragments located at each side. A few months later, using an anti-melanoma CTL clone, we identified a peptide that was derived from the melanoma differentiation antigen gp100 and was also composed of two non-contiguous fragments of the protein [8]. In this case, the fragments to be spliced were separated by only four amino acids (Figure 3). The peptide was shown to be present at the surface of tumor cells. The treatment of tumor cells with proteasome inhibitors blocked presentation of the peptide, suggesting that the proteasome could be involved in the peptide splicing reaction [8]. To confirm this hypothesis, we incubated purified 20S proteasomes with unspliced peptide precursor RTKAWNRQLYPEW. This led to the production of spliced peptide

RTK_QLYPEW, which could be detected in the digest by mass spectrometry or with the CTL when the digests were pulsed on antigen-presenting cells. We showed that the splicing reaction occured by transpeptidation in the proteasome according to the following events (Figure 2C): (1) in the course of protein degradation the peptide bond between fragment RTK and AWNR is attacked by the hydroxyl group of a catalytic subunit; (2) this leads to the production of an acyl-enzyme intermediate where fragment RTK remains attached, through an ester link, to the catalytic site of proteasome; (3) this acyl-enzyme intermediate is then subjected to a nucleophilic attack by the N-terminal group of the fragment QLYPEW, which was released in the proteasome chamber following cleavage of the peptide bond between (RTK)AWNR and QLYPEW; (4) this leads to the production of a spliced peptide composed by the assembly of fragments RTK and QLYPEW (RTK_QLYPEW). We later confirmed that the first spliced peptide described by Hanada *et al.*, which was derived from FGF-5, was also spliced by the proteasome through a reaction of transpeptidation, despite the fact that in this case the fragments were separated by a much longer intervening sequence (40 amino acids) [28]. To determine whether the length of the intervening sequence could affect the efficiency of the splicing reaction, we transfected constructs in which the peptide fragments to be spliced were separated by intervening sequences of different lengths [28]. We found that the shorter was the intervening sequence, the more efficient was the peptide splicing process. In addition, to test whether peptide splicing could occur between peptide fragments located on distinct proteins, we transfected COS-7 cells with pairs of constructs that were mutated in such a way that the only way to produce the relevant antigenic peptide was to splice fragments located on two distinct proteins [28]. Because transfected cells were hardly recognized by the CTL despite the fact that constructs were overexpressed in this system, we concluded that trans-splicing (i.e. the splicing of peptide fragments originating from different proteins) could not occur in physiological conditions.

A few years later, our group studied a CTL isolated from a cancer patient who had received an allogeneic hematopoietic cell transplantation from an MHC-matched donor [9]. This CTL recognized a minor histocompatibility antigen in the SP110 protein, which was composed of two peptide fragments that were not contiguous in the SP110 protein (Figure 3). Interestingly, in this case, the fragments were spliced to each other in the reverse order to that in which they occur in the SP110 protein. Again, splicing of this peptide was shown to take place in the proteasome by transpeptidation.

Searching for the peptide recognized by tumor infiltrating lymphocytes (TIL) of a melanoma patient, we then identified a fourth example of spliced peptide. This peptide contained two fragments of the tyrosinase protein spliced in the reverse order [29]. Interestingly, this peptide contained two additional post-translational modifications, resulting from the conversion of two asparagine residues into aspartates by deamidation (Figure 3). The deamidation was shown to occur upon deglycosylation of the N-glycosylated tyrosinase in the cytosol by peptide N-glycanase (PNGase): treatment of melanoma cells by the PNGase inhibitor z-VAD-fmk blocked the production of the antigenic peptide. Following deamidation, the aspartate-containing protein was degraded by the proteasome, which then spliced together the aspartate-containing fragments (Figure 4).

More recently, we identified a fifth spliced peptide, which was derived from melanoma differentiation protein gp100 and recognized by a CTL clone on melanoma cells (Figure 3). In contrast to previous examples of spliced peptides, this peptide comprised an 8 amino acid-long peptide spliced to a single amino acid, an arginine residue that was added at the C-terminus [30]. To test whether a single amino acid could be spliced as such or whether it had to be included into longer peptide fragments, we incubated purified proteasomes with pairs of fragments of different lengths. This led to the observation that the fragment performing the nucleophilic attack on the acyl-enzyme intermediate (i.e. the C-terminal spliced reactant) had

to contain at least 3 amino acids in order for the splicing reaction to occur. In the case of this gp100 spliced peptide, the splicing reaction therefore produces a spliced peptide that is elongated at the C-terminus and is further cleaved, most likely by the proteasome, to produce the final antigenic peptide. This mechanism is likely responsible for the production of spliced peptides comprising C-terminal fragments shorter than three amino acids. It is expected that the size of the N-terminal spliced reactant might also be restrictive. However, N-terminal trimming of peptides is an efficient process that takes place in the ER and is orchestrated by ERAP1/2 (Figure 1) [3-5]. Therefore, spliced peptides with a short N-terminal fragment likely exist and might result from the production of N-terminally extended peptides that are further trimmed in the ER by aminopeptidases.

Help of computational approaches for the identification of spliced peptides

In principle, peptide splicing could involve any fragment produced by proteasome degradation of a given protein, and should therefore produce a highly diverse repertoire of spliced peptides. However, thus far, the identification of spliced peptides was limited by the laborious isolation of T cell clones from the blood or tumor of cancer patients. In order to facilitate the identification of spliced peptides, Liepe *et al.*, developed *Splicemet*, an algorithm enabling the identification of spliced peptides by mass spectrometry in proteasome digests performed with a given precursor peptide [31]. *Splicemet* computes the m/z values of any spliced and non-spliced peptide theoretically generated from the precursor peptide, and these m/z values are then compared to the m/z values of the ions detected by MS in the digest. The presence of a specific peptide is then verified by MS/MS. This led to the identification of peptide QLYPEW_RTK, which is derived from the gp100 protein and composed of fragments that are identical to those contained in the previously described peptide RTK_QLYPEW, but assembled in the reverse order [32]. A CTL clone recognizing peptide QLYPEW_RTK on HLA-A3 was isolated and shown to recognize tumor cells expressing both gp100 and HLA-A3,

demonstrating that this peptide was indeed expressed at the surface of tumor cells [32]. Because the authors observed that peptide QLYPEW_RTK could be produced *in vitro* by incubating proteasomes with peptides RTK and QLYPEW, they suggested that this peptide could be produced by proteasome-catalyzed condensation rather than transpeptidation: this is because fragment QLYPEW has already its final size and cannot produce an acyl-enzyme intermediate, which is necessary for the transpeptidation reaction. Thus far, however, the experimental evidence supporting splicing by condensation is limited to proteasome digests and to the use of cells overexpressing appropriately truncated constructs, but is absent for cells expressing full-length proteins. Moreover, the prominent role of transpeptidation over condensation in the splicing reaction was confirmed by mass spectrometry analysis of peptide digests performed in the presence of deuterium-labeled water [33].

Recent progresses have greatly improved the sensitivity of mass spectrometry and facilitated the identification of peptides that can be eluted from MHC class I molecules expressed by tumor cells. When coupled to RNAseq analysis, this even enables the identification patient-specific neoantigens. However, when it comes to spliced peptides, a major limitation lies in the fact that the peptide databases that are used to match the m/z values and identify the peptides only contain linear fragments of the proteome.

As a result, the abundance and diversity of spliced peptides among the HLA-bound peptidome could not be assessed, and it has remained unclear whether spliced peptides represented infrequent and peculiar cases, or whether they contributed for a significant fraction of the HLA peptidome. Solving this issue would require the creation of dedicated databases containing all possible spliced peptides. Such databases would be enormous in size and their management beyond the capacity of current computing power.

Recently, however, Liepe *et al.* overcame this bioinformatical obstacle and created such a custom peptide database [34]. They limited the size of the database by computing only spliced

peptides predicted from the proteins encoded by genes expressed in the cell lines they studied. They only considered peptide spliced by cis-splicing (i.e. produced by splicing fragments originating from the same protein), they imposed a minimal size of 3 amino acids for the C-terminal spliced reactant and limited the length of the intervening sequence to 25 amino acids - a likely underestimation since previous work has described splicing of fragments located 40 amino acids apart [7]. Surprisingly, using such databases, they observed that spliced peptides represented about 25% of the diversity of the HLA peptidome in the different lymphoid and fibroblast cell lines they studied. In terms of abundance, spliced peptides even accounted for about 30% of the HLA peptidome in those cell lines [34].

