Are There Indeed Spliced Peptides in the Immunopeptidome?

Author
Arie Admon

Correspondence
admon@technion.ac.il

In Brief
Peptide splicing was suggested to significantly contribute ligands to the immunopeptidome. This article argues that peptide splicing is at most very rare, even if it happens at all. Considerations against peptide splicing are based on bioinformatics calculations related to the analysis of the LC-MS/MS data, and on the abundance of water in the cells, which should compete effectively with the transpeptidation reaction, needed for peptide splicing.

Graphical Abstract

(peptide + water ↔ peptide + peptide)
is more likely than
(peptide + nucleophile spliced-in peptide ↔ spliced peptide + remainder peptide).

Highlights

- Peptide splicing was suggested to contribute to the immunopeptidome.
- I suggest that this idea should be reconsidered based on new evidences.
- Both biochemical and bioinformatics considerations argue against peptide splicing.

2021, Mol Cell Proteomics *20*, 100099

https://doi.org/10.1016/j.mcpro.2021.100099

Are There Indeed Spliced Peptides in the Immunopeptidome?

Arie Admon*

The claims that a large fraction of the immunopeptidome is composed of spliced major histocompatibility complex (MHC) peptides have stirred significant excitement and raised controversy. Here, I suggest that there are likely no spliced peptides in the immunopeptidome, and if they exist at all, they are extremely rare. I base this claim on both biochemical and bioinformatics considerations. First, as a reactant in normal proteolytic reactions, water will compete with transpeptidation, which has been suggested as the mechanism of peptide splicing. The high mobility and abundance of water in aqueous solutions renders transpeptidation very inefficient and therefore unlikely to occur. Second, new studies have refuted the bioinformatics assignments to spliced peptides of most of the immunopeptidome MS data, suggesting that the correct assignments are likely other canonical, noncanonical, and post-translationally modified peptides. Therefore, I call for rigorous experimental methodology using heavy stable isotope peptides spiking into the immunoaffinity-purified mixtures of natural MHC peptides and analysis by the highly reliable targeted MS, to claim that MHC peptides are indeed spliced.

The major histocompatibility complex (MHC) molecules transport peptides from the cell interior and 'present' them at the cell surface, enabling immune scrutiny of the health status of the cells. MHC class I molecules are expressed on most nucleated cells of vertebrates and present peptides derived from proteolysis of intracellular proteins. MHC class II molecules are expressed on immune cells and present peptides originating from intracellular proteins and from proteins taken up from outside the cells. Presentation of just a few molecules of pathogen-derived (or cancer-specific) peptides is sufficient to elicit strong immune responses, to cytotoxic killing of the diseased cells, and to stimulation, expansion, and differentiation of B cells into plasma cells, leading to the production of anti-pathogen antibodies (1, 2).

A significant degree of excitement has arisen in the field of MHC peptidome (immunopeptidome) research, with the claim that some MHC-presented peptides derive from peptide splicing (3–5), as reviewed (6–12). This concept has stirred the field tremendously; if indeed, some peptides are spliced before loading onto the MHC, the size of the required protein sequence database used for searching peptide ligands, is many folds larger than the currently used canonical protein sequence database. It would also suggest that many of the currently identified sequences of nonspliced human leukocyte antigen (HLA) peptides are incorrect. This possibility may also complicate the central dogma of immune tolerance, namely, that presentation of 'self' peptides in the thymus leads to deletion of the 'anti-self' T cells during the development of the adaptive immune system (1, 2). To enable such immune tolerance, identical peptide splicing reactions must occur both in the thymus and in the periphery (8). The proteasome was suggested as the likely site of these proteasome-catalyzed peptide splicing because it is considered the main producer of peptides for MHC presentation (13–18) and treatment of the cells with proteasome inhibitors modulated the presentation level of the supposedly spliced peptides (3, 19, 20). The notion that proteasome-catalyzed peptide splicing produces larger pools of ligands for MHC peptidome presentation was considerably promoted after the implementation of dedicated software tools that use LC-MS/MS data to identify spliced peptides (SpliceMet) (21).

The proteasomes have been suggested to perform the splicing reaction by transpeptidation (3, 5, 22–24), with *cis*-splicing preferred over *trans*-splicing (splicing fragments of the peptides from one versus two different proteins) (23, 25). Splicing of peptides in the reverse order was also suggested (5, 26), as well as splicing with isopeptide bonds (27). Differential effects of the standard proteasomes and immunoproteasomes on peptide splicing were also demonstrated (19, 22, 28). In addition, hybrid insulin peptides were suggested to be processed in the β cell crinophagic bodies (crinosomes), where insulin degradation products are fused and then presented by MHC class II molecules, leading to diabetes and autoimmunity (11, 29–32). More recently, it was also suggested that fusion between human and viral peptides with subsequent splicing may increase the range of the immunopeptidomes even further, and induce immunogenic reactions that can lead to diabetes (33, 34). Furthermore, in very careful

From the Faculty of Biology, Technion-Israel Institute of Technology, Haifa, Israel
*For correspondence: Arie Admon, admon@technion.ac.il.

https://doi.org/10.1016/j.mcpro.2021.100099

and meticulous analyses, the authors of the first publications on spliced MHC peptides (3, 4) demonstrated that only the spliced peptide sequences elicited immune responses by the T cells clones, which were oblivious to the nonspliced peptides. The immunopeptidome research community has generally related to these early findings as an exciting possibility and awaited further studies and the development of methodologies to discover the few spliced peptides among the large immunopeptidomes.

A major twist in the story of spliced HLA peptides occurred when Liepe and Mishto in the Kloetzel and Heck labs (20) performed a large-scale search for spliced MHC peptides in their own immunopeptidome LC-MS/MS data, as well as the data of Bassani-Sternberg *et al.* (35), and have shown that approximately 25% of the HLA peptidome is composed of spliced peptides. Supportive evidence was obtained using additional cultured human cell lines, primary cancer cells (36), new software tools (37–39), and cells infected with pathogens (40–42), as reviewed (8–10, 43, 44). Yet, other laboratories disagreed with the claim that a large percentage of the immunopeptidome is spliced and that many of the identifications in the canonical immunopeptidome data are incorrect. The reanalysis of the data of Bassani-Sternberg *et al.* (35) used by Liepe *et al.* (20) concluded that, at most, the percentage of spliced peptides does not exceed 2 to 6% of the canonical immunopeptidome (45). First, the spliced peptides' sequences do not fit the canonical sequence motifs of the cells (lengths and anchor residues), as do the canonical peptide sequences identified with the same MS/MS spectra. The spliced peptide sequences had worse MS/MS identification scores than the canonical, had higher mass errors, and fitted less the synthetic peptides' MS/MS relative to the canonical (UniProt) peptides. Furthermore, many of the MS/MS of the supposedly spliced peptides fitted better post-translationally modified canonical peptides (45). Subsequently, the published spliced peptide data of Faridi *et al.* (37) that claimed 13 to 45% of the immunopeptidome is composed of spliced peptides were reanalyzed by Rolfs *et al.* (38) using newly developed software tool (Neo-Fusion) and determined that the percentage of spliced peptides was at most 1 to 6%. The canonical peptides have better scores than the spliced peptides, and between calculated and observed chromatography retention times. Reanalysis by Lichti (46) of the data of Wan *et al.*, Bassani-Sternberg *et al.*, and Faridi *et al.* (31, 35, 37) reached similar conclusions, namely that only a small subset of the immunopeptidome is possibly spliced and that many of the previous spliced peptide identifications were actually of canonical post-translationally modified peptides (except the HLA-A*01:01 cell line of Faridi *et al*. (37) where up to 17% of the peptides could be explained by splicing). A series of comments and responses that ensued these publications (47, 48) did not result in an agreement, and the debate is still going on (44) and becoming even more emotional. Recently, Willimsky *et al.* (49) also concluded that the frequency of spliced epitopes is largely overestimated.

As with proteomics, MHC peptides are identified by LC-MS/MS using their observed precise masses and the masses of their gas-phase generated fragments (50), followed by a comparison of the data to *in silico*–generated simulated spectra of 'no-enzyme' proteolysis of the known protein sequences of the animal (51–56). When peptide splicing is considered, a significantly larger protein sequence database than the canonical should be used to include all potential spliced sequences. The larger sequence databases increase the chance of assignment of erroneous peptide sequences, even under a stringent false discovery rate (FDR) of 1%, and reduces the number of correctly identified peptides (57–59). As described above, the efforts to overcome these limitations included the development of new software tools to identify spliced peptides, such as that designed by Rolfs *et al.* (38), who evaluated different immunopeptidome LC-MS/MS datasets and came to the conclusion that indeed the percentage of the spliced peptides among the immunopeptidome is only about 1% of the identified peptidome, and by Mylonas *et al.* (45) who concluded about 2 to 6% spliced peptides. In contrast, Faridi *et al.* (37) performed a similar analysis, using their LC-MS/MS search procedure, combined with *de novo* sequencing with the PEAKS Studio software tool (60), but using a significantly smaller decoy database, and came to a similar conclusion as did Liepe *et al.* (20), namely that a large fraction of the immunopeptidome is indeed composed of spliced peptides. Importantly, Mishto *et al.* (22) analyzed peptide splicing using isolated proteasomes (of yeasts) in $H_2{}^{18}O$ buffer and pools of synthetic peptides and concluded that 1 to 2% of the processed peptides are spliced. In addition, Berkers *et al.* (61), Paes *et al.* (62), and Specht *et al.* (25) used a set of synthetic peptides to elucidate the patterns and degree of peptide splicing by isolated human proteasomes, while Kuckelkorn *et al.* (63) have performed this analysis also with thymoproteasomes and immunoproteasomes, and both studies have detected a significant fraction of the peptides as being spliced peptides. Recently, Erhard *et al.* (64) performed a similar search using a new software tool (Peptide-PRISM) that also uses the PEAKS *de novo* search engine and suggested that most of the supposedly spliced peptides in the study by Liepe *et al.* (20) are actually products of canonical peptides and noncanonical translations from previously unidentified ORFs. They concluded that the percentage of spliced MHC peptides is less than 1% of the immunopeptidome. The authors provide a detailed explanation on the problem of erroneous identification of peptide sequences because of problems with assigning the correct FDR by the target-decoy approach in the Supplementary notes of Erhard *et al.* (64).

Here, I argue that peptide splicing is, at most, an extremely rare event and likely does not happen at all. I base this claim on two main arguments. The first is a biochemical consideration. For peptide splicing to be catalyzed by a protease, such as a proteasome, by transpeptidation during normal

proteolysis, water must be excluded from its active site. Peptide splicing requires that the highly unstable acyl-enzyme reaction intermediate be attacked by the nucleophile N terminus of another peptide that eventually forms the spliced-in partner of the final product. During regular proteolysis, water molecules provide this nucleophile, attack the acyl-enzyme intermediate, and release the proteolysis product from the active site (8, 22, 24). The water molecules are eventually split to hydrolyze the peptide bond and contribute hydroxyl and hydrogen to each of the products. Water is (obviously) very abundant in aqueous solutions, with a concentration of about 55 M, whereas the peptide substrate candidates for splicing are present at concentrations many orders of magnitude lower, even within the proteasome cavity. It is true that within the confinement of the active site of the proteasome, the local concentration of free peptides can be much higher than their concentrations in the surrounding cytoplasm. However, the proteasome cavity is a large and open cavity, with an abundance of free water molecules (65–67). Similarly, in the beta cell granules or crinophagic bodies, the concentration of proteins and peptides can be very high, thus favoring peptide splicing (11). Therefore, to achieve effective peptide splicing, the nucleophile N terminus of another free peptide must diffuse in, and attack the bound unstable acyl-enzyme intermediate faster than the abundant and highly mobile water molecules present within the same cavity. Alternatively, the nucleophile peptides may occupy simultaneously the protease active site, with the protein substrate, thus preventing water from diffusing in. No evidence for concurrent occupation of the proteasome active sites by two peptides has been demonstrated so far. The diffusion rates of peptides and proteins are orders of magnitude lower than those of water molecules. Therefore, enzymatic reactions that have a protein (or peptide) and water as substrates, and two peptides as products (peptide + water ↔ peptide + peptide), are much more likely to occur than a splicing reaction using two peptides as substrates and producing two other peptides as products (long peptide + nucleophile spliced-in peptide ↔ spliced peptide + remainder peptide).

The second argument is based on bioinformatics considerations. Although MS is a truly advanced methodology, analysis of LC-MS/MS data is still prone to some misidentifications. The level of error in identification of spliced peptides using LC-MS/MS data is large because of the sheer size of the sequence databases used for the search, which are many folds larger than the normal protein sequence database used for the identification of normal peptides. These databases are often so large that they cannot be deposited in the public repositories and are not available for evaluation by peers (20, 36, 37, 39). Using very large sequence databases or bioinformatics tools that create large numbers of alternative sequence variants results in the identification of large numbers of erroneous identifications, even when the FDR is set to the low 1%. This problem requires employing of stringent bioinformatics and experimental routines for correct assignment of peptide sequences (45, 47, 55, 68) and thus, such 'Extraordinary claims require extraordinary evidence' (attributed to Carl Sagan), as reviewed in (69). The claim of a large percentage of spliced peptides within the immunopeptidome requires extraordinary validations, such as analysis of synthetic versions of the selected natural peptides (see, for example, (70)), as was indeed done in the original claims of HLA peptide splicing (3–5, 7) and in the more recent studies (20, 37, 40, 42, 45, 71) on noncanonical cryptic peptides (64). There is also a need to overcome the risk of circular arguments; if a peptide sequence is predicted from a particular MS/MS spectrum and a synthetic peptide is produced based on the predicted sequence, the synthetic peptide will likely fragment similarly to the observed natural peptide. This does not provide full proof that their sequences are identical because similar fragmentation patterns may result from similar but nonidentical sequences, such as peptides with leucine or isoleucine that have the exact same mass. Sound proof that the synthetic peptide is identical to the natural one requires additional similarities other than their MS/MS patterns. For example, both synthetic and natural peptides and their fragments should have identical chromatography retention times. Their fragmentation patterns should be identical, both with regard to the fragment composition and in the relative intensities of the fragments throughout the elution profile of their chromatography peaks. Currently, the safer way to prove such correct assignment of peptide sequences is by spiking natural peptide mixtures with heavy stable isotope-labeled synthetic peptides and demonstrating that these heavy peptides coelute and cofragment exactly like their natural peptide counterparts. Performing this analysis with unlabeled, light isotope–labeled synthetic peptides is not possible because they are indistinguishable in their masses and therefore mask the signals of natural peptides, when spiked in. To perform this analysis properly, targeted MS should be used, in which the mass spectrometers are instructed to fragment multiple times, both the natural and spiked-in heavy isotope–labeled synthetic peptides during their elution peak. These methods are commonly used for targeted proteomics (72, 73) and provide the best evidence that both natural and synthetic peptides have identical retention times and fragmentation patterns (51, 74–77). However, even this approach is not completely 'problem-free', heavy isotope–labeled peptides are costly and not readily available and therefore not relevant for validation of thousands of potentially spliced peptides. A sequence that matches perfectly a spliced peptide might still be due to RNA splicing events, a phenomenon that can be discovered by performing Ribo-seq with ribosomes collected from the same cell culture (64). Importantly, heavy peptides should be obtained only from reputable sources, to ensure absence of even trace amounts of contaminating 'light' synthetic peptides that might be mistaken for the natural (spliced) peptides. Therefore, realistically, validations should only be

performed with randomly selected subsets of spliced peptide candidates, using well-annotated canonical peptides as the controls.

CONCLUSIONS

I suggest that the elegant idea of a large or even a small fraction of the MHC peptidome is spliced, is likely incorrect. While I accept that it is not possible to fully rule out the existence of some spliced peptides, the difficulty to rule out the phenomenon does not mean that they do exist. It is more likely that peptides previously thought to be spliced are actually nonspliced (canonical and noncanonical) peptides that have not yet been classified as such because of more exotic post-translational modifications (46) or noncanonical transcripts that remain to be identified (64, 78, 79). Future studies on spliced MHC peptides must resort to extraordinary validations, deposit the amino acid sequence databases used for the search, or at least provide their size, and such studies should not disregard the high abundance of water in the aqueous cytoplasm.

Acknowledgments—Research in the Admon lab is supported by the Israel Science Foundation, Israel (ISF 1435/16), and by the Koppel SCLC foundation.

Author contributions—A. A. conceived and wrote this review.

Conflict of interest—The author declares no competing interests.

Abbreviations—The abbreviations used are: FDR, false discovery rate; HLA, human leukocyte antigen.

Received December 16, 2020, and in revised from, April 13, 2021
Published, MCPRO Papers in Press, May 20, 2021, https://doi.org/10.1016/j.mcpro.2021.100099

REFERENCES

1. Rock, K. L., Reits, E., and Neefjes, J. (2016) Present yourself! By MHC class I and MHC class II molecules. *Trends Immunol.* **37**, 724–737
2. Roche, P. A., and Cresswell, P. (2016) Antigen processing and presentation mechanisms in myeloid cells. *Microbiol. Spectr.* **4**
3. Vigneron, N., Stroobant, V., Chapiro, J., Ooms, A., Degiovanni, G., Morel, S., van der Bruggen, P., Boon, T., and Van den Eynde, B. J. (2004) An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587–590
4. Hanada, K.-I., Yewdell, J. W., and Yang, J. C. (2004) Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252–256
5. Warren, E. H., Vigneron, N. J., Gavin, M. A., Coulie, P. G., Stroobant, V., Dalet, A., Tykodi, S. S., Xuereb, S. M., Mito, J. K., Riddell, S. R., and Van den Eynde, B. J. (2006) An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* **313**, 1444–1447
6. Melief, C. J. M., and Kessler, J. H. (2017) Novel insights into the HLA class I immunopeptidome and T-cell immunosurveillance. *Genome Med.* **9**, 44
7. Vigneron, N., Ferrari, V., Stroobant, V., Abi Habib, J., and Van den Eynde, B. J. (2017) Peptide splicing by the proteasome. *J. Biol. Chem.* **292**, 21170–21179
8. Vigneron, N., Stroobant, V., Ferrari, V., Abi Habib, J., and Van den Eynde, B. J. (2019) Production of spliced peptides by the proteasome. *Mol. Immunol.* **113**, 93–102
9. Liepe, J., Ovaa, H., and Mishto, M. (2018) Why do proteases mess up with antigen presentation by re-shuffling antigen sequences? *Curr. Opin. Immunol.* **52**, 81–86
10. Mannering, S. I., So, M., Elso, C. M., and Kay, T. W. H. (2018) Shuffling peptides to create T-cell epitopes: Does the immune system play cards? *Immunol. Cell Biol.* **96**, 34–40
11. Wiles, T. A., and Delong, T. (2019) HIPs and HIP-reactive T cells. *Clin. Exp. Immunol.* **198**, 306–313
12. Faridi, P., Dorvash, M., and Purcell, A. W. (2021) Spliced HLA-bound peptides: A black swan event in immunology. *Clin. Exp. Immunol.* **204**, 179–188
13. Goldberg, A. L., and Rock, K. L. (1992) Proteolysis, proteasomes and antigen presentation. *Nature* **357**, 375–379
14. Gaczynska, M., Rock, K. L., and Goldberg, A. L. (1993) Role of proteasomes in antigen presentation. *Enzyme Protein* **47**, 354–369
15. Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* **78**, 761–771
16. Rock, K. L., and Goldberg, A. L. (1999) Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* **17**, 739–779
17. Sijts, E. J. A. M., and Kloetzel, P. M. (2011) The role of the proteasome in the generation of MHC class I ligands and immune responses. *Cell. Mol. Life Sci.* **68**, 1491–1502
18. Blum, J. S., Wearsch, P. A., and Cresswell, P. (2013) Pathways of antigen processing. *Annu. Rev. Immunol.* **31**, 443–473
19. Ebstein, F., Textoris-Taube, K., Keller, C., Golnik, R., Vigneron, N., Van den Eynde, B. J., Schuler-Thurner, B., Schadendorf, D., Lorenz, F. K. M., Uckert, W., Urban, S., Lehmann, A., Albrecht-Koepke, N., Janek, K., Henklein, P., *et al*. (2016) Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes. *Sci. Rep.* **6**, 24032
20. Liepe, J., Marino, F., Sidney, J., Jeko, A., Bunting, D. E., Sette, A., Kloetzel, P. M., Stumpf, M. P. H., Heck, A. J. R., and Mishto, M. (2016) A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* **354**, 354–358
21. Liepe, J., Mishto, M., Textoris-Taube, K., Janek, K., Keller, C., Henklein, P., Kloetzel, P. M., and Zaikin, A. (2010) The 20S proteasome splicing activity discovered by SpliceMet. *PLoS Comput. Biol.* **6**, e1000830
22. Mishto, M., Goede, A., Taube, K. T., Keller, C., Janek, K., Henklein, P., Niewienda, A., Kloss, A., Gohlke, S., Dahlmann, B., Enenkel, C., and Kloetzel, P. M. (2012) Driving forces of proteasome-catalyzed peptide splicing in yeast and humans. *Mol. Cell. Proteomics* **11**, 1008–1023
23. Dalet, A., Vigneron, N., Stroobant, V., Hanada, K.-I., and Van den Eynde, B. J. (2010) 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**, 3016–3024
24. Berkers, C. R., de Jong, A., Ovaa, H., and Rodenko, B. (2009) Transpeptidation and reverse proteolysis and their consequences for immunity. *Int. J. Biochem. Cell Biol.* **41**, 66–71
25. Specht, G., Roetschke, H. P., Mansurkhodzhaev, A., Henklein, P., Textoris-Taube, K., Urlaub, H., Mishto, M., and Liepe, J. (2020) Large database for the analysis and prediction of spliced and non-spliced peptide generation by proteasomes. *Sci. Data* **7**, 146
26. Dalet, A., Robbins, P. F., Stroobant, V., Vigneron, N., Li, Y. F., El-Gamil, M., Hanada, K., Yang, J. C., Rosenberg, S. A., and Van den Eynde, B. J. (2011) An antigenic peptide produced by reverse splicing and double asparagine deamidation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E323–E331
27. Berkers, C. R., de Jong, A., Schuurman, K. G., Linnemann, C., Geenevasen, J. A. J., Schumacher, T. N. M., Rodenko, B., Ovaa, H., De Jong, A., Karianne, G., Linnemann, C., Geenevasen, J. A. J., Schumacher, N. M., Rodenko, B., Ovaa, H., *et al*. (2015) Peptide splicing in the proteasome creates a novel type of antigen with an isopeptide linkage. *J. Immunol.* **195**, 4075–4084
28. Dalet, A., Stroobant, V., Vigneron, N., and Van den Eynde, B. J. (2011) Differences in the production of spliced antigenic peptides by the

standard proteasome and the immunoproteasome. *Eur. J. Immunol.* **41**, 39–46
29. Delong, T., Wiles, T. A., Baker, R. L., Bradley, B., Barbour, G., Reisdorph, R., Armstrong, M., Powell, R. L., Reisdorph, N., Kumar, N., Elso, C. M., De Nicola, M., Bottino, R., Powers, A. C., Harlan, D. M., *et al*. (2016) Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* **351**, 711–714
30. Wiles, T. A., Powell, R., Michel, C., Beard, K. S., Hohenstein, A., Bradley, B., Reisdorph, N., Haskins, K., and Delong, T. (2019) Identification of hybrid insulin peptides (HIPs) in mouse and human islets by mass spectrometry. *J. Proteome Res.* **18**, 814–825
31. Wan, X., Vomund, A. N., Peterson, O. J., Chervonsky, A. V., Lichti, C. F., and Unanue, E. R. (2020) The MHC-II peptidome of pancreatic islets identifies key features of autoimmune peptides. *Nat. Immunol.* **21**, 455–463
32. Reed, B., Crawford, F., Hill, R. C., Jin, N., White, J., Krovi, S. H., Marrack, P., Hansen, K., and Kappler, J. W. (2021) Lysosomal cathepsin creates chimeric epitopes for diabetogenic CD4 T cells via transpeptidation. *J. Exp. Med.* **218**, e20192135
33. Mishto, M., Mansurkhodzhaev, A., Rodriguez-Calvo, T., and Liepe, J. (2021) Potential mimicry of viral and pancreatic β cell antigens through non-spliced and cis-spliced zwitter epitope candidates in type 1 diabetes. *Front Immunol* **12**, 656451
34. Mansurkhodzhaev, A., Barbosa, C. R. R., Mishto, M., and Liepe, J. (2021) Proteasome-generated cis-spliced peptides and their potential role in CD8+ T cell tolerance. *Front Immunol* **12**, 614276
35. Bassani-Sternberg, M., Pletscher-Frankild, S., Jensen, L. J., and Mann, M. (2015) Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell. Proteomics* **14**, 658–673
36. Faridi, P., Woods, K., Ostrouska, S., Deceneux, C., Aranha, R., Duscharla, D., Wong, S. Q., Chen, W., Ramarathinam, S. H., Lim Kam Sian, T. C. C. C., Croft, N. P., Li, C., Ayala, R., Cebon, J. S., Purcell, A. W., *et al*. (2020) Spliced peptides and cytokine-driven changes in the immunopeptidome of melanoma. *Cancer Immunol. Res.* **8**, 1322–1334
37. Faridi, P., Li, C., Ramarathinam, S. H., Vivian, J. P., Illing, P. T., Mifsud, N. A., Ayala, R., Song, J., Gearing, L. J., Hertzog, P. J., Ternette, N., Rossjohn, J., Croft, N. P., and Purcell, A. W. (2018) A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. *Sci. Immunol.* **3**, eaar3947
38. Rolfs, Z., SoIntsev, S. K., Shortreed, M. R., Frey, B. L., and Smith, L. M. (2019) Global identification of post-translationally spliced peptides with neo-fusion. *J. Proteome Res.* **18**, 349–358
39. Liepe, J., Sidney, J., Lorenz, F. K. M., Sette, A., and Mishto, M. (2019) Mapping the MHC class I-spliced immunopeptidome of cancer cells. *Cancer Immunol. Res.* **7**, 62–76
40. Platteel, A. C. M., Liepe, J., Textoris-Taube, K., Keller, C., Henklein, P., Schalkwijk, H. H., Cardoso, R., Kloetzel, P. M., Mishto, M., and Sijts, A. J. A. M. (2017) Multi-level strategy for identifying proteasome-catalyzed spliced epitopes targeted by CD8+ T cells during bacterial infection. *Cell Rep.* **20**, 1242–1253
41. Paes, W., Leonov, G., Partridge, T., Chikata, T., Murakoshi, H., Frangou, A., Brackenridge, S., Nicastri, A., Smith, A. G., Learn, G. H., Li, Y., Parker, R., Oka, S., Pellegrino, P., Williams, I., *et al*. (2019) Contribution of proteasome-catalyzed peptide cis-splicing to viral targeting by CD8+ T cells in HIV-1 infection. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24748–24759
42. Platteel, A. C. M., Mishto, M., Textoris-Taube, K., Keller, C., Liepe, J., Busch, D. H., Kloetzel, P. M., and Sijts, A. J. A. M. (2016) CD8(+) T cells of Listeria monocytogenes-infected mice recognize both linear and spliced proteasome products. *Eur. J. Immunol.* **46**, 1109–1118
43. Mishto, M., and Liepe, J. (2017) Post-translational peptide splicing and T cell responses. *Trends Immunol.* **38**, 904–915
44. Mishto, M. (2020) What we see, what we do not see, and what we do not want to see in HLA class I immunopeptidomes. *Proteomics*, e2000112
