A guide to antigen processing and presentation

Novalia Pishesha[1,2,3], Thibault J. Harmand[1] and Hidde L. Ploegh[1] ✉

Abstract | Antigen processing and presentation are the cornerstones of adaptive immunity. B cells cannot generate high-affinity antibodies without T cell help. CD4$^+$ T cells, which provide such help, use antigen-specific receptors that recognize major histocompatibility complex (MHC) molecules in complex with peptide cargo. Similarly, eradication of virus-infected cells often depends on cytotoxic CD8$^+$ T cells, which rely on the recognition of peptide–MHC complexes for their action. The two major classes of glycoproteins entrusted with antigen presentation are the MHC class I and class II molecules, which present antigenic peptides to CD8$^+$ T cells and CD4$^+$ T cells, respectively. This Review describes the essentials of antigen processing and presentation. These pathways are divided into six discrete steps that allow a comparison of the various means by which antigens destined for presentation are acquired and how the source proteins for these antigens are tagged for degradation, destroyed and ultimately displayed as peptides in complex with MHC molecules for T cell recognition.

Adaptive immune responses require the involvement of the antigen processing and presentation pathways. These pathways enable glycoproteins encoded by the major histocompatibility complex (MHC) to be loaded with their proper ligands[1–3]. Only liganded MHC molecules expressed on the surface of antigen-presenting cells can activate T cells, which then execute effector functions, such as cytotoxicity, provision of help to B cells and cytokine production[4–6]. Loss of MHC molecules owing to mutations incurs the risk of infectious disease and can prevent elimination of malignant growths[7,8]. This is compelling evidence for the biological importance of the pathways and molecules involved. The strong associations of autoimmune diseases with particular allelic forms of MHC molecules lend yet further support[9]. How MHC molecules acquire their ligands therefore occupies a central position in immunology. Put differently, the adaptive immune system views the world of antigens through the lens of MHC molecules, MHC-like molecules and the cargo they present. While the literature on the subject is voluminous and has been expertly reviewed, this guide will focus on the essentials[1,10–13].

Perhaps the most important concept to emerge from our understanding of antigen presentation is this: it provides a means for the adaptive immune system to survey the host proteome and its modifications. Any changes that might indicate the presence of undesirable intruders, such as viruses and bacteria, or the occurrence of mutations that contribute to malignancy, can be detected. In the absence of infectious agents or cancerous cells, the constitutive nature of many of the biochemical and cell biological processes that dictate antigen presentation provides a reference for what the immune system should consider as harmless 'self'[14,15]. Such imposition of tolerance to 'self' operates not only in thymic T cell development and selection but also in homeostasis[16]. Antigen presentation is thus indelibly associated with establishing a functioning, self-tolerant T cell repertoire.

MHC molecules and their antigens

Structure of MHC molecules. In reference to the history of their discovery as drivers of transplant rejection, MHC molecules are classified as 'classical' or — in the case of the more recently characterized molecules — as 'non-classical'[17,18]. The unifying principle of MHC molecules that present antigens is their similar overall three-dimensional structure[19]. A co-dominantly expressed set of genes encode the classical MHC molecules in humans and mice[20–23] (FIG. 1). The vertebrate MHCs show strong evolutionary conservation, but with variations in the number of genes, their relative order and the number of allelic variants for each MHC locus. The structure of MHC molecules determines which antigens are presented. MHC class I (MHC-I) molecules are composed of a membrane-anchored, MHC-encoded glycoprotein heavy chain in association with β_2-microglobulin, a small soluble protein. MHC-I molecules bind peptides, typically in the 8–10 amino acid residue range, but longer peptide ligands may bind as well[24–26] (FIG. 2a). In both humans and mice, the genes that encode the MHC-I-like CD1 proteins — which present lipid antigens to innate-like

[1]*Program in Cellular and Molecular Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA.*

[2]*Society of Fellows, Harvard University, Cambridge, MA, USA.*

[3]*Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA.*

✉*e-mail: hidde.ploegh@childrens.harvard.edu*

https://doi.org/10.1038/s41577-022-00707-2

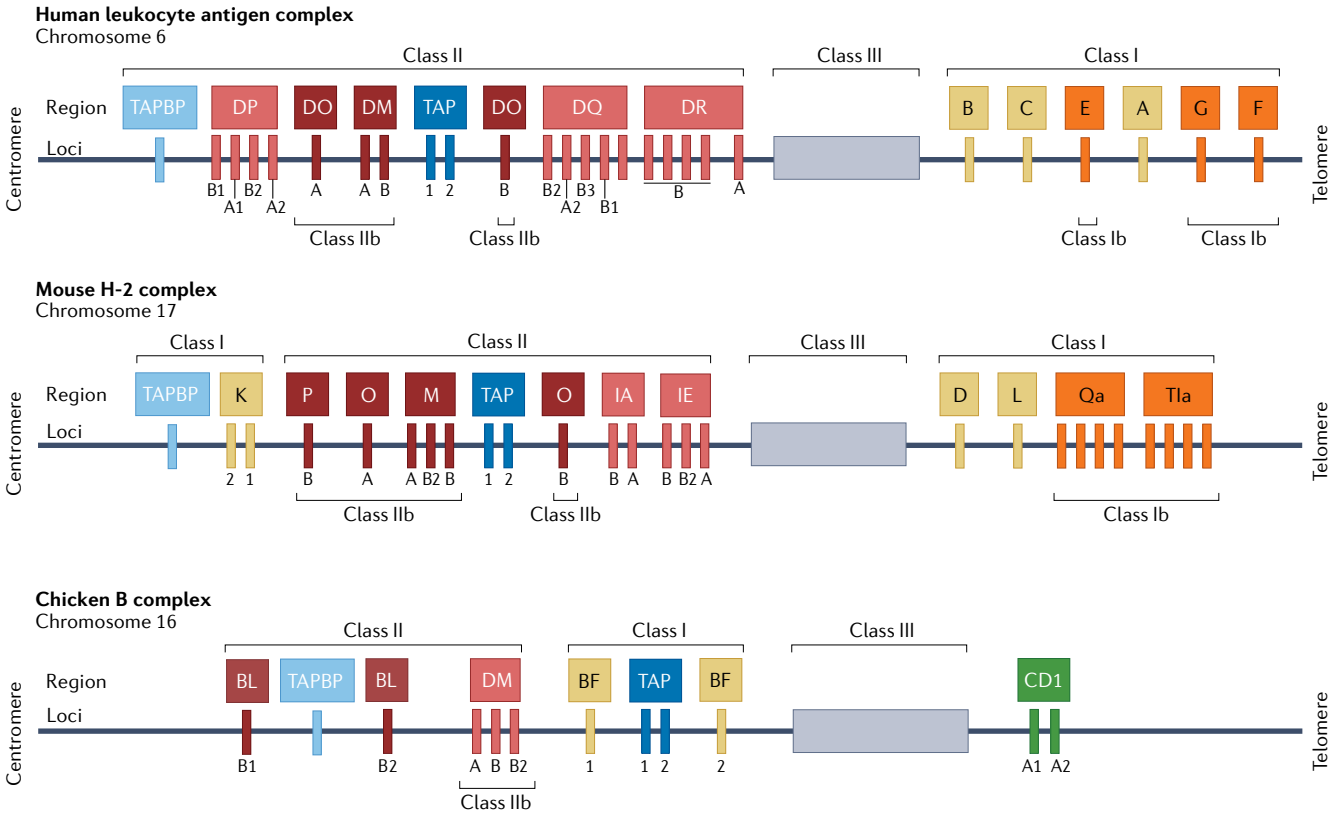

Fig. 1 | **Simplified schematic maps of the MHC complex in humans, mice and chickens.** The main major histocompatibility complex (MHC) genes are shown, including classical class I, Ib, II, IIb and III genes. The genes for the transporter associated with antigen processing (TAP) and tapasin (also known as TAP-binding protein (TAPBP)) are also included. Interestingly, the CD1 gene is located on the same chromosome as the MHC complex in chickens, but not in humans or mice[22,23,29].

T cells known as natural killer T cells — are not linked to the MHC, but in chickens and other vertebrates, CD1 genes are linked to the MHC[27–29]. The evolutionarily conserved, structurally related MHC-I-like protein MR1 is also not encoded by the MHC but is dedicated to the presentation of small molecules to populations of innate-like T cells known as mucosal-associated invariant T cells[27]. In addition to these 'true' antigen-presenting molecules, the MHC-I fold is present also in proteins of unrelated function (Supplementary Fig. 1).

MHC class II (MHC-II) molecules comprise MHC-encoded, membrane-anchored glycosylated α and β subunits. MHC-II molecules bind peptides of a slightly larger average size than those bound by MHC-I molecules — normally in the 10–15 amino acid residue range (FIG. 2b) — yet can also interact with far longer stretches of amino acids[30]. The biosynthesis of MHC molecules and their structural lookalikes is intimately linked to the relevant processing and presentation pathways, described later herein.

Polymorphism of MHC molecules is their hallmark[31]. The MHC molecules expressed in a given individual — who in an outbred population is usually heterozygous for the MHC — can present complex sets of peptide antigens. Each individual MHC molecule can present hundreds if not thousands of peptides. Evolution and selection drive this diversity to keep pace with the evolving spectrum of pathogens[32]. While the extent of polymorphism differs for various vertebrate species, the combined total number of molecularly defined allelic variants for human MHC-I and MHC-II molecules is in the thousands[33–35] (Supplementary Table 1). The genes for MHC molecules are located on a 1–2-Mb stretch of DNA that can comprise dozens of genes, including the genes that encode proteins involved in antigen processing and presentation. Examples include genes encoding a component of the peptide-loading complex (PLC), which orchestrates peptide translocation and editing, inducibly expressed proteasome subunits and the two subunits of the peptide transporter, all of which are important for antigen presentation via MHC-I molecules, as we discuss later[36]. Also encoded by the MHC are HLA-DM, a protein that serves as a peptide editor for MHC-II-bound ligands, and its modifier HLA-DO[36]. Finally, encoded elsewhere in the genome there are MHC-related proteins with functions not obviously related to antigen presentation[18]. The human MHC and the mouse MHC encode proteins referred to as non-classical MHC-I molecules or MHC-Ib molecules. These include HLA-E, HLA-F and HLA-G for the human MHC, and Qa-1, TL and related proteins in the mouse[18,37,38]. The HLA-E molecules found in rhesus macaques are of particular interest in view of their proposed superior capacity to bind and present a far wider array of peptides than their classical counterparts[39]. The MHC of rhesus macaques encodes several such

HLA-E-like proteins. On the other hand, HLA-F is less well studied, and has only recently been suggested to be a ligand for receptors on natural killer cells (NK cells) and to be capable of presenting peptides to T cells[40].

The structures of MHC-I and MHC-II molecules are well understood, including the features of their peptide-binding pockets that dictate which peptides they can accommodate. Recovery of peptides from

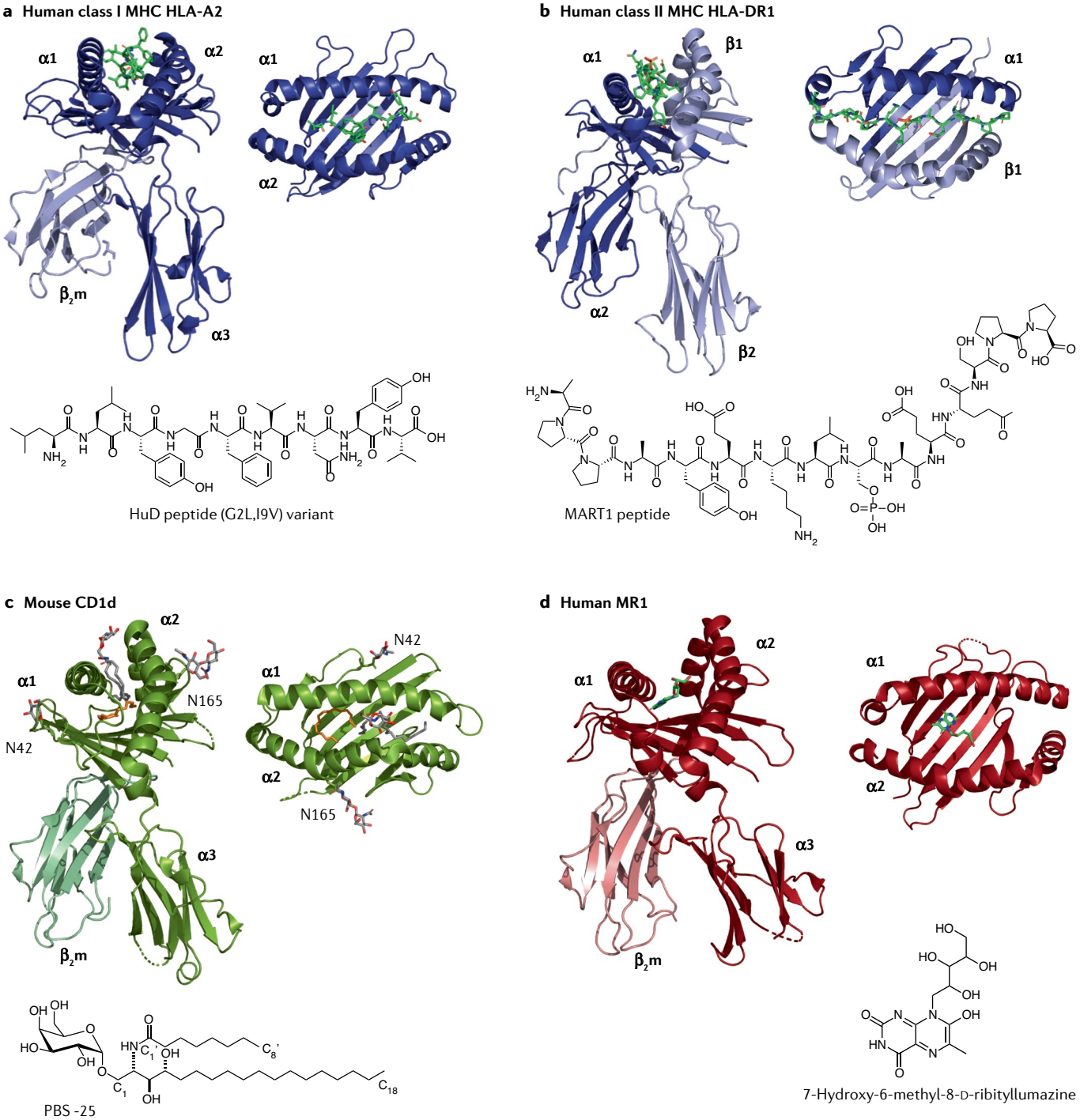

Fig. 2 | **Molecular models of classical MHC molecules and MHC-like molecules. a** | Side view and top view (rotated by 90°) of a crystal structure of a human major histocompatibility complex (MHC) class I (MHC-I)–peptide complex (Protein Data Bank (PDB) ID 3PWJ). The chemical structure of the peptide presented by this MHC-I molecule is shown under the top view[154]. **b** | Side view and top view (rotated by 90°) of a crystal structure of a human MHC class II (MHC-II)–peptide complex (PDB ID 3L6F). The chemical structure of the peptide presented by this MHC-II molecule is shown under the top view[155]. **c** | Side view and top view (rotated by 90°) of a crystal structure of the mouse MHC-I-like molecule, CD1, with a non-peptide lipid ligand (PDB ID 1Z5L). The chemical structure of the glycolipid is shown under the top view[156]. **d** | Side view and top view (rotated by 90°) of a crystal structure of a human MHC-I-like molecule, MR1, with a small molecule vitamin B metabolite (PDB ID 4L4V). The chemical structure of the small molecule is shown under the top view[157]. β_2m, β_2-microglobulin.

purified MHC molecules and the determination of their sequence by mass spectrometry is a technique referred to as '(immuno)peptidomics'[41]. These bound peptides are the actual products of processing and presentation pathways and are not based on predictions. X-ray structures of MHC molecules can further help define peptide-binding motifs on the basis of their molecular architecture[42,43]. Computational methods to predict peptides that can bind to a given MHC allele continue to improve but cannot replace mass spectrometry of recovered peptides and actual binding measurements[44].

Antigens. What are the types of antigens that must be processed for presentation at the cell surface? The ligands for classical MHC molecules are protein-derived peptides and modified versions of such peptides, for example those containing glycans, phosphates or disulfide-linked substituents[19]. Non-classical MHC-I molecules and MHC-I-like molecules sample other classes of molecules, including lipids and small molecules[27]. The ensemble of MHC-I and MHC-II ligands is referred to as the '(immuno)peptidome'. Proteolysis, as will be discussed later herein, is an essential step in the generation of the (immuno)peptidome, and multiple types of proteases contribute. For lipids, glycolipids and other small molecules presented by MHC-like molecules, the requisite molecular transformations may be on a more limited scale and may require enzymes of types different from those that act on proteins and peptides[27].

Stability and turnover of MHC–peptide complexes. Stability of MHC–peptide complexes is important to ensure that T cells that recognize them are not led astray. The absence of an appropriate ligand tends to destabilize MHC-I molecules, both in the case of cell surface-displayed molecules and for those in transit during their synthesis[10]. If a source of a ligand is withheld, either genetically or through pharmacological intervention, the assembly and surface expression of MHC-I molecules are compromised to various degrees[45]. Should a peptide dissociate, the resulting 'empty' MHC-I molecule assumes a less stable conformation, loses its tightly associated β_2-microglobulin subunit and becomes prone to internalization and proteolysis. Timely destruction of an MHC-I molecule that has lost its ligand reduces the possibility of peptide capture from the circulation or from virus-infected cells in close proximity. Peptide capture in *trans*, as distinct from peptides being sourced internally from the same cell that displays the MHC-I molecule in question, entails the risk of inviting inappropriate attack by cytotoxic $CD8^+$ T cells. Continuous protein turnover ensures that changes in gene expression are reported out by the affected cell as an altered MHC-I peptidome. By contrast, the biosynthesis of MHC-II molecules yields exceptionally stable peptide–MHC-II complexes that do not usually serve as a target for cytotoxic T cells and may not require such a molecular timer for self-immolation[46]. Over time, MHC-II molecules on antigen-presenting cells may thus activate many appropriately specific T cells. Recognition of an MHC-I–peptide complex on a virus-infected cells by a single $CD8^+$ cytotoxic T cell is all that is needed to eliminate that target[47].

Six keys steps in antigen presentation

The antigen processing and presentation pathways fit our understanding of biochemical and cell biological pathways common to all eukaryotes. MHC molecules are dispensable for survival of the cell that expresses them, which has facilitated the analysis of their biosynthesis and intracellular trafficking[48,49]. The division between the cytosol (and its topologically equivalent compartments) and the extracellular space (and its topological equivalents) is a helpful first distinction to understand why there are the two major classes of MHC molecule. To a first approximation, MHC-I molecules exploit the products of cytosolic proteolysis (with fine-tuning in the endoplasmic reticulum (ER)), whereas MHC-II molecules mostly sample the extracellular world, by presenting antigens captured and processed in endolysosomal compartments[4]. Compartmentalization of proteolysis and the distinct intracellular distributions of MHC molecules are the primary drivers of antigen processing and presentation. The entire process can be broken down into six discrete steps, which we summarize next.

Step 1: acquisition of antigens. Cells engaged in MHC-II-restricted presentation acquire antigens through phagocytosis and endocytosis, processes in which receptors with various degrees of specificity participate[46,50]. These include the B cell receptor (BCR) for an antigen, the mannose receptor, complement receptors, Fc receptors and probably scavenger receptors, all of which deliver the captured antigen to the various compartments of the endocytic pathway[51]. The BCR or surface immunoglobulin is a special case: not only does the BCR deliver an activation signal to B cells when it binds an antigen, it also internalizes the resulting immune complex for processing and presentation via MHC-II molecules as a call for $CD4^+$ T cell help[52,53]. Decoration of antigens with antibodies or covalent modification with components of the complement pathway (opsonization) can facilitate antigen acquisition via Fc and complement receptors. Macropinocytosis and micropinocytosis deliver captured extracellular fluid and its contents to endocytic compartments[54]. Finally, autophagy targets damaged cytoplasmic organelles such as mitochondria to a lysosomal compartment for degradation[55]. Autophagy can also capture cytoplasmic viral replication machineries and therefore the viral products they contain.

From an immunological perspective, the endolysosomal pathway should be viewed as a series of interconnected compartments, all of which can serve as a source of MHC ligands. Under homeostatic conditions, proteins that face or reside in the extracellular space, including membrane proteins, are continuously internalized and turned over through delivery to endolysosomal compartments. Consequently, in the absence of an infection with a pathogen, the source of peptide ligands for MHC-II molecules is the host's own proteins and possibly those obtained from commensals and ingested food[56,57]. The MHC-I-like molecules of the CD1 family (CD1a–CD1e in humans; CD1d in mice) behave like MHC-II molecules in their trafficking behaviour and traverse endolysosomal compartments[58] (FIG. 2c). Members of the CD1 family present a heterogeneous set of lipids, often

mimicked in experimental settings by the artificial CD1 ligand α-galactosyl ceramide[59]. CD1 ligands are acquired from an external source via lipoprotein receptors or generated endogenously through metabolic conversion of the appropriate precursor molecules[60].

MHC-I molecules sample the proteome of MHC-I-positive cells. Cytosolic proteins are also continuously turning over, albeit at widely different rates, and this involves proteolysis, predominantly by the ubiquitin–proteasome system[61]. All nucleated cells express MHC-I molecules or can be induced to do so[62]. Replication of viral pathogens occurs in nucleated cells, which are capable of nucleic acid and protein synthesis, unlike red blood cells or platelets. In the case of a virus infection, viral protein synthesis contributes to the proteome of the infected cell and serves as a source of antigens for MHC-I molecules[63]. Proteins in the extracellular space or its equivalents can also contribute to peptides presented by MHC-I molecules. This is referred to as 'cross-presentation' and is discussed in more detail later[64]. Membrane and secreted proteins that fail to pass quality control in the ER are expelled into the cytosol in a process referred to as 'dislocation' or 'ER-associated degradation'[65]. These discarded proteins likewise contribute to the pool of antigens sampled by MHC-I molecules.

Non-classical MHC-I molecules include HLA-E and its murine counterpart Qa-1, which predominantly present ligands derived from signal peptides[66]. These ligands are created in the ER through the combined action of signal peptidase and signal peptide peptidase.

Step 2: tagging antigens for proteolysis into peptides. MHC-II molecules mostly rely on lysosomal proteolysis to convert proteins into peptides suitable for presentation[11,56]. The MHC-II presentation pathway exploits physico-chemical means to tag proteins for degradation. The protease-rich compartments of the endocytic pathway are acidified to various degrees by H^+-pumping vacuolar ATPases. The resulting drop in pH can destabilize proteins not specifically designed to survive at acidic pH. Subtle changes in pH may also affect the outcome of proteolysis. What are believed to be the processing compartments in dendritic cells are less acidic than, for example, lysosomes in macrophages, a trait that may favour the generation of MHC-II ligands[67]. Endocytic compartments are also rich in reduced glutathione. This creates an environment that promotes reduction of disulfide bonds that stabilize many secreted and extracellular proteins. An interferon-γ (IFNγ)-inducible oxidoreductase, GILT, assists in the reduction of S–S bonds[68]. The combination of low pH and a reducing environment helps unfold proteins, making them better substrates for endolysosomal proteases.

For MHC-I molecules, those proteins that serve as a source of MHC-I ligands are tagged for degradation by ubiquitylation for entry into the ubiquitin–proteasome system. How do proteins enter that pathway? Errors in translation are manifold. Misacylation of tRNAs, misreading of a codon, translational frameshifts, failure to remove introns from pre-mRNA, inappropriate initiation of translation, premature translational stops, a failure of nascent chains to engage chaperones in a timely manner, missing partner subunits and oxidative damage can all create defective polypeptides (defective ribosomal products) that must be cleared to prevent them from exerting harmful effects[69–73]. Frameshifts not only yield defective translation products with a high probability of misfolding and susceptibility to proteolysis, but in a tumour setting they might also serve as a source of neoantigens because of the aberrant sequence introduced by the frameshift. Recall that secretory and membrane proteins that fail to fold correctly in the ER are subject to dislocation to the cytosol. The ubiquitin–proteasome machinery recognizes all these misfits and rapidly destroys them, presumably because they display structural features distinct from their properly folded counterparts[74,75].

The ubiquitin–proteasome system evolved to deal with the complexity of the proteome, which is reflected in the make-up of the machinery responsible for covalent attachment of the 76-residue protein modifier ubiquitin[76]. It comprises a cascade of E1, E2 and many (hundreds of) E3-type ubiquitin ligase activities[77]. The rules that govern protein (mis)folding and degradation apply to host proteins and pathogen-derived proteins alike. As is the case for MHC-II molecules, in the absence of an intruder, the sets of proteins tagged for degradation are the host's own proteins[4]. Upon infection with a virus, host cells will make errors in translation or folding in both cellular and viral proteins. Both will therefore contribute to the immunopeptidome. The built-in timer of turnover of MHC-I molecules further ensures that the bound immunopeptidome reflects changes in gene expression in the MHC-I-expressing cell[10].

Step 3: proteolysis. Proteolysis by lysosomal proteases and proteasomes generally involves the formation of acyl-enzyme reaction intermediates. Hydrolysis is the resolution of such covalent intermediates in a nucleophilic attack by a water molecule. However, peptides can also serve as the incoming nucleophile in a transpeptidation reaction, which then creates 'spliced' peptides[78] (FIG. 3a).

The generation of peptides destined for presentation by MHC-I molecules is mostly the purview of the ubiquitin–proteasome system. Not all ubiquitin-modified proteins are necessarily destroyed, and not all proteolysis by the ubiquitin–proteasome system is constitutive: for example, proteasomal degradation of proteins that control the cell cycle is tightly regulated[79]. The proteasome is a well-studied, abundant protein megacomplex, the core component of which consists of four rings of seven subunits each: a ring composed of seven α-subunits at the ends and two seven-membered rings, each with three catalytically active β-subunits, sandwiched between the α-subunit rings[80] (FIG. 3b,c). The ends of this core particle, in turn, form complexes — called 'cap complexes' — with numerous adapter proteins. These are tasked with recognition of ubiquitin chains and the removal of ubiquitin. They also include ATPases that assist in unfolding to thread substrate proteins destined for degradation through the central pore of the core particle. The three catalytically active β-subunits, β1, β2 and β5, line the chamber where proteolysis occurs. Proteasomal

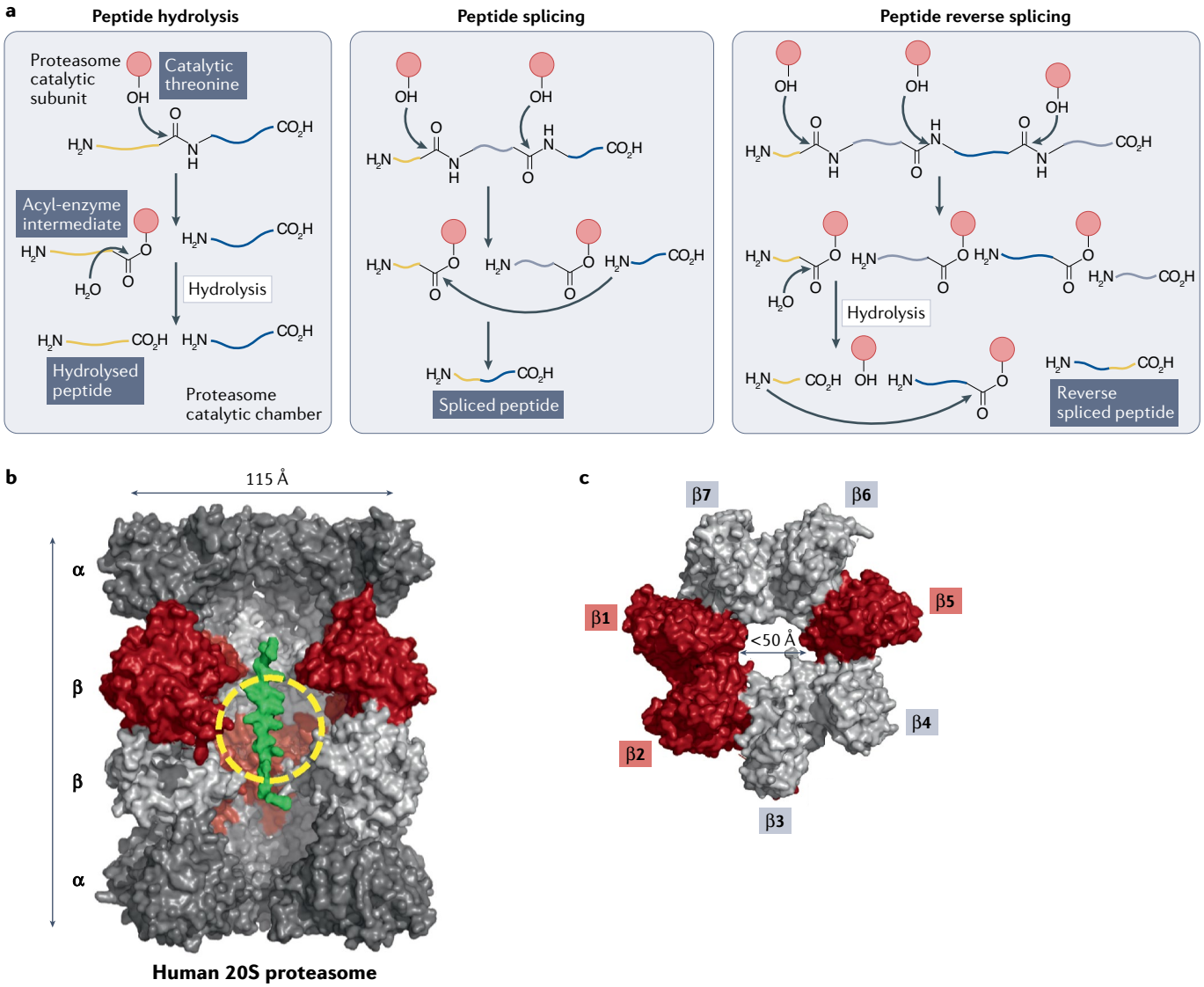

Fig. 3 | **Peptide splicing chemical reactions and anatomy of the proteasome. a** | Peptides entering the proteasome can undergo different types of enzymatic reactions. Simple hydrolysis of the acyl intermediate generated by the catalytic threonine leads to short peptides. However, because of the restricted volume of the catalytic chamber, a newly generated free peptide can react with an acyl intermediate, juxtaposing sequence elements that were at some distance from each other in the source protein. These 'spliced' fragments can switch order. **b** | The 3.5-Å cryo-electron microscopy structure of a 20S human proteasome complex. Side view of the proteasome showing the two seven-membered α and β rings. The C_2 symmetry axis and the dimensions are indicated. **c** | Top view of the α and a β ring. The catalytic subunits β1, β2 and β5 are in dark red. The dimension of the internal catalytic chamber is indicated (Protein Data Bank ID for parts **b** and **c** 5A0Q)[158].

proteolysis is processive: with few exceptions, no processing intermediates escape[81]. The rate of proteolysis must keep up with the rate of protein synthesis, in view of the numerous misfits that are the unavoidable consequence of errors in translation, protein folding and assembly. The confined space of the proteasome's catalytic chamber allows transpeptidation reactions — in addition to hydrolysis — to occur, because the local peptide concentration there is high. These reactions involve the formation of a covalent acyl-β subunit intermediate, resolved by attack via the NH_2 terminus of a peptide in the central cavity to generate *trans*-spliced peptides. The limited space in the catalytic chamber is unlikely to accommodate more than a single polypeptide during its destruction[82]. Peptides derived from one and the same polypeptide are therefore more likely to engage in *trans*-splicing than peptides derived from different polypeptides. The distance between the discontinuous sequences joined by peptide splicing can range from 4 to 40 residues[83,84]. The presence of an NH_2-terminal lysine residue for the COOH-terminal fusion partner is favoured and even allows the formation of an isopeptide bond if ε-NH_2 instead of α-NH_2 of the lysine residue serves as the nucleophile. Such *trans*-spliced peptides are bona fide, immunogenic MHC-I ligands[78]. Computational prediction tools that rely exclusively on 'standard' proteasomal cleavage patterns are unlikely to capture such events, the observation of which has relied until now on experiment.

The β1, β2 and β5 subunits characterize the 'standard' or constitutive version of the proteasome. Each subunit is associated with a preferred cleavage specificity:

trypsin-like (Lys and Arg), chymotrypsin-like (Phe, Tyr and Trp) and post-glutamyl peptidase (Glu and Asp) activities, respectively[82]. When the standard β catalytic subunits are replaced by MHC-encoded IFNγ-inducible β1i, β2i and β5i subunits, the resulting structure is referred to as the 'immunoproteasome'[85]. Both types of proteasome produce peptides that bracket the size range appropriate for recognition by the MHC-encoded peptide transporter (discussed later). A thymus-specific proteasome β-subunit (β5t) contributes to the immunopeptidome in the course of T cell development[85].

The immunoproteasome skews the distribution of peptides towards those with a higher probability of being proper MHC-I ligands[86]. Aminopeptidases can act on the peptides created by proteasomal proteolysis. Of particular relevance for the MHC-I presentation pathway are the ER-associated aminopeptidases ERAAP1 and ERAAP2, of which several allelic variants exist, each with a distinct substrate preference that impacts the immunopeptidome[87]. Glycoproteins subject to dislocation lose their N-linked glycans in the cytosol in an *N*-glycanase-catalysed reaction[88]. This converts the Asn residue that participates in the glycoamide bond to an Asp residue, a conversion essential for recognition of certain tumour antigens[13].