Moreover, when analyzing the sequences of the eluted peptides, the authors found that some proteins were only represented by one or several spliced peptides in the HLA ligandome but not by any linear peptide [34]. Although this can result from the limited sensitivity of the MS approach, this finding suggests that such proteins could lack linear peptides with adequate HLA-binding motifs or produce peptides that are too small to bind to HLA class I molecules. In that case, peptide splicing might be a way to produce peptides with adequate sequence and length to bind to HLA class I. Altogether, this pioneering work established the importance of spliced antigenic peptides in the peptide repertoire presented by MHC class I molecules, and led to the unexpected conclusion that they represent a significant fraction of the HLA peptidome. It will be interesting to see whether such findings can be confirmed in other cell lines and by other groups.

Spliced peptides and infectious diseases

A similar computerized approach was recently used to identify spliced peptides derived from the phosphatidylcholine-prefering phospholipase C (plcB) protein of *Listeria monocytogenes* [35]. In this case, the authors computed any potential spliced peptide derived from this protein and applied several additional steps to restrain the size of the peptide database produced: they

limited the size of the intervening sequence to 40 amino acids, and only considered peptides produced by cis-splicing. From the list of peptides obtained, they selected those that displayed the highest affinity for H2-K^b and then tested their recognition by T cells isolated from mice infected with *Listeria*. Two of the spliced peptides tested were recognized. These peptides were produced by the proteasome and were amongst those that displayed the highest affinity for H2-K^b.

Peptide splicing and proteasome subtypes

Standard proteasome and immunoproteasome

In immune cells or cells exposed to an inflammatory environment, the catalytic subunits of the proteasome are replaced by three alternative subunits, LMP2 (β1i), MECL-1 (β2i) and LMP7 (β5i) whose expression is induced by interferon-γ (reviewed in [36]). Because these subunits are preferentially incorporated into proteasome in place of their standard counterparts β1, β2 and β5, they form a new subtype of proteasome called immunoproteasome. Based on the study of proteasome digests of proteins, peptides and fluorogenic peptides, three main proteasome catalytic activity were described: the chymotrypsin-like activity, the trypsin-like activity and the caspase-like activity, which respectively cleave after hydrophobic, basic or acidic amino acids [37]. Other proteasome activities such as the branched chain amino acid preferring peptidase (BRAAP) or the small neutral amino acid preferring (SNAAP) activities were also described [38]. Mutational studies and analysis of proteasome crystal structure have assigned each of these activities to one or the other proteasome subunit: the caspase-like activity is linked to the β1 subunit, while the trypsin-like and chymotrypsin-like activities are respectively linked the β2 and the β5 subunits in the standard proteasome [37]. Replacement of standard proteasome subunits by immunoproteasome subunits increases cleavages after basic and hydrophobic residues while drastically decreasing cleavage after acidic residues [39-41]. The BRAAP activity can be mediated by either β1, β5 or β1i and the SNAAP acivity by β5 [42].

These assignments are however not absolute because the nature of the amino acids surrounding the cleavage sites also influence cleavage efficacy. It is indeed the nature of the amino acids that line the catalytic pockets, as well as the size and structure of these pockets that will determine which sequences will be cleaved. In line with this, subtle conformational differences in the substrate-binding pockets of the catalytic subunits of standard and immunoproteasome have been observed [43]: the S1 substrate-binding pocket (which interacts with the amino acid residue forming the C-terminal end of the cleaved peptide) of β1i is smaller, more hydrophobic, and lacks a positively charged residue (R45) when compared to that of β1, probably explaining the reduced caspase-like activity of β1i and its higher cleavage activity after small hydrophobic and branched residues [43]. On the other hand, the S1 pocket of β5i is more spacious than that of β5, favouring cleavage after large non-polar residues such as tyrosine, tryptophan or phenylalanine [43]. Because of its increased propensity to cleave after basic and hydrophobic amino acids, the immunoproteasome was predicted to be more efficient at producing HLA-class I-binding peptides. This idea was supported by the study of mice knockout for all three immunosubunits in which presentation of a number of MHC class I epitopes was decreased [44]. This is not always the case, however, as some antigens were shown to be better produced by the standard proteasome than the immunoproteasome [45, 46]. Overall, the capacity of one proteasome type to produce an antigenic peptide will depend on its capacity to do the cleavages that release the peptide and to do internal cleavages that destroy the peptide [47]. Several years ago, we assessed the ability of the standard proteasome and the immunoproteasome to produce spliced peptides [48]. The peptides derived from gp100 (RTK_QLYPEW), FGF-5 and tyrosinase were all better produced by the standard proteasome. In contrast, the SP110 peptide was better produced by the immunoproteasome. When looking at the cleavages exerted by the two proteasomes in those precursor peptides, we confirmed the prediction that was made by the transpeptidation model of the splicing reaction: the capacity of one proteasome type to produce

a spliced peptide depends on its ability produce the adequate acyl-enzyme intermediate and the C-terminal splice reactant. This observation suggested that the rules that govern peptide splicing are dependent on the rules that determine cleavage by the proteasome.

Intermediate proteasomes β5i and β1iβ5i

In addition to the standard proteasome and the immunoproteasome, two other proteasome types were described that contain one (β1 β2 β5i) or two (β1i β2 β5i) of the three immunoproteasome subunits [49]. These intermediate proteasomes are abundant in normal tissues, where they represent 30 to 50% of the proteasome found in liver, kidney, small bowel, colon and dendritic cells. In tumors, they represent around 10-20% of the total proteasome. The physiological role of intermediate proteasomes has not yet been elucidated. Because of their particular proteasome subunit composition, intermediate proteasomes show cleavage specificities that are different from that of standard or immunoproteasomes [49]. As a consequence, they produce their own peptide repertoire, which contain peptides that are unique to these proteasome subtypes [46, 49].

Thymoproteasome

Positive selection takes place in the thymic cortex and is mediated by cortical epithelial cells (cTEC). cTEC express a specific subtype of proteasome, called thymoproteasome, which contains catalytic subunit β5t, in place of β5-β5i [50]. Subunit β5t is associated to β1i and β2i in the thymoproteasome, which displays a low chymotrypsin-like activity and appears to produce a distinct peptide repertoire as compared to the immunoproteasome [51]. The importance of the thymoproteasome in the development of the $CD8^+$ T cells has been highlighted by the fact that β5t-knocked-out mice display an altered CTL repertoire and fail to respond efficiently to viral infections [50, 51]. However, the exact function of the thymoproteasome in the process of positive selection is not yet clear. Whether the thymoproteasome is able to produce spliced peptides still needs to be demonstrated, but the

inherent way by which spliced peptides are produced implies that the thymoproteasome should also produce spliced peptides, as long as the appropriate acyl-enzyme intermediate(s) and C-terminal reactant(s) can be formed.

Peptide splicing: are there any rules?

As mentioned above, the splicing mechanism implies that the rules that govern splicing are not different from the rules that determine proteasome cleavage. This is because the first steps of the splicing reaction are in fact producing the splice reactants. In line with this, Mishto *et al.* observed, by in depth analysis of proteasome digests, that one of the rate-limiting factors of the peptide splicing reaction is the amount of the spliced reactants involved, and in particular the amount of the less abundant peptide fragment, regardless of its position in the final antigenic peptide [33]. In contrast with this finding, the same authors analysing the same digests found no correlation between the strength of proteasome cleavage at one position, and the frequency at which the corresponding amino acid was used to produce either the N- or the C-terminal residue of the spliced reactants (i.e. P1' and P1 respectively) (Figure 5). Rather, the splicing reaction appeared to often involve fragments released from minor cleavage sites [33]. However, the meaning of this observation is unclear, and it should be stressed that it was based on MS quantification of different peptide fragments, which is imprecise due to variation in the ionization of the different fragments. In addition, the abundance of a splice reactant depends not only on the strength of the cleavage releasing it, but also on the strength of internal cleavages destroying it. Indeed, multiple cleavages are often observed in these *in vitro* digests, in which peptides can undergo successive cycles of degradation. Yet, in theory it is not impossible that splicing could be favoured by certain parameters, such as the sequence of the spliced reactant and/or its ability to be positioned adequately in the catalytic site so as to complete the splicing reaction. Using short peptide libraries, Berkers *et al.* tried to identify amino acid sequences that

could promote splicing by the proteasome [52]. Their results indicated that the peptide splicing reaction was driven by the sequence of the N-terminal reactant and by the concentration rather than the sequence of the C-terminal reactant. It has to be noted however, that it remains difficult to determine whether the effect of interchanging amino acids in the N-terminal spliced reactant really affects the splicing reaction itself or only the affinity of the N-terminal reactant for the non-primed binding site of the catalytic subunit and thereby its ability to form the acyl-enzyme intermediate. This is especially important as proteasomes contain three different types of catalytic subunits: the sequence requirements for splicing are likely to differ between the three catalytic subunits of the proteasome as much as they do for cleavage. Hence, changing one amino acid in the N-terminal splice reactant may redirect the fragment to another catalytic subunit. In addition, the C-terminal reactant itself might not access the primed binding site of the different catalytic subunits with the same efficiency.

