ENDOPLASMIC RETICULUM AMINO PEPTIDASE 1 (ERAP1) IN ANKYLOSING SPONDYLITIS

Marius Cherciu¹, Mihai Bojinca², Laura Ioana Cherciu¹, Constantin Bara¹, Olivia Mihaela Popa¹
¹Department of Pathophysiology and Immunology, Faculty of Medicine, Carol Davila University of Medicine and Pharmacy, Bucharest
²Department of Rheumatology and Internal Medicine, Faculty of Medicine, Carol Davila University of Medicine and Pharmacy, I.C. Cantacuzino Hospital, Bucharest

Abstract

Recently, endoplasmic reticulum aminopeptidase 1 (ERAP1) gene polymorphisms were associated with ankylosing spondylitis (AS) in genome-wide association studies (GWAS), associations which later were replicated across different populations, including the Romanian population. ERAP1 enzyme was proved to be involved in the final trimming, to the optimal length, of the antigenic peptides before loading them to human leucocyte antigen (HLA) class I molecules, different variants of ERAP1 being able to modify the types of peptides presented by HLA I molecules (including HLA-B27) to the CD8+ T cells. This article presents a detailed description of the functions, structure and the pathogenic potential of ERAP1 in AS.

Keywords: ERAP1, ankylosing spondylitis, HLA-B27, single nucleotide polymorphisms (SNPs)

Over time, endoplasmic reticulum aminopeptidase 1 (ERAP1) has had different names, among which we mention: aminopeptidase regulating tumor necrosis factor receptor I shedding (ARTS1), adipocyte-derived leucine aminopeptidase (A-LAP), puromycin-insensitive leucine aminopeptidase (PILS-AP) or endoplasmic reticulum aminopeptidase associated with antigen processing (ERAAP). Currently, ERAP1 is the term accepted by Human Genome Organisation (HUGO) (1).

ERAP1 protein was isolated in 2002 from the HeLa cell line and it contains 941 amino acids (2). ERAP1 gene, which encodes all of these amino acids, is located on the long arm of chromosome 5 in position 15, along with ERAP2 gene. Both chromosomal copies of ERAP1 gene are codominantly expressed in humans (3). ERAP2 molecule has a similar function, evolutionary studies suggesting that the gene was formed after a relatively recent duplication of ERAP1 (4). ERAP1/ERAP2 complex seems to work synergistically and with high efficiency (5), but to date there are few studies regarding strictly ERAP2 function (6).

ERAP1 FUNCTIONS

Several biological functions of ERAP1 are known until now. Some functions belong to the metallopeptidase class which is the class ERAP1 is part of. ERAP1 is involved in angiogenesis and also helps regulating blood pressure (7,8). Recently, it was discovered that ERAP1 is secreted in the macrophages, as a response to the lipopolysaccharide and interferon γ activation, resulting in the amplification of the phagocytic activity of the respective macrophages (9).

Another important function, which can play a role in the etiopathogenesis of ankylosing spondylitis (AS), is the cleavage of the cell surface receptors of various pro-inflammatory cytokines: TNF-R1 (10), IL1-R1I (11), IL-6 Rα (12) – a process called shedding. The decrease of surface receptors for pro-inflammatory cytokines lowers intracellular signals, which theoretically means that malfunctions of ERAP1 can amplify inflammatory signals and maintain chronic inflammation.
But perhaps the most important function that ERAP1 has, is to adjust antigenic peptides to be loaded onto human leukocyte antigen (HLA) class I (including HLA-B27), thus playing a key role in generating the antigenic peptides which are presented by these molecules to CD8+ T cells and natural killer lymphocytes (13-16). ERAP1 is dealing with the final phase of the antigenic epitopes processing which are thus brought to a proper length of 8-9 amino acids.

Original antigenic protein complexes are degraded by the proteasome complex, which generates fragments of up to 25 amino acids (17), which are then transported into the endoplasmic reticulum by the transporter associated with antigen processing (TAP), which carries polypeptide fragments of up to 16 amino acids, then taken by ERAP1, cut to the optimum length of 8-9 amino acids and loaded onto nascent molecules of the major histocompatibility complex (MHC) class I to be transported to the cell surface.

