Antigen Processing

Johan K Sandberg, *Karolinska Institute, Stockholm, Sweden*
Rickard Glas, *Harvard University, Boston, Massachusetts, USA*

Processing of antigenic proteins into short peptide fragments is required for antigen recognition by T lymphocytes. Antigen recognition is mediated through cell surface presentation of the processed peptide fragments by major histocompatibility complex (MHC) gene products.

Secondary article

Article Contents

Requirement for Antigen Processing

Infectious organisms such as bacteria and viruses produce proteins that are foreign (nonself) to the host immune system, and therefore serve as potential target structures for an immune response. The humoral immune response is effective against extracellular antigens, and is mediated by B lymphocyte-derived antibodies which bind the native conformation of the invading microorganism's proteins. Thus, antibody-mediated recognition of native protein requires no antigen processing. In contrast, T lymphocytes recognize fragments of proteins that are processed into short peptides by proteases within the antigen-presenting cell (APC). These peptides are presented at the cell surface bound to major histocompatibility complex (MHC) molecules and represent a sampling of the intracellular and extracellular milieu. The MHC-presented peptides are scanned and recognized by circulating T lymphocytes via their antigen-specific cell surface T-cell receptors (TCRs). For these reasons, processing of antigens is crucial for activation of T lymphocytes during infection, and is therefore the key to a successful adaptive immune response.

Antigen processing and presentation occurs by two major pathways: the MHC class I and the MHC class II pathway. Antigens presented by MHC class I activate CD8+ cytotoxic T lymphocytes (CTLs) to kill infected cells, whereas MHC class II-presented antigens activate CD4+ helper T lymphocytes to perform their key roles in controlling humoral, CTL-mediated and inflammatory immune responses.

Uptake of Exogenous Antigen

MHC-dependent immune responses are induced by specialized APCs, which endocytose exogenous antigens for presentation on the cell surface by MHC molecules (reviewed in Watts, 1997). Specialized APCs utilize three main mechanisms for endocytosis: phagocytosis and pinocytosis by macrophages and dendritic cells, and specific receptor-mediated uptake by B lymphocytes.

Phagocytosis and pinocytosis

The phagocytic pathway of antigen uptake is a form of endocytosis in which large particles such as microorganisms are ingested by macrophages into large endocytic vesicles called phagosomes. Phagocytosis is dependent on binding of antibodies, components of the complement system or liver-derived acute-phase proteins to the surface of the microorganism. The process of coating a pathogen with these host proteins is called opsonization. Phagocytosis is initiated by binding of opsonized particles to receptors for the constant part of an antibody (Fc receptors), complement receptors or receptors for acute-phase proteins on the surface of a phagocytic cell. These particles are engulfed to form a phagosome, which subsequently fuses with lysosomes to form a phagolysosome. This compartment possesses a low pH coupled with lysosomal proteases which attack the phagocytosed particle. Macrophages are probably the most important phagocytic APCs *in vivo*, although several cell types, such as dendritic cells, B lymphocytes and neutrophils, are also capable of phagocytosis.

Antigen uptake through pinocytosis occurs by formation of small clathrin-coated pits at the plasma membrane. By internalization and formation of pinosomes, this mechanism mediates uptake of soluble proteins. Dendritic cells exhibit a specialized form of constitutive macropinocytosis, which make them particularly efficient at filtering the extracellular fluid.

Receptor-mediated uptake by B lymphocytes

Membrane-bound immunoglobulins (mIg) on the cell surface of B lymphocytes are associated with the disulfide-linked Igα–Igβ heterodimer to form the B-cell receptor (BCR). An antigen that binds to the BCR is endocytosed via clathrin-coated pits, and uptake requires signalling by either the cytoplasmic tail of the mIg or the Igα–Igβ dimer. After leaving the plasma membrane, clathrin-coated vesicles lose their membrane coats and fuse with endocytic compartments. Antigen-derived epitopes are processed in B lymphocytes and presented in the

context of the class II MHC to helper T lymphocytes. In turn, activated helper T lymphocytes activate the antigen-specific B lymphocyte to produce antibodies. Owing to the antigen specificity and clonal distribution of the BCR, antigen capture by this mechanism is an important step in controlling the specificity of humoral immune responses. Only antigen-specific B lymphocytes acquire help from T lymphocytes.

