Nature's Shotgun Proteomics

George W. Preston

King's College London, 150 Stamford Street, London, SE1 9NH, United Kingdom; e-mail: george.preston@kcl.ac.uk.

INTRODUCTION

The mammalian immune system employs a number of different strategies to rid the body of infectious agents (e.g., viruses), and to make sure that infections, once neutralised, cannot return. Some of the strategies involve the cell-mediated immune response, a process whereby T cells are stimulated to differentiate (Alberts *et al*., 2002). The differentiated T cells are capable of either killing infected cells (cytotoxic T cells), or spreading news of the infection to other parts of the immune system (helper T cells).

Viruses propagate within cells, hijacking the translational machinery and using it to make more virus particles. An infected cell must inform the immune system of its predicament, so that the threat posed by the virus can be neutralised. In doing so, the cell performs something akin to a shotgun proteomics experiment. Viral proteins are digested, and the resulting peptides are dispersed—as a mixture—across the cell surface (Fig. 1). These processes of *antigen processing* and *antigen presentation* are somewhat analogous to the sample preparation steps of a proteomics workflow. As will be seen, it is possible for a real-life proteomics experiment to pick up from where antigen processing leaves off (i.e., by analysing the peptides).

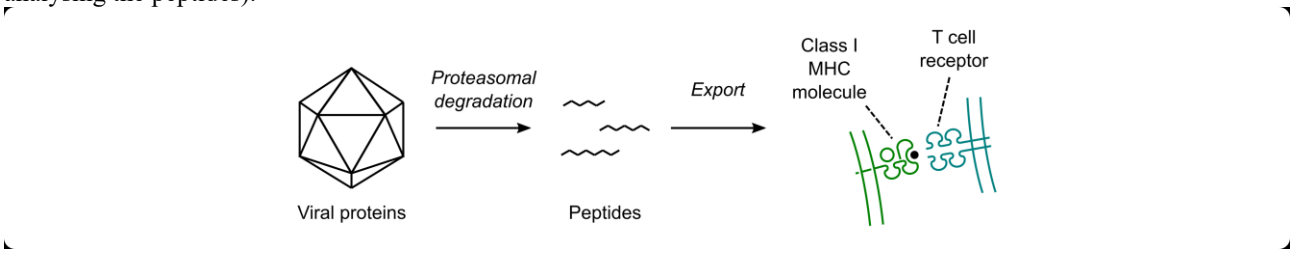

Fig. 1. Antigen processing and presentation. Viral proteins are degraded intracellularly and the resulting peptides are exported to the cell surface by class I MHC molecules. Here, the peptides—still bound to their MHC molecules—are recognised by receptors on the surfaces of T cells.

APPARATUS: CLASS I MHC MOLECULES

The most important pieces of 'apparatus' in the processing and presentation of antigens are class I MHC (*major histocompatibility complex*) protein molecules. These are the molecules that bind peptides and present them at the cell surface (Jewell and Wilson, 1996). Class I MHC molecules are integral membrane proteins each consisting of two subunits. The larger of the subunits (the α-subunit) spans the membrane, and is responsible for binding the peptides. The smaller subunit, known as β_2-microglobulin, is associated with the extracellular part of the α-subunit, but is not directly involved in peptide binding (Alberts *et al*., 2002). Peptides bind to the α-subunit within a groove formed at the interface of two domains. The groove is long and narrow, meaning that the bound peptide must adopt an extended conformation (Jewell and Wilson, 1996). Since the immune system must be able to deal with a variety of threats (some to which it has had no prior exposure) class I MHC molecules must be able to accommodate a variety of peptides (Alberts *et al*., 2002). The peptides have in common their lengths (typically eight or nine amino acid residues) and certain residues that anchor them to the MHC molecule.

The term *major histocompatibility complex* actually refers to a cluster of *genes* rather than, as the name might suggest, a group of noncovalently-associated protein subunits (although, clearly, class I MHC proteins would also fit that description). Genes of the MHC encode two other very important pieces of apparatus: proteasomal subunits and peptide transporters (Monaco, 1996). These are used for 'sample preparation', as described in the next section.

SAMPLE PREPARATION: DEGRADATION AND TRANSPORT

The peptides displayed by class I MHC molecules derive from proteins in the cytoplasm. How do *proteins in the cytoplasm* become *peptides on the cell surface*? The first stage of this 'sample preparation' is proteolytic digestion, which occurs in the cytoplasm (Monaco, 1996). Digestion is performed by proteasomes—large, hollow, cylindrical protein complexes with hydrolytic active sites on their inner surfaces (Mishto and Liepe, 2017). Substrates are retained in proteasomes until completely digested, and the resulting peptides are then translocated to the lumen of the endoplasmic reticulum, where empty class I MHC α-subunits are waiting to bind them. The peptides are translocated by a heterodimeric integral membrane protein called TAP, the *transporter of antigenic peptides*. TAP appears to pre-select those peptides that will make good binding partners for class I MHC molecules (Monaco, 1996).

ANALYSIS: T CELL ACTIVATION

In a sense, the physical dispersal of peptides among class I MHC molecules is the first stage of the 'analysis'. In the analogous proteomics experiment, physical separation of peptides is achieved using liquid chromatography, and peptides are eluted into a mass spectrometer. Spectra are acquired and searched against a database of protein sequences (Cottrell, 2011). In the cell-mediated immune response, bound peptides are 'searched' against a 'database' of immature T cells or—to be specific—their T cell receptors. T cell receptors are structurally diverse; and they need to be, because of the diversity of their potential binding partners. The 'search hits' (i.e., T cells with receptors that bind the peptide-occupied class I MHC molecule) are induced to differentiate into cytotoxic T cells. These interact with infected cells and induce them to undergo programmed cell death (Alberts *et al*., 2002).

PEPTIDE SPLICING

No sooner had the basic mechanism of antigen processing been established than a puzzling and unexpected phenomenon began to be observed (Vigneron *et al*., 2004). Investigators noticed that some of the peptides presented by class I MHC molecules had segments missing (Mishto and Liepe, 2017). In each case, a segment of a processed protein had apparently been excised, and the peptides flanking the excised segment had been spliced. Vigneron *et al*., having observed the formation of RTKQLYPEW from RTKAWNRQLYPEW (excised segment underlined), postulated that the splicing was a result of proteasome-catalysed transpeptidation, and demonstrated that the phenomenon could be reproduced *in vitro* (Vigneron *et al*., 2004).

The results of a recent study have indicated that splicing is much more widespread than was initially apparent (Liepe *et al*., 2016). Liepe *et al*. isolated peptides from class I MHC molecules and analysed them using liquid chromatography and mass spectrometry (i.e., a proteomic workflow without the proteolytic digestion). Using database-search methods, the authors found evidence of splicing in a third of the peptides. This surprising result has implications for the way in which we perceive the cell-mediated immune response (shouldn't splicing hinder it?) and the way in which antigens are analysed (is the search space big enough?) (Liepe *et al*., 2016; Mishto and Liepe, 2017).

REFERENCES

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The Adaptive Immune System, in: *Molecular Biology of the Cell*, 4th ed., Garland Science, New York, 2002.

Cottrell JS (2011) Protein identification using MS/MS data. *J. Proteomics* 74, 1842-1851. DOI: 10.1016/j.jprot.2011.05.014

Jewell DA, Wilson IA. Structure and function of MHC class I and class II antigens, in: Hames BD, Glover DM (Eds) *Molecular Immunology*, 2nd ed., Oxford University Press, New York, 1996.

Liepe J, Marino F, Sidney J, Jeko A, Bunting DE, Sette A, Kloetzel PM, Stumpf MPH, Heck AJR, Mishto M (2016) A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* 354, 354-358. DOI: 10.1126/science.aaf4384

Mishto M, Liepe J (2017) Post-Translational Peptide Splicing and T Cell Responses. *Trends Immunol.* 38, 904-915. DOI: 10.1016/j.it.2017.07.011

Monaco JJ. Molecular mechanisms of antigen processing, in: Hames BD, Glover DM (Eds) *Molecular Immunology*, 2nd ed., Oxford University Press, New York, 1996.

Vigneron N, Stroobant V, Chapiro J, Ooms A, Degiovanni G, Morel S, van der Bruggen P, Boon T, Van den Eynde BJ (2004) An Antigenic Peptide Produced by Peptide Splicing in the Proteasome. *Science* 304, 587-590. DOI: 10.1126/science.1095522