**ERAP1 STRUCTURE**

ERAP1 belongs to the M1 aminopeptidases family (along with aminopeptidase A, placental leucine aminopeptidase and another 8 aminopeptidases) (18), zinc-metallopeptidases which present the aminoacid patterns H-E-x-x-H-(x)18-E (the Zn linking site) and G-A-M-E-N that are crucial for their enzymatic function (19).

The three-dimensional crystallographic structure of ERAP1 initially generated was based on the crystallographic structure of a similar protein, TIFF3 (tricorn-interacting factor F3), which resembles ERAP1 up to 82% between the amino acids in the catalytic site (amino acids 280-486) (20). In 2011, two independent groups of researchers (Kochan et al. and Nguyen et al.) determined the crystallographic structure of ERAP1 in the two states (open and closed), paving the way towards understanding the relationship between the molecule structure and its enzymatic activity (18,21). The molecule, very similar to the model initially generated, presents four domains labeled from I to IV, from the amino-terminus. The first domain contains the first 254 amino acids, the second domain, the 255-527 amino acids, the third domain, the 528-613 amino acids while the fourth domain contains the last amino acids (614-941) – representing the carboxyl-terminus of the molecule (Fig. 1).

*FIGURE 1. ERAP1 crystallographic structure*

The open conformation (left) and closed conformation (right). Zn atom is represented in black. Adapted by permission from Macmillan Publishers Ltd: Nature structural & molecular biology, Nguyen et al. Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1, 18(5):604-13., copyright (2011)
The catalytic domain of the molecule is the second domain, which presents the amino acid sequences H-E-x-x-H-(x)₁₈-E and G-A-M-E-N specific to the aminopeptidases and it is also containing the Zinc atom. The active site of ERAP1 consists of five structures: H2 and H3 helices (containing H-E-x-x-H-(x)₁₈-E), G-A-M-E-N loop, domain I loop and H5 helix (18,21). Domain I is situated above the catalytic site and it seems to be the anchorage spot for the substrate’s amino-terminus. The Domain III is a short domain with less than 100 amino acids, which connects with the domain IV whose helical structures (16 in number) are arranged in a large conical structure near the active site (21). In the open conformation the domain IV is situated at considerable distance from the catalytic site, but it is closer in the closed conformation of the molecule, the domain III thus acting like a hinge structure. The cavity which is formed between the domains II and IV is narrower towards the catalytic site and the wider towards the domain IV where the carboxy-terminal portion of the polypeptide substrate can bind. The site of cleavage of ERAP1 allows the attachment of the amino-terminal end of the peptidic molecules, in particular the hydrophobic amino acids like leucine and methionine, while the hydrophilic amino acid threonine, glutamic acid, tryptophan, arginine and cysteine are attached less accurately, thus the processing of such epitopes containing amino acids at the amino-terminal end is more difficultly done (5,20,22). On the other hand, if the epitope has proline as the second amino acid of the amino-terminal end, ERAP1 is not able to carry out the cleavage (23), the docking of that specific polypeptide at the catalytic site being made only in terms of disruption of G-A-M-E-N amino acids with catalytic role. It is thus observed that ERAP1 molecule presents high substrate specificity. This specificity is not limited only to the amino-terminus end of the molecule. The determination of the surface potential of the cavity formed by the second and fourth domains shows a highly electronegative area (21), which explains the affinity for the cleavage of the electropositive polypeptides previously observed (20).

ERAP1 can accept substrates up to 16 amino acids, having an optimum activity when the peptides are 10 amino acids long, while the peptidic substrates of less than 8 amino acids are not efficiently processed (24). This type of activity is not common among peptidases, which are usually more active against small peptides (22), but it is perfectly suited to the role that ERAP1 plays in antigen processing (24).

