



BIROn - Birkbeck Institutional Research Online

Barbosa, C.R.R. and BARTON, Justin and Shepherd, Adrian and Mishto, M. (2021) Mechanistic diversity in MHC class I antigen recognition. *Biochemical Journal* 478 (24), pp. 4187-4202. ISSN 0264-6021.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/46845/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

Review Article

Mechanistic diversity in MHC class I antigen recognition

Camila R. R. Barbosa^{1,2}, Justin Barton³, Adrian J. Shepherd³ and  Michele Mishto^{1,2}

¹Centre for Inflammation Biology and Cancer Immunology (CIBC) & Peter Gorer Department of Immunobiology, King's College London, SE1 1UL London, U.K.; ²Francis Crick Institute, NW1 1AT London, U.K.; ³Department of Biological Sciences and Institute of Structural and Molecular Biology, Birkbeck, University of London, WC1E 7HX London, U.K.

Correspondence: Michele Mishto (michele.mishto@crick.ac.uk) or Camila Rodrigues Barbosa (camila.rodrigues_barbosa@kcl.ac.uk)



Throughout its evolution, the human immune system has developed a plethora of strategies to diversify the antigenic peptide sequences that can be targeted by the CD8⁺ T cell response against pathogens and aberrations of self. Here we provide a general overview of the mechanisms that lead to the diversity of antigens presented by MHC class I complexes and their recognition by CD8⁺ T cells, together with a more detailed analysis of recent progress in two important areas that are highly controversial: the prevalence and immunological relevance of unconventional antigen peptides; and cross-recognition of antigenic peptides by the T cell receptors of CD8⁺ T cells.

Immune response in cancer and infectious diseases

The immune response to infectious diseases or cancer shares similar systematic properties: the immune system can detect ‘non-self’ antigens that occur in both infected and tumour cells; it encompasses effector functions to specifically target and destroy the infected or tumour cells while protecting the host; it develops immunological memory via the adaptive immune responses that facilitates subsequent rapid defence mechanisms [1].

There are two different categories of immune response that afford protection and maintaining the host's normal state of homeostatic state: innate and adaptive. Immediate and generic immune responses are classified as innate due to their fast-acting nonspecific response against foreign antigens such as allergenic antigens, or non-self proteins and molecules [2]. The key players in cell-mediated innate immune responses are phagocytes and natural killer (NK) cells. These phagocytes (neutrophils, monocytes, and macrophages) facilitate immediate host protection by engulfing cells that express non-self-antigens or altered self-antigens and killing them with lysosomal enzymes [3].

In common with innate immunity, the adaptive immune response involves several components that can eliminate both pathogen-infected cells and tumour cells. Adaptive immunity can target antigens in infected and tumour cells by exploiting the effector functions of antibodies, T cells, B cells, and professional antigen-presenting cells (APCs) [2]. In general, endogenous antigens derived from intracellular bacteria, viruses and tumour cells are presented by Major Histocompatibility Complex class I (MHC-I) molecules at the cell surface and recognised by the T cell receptor (TCR) $\alpha\beta$ heterodimers expressed at the cell surface of CD8⁺ T cells. Most cells express MHC-I molecules, and the reduced or null expression of MHC-I complexes at the cell surface can be patrolled by NK cells. Conversely, MHC-II molecules usually present exogenous peptides to TCR $\alpha\beta$ of CD4⁺ T cells. MHC-II expression is usually limited to professional APCs [4].

CD8⁺ T cell-mediated immunity and MHC-I antigen processing and presentation (APP) pathway

To participate in the adaptive immune response, a naive CD8⁺ T cell must first encounter an antigen, engage with professional APCs, and then be induced to differentiate into effector cells capable of removing the cells presenting the antigens, such as infected or tumour cells. In addition to providing

Received: 17 September 2021
 Revised: 16 November 2021
 Accepted: 18 November 2021

Version of Record published:
 23 December 2021

effector T cells, a primary CD8⁺ T cell response generates immunological memory, giving protection from subsequent challenge by the same pathogen. This latter process is exploited in prophylactic vaccine development for infectious diseases [5] and therapeutic cancer vaccines targeting specific tumour-associated antigens [6,7].

CD8⁺ Cytotoxic T Lymphocytes (CTL) are a subset of T cells that have a crucial role in killing host cells infected with viruses, intracellular bacteria, and protozoans such as malaria [8–10]. In the anti-tumour immune response, CTLs such as CD8⁺ tumour-infiltrating lymphocytes (TILs) may make a pivotal contribution to the killing of MHC-I expressing tumour cells [11]. Although antibodies can recognise linear epitopes, they mainly recognise conformational epitopes exposed on the surfaces of proteins with tertiary structure. In contrast, CD8⁺ T cells recognise linear peptide fragments derived from the antigen, which are presented in an extended conformation. These antigenic peptides are presented by MHC-I molecules, and are preferentially 8–15 amino acid residues in length [12]. These peptides are used to discriminate non-self from self antigens [13], although there is a partial overlap between the pool of peptides derived from self- and non-self-antigens. We named those peptides that can be derived from both pathogens and from self as self/non-self *zwitter* antigenic peptides. *Zwitter* is the German word for ‘hybrid’ and ‘hermaphrodites’, originating from *zwi-*, meaning ‘duplex’ [14–16]. The real frequency of these peptides still has to be clarified, although various groups have made *in silico* estimates of self/non-self *zwitter* antigenic peptide frequency, taking into account the origins of unconventional antigenic peptide origins [15–23]. It has been hypothesised that they are involved in creating ‘holes’ in the T cell repertoire, involving mechanisms of central and peripheral tolerance, and in triggering autoimmune responses upon pathogen infection (see also [15,16] for details).

The generation and loading of self and non-self peptides onto MHC-I molecules is carried out by the antigen processing and presentation (APP) machinery, and involves multiple steps: (a) peptide generation by proteasomes and other proteases; (b) peptide trimming by aminopeptidases; (c) peptide transport into the endoplasmic reticulum (ER); (d) assembly of the MHC-I-peptide loading complex (PLC) in the ER; and (e) MHC-I-peptide presentation at the cell surface, with the set of presented peptides known collectively MHC-I immunopeptidome [24] (Figure 1).

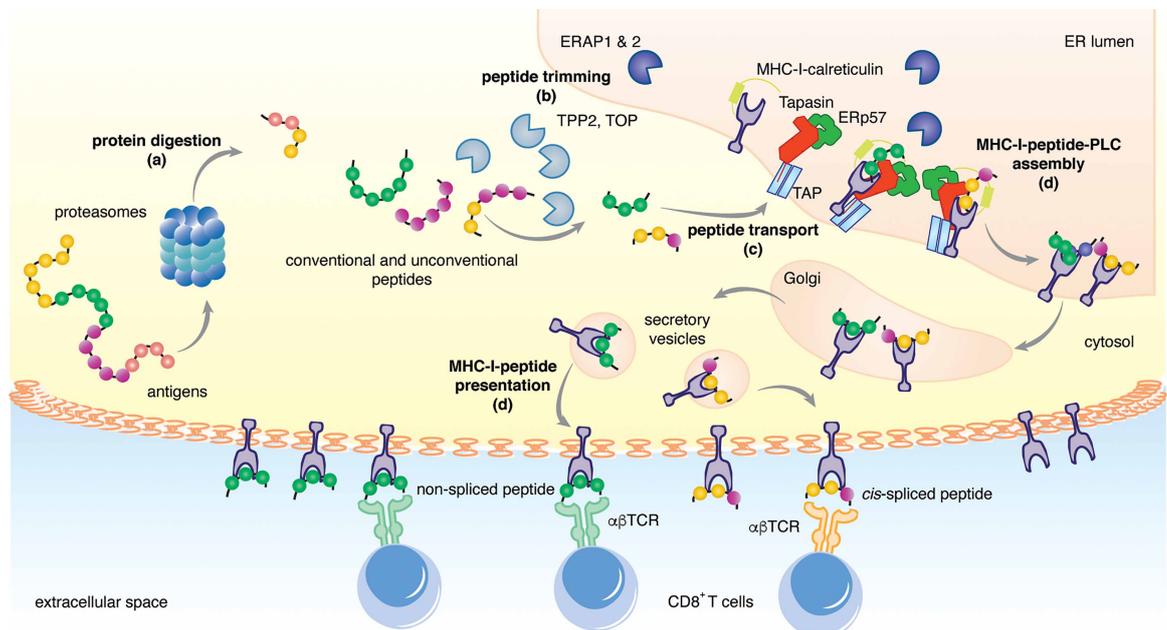


Figure 1. Key steps in the diversity of peptides presented via the MHC-I antigen processing and presentation (APP) pathway.

The MHC-I APP pathway consists of a sequence of steps: (a) peptide generation mostly by proteasomes; (b) peptide trimming by aminopeptidases in the cytosol (e.g. TPP2) and in the ER (e.g. ERAP1); (c) peptide transport into the ER, mostly via TAP; (d) assembly of the MHC-I-peptide-PLC in the ER; (e) MHC-I-peptide presentation at the cell surface. All these steps modulate MHC-I immunopeptidome diversity.

Most of the APP steps impinge upon the potential variety of peptide sequences in MHC-I immunopeptidomes by selecting which peptides are successfully presented. All molecules involved in MHC-I APP pathways can occur in different isoforms with different peptide sequence preference. Variation of the metabolic status of a cell, epigenetic regulation of gene expression and other cellular and molecular events can strongly affect the function and expression of all these molecules, thereby leading to a large diversity of MHC-I immunopeptidomes derived from cells of the same human being. Most of the molecules that are involved in the MHC-I APP pathway — e.g. proteasomes, ER aminopeptidases (ERAPs) and MHC-I — can also carry different polymorphisms or mutations that may strongly influence their activity [25–29]. In addition, some of the MHC-I APP steps can modify the original peptide sequence by ligating non-contiguous peptide fragments (peptide splicing) or modifying single amino acids through post-translational modifications (PTMs; e.g. oxidation, glycosylation, etc.), thereby generating antigenic peptides capable of being recognised by CTLs that cannot recognise the ‘parent’ sequence(s) from which that peptide was derived (Figure 1).