45. Mylonas, R., Beer, I., Iseli, C., Chong, C., Pak, H.-S., Gfeller, D., Coukos, G., Xenarios, I., Müller, M., and Bassani-Sternberg, M. (2018) Estimating the contribution of proteasomal spliced peptides to the HLA-I ligandome. *Mol. Cell. Proteomics* **17**, 2347–2357
46. Lichti, C. F. (2021) Identification of spliced peptides in pancreatic islets uncovers errors leading to false assignments. *Proteomics* **21**, e2000176
47. Rolfs, Z., Müller, M., Shortreed, M. R., Smith, L. M., and Bassani-Sternberg, M. (2019) Comment on "A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands". *Sci. Immunol.* **4**, eaaw1622
48. Faridi, P., Li, C., Ramarathinam, S. H., Illing, P. T., Mifsud, N. A., Ayala, R., Song, J., Gearing, L. J., Croft, N. P., and Purcell, A. W. (2019) Response to comment on "A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands". *Sci. Immunol.* **4**, eaaw8457
49. Willimsky, G., Beier, C., Immisch, L., Papafotiou, G., Scheuplein, V., Goede, A., Holzhütter, H.-G., Blankenstein, T., and Kloetzel, P. M. (2021) In vitro proteasome processing of neo-splicetopes does not predict their presentation in vivo. *Elife* **10**, e62019
50. Aebersold, R., and Mann, M. (2016) Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355
51. Caron, E., Kowalewski, D. J., Chiek Koh, C., Sturm, T., Schuster, H., and Aebersold, R. (2015) Analysis of major histocompatibility complex (MHC) immunopeptidomes using mass spectrometry. *Mol. Cell. Proteomics* **14**, 3105–3117
52. Bassani-Sternberg, M., and Coukos, G. (2016) Mass spectrometry-based antigen discovery for cancer immunotherapy. *Curr. Opin. Immunol.* **41**, 9–17
53. Schumacher, F.-R., Delamarre, L., Jhunjhunwala, S., Modrusan, Z., Phung, Q. T., Elias, J. E., and Lill, J. R. (2017) Building proteomic tool boxes to monitor MHC class I and class II peptides. *Proteomics* **17**, 1600061
54. Creech, A. L., Ting, Y. S., Goulding, S. P., Sauld, J. F. K., Barthelme, D., Rooney, M. S., Addona, T. A., and Abelin, J. G. (2018) The role of mass spectrometry and proteogenomics in the advancement of HLA epitope prediction. *Proteomics* **18**, e1700259
55. Bassani-Sternberg, M. (2018) Mass spectrometry based immunopeptidomics for the discovery of cancer neoantigens. *Methods Mol. Biol.* **1719**, 209–221
56. Kote, S., Pirog, A., Bedran, G., Alfaro, J., and Dapic, I. (2020) Mass spectrometry-based identification of MHC-associated peptides. *Cancers (Basel)* **12**, 535
57. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214
58. Jeong, K., Kim, S., and Bandeira, N. (2012) False discovery rates in spectral identification. *BMC Bioinformatics* **13 Suppl 1**, S2
59. Savitski, M. M., Wilhelm, M., Hahne, H., Kuster, B., and Bantscheff, M. (2015) A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell. Proteomics* **14**, 2394–2404
60. Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003) PEAKS: Powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 2337–2342
61. Berkers, C. R., de Jong, A., Schuurman, K. G., Linnemann, C., Meiring, H. D., Janssen, L., Neefjes, J. J., Schumacher, T. N. M., Rodenko, B., and Ovaa, H. (2015) Definition of proteasomal peptide splicing rules for high-efficiency spliced peptide presentation by MHC class I molecules. *J immuno* **195**, 4085–4895
62. Paes, W., Leonov, G., Partridge, T., Nicastri, A., Ternette, N., and Borrow, P. (2020) Elucidation of the signatures of proteasome-catalyzed peptide splicing. *Front. Immunol.* **11**, 563800
63. Kuckelkorn, U., Stübler, S., Textoris-Taube, K., Kilian, C., Niewienda, A., Henklein, P., Janek, K., Stumpf, M. P. H., Mishto, M., and Liepe, J. (2019) Proteolytic dynamics of human 20S thymoproteasome. *J. Biol. Chem.* **294**, 7740–7754
64. Erhard, F., Dölken, L., Schilling, B., and Schlosser, A. (2020) Identification of the cryptic HLA-I immunopeptidome. *Cancer Immunol. Res.* **8**, 1018–1026
65. Budenholzer, L., Cheng, C. L., Li, Y., and Hochstrasser, M. (2017) Proteasome structure and assembly. *J. Mol. Biol.* **429**, 3500–3524
66. Bard, J. A. M., Goodall, E. A., Greene, E. R., Jonsson, E., Dong, K. C., and Martin, A. (2018) Structure and function of the 26S proteasome. *Annu. Rev. Biochem.* **87**, 697–724