The proposed model for understanding this type of response to polypeptide chains considered large (9-16 amino acids) versus the response to small polypeptide chains (5-6 amino acids) is based on the existence of a regulatory site and a catalytic site. Thus, the binding of the carboxyl terminal end of the polypeptides to regulatory site causes a conformational change in the ERAP1 molecule which switches from the open form to the enzymatically active closed form. On the other hand the small peptides can not connect to the catalytic site and the regulatory site at the same time, therefore not being able to stabilize the closed conformation of the molecule and thus the small peptides will be processed at a much lower rate (21). (Fig. 2)

**FIGURE 2. Model for ERAP1 length-dependent cleavage activity**

a) short peptide substrate  
b) long peptide substrate

GENETIC ASSOCIATION WITH AS AND PATHOGENIC HYPOTHESES

Recent studies show that ERAP1 gene is highly polymorphic, with at least 70 nonsynonymous polymorphisms described Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. (dbSNP Human Build ID: 1420 – http://www.ncbi.nlm.nih.gov/SNP/). Many of these variations are very rare. The combination of the amino acids from these variable positions gives rise to multiple allotypes of ERAP1 protein, allotypes that have different trimming activity, mainly characterized as “normal”, “hypo” or “hiper” trimmers. Reeves et al. have found 13 different ERAP1 molecules in a group of 72 individuals (38 controls and 34 AS cases). The analysis of the combinations of these allotypes in pairs (one protein from each of the two chromosomes) in the studied population revealed that AS patients presented combinations not found in controls and vice-versa (3). More studies are needed to confirm these findings.

More Genome Wide Association Studies (GWAS) in 2007, 2010 and 2011 have shown a strong association between the most common functional variants of ERAP1 gene (25-27) and susceptibility to AS, identifying several single nucleotide polymorphisms including: rs2287987 (Met349Val), rs10050860 (Asp575Asn), rs30187 (Lys528Arg), rs17482078 (Arg725Gln) and rs27044 (Gln730Glu) which were confirmed in multiple subsequent case-control studies conducted in different populations (see Table 1).

A meta-analysis conducted in 2012 and focused on studies from both Europe and Asia, involving over 8,500 patients and more than 12,000 controls, confirmed several single nucleotide polymorphisms of ERAP1 gene to be associated with AS: rs27044 (OR 1.57, P < 0.001), rs17482078 (OR 1.271, P < 0.001), rs10050860 (OR 0.772, P = 0.006), rs30187 (OR 1.348, P < 0.001), rs2287987 (OR 0.746, P < 0.001) and rs27037 (OR 1.257, P = 0.001) (35).

Moreover, the association of these single nucleotide polymorphisms of ERAP1 with AS is valid only for HLA-B27 positive individuals (36, 37), which proves the consistent role of both genetic associations in AS. Reveille et al. showed that ERAP1 presents the most important association with AS after HLA-B27, with an attributable risk of 26% (36).

The data from a personal study demonstrates the association of rs30187 and rs27044 gene polymorphisms of ERAP1 gene with the risk of spondyloarthritis in general and AS in particular, in HLA-B27 positive Romanian individuals (34).

The two most common nucleotide polymorphisms associated with the AS, rs30187 and rs27044, cause changes in the structure of the ERAP1 polypeptide: arginine is replaced by lysine in 528 position (Lys528Arg) and glutamine replaces glutamic acid in 730 position (Gln730Glu), yet these changes are not directly connected with the catalytic site of ERAP1, similar to the changes determined by rs2287987 (Met349Val), for example.

However several explanations for the association of these polymorphisms with AS can be found. The single nucleotide polymorphisms rs30187 (Lys528Arg) causes an amino acid change localized in the domain III, and this seems to influence the conformational changes that occur during the transition from the open form to the closed, enzymatically active form of the molecule (18, 21).

Table 1.

<table>
<thead>
<tr>
<th>Population, Year</th>
<th>Study type</th>
<th>ERAP1 polymorphisms associated with AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom, 2007 (25)</td>
<td>GWAS</td>
<td>rs30187, rs27044, rs2287987, rs10050860, rs17482078</td>
</tr>
<tr>
<td>Australia, United Kingdom, USA, 2010 (26)</td>
<td>GWAS</td>
<td>rs27434, rs27037</td>
</tr>
<tr>
<td>Australia, United Kingdom, Canada, 2011 (27)</td>
<td>GWAS</td>
<td>rs30187</td>
</tr>
<tr>
<td>Portugal, 2009 (28)</td>
<td>Case-control</td>
<td>rs30187, rs27044</td>
</tr>
<tr>
<td>United Kingdom, 2009 (29)</td>
<td>Case-control</td>
<td>rs28366066, rs26653, rs2287987, rs27434, rs30187, rs10050860, rs469783, rs17482078, rs1065407, rs13167972</td>
</tr>
<tr>
<td>South Korea, 2010 (30)</td>
<td>Case-control</td>
<td>rs27044, rs30187</td>
</tr>
<tr>
<td>Spain, 2011 (31)</td>
<td>Case-control</td>
<td>rs17482078, rs30187, rs2287987, rs26653, rs10050860</td>
</tr>
<tr>
<td>China, 2011 (32)</td>
<td>Case-control</td>
<td>rs27434, rs27529</td>
</tr>
<tr>
<td>Iran, 2012 (33)</td>
<td>Case-control</td>
<td>rs30187, rs27434</td>
</tr>
<tr>
<td>Romania, 2013 (34)</td>
<td>Case-control</td>
<td>rs27044, rs30187</td>
</tr>
</tbody>
</table>
the carboxyl-terminus end of the peptidic antigen and can thus affect the substrate specificity of ERAP1.