The starting point for the MHC-I APP pathway and the generation of MHC-I immunopeptidome diversity is the availability and selection of protein and polypeptide substrates. They can vary from cell to cell and over time within the same cell because of variations in cell metabolism, in the cell cycle, and in the activity of the ubiquitin–proteasome system (UPS), which is the main proteolytic pathway of cytoplasmatic proteins. In eukaryotic cells, the complete inhibition of proteasome activity causes cell death through necrosis and apoptosis. The multiple subunit 20S proteasomes are a pool of proteasome isoforms that are all associated with a common structure, characterised by two external α rings and two internal β rings. Each ring contains seven subunits. Each β ring contains three catalytic subunits, characterised by a Thr₁ as active site [30] (Figure 2A). 20S standard proteasomes contain catalytic subunits β 1, β 2 and β 5, whereas immunoproteasomes contain β 1i/LMP2, β 2i/MECL1 and β 5i/LMP7 subunits, and thymoproteasomes contain β 1i, β 2i and β 5t subunits (Figure 2B). Immunoproteasomes are expressed in immune cells such as mature dendritic cells (DCs), medullary thymic epithelial cells (mTECs) and in cells stimulated by IFN- γ and other inflammatory stimuli. Thymoproteasomes are almost exclusively expressed in cortical TECs (cTECs). There are also intermediate-type

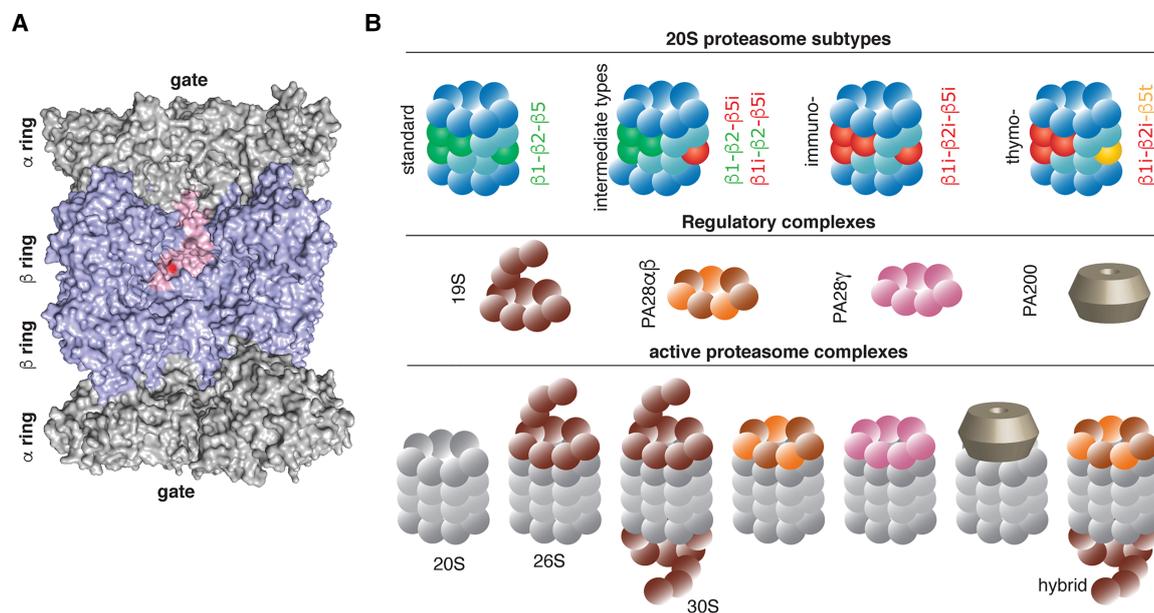


Figure 2. Proteasome isoforms.

(A) The human 20S proteasome is shown based on the structure generated by Schrader et al. [30]. The chains B, C, H, I, J, Q, R, S, Y and Z are hidden from the structure to show the inner proteasome cavities with the central chamber and its two antechambers. The α and β subunits are coloured in grey and blue, respectively. The β 2 subunit is shown in pink with its active site Thr₁ in red, as an example of a catalytic subunit. (B) Proteasome isoforms. 20S proteasomes can be present in many isoforms, which vary on the base of their catalytic subunits. 20S proteasomes can binds to several regulatory complexes, which change both the conformation and the activity of the 20S proteasome core.

20S proteasomes, which carry a different combination of standard-, immuno- and thymo-subunits [31]. Of note, cells generally contain a mixture of different 20S proteasome isoforms.

The effects of changes in the catalytic subunit composition of proteasome isoforms have been studied using a wide variety of methods. Modifications to the catalytic subunit composition alter the catalytic pocket and thus the substrate sequence preferences, and also induce subtle changes in proteasome conformation that affect its proteolytic dynamics [32,33]. All 20S proteasome isoforms can cleave after any substrate amino acid residue, but with very different frequencies, which greatly impacts the degradation rate of specific proteins and the amount of antigenic peptide produced [34–38]. These differences are key factors underpinning MHC-I immunopeptide diversity, as proteasome isoforms can affect both cell metabolism — and thus what proteins are processed by UPS — and directly the antigenic peptides produced [39–45].

Although a proteolytic active role has been ascribed to 20S proteasomes in isolation under various conditions [46,47], this protease is often associated with regulatory particles such as 19S, PA28 $\alpha\beta$, PA28 γ and PA200, which affect proteasome conformation and activity [31,48–50] (Figure 2B). In 26S proteasomes, formed from the complex of 20S and the 19S regulator, the latter activates the 20S core, and binds to and unfolds ubiquitylated substrates. Proteins are poly-ubiquitylated through a cascade of E1, E2 and E3 enzymes, which activates, conjugates and transfers multiple ubiquitin moieties to protein substrates for degradation [51]. The other well-studied regulatory complex, PA28 $\alpha\beta$, binds both standard- and immuno-proteasomes, although it has been hypothesised that it is more frequently associated with immunoproteasomes. The PA28 $\alpha\beta$ -20S proteasome complexes can degrade non-polyubiquitinated proteins [52,53].

Which regulatory complexes are bound to 20S proteasomes affects the features and sequences of antigenic peptides that are produced by the MHC-I APP pathway [54–58].

It is also worth noting that proteasomes are not the only proteases that generate peptides presented by MHC-I molecules, and some scientists have hypothesised that proteasomes make a less significant contribution to MHC-I APP than is generally supposed [59].

The second step of the MHC-I APP pathway that leads to diversity in the MHC-I immunopeptides and the presentation of epitopes is the trimming of peptides by amino peptidases. This can occur in the cytosol — e.g. by tripeptidyl peptidase 2 (TPP2) — as well as in the ER by ERAP1 and ERAP2. Aminopeptidases can shorten peptides, thereby generating the minimal epitopes that can be recognised by CD8⁺ T cells, as well as destroying potential antigenic peptides [60–62]. TPP2 and neurolysin can cleave large peptide fragments. Smaller peptides of 8–15 amino acids are handled by thymet oligopeptidase (TOP), and even smaller ones by other peptidases [63]. Binding of peptides longer than 8–9 amino acids, but not shorter ones, triggers a conformational change in ERAP1 that activates its hydrolysis [64]. ERAP1/2 polymorphisms, such as the ERAP1 Lys/Arg528 polymorphism, can affect MHC-I immunopeptides, the diversity of the response against infections and the onset of autoimmune responses such as Ankylosing Spondylitis [29,65–69].

The third step of the MHC-I APP pathway that contributes to diversity in the MHC-I immunopeptides involves the transportation of peptides into the ER. Peptides have a half-life of 6–10 s in the cytosol of living cells, therefore the intracellular trafficking of peptides from the cytosol into the ER is responsible for the selection of peptides matching in length and sequence specificity to the respective MHC-I molecule [63]. The transporter associated with antigen processing (TAP) heterocomplex — formed by TAP1 and TAP2 — is a key molecule in this process [70]. Peptides with a length of 8–16 amino acids are preferentially transported by TAP, although longer peptides can also be transported by TAP, but with less efficiency [71]. The TAP complex has peptide sequence preferences, which may reduce the transportation of ‘optimal’ epitopes in favour of their N-extended precursors [72,73]. The TAP-independent transport of peptides into the ER has also been described [62,63,74]. TAP forms a molecular complex with other proteins of the peptide loading complex (PLC) such as the oxidoreductase Erp57, the MHC-I heterodimer, and the chaperones tapasin and calreticulin [75]. The latter hold the empty MHC-I complexes in a peptide-receptive state, and tapasin promotes MHC-I-peptide binding with a slow off rate, thereby helping to shape MHC-I immunopeptides [64].

The last step of the MHC-I APP pathway that affects MHC-I immunopeptide diversity, which we consider in this review, is the binding of peptides to the complex that presents antigenic peptides to the TCRs of CD8⁺ T cells, i.e. the MHC-I complex. The binding affinity and stability of the complex formed between the MHC-I complex and an antigenic peptide depends on two factors. First, the peptide that binds to MHC-I molecules forms a hydrogen bonding network that is conserved in the N- and C- terminal regions of the peptide and the corresponding heavy chain residues of the MHC-I, namely those associated with pockets A and F [76]. The interaction between the peptide’s N- and C-termini with the MHC-I groove contributes more

to the binding energy of the MHC-I-peptide complexes than interactions involving the peptide side chains [77–79]. This is regarded as the main limiting factor in terms of the length of the peptides presented by MHC-I molecules, which is a predominantly in the range 9 to 12 amino acids in humans and 8 to 10 amino acids in mice, although longer peptides can also bind to MHC-I complexes with lower frequency [80–87]. Second, the MHC-I binding groove has two (or less commonly three) pockets that show a preference for one to five (more often one or two) of the 20 possible amino acid side chains [88]. One of these pockets always accommodates the C-terminus of the peptide. The residues lodged by the other pocket vary, depending on the shape, but are almost always the second, third or fifth N-terminal residue of the peptide [89]. In human beings, MHC-I heavy chains are encoded by three genes (*HLA-A*, *-B* and *-C*) located in chromosome 6. These genes are the most polymorphic genes in the human genome. The allelic variation mainly affects the nature and composition of the peptide-binding groove and thus modulate MHC-I immunopeptidomes [76].

Conventional and unconventional antigenic peptides

One of the mechanisms underpinning MHC-I immunopeptidome diversity is the generation of antigenic peptides that are not derived straightforwardly via the translation of canonical exon sequences. These unconventional peptides were substantially overlooked until the beginning of the millennium and their discovery is gaining momentum through the continuous development of novel methods for their identification. A sizable portion of these peptides is represented by peptide sequences that derive from putative non-coding regions of the genomes, such as 5'- and 3'-UTRs, introns, intergenic and endogenous retroviral regions [90–96]. We also have peptide sequences that derived from non-canonical open reading frames (ORFs), frameshifted canonical genes, and alternative RNA splicing [97–99]. Other sources of unconventional antigenic peptides are protein fragments that contain single amino acid polymorphisms or somatic mutations. The former are examples of the minor histocompatibility antigens and play a role in organ transplantation. The latter, which are known as neoepitopes when they trigger a CTL response, are particularly relevant in cancer and are proving to be attractive targets for immunotherapies [100–105]. All these are unconventional peptides that can be produced by proteasomes via peptide hydrolysis and do not require any PTM. In addition, peptides may either acquire PTMs — e.g. via phosphorylation [106–109] — or the original antigen sequences may be reshuffled by the ligation of non-contiguous peptide fragments. This latter process, called peptide splicing, has been described in the context of both MHC-I and -II pathways. In the former, the process is mainly catalysed by proteasomes, whereas in the latter, the role of lysosomal proteases such as cathepsin L has been hypothesised and can explain the production of hybrid insulin peptides [110–112]. Proteasome-catalysed peptide splicing (PCPS) can occur by combining two fragments of the same protein (*cis*-PCPS) or two distinct protein molecules (*trans*-PCPS; see Figure 3A). The main biochemical mechanism of PCPS is transpeptidation, as proposed by Vigneron et al. [113], in the first study describing PCPS in the context of tumour. During transpeptidation, a peptide is cleaved. The N-terminal fragment forms an acyl-enzyme intermediate with Thr₁ of a proteasome catalytic subunit. This intermediate can interact with water and then be released as canonical non-spliced peptide by peptide hydrolysis. Alternatively, the acyl-enzyme intermediate can interact with another peptide fragment and form a spliced peptide through PCPS (Figure 3B). PCPS via transpeptidation has been demonstrated using various biochemical and cellular biology strategies [114–117]. A single example of condensation reaction has also been described as an alternative PCPS mechanism [118].

First discovered by Van den Eynde et al. [113,119], proteasome-generated spliced peptides have since proved controversial. In the last decade, one common opinion (to quote a reviewer of a grant application rejected a decade ago) was that peptide splicing is 'a rare event and a curiosity of nature', although this opinion was not supported by experimental evidence. The subsequent identification, in MHC-I immunopeptidomes, of *cis*-spliced peptides by Mishto et al. [81], and *trans*-spliced peptides by Faridi, Croft, Purcell and colleagues [80] has served to heighten the controversy. Both of these studies have been intensively scrutinised by other groups working on immunopeptidomics, and the ensuing disagreements have left many in the scientific community uncertain about the true frequency of spliced peptides in MHC-I immunopeptidomes [120–123]. This controversy spanned the fields of mass spectrometry and immunopeptidomics — perhaps nurtured by the same general scepticism that welcomed the pioneering research on PCPS — and matured in the hypothesis that PCPS 'is, at most, an extremely rare event and likely does not happen at all' [120]. As discussed in a recent commentary [124], this hypothesis clashes with a good deal of experimental evidence demonstrating that proteasomes can efficiently catalyse both peptide hydrolysis and splicing. For instance, *cis*-spliced epitopes can

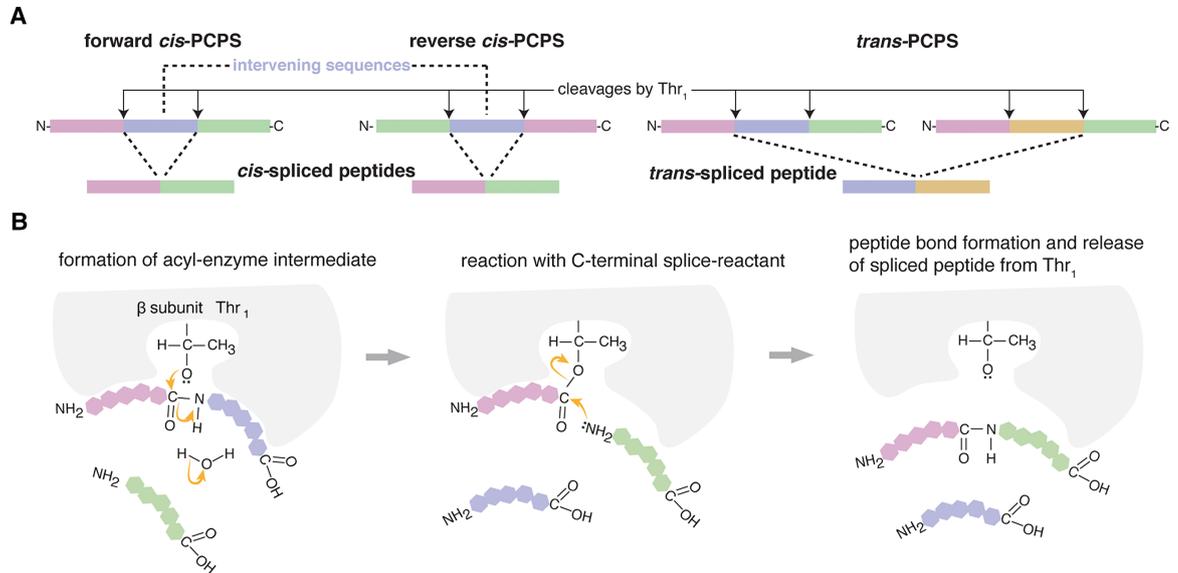


Figure 3. Proteasome-catalysed peptide splicing (PCPS).

(A) Spliced peptides can be formed by: (i) *cis*-PCPS, i.e. when the two splice-reactants derive from the same protein. The ligation of the splice-reactants can occur in normal order, i.e. following the orientation from N- to C-terminus of the parental protein (forward *cis*-PCPS), or in the reverse order (reverse *cis*-PCPS); (ii) *trans*-PCPS, when the two splice-reactants originate from two distinct proteins. (B) Transpeptidation. Proteasome's catalytic Thr₁ breaks the peptide bond of the residue (P1) of the protein - thereby forming an acyl-enzyme intermediate with the N-terminal splice-reactant, and releasing the C-terminal peptide fragment. The acyl-enzyme intermediate can then interact with another peptide fragment (the C-terminal splice-reactant) and form a new peptide bond between the P1 residue of the N-terminal splice-reactant and the residue P1' of the C-terminal splice-reactant.

trigger a specific CD8⁺ T cell response against tumour-associated [40,113,117–119,125,126], Type1 Diabetes-associated [99] and *Listeria monocytogenes*-derived [127] antigens. *Cis*-spliced epitopes can also stimulate CTL cross-recognition during infections through TCR degeneracy (see below) [128,129]. Successful application of immunotherapies targeting *cis*-spliced epitopes has been described in the treatment of a metastatic melanoma patient using an autologous tumour-infiltrating lymphocyte clone [125,130], in nonobese diabetic/severe combined immunodeficient mice leading to engraftment inhibition of human acute myelogenous leukaemia cells [117,131], as well as in a mouse model of glioblastoma via peptide vaccination [132]. Although *cis*-PCPS might theoretically generate a very large number of peptide sequences that, preliminary estimates suggest it has a limited role in T cell tolerance and viral-driven autoimmunity [15,16]. The biochemical aspect of PCPS can be studied in relatively simple experiments, wherein synthetic polypeptides are processed by purified proteasome isoforms, and their products are detected via mass spectrometry. The largest database of non-spliced and spliced peptides produced by proteasomes in these *in vitro* assays has been published by Specht and colleagues [133], and contains almost 15 000 unique peptides. In this database, *cis*-spliced and *trans*-spliced peptides represent each one third of the peptide product variety [133], although this experimental set up might artificially favour PCPS [124]. These results seem to confirm that *cis*-spliced peptides may make a substantial contribution to MHC-I diversity, as suggested by *in cellula* experiments [40,80,132,134]. However, unbiased evaluation of all unconventional antigenic peptides in relation to conventional peptides still needs to be carried out, and may ultimately resolve current disputes about this topic.

T cell receptor (TCR)

Whereas the MHC-I APP pathway can be viewed as the cornerstone of the MHC-I immunopeptidomes diversity, TCRs expressed on the surface of CD8⁺ T cells arguably play a comparable role with respect to the diversity of antigen recognition. Those TCRs that recognise MHC-I-peptide complexes are heterodimeric proteins, composed of an α chain and a β chain (Figure 4A). Other types of TCRs with different function and

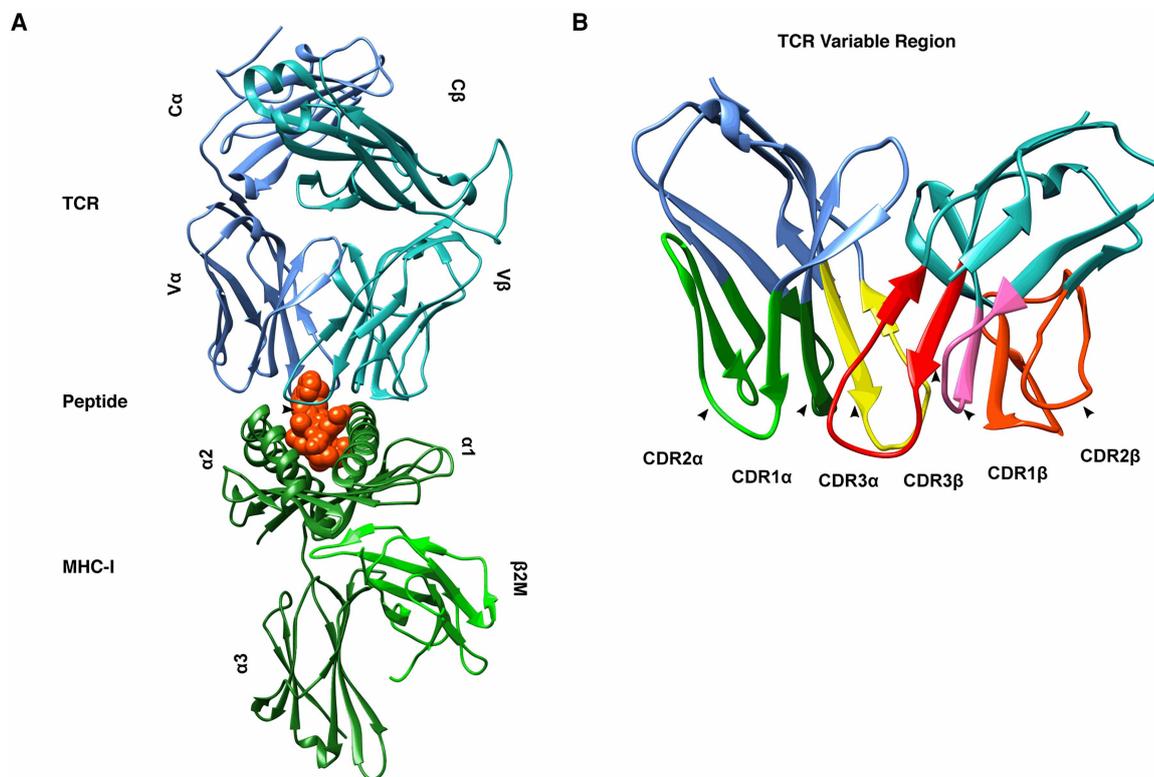


Figure 4. Structure of a CD8⁺ TCR.

(A) A CD8⁺ TCR bound to a cognate MHC-I-peptide complex. The TCR is shown in blue (α chain) and cyan (β chain) with the constant (C) and variable (V) regions annotated. The peptide is shown in red and the MHC-I in green with the α -helices ($\alpha 1$ through $\alpha 3$) and $\beta 2$ microglobulin ($\beta 2M$) annotated. (B) Looking more closely at the variable region of the TCR reveals the six complementarity determining regions (CDRs) at the interface with the MHC-I-peptide complex. Note the central position of the CDR3s on both the α and β chains; these are the CDRs most associated with peptide contacts.

specificities have also been described [135]. Each α and β chain have three protruding loops that contact the MHC-I-peptide complex (Figure 4B). Owing to their role in binding the MHC-I-peptide complex, these six loops are referred to as complementarity determining regions (CDRs). Of these six CDRs, analyses of solved crystal structures suggest that the third loop on the β chain, referred to as CDR3 β , has the greatest impact on the interaction of TCR and MHC-I-peptide complex, and the antigenic peptide in particular [136].

TCR sequences are generated by recombination of variable (V), diversity (D), and joining (J) gene segments during T cell maturation in the thymus. This VDJ recombination process leads to a vast space of possible TCRs, with estimates of 10^{15} distinct possible combinations [137,138]. Of course, not all possible combinations are viable nor are all viable combinations observed with equal frequency. Recombinatorial biases arise where proximal sites in some loci undergo recombination much more frequently than distal sites. This leads to biased and constrained TCR ‘repertoires’ [139]. The set of TCRs within an individual subject is referred to as the TCR repertoire. The TCR repertoire is refined by a selection process that promotes useful TCRs that successfully bind antigens and down-regulates potentially harmful TCRs which bind self-antigens. This selection process is often referred to as ‘thymic selection’ since it takes place in the thymus. The TCR repertoire is thus shaped by the history of pathogens that a host has been exposed to [140].

Given the apparent primacy of CDR3 β in determining the specificity of TCR binding with respect to a given antigenic peptide, many popular algorithms have been developed which cluster TCRs based on the sequence similarity of the CDR3 β region [141–144]. These algorithms purport to identify clusters of TCRs with similar antigen specificity. More recently, there has been evidence suggesting that the CDRs most actively involved in MHC-I-peptide binding can vary by MHC-I-peptide complexes. This is likely due to the complex conformational changes in both the TCR and the MHC-I-peptide complex that can facilitate different binding

arrangements [145,146]. This conformational plasticity is one mechanism by which a single TCR can potentially recognise many different antigen-derived peptides.

TCR cross-reactivity

A single TCR can potentially recognise more than a million different peptides [147]. Conversely, a peptide can elicit a response from millions of TCRs [138]. This flexibility is referred to as ‘cross-reactivity’, or ‘degeneracy’.

The number of peptides which can be presented in MHC-I-peptide complexes is estimated to be greater than 10^{15} [138]. This estimate is derived by calculating all the possible combinatorial permutations of the 20 common amino acids for peptide lengths that can be presented by MHC molecules (8 to 14 residues), then estimating that 1% to 3% of these possible peptides will bind with and be presented by MHC molecules. On the other hand, the number of unique TCRs within an individual’s repertoire is estimated to be in the range of 10^6 [148] to 10^8 [149]. This indicates that an individual’s ability to mount a successful immune response to a broad proportion of the potential pathogen space is partially reliant on cross-reactivity, as it facilitates a broader range of specificity than would otherwise be possible within the physical constraints of the adaptive immune system. In the context of the vast universe of presented peptides, however, even a highly cross-reactive TCR only recognises a tiny fraction. Ishizuka et al. [19] empirically estimated the likelihood of unrelated peptides being recognised by the same TCR as 1 in 30 000.

Cross-reactivity comes at a cost, however. Cross-reactivity can mean that response to a particular antigen has off-target effects, as is often the case in autoimmunity and allergy. Unanticipated cross-reactivity is a major concern, since on-target and off-target toxicity can be extremely damaging, and even fatal. On-target toxicity occurs when introduction of an antigen causes unintentional cytotoxic destruction of healthy tissues. Off-target toxicity occurs in situations such as molecular mimicry where different antigens are presented as MHC-I-peptide complexes with similar key structural and chemical features and bind with the same TCRs, potentially leading to unintended collateral damage [150]. The self-non-self *zwitter* epitopes described above are a particular example of cross-recognition. Indeed, both self-derived and non-self-derived *zwitter* epitopes can be recognised by the same TCR, since they have exactly the same peptide sequence.

A corollary of cross-reactivity is what is called ‘heterologous immunity’. Heterologous immunity occurs when previous infection by one pathogen confers immunity against an unrelated pathogen. It is important to note that heterologous immunity is not always bi-directional. For example, infection with Influenza A in mice confers protection against Vaccinia, but the reverse is not true [151].

As one might expect, instances of cross-reactivity have been observed between peptides with a high degree of sequence similarity. Perhaps less intuitively, instances have also been observed of peptides which bind to the same TCR but share almost no sequence similarity [152].

Cross-reactivity is driven in large part by conformational flexibility in the TCR, peptide, and MHC, leading to a high diversity in the interfaces between the TCR and MHC-I-peptide complexes [153]. These different interfaces can have very different contact residues and conformational geometries. There are five known mechanisms that can lead to these varied interfaces: TCR plasticity, MHC-I-peptide complex plasticity, TCR docking, molecular mimicry, and structural degeneracy [154] (Figure 5).

Perhaps the most broadly discussed of these mechanisms is the conformational plasticity of the TCR, particularly in the CDR loops (Figure 5A). This flexibility allows the TCR to accommodate different MHC-I-peptide complexes without changing the docking orientation. Several studies have observed and quantified just how variable the conformations of the CDR loops can be. Distances between the bound and unbound states typically range from 0.3 Å to 11.4 Å with the CDR3 loop exhibiting the greatest range of motion [155].

In a similar fashion, MHC-I-peptide complexes can exhibit varying degrees of conformational flexibility, which can likewise allow for substantial rearrangement at the time of TCR binding [154] (Figure 5B). At the present time this mechanism has been less studied and is less well-understood, though anecdotal examples are presented in [156,157]. Riley et al. [158] reported a particularly interesting example of unexpected cross-reactivity between a TCR and peptides sharing few physiochemical commonalities. Binding with the TCR induced substantial conformational changes in one MHC-I-peptide complex, including a register shift that resulted in a 3.5 Å root-mean-square deviation (RMSD) movement for all peptide atoms. A ‘register shift’ refers to a shift in the peptide such that residues lie in different positions within the MHC groove, potentially resulting in attendant changes in peptide bulges and side chain orientations.

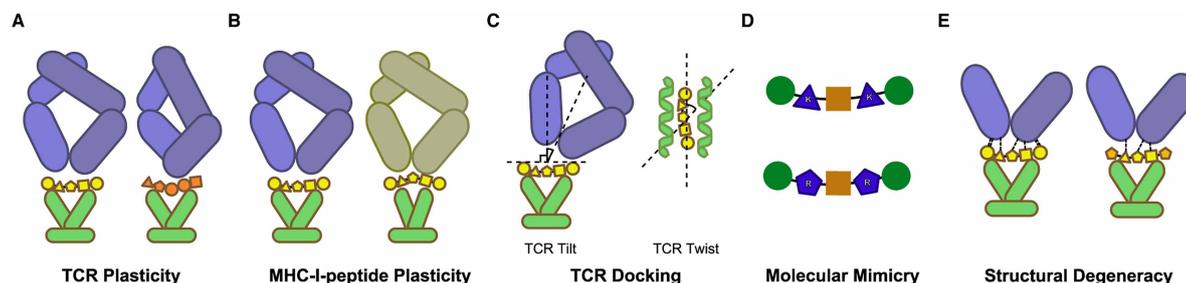


Figure 5. Structural mechanisms of T cell cross-reactivity.

(A) Cross-reactivity can be facilitated by the significant conformational plasticity of the TCR, particularly in the CDR loops. Here a TCR (blue) changes conformation to bind different MHC-I-peptide complexes. (B) Similarly, the MHC-I-peptide complex can undergo conformational changes that promote binding with different TCRs. (C) Variable TCR docking angles can alter the interface and allow for binding a broader range of MHC-I-peptide complexes. TCR docking angles vary substantially in the vertical (TCR tilt) and horizontal (TCR twist) planes. (D) Different peptides can bind with the same TCRs by presenting very similar physicochemical surfaces at the interface. Here peptides are shown substituting arginine (R) residues for lysine (K) residues, which conserve physicochemical features such as positive charge and polarity. (E) TCR-MHC-I-peptide binding is an interplay of multiple factors including shape complementarities, hydrogen bonds, salt bridges, and van der Waals interactions. This confederation of forces includes some redundancy and often binding can be achieved with a subset of the factors. This allows binding to be resilient to substitutions in peptide sequences.

Variable docking orientation is another mechanism by which a TCR can bind to different MHC-I-peptide complexes (Figure 5C). These docking angles can be quite variable in both the vertical (referred to as ‘TCR tilt’) and horizontal (referred to as ‘TCR twist’) planes. To-date incident angles (measuring TCR tilt) have mostly been observed to fall within the range 0° to 30° range (relative to the MHC normal vector) and crossing angles (measuring TCR twist) within the range 22° to 69° range (relative to the MHC groove vector) though outliers have been observed substantially outside of these ranges and the sample size of solved structures remains relatively small [157,159]. The TCR3d database [160] provides known TCR structures including their docking angles.

‘Molecular mimicry’ refers to instances where MHC-I-peptide complexes that are unrelated can share key structural and chemical features, thereby allowing a TCR to recognise both complexes (Figure 5D). It has been demonstrated that even peptides with low sequence and physicochemical similarity can present very similar MHC-I-peptide surfaces at their interface with TCRs. In an extreme case, Zhang et al. [161] presented an example of cross-reactivity between peptides with no sequence overlap. Cross-reactivity driven by molecular mimicry has been successfully predicted using machine learning approaches trained on feature data including surface electrostatic charges and accessible surface area of selected residues in MHC-I-peptide complexes [152,162], at least for a specific MHC-I allele (HLA-A*02:01).

In some cases, TCRs can bind to MHC-I-peptide complexes with only a subset of the binding mechanisms that might otherwise be present, including some cases with no hydrogen bonds or salt bridges and poor shape complementarity (Figure 5E). This degeneracy allows TCR binding to be resilient to substitutions in the contact residues of related peptides.

Note that the mechanisms of TCR cross-reactivity discussed here are laid out for conceptual understanding and are by no means mutually exclusive. Combinations of these mechanisms have been observed and it is likely that cross-reactivity dynamics include a complex interplay of various mechanisms. Note also that peptide recognition is not a binary event, and affinity can vary substantially. At the interface between the TCR and MHC-I-peptide complexes, several residues may participate in binding, influencing the affinity and the specific level of T-cell stimulation that will be triggered by each MHC-I-peptide complex. Due to limitations in current experimental methods, available data may be biased towards higher-affinity interactions as low-affinity interactions can escape detection.

Though binding affinity is an important consideration, it is worth noting that activation is not always correlated with affinity. In fact, TCR-MHC-I-peptide complexes with high binding affinity but no detectable T cell activation occur frequently *in vivo* in humans [163]. Activation is also not strictly a binary event, as the consequence of T cell activation can vary, for example with respect to differences in cytokine production.

Cross-reactivity is a key component in the functional dynamics of adaptive immunity. It has direct implications for immunotherapies, autoimmunity, and heterologous immune response. As immunotherapies grow in popularity, the need to reliably predict cross-reactivity will be key to safe and effective interventions.

Our limited knowledge and future directions

In conclusion, there has been a giant leap in our understanding of the mechanistic diversity of MHC-I antigen presentation and recognition during recent decades. However, the journey is still long, and areas that are still neglected and controversial, such as the presence and features of unconventional epitopes and the drivers of TCR degeneracy, may represent important directions for future research, and contribute to the development of novel immunotherapies against cancer and pathogens.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was financed in part by: (i) Cancer Research UK [C67500/A29686] and National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's as well as St Thomas' NHS Foundation Trust and King's College London and/or the NIHR Clinical Research Facility to M.M.

Acknowledgements

Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualisation, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

Abbreviations

APC, antigen-presenting cells; APP, antigen processing and presentation; CDR, complementarity determining regions; cTECs, cortical thymic epithelial cells; CTL, CD8⁺ Cytotoxic T Lymphocytes; DC, dendritic cells; ER, endoplasmic reticulum; ERAP, ER aminopeptidases; ERp57, endoplasmic reticulum resident protein 57; LMP2, low-molecular mass polypeptide 2; LMP7, low-molecular mass polypeptide 7; MECL-1, multicatalytic endopeptidase complex-like-1; MHC-I, Major Histocompatibility Complex class I; mTECs, medullary thymic epithelial cells; NK, natural killer; PCPS, proteasome-catalysed peptide splicing; PLC, peptide loading complex; PTM, post-translational modifications; RMSD, root-mean-square deviation; TAP, transporter associated with antigen processing; TCR, T cell receptor; TILs, tumour-infiltrating lymphocytes; TOP, thymet oligopeptidase; TPP2, tripeptidyl peptidase 2; UPS, ubiquitin–proteasome system; β 2M, β 2 microglobulin.

References

- 1 Janeway, C.A., Travers, P., Walport, M. and Shlomchik, M.J. (2001) T cell-mediated cytotoxicity. In *Immunobiology: The Immune System in Health and Disease*, Garland Science, New York
- 2 Pandya, P.H., Murray, M.E., Pollok, K.E. and Renbarger, J.L. (2016) The immune system in cancer pathogenesis: potential therapeutic approaches. *J. Immunol. Res.* **2016**, 4273943 <https://doi.org/10.1155/2016/4273943>
- 3 Sun, J.C. and Lanier, L.L. (2009) Natural killer cells remember: an evolutionary bridge between innate and adaptive immunity? *Eur. J. Immunol.* **39**, 2059–2064 <https://doi.org/10.1002/eji.200939435>
- 4 Chen, D.S. and Mellman, I. (2013) Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1–10 <https://doi.org/10.1016/j.immuni.2013.07.012>
- 5 Pollard, A.J. and Bijker, E.M. (2021) A guide to vaccinology: from basic principles to new developments. *Nat. Rev. Immunol.* **21**, 83–100 <https://doi.org/10.1038/s41577-020-00479-7>
- 6 Bilusic, M. and Madan, R.A. (2012) Therapeutic cancer vaccines. *Am. J. Ther.* **19**, e172–e181 <https://doi.org/10.1097/MJT.0b013e3182068cdb>
- 7 Li, L., Goedegebuure, S.P. and Gillanders, W. (2020) Cancer vaccines: shared tumor antigens return to the spotlight. *Signal Transduct. Target Ther.* **5**, 251 <https://doi.org/10.1038/s41392-020-00364-8>
- 8 Bahjat, K.S., Liu, W., Lemmens, E.E., Schoenberger, S.P., Portnoy, D.A., Dubensky, T.W. et al. (2006) Cytosolic entry controls CD8⁺-T-cell potency during bacterial infection. *Infect. Immun.* **74**, 6387–6397 <https://doi.org/10.1128/IAI.01088-06>
- 9 Wiesel, M., Walton, S., Richter, K. and Oxenius, A. (2009) Virus-specific CD8T cells: activation, differentiation and memory formation. *APMIS* **117**, 356–381 <https://doi.org/10.1111/j.1600-0463.2009.02459.x>
- 10 Junqueira, C., Barbosa, C.R.R., Costa, P.A.C., Teixeira-Carvalho, A., Castro, G., Sen Santana, S. et al. (2018) Cytotoxic CD8(+) T cells recognize and kill Plasmodium vivax-infected reticulocytes. *Nat. Med.* **24**, 1330–1336 <https://doi.org/10.1038/s41591-018-0117-4>
- 11 Farhood, B., Najafi, M. and Mortezaee, K. (2019) CD8⁺ cytotoxic t lymphocytes in cancer immunotherapy: a review. *J. Cell. Physiol.* **234**, 8509–8521 <https://doi.org/10.1002/jcp.27782>

- 12 Sanchez-Trincado, J.L., Gomez-Perosanz, M. and Reche, P.A. (2017) Fundamentals and methods for T- and B-cell epitope prediction. *J. Immunol. Res.* **2017**, 2680160 <https://doi.org/10.1155/2017/2680160>
- 13 Klein, L., Kyewski, B., Allen, P.M. and Hogquist, K.A. (2014) Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat. Rev. Immunol.* **14**, 377–391 <https://doi.org/10.1038/nri3667>
- 14 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 <https://doi.org/10.1016/j.coi.2018.04.016>
- 15 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 <https://doi.org/10.3389/fimmu.2021.614276>
- 16 Mishto, M., Mansurkhodzhaev, A., Rodriguez-Calvo, T. and Liepe, J. (2021) Potential mimicry of viral and pancreatic beta cell antigens through non-spliced and cis-spliced zwitter epitope candidates in type 1 diabetes. *Front. Immunol.* **12**, 656451 <https://doi.org/10.3389/fimmu.2021.656451>
- 17 Frankild, S., de Boer, R.J., Lund, O., Nielsen, M. and Kesmir, C. (2008) Amino acid similarity accounts for T cell cross-reactivity and for 'holes' in the T cell repertoire. *PLoS ONE* **3**, e1831 <https://doi.org/10.1371/journal.pone.0001831>
- 18 Calis, J.J., de Boer, R.J. and Kesmir, C. (2012) Degenerate T-cell recognition of peptides on MHC molecules creates large holes in the T-cell repertoire. *PLoS Comput. Biol.* **8**, e1002412 <https://doi.org/10.1371/journal.pcbi.1002412>
- 19 Ishizuka, J., Grebe, K., Shenderov, E., Peters, B., Chen, Q., Peng, Y. et al. (2009) Quantitating T cell cross-reactivity for unrelated peptide antigens. *J. Immunol.* **183**, 4337–4345 <https://doi.org/10.4049/jimmunol.0901607>
- 20 Kanduc, D., Stufano, A., Lucchese, G. and Kusalik, A. (2008) Massive peptide sharing between viral and human proteomes. *Peptides* **29**, 1755–1766 <https://doi.org/10.1016/j.peptides.2008.05.022>
- 21 Kusalik, A., Bickis, M., Lewis, C., Li, Y., Lucchese, G., Marincola, F.M. et al. (2007) Widespread and ample peptide overlapping between HCV and homo sapiens proteomes. *Peptides* **28**, 1260–1267 <https://doi.org/10.1016/j.peptides.2007.04.001>
- 22 Trost, B., Kusalik, A., Lucchese, G. and Kanduc, D. (2010) Bacterial peptides are intensively present throughout the human proteome. *Self Nonself.* **1**, 71–74 <https://doi.org/10.4161/self.1.1.9588>
- 23 Trost, B., Lucchese, G., Stufano, A., Bickis, M., Kusalik, A. and Kanduc, D. (2010) No human protein is exempt from bacterial motifs, not even one. *Self Nonself.* **1**, 328–334 <https://doi.org/10.4161/self.1.4.13315>
- 24 Mishto, M. and Liepe, J. (2017) Post-Translational peptide splicing and T cell responses. *Trends Immunol.* **38**, 904–915 <https://doi.org/10.1016/j.it.2017.07.011>
- 25 Mishto, M., Bellavista, E., Ligorio, C., Textoris-Taube, K., Santoro, A., Giordano, M. et al. (2010) Immunoproteasome LMP2 60HH variant alters MBP epitope generation and reduces the risk to develop multiple sclerosis in Italian female population. *PLoS ONE* **5**, e9287 <https://doi.org/10.1371/journal.pone.0009287>
- 26 Mishto, M., Santoro, A., Bellavista, E., Sessions, R., Textoris-Taube, K., Dal Piaz, F. et al. (2006) A structural model of 20S immunoproteasomes: effect of LMP2 codon 60 polymorphism on expression, activity, intracellular localisation and insight into the regulatory mechanisms. *Biol. Chem.* **387**, 417–429 <https://doi.org/10.1515/BC.2006.056>
- 27 Costantino, F., Talpin, A., Evnouchidou, I., Kadi, A., Leboime, A., Said-Nahal, R. et al. (2015) ERAP1 gene expression is influenced by nonsynonymous polymorphisms associated with predisposition to spondyloarthritis. *Arthritis Rheumatol.* **67**, 1525–1534 <https://doi.org/10.1002/art.39072>
- 28 Weimershaus, M., Evnouchidou, I., Li, L., van Endert, P. and Bouvier, M. (2019) Trimming of MHC class I ligands by ERAP aminopeptidases. *Methods Mol. Biol.* **1988**, 31–43 https://doi.org/10.1007/978-1-4939-9450-2_3
- 29 Wisniewski, A., Sobczynski, M., Pawelczyk, K., Porebska, I., Jasek, M., Wagner, M. et al. (2021) Polymorphisms of antigen-presenting machinery genes in non-small cell lung cancer: different impact on disease risk and clinical parameters in smokers and never-smokers. *Front. Immunol.* **12**, 664474 <https://doi.org/10.3389/fimmu.2021.664474>
- 30 Schrader, J., Henneberg, F., Mata, R.A., Tittmann, K., Schneider, T.R., Stark, H. et al. (2016) The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. *Science* **353**, 594–598 <https://doi.org/10.1126/science.aaf8993>
- 31 Vigneron, N. and Van den Eynde, B.J. (2014) Proteasome subtypes and regulators in the processing of antigenic peptides presented by class I molecules of the major histocompatibility complex. *Biomolecules* **4**, 994–1025 <https://doi.org/10.3390/biom4040994>
- 32 Kuckelkorn, U., Stubler, S., Textoris-Taube, K., Kilian, C., Niewianda, A., Henklein, P. et al. (2019) Proteolytic dynamics of human 20S thymoproteasome. *J. Biol. Chem.* **294**, 7740–7754 <https://doi.org/10.1074/jbc.RA118.007347>
- 33 Liepe, J., Holzthutter, H.G., Bellavista, E., Kloetzel, P.M., Stumpf, M.P. and Mishto, M. (2015) Quantitative time-resolved analysis reveals intricate, differential regulation of standard- and immuno-proteasomes. *eLife* **4**, e07545 <https://doi.org/10.7554/eLife.07545>
- 34 Mishto, M., Liepe, J., Textoris-Taube, K., Keller, C., Henklein, P., Weberuss, M. et al. (2014) Proteasome isoforms exhibit only quantitative differences in cleavage and epitope generation. *Eur. J. Immunol.* **44**, 3508–3521 <https://doi.org/10.1002/eji.201444902>
- 35 Arciniega, M., Beck, P., Lange, O.F., Groll, M. and Huber, R. (2014) Differential global structural changes in the core particle of yeast and mouse proteasome induced by ligand binding. *Proc. Natl Acad. Sci. U.S.A.* **111**, 9479–9484 <https://doi.org/10.1073/pnas.1408018111>
- 36 Guillaume, B., Stroobant, V., Bousquet-Dubouch, M.P., Colau, D., Chapiro, J., Parmentier, N. et al. (2012) Analysis of the processing of seven human tumor antigens by intermediate proteasomes. *J. Immunol.* **189**, 3538–3547 <https://doi.org/10.4049/jimmunol.1103213>
- 37 Guillaume, B., Chapiro, J., Stroobant, V., Colau, D., Van Holle, B., Parvizi, G. et al. (2010) Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proc. Natl Acad. Sci. U.S.A.* **107**, 18599–18604 <https://doi.org/10.1073/pnas.1009778107>
- 38 Textoris-Taube, K., Keller, C., Liepe, J., Henklein, P., Sidney, J., Sette, A. et al. (2015) The T210M substitution in the HLA-a*02:01 gp100 epitope strongly affects overall proteasomal cleavage site usage and antigen processing. *J. Biol. Chem.* **290**, 30417–30428 <https://doi.org/10.1074/jbc.M115.695189>
- 39 Goncalves, G., Mullan, K.A., Duscharla, D., Ayala, R., Croft, N.P., Faridi, P. et al. (2021) IFN γ modulates the immunopeptidome of triple negative breast cancer cells by enhancing and diversifying antigen processing and presentation. *Front. Immunol.* **12**, 645770 <https://doi.org/10.3389/fimmu.2021.645770>
- 40 Faridi, P., Woods, K., Ostrouska, S., Deceneux, C., Aranha, R., Duscharla, D. et al. (2020) Spliced peptides and cytokine-driven changes in the immunopeptidome of melanoma. *Cancer Immunol. Res.* **8**, 1322–1334 <https://doi.org/10.1158/2326-6066.CIR-19-0894>

- 41 Apavaloai, A., Brochu, S., Dong, M., Rouette, A., Hardy, M.P., Villafano, G. et al. (2019) PSMB11 orchestrates the development of CD4 and CD8 thymocytes via regulation of gene expression in cortical thymic epithelial cells. *J. Immunol.* **202**, 966–978 <https://doi.org/10.4049/jimmunol.1801288>
- 42 St-Pierre, C., Morgand, E., Benhamadi, M., Rouette, A., Hardy, M.P., Gaboury, L. et al. (2017) Immunoproteasomes control the homeostasis of medullary thymic epithelial cells by alleviating proteotoxic stress. *Cell Rep.* **21**, 2558–2570 <https://doi.org/10.1016/j.celrep.2017.10.121>
- 43 Kincaid, E.Z., Murata, S., Tanaka, K. and Rock, K.L. (2016) Specialized proteasome subunits have an essential role in the thymic selection of CD8(+) T cells. *Nat. Immunol.* **17**, 938–945 <https://doi.org/10.1038/ni.3480>
- 44 Sijts, A.J., Ruppert, T., Reherrmann, B., Schmidt, M., Koszinowski, U. and Kloetzel, P.M. (2000) Efficient generation of a hepatitis B virus cytotoxic T lymphocyte epitope requires the structural features of immunoproteasomes. *J. Exp. Med.* **191**, 503–514 <https://doi.org/10.1084/jem.191.3.503>
- 45 Sijts, A.J., Standera, S., Toes, R.E., Ruppert, T., Beekman, N.J., van Veelen, P.A. et al. (2000) MHC class I antigen processing of an adenovirus CTL epitope is linked to the levels of immunoproteasomes in infected cells. *J. Immunol.* **164**, 4500–4506 <https://doi.org/10.4049/jimmunol.164.9.4500>
- 46 Ben-Nissan, G. and Sharon, M. (2014) Regulating the 20S proteasome ubiquitin-independent degradation pathway. *Biomolecules* **4**, 862–884 <https://doi.org/10.3390/biom4030862>
- 47 Fabre, B., Lambour, T., Garrigues, L., Ducoux-Petit, M., Amalric, F., Monsarrat, B. et al. (2014) Label-free quantitative proteomics reveals the dynamics of proteasome complexes composition and stoichiometry in a wide range of human cell lines. *J. Proteome Res.* **13**, 3027–3037 <https://doi.org/10.1021/pr500193k>
- 48 Ruschak, A.M. and Kay, L.E. (2012) Proteasome allostery as a population shift between interchanging conformers. *Proc. Natl Acad. Sci. U.S.A.* **109**, E3454–E3462 <https://doi.org/10.1073/pnas.1213640109>
- 49 Dick, T.P., Ruppert, T., Groettrup, M., Kloetzel, P.M., Kuehn, L., Koszinowski, U.H. et al. (1996) Coordinated dual cleavages induced by the proteasome regulator PA28 lead to dominant MHC ligands. *Cell.* **86**, 253–262 [https://doi.org/10.1016/S0092-8674\(00\)80097-5](https://doi.org/10.1016/S0092-8674(00)80097-5)
- 50 Emmerich, N.P., Nussbaum, A.K., Stevanovic, S., Priemer, M., Toes, R.E., Rammensee, H.G. et al. (2000) The human 26 S and 20 S proteasomes generate overlapping but different sets of peptide fragments from a model protein substrate. *J. Biol. Chem.* **275**, 21140–21148 <https://doi.org/10.1074/jbc.M000740200>
- 51 VerPlank, J.J.S. and Goldberg, A.L. (2017) Regulating protein breakdown through proteasome phosphorylation. *Biochem. J.* **474**, 3355–3371 <https://doi.org/10.1042/BCJ20160809>
- 52 Fabre, B., Lambour, T., Garrigues, L., Amalric, F., Vigneron, N., Menneteau, T. et al. (2015) Deciphering preferential interactions within supramolecular protein complexes: the proteasome case. *Mol. Syst. Biol.* **11**, 771 <https://doi.org/10.15252/msb.20145497>
- 53 Lesne, J., Locard-Paulet, M., Parra, J., Zivkovic, D., Menneteau, T., Bousquet, M.P. et al. (2020) Conformational maps of human 20S proteasomes reveal PA28- and immuno-dependent inter-ring crosstalks. *Nat. Commun.* **11**, 6140 <https://doi.org/10.1038/s41467-020-19934-z>
- 54 Mishto, M., Luciani, F., Holzhtuter, H.G., Bellavista, E., Santoro, A., Textoris-Taube, K. et al. (2008) Modeling the in vitro 20S proteasome activity: the effect of PA28-alpha-beta and of the sequence and length of polypeptides on the degradation kinetics. *J. Mol. Biol.* **377**, 1607–1617 <https://doi.org/10.1016/j.jmb.2008.01.086>
- 55 de Graaf, N., van Helden, M.J., Textoris-Taube, K., Chiba, T., Topham, D.J., Kloetzel, P.M. et al. (2011) PA28 and the proteasome immunosubunits play a central and independent role in the production of MHC class I-binding peptides in vivo. *Eur. J. Immunol.* **41**, 926–935 <https://doi.org/10.1002/eji.201041040>
- 56 Huang, X., Seifert, U., Salzmann, U., Henklein, P., Preissner, R., Henke, W. et al. (2002) The RTP site shared by the HIV-1 Tat protein and the 11S regulator subunit alpha is crucial for their effects on proteasome function including antigen processing. *J. Mol. Biol.* **323**, 771–782 [https://doi.org/10.1016/S0022-2836\(02\)00998-1](https://doi.org/10.1016/S0022-2836(02)00998-1)
- 57 Respondek, D., Voss, M., Kuhlewindt, I., Klingel, K., Kruger, E. and Beling, A. (2017) PA28 modulates antigen processing and viral replication during coxsackievirus B3 infection. *PLoS ONE* **12**, e0173259 <https://doi.org/10.1371/journal.pone.0173259>
- 58 Raule, M., Cerruti, F., Benaroudj, N., Migotti, R., Kikuchi, J., Bachi, A. et al. (2014) PA28alpha-beta reduces size and increases hydrophilicity of 20S immunoproteasome peptide products. *Chem. Biol.* **21**, 470–480 <https://doi.org/10.1016/j.chembiol.2014.02.006>
- 59 Milner, E., Gutter-Kapon, L., Bassani-Strenberg, M., Barnea, E., Beer, I. and Admon, A. (2013) The effect of proteasome inhibition on the generation of the human leukocyte antigen (HLA) peptidome. *Mol. Cell. Proteomics* **12**, 1853–1864 <https://doi.org/10.1074/mcp.M112.026013>
- 60 Seifert, U., Maranon, C., Shmueli, A., Desoutter, J.F., Wesoloski, L., Janek, K. et al. (2003) An essential role for tripeptidyl peptidase in the generation of an MHC class I epitope. *Nat. Immunol.* **4**, 375–379 <https://doi.org/10.1038/ni905>
- 61 Textoris-Taube, K., Cammann, C., Henklein, P., Topfstedt, E., Ebstein, F., Henze, S. et al. (2020) ER-aminopeptidase 1 determines the processing and presentation of an immunotherapy-relevant melanoma epitope. *Eur. J. Immunol.* **50**, 270–283 <https://doi.org/10.1002/eji.201948116>
- 62 Vigneron, N., Ferrari, V., Van den Eynde, B.J., Cresswell, P. and Leonhardt, R.M. (2018) Cytosolic processing governs TAP-independent presentation of a critical melanoma antigen. *J. Immunol.* **201**, 1875–1888 <https://doi.org/10.4049/jimmunol.1701479>
- 63 Neeffes, J., Jongasma, M.L., Paul, P. and Bakke, O. (2011) Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836 <https://doi.org/10.1038/nri3084>
- 64 Rock, K.L., Reits, E. and Neeffes, J. (2016) Present yourself! by MHC class I and MHC class II molecules. *Trends Immunol.* **37**, 724–737 <https://doi.org/10.1016/j.it.2016.08.010>
- 65 Maben, Z., Arya, R., Georgiadis, D., Stratikos, E. and Stern, L.J. (2021) Conformational dynamics linked to domain closure and substrate binding explain the ERAP1 allosteric regulation mechanism. *Nat. Commun.* **12**, 5302 <https://doi.org/10.1038/s41467-021-25564-w>
- 66 Niepieklo-Miniowska, W., Matusiak, L., Narbutt, J., Lesiak, A., Kuna, P., Wisniewski, A. et al. (2021) Contribution of antigen-processing machinery genetic polymorphisms to atopic dermatitis. *Life (Basel)* **11**, 333 <https://doi.org/10.3390/life11040333>
- 67 Reeves, E., Wood, O., Ottensmeier, C.H., King, E.V., Thomas, G.J., Elliott, T. et al. (2019) HPV epitope processing differences correlate with ERAP1 allotype and extent of CD8(+) T-cell tumor infiltration in OPSCC. *Cancer Immunol. Res.* **7**, 1202–1213 <https://doi.org/10.1158/2326-6066.CIR-18-0498>
- 68 Hanson, A.L., International Genetics of Ankylosing Spondylitis Consortium, Vukcevic, D., Leslie, S., Harris, J., Le Cao, K.A., Kenna, T.J. et al. (2020) Epistatic interactions between killer immunoglobulin-like receptors and human leukocyte antigen ligands are associated with ankylosing spondylitis. *PLoS Genet.* **16**, e1008906 <https://doi.org/10.1371/journal.pgen.1008906>