The MHC-I-like molecules (sometimes also referred to as 'class Ib molecules') that capture signal peptides and fragments derived from them do not rely on proteasomal proteolysis[89]. Their ligands may likewise undergo modification by aminopeptidases and carboxypeptidases. Members of the CD1 family rely on enzymes involved in lipid transport and metabolism[27]. Finally, MR1 relies on the provision of microbial metabolites for its exit from the ER[90]. These metabolites when bound to MR1 can have an inhibitory or a stimulatory effect on mucosal-associated invariant T cells[27] (FIG. 2d).

Protein degradation in endolysosomal compartments is catalysed by members of the cathepsin family[91,92]. These lysosomal enzymes belong to the class of aspartyl or cysteinyl proteases, with an acidic pH optimum. They can reach millimolar concentrations in the compartments where they reside. Aminopeptidases and carboxypeptidases can further trim the digestion products generated by cathepsins. Exceptionally, peptides serve as nucleophiles in a transpeptidation reaction, as for proteasomal proteolysis. This creates fused products composed of two distinct peptides, which may then be derived from completely unrelated proteins, given the relatively large volume of the compartments where endolysosomal proteolysis occurs. While there is evidence that such products exist and can be recognized by appropriately specific T cells[84,93], the prevalence of this mode of peptide production continues to be explored. During proteolysis, post-translational modifications such as N-linked and O-linked glycans can be modified or even removed completely by (amino)glycohydrolases[94].

Step 4: delivery of antigens to MHC molecules. Because the bulk of peptides destined for presentation by MHC-I molecules is generated cytosolically, transport across a membrane is required for peptides to reach the peptide-binding portion of newly synthesized MHC-I molecules, which resides in the lumen of the ER[4] (FIG. 4). Peptide translocation is entrusted to a heterodimeric member of the ATP-binding cassette-containing (ABC) family of transporters, the MHC-encoded transporter associated with antigen processing (TAP) complex[95–98]. The TAP1–TAP2 heterodimer transports peptides into the lumen of the ER at the expense of ATP hydrolysis and accepts peptides in the size range appropriate for binding to MHC-I molecules. The peptide's COOH terminus is an important determinant for acceptance by TAP and mostly fits the preference of the corresponding pocket in the MHC-I molecule's peptide-binding groove[99,100]. Peptides may require further trimming by ERAAP to enable their binding to the MHC-I molecules[10,101–103].

As mentioned already, MHC-II molecules sample peptides generated mostly in endocytic compartments, which are topologically equivalent to extracellular space. MHC-II molecules reach endocytic compartments by virtue of their association with the MHC-II-associated invariant chain (Ii), a type II membrane protein[10,104] (FIG. 5). The cytoplasmic NH_2-terminal tail of Ii contains sorting information that ensures delivery of the associated MHC-II molecules to endolysosomal compartments. The fraction of MHC-II molecules that escapes such sorting is delivered to the cell surface, from where it can be internalized into endosomal compartments, likewise by signals carried in Ii's cytoplasmic tail[46]. It is the delivery of MHC-II molecules to these compartments that ensures their exposure to the pool of possible peptide ligands. The MHC-II molecules themselves are quite resistant to proteolysis.

A cytokine-inducible asparagine endoprotease is involved in the generation of functional MHC-II molecules and appears to attack Ii preferentially[105]. Complete proteolysis of Ii culminates in the formation of an Ii-derived fragment, the peptide CLIP, which remains associated with the MHC-II molecules and prevents peptide binding until its removal under the agency of the HLA-DM molecule, as is discussed in the following section[11].

Step 5: loading of peptides onto MHC molecules. Loading of peptides onto MHC molecules requires the assistance of accessory proteins and should not be considered a spontaneous process. In the case of MHC-I molecules, the participation of a set of proteins known as the peptide-loading complex (PLC) is key[3]. The molecular structure of the PLC has been determined[2]. It minimally includes the MHC-I molecules themselves, TAP, tapasin (also known as TAP-binding protein (TAPBP)), ERp57 and the more loosely associated calreticulin and possibly calnexin. The last two are part of the constitutive glycoprotein quality control machinery[106] (FIG. 6a). The PLC then catalyses loading of peptides onto MHC-I molecules, a process in which tapasin plays a key role[107]. Tapasin preferentially interacts with MHC-I molecules that have yet to receive peptides; hence, its genetic ablation leads to enhanced surface display of peptide-receptive MHC-I molecules[12,108].

For MHC-II molecules, the removal of Ii and its remnants (CLIP) involves the action of the MHC-II homologue HLA-DM, which is structurally similar to classical MHC-II molecules but with the distinction of having

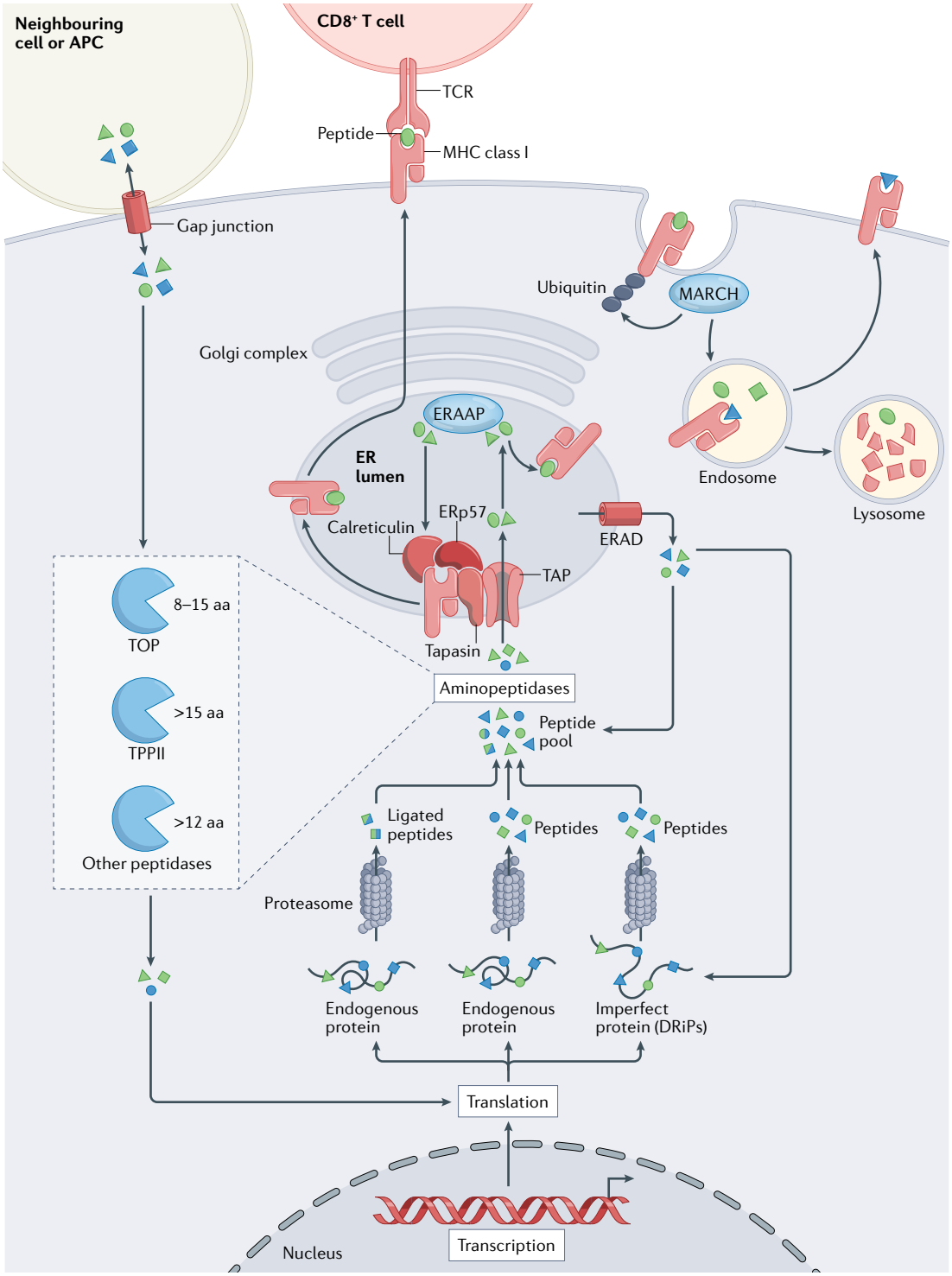

Fig. 4 | **Trafficking pathways for MHC class I molecules.** Antigens are derived from polypeptides with imperfections (defective ribosomal products (DRiPs)). DRiPs are degraded by the proteasome. The resulting peptides may be further processed by cytosolic aminopeptidases, such as thimet oligopeptidase (TOP) and tripeptidyl peptidase II (TPPII). Peptides can escape complete destruction by translocation into the endoplasmic reticulum (ER) lumen via the transporter associated with antigen processing (TAP), part of the peptide-loading complex. In the ER, peptides can associate directly with a major histocompatibility complex (MHC) class I molecule or undergo further enzymatic reactions involving the ER aminopeptidase associated with antigen processing (ERAAP) before binding to the MHC class I molecule via the peptide-loading complex. Peptide–MHC class I complexes are transported to the plasma membrane for antigen presentation to $CD8^+$ T cells. Some peptides may be transferred to neighbouring antigen-presenting cells (APCs) through gap junctions. Internalization and degradation of MHC class I molecules can occur in a ubiquitin-dependent manner. Some endocytosed MHC class I molecules may recycle to the surface after peptide exchange with endosomal peptides. aa, amino acids; ERAD, endoplasmic reticulum-associated degradation; TCR, T cell receptor. Figure adapted from REF.[10], Springer Nature Limited.

an occluded peptide-binding pocket[11] (FIG. 6c). MHC-II–HLA-DM interactions expel bound peptides and allow replacement with other suitable ligands until a peptide has bound that can no longer be displaced by HLA-DM: this peptide editing process converges on exceptionally stable and long-lived MHC-II–peptide complexes[109]. These often resist even elevated temperatures in the presence of sodium dodecyl sulfate. HLA-DO is another unconventional MHC-II molecule that modulates the activity of HLA-DM: its absence generates an altered MHC-II-bound peptide repertoire that can be recognized by HLA-DO-proficient T cells[110].

Step 6: display of MHC molecules at the cell surface. Most of the steps of antigen processing and presentation outlined so far are constitutive in cells that express MHC molecules. This has obvious implications for T cell development (in the positive and negative selection of thymocytes) and autoimmunity[111,112]. Because protein synthesis and protein turnover are continuously ongoing, surface MHC-I molecules provide a display of what a given cell's proteome can contribute to the immunopeptidome. It is on this massively diverse set of self-peptides that T cell repertoires are test-driven in the thymus to weed out potentially self-reactive T cells (negative selection), while allowing maturation of T cells with some measure of reactivity with the set of peptide-loaded MHC molecules that are present (positive selection)[112]. A given T cell receptor (TCR) can recognize many different MHC–peptide combinations, depending on their abundance and the affinity with which they bind to the TCR: T cells are inherently cross-reactive[113]. When a

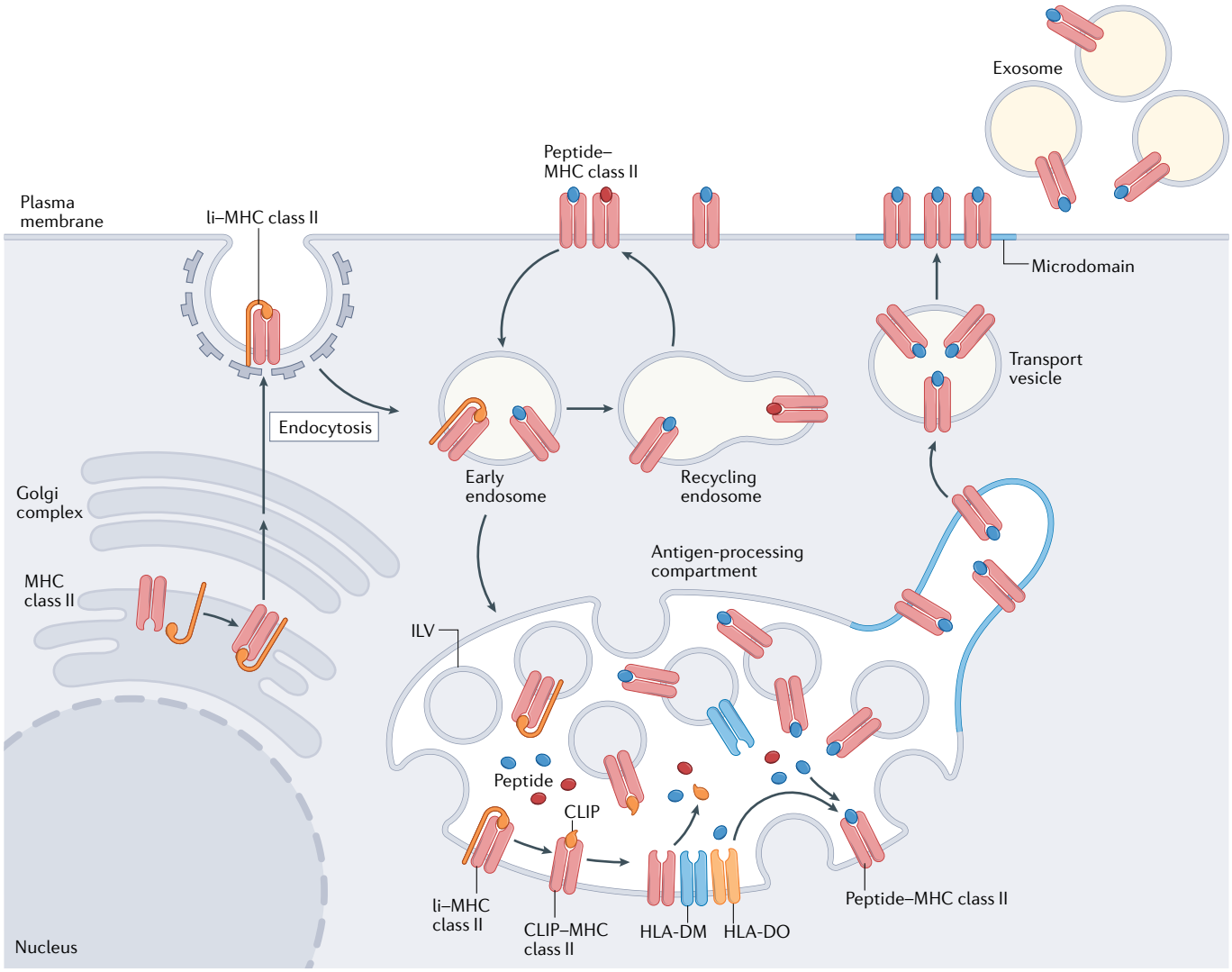

Fig. 5 | **Trafficking pathways for MHC class II molecules.** Major histocompatibility complex (MHC) class II molecules associate with the invariant chain (Ii) in the endoplasmic reticulum (ER) and travel to the cell surface via the Golgi apparatus. From there, the complex is delivered to endosomal compartments or reaches the cell surface, from where it may undergo internalization and delivery to antigen-processing compartments. Ii is then enzymatically degraded to yield a fragment of Ii (class II-associated invariant chain peptide (CLIP)) that remains bound to the peptide-binding groove of MHC class II. CLIP is removed in an exchange for an antigenic peptide in a reaction assisted by the chaperone HLA-DM. The activity of HLA-DM is regulated by HLA-DO, but the mechanism of regulation remains unknown. MHC class II molecules are then transported to the plasma membrane. Peptide–MHC class II complexes present on the cell surface may be internalized and undergo lysosomal degradation or recycle back to the plasma membrane. ILV, intraluminal vesicle. Figure adapted from REF.[46], Springer Nature Limited.

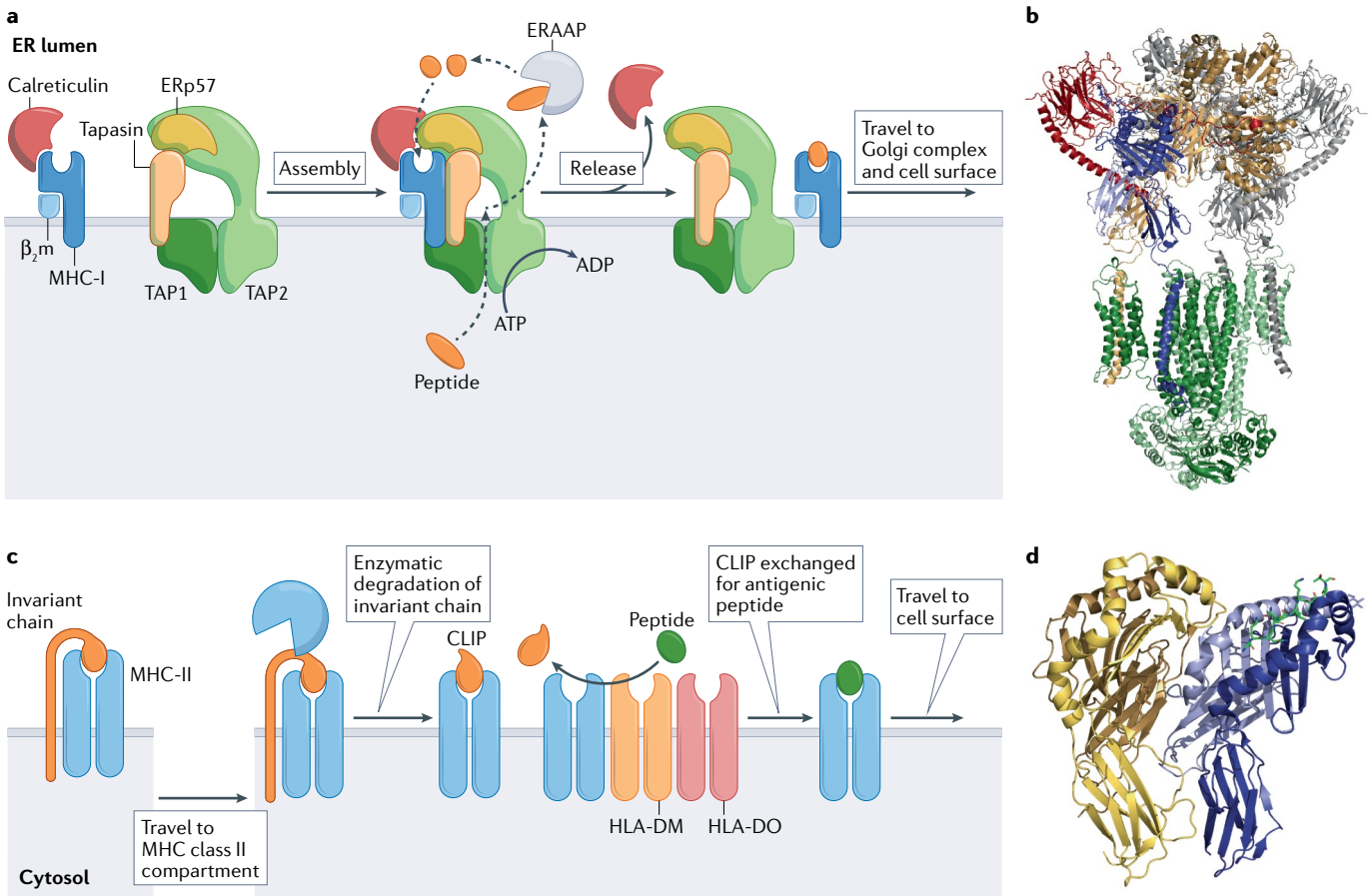

Fig. 6 | **Organization and structure of the MHC peptide-loading complexes. a** | The major histocompatibility complex (MHC) class I (MHC-I) peptide-loading complex (PLC). The MHC-I molecule is first recruited by calreticulin to an asymmetric PLC and then forms a symmetric, fully functional PLC[2]. The transporter associated with antigen processing (TAP) transports antigenic peptides from the cytosol into the endoplasmic reticulum (ER). Peptides may then be further modified by the ER aminopeptidase associated with antigen processing (ERAAP) and loaded onto MHC-I molecules. Stable MHC-I–peptide complexes are released from the PLC and travel to the cell surface via the Golgi complex. **b** | Cryo-electron microscopy analysis of the human MHC-I PLC[2,159]. Colours correspond to those used in part **a**. **c** | The MHC class II (MHC-II) PLC. The α and β chains of the MHC-II molecule are synthesized in the ER, where they form a stable complex with the invariant chain. This complex is then transported to the MHC-II compartment, where the invariant chain is degraded by proteases to yield the class II-associated invariant chain peptide (CLIP), with the latter still bound to the MHC-II pocket. The chaperone protein HLA-DM catalyses replacement of CLIP with an antigenic peptide. The newly formed stable complex is then moved to the cell surface. **d** | Crystal structure of the HLA-DM–MHC II complex (Protein Data Bank number 4GBX)[160]. Colours correspond to those used in part **c**. β_2m, β_2-microglobulin.

peptide–MHC complex not previously encountered makes its appearance in the periphery, as in the case of a virus-infected cell, or when a somatic mutation has occurred, or perhaps as the consequence of cellular stress, then T cells selected on a complex and diverse set of MHC molecules loaded with self-peptides might have sufficient affinity for these newly appearing peptide–MHC complexes to be properly activated[114]. Timely acquisition of a ligand is also important for the surface display of non-classical MHC-I molecules: the relevant ligands here include the signal sequences for HLA-E and Qa-1, and small molecule vitamin derivatives and microbial metabolites for MR1 (REFS[27,66]).

Dendritic cells retain most of their newly synthesized MHC-II molecules intracellularly, and they populate compartments where endolysosomal proteolysis and peptide loading occur[115]. When dendritic cells are activated (for example, following the sensing of viral or bacterial products through Toll-like receptors, these intracellular MHC-II molecules translocate to the cell surface for display to T cells[116,117]. Exposure to cytokines such as IFNγ can enhance cell surface display of MHC molecules, both for MHC-I and MHC-II, by a variety of mechanisms, including transcriptional control of MHC genes and their accessories, whereas IL-10 can inhibit translocation of MHC-II molecules to the cell surface[118].

The field advances: the toolbox

Genome-wide screens continue to refine our understanding of antigen processing and presentation through the identification of new molecular components, but without requiring drastic revisions of the global outline summarized in the six steps[119]. Most, if not all, components of the processing and presentation pathways are upregulated in the context of infectious events, in response to cytokines or by inflammatory stimuli. Many pathogens have developed appropriate countermeasures

to escape antigen processing and presentation[120,121]. These countermeasures can be mapped onto the six-step pathway outlined herein, but also include interference with transcription. Most prominently, members of the herpesvirus, adenovirus and poxvirus families encode proteins that bind and detain MHC-I molecules intracellularly or encode proteins that interfere with the activity of the MHC-encoded TAP molecules[121–124]. Some herpesviruses also encode proteins that structurally resemble MHC-I molecules[125]. These can serve as decoys for receptors on NK cells, which can recognize and kill cells of haematopoietic origin that have lost expression of classical MHC-I molecules[126]. Combined, these viral evasion strategies allow infected cells to avoid killing by both $CD8^+$ cytotoxic T cells and NK cells[120]. Cancer cells under immune selection by T cells are also prone to loss of expression of MHC-I molecules[127]. This Achilles heel can be exploited through therapy with chimeric antigen receptor T cells, which are not MHC restricted and are designed to recognize lineage markers or tumour-specific antigens[128].

For the experimentalist, there are tools to interfere with each of the six steps with various degrees of specificity and precision. These tools can be genetic (knockout mice and modified cell lines), physical (temperature and pH) or pharmacological (various drugs, including protease inhibitors, lysosomotropic agents such as proton pump inhibitors and inhibitors of endocytosis such as dynasore)[10]. Outcomes are measured by expression of MHC molecules themselves, and in functional assays (T cell activation, antibody production and protection against pathogens), both in vitro and in vivo. A few examples are given next: elimination of β_2-microglobulin blocks assembly of most MHC-I molecules and thus prevents antigen presentation[49]. Knockout of immunoproteasome subunits impedes MHC-I-restricted presentation but may do so selectively[129]. Elimination of ERAAP can produce subtle defects in the presentation of some antigens[129,130]. Ablation of tapasin allows the escape from the ER and surface display of MHC-I molecules deprived of peptides, which tends to destabilize them[108]. Proteasome inhibitors prevent the generation of the appropriate peptides and reduce expression of MHC-I molecules, as do TAP inhibitors, which preclude delivery of cytosolic peptides to the lumen of the ER[131,132]. A TAP deficiency, either genetically imposed or the result of viral interference, compromises peptide loading and surface expression of MHC-I molecules[63,133]. In the case of cancerous cells, knowledge of the defect responsible for loss of MHC expression may be important for therapeutic approaches[127]. Can the reduction or loss of MHC expression be counteracted and render tumours susceptible to attack by cytotoxic T cells, for example through administration of cytokines or by manipulation of epigenetic writers and erasers?

Cross-presentation delivers extracellular antigens to the MHC-I pathway[134]. While autophagy is well understood, thanks to the extensive genetic analyses that have clarified the process, cross-presentation remains less well defined from a cell biological/biochemical perspective[135]. The underlying mechanisms remain to be defined in molecular terms, are likely to be diverse[136–140] and will depend on the type of antigen-presenting cell. Proteins involved in intracellular trafficking such as RAB7 and SEC22 have been shown to be important for efficient cross-presentation[138,141]. Dendritic cells can engage in cross-presentation using the ER–Golgi intermediate compartment when TAP is dysfunctional[142]. Toll-like receptor signalling enhances cross-presentation by modulating accumulation of MHC-I molecules in the phagosomes that carry microbial components and their digestion products[143]. It is also important to distinguish between cross-presentation required to initiate a $CD8^+$ T cell response — a process that almost certainly involves specialized dendritic cells — and cross-presentation to allow target recognition by activated $CD8^+$ T cells[120,144]. Many different cell types may be capable of the latter. Autophagy captures damaged intracellular organelles such as mitochondria and targets them for lysosomal destruction, and so delivers endogenous proteins and proteins of cytoplasmic disposition to the MHC-II processing and presentation machinery[46].

Conclusions and practical implications

The field of antigen processing and presentation serves two distinct but overlapping communities of practitioners. Cell biologists and biochemists appreciate the well-equipped immunological and genetic toolbox to study protein quality control, protein degradation, glycoprotein trafficking and vesicular transport. We have a detailed understanding of the vesicular trafficking pathways that contribute to antigen processing and presentation, in particular for MHC-II molecules[119,145]. Supported by a wealth of structural data, we also have detailed pictures of the molecular machines that recognize and destroy damaged or defective proteins. The discovery of proteasomal *trans*-splicing owed much to the exquisite specificity and sensitivity of T cell recognition. Images of the PLC show protein assemblies comparable in their complexity to the molecular machines responsible for protein import into various organelles (FIG. 6). The network of interactors, with MHC-encoded molecules at its core, thus continues to expand.

Using ever more refined genetic approaches, new factors that regulate the rate and specificity of antigen processing and presentation will continue to be added to the list of known players[119]. Creative methods that deploy reporter cell lines to express in a genome-wide manner segments of antigens in a high-throughput configuration allow the detection of antigens that can stimulate $CD8^+$ T cells of any particular TCR configuration[146]. Because so many MHC molecules have been characterized by crystallography, the possibility of creating recombinant MHC molecules, loaded with defined peptides or expressed as 'single-chain' constructs comprising all component parts, has become a tool to better understand interactions of MHC molecules with TCRs[147]. Continued improvements in analytical methods such as mass spectrometry have shown their value in the field of immunopeptidomics[148]. The relative importance of peptide *trans*-splicing and the contribution of (post-translational) modifications to the side chains of peptides bound to MHC molecules are relatively recent, yet important refinements in our understanding of

antigen presentation. Cross-presentation clearly occurs, but how it occurs remains a contentious issue and raises a number of questions: are MHC-I molecules directed to the endolysosomal pathway for acquisition of peptides? Do antigens escape from the endolysosomal compartment to reach the cytosol (for example, through rupture of these compartments) or are specific transporters involved? If so, are these transporters in any way related to the TAP1–TAP2 complex and other ABC-type transporters, or should we anticipate the repurposing of other solute transporters? Do all of the above apply? Do different cell types use different mechanisms of cross-presentation?

Immunologists are more concerned with the consequences of (aberrant) antigen presentation and its practical applications. Tumour biologists, under the flag of cancer as a cellular genetic aberration, seek to identify neoepitopes that arise because of the mutational spectrum associated with particular tumour types[149]. Some of these mutations have long been known as primary drivers of malignancy, but there is wide appreciation that passenger mutations can serve as a source of neoepitopes as well[149]. Because peptides carrying such mutations were not presented to T cells during thymic development, the immune system can and does treat them as 'foreign'. Even proteins with no ostensible mutations or defects could serve as a source of an antigen if the levels of the relevant processed product exceed the threshold required to impose tolerance in the course of T cell selection[150]. Simply making a lot more of such peptide–MHC complexes, as might be caused by cellular stress or malignant transformation, may suffice to trigger T cells that escaped negative selection, yet that otherwise would have remained 'silent'.

Selection exerted by the immune system favours loss of MHC expression in tumours[8]. Even if a suitable set of neoepitopes can be predicted on the basis of whole-genome sequencing, such epitopes still require processing and presentation for their presence. Defects in these pathways may underlie defects in MHC expression. Upregulation of MHC molecules and of the components in the pathways that generate epitopes, as might be achieved with cytokine treatment[149,151] or perhaps even irradiation[152], may be viable therapeutic strategies. Time will tell whether single cell-based methods or knowledge of spatially defined expression of particular peptide–MHC complexes will add to the picture in a meaningful manner. To paraphrase Janeway[153], the field of antigen processing and presentation may be approaching the asymptote.

Published online 13 April 2022

1. Vyas, J. M., Van der Veen, A. G. & Ploegh, H. L. The known unknowns of antigen processing and presentation. *Nat. Rev. Immunol.* **8**, 607–618 (2008).
2. Blees, A. et al. Structure of the human MHC-I peptide-loading complex. *Nature* **551**, 525–528 (2017).
3. Trowitzsch, S. & Tampe, R. Multifunctional chaperone and quality control complexes in adaptive immunity. *Annu. Rev. Biophys.* **49**, 135–161 (2020).
4. Jensen, P. E. Recent advances in antigen processing and presentation. *Nat. Immunol.* **8**, 1041–1048 (2007).
5. Call, M. E. & Wucherpfennig, K. W. The T cell receptor: critical role of the membrane environment in receptor assembly and function. *Annu. Rev. Immunol.* **23**, 101–125 (2005).
6. Martin, F. & Chan, A. C. B cell immunobiology in disease: evolving concepts from the clinic. *Annu. Rev. Immunol.* **24**, 467–496 (2006).
7. Lizee, G. et al. Control of dendritic cell cross-presentation by the major histocompatibility complex class I cytoplasmic domain. *Nat. Immunol.* **4**, 1065–1073 (2003).
8. Reeves, E. & James, E. Antigen processing and immune regulation in the response to tumours. *Immunology* **150**, 16–24 (2017).
9. Fernando, M. M. et al. Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet.* **4**, e1000024 (2008).
10. Neefjes, J. et al. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat. Rev. Immunol.* **11**, 823–836 (2011).
11. Blum, J. S., Wearsch, P. A. & Cresswell, P. Pathways of antigen processing. *Annu. Rev. Immunol.* **31**, 443–473 (2013).
12. Cresswell, P. et al. Mechanisms of MHC class I-restricted antigen processing and cross-presentation. *Immunol. Rev.* **207**, 145–157 (2005).
13. van Kasteren, S. I. et al. Chemical biology of antigen presentation by MHC molecules. *Curr. Opin. Immunol.* **26**, 21–31 (2014).
14. Kyewski, B. & Klein, L. A central role for central tolerance. *Annu. Rev. Immunol.* **24**, 571–606 (2006).
15. Lee, J. W. et al. Peripheral antigen display by lymph node stroma promotes T cell tolerance to intestinal self. *Nat. Immunol.* **8**, 181–190 (2007).
16. Kyewski, B. & Derbinski, J. Self-representation in the thymus: an extended view. *Nat. Rev. Immunol.* **4**, 688–698 (2004).
17. Mayassi, T. et al. A multilayered immune system through the lens of unconventional T cells. *Nature* **595**, 501–510 (2021).
18. Adams, E. J. & Luoma, A. M. The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules. *Annu. Rev. Immunol.* **31**, 529–561 (2013).
19. Stern, L. J. & Wiley, D. C. Antigenic peptide binding by class I and class II histocompatibility proteins. *Behring Inst. Mitt.* **2**, 245–251 (1994).
20. Tumer, G., Simpson, B. and Roberts, T. K. Genetics, Human Major Histocompatibility Complex (MHC) (StatPearls, 2021)
21. Choo, S. Y. The HLA system: genetics, immunology, clinical testing, and clinical implications. *Yonsei Med. J.* **48**, 11–23 (2007).
22. Shiina, T. et al. The HLA genomic loci map: expression, interaction, diversity and disease. *J. Hum. Genet.* **54**, 15–39 (2009).
23. Shiina, T. et al. Comparative genomics of the human, macaque and mouse major histocompatibility complex. *Immunology* **150**, 127–138 (2017).
24. Matsumura, M. et al. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science* **257**, 927–934 (1992).
25. Bouvier, M. & Wiley, D. C. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science* **265**, 398–402 (1994).
26. Zacharias, M. & Springer, S. Conformational flexibility of the MHC class I α_1-α_2 domain in peptide bound and free states: a molecular dynamics simulation study. *Biophys. J.* **87**, 2203–2214 (2004).
27. Van Rhijn, I. et al. Lipid and small-molecule display by CD1 and MR1. *Nat. Rev. Immunol.* **15**, 643–654 (2015).
28. Silva, A. P. D. & Gallardo, R. A. The chicken MHC: insights into genetic resistance, immunity, and inflammation following infectious bronchitis virus infections. *Vaccines (Basel)* **8**, 637 (2020).
29. Miller, M. M. & Taylor, R. L. Jr. Brief review of the chicken major histocompatibility complex: the genes, their distribution on chromosome 16, and their contributions to disease resistance. *Poult. Sci.* **95**, 375–392 (2016).
30. Chicz, R. M. et al. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* **358**, 764–768 (1992).
31. Abualrous, E. T., Sticht, J. & Freund, C. Major histocompatibility complex (MHC) class I and class II proteins: impact of polymorphism on antigen presentation. *Curr. Opin. Immunol.* **70**, 95–104 (2021).
32. Flajnik, M. F. & Kasahara, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* **11**, 47–59 (2010).
33. Bontrop, R. E. Comparative genetics of MHC polymorphisms in different primate species: duplications and deletions. *Hum. Immunol.* **67**, 388–397 (2006).
34. Kasahara, M. The chromosomal duplication model of the major histocompatibility complex. *Immunol. Rev.* **167**, 17–32 (1999).
35. Robinson, J. et al. IPD-IMGT/HLA Database. *Nucleic Acids Res.* **48**, D948–D955 (2020).
36. Kelly, A. & Trowsdale, J. Genetics of antigen processing and presentation. *Immunogenetics* **71**, 161–170 (2019).
37. Tomasec, P. et al. Surface expression of HLA-E, an inhibitor of natural killer cells, enhanced by human cytomegalovirus gpUL40. *Science* **287**, 1031 (2000).
38. Larsen, M. H. & Hviid, T. V. Human leukocyte antigen-G polymorphism in relation to expression, function, and disease. *Hum. Immunol.* **70**, 1026–1034 (2009).
39. Wu, H. L. et al. The role of MHC-E in T cell immunity is conserved among humans, rhesus macaques, and cynomolgus macaques. *J. Immunol.* **200**, 49–60 (2018).
40. Dulberger, C. L. et al. Human leukocyte antigen F presents peptides and regulates immunity through interactions with NK cell receptors. *Immunity* **46**, 1018–1029 e7 (2017).
41. Creech, A. L. et al. The role of mass spectrometry and proteogenomics in the advancement of HLA epitope prediction. *Proteomics* **18**, e1700259 (2018).
42. Thomas, C. & Tampe, R. MHC I chaperone complexes shaping immunity. *Curr. Opin. Immunol.* **58**, 9–15 (2019).
43. Wieczorek, M. et al. Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation. *Front. Immunol.* **8**, 292 (2017).
44. Truong, H. V. & Sgourakis, N. G. Dynamics of MHC-I molecules in the antigen processing and presentation pathway. *Curr. Opin. Immunol.* **70**, 122–128 (2021).
45. Zaitoua, A. J., Kaur, A. & Raghavan, M. Variations in MHC class I antigen presentation and immunopeptidome selection pathways. *F1000Res.* https://doi.org/10.12688/f1000research.26935.1 (2020).

46. Roche, P. A. & Furuta, K. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat. Rev. Immunol.* **15**, 203–216 (2015).
47. Jiang, J., Natarajan, K. & Margules, D. H. MHC molecules, T cell receptors, natural killer cell receptors, and viral immunoevasins-key elements of adaptive and innate immunity. *Adv. Exp. Med. Biol.* **1172**, 21–62 (2019).
48. Cosgrove, D. et al. Mice lacking MHC class II molecules. *Cell* **66**, 1051–1066 (1991).
49. Koller, B. H. et al. Normal development of mice deficient in β_2M, MHC class I proteins, and CD8$^+$ T cells. *Science* **248**, 1227–1230 (1990).
50. Stuart, L. M. & Ezekowitz, R. A. Phagocytosis: elegant complexity. *Immunity* **22**, 539–550 (2005).
51. Tse, S. M. et al. Differential role of actin, clathrin, and dynamin in Fc gamma receptor-mediated endocytosis and phagocytosis. *J. Biol. Chem.* **278**, 3331–3338 (2003).
52. Siemasko, K. et al. Cutting edge: signals from the B lymphocyte antigen receptor regulate MHC class II containing late endosomes. *J. Immunol.* **160**, 5203–5208 (1998).
53. Lankar, D. et al. Dynamics of major histocompatibility complex class II compartments during B cell receptor-mediated cell activation. *J. Exp. Med.* **195**, 461–472 (2002).
54. Lim, J. P. & Gleeson, P. A. Macropinocytosis: an endocytic pathway for internalising large gulps. *Immunol. Cell Biol.* **89**, 836–843 (2011).
55. Schmid, D., Pypaert, M. & Munz, C. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* **26**, 79–92 (2007).
56. Trombetta, E. S. & Mellman, I. Cell biology of antigen processing in vitro and in vivo. *Annu. Rev. Immunol.* **23**, 975–1028 (2005).
57. Wilson, N. S., El-Sukkari, D. & Villadangos, J. A. Dendritic cells constitutively present self antigens in their immature state in vivo and regulate antigen presentation by controlling the rates of MHC class II synthesis and endocytosis. *Blood* **103**, 2187–2195 (2004).
58. Chancellor, A., Gadola, S. D. & Mansour, S. The versatility of the CD1 lipid antigen presentation pathway. *Immunology* **154**, 196–203 (2018).
59. Koch, M. et al. The crystal structure of human CD1d with and without alpha-galactosylceramide. *Nat. Immunol.* **6**, 819–826 (2005).
60. Wu, D., Fujio, M. & Wong, C. H. Glycolipids as immunostimulating agents. *Bioorg. Med. Chem.* **16**, 1073–1083 (2008).
61. Rock, K. L. & Goldberg, A. L. Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* **17**, 739–779 (1999).
62. Hewitt, E. W. The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology* **110**, 163–169 (2003).
63. Hill, A. & Ploegh, H. Getting the inside out: the transporter associated with antigen processing (TAP) and the presentation of viral antigen. *Proc. Natl Acad. Sci. USA* **92**, 341–343 (1995).
64. Cruz, F. M. et al. The biology and underlying mechanisms of cross-presentation of exogenous antigens on MHC-I molecules. *Annu. Rev. Immunol.* **35**, 149–176 (2017).
65. Hughes, E. A., Hammond, C. & Cresswell, P. Misfolded major histocompatibility complex class I heavy chains are translocated into the cytoplasm and degraded by the proteasome. *Proc. Natl Acad. Sci. USA* **94**, 1896–1901 (1997).
66. van Hall, T. et al. The other Janus face of Qa-1 and HLA-E: diverse peptide repertoires in times of stress. *Microbes Infect.* **12**, 910–918 (2010).
67. Delamarre, L. et al. Differential lysosomal proteolysis in antigen-presenting cells determines antigen fate. *Science* **307**, 1630–1634 (2005).
68. West, L. C. & Cresswell, P. Expanding roles for GILT in immunity. *Curr. Opin. Immunol.* **25**, 103–108 (2013).
69. Li, M. et al. Widespread RNA and DNA sequence differences in the human transcriptome. *Science* **333**, 53–58 (2011).
70. Yewdell, J. W. & Hickman, H. D. New lane in the information highway: alternative reading frame peptides elicit T cells with potent antiretrovirus activity. *J. Exp. Med.* **204**, 2501–2504 (2007).
71. Berglund, P. et al. Viral alteration of cellular translational machinery increases defective ribosomal products. *J. Virol.* **81**, 7220–7229 (2007).
72. Netzer, N. et al. Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* **462**, 522–526 (2009).
73. Dolan, B. P. et al. Distinct pathways generate peptides from defective ribosomal products for CD8+T cell immunosurveillance. *J. Immunol.* **186**, 2065–2072 (2011).
74. Schubert, U. et al. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **404**, 770–774 (2000).
75. Reits, E. A. et al. The major substrates for TAP in vivo are derived from newly synthesized proteins. *Nature* **404**, 774–778 (2000).
76. Welchman, R. L., Gordon, C. & Mayer, R. J. Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat. Rev. Mol. Cell Biol.* **6**, 599–609 (2005).
77. Ciechanover, A. Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nat. Rev. Mol. Cell Biol.* **6**, 79–87 (2005).
78. Vigneron, N. et al. Peptide splicing by the proteasome. *J. Biol. Chem.* **292**, 21170–21179 (2017).
79. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).
80. Bard, J. A. M. et al. Structure and function of the 26S proteasome. *Annu. Rev. Biochem.* **87**, 697–724 (2018).
81. Dikic, I. Proteasomal and autophagic degradation systems. *Annu. Rev. Biochem.* **86**, 193–224 (2017).
82. Tanaka, K., Mizushima, T. & Saeki, Y. The proteasome: molecular machinery and pathophysiological roles. *Biol. Chem.* **393**, 217–234 (2012).
83. Hanada, K., Yewdell, J. W. & Yang, J. C. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252–256 (2004).
84. Vigneron, N. et al. An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587–590 (2004).
85. Murata, S. et al. The immunoproteasome and thymoproteasome: functions, evolution and human disease. *Nat. Immunol.* **19**, 923–931 (2018).
86. Tanaka, K. & Kasahara, M. The MHC class I ligand-generating system: roles of immunoproteasomes and the interferon-gamma-inducible proteasome activator PA28. *Immunol. Rev.* **163**, 161–176 (1998).
87. Hattori, A. & Tsujimoto, M. Endoplasmic reticulum aminopeptidases: biochemistry, physiology and pathology. *J. Biochem.* **154**, 219–228 (2013).
88. Suzuki, T., Huang, C. & Fujihira, H. The cytoplasmic peptide:N-glycanase (NGLY1) - structure, expression and cellular functions. *Gene* **577**, 1–7 (2016).
89. Rodgers, J. R. & Cook, R. G. MHC class Ib molecules bridge innate and acquired immunity. *Nat. Rev. Immunol.* **5**, 459–471 (2005).
90. McWilliam, H. E. G. et al. Endoplasmic reticulum chaperones stabilize ligand-receptive MR1 molecules for efficient presentation of metabolite antigens. *Proc. Natl Acad. Sci. USA* **117**, 24974–24985 (2020).
91. Villadangos, J. A. et al. Proteases involved in MHC class II antigen presentation. *Immunol. Rev.* **172**, 109–120 (1999).
92. Honey, K. & Rudensky, A. Y. Lysosomal cysteine proteases regulate antigen presentation. *Nat. Rev. Immunol.* **3**, 472–482 (2003).
93. Wang, Y. et al. How C-terminal additions to insulin B-chain fragments create superagonists for T cells in mouse and human type 1 diabetes. *Sci. Immunol.* **4**, eaav7517 (2019).
94. Winchester, B. Lysosomal metabolism of glycoproteins. *Glycobiology* **15**, 1R–15R (2005).
95. Parcej, D. & Tampe, R. ABC proteins in antigen translocation and viral inhibition. *Nat. Chem. Biol.* **6**, 572–580 (2010).
96. Eggensperger, S. & Tampe, R. The transporter associated with antigen processing: a key player in adaptive immunity. *Biol. Chem.* **396**, 1059–1072 (2015).
97. Thomas, C. & Tampe, R. Structural and mechanistic principles of ABC transporters. *Annu. Rev. Biochem.* **89**, 605–636 (2020).
98. Grossmann, N. et al. Mechanistic determinants of the directionality and energetics of active export by a heterodimeric ABC transporter. *Nat. Commun.* **5**, 5419 (2014).
99. Gubler, B. et al. Substrate selection by transporters associated with antigen processing occurs during peptide binding to TAP. *Mol. Immunol.* **35**, 427–433 (1998).
100. Uebel, S. et al. Recognition principle of the TAP transporter disclosed by combinatorial peptide libraries. *Proc. Natl Acad. Sci. USA* **94**, 8976–8981 (1997).
101. Serwold, T. et al. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* **419**, 480–483 (2002).
102. Saric, T. et al. An IFN-γ-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol.* **3**, 1169–1176 (2002).
103. York, I. A. et al. The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nat. Immunol.* **3**, 1177–1184 (2002).
104. Roche, P. A. & Cresswell, P. Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature* **345**, 615–618 (1990).
105. Reich, M. et al. Invariant chain processing is independent of cathepsin variation between primary human B cells/dendritic cells and B-lymphoblastoid cells. *Cell Immunol.* **269**, 96–103 (2011).
106. Williams, D. B. Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *J. Cell Sci.* **119**, 615–623 (2006).
107. Raghavan, M. et al. MHC class I assembly: out and about. *Trends Immunol.* **29**, 436–443 (2008).
108. Garbi, N. et al. Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. *Nat. Immunol.* **1**, 234–238 (2000).
109. Busch, R. et al. Achieving stability through editing and chaperoning: regulation of MHC class II peptide binding and expression. *Immunol. Rev.* **207**, 242–260 (2005).
110. Poluektov, Y. O., Kim, A. & Sadegh-Nasseri, S. HLA-DO and its role in MHC class II antigen presentation. *Front. Immunol.* **4**, 260 (2013).
111. Sollid, L. M., Pos, W. & Wucherpfennig, K. W. Molecular mechanisms for contribution of MHC molecules to autoimmune diseases. *Curr. Opin. Immunol.* **31**, 24–30 (2014).
112. Klein, L. et al. Antigen presentation in the thymus for positive selection and central tolerance induction. *Nat. Rev. Immunol.* **9**, 833–844 (2009).
113. Baker, B. M. et al. Structural and dynamic control of T-cell receptor specificity, cross-reactivity, and binding mechanism. *Immunol. Rev.* **250**, 10–31 (2012).
114. Smith-Garvin, J. E., Koretzky, G. A. & Jordan, M. S. T cell activation. *Annu. Rev. Immunol.* **27**, 591–619 (2009).
115. ten Broeke, T., Wubbolts, R. & Stoorvogel, W. MHC class II antigen presentation by dendritic cells regulated through endosomal sorting. *Cold Spring Harb. Perspect. Biol.* **5**, a016873 (2013).
116. Watts, C., West, M. A. & Zaru, R. TLR signalling regulated antigen presentation in dendritic cells. *Curr. Opin. Immunol.* **22**, 124–130 (2010).
117. Bhati, M. et al. The versatility of the αβ T-cell antigen receptor. *Protein Sci.* **23**, 260–272 (2014).
118. Mittal, S. K. & Roche, P. A. Suppression of antigen presentation by IL-10. *Curr. Opin. Immunol.* **34**, 22–27 (2015).
119. Paul, P. et al. A Genome-wide multidimensional RNAi screen reveals pathways controlling MHC class II antigen presentation. *Cell* **145**, 268–283 (2011).
120. van de Weijer, M. L., Luteijn, R. D. & Wiertz, E. J. Viral immune evasion: lessons in MHC class I antigen presentation. *Semin. Immunol.* **27**, 125–137 (2015).
121. Loureiro, J. & Ploegh, H. L. Antigen presentation and the ubiquitin-proteasome system in host-pathogen interactions. *Adv. Immunol.* **92**, 225–305 (2006).
122. Bauer, D. & Tampe, R. Herpes viral proteins blocking the transporter associated with antigen processing TAP — from genes to function and structure. *Curr. Top. Microbiol. Immunol.* **269**, 87–99 (2002).
123. Berry, R. et al. Modulation of innate and adaptive immunity by cytomegaloviruses. *Nat. Rev. Immunol.* **20**, 113–127 (2020).
124. Lin, J. et al. A negative feedback modulator of antigen processing evolved from a frameshift in the cowpox virus genome. *PLoS Pathog.* **10**, e1004554 (2014).
125. Browne, H. et al. A complex between the MHC class I homologue encoded by human cytomegalovirus and β_2 microglobulin. *Nature* **347**, 770–772 (1990).
126. Farrell, H. E. et al. Inhibition of natural killer cells by a cytomegalovirus MHC class I homologue in vivo. *Nature* **386**, 510–514 (1997).
127. Dhatchinamoorthy, K., Colbert, J. D. & Rock, K. L. Cancer immune evasion through loss of MHC class I antigen presentation. *Front. Immunol.* **12**, 636568 (2021).
128. Hong, M., Clubb, J. D. & Chen, Y. Y. Engineering CAR-T cells for next-generation cancer therapy. *Cancer Cell* **38**, 473–488 (2020).
129. Kincaid, E. Z. et al. Mice completely lacking immunoproteasomes show major changes in antigen presentation. *Nat. Immunol.* **13**, 129–135 (2011).
130. Yan, J. et al. In vivo role of ER-associated peptidase activity in tailoring peptides for presentation by MHC class Ia and class Ib molecules. *J. Exp. Med.* **203**, 647–659 (2006).

131. Miller, Z. et al. Inhibitors of the immunoproteasome: current status and future directions. *Curr. Pharm. Des.* **19**, 4140–4151 (2013).
132. Rock, K. L. et al. 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 (1994).
133. Van Kaer, L. et al. *TAP1* mutant mice are deficient in antigen presentation, surface class I molecules, and $CD4^{-}8^{+}$ T cells. *Cell* **71**, 1205–1214 (1992).
134. Colbert, J. D., Cruz, F. M. & Rock, K. L. Cross-presentation of exogenous antigens on MHC I molecules. *Curr. Opin. Immunol.* **64**, 1–8 (2020).
135. Mizushima, N. & Komatsu, M. Autophagy: renovation of cells and tissues. *Cell* **147**, 728–741 (2011).
136. Parekh, V. V. et al. Autophagy-related protein Vps34 controls the homeostasis and function of antigen cross-presenting $CD8\alpha^{+}$ dendritic cells. *Proc. Natl Acad. Sci. USA* **114**, E6371–E6380 (2017).
137. Mintern, J. D. et al. Differential use of autophagy by primary dendritic cells specialized in cross-presentation. *Autophagy* **11**, 906–917 (2015).
138. Blander, J. M. Regulation of the cell biology of antigen cross-presentation. *Annu. Rev. Immunol.* **36**, 717–753 (2018).
139. Blander, J. M. The comings and goings of MHC class I molecules herald a new dawn in cross-presentation. *Immunol. Rev.* **272**, 65–79 (2016).
140. Theisen, D. J. et al. WDFY4 is required for cross-presentation in response to viral and tumor antigens. *Science* **362**, 694–699 (2018).
141. Cebrian, I. et al. Sec22b regulates phagosomal maturation and antigen crosspresentation by dendritic cells. *Cell* **147**, 1355–1368 (2011).
142. Barbet, G. et al. TAP dysfunction in dendritic cells enables noncanonical cross-presentation for T cell priming. *Nat. Immunol.* **22**, 497–509 (2021).
143. Nair-Gupta, P. et al. TLR signals induce phagosomal MHC-I delivery from the endosomal recycling compartment to allow cross-presentation. *Cell* **158**, 506–521 (2014).
144. Segura, E. & Amigorena, S. Cross-presentation in mouse and human dendritic cells. *Adv. Immunol.* **127**, 1–31 (2015).
145. Jongsma, M. L. et al. An ER-associated pathway defines endosomal architecture for controlled cargo transport. *Cell* **166**, 152–166 (2016).
146. Kula, T. et al. T-scan: a genome-wide method for the systematic discovery of T cell epitopes. *Cell* **178**, 1016–1028 e13 (2019).
147. Woodham, A. W. et al. In vivo detection of antigen-specific $CD8^{+}$ T cells by immuno-positron emission tomography. *Nat. Methods* **17**, 1025–1032 (2020).
148. Stopfer, L. E. et al. Multiplexed relative and absolute quantitative immunopeptidomics reveals MHC I repertoire alterations induced by CDK4/6 inhibition. *Nat. Commun.* **11**, 2760 (2020).
149. Jhunjhunwala, S., Hammer, C. & Delamarre, L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. *Nat. Rev. Cancer* **21**, 298–312 (2021).
150. Bassani-Sternberg, M. et al. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell Proteom.* **14**, 658–673 (2015).
151. Dunn, G. P., Koebel, C. M. & Schreiber, R. D. Interferons, immunity and cancer immunoediting. *Nat. Rev. Immunol.* **6**, 836–848 (2006).
152. Bhalla, N., Brooker, R. & Brada, M. Combining immunotherapy and radiotherapy in lung cancer. *J. Thorac. Dis.* **10**, S1447–S1460 (2018).
153. Janeway, C. A. Jr Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.* **54**, 1–13 (1989).
154. Borbulevych, O. Y., Piepenbrink, K. H. & Baker, B. M. Conformational melding permits a conserved binding geometry in TCR recognition of foreign and self molecular mimics. *J. Immunol.* **186**, 2950–2958 (2011).
155. Li, Y. et al. Structural basis for the presentation of tumor-associated MHC class II-restricted phosphopeptides to $CD4^{+}$ T cells. *J. Mol. Biol.* **399**, 596–603 (2010).
156. Zajonc, D. M. et al. Structure and function of a potent agonist for the semi-invariant natural killer T cell receptor. *Nat. Immunol.* **6**, 810–818 (2005).
157. Patel, O. et al. Recognition of vitamin B metabolites by mucosal-associated invariant T cells. *Nat. Commun.* **4**, 2142 (2013).
158. da Fonseca, P. C. & Morris, E. P. Cryo-EM reveals the conformation of a substrate analogue in the human 20S proteasome core. *Nat. Commun.* **6**, 7573 (2015).
159. Fisette, O., Schroder, G. F. & Schafer, L. V. Atomistic structure and dynamics of the human MHC-I peptide-loading complex. *Proc. Natl Acad. Sci. USA* **117**, 20597–20606 (2020).
160. Pos, W. et al. Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. *Cell* **151**, 1557–1568 (2012).

Acknowledgements
The authors thank P. Cresswell and J. Strominger for reading an earlier version of the manuscript. They also thank members of the Ploegh laboratory for their input.

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Immunology thanks L. Fugger, R. Tampe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Supplementary information
The online version contains supplementary material available at https://doi.org/10.1038/s41577-022-00707-2.