As expected, Liepe *et al.* observed significant differences in the nature of the amino acids found at positions P_N, P_1, $P_{1'}$, and P_C (P_N and P_C being the N-terminal and the C-terminal residues of the eluted spliced peptides, respectively) (Figure 5) of the spliced peptides when these were eluted from distinct cell lines [34]. This is likely because the nature of the amino acid found at these positions is also influenced by the HLA molecule to which these peptides bind [34], a bias that clearly needs to be taken into account when drawing conclusions about the sequence requirement for peptide splicing. Altogether, so far, the evidence supporting the existence of rules determining the splicing reaction is weak, and we are left with the assumption that splicing is mostly driven by the efficiency of the cleavages producing the splice reactants.

Spliced isopeptides

Recently, Berkers et *al.* showed that the proteasome could also use the ε-amino group of the N-terminal lysine to perform a transpeptidation reaction [53]. This leads to the creation of a spliced

peptide containing an isopeptide bond (hereafter called spliced isopeptide). Surprisingly, splicing using the ε-amino group of the N-terminal lysine is only 10 times less efficient than splicing using the α-amino group. However, the relative efficiency of ε-amino versus α-amino splicing might have been overestimated in these experiments, because the authors also observed that isopeptides were much more resistant to proteasome degradation than classical peptides [53]. Hence, in the digests analysed, secondary degradation of spliced peptides was favoured over that of spliced isopeptides and likely increased the ratio of spliced isopeptides over spliced peptides. Spliced isopeptides can efficiently bind HLA class I molecules, as long as the isopeptide bond is located at in the middle of the peptide or at its N-terminus. However, the presence of spliced isopeptides at the cell surface remains to be demonstrated in order to establish their physiological relevance.

Splicing of MHC class II-restricted peptides

Using a peptide library and proteomic analysis, Delong *et al.* showed that, in non obese diabetic mice (NOD), diabetogenic $CD4^+$ T cells clones recognized hybrid peptides produced by splicing of fragments originated from two different proteins found in abundance in the secretory granules of β cells [54, 55]. One of these peptides was composed of an Chromogranin A-derived peptide at the C-terminal side and insulin II C-peptide fragment at the N-terminal side, while the other was produced by the assembly of fragments derived from the islet amyloid polypeptide at the C-terminal side and insulin II C-peptides at the N-terminal side. The presence of these peptides in β cells extracts was confirmed by mass spectrometry. The enzyme responsible for the splicing of these peptides was not identified. Because these peptides are presented by MHC class II molecules, their processing does not take place in the cytosol but in the secretory and/or endocytic compartment. Therefore, they are not produced by the proteasome but by a peptidase present in this compartment. It is likely that the reversed proteolytic transpeptidation is favoured in this case by the high concentration of the peptides

inside secretory granules. In contrast to proteasome-catalyzed peptide splicing, splicing of such abundant secretory peptides may therefore not require enzyme-procured compartmentalization. This probably explains why production of antigenic peptides by trans-splicing is physiologically possible in this compartment. In humans, $CD4^+$ T cells isolated from the islets of type I diabetic patients were also shown to recognize such hybrid peptides [54, 56].

Immunological relevance of peptide splicing

Because spliced peptides are byproducts of the proteasome activity, any type of proteasome should be able to produce spliced peptides. Indeed, both the standard and the immunoproteasome can splice peptides [48], and so does the yeast proteasome [33]. As a result, not only tumor cells but also healthy cells produce spliced peptides, as demonstrated in normal melanocytes [29]. The ability of normal cells to splice peptides should be kept in mind when considering the immunogenicity of spliced peptides. Because spliced peptides are not encoded in the self-genome, they are sometimes considered as non-self antigens with high immunogenicity, similar to neoantigens caused by somatic mutations on cancer cells. Yet, similarly to what occurs for other non-mutated tumor antigens, the immunogenicity of spliced peptides will depend on whether the antigenic peptide is expressed in the thymic cells driving negative selection and/or by dendritic cells in the periphery, which can either tolerize or activate T cells against particular antigenic peptides, depending on the context in which the antigen presentation takes place.

Central tolerance to spliced peptides

Central tolerance to self antigens is established in the thymus during T cell development, through the sequential steps of positive selection, which occurs in the thymic cortex and retains only T cells whose T cell receptor (TCR) can interact with self MHC, and negative selection, which eliminates T cells that recognize self peptides presented in the thymic medulla [57]. This tightly regulated selection releases a highly diverse T cell population, which is purged of most

self-reactive T cells and can be primed upon infection or cancer initiation. Although the exact subtypes of proteasome present in epithelial and dendritic cells of the thymic medulla have not been defined, it is likely that those cells can also produce spliced peptides. Therefore, provided that the parental protein is expressed in these cells, spliced peptides derived from self-proteins will be produced, and T cells recognizing those will be eliminated. Hence, spliced peptides are not intrinsically more immunogenic than non-spliced peptides, as central tolerance applies equally, and the majority of spliced peptides displayed at the cell surface are immunologically silent. However, as for non-spliced peptides, it is likely that some specific spliced peptides escape this process of tolerance, either because the parental protein is not expressed at a sufficiently high level in thymic cells, or because the proteasome subtype present in these cells does not allow the production of these peptides. This could explain how T cells directed against spliced self-peptides were found in cancer patients.

Presentation of spliced peptides by dendritic cells

Dendritic cells (DC), which are the most efficient antigen-presenting cells of the immune system, are key players in the regulation of T cell immunity. DC have the unique ability to engulf dead cells carrying antigens and to present peptides derived from these antigens on their MHC class II (conventional presentation) or MHC class I molecules (cross-presentation) [58]. Early evidences have shown that cross-presentation is essential for the cross-priming of naïve $CD8^+$ T cells against antigens derived from tumors or from viruses that cannot infect DC [59, 60]. Productive activation of T cells will take place when cross-presentation is accompanied by the expression of co-stimulatory ligands on the DC and the release of pro-inflammatory cytokines [61, 62]. Lacking either of these signals leads to cross-tolerization, which is essential to the maintenance of peripheral tolerance to tissue-specific antigens [63, 64]. In addition, cross-presentation also appears to participate to negative selection of autoreactive T cells by bone-marrow-derived DC from the thymic medulla [65]. The intracellular processes involved

in cross-presentation are not yet clear but currently involve two main pathways: the vacuolar pathway and the cytosolic pathway. In the vacuolar pathway, exogenous antigens are processed inside the endosomal compartment, where peptides are directly loaded on endosomal MHC class I molecules. The cytosolic pathway involves the transfer of antigens from the lumen of the phagocytic compartment to the cytosol, where antigens are then degraded by proteasome to produce antigenic peptides, which can then be further transported by TAP into the ER lumen to be loaded onto MHC class I molecules. Cross-presentation of spliced peptide has not yet been demonstrated, but the fact that CTL can be expanded from the blood of cancer patients simply by incubating peripheral lymphocytes with irradiated tumor cells suggests that a priming step might have occurred *in vivo* prior to the *in vitro* culture, implying that cross-presentation of spliced antigen might take place *in vivo*. However, further experiments will be needed to test whether spliced peptides can indeed be efficiently cross-presented by DC. Because spliced peptides are specifically produced by the proteasome, one would expect that, if it occurs, cross-presentation takes place through the cytosolic pathway. Indeed, it is likely that the concentration of the antigen in the endosomal compartments would not be high enough to enable the production of spliced peptide by other proteases in these compartments. In that context, defining the type of proteasome able to produce a spliced peptide is essential, in order to better define the vaccination strategies approach to use in the frame of cancer immunotherapy. For example, immunization strategies targeting spliced tumor antigens that are better produced by the standard proteasome should rely on short peptides or peptide-encoding minigene constructs, which bypass intracellular processing: dendritic cells mostly contain immunoproteasome and antigens processed by the standard proteasome will not be efficiently processed in these cells when using full-length constructs [66]. This assumption is not specific to spliced peptides, but rather relates to vaccination involving any peptide that is more efficiently produced by the standard proteasome.

Concluding remarks

The recent and dramatic progresses in sequencing, mass spectrometry and bioinformatics have opened new perspectives for the identification of peptides recognized by T cells on tumors. These technologies will help identifying peptides derived from mutated antigens but also tumor-specific proteins, which could be of great help to render vaccine-based therapies accessible to large-scale populations of patients sharing similar HLA class I molecules. One should keep in mind that identifying peptides by mass spectrometry is not sufficient to validate these peptides for use in immunotherapeutic approaches. It remains essential to isolate T cell clones recognizing the peptides of interest in order to validate their immunogenicity and to further characterize their processing and presentation by tumor cells. From the recent work of Liepe *et al*, it appears that some proteins are solely represented by spliced peptides, which suggests that such peptides might be the only possibility to target some proteins [34]. This might be inherent to the type of amino acid sequence composing these proteins, which might not be appropriate to produce peptides binding the MHC class I molecules involved: either because the proteins cannot be cleaved by the proteasome found in those cells or because peptides with the adequate HLA class I binding motifs are simply absent from the protein sequence. In such scenario, splicing together small peptide sequences distant in the protein might be the only way to produce adequate HLA-binding peptides, because of the increased combinatorial diversity afforded by spliced peptides.