The base for AS susceptibility is due to variations in the enzymatic activity regarding epitope processing that these allelic variants of ERAP1 present. In a 2011 study, Evans et al. demonstrates that rs30187 and rs17482078 alleles present a proteolytic activity decreased by 40% compared to the ancestral, wild-type allele (27). On the other hand, Evnouchidou et al. compared the enzymatic activity of ERAP1 gene variants, discovering that the same polypeptide can be processed at different speeds depending not only on the type of polypeptide precursor but also on its concentration (38). When the substrate concentration was low the rs27044 allele (Gln730Glu) processed the LSRHHAFSFR polypeptide 1.5 times faster than the ancestral allele, but when the substrate concentration was increased, probably due to inhibition of substrate, the same polypeptide was processed more slowly by the rs27044 allele (Gln730Glu) compared with the ancestral allele (38).

It is assumed that in the endoplasmic reticulum, where ERAP1 operates, the peptide loading complex (PLC) – a macromolecular structure to which ERAP1 belongs, concentrates the antigen precursors in its vicinity, which causes the enzyme to work in saturation conditions and therefore its activity to be influenced by substrate inhibition, the allelic variant rs27044 (Gln730Glu) having therefore, an activity decreased by 50% compared to the ancestral variant in vivo (38).

Studies on animal models of transgenic mice lacking the ERAP1 gene shows that the antigen repertoire presented by MHC I to CD8 + lymphocytes and natural killer (NK), is different when compared with the normal mice. Thus, the antigenic peptides presented by MHC I have an increased length, leading to the activation of a certain CD8 + lymphocyte population which specifically recognises the amino-terminus of these epitopes (39). Also, the quality of the complex peptide - MHC I molecules is altered, on one hand the loss of a certain complexes peptides - MHC I with up to 20% takes place, and on the other hand, new complexes peptide - MHC I molecules emerge in the absence of ERAP1 gene but also in the absence of the tapasin encoding gene (40). This suggests that certain peptides which are normally removed become present on the cell surface expressed by MHC I, when ERAP1 levels are inadequate. In wild-type mice (mice which have the ERAP1 gene present) immunized with ERAP1 deficient cells, the antigenic repertoire associated with both classical MHC I molecules and non-classical MHC Ib molecules is immunogenic and elicits a CD8 + cytotoxic T lymphocytes cell-mediated immune response. Moreover, the CD8 + T lymphocytes recognize the novel peptides presented by MHC I and Ib on the surface of the ERAP1 gene deficient cells, removing them both in vitro and in vivo (41, 42). ERAP1 gene variants may contribute to the pathogenesis of AS, by generating altered peptide-HLA I complexes, which will lead to an abnormal immune response mediated by CD8 + T lymphocytes and natural killer.

It seems that abnormal processing of antigenic peptides by allelic variants of ERAP1 gene can lead to unstable peptides-HLA-B27 complexes, which can oligomerize and form another complexes with chaperone proteins in the endoplasmic reticulum, property called misfolding, which leads to accumulation of folded HLA-B27 heavy chains, triggering the unfolded protein response (UPR), and finally, the onset of an inflammatory response mediated by IL-1, TNF-α (43), which may explain the persistence of the inflammatory process in AS. A recent study though seems to refute this working hypothesis (44). ERAP1 mononucleotide polymorphisms can influence not only the expression of HLA-B27-peