Chapter 3.1

Transpeptidation and reverse proteolysis and their consequences for immunity

International Journal of Biochemistry and Cell Biology
Volume 41, 2009, Pages 66-71
Transpeptidation and reverse proteolysis and their consequences for immunity

Celia R. Berkers, Annemieke de Jong, Huib Ovaa & Boris Rodenko

Division of Cellular Biochemistry, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

Reverse proteolysis and transpeptidation lead to the generation of polypeptide sequences that cannot be inferred directly from genome sequences as they are post-translational phenomena. These phenomena have so far received little attention although the physiological consequences may reach far. The protease mediated synthesis of several immunodominant MHC class I antigens was recently reported, underscoring its importance to immunity. Reverse proteolytic and transpeptidation mechanisms as well as conditions that favor successful protease-catalyzed synthetic events are discussed here.

INTRODUCTION

A broad repertoire of antigenic peptides is presented at the cell surface to the immune system by major histocompatibility complexes (MHCs) class I and class II, which are known in humans as human leukocyte antigens (HLAs) class I and class II. Intracellular proteins form the source of antigens that are presented by MHC class I, enabling the immune system to monitor changes in intracellular protein content and sense malignant transformation or the presence of viral proteins. Recognition of these foreign or abnormal peptides by CD8+ cytotoxic T cells ultimately results in the elimination of the antigen presenting cell.1,2 MHC class I antigens are primarily generated in the cytosol, where proteins are initially degraded by the proteasome into smaller fragments. In order to fit the MHC class I binding groove, these peptides undergo further N-terminal trimming by various aminopeptidases,3,4 that include tripeptidylpeptidase II (TPP II), bleomycin hydrolase, thymet oligopeptidase, puromycin sensitive aminopeptidase, neurolysin, and leucine aminopeptidase, although many other proteases may be involved in this trimming process. After processing, these peptides are transported into the endoplasmic reticulum (ER), where some can be loaded onto MHC class I directly, while others may require additional N-terminal trimming by ER-associated aminopeptidase (ERAAP).5

Antigens generated from exogenous pathogens are presented via the MHC class II pathway, which leads to the activation of CD4+ T helper cells. Following endocytosis by professional antigen presenting cells, exogenous proteins are degraded in early endosomes, late endosomes and lysosomes in a pH-dependent manner. Association of thus formed peptides with MHC class II can take place in any of these compartments,6 but occurs mainly in late endocytic structures called MIIC compartments. Cysteine proteases, including cathepsins S, L, B, F, H and V as well as aspar-

5These authors contributed equally to this work. *Correspondence should be addressed to H.O. (h.ovaa@nki.nl).
agrine endopeptidase (AEP), have mainly been implicated in antigen processing, but various metallo-, aspartic acid- and serine-proteases have been shown to be involved in MHC class II antigen processing as well.\textsuperscript{6}

Recently, it has become apparent that not only contiguous antigens, but also non-contiguous antigens, that consist of two post-translationally fused peptides, can be presented to the immune system.\textsuperscript{1,7,8} The ligation that produces these epitopes is thought to be catalyzed by proteases \textit{in vivo}. Here, we summarize how and under which conditions proteases catalyze peptide ligation, as well as the consequences of this process for immunity.

**PROTEASE-CATALYZED PEPTIDE LIGATION**

Based on their proteolytic mechanism, proteases can be divided into five classes,\textsuperscript{9,10} and broadly form two groups: cysteine, serine, and threonine proteases, that hydrolyze peptide bonds via the formation of an acyl-enzyme intermediate (Figure 1), and aspartic- and metalloproteases, that act via the direct activation of a water molecule that acts as the primary hydrolyzing species. Since proteases are catalysts, they alter the rate at which the thermodynamic equilibrium of the protein hydrolysis reaction is reached, but they do not change the overall thermodynamic equilibrium between starting materials and proteolytic products.\textsuperscript{11,12} Consequently, all proteases are also able to catalyze the reverse reaction, \textit{i.e.} peptide bond synthesis, provided that the reaction conditions favor this process, which is referred to as reverse proteolysis (Scheme 1). Furthermore, proteases that act via an acyl-enzyme intermediate can catalyze peptide bond formation via a transeptidation mechanism, which competes with hydrolysis (Figure 1).

**REVERSE PROTEOLYSIS**

Peptide ligation by reverse proteolysis was first reported in the 1930s, when papain and chymotrypsin, a cysteine and serine protease, respectively, were used as catalysts for enzymatic synthesis.\textsuperscript{13-16} Reverse proteolysis involves the direct reversal of catalytic hydrolysis by proteases, as shown in Scheme 1, in which $K_{\text{ion}}$ is the equilibrium constant of ionization and $K_{\text{pept}}$ is the equilibrium constant of peptidation.\textsuperscript{11,12,17} In an aqueous environment, the reaction will be driven in the direction of hydrolysis.\textsuperscript{11} Furthermore, the ligation of peptide fragments is an endergonic process (entropy is lost) and is therefore energetically unfavorable.\textsuperscript{17} One strategy to direct the equilibrium towards ligation is based on the law of mass action.\textsuperscript{12,17,18} \textit{In vitro}, product removal by precipitation,\textsuperscript{19-22} the use of biphasic systems,\textsuperscript{23} the use of a large molar excess of one of the substrates,\textsuperscript{11,17} or a conformational trap\textsuperscript{22,24,25} have all been shown to promote ligation. Alternatively, the equilibrium of ionization can be manipulated to enhance ligation.\textsuperscript{12,17,18} Peptide bond formation is dependent on the concentration of uncharged substrates formed in the ionization equilibrium, which is dependent on their $pK_a$ values. Therefore, the pH optimum for ligation lies in the range between the $pK_a$ of the

![Scheme 1 | Mechanism of reverse proteolysis.](image)
Chapter 3.1   |   Transpeptidation and reverse proteolysis

Figure 1 | General mechanism of protease mediated hydrolysis and transpeptidation via an acyl-enzyme intermediate. (i) The protein or peptide substrate with its P₁, P₂, P₃, P₁', P₂' and P₃' residues docks into the complementary S₁, S₂, S₃ and S₁', S₂', and S₃' pockets of the protease. (ii) Nucleophilic attack of the reactive residue on the enzyme results in formation of the acyl-enzyme intermediate. (iii) The cleaved C-terminal substrate fragment leaves the active site and hydrolysis of the acyl-enzyme intermediate can take place - path a - under release of the N-terminal peptide or protein fragment. (iv) Alternatively, the vacated S₁', S₂' and S₃' binding sites are subsequently occupied by a protein or peptide fragment containing complementary P₁', P₂' and P₃' residues. Aminolysis by the free N-terminus - path b - results in net transpeptidation, and competes with hydrolysis - path a.
α-carboxylic group of the N-terminal component and the amino group of the C-terminal substrate, which is normally found between pH 6 and 7. Several water-miscible organic co-solvents can reduce the acidity of the α-carboxy group of the N-terminal component, which shifts $K_{\text{ion}}$ in favor of ligation, and have therefore been used in vitro to influence the ionization equilibrium. Furthermore, organic co-solvents reduce the activity of water by lowering its concentration, thereby hampering the hydrolysis reaction. Under physiological conditions, the equilibrium as shown in Scheme 1 favors hydrolysis and ligation is thought by many to be negligible in vivo. It has been postulated, however, that macromolecular crowding, defined by the high total concentration of background molecules, such as lipids, proteins, nucleic acids and carbohydrates that are present in cells, shifts the equilibrium towards ligation in vivo. Due to non-specific steric repulsions, the presence of a significant volume fraction of macromolecules places constraints on the placement of additional macromolecules. Macromolecules therefore exclude volume to each other. This excluded volume effect reduces the configurational entropy, which is energetically unfavorable. Therefore, processes that are accompanied by a reduction in excluded volume, such as macromolecular compaction and association, are facilitated in a macromolecularly crowded environment. Consequently, molecular crowding has been shown to promote peptide ligation as well as the synthesis of proteins from partly structured large polypeptides, in cases where ligation led to a large reduction in excluded volume. This suggests that in vivo, proteases might also catalyze peptide bond synthesis in addition to their ability to function as hydrolases.

**TRANSPSEPTIDATION**

Transpeptidation occurs when the rate of aminolysis is faster than or comparable to the rate of hydrolysis, which is dependent on the competition between nucleophilic components and water for the nucleophilic attack on the acyl-enzyme intermediate (Figure 1). In vitro, the use of water-miscible organic co-solvents or biphasic systems can help to reduce the water content and therefore favor aminolysis over hydrolysis. High reactant concentrations and a high $p\text{H}$, preferably higher than the $pK_a$ of the attacking nucleophile to ensure its unprotonated state, have also been shown to promote transpeptidation. Evidently, the specificity of the protease for the reactants, is also a crucial factor as its active site has to be able to accommodate reactants for productive reactions. Both in vitro studies, in which different amino acids or peptides were used as nucleophiles, and an in vivo study (see below) indicate that ligation efficiencies are highly dependent on the fit of the amino acids at the $P_1'$, $P_2'$ and/or $P_3'$ positions in the $S'$ binding sites (Figure 1).

Examples of protease-catalyzed transpeptidation occurring in vivo are scarce, but few examples have been reported. Asparaginyl endopeptidases (AEPs) have been implicated in the catalysis of two distinct transpeptidation events in plants: (1) the post-translational processing of the lectin concanavalin A (con A) in maturing jack beans and (2) the biosynthesis of the backbone of plant-produced cyclotides, the largest family of circular proteins that are notoriously stable, due to the combination of a cyclic backbone and a cysteine knot of three disulfide bonds. During cleavage of pro-con A into smaller fragments, two of these fragments are ligated in the reverse order by AEP in a transpeptidation reaction to form mature con A. Both the
increased stability of mature con A compared to its unligated precursor\textsuperscript{48} and the occurrence of a proximity effect may explain why transpeptidation is favored over hydrolysis \textit{in vivo}. If protein complexation or conformation positions the amine-donating component involved in transpeptidation in close proximity to the acyl ester component, this will facilitate its nucleophilic attack on the acyl-enzyme intermediate and hence favor aminolysis. Although such an effect has never been implicated in con A transpeptidation, the proposed structure of pro-con A\textsuperscript{50} indicates this is likely the case.

Plant-produced cyclotides exhibit a broad range of bioactivities, amongst which are antimicrobial-, hemolytic-, anti-HIV- and insecticidal activities.\textsuperscript{46,49} Although the mechanism is not fully understood, cyclization of these proteins \textit{in vivo} is thought to be the result of an AEP-catalyzed transpeptidation reaction.\textsuperscript{46,49} The cysteine knots are believed to keep the amine-donating and carboxy-donating components involved in transpeptidation in close proximity to each other, indicating that a proximity effect facilitates ligation \textit{in vivo}. Furthermore, a tripeptide binding motif at the N-terminus of the AEP bound propeptide binds, after release of the cleaved C-terminal propeptide fragment, to the vacated S\textsubscript{1}', S\textsubscript{2}', and S\textsubscript{3}' subsites, which facilitates cyclization.\textsuperscript{46} This suggests that substrate specificity further enhances transpeptidation.

The proteasome, a threonine protease, has been shown to catalyze transpeptidation during the production of antigens for presentation by MHC class I. To date, three non-contiguous immunodominant antigens, formed by the ligation of two peptides from the same parental protein during proteasomal degradation, have been identified. The first non-contiguous antigen, which resulted from the fusion of residues 172–176 and 217–220 from fibroblast growth factor-5 was identified from a patient with renal cell carcinoma.\textsuperscript{8} Vigneron et al. identified a second non-contiguous tumor antigen, composed of residues 40–42 and 47–52 of the melanocytic glycoprotein gp100, in a melanoma patient.\textsuperscript{7} Subsequently, an antigen was identified in a recipient of MHC-matched hematopoietic cell transplantation, consisting of two peptides from the SP110 nuclear body protein, ligated in the reverse order.\textsuperscript{1} Proteasomal involvement in the ligation events was suggested,\textsuperscript{1,7,8} and \textit{in vitro} digestion experiments with purified 20S proteasome confirmed this notion.\textsuperscript{1,7} Further experiments showed that both cleavage of the carboxy component and a free N-terminus at the amine-donating reaction component were required for ligation, indicating that splicing in the proteasome occurs via a transpeptidation mechanism.\textsuperscript{1,7} In most cytosolic proteases the reaction products diffuse away too quickly under physiological conditions to promote ligation. However, in barrel-shaped, self-compartmentalizing proteases such as the proteasome,\textsuperscript{51} reaction partners are generated \textit{in situ} and remain confined to a small volume, resulting in their high local concentrations and a high degree of substrate specificity, which likely facilitates peptide ligation \textit{in vivo}.

\textbf{CONSEQUENCES FOR IMMUNITY}

Several reports show that transpeptidation occurs readily under physiological conditions \textit{in vitro}, during digestion of proteins for structural analysis. Trypsin,\textsuperscript{52-56} Staphylococcus aureus V8 protease,\textsuperscript{52,56} and Lys-C endoprotease\textsuperscript{56} were all shown to catalyze transpeptidation. Based on their observations, Fodor and Zhang stated that transpeptidation occurs more often than was known previously.\textsuperscript{56} Schaefer et al. emphasized that available database search algorithms fail to detect transpeptidation products, as they cannot be inferred from the...
genome. In vivo, the extent to which proteases catalyze transpeptidation rather than hydrolysis is fully dependent on specific local conditions. γ-Glutamyltranspeptidases, for example, which bind glutathione and function predominantly as hydrolases in vitro, are assumed by many to mainly catalyze transpeptidation reactions in vivo due to the presence of high concentrations of amino acids in kidney as amine-donating reaction components. The conditions promoting ligation described above, including molecular crowding, proximity effects, a high degree of substrate specificity or confinement can certainly be met by proteases other than AEPs and the proteasome, suggesting that protease-catalyzed peptide ligation represents a more widespread and common mechanism than currently assumed. Macromolecular crowding in the cytosol may facilitate ligation of non-contiguous polypeptides to form novel proteins. During proteasomal degradation of cytosolic proteins into peptides, transpeptidation produces spliced and reordered antigens. Notably, both bleomycin hydrolase and TPPII, peptidases that process these antigens downstream of the proteasome, are thought to assemble into self-compartmentalizing proteases, which may facilitate transpeptidation. Bleomycin hydrolase has been shown to be able to convert itself into a peptide ligase in vitro. This suggests that transpeptidation is likely to occur during MHC class I epitope trimming downstream of the proteasome. Furthermore, functional homologs of mammalian AEP, which is actively involved in MHC class II antigen production, have been shown to catalyze transpeptidation in plants.

In our current understanding of the cellular antigen presenting pathways, proteins are cleaved in contiguous sequences, creating antigens that can traditionally be inferred from genomic sequences, and in which reverse proteolysis and transpeptidation have not been taken into account. The repertoire and diversity of both foreign and self antigens in the MHC routes may thus be significantly broader than currently thought. Amplifying the diversity of epitopes increases the chance that one or more of these epitopes are recognized by CD4+ and CD8+ T cells, which ultimately results in the elimination of infected or malignant cells. It is thus likely that the expansion of the epitope repertoire as a result of transpeptidation or reverse proteolysis events, results in enhanced anti-tumor immunity as well as in better protection from pathogens. Increasing diversity by V(D)J recombination is essential for protection from a constantly changing bacterial and viral repertoire. V(D)J recombination results in a large variety of antibodies and T cell receptors, while differential splicing of RNA in B lymphocytes results in the production of either membrane bound or secreted antibody molecules. The protease-catalyzed ligation of peptides or even proteins could be yet another way of the immune system to increase diversity and enhance protection.

The increased diversity of self-antigens on the other hand demands additional negative selection to prevent autoimmunity. During T cell development, the thymus is responsible for the negative selection of T cells, resulting in a T cell repertoire that only recognizes foreign antigens. Autoimmunity against spliced antigens may occur if self-antigens are spliced differently in antigen presenting cells compared to cells in the thymus, thereby evading negative selection. Differential splicing in different tissue types will occur if splicing is a random process. However, both transpeptidation and reverse proteolysis reactions only take place under specific circumstances, requiring, for instance, a high degree of substrate specificity. Therefore, it is likely that splicing is a controlled process that is performed identically in different cell types. Little data is available
to confirm this notion, but notably, Vigneron et al. have found identical results with proteasomes purified from either a human renal cell carcinoma line or human erythrocytes.\(^7\) In splicing experiments performed by Warren et al. also constitutive and immunoproteasomes seem to behave in a similar fashion.\(^1\) These findings may indicate that proteasomes from different cell types all behave similarly and point to a tightly controlled splicing process, which is carried out identically in different cell types or even with different proteasome compositions. The spliced antigens produced by both constitutive and immunoproteasomes in the thymus are therefore not likely to be different from antigens produced in other tissues. It should be noted that Murata et al. have discovered a thymus-specific proteasome type, which contains a novel catalytic subunit (β5t) and displays different catalytic properties.\(^6,7\) However, these thymoproteasomes are not involved in negative T cell selection,\(^6,7\) and the antigens produced by this type of proteasome are therefore not interfering with autoimmunity control. Together, these reports suggest that transpeptidation and reverse proteolysis reactions in the cell are not likely to lead to an increase in the recognition of self-antigens.

It is imperative to know the identity of the antigens presented by the MHC pathways, as they mediate T cell responses to cancer and transplants, as well as bacterial and viral infection. The characterization of novel antigens is therefore important for vaccine development, T cell-based therapeutic strategies as well as pharmacological immunosuppression. Further studies need to show the contribution of different proteases to the production of noncontiguous antigens and to elucidate their role in immunity.

**REFERENCES**

45. Stein, R.L. & Strimpler, A.M. Catalysis by human