Although some rules might apply to the peptide splicing reaction, we believe that spliced peptides are simply byproducts of the main proteasome activity, which is to maintain cellular protein homeostasis and allow amino-acid recycling by destroying cellular proteins. The immune system takes advantage of this proteasome function to display a repertoire of peptides

at the cell surface, allowing T cells to monitor the integrity of the proteome. As discussed above, spliced peptides are unlikely to be more immunogenic than classical peptides, nor is there any *a priori* reason to think that the immune system would favor spliced over classical peptides. However, the important implication of the existence of spliced peptides is a significant increase in the size of the peptide repertoire presented by MHC class I molecules. Taking advantage of the combinatorial diversity of peptide fragments, splicing might be key to ensure that any protein gives rise to at least one peptide binding to any MHC specificity. These considerations may shed some light on the apparent paradox between the intrinsically low efficiency of the peptide splicing reaction - 1 to 2% of the fragments in a peptide/proteasome digest represent spliced peptides [33] - and the overrepresentation of spliced peptides in the surface peptide repertoire (up to 25%) [34]. Further work will be required to confirm this high proportion of spliced peptides at the cell surface and determine whether it results only from the high combinatorial diversity that makes spliced peptides more likely to bind to any MHC specificity, or whether other factors favor the processing of spliced peptides.

Acknowledgments

We are grateful to Mrs Auriane Sibille for her precious help in the preparation of this manuscript. VF is supported by a fellowship from the *Fonds National de la Recherche Scientifique* (FRIA grant No. 1.E091.14). JAH is supported by a fellowship from the *Fonds National de la Recherche Scientifique* (TELEVIE grant No. 7455115F).

Legends.

Figure 1. The MHC class I processing pathway. Peptides recognized by CTL generally derive from the degradation of cellular proteins by the proteasome. Further processing by other cytosolic proteases can also occur in the cytosol. The peptides generated are then transported

in the lumen of the ER by the TAP transporter. Once in the ER they can be further trimmed by ER-associated peptidase ERAP1 or ERAP2 and loaded onto the MHC class I molecule with the help of the peptide loading complex, which is composed of TAP, TAPASIN, Erp57 and Calretculin. Peptide/MHC complexes that are sufficiently stable can then exit the ER lumen and reach the cell surface through the secretory pathway.

Figure 2. **Proteasome catalytic activities.** ***(A) Peptides generated by the proteasome.*** Unspliced peptides are generated following hydrolysis of the peptide bonds located at both extremities of the released peptide. Spliced peptides are generated by the assembly of two peptide fragments that are originally distant in the parental protein. This involves the creation of a peptide bond between the fragments to splice. These fragments can be spliced either in the same order, or in the reverse order to that in which the occur in the parent protein. ***(B) Peptide bond hydrolysis by the proteasome.*** The hydroxyl group of the N-terminal threonine of the catalytic subunits of the proteasome produces a nucleophilic attack on the carbonyl group of the peptide bond, leading to the production of an acyl-enzyme intermediate, in which a peptide fragment remains attached to the proteasome by an ester link. A water molecule hydrolyses the ester link between the peptide and the threonine residue, thereby restoring the hydroxyl group of the catalytic threonine and producing the C-terminal end of the antigenic peptide. ***(C) Peptide splicing in the proteasome.*** Here, the splicing of the antigenic peptide RTK_QLYPEW derived from the gp100 is shown. Following formation of the acyl-enzyme intermediate involving the fragment RTK, the free N-terminal amino-group of peptide QLYPEW present in the proteasome chamber attacks the acyl-enzyme intermediate. This leads to the creation of a new peptide bond and the creation of peptide RTKQLYPEW.

Figure 3. Production of spliced peptides recognized by anti-tumor CTL. Six peptides recognized by anti-tumor CTL were identified thus far. Three peptides are derived from gp100 and three others are derived from either FGF-5, SP110 or tyrosinase. The tyrosinase peptide

produced by reverse splicing contains two additional post-translational modifications based on the conversion of two asparagine residues into aspartates by deamidation.

Figure 4. Production of the tyrosinase peptide IYMDGTADFSF. Upon entry in the ER, tyrosinase is glycosylated at specific asparagine residues (A). Some of the glycosylated tyrosinase protein is then retrotranslocated from the ER into the cytosol, presumably due to misfolding (B). Upon deglycosylation by PNGase (peptide N-glycanase), asparagines are deamidated into aspartate residues (C). The aspartate-containing tyrosinase is then degraded by the proteasome (D), a process that leads to the reverse splicing of two fragments, each containing a post-translationally modified aspartate. Spliced, deamidated peptide then reaches the ER lumen using the TAP transporter (E), to be loaded on MHC class I molecules.

Figure 5. Schematic representation of the splicing reaction occurring at the catalytic site of the proteasome. The peptide to be cleaved is represented by a necklace with pearls indicating individual amino acids. The non-primed binding site is indicated in blue and the primed binding site in red. Each binding site is subdivided in different pockets. Three pockets are here labelled: S1 to S3 for the non-primed site and S1' to S3' for the primed binding site. Amino acids binding to the non-primed binding site are labelled P_1 to P_n and those binding to the primed binding site are labelled P_1'to P_n'. The peptide splicing reaction starts with the formation of an acyl-enzyme intermediate with the peptide located at the non-primed binding site. In the course of the peptide splicing reaction, the C-terminal splice reactant (grey) comes into vicinity of the acyl-enzyme intermediate, most likely by binding to the primed-binding site. An alternative pocket that could fit the C-terminal reactant has been postulated but is not represented here. The amine group of the C-terminal splice reactant then produces a nucleophilic attack on the ester bond of the acyl-enzyme intermediate to create the spliced peptide, which then exits the catalytic site.

References

[1] K.L. Rock, A.L. Goldberg, Degradation of cell proteins and the generation of MHC class I-presented peptides, Annu. Rev. Immunol., 17 (1999) 739-779.
[2] M.J. Androlewicz, K.S. Anderson, P. Cresswell, Evidence that transporters associated with antigen processing translocate a major histocompatibility complex class I-binding peptide into the endoplasmic reticulum in an ATP-dependent manner, Proc. Natl. Acad. Sci. U.S.A., 90 (1993) 9130-9134.
[3] T. Saric, S.C. Chang, A. Hattori, I.A. York, S. Markant, K.L. Rock, M. Tsujimoto, A.L. Goldberg, An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides, Nat. Immunol., 3 (2002) 1169-1176.
[4] L. Saveanu, O. Carroll, V. Lindo, M. Del Val, D. Lopez, Y. Lepelletier, F. Greer, L. Schomburg, D. Fruci, G. Niedermann, P.M. van Endert, Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum, Nat. Immunol., 6 (2005) 689-697.
[5] I.A. York, S.C. Chang, T. Saric, J.A. Keys, J.M. Favreau, A.L. Goldberg, K.L. Rock, The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues, Nature Immunology, 3 (2002) 1177-1184.
[6] E. Rufer, R.M. Leonhardt, M.R. Knittler, Molecular architecture of the TAP-associated MHC class I peptide-loading complex, J. Immunol., 179 (2007) 5717-5727.
[7] K. Hanada, J.W. Yewdell, J.C. Yang, Immune recognition of a human renal cancer antigen through post-translational protein splicing, Nature, 427 (2004) 252-256.
[8] N. Vigneron, V. Stroobant, J. Chapiro, A. Ooms, G. Degiovanni, S. Morel, P. van der Bruggen, T. Boon, B. Van den Eynde, An antigenic peptide produced by peptide splicing in the proteasome, Science, 304 (2004) 587-590.
[9] E.H. Warren, N.J. Vigneron, M.A. Gavin, P.G. Coulie, V. Stroobant, A. Dalet, S.S. Tykodi, S.M. Xuereb, J.K. Mito, S.R. Riddell, B.J. Van den Eynde, An antigen produced by splicing of noncontiguous peptides in the reverse order, Science, 313 (2006) 1444-1447.
[10] J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, R. Huber, Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution, Science, 268 (1995) 533-539.
[11] M. Groll, L. Ditzel, J. Löwe, D. Stock, M. Bochtler, H.D. Bartunik, R. Huber, Structure of 20S proteasome from yeast at 2.4 Å resolution, Nature, 386 (1997) 463-471.
[12] M. Unno, T. Mizushima, Y. Morimoto, Y. Tomisugi, K. Tanaka, N. Yasuoka, T. Tsukihara, The structure of the mammalian 20S proteasome at 2.75 A resolution, Structure, 10 (2002) 609-618.
[13] M. Groll, T. Clausen, Molecular shredders: how proteasomes fulfill their role, Curr. Opin. Struct. Biol., 13 (2003) 665-673.
[14] K.J. Davies, Degradation of oxidized proteins by the 20S proteasome, Biochimie, 83 (2001) 301-310.
[15] A.M. Pickering, A.L. Koop, C.Y. Teoh, G. Ermak, T. Grune, K.J. Davies, The immunoproteasome, the 20S proteasome and the PA28αβ proteasome regulator are oxidative-stress-adaptive proteolytic complexes, Biochem. J., 432 (2010) 585-594.
[16] J.E. Whittier, Y. Xiong, M.C. Rechsteiner, T.C. Squier, Hsp90 enhances degradation of oxidized calmodulin by the 20 S proteasome, J. Biol. Chem., 279 (2004) 46135-46142.
[17] M. Orlowski, Selective activation of the 20 S proteasome (multicatalytic proteinase complex) by histone h3, Biochemistry, 40 (2001) 15318-15326.
[18] A.F. Kisselev, D. Kaganovich, A.L. Goldberg, Binding of hydrophobic peptides to several non-catalytic sites promotes peptide hydrolysis by all active sites of 20 S

proteasomes. Evidence for peptide-induced channel opening in the α-rings, J. Biol. Chem., 277 (2002) 22260-22270.
[19] J. Rabl, D.M. Smith, Y. Yu, S.C. Chang, A.L. Goldberg, Y. Cheng, Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases, Mol. Cell, 30 (2008) 360-368.
[20] N. Vigneron, B.J. Van den Eynde, Proteasome Subtypes and Regulators in the Processing of Antigenic Peptides Presented by Class I Molecules of the Major Histocompatibility Complex, Biomolecules, 4 (2014) 994-1025.
[21] M. Chu-Ping, C.A. Slaughter, G.N. DeMartino, Purification and characterization of a protein inhibitor of the 20S proteasome (macropain), Biochim. Biophys. Acta, 1119 (1992) 303-311.
[22] M. Bader, S. Benjamin, O.L. Wapinski, D.M. Smith, A.L. Goldberg, H. Steller, A conserved F box regulatory complex controls proteasome activity in Drosophila, Cell, 145 (2011) 371-382.
[23] X. Li, D. Thompson, B. Kumar, G.N. DeMartino, Molecular and cellular roles of PI31 (PSMF1) protein in regulation of proteasome function, J. Biol. Chem., 289 (2014) 17392-17405.
[24] Y. Sun, A.J. Sijts, M. Song, K. Janek, A.K. Nussbaum, S. Kral, M. Schirle, S. Stevanovic, A. Paschen, H. Schild, P.M. Kloetzel, D. Schadendorf, Expression of the proteasome activator PA28 rescues the presentation of a cytotoxic T lymphocyte epitope on melanoma cells, Cancer. Res., 62 (2002) 2875-2882.
[25] M. Keller, F. Ebstein, E. Burger, K. Textoris-Taube, X. Gorny, S. Urban, F. Zhao, T. Dannenberg, A. Sucker, C. Keller, L. Saveanu, E. Kruger, H.J. Rothkotter, B. Dahlmann, P. Henklein, A. Voigt, U. Kuckelkorn, A. Paschen, P.M. Kloetzel, U. Seifert, The proteasome immunosubunits, PA28 and ER-aminopeptidase 1 protect melanoma cells from efficient MART-126-35 -specific T-cell recognition, Eur. J. Immunol., 45 (2015) 3257-3268.
[26] D. Respondek, M. Voss, I. Kuhlewindt, K. Klingel, E. Kruger, A. Beling, PA28 modulates antigen processing and viral replication during coxsackievirus B3 infection, PLoS One, 12 (2017) e0173259.
[27] E. Seemuller, A. Lupas, D. Stock, J. Lowe, R. Huber, W. Baumeister, Proteasome from Thermoplasma acidophilum: a threonine protease, Science, 268 (1995) 579-582.
[28] A. Dalet, N. Vigneron, V. Stroobant, K. Hanada, B.J. Van den Eynde, Splicing of distant peptide fragments occurs in the proteasome by transpeptidation and produces the spliced antigenic peptide derived from fibroblast growth factor-5, J. Immunol., 184 (2010) 3016-3024.
[29] A. Dalet, P.F. Robbins, V. Stroobant, N. Vigneron, Y.F. Li, M. El-Gamil, K. Hanada, J.C. Yang, S.A. Rosenberg, B.J. Van den Eynde, An antigenic peptide produced by reverse splicing and double asparagine deamidation, Proc. Natl. Acad. Sci. USA, 108 (2011) E323-331.
[30] A. Michaux, P. Larrieu, V. Stroobant, J.F. Fonteneau, F. Jotereau, B.J. Van den Eynde, A. Moreau-Aubry, N. Vigneron, A Spliced Antigenic Peptide Comprising a Single Spliced Amino Acid Is Produced in the Proteasome by Reverse Splicing of a Longer Peptide Fragment followed by Trimming, J. Immunol., 192 (2014) 1962-1971.
[31] J. Liepe, M. Mishto, K. Textoris-Taube, K. Janek, C. Keller, P. Henklein, P.M. Kloetzel, A. Zaikin, The 20S proteasome splicing activity discovered by SpliceMet, PLoS Comput. Biol., 6 (2010) e1000830.
[32] F. Ebstein, K. Textoris-Taube, C. Keller, R. Golnik, N. Vigneron, B.J. Van den Eynde, B. Schuler-Thurner, D. Schadendorf, F.K. Lorenz, W. Uckert, S. Urban, A. Lehmann, N. Albrecht-Koepke, K. Janek, P. Henklein, A. Niewienda, P.M. Kloetzel, M. Mishto,

Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes, Sci. Rep., 6 (2016) 24032.
[33] M. Mishto, A. Goede, K.T. Taube, C. Keller, K. Janek, P. Henklein, A. Niewienda, A. Kloss, S. Gohlke, B. Dahlmann, C. Enenkel, P.M. Kloetzel, Driving forces of proteasome-catalyzed peptide splicing in yeast and humans, Mol. Cell. Proteomics, 11 (2012) 1008-1023.
[34] J. Liepe, F. Marino, J. Sidney, A. Jeko, D.E. Bunting, A. Sette, P.M. Kloetzel, M.P. Stumpf, A.J. Heck, M. Mishto, A large fraction of HLA class I ligands are proteasome-generated spliced peptides, Science, 354 (2016) 354-358.
[35] A.C.M. Platteel, J. Liepe, K. Textoris-Taube, C. Keller, P. Henklein, H.H. Schalkwijk, R. Cardoso, P.M. Kloetzel, M. Mishto, A. Sijts, Multi-level Strategy for Identifying Proteasome-Catalyzed Spliced Epitopes Targeted by CD8+ T Cells during Bacterial Infection, Cell. Rep., 20 (2017) 1242-1253.
[36] N. Vigneron, B.J. Van den Eynde, Proteasome subtypes and the processing of tumor antigens: increasing antigenic diversity, Curr. Opin. Immunol., 24 (2012) 84-91.
[37] M. Orlowski, S. Wilk, Catalytic activities of the 20 S proteasome, a multicatalytic proteinase complex, Arch. Biochem. Biophys., 383 (2000) 1-16.
[38] M. Orlowski, C. Cardozo, C. Michaud, Evidence for the presence of five distinct proteolytic components in the pituitary multicatalytic proteinase complex. Properties of two components cleaving bonds on the carboxyl side of branched chain and small neutral amino acids, Biochemistry, 32 (1993) 1563-1572.
[39] M. Gaczynska, A.L. Goldberg, K. Tanaka, K.B. Hendil, K.L. Rock, Proteasome subunits X and Y alter peptidase activities in opposite ways to the interferon-g-induced subunits LMP2 and LMP7, J. Biol. Chem., 271 (1996) 17275-17280.
[40] M. Gaczynska, K.L. Rock, A.L. Goldberg, γ-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes, Nature, 365 (1993) 264-267.
[41] M. Gaczynska, K.L. Rock, T. Spies, A.L. Goldberg, Peptidase activities of proteasomes are differentially regulated by the major histocompatibility complex-encoded genes for LMP2 and LMP7, Proc. Natl. Acad. Sci. USA, 91 (1994) 9213-9217.
[42] M. Groll, R. Huber, Inhibitors of the eukaryotic 20S proteasome core particle: a structural approach, Biochim. Biophys. Acta, 1695 (2004) 33-44.
[43] E.M. Huber, M. Basler, R. Schwab, W. Heinemeyer, C.J. Kirk, M. Groettrup, M. Groll, Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity, Cell, 148 (2012) 727-738.
[44] E.Z. Kincaid, J.W. Che, I. York, H. Escobar, E. Reyes-Vargas, J.C. Delgado, R.M. Welsh, M.L. Karow, A.J. Murphy, D.M. Valenzuela, G.D. Yancopoulos, K.L. Rock, Mice completely lacking immunoproteasomes show major changes in antigen presentation, Nat. Immunol., 13 (2012) 129-135.
[45] S. Morel, F. Lévy, O. Burlet-Schiltz, F. Brasseur, M. Probst-Kepper, A.-L. Peitrequin, B. Monsarrat, R. Van Velthoven, J.-C. Cerottini, T. Boon, J.E. Gairin, B.J. Van den Eynde, Processing of some antigens by the standard proteasome but not by the immunoproteasome results in poor presentation by dendritic cells, Immunity, 12 (2000) 107-117.
[46] B. Guillaume, V. Stroobant, M.P. Bousquet-Dubouch, D. Colau, J. Chapiro, N. Parmentier, A. Dalet, B.J. Van den Eynde, Analysis of the processing of seven human tumor antigens by intermediate proteasomes, J. Immunol., 189 (2012) 3538-3547.
[47] J. Chapiro, S. Claverol, F. Piette, W. Ma, V. Stroobant, B. Guillaume, J.-E. Gairin, S. Morel, O. Burlet-Schiltz, B. Monsarrat, T. Boon, B. Van den Eynde, Destructive cleavage of

antigenic peptides either by the immunoproteasome or by the standard proteasome results in differential antigen presentation, J. Immunol., 176 (2006) 1053-1061.
[48] A. Dalet, V. Stroobant, N. Vigneron, B.J. Van den Eynde, Differences in the production of spliced antigenic peptides by the standard proteasome and the immunoproteasome, Eur. J. Immunol., 41 (2011) 39-46.
[49] B. Guillaume, J. Chapiro, V. Stroobant, D. Colau, B. Van Holle, G. Parvizi, M.P. Bousquet-Dubouch, I. Theate, N. Parmentier, B.J. Van den Eynde, Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules, Proc. Natl. Acad. Sci. U.S.A., 107 (2010) 18599-18604.
[50] S. Murata, K. Sasaki, T. Kishimoto, S. Niwa, H. Hayashi, Y. Takahama, K. Tanaka, Regulation of CD8+ T cell development by thymus-specific proteasomes, Science, 316 (2007) 1349-1353.
[51] T. Nitta, S. Murata, K. Sasaki, H. Fujii, A.M. Ripen, N. Ishimaru, S. Koyasu, K. Tanaka, Y. Takahama, Thymoproteasome shapes immunocompetent repertoire of CD8+ T cells, Immunity, 32 (2010) 29-40.
[52] C.R. Berkers, A. de Jong, K.G. Schuurman, C. Linnemann, H.D. Meiring, L. Janssen, J.J. Neefjes, T.N. Schumacher, B. Rodenko, H. Ovaa, Definition of Proteasomal Peptide Splicing Rules for High-Efficiency Spliced Peptide Presentation by MHC Class I Molecules, J. Immunol., 195 (2015) 4085-4095.
[53] C.R. Berkers, A. de Jong, K.G. Schuurman, C. Linnemann, J.A. Geenevasen, T.N. Schumacher, B. Rodenko, H. Ovaa, Peptide Splicing in the Proteasome Creates a Novel Type of Antigen with an Isopeptide Linkage, J. Immunol., 195 (2015) 4075-4084.
[54] T. Delong, T.A. Wiles, R.L. Baker, B. Bradley, G. Barbour, R. Reisdorph, M. Armstrong, R.L. Powell, N. Reisdorph, N. Kumar, C.M. Elso, M. DeNicola, R. Bottino, A.C. Powers, D.M. Harlan, S.C. Kent, S.I. Mannering, K. Haskins, Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion, Science, 351 (2016) 711-714.
[55] T.A. Wiles, T. Delong, R.L. Baker, B. Bradley, G. Barbour, R.L. Powell, N. Reisdorph, K. Haskins, An insulin-IAPP hybrid peptide is an endogenous antigen for CD4 T cells in the non-obese diabetic mouse, J. Autoimmun., 78 (2017) 11-18.
[56] J.A. Babon, M.E. DeNicola, D.M. Blodgett, I. Crevecoeur, T.S. Buttrick, R. Maehr, R. Bottino, A. Naji, J. Kaddis, W. Elyaman, E.A. James, R. Haliyur, M. Brissova, L. Overbergh, C. Mathieu, T. Delong, K. Haskins, A. Pugliese, M. Campbell-Thompson, C. Mathews, M.A. Atkinson, A.C. Powers, D.M. Harlan, S.C. Kent, Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes, Nat. Med., 22 (2016) 1482-1487.
[57] L. Klein, B. Kyewski, P.M. Allen, K.A. Hogquist, Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see), Nature reviews. Immunology, 14 (2014) 377-391.
[58] W.R. Heath, F.R. Carbone, Cross-presentation in viral immunity and self-tolerance, Nature reviews. Immunology, 1 (2001) 126-134.
[59] A.Y. Huang, P. Golumbek, M. Ahmadzadeh, E. Jaffee, D. Pardoll, H. Levitsky, Role of bone marrow-derived cells in presenting MHC class I-restricted tumor antigens, Science, 264 (1994) 961-965.
[60] L.J. Sigal, S. Crotty, R. Andino, K.L. Rock, Cytotoxic T-cell immunity to virus-infected non-haematopoietic cells requires presentation of exogenous antigen, Nature, 398 (1999) 77-80.
[61] W.R. Heath, F.R. Carbone, Cross-presentation, dendritic cells, tolerance and immunity, Annu. Rev. Immunol., 19 (2001) 47-64.

[62] A.D. Garg, P.G. Coulie, B.J. Van den Eynde, P. Agostinis, Integrating Next-Generation Dendritic Cell Vaccines into the Current Cancer Immunotherapy Landscape, Trends Immunol., 38 (2017) 577-593.
[63] C. Kurts, H. Kosaka, F.R. Carbone, J.F. Miller, W.R. Heath, Class I-restricted cross-presentation of exogenous self-antigens leads to deletion of autoreactive CD8(+) T cells, J. Exp. Med., 186 (1997) 239-245.
[64] N. Luckashenak, S. Schroeder, K. Endt, D. Schmidt, K. Mahnke, M.F. Bachmann, P. Marconi, C.A. Deeg, T. Brocker, Constitutive crosspresentation of tissue antigens by dendritic cells controls CD8+ T cell tolerance in vivo, Immunity, 28 (2008) 521-532.
[65] A.M. Gallegos, M.J. Bevan, Central tolerance to tissue-specific antigens mediated by direct and indirect antigen presentation, J. Exp. Med., 200 (2004) 1039-1049.
[66] L. Chapatte, M. Ayyoub, S. Morel, A.L. Peitrequin, N. Levy, C. Servis, B.J. Van den Eynde, D. Valmori, F. Levy, Processing of tumor-associated antigen by the proteasomes of dendritic cells controls *in vivo* T-cell responses, Cancer Res., 66 (2006) 5461-5468.

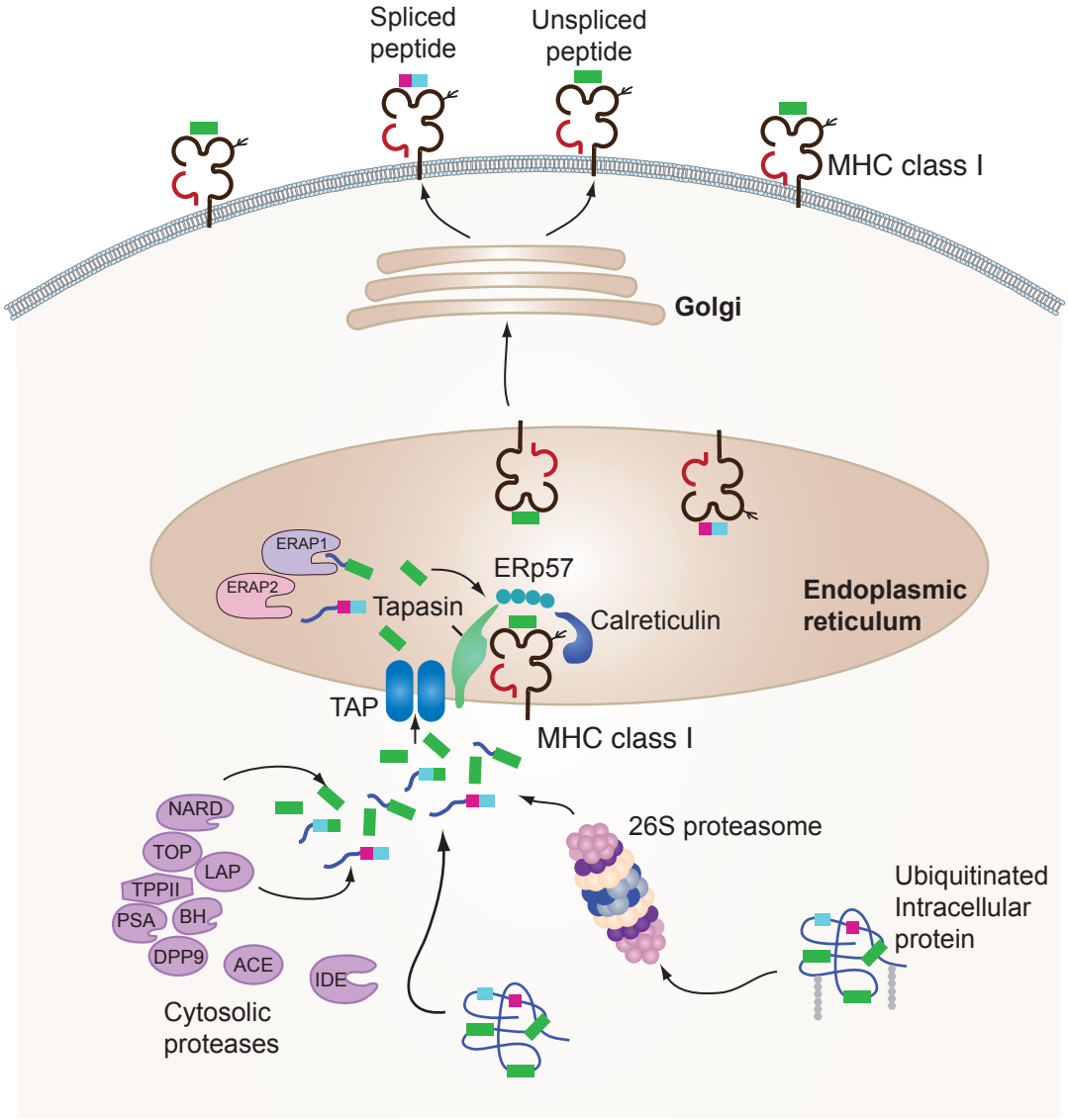
Spliced peptide
Unspliced peptide
MHC class I
Golgi
ERAP1
ERAP2
Tapasin
ERp57
Calreticulin
Endoplasmic reticulum
TAP
MHC class I
NARD
TOP
LAP
TPPII
PSA
BH
DPP9
ACE
IDE
Cytosolic proteases
26S proteasome
Ubiquitinated Intracellular protein

Figure 1.

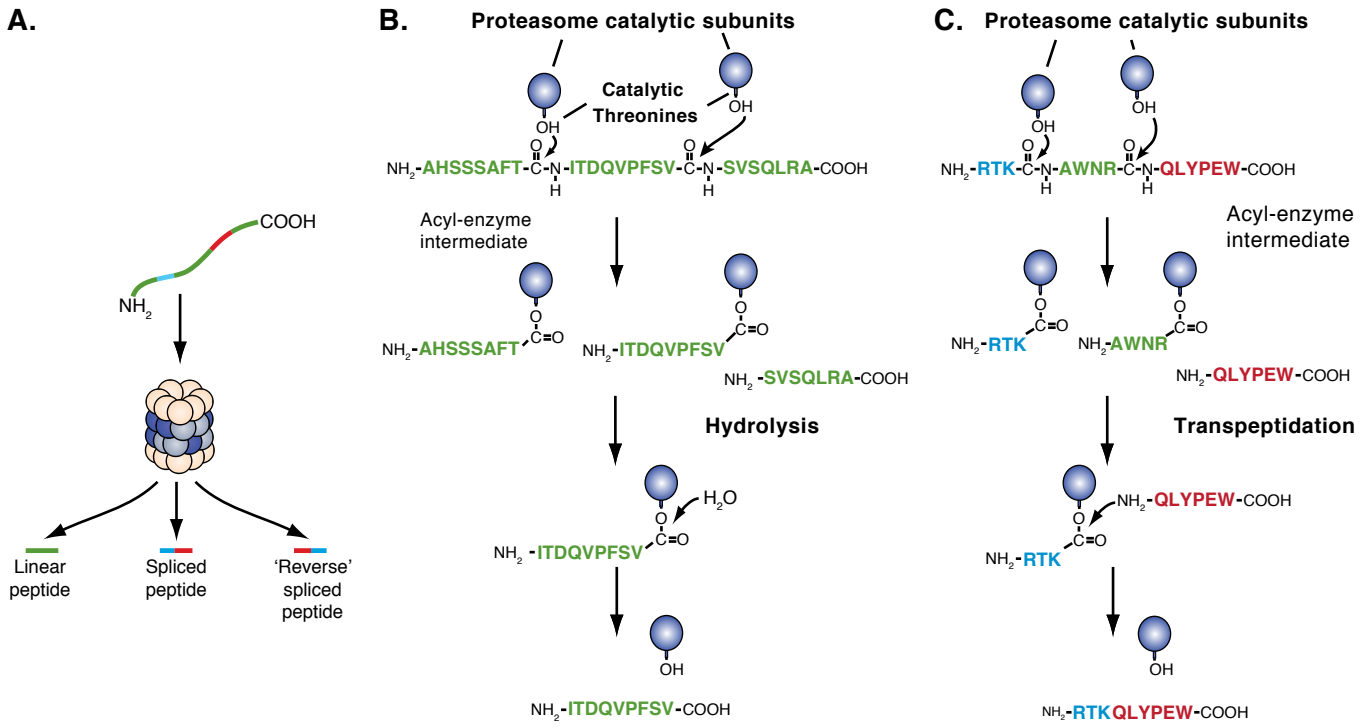
A.
COOH
NH2
Linear peptide
Spliced peptide
'Reverse' spliced peptide
B.
Proteasome catalytic subunits
Catalytic Threonines
OH
OH
NH2-AHSSSAFT-C-N-ITDQVPFSV-C-N-SVSQLRA-COOH
H
H
Acyl-enzyme intermediate
NH2-AHSSSAFT
NH2-ITDQVPFSV
NH2-SVSQLRA-COOH
Hydrolysis
H2O
NH2-ITDQVPFSV
NH2-ITDQVPFSV-COOH
C.
Proteasome catalytic subunits
NH2-RTK-C-N-AWNR-C-N-QLYPEW-COOH
Acyl-enzyme intermediate
NH2-RTK
NH2-AWNR
NH2-QLYPEW-COOH
Transpeptidation
NH2-QLYPEW-COOH
NH2-RTK
NH2-RTKQLYPEW-COOH

Figure 2.

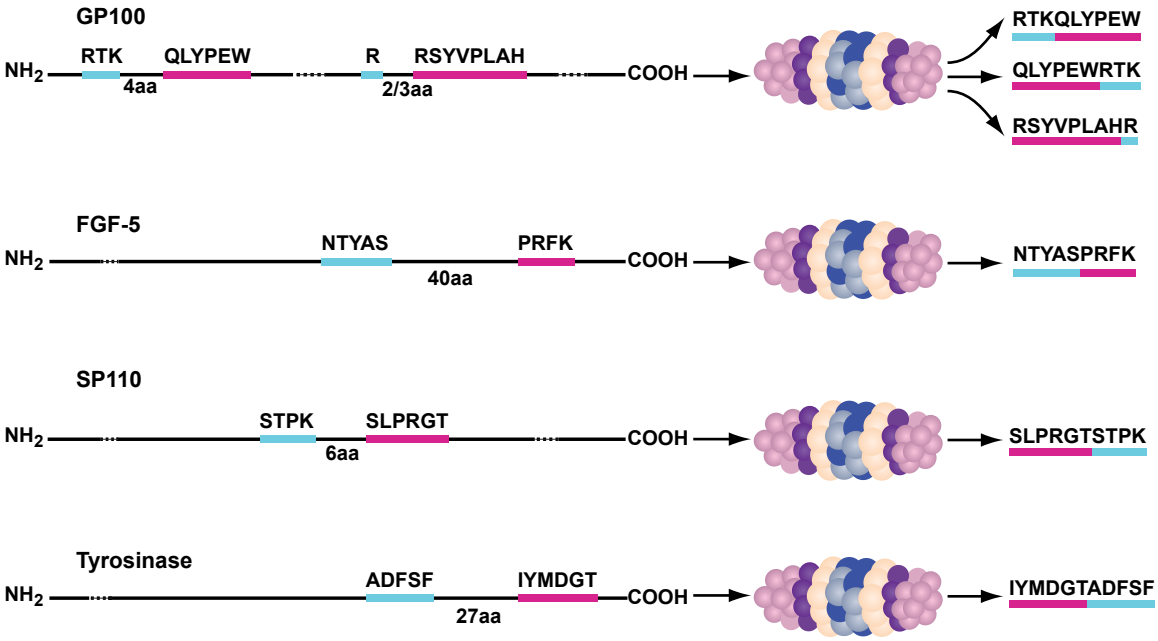
GP100
NH2
RTK
4aa
QLYPEW
R
2/3aa
RSYVPLAH
COOH
RTKQLYPEW
QLYPEWRTK
RSYVPLAHR
FGF-5
NH2
NTYAS
40aa
PRFK
COOH
NTYASPRFK
SP110
NH2
STPK
6aa
SLPRGT
COOH
SLPRGTSTPK
Tyrosinase
NH2
ADFSF
27aa
IYMDGT
COOH
IYMDGTADFSF

Figure 3

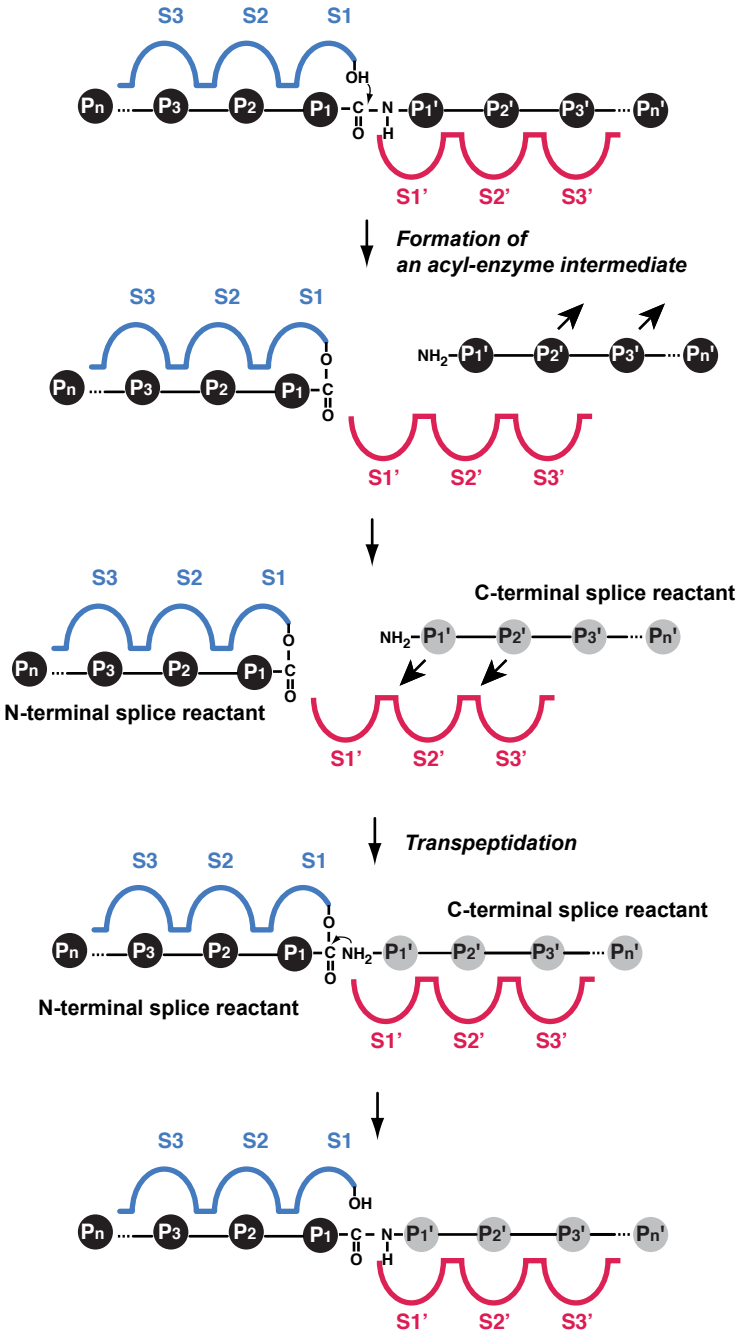
S3
S2
S1
Pn
P3
P2
P1
OH
C
O
N
H
P1'
P2'
P3'
Pn'
S1'
S2'
S3'
Formation of
an acyl-enzyme intermediate
NH2
C-terminal splice reactant
N-terminal splice reactant
Transpeptidation

Figure 5

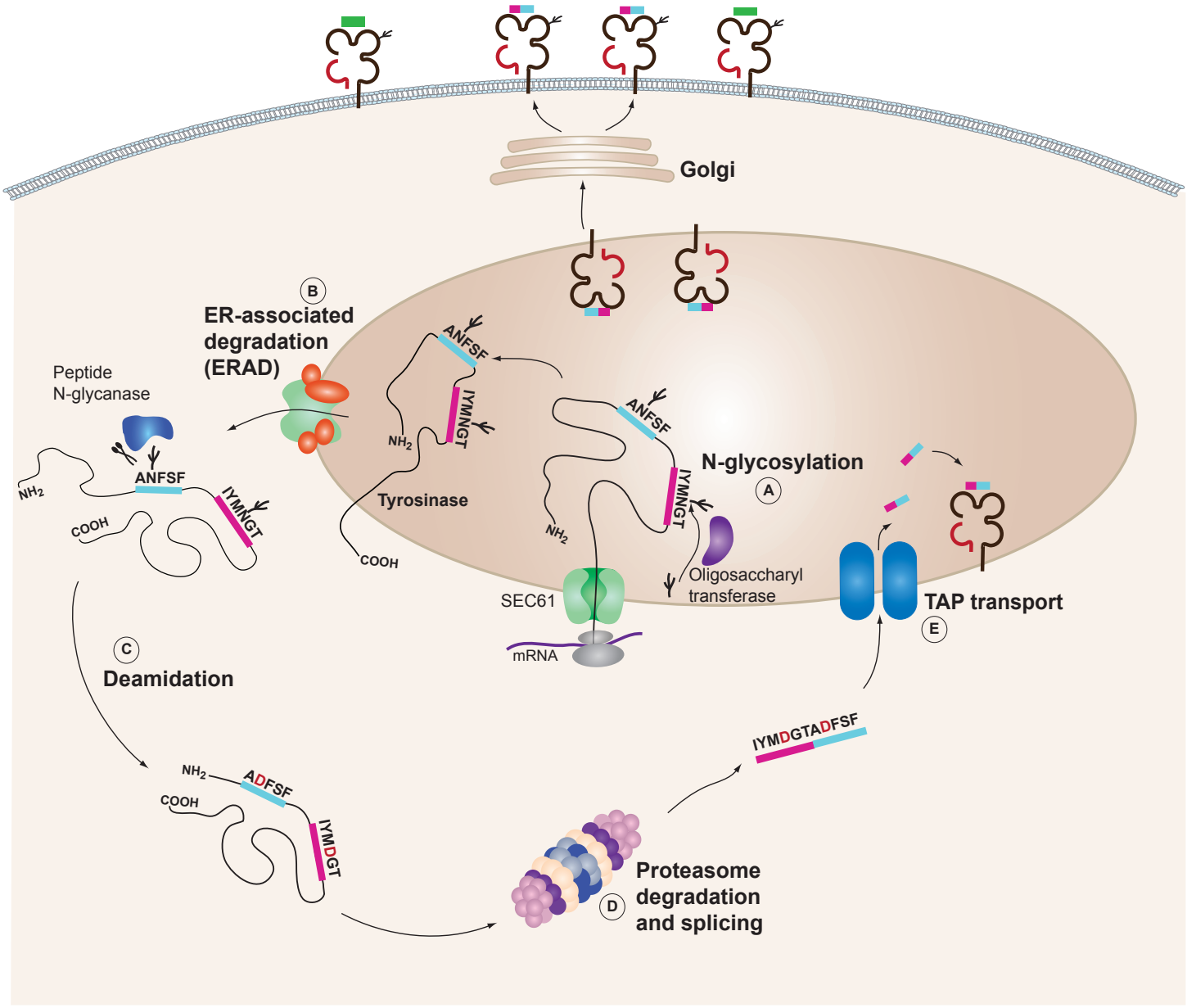
Golgi
B
ER-associated
degradation
(ERAD)
Peptide
N-glycanase
ANFSF
IYMNGT
NH2
COOH
Tyrosinase
N-glycosylation
A
Oligosaccharyl
transferase
SEC61
mRNA
TAP transport
E
C
Deamidation
ADFSF
IYMDGT
IYMDGTADFSF
D
Proteasome
degradation
and splicing

Figure 4