67. Collins, G. A., and Goldberg, A. L. (2017) The logic of the 26S proteasome. *Cell* **169**, 792–806
68. Faridi, P., Purcell, A. W., and Croft, N. P. (2018) Immunopeptidomics we need a sniper instead of a shotgun. *Proteomics* **18**, e1700464
69. Deming, D. (2016) Do extraordinary claims require extraordinary evidence? *Philosophia* **44**, 1319–1331
70. Blatnik, R., Mohan, N., Bonsack, M., Falkenby, L. G., Hoppe, S., Josef, K., Steinbach, A., Becker, S., Nadler, W. M., Rucevic, M., Larsen, M. R., Salek, M., and Riemer, A. B. (2018) A targeted LC-MS strategy for low-abundant HLA class-I-presented peptide detection identifies novel human papillomavirus T-cell epitopes. *Proteomics* **18**, e1700390
71. Mishto, M., Mansurkhodzhaev, A., Ying, G., Bitra, A., Cordfunke, R. A., Henze, S., Paul, D., Sidney, J., Urlaub, H., Neefjes, J., Sette, A., Zajonc, D. M., and Liepe, J. (2019) An in silico-in vitro pipeline identifying an HLA-A*02:01+ KRAS G12V+ spliced epitope candidate for a broad tumor-immune response in cancer patients. *Front. Immunol.* **10**, 2572
72. Gallien, S., Bourmaud, A., Kim, S. Y., and Domon, B. (2014) Technical considerations for large-scale parallel reaction monitoring analysis. *J. Proteomics* **100**, 147–159
73. Shi, T., Song, E., Nie, S., Rodland, K. D., Liu, T., Qian, W. J., and Smith, R. D. (2016) Advances in targeted proteomics and applications to biomedical research. *Proteomics* **16**, 2160–2182
74. Hogan, K. T., Sutton, J. N., Chu, K. U., Busby, J. A. C., Shabanowitz, J., Hunt, D. F., and Slingluff, C. L. (2005) Use of selected reaction monitoring mass spectrometry for the detection of specific MHC class I peptide antigens on A3 supertype family members. *Cancer Immunol. Immunother.* **54**, 359–371
75. Tan, C. T., Croft, N. P., Dudek, N. L., Williamson, N. A., and Purcell, A. W. (2011) Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. *Proteomics* **11**, 2336–2340
76. Bijen, H. M., Hassan, C., Kester, M. G. D., Janssen, G. M. C., Hombrink, P., de Ru, A. H., Drijfhout, J. W., Meiring, H. D., de Jong, A. P., Falkenburg, J. H. F., Jimenez, C. R., Heemskerk, M. H. M., and van Veelen, P. A. (2018) Specific T cell responses against minor histocompatibility antigens cannot generally be explained by absence of their allelic counterparts on the cell surface. *Proteomics* **18**, e1700250
77. Croft, N. P., Purcell, A. W., and Tscharke, D. C. (2015) Quantifying epitope presentation using mass spectrometry. *Mol. Immunol.* **68**, 77–80
78. Starck, S. R., and Shastri, N. (2016) Nowhere to hide: Unconventional translation yields cryptic peptides for immune surveillance. *Immunol. Rev.* **272**, 8–16
79. Laumont, C. M., Daouda, T., Laverdure, J.-P., Bonneil, É., Caron-Lizotte, O., Hardy, M.-P., Granados, D. P., Durette, C., Lemieux, S., Thibault, P., and Perreault, C. (2016) Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat. Commun.* **7**, 10238