Endocytic compartments involved in antigen processing

The endocytic system consists of an interactive network of vacuoles, vesicles and tubules that also interact with the Golgi apparatus and the plasma membrane. Although a precise definition of a subcompartment that processes extracellular antigens is unclear, a few specific compartments are believed to be important. Endocytosed antigens are delivered to the early endosome, a mildly acidic compartment possessing low proteolytic activity. A recycling plasma membrane protein, the transferrin receptor, is often used as marker for early endosomes. Antigens are then transported to the late endosomes, which exhibit a degradative microenvironment with an acidic pH and active proteases. The late endosomal compartment is the intracellular site where endocytosed antigens are degraded.

Differences in Antigen Processing for MHC Class I and Class II Presentation

The two major classes of MHC molecules, class I and II, show a differential distribution and function (**Figure 1**). While MHC class I is expressed on virtually all nucleated cells, expression of class II molecules is restricted to specialized APCs, including macrophages, dendritic cells and B lymphocytes. Antigen degradation is essential for both MHC class I and class II presentation. However, the processes and mechanisms involved in these two pathways are fundamentally different. Also, the immunological outcome of class I and class II presentation differs, resulting in CD8+ and CD4+ T-lymphocyte activation respectively.

Processing for presentation by MHC class I

Epitopes presented on MHC class I molecules are predominantly derived from intracellular proteins, and all proteins present in the cytosol are potential sources of peptides for this pathway (**Figure 1a**) (Pamer and Cresswell, 1998). A cytosolic multisubunit proteolytic complex called the proteasome is responsible for the generation of many such peptides (see below). The peptides are transported into the endoplasmic reticulum (ER) by the adenosine triphosphate-dependent transporters associated with antigen processing (TAPs), where they associate with

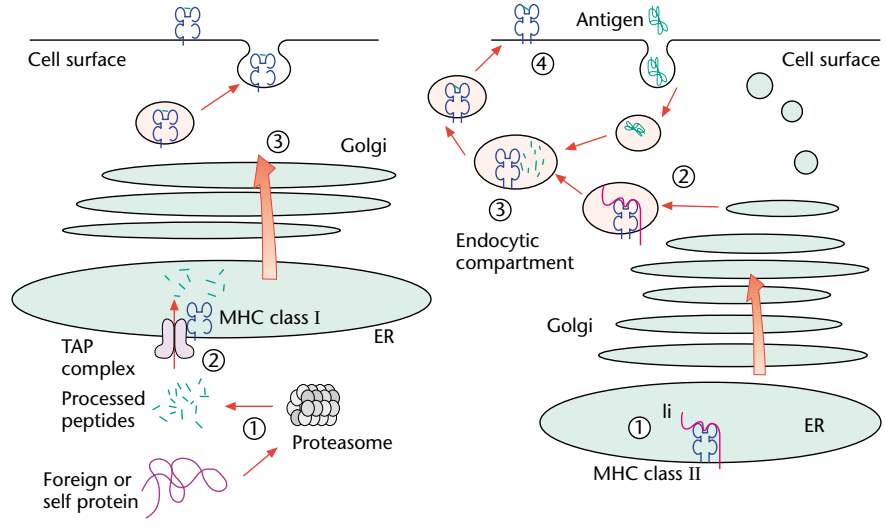

Figure 1 Schematic representation of the major histocompatibility complex (MHC) class I and class II pathways for antigen processing and presentation. (a) Peptides generated in the cytosol by the proteasome (1) are translocated into the endoplasmic reticulum (ER) by the peptide transporter associated with antigen processing (TAP) (2) and loaded on to MHC class I molecules. Subsequently, the class I complexes are transported through the Golgi apparatus (3) to the cell surface. (b) MHC class II molecules are associated with the invariant chain (Ii) in the ER (1). The Ii-associated complex is targeted to the endosomal compartment (2), where peptides generated from endocytosed proteins are loaded into MHC class II molecules (3) and transferred to the cell surface (4).

heterodimers of MHC class I heavy chain and β_2-microglobulin. When stably assembled, the trimeric complexes are transported to the cell surface via the ER and Golgi network. The peptide-binding cleft of MHC class I molecules has a strict peptide length requirement of 8–11 amino acids. Once the peptide is bound deep within the cleft, further processing is not likely to occur. Despite examples of peptide trimming in the ER, it has been established that the cytosol is the major site of processing of class I-presented epitopes. This pathway operates in almost all types of cells, enabling the CD8+ CTL to detect intracellular antigens and destroy infected cells.

It should be noted that, although intracellular antigens are the main sources of MHC class I peptide ligands, processing of exogenous antigens for presentation on class I molecules has been described (reviewed in Jondal *et al.*, 1996). In fact, MHC class I-mediated presentation of exogenous antigens by the specialized APCs is probably necessary for induction of CTL responses against microorganisms that do not infect the specialized APCs themselves.

Processing for presentation by MHC class II

In contrast to MHC class I-mediated processing and presentation, the MHC class II pathway is restricted to specialized APCs. The MHC class II molecules present peptides derived from the extracellular environment of the APC. Proteins, microorganisms and apoptotic material are taken up by the APC and processed by a number of proteases in the endocytic compartment (**Figure 1b**) (Chapman, 1998).

The newly biosynthesized MHC class II αβ heterodimers are associated in the ER with the third component of the early class II molecule, the nonpolymorphic invariant chain (Ii). Ii acts as a chaperone for class II folding and prohibits peptide loading in the ER by occupying the peptide-binding groove. Ii is anchored in the ER membrane, and the cytosolic portion of the molecule directs intracellular sorting of class II molecules through the Golgi to the MHC class II compartment (MIIC) (**Figure 2**). Proteolytic degradation of Ii in endosomes leaves the class II-associated invariant peptide (CLIP) in the peptide-binding groove. MHC class II–CLIP heterotrimers interact with a nonpolymorphic MHC-like dimer, human leucocyte antigen (HLA)-DM (H-2M in the mouse), which catalyses the dissociation of CLIP from the class II molecule. Peptides from endocytosed or phagocytosed material are then loaded on to the now empty MHC class II αβ dimer. Thus, endosomal proteins destined to generate antigenic epitopes are fragmented by proteases residing in the same compartment as the MHC class II molecules themselves. Moreover, since the ends of the peptide-binding groove in the MHC class II are open, these heterodimers can bind a variety of partially fragmented

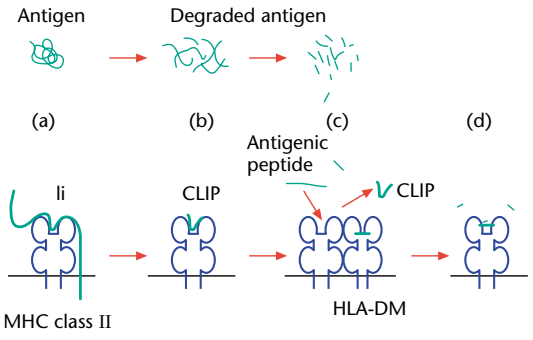

Figure 2 Endosomal processing of antigens and major histocompatibility complex (MHC) class II molecules for presentation of endocytosed antigens. (a) The invariant chain (Ii) targets a MHC class II αβ dimer to the MHC class II endocytic compartment. (b) Proteolytic processing of antigens and of Ii to generate MHC class II with bound class II-associated invariant peptide (CLIP). (c) Human leucocyte antigen (HLA)-DM catalysed exchange of CLIP for antigenic peptide. (d) Proteolytic trimming of the peptide termini and transport of the complex to the cell surface.

proteins of different lengths. While the core epitope bound by the class II molecule is protected from further degradation, the peptide termini protruding from both sides of the peptide-binding groove are trimmed by proteases to yield the final epitope. Thus, in contrast to class I molecules, class II molecules may participate directly in the specificity of processing to determine which epitopes are presented on the cell surface.

In APCs, the site of antigenic peptide loading into MHC class II molecules, the MIIC, is probably similar to late endosomes (Tulp *et al.*, 1994). Antigens endocytosed by pinocytosis or via receptor-mediated uptake probably enter the MIIC via early and late endosomes. However, in the case of phagocytosis it is more unclear how antigens reach the MIIC.

Processing for presentation by MHC class Ib and CD1

In addition to the classical MHC class I and II pathways, antigen processing is also required for presentation by MHC class Ib and CD1 molecules. MHC class Ib molecules are associated with β_2-microglobulin, and are structurally similar to the classical MHC class I. Some MHC class Ib molecules present bacterially derived peptides that contain *N*-formylmethionine at the N-terminus (Lenz *et al.*, 1996). However, little is known concerning the processing of these epitopes.

CD1 molecules share structural features with MHC molecules and require β_2-microglobulin for cell surface expression, although they are encoded by genes located outside the MHC. Processing for presentation by CD1 has similarities with the MHC class II pathway, as CD1 is targeted to the MIIC compartment. Remarkably, CD1 presents nonpeptide lipoglycan antigens which are taken

up via the macrophage mannose receptor (Prigozy *et al*., 1997).

Enzymology of Processing

Generation of short peptides from proteins requires proteolysis. The proteases responsible for processing of antigens for MHC class I and II molecules reside in different cellular compartments and use different enzymatic mechanisms in the hydrolysis of peptide bonds.

Proteases involved in generation of peptide ligands for MHC class I molecules

The short peptides that bind to MHC class I molecules are derived almost exclusively from cytosolic and nuclear proteins degraded in the cytosol. Thus, processing of these peptides occurs in a separate compartment from the loading of peptides on to MHC class I molecules. Controversy exists as to how these peptides are generated, but a substantial fraction are generated by the proteasome, the main nonlysosomal proteolytic system of the cell (Coux *et al*., 1996). Blocking of proteasomal activity in living cells leads to a substantial (50–90%) reduction in assembly of MHC class I molecules, a process that is dependent on such peptide ligands. This suggests that a significant proportion of MHC class I-bound peptides are provided by non-proteasomal proteases, but no such protease has yet been identified (Glas *et al*., 1998).

Cytosolic peptides that are not bound to MHC class I molecules are exposed to proteases and subject to degradation. Soluble short peptides can be degraded rapidly by cytosolic proteases, and it has even been demonstrated that proteasomal inhibition may lead to increased processing of an epitope from *Influenza virus*. Thus, cytosolic processing of peptides for MHC class I molecules is a balance between proteases that can either generate or destroy antigenic peptides (Luckey *et al*., 1998).

Some antigens appear to require only limited processing before binding to MHC class I molecules. Signal peptides target proteins to be translated across the ER membrane, and some such peptides bind to MHC class I. These signal peptides reside in the first 20 amino acids of a polypeptide and are cleaved off in the ER by signal peptidase during translation, and can then be loaded on to MHC class I molecules, even in the absence of both cytosolic proteolysis and TAP transport. Another example may be given by the formation of short defective ribosomal translation products (DRiPs). It is inevitable that ribosomal translation products will sometimes terminate prematurely, and this will produce a pool of unfolded proteins in the cytosol. These would then serve as a major source of precursors in the generation of peptides for MHC class I molecules, but this hypothesis remains to be proven.

Proteases involved in generation of peptide ligands for MHC class II molecules

Cathepsins, a class of proteases that generate MHC class II-bound peptides, are comprised of cysteine and aspartyl

Table 1 Proteases in the endocytic compartment of antigen-presenting cells

Enzyme	Substrate specificity (P4-P3-P2-P1-P1′-P2′)[a]	Ii degradation	Antigen degradation
Endoproteases			
Cathepsin D	X-X-H-H-H-Charged	No	Yes
Cathepsin E	Pro-X-X-H-H-X	?	?
Cathepsin L	X-X-H-X-X-X[b]	Weak	Yes
Cathepsin S	X-X-H-X-X-X[c]	Yes	?
Exopeptidases			
Cathepsin A	Carboxypeptidase	No	?
Cathepsin B	Carboxypeptidase	No	Yes
Cathepsin H	Aminopeptidase	No	Yes
Peptidylpeptidase I	Aminopeptidase	No	?
Peptidylpeptidase II	Prolyl-carboxypeptidase	No	?
Peptidylpeptidase IV	Proline–alanine-aminopeptidase	No	?

Adapted from Chapman (1998)

[a]Amino acid residues in the site cleaved by the protease. P positions refer to amino acid residues in the substrate relative to the target amide bond, which is designated as P1-P1′. The letter X refers to any amino acid, and H to any hydrophobic amino acid.

[b]Cathepsin L has a preference for aromatic amino acids and leucine.

[c]Cathepsin S has a preference for branched aliphatic amino acids and methionine.

proteases (**Table 1**) (Chapman, 1998). In addition to antigen degradation the cathepsins also degrade Ii, a process that is necessary to make the antigen-binding cleft accessible for peptide binding (**Figure 2**). However, the MHC class II molecules themselves are remarkably resistant to proteolysis. The family of cathepsins contains both endoproteases and exopeptidases; endoproteases hydrolyse internal amide bonds in a polypeptide, whereas exopeptidases can only hydrolyse bonds one or two amino acids from the N- or C-terminus. There is redundancy among cathepsins, such that the APC can present antigens even in the absence of either of the two most abundant cathepsins, B and D. However, cathepsin S is in some mouse strains necessary for the degradation of Ii, and inhibition of this protease will in these situations inhibit antigen presentation by MHC class II molecules. As most of the cathepsins are members of the cysteine protease family, they share a similar enzymatic mechanism with serine proteases and the proteasome. These proteases all belong to the group of N-terminal hydrolases which use their N-terminal amino acid as an active nucleophile.

Role of the Proteasome

The proteasome is a main provider of peptides for MHC class I molecules owing to its abundance in the cytosol, as well as its ability to degrade proteins into short peptides (Coux *et al.*, 1996). This protease is a barrel-shaped complex with four stacked rings of noncatalytic α subunits and catalytic β subunits (**Figure 3a**), which have endoprotease activity and can cleave proteins after hydrophobic, basic and acidic amino acid residues. Proteasomes are present in cells from mammals, plants and even certain species of archaebacteria. Note that many of these species do not express MHC molecules, indicating that the immune system probably evolved to use the proteolytic systems already present in the cell. The physiological role of the proteasome includes regulation of the cell cycle, processing of certain transcription factors and the turnover of cellular proteins. MHC-bound peptides produced by the proteasome can be regarded as proteolytic byproducts, used by the immune system to scan for intracellular proteins of microbial origin.

Characteristic of proteases, the proteasome has active sites containing residues that are able to catalyse the hydrolysis of peptide bonds. Proteasomes from eukaryotic cells have several different active sites exhibiting differential specificity, located within the core of the complex (**Figure 3a**). The proteasome is a member of the N-terminal hydrolase family, and uses an N-terminal threonine as a nucleophile. The N-terminal nucleophilic threonine residue forms a covalent intermediate with the substrate, which is subsequently hydrolysed.

In recent years inhibitors of the proteasome have been obtained. Peptide-aldehydes and peptide-vinylsulfones are efficient, and relatively specific, inhibitors of the proteasome, and a variation in the inhibitor peptide sequence may alter their specificity for different active sites of the proteasome. In addition, the fungal metabolite lactacystin inhibits proteasome function by covalent modification of the X subunit. Such inhibitors are of low molecular weight, and can be used to inhibit proteasomes *in vivo*. Inhibition of the proteasome in living cells has confirmed its crucial role for antigen processing as well as for other essential pathways in the cell.

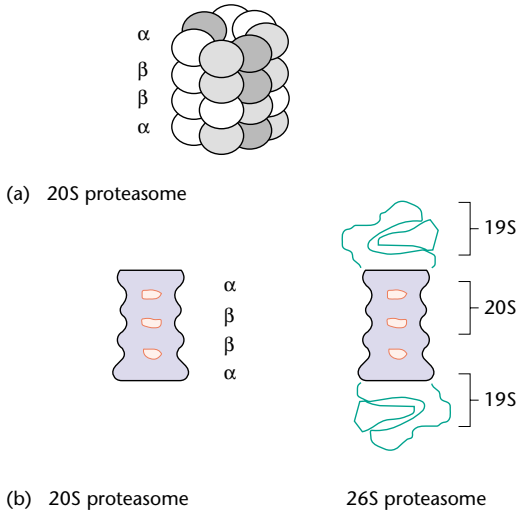

Figure 3 The proteasome, a cytosolic protease involved in production of peptide ligands for major histocompatibility complex (MHC) class I molecules. (a) The eukaryotic 20S proteasome consists of four stacked rings of α and β subunits (7α, 7β, 7β, 7α). (b) Schematic representation of the 20S and 26S proteasomes. The 26S proteasome is formed by the addition of two 19S complexes to the 20S proteasome.

Regulation of Processing

Several components of the antigen-processing machinery are induced by the cytokine interferon (IFN)-γ, which is secreted by activated T lymphocytes and natural killer cells. IFNγ-induced components include MHC class I and II heavy chains, TAP and the proteasomal β subunits LMP-2, LMP-7 and MECL-1. Upon IFNγ induction, the constitutively expressed proteasomal subunits Y, X and Z are replaced by LMP-2, LMP-7 and MECL-1. This substitution shifts proteasomal cleavage to occur after hydrophobic amino acid residues. The peptide sequence requirement for MHC class I binding is dependent on the MHC class I allele, but most alleles require a hydrophobic C-terminus in the peptide for binding. This motif is consistent with the specificity of the TAP peptide transporter and the peptides generated by proteasomes possessing the IFNγ-inducible proteasomal β subunits.

Thus, proteasomal degradation products from cells treated with IFNγ will be more likely to serve as MHC class I-presented antigens.

Proteasomes are regulated by their association with various accessory protein complexes. The four stacked rings of α and β subunits together form the 20S proteasome, but the accessory 19S subunit is required for ubiquitin-dependent degradation. When associated, the 20S and 19S subunits form the 26S proteasome (**Figure 3b**). In addition to the 19S subunit, the 20S proteasome can also associate with PA28 (11S subunit), thereby enhancing the spectrum of produced peptides, which appears to improve antigen presentation. The PA28 subunit also has a *heptameric* configuration of alternating α and β subunits, and is inducible by IFNγ.

The 26S proteasome degrades ubiquitin-conjugated proteins. There are a number of enzymes that catalyse the covalent conjugation of ubiquitin to lysine residues within proteins via formation of isopeptide bonds. This leads to protein degradation by the proteasome, and ubiquitin is thereby used as a tag to target the protein for proteasomal degradation. The role for this pathway is not entirely clear, but some antigens may require ubiquitin conjugation for antigen processing.

Regulation of processing in the MHC class II pathway is largely dependent on regulation of cathepsin activity in the endosomal compartment. Two major factors controlling activity of these cysteine proteases are the pH and the presence of endogenous cysteine protease inhibitors. Most cysteine proteases are unstable and have weak activity at neutral pH, and are optimized to function in acidic endosomal compartments. The cystatin family of endogenous protease inhibitors is found in the cytoplasm and extracellular space, and the main role of these inhibitors is to ensure that proteases that escape compartmentalization remain inactive. By these mechanisms, cathepsin activity is directed to the endosomal compartment where antigen processing and MHC class II loading occurs.

Concluding Remarks

Although the last 10 years have seen a tremendous increase in knowledge about the mechanisms of antigen processing, many aspects remain unclear. One challenge ahead is to determine the role of different proteases in both cytosolic and endosomal processing. Which cytosolic proteases, other than the proteasome, can generate or destroy antigenic peptides? What roles do the apparently redundant cathepsins play in endosomal processing? The answers to these questions will have important implications for our understanding of how pathogens can evade the immune system by inhibiting antigen processing, and for development of more efficient vaccines.

References

Chapman HA (1998) Endosomal proteolysis and MHC class II function. *Current Opinion in Immunology* **10**: 93–102.

Coux O, Tanaka K and Goldberg AL (1996) Structure and functions of the 20S and 26S proteasomes. *Annual Review of Biochemistry* **65**: 801–847.

Glas R, Bogyo M, McMaster JS, Gaczynska M and Ploegh HL (1998) A proteolytic system that compensates for loss of proteasome function. *Nature* **392**: 618–622.

Jondal M, Schirmbeck R and Reimann J (1996) MHC class I restricted CTL responses to exogenous antigens. *Immunity* **5**: 295–302.

Lenz LL, Dere B and Bevan MJ (1996) Identification of an H2-M3 restricted *Listeria* epitope: implications for antigen presentation by M3. *Immunity* **5**: 63–72.

Luckey CJ, King GM, Marto JA *et al*. (1998) Proteasomes can either generate or destroy MHC class I epitopes: evidence for nonproteasomal epitope generation in the cytosol. *Journal of Immunology* **161**: 112–121.

Pamer E and Cresswell P (1998) Mechanisms of MHC class I-restricted antigen processing. *Annual Review of Immunology* **16**: 323–358.

Prigozy TI, Sieling PA, Clemens D *et al*. (1997) The mannose receptor delivers lipoglycan antigens to endosomes for presentation to T cells by CD1b molecules. *Immunity* **6**: 187–197.

Tulp A, Verwoerd D, Dobberstein B, Ploegh HL and Pieters J (1994) Isolation and characterization of the intracellular MHC class II compartment. *Nature* **369**: 120–126.

Watts C (1997) Capture and processing of exogenous antigens for presentation on MHC molecules. *Annual Review of Immunology* **15**: 821–850.

Further Reading

Chapman HA (1998) Endosomal proteolysis and MHC class II function. *Current Opinion in Immunology* **10**: 93–102.

Janeway CA Jr and Travers P (1997) *Immunobiology: The Immune System in Health and Disease*, 3rd edn. Current Biology–Garland. London and New York.

Pamer E and Cresswell P (1998) Mechanisms of MHC class I-restricted antigen processing. *Annual Review of Immunology* **16**: 323–358.

Roitt IM (1997) *Essential Immunology*, 9th edn. Oxford: Blackwell Science.

Watts C (1997) Capture and processing of exogenous antigens for presentation on MHC molecules. *Annual Review of Immunology* **15**: 821–850.

Wolf PR and Ploegh HL (1995) How MHC class II molecules acquire peptide cargo: biosynthesis and trafficking through the endocytic pathway. *Annual Review of Cell and Developmental Biology* **11**: 267–306.