- 69 Hanson, A.L., Nel, H.J., Bradbury, L., Phipps, J., Thomas, R., Le Cao, K.A. et al. (2020) Altered repertoire diversity and disease-associated clonal expansions revealed by T cell receptor immunosequencing in ankylosing spondylitis patients. *Arthritis Rheumatol.* **72**, 1289–1302 <https://doi.org/10.1002/art.41252>
- 70 Rupert Abele, R.T. (1999) Function of the transport complex TAP in cellular immune recognition. *Biochim. Biophys. Acta* **1461**, 405–419 [https://doi.org/10.1016/S0005-2736\(99\)00171-6](https://doi.org/10.1016/S0005-2736(99)00171-6)
- 71 Koopmann, J.-O., Post, M., Neefjes, J.J., Hammerling, G.J. and Momburg, F. (1996) Translocation of long peptides by transporters associated with antigen processing (TAP). *Eur. J. Immunol.* **26**, 1720–1728 <https://doi.org/10.1002/eji.1830260809>
- 72 Neefjes, J., Gottfried, E., Roelse, J., Gromme, M., Obst, R., Hammerling, G.J. et al. (1995) Analysis of the fine specificity of rat, mouse and human TAP peptide transporters. *Eur. J. Immunol.* **25**, 1133–1136 <https://doi.org/10.1002/eji.1830250444>
- 73 Neisig, A., Roelse, J., Sijts, A.J., Ossendorp, F., Feltkamp, M.C., Kast, W.M. et al. (1995) Major differences in transporter associated with antigen presentation (TAP)-dependent translocation of MHC class I-presentable peptides and the effect of flanking sequences. *J. Immunol.* **154**, 1273–1279 PMID:7822796
- 74 Weinzierl, A.O., Rudolf, D., Hillen, N., Tenzer, S., van Ender, P., Schild, H. et al. (2008) Features of TAP-independent MHC class I ligands revealed by quantitative mass spectrometry. *Eur. J. Immunol.* **38**, 1503–1510 <https://doi.org/10.1002/eji.200838136>
- 75 Bleses, A., Janulienė, D., Hofmann, T., Koller, N., Schmidt, C., Trowitzsch, S. et al. (2017) Structure of the human MHC-I peptide-loading complex. *Nature* **551**, 525–528 <https://doi.org/10.1038/nature24627>
- 76 Wieczorek, M., Abualrous, E.T., Sticht, J., Alvaro-Benito, M., Stolzenberg, S., Noe, F. et al. (2017) Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation. *Front. Immunol.* **8**, 292 <https://doi.org/10.3389/fimmu.2017.00292>
- 77 Abualrous, E.T., Saini, S.K., Ramnarayan, V.R., Ilca, F.T., Zacharias, M. and Springer, S. (2015) The carboxy terminus of the ligand peptide determines the stability of the MHC class I molecule H-2Kb: a combined molecular dynamics and experimental study. *PLoS ONE* **10**, e0135421 <https://doi.org/10.1371/journal.pone.0135421>
- 78 Rist, M.J., Theodossis, A., Croft, N.P., Neller, M.A., Welland, A., Chen, Z. et al. (2013) HLA peptide length preferences control CD8+T cell responses. *J. Immunol.* **191**, 561–571 <https://doi.org/10.4049/jimmunol.1300292>
- 79 Sidney, J., Peters, B., Frahm, N., Brander, C. and Sette, A. (2008) HLA class I supertypes: a revised and updated classification. *BMC Immunol.* **9**, 1 <https://doi.org/10.1186/1471-2172-9-1>
- 80 Faridi, P., Li, C., Ramarathinam, S.H., Vivian, J.P., Illing, P.T., Mifsud, N.A. et al. (2018) A subset of HLA-I peptides are not genomically templated: evidence for cis- and trans-spliced peptide ligands. *Sci. Immunol.* **3**, eaar3947 <https://doi.org/10.1126/sciimmunol.aar3947>
- 81 Liepe, J., Marino, F., Sidney, J., Jeko, A., Bunting, D.E., Sette, A. et al. (2016) A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* **354**, 354–358 <https://doi.org/10.1126/science.aaf4384>
- 82 Bassani-Sternberg, M., Braunlein, E., Klar, R., Engleitner, T., Sinitcyn, P., Audehm, S. et al. (2016) Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat. Commun.* **7**, 13404 <https://doi.org/10.1038/ncomms13404>
- 83 Berlin, C., Kowalewski, D.J., Schuster, H., Mirza, N., Walz, S., Handel, M. et al. (2015) Mapping the HLA ligandome landscape of acute myeloid leukemia: a targeted approach toward peptide-based immunotherapy. *Leukemia* **29**, 647–659 <https://doi.org/10.1038/leu.2014.233>
- 84 Schuster, H., Shao, W., Weiss, T., Pedrioli, P.G.A., Roth, P., Weller, M. et al. (2018) A tissue-based draft map of the murine MHC class I immunopeptidome. *Sci. Data* **5**, 180157 <https://doi.org/10.1038/sdata.2018.157>
- 85 Sarkizova, S., Klaeger, S., Le, P.M., Li, L.W., Oliveira, G., Keshishian, H. et al. (2020) A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat. Biotechnol.* **38**, 199–209 <https://doi.org/10.1038/s41587-019-0322-9>
- 86 Remesh, S.G., Andreatta, M., Ying, G., Kaever, T., Nielsen, M., McMurtrey, C. et al. (2017) Unconventional peptide presentation by major histocompatibility complex (MHC) class I allele HLA-A*02:01: breaking confinement. *J. Biol. Chem.* **292**, 5262–5270 <https://doi.org/10.1074/jbc.M117.776542>
- 87 McMurtrey, C., Trolle, T., Sansom, T., Remesh, S.G., Kaever, T., Bardet, W. et al. (2016) *Toxoplasma gondii* peptide ligands open the gate of the HLA class I binding groove. *eLife* **5**, e12556 <https://doi.org/10.7554/eLife.12556>
- 88 Daniel, C., Horvath, S. and Allen, P.M. (1998) A basis for alloreactivity: MHC helical residues broaden peptide recognition by the TCR clade. *Immunity* **8**, 543–552 [https://doi.org/10.1016/S1074-7613\(00\)80559-2](https://doi.org/10.1016/S1074-7613(00)80559-2)
- 89 Yewdell, J.W. and Bennink, J.R. (1999) IMMUNODOMINANCE IN MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I-RESTRICTED T LYMPHOCYTE RESPONSES. *Annu. Rev. Immunol.* **17**, 51–88 <https://doi.org/10.1146/annurev.immunol.17.1.51>
- 90 Andersen, R.S., Andersen, S.R., Hjortso, M.D., Lyngaa, R., Idorn, M., Kollgard, T.M. et al. (2013) High frequency of T cells specific for cryptic epitopes in melanoma patients. *Oncoimmunology* **2**, e25374 <https://doi.org/10.4161/onci.25374>
- 91 Apcher, S., Millot, G., Daskalogianni, C., Scherl, A., Manoury, B. and Fahraeus, R. (2013) Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway. *Proc. Natl Acad. Sci. U.S.A.* **110**, 17951–17956 <https://doi.org/10.1073/pnas.1309956110>
- 92 Ehx, G., Larouche, J.D., Durette, C., Laverdure, J.P., Hesnard, L., Vincent, K. et al. (2021) Atypical acute myeloid leukemia-specific transcripts generate shared and immunogenic MHC class-I-associated epitopes. *Immunity* **54**, 737–752 e710 <https://doi.org/10.1016/j.immuni.2021.03.001>
- 93 Laumont, C.M., Vincent, K., Hesnard, L., Audemard, E., Bonneil, E., Laverdure, J.P. et al. (2018) Noncoding regions are the main source of targetable tumor-specific antigens. *Sci. Transl. Med.* **10**, eaau5516 <https://doi.org/10.1126/scitranslmed.aau5516>
- 94 Attig, J., Young, G.R., Hosie, L., Perkins, D., Encheva-Yokoya, V., Stoye, J.P. et al. (2019) LTR retroelement expansion of the human cancer transcriptome and immunopeptidome revealed by de novo transcript assembly. *Genome Res.* **29**, 1578–1590 <https://doi.org/10.1101/gr.248922.119>
- 95 Chong, C., Muller, M., Pak, H., Harnett, D., Huber, F., Grun, D. et al. (2020) Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat. Commun.* **11**, 1293 <https://doi.org/10.1038/s41467-020-14968-9>
- 96 Smart, A.C., Margolis, C.A., Pimentel, H., He, M.X., Miao, D., Adeegbe, D. et al. (2018) Intron retention is a source of neoepitopes in cancer. *Nat. Biotechnol.* **36**, 1056–1058 <https://doi.org/10.1038/nbt.4239>
- 97 Ruiz Cuevas, M.V., Hardy, M.P., Holly, J., Bonneil, E., Durette, C., Courcelles, M. et al. (2021) Most non-canonical proteins uniquely populate the proteome or immunopeptidome. *Cell Rep.* **34**, 108815 <https://doi.org/10.1016/j.celrep.2021.108815>
- 98 Laumont, C.M., Daouda, T., Laverdure, J.P., Bonneil, E., Caron-Lizotte, O., Hardy, M.P. et al. (2016) Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat. Commun.* **7**, 10238 <https://doi.org/10.1038/ncomms10238>

- 99 Gonzalez-Duque, S., Azoury, M.E., Colli, M.L., Afonso, G., Turatsinze, J.V., Nigi, L. et al. (2018) Conventional and neo-antigenic peptides presented by beta cells Are targeted by circulating naive CD8+ T cells in type 1 diabetic and healthy donors. *Cell Metab.* **28**, 946–960 <https://doi.org/10.1016/j.cmet.2018.07.007>
- 100 Dersh, D., Holly, J. and Yewdell, J.W. (2021) A few good peptides: MHC class I-based cancer immunosurveillance and immunoevasion. *Nat. Rev. Immunol.* **21**, 116–128 <https://doi.org/10.1038/s41577-020-0390-6>
- 101 Granados, D.P., Sriranganadane, D., Daouda, T., Zieger, A., Laumont, C.M., Caron-Lizotte, O. et al. (2014) Impact of genomic polymorphisms on the repertoire of human MHC class I-associated peptides. *Nat. Commun.* **5**, 3600 <https://doi.org/10.1038/ncomms4600>
- 102 Mishto, M., Rodriguez-Hernandez, G., Neefjes, J., Urlaub, H. and Liepe, J. (2021) Response: commentary: 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.* **12**, 679836 <https://doi.org/10.3389/fimmu.2021.679836>
- 103 Mishto, M., Mansurkhodzhaev, A., Ying, G., Bitra, A., Cordfunke, R.A., Henze, S. et al. (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 <https://doi.org/10.3389/fimmu.2019.02572>
- 104 Litchfield, K., Reading, J.L., Puttick, C., Thakkar, K., Abbosh, C., Bentham, R. et al. (2021) Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* **184**, 596–614 e514 <https://doi.org/10.1016/j.cell.2021.01.002>
- 105 McGranahan, N., Furness, A.J., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S.K. et al. (2016) Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 <https://doi.org/10.1126/science.aaf1490>
- 106 Zarling, A.L., Polefrone, J.M., Evans, A.M., Mikesh, L.M., Shabanowitz, J., Lewis, S.T. et al. (2006) Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy. *Proc. Natl Acad. Sci. U.S.A.* **103**, 14889–14894 <https://doi.org/10.1073/pnas.0604045103>
- 107 Engelhard, V.H., Altrich-Vanlith, M., Ostankovitch, M. and Zarling, A.L. (2006) Post-translational modifications of naturally processed MHC-binding epitopes. *Curr. Opin. Immunol.* **18**, 92–97 <https://doi.org/10.1016/j.coi.2005.11.015>
- 108 Engelhard, V.H., Obeng, R.C., Cummings, K.L., Petroni, G.R., Ambakhtwala, A.L., Chianese-Bullock, K.A. et al. (2020) MHC-restricted phosphopeptide antigens: preclinical validation and first-in-humans clinical trial in participants with high-risk melanoma. *J. Immunother. Cancer* **8**, e000262 <https://doi.org/10.1136/jitc-2019-000262>
- 109 Alpizar, A., Marino, F., Ramos-Fernandez, A., Lombardia, M., Jeko, A., Pazos, F. et al. (2017) A molecular basis for the presentation of phosphorylated peptides by HLA-B antigens. *Mol. Cell. Proteomics* **16**, 181–193 <https://doi.org/10.1074/mcp.M116.063800>
- 110 Reed, B., Crawford, F., Hill, R.C., Jin, N., White, J., Krovi, S.H. et al. (2021) Lysosomal cathepsin creates chimeric epitopes for diabetogenic CD4T cells via transpeptidation. *J. Exp. Med.* **218**, e20192135 <https://doi.org/10.1084/jem.20192135>
- 111 Arribas-Layton, D., Guyer, P., Delong, T., Dang, M., Chow, I.T., Speake, C. et al. (2020) Hybrid insulin peptides Are recognized by human T cells in the context of DRB1*04:01. *Diabetes* **69**, 1492–1502 <https://doi.org/10.2337/db19-0620>
- 112 Wiles, T.A., Powell, R., Michel, R., Beard, K.S., Hohenstein, A., Bradley, B. et al. (2019) Identification of hybrid insulin peptides (HIPs) in mouse and human islets by mass spectrometry. *J. Proteome Res.* **18**, 814–825 <https://doi.org/10.1021/acs.jproteome.8b00875>
- 113 Vigneron, N., Stroobant, V., Chapiro, J., Ooms, A., Degiovanni, G., Morel, S. et al. (2004) An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587–590 <https://doi.org/10.1126/science.1095522>
- 114 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 <https://doi.org/10.1002/eji.201040750>
- 115 Dalet, A., Vigneron, N., Stroobant, V., Hanada, K. 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 <https://doi.org/10.4049/jimmunol.0901277>
- 116 Mishto, M., Goede, A., Taube, K.T., Keller, C., Janek, K., Henklein, P. et al. (2012) Driving forces of proteasome-catalyzed peptide splicing in yeast and humans. *Mol. Cell. Proteomics* **11**, 1008–1023 <https://doi.org/10.1074/mcp.M112.020164>
- 117 Warren, E.H., Vigneron, N.J., Gavin, M.A., Coulie, P.G., Stroobant, V., Dalet, A. et al. (2006) An antigen produced by splicing of noncontiguous peptides in the reverse order. *Science* **313**, 1444–1447 <https://doi.org/10.1126/science.1130660>
- 118 Ebstein, F., Textoris-Taube, K., Keller, C., Golnik, R., Vigneron, N., Van den Eynde, B.J. et al. (2016) Proteasomes generate spliced epitopes by two different mechanisms and as efficiently as non-spliced epitopes. *Sci. Rep.* **6**, 24032 <https://doi.org/10.1038/srep24032>
- 119 Hanada, K., 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 <https://doi.org/10.1038/nature02240>
- 120 Admon, A. (2021) Are there indeed spliced peptides in the immunopeptidome? *Mol. Cell. Proteomics* **20**, 100099 <https://doi.org/10.1016/j.mcpro.2021.100099>
- 121 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 <https://doi.org/10.1111/cei.13589>
- 122 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* **20**, e2000112 <https://doi.org/10.1002/pmic.202000112>
- 123 Purcell, A.W. (2021) Is the immunopeptidome getting darker?: a commentary on the discussion around mishto et al., 2019. *Front. Immunol.* **12**, 720811 <https://doi.org/10.3389/fimmu.2021.720811>
- 124 Mishto, M. (2021) Commentary: are there indeed spliced peptides in the immunopeptidome? *Mol. Cell. Proteomics* **20**, 100158 <https://doi.org/10.1016/j.mcpro.2021.100158>
- 125 Dalet, A., Robbins, P.F., Stroobant, V., Vigneron, N., Li, Y.F., El-Gamil, M. et al. (2011) An antigenic peptide produced by reverse splicing and double asparagine deamidation. *Proc. Natl Acad. Sci. U.S.A.* **108**, E323–E331 <https://doi.org/10.1073/pnas.1101892108>
- 126 Michaux, A., Larrieu, P., Stroobant, V., Fonteneau, J.F., Jotereau, F., Van den Eynde, B.J. et al. (2014) 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**, 1962–1971 <https://doi.org/10.4049/jimmunol.1302032>

- 127 Platteel, A.C.M., Liepe, J., Textoris-Taube, K., Keller, C., Henklein, P., Schalkwijk, H.H. et al. (2017) Multi-level strategy for identifying proteasome-catalyzed spliced epitopes targeted by CD8+ T cells during bacterial infection. *Cell Rep.* **20**, 1242–1253 <https://doi.org/10.1016/j.celrep.2017.07.026>
- 128 Paes, W., Leonov, G., Partridge, T., Chikata, T., Murakoshi, H., Frangou, A. 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 <https://doi.org/10.1073/pnas.1911622116>
- 129 Platteel, A.C., Mishto, M., Textoris-Taube, K., Keller, C., Liepe, J., Busch, D.H. et al. (2016) CD8(+) t cells of Listeria monocytogenes-infected mice recognize both linear and spliced proteasome products. *Eur. J. Immunol.* **46**, 1109–1118 <https://doi.org/10.1002/eji.201545989>
- 130 Robbins, P.F., el-Gamil, M., Kawakami, Y., Stevens, E., Yannelli, J.R. and Rosenberg, S.A. (1994) Recognition of tyrosinase by tumor-infiltrating lymphocytes from a patient responding to immunotherapy. *Cancer Res.* **54**, 3124–3126 PMID:8205528
- 131 Bonnet, D., Warren, E.H., Greenberg, P.D., Dick, J.E. and Riddell, S.R. (1999) CD8(+) minor histocompatibility antigen-specific cytotoxic T lymphocyte clones eliminate human acute myeloid leukemia stem cells. *Proc. Natl Acad. Sci. U.S.A.* **96**, 8639–8644 <https://doi.org/10.1073/pnas.96.15.8639>
- 132 Fidanza, M., Gupta, P., Sayana, A., Shanker, V., Pahlke, S.M., Vu, B. et al. (2021) Enhancing proteasomal processing improves survival for a peptide vaccine used to treat glioblastoma. *Sci. Transl. Med.* **13**, eaax4100 <https://doi.org/10.1126/scitranslmed.aax4100>
- 133 Specht, G., Roetschke, H.P., Mansurkhodzhaev, A., Henklein, P., Textoris-Taube, K., Urlaub, H. et al. (2020) Large database for the analysis and prediction of spliced and non-spliced peptide generation by proteasomes. *Sci. Data* **7**, 146 <https://doi.org/10.1038/s41597-020-0487-6>
- 134 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 <https://doi.org/10.1158/2326-6066.CIR-18-0424>
- 135 Hayday, A.C. and Vantourout, P. (2020) The innate biologies of adaptive antigen receptors. *Annu. Rev. Immunol.* **38**, 487–510 <https://doi.org/10.1146/annurev-immunol-102819-023144>
- 136 Lanzarotti, E., Marcatili, P. and Nielsen, M. (2019) T-Cell receptor cognate target prediction based on paired α and β chain sequence and structural CDR loop similarities. *Front. Immunol.* **10**, 2080 <https://doi.org/10.3389/fimmu.2019.02080>
- 137 Zarnitsyna, V.I., Evavold, B.D., Schoettle, L.N., Blattman, J.N. and Antia, R. (2013) Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front. Immunol.* **4**, 485 <https://doi.org/10.3389/fimmu.2013.00485>
- 138 Sewell, A.K. (2012) Why must T cells be cross-reactive? *Nat. Rev. Immunol.* **12**, 669–677 <https://doi.org/10.1038/nri3279>
- 139 Madi, A., Poran, A., Shifrut, E., Reich-Zeliger, S., Greenstein, E., Zaretsky, I. et al. (2017) T cell receptor repertoires of mice and humans are clustered in similarity networks around conserved public CDR3 sequences. *eLife* **6**, e22057 <https://doi.org/10.7554/eLife.22057>
- 140 DeWitt, W.S., Smith, A., Schoch, G., Hansen, J.A., Matsen, F.A. and Bradley, P. (2018) Human T cell receptor occurrence patterns encode immune history, genetic background, and receptor specificity. *eLife* **7**, e38358 <https://doi.org/10.7554/eLife.38358>
- 141 Dash, P., Fiore-Gartland, A.J., Hertz, T., Wang, G.C., Sharma, S., Souquette, A. et al. (2017) Quantifiable predictive features define epitope specific T cell receptor repertoires. *Nature*. **547**, 89–93 <https://doi.org/10.1038/nature22383>
- 142 Glanville, J., Huang, H., Nau, A., Hatton, O., Wagar, L.E., Rubelt, F. et al. (2017) Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**, 94–98 <https://doi.org/10.1038/nature22976>
- 143 Pogorely, M.V. and Shugay, M. (2019) A framework for annotation of antigen specificities in high-throughput T-cell repertoire sequencing studies. *Front. Immunol.* **10**, 2159 <https://doi.org/10.3389/fimmu.2019.02159>
- 144 Ritvo, P.-G., Saadawi, A., Barennes, P., Quiniou, V., Chaara, W., Soufi, K.E. et al. (2018) High-resolution repertoire analysis reveals a major bystander activation of Tfh and Tfr cells. *Proc. Natl Acad. Sci. U.S.A.* **115**, 9604–9609 <https://doi.org/10.1073/pnas.1808594115>
- 145 Baker, B.M., Scott, D.R., Blevins, S.J. and Hawse, W.F. (2012) Structural and dynamic control of T-cell receptor specificity, cross-reactivity, and binding mechanism. *Immunol. Rev.* **250**, 10–31 <https://doi.org/10.1111/j.1600-065X.2012.01165.x>
- 146 Rossjohn, J., Gras, S., Miles, J.J., Turner, S.J., Godfrey, D.I. and McCluskey, J. (2015) T cell antigen receptor recognition of antigen-Presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200 <https://doi.org/10.1146/annurev-immunol-032414-112334>
- 147 Wooldridge, L., Ekeruche-Makinde, J., van den Berg, H.A., Skowera, A., Miles, J.J., Tan, M.P. et al. (2012) A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177 <https://doi.org/10.1074/jbc.M111.289488>
- 148 Arstila, T.P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J. and Kourlisky, P. (1999) A direct estimate of the human $\alpha\beta$ T cell receptor diversity. *Science* **286**, 958–961 <https://doi.org/10.1126/science.286.5441.958>
- 149 Qi, Q., Liu, Y., Cheng, Y., Glanville, J., Zhang, D., Lee, J.Y. et al. (2014) Diversity and clonal selection in the human T-cell repertoire. *Proc. Natl Acad. Sci. U.S.A.* **111**, 13139–13144 <https://doi.org/10.1073/pnas.1409155111>
- 150 Mösch, A., Raffeggerst, S., Weis, M., Schendel, D.J. and Frishman, D. (2019) Machine learning for cancer immunotherapies based on epitope recognition by T cell receptors. *Front. Genet.* **10**, 1141 <https://doi.org/10.3389/fgene.2019.01141>
- 151 Welsh, R.M., Che, J.W., Brehm, M.A. and Selin, L.K. (2010) Heterologous immunity between viruses. *Immunol. Rev.* **235**, 244–266 <https://doi.org/10.1111/j.0105-2896.2010.00897.x>
- 152 Mendes, M.F.A., Antunes, D.A., Rigo, M.M., Sinigaglia, M. and Vieira, G.F. (2015) Improved structural method for T-cell cross-reactivity prediction. *Mol. Immunol.* **67**, 303–310 <https://doi.org/10.1016/j.molimm.2015.06.017>
- 153 Borbulevych, O.Y., Piepenbrink, K.H., Gloor, B.E., Scott, D.R., Sommese, R.F., Cole, D.K. et al. (2009) T cell receptor cross-reactivity directed by antigen-dependent tuning of peptide-MHC molecular flexibility. *Immunity* **31**, 885–896 <https://doi.org/10.1016/j.immuni.2009.11.003>
- 154 Yin, Y. and Mariuzza, R.A. (2009) The multiple mechanisms of T cell receptor cross-reactivity. *Immunity* **31**, 849–851 <https://doi.org/10.1016/j.immuni.2009.12.002>
- 155 Petrova, G., Ferrante, A. and Gorski, J. (2012) Cross-Reactivity of T cells and its role in the immune system. *Crit. Rev. Immunol.* **32**, 349–372 <https://doi.org/10.1615/CritRevImmunol.v32.i4.50>
- 156 Fodor, J., Riley, B.T., Borg, N.A. and Buckle, A.M. (2018) Previously hidden dynamics at the TCR–peptide–MHC interface revealed. *J. Immunol.* **200**, 4134–4145 <https://doi.org/10.4049/jimmunol.1800315>
- 157 Pierce, B.G. and Weng, Z. (2013) A flexible docking approach for prediction of T cell receptor–peptide–MHC complexes. *Protein Sci.* **22**, 35–46 <https://doi.org/10.1002/pro.2181>
- 158 Riley, T.P., Hellman, L.M., Gee, M.H., Mendoza, J.L., Alonso, J.A., Foley, K.C. et al. (2018) T cell receptor cross-reactivity expanded by dramatic peptide–MHC adaptability. *Nat. Chem. Biol.* **14**, 934–942 <https://doi.org/10.1038/s41589-018-0130-4>

- 159 Rudolph, M.G., Stanfield, R.L. and Wilson, I.A. (2006) How tcrs bind mhcs, peptides, and coreceptors. *Annu. Rev. Immunol.* **24**, 419–466 <https://doi.org/10.1146/annurev.immunol.23.021704.115658>
- 160 Gowthaman, R. and Pierce, B.G. (2019) TCR3d: the T cell receptor structural repertoire database. *Bioinformatics* **35**, 5323–5325 <https://doi.org/10.1093/bioinformatics/btz517>
- 161 Zhang, S., Bakshi, R.K., Suneetha, P.V., Fytilli, P., Antunes, D.A., Vieira, G.F. et al. (2015) Frequency, private specificity, and cross-reactivity of preexisting hepatitis C virus (HCV)-specific CD8+ T cells in HCV-seronegative individuals: implications for vaccine responses. *J. Virol.* **89**, 8304–8317 <https://doi.org/10.1128/JVI.00539-15>
- 162 Antunes, D.A., Rigo, M.M., Freitas, M.V., Mendes, M.F.A., Sinigaglia, M., Lizée, G. et al. (2017) Interpreting T-cell cross-reactivity through structure: implications for TCR-based cancer immunotherapy. *Front. Immunol.* **8**, 1210 <https://doi.org/10.3389/fimmu.2017.01210>
- 163 Sibener, L.V., Fernandes, R.A., Kolawole, E.M., Carbone, C.B., Liu, F., McAfee, D. et al. (2018) Isolation of a structural mechanism for uncoupling T cell receptor signaling from peptide-MHC binding. *Cell* **174**, 672–687.e627 <https://doi.org/10.1016/j.cell.2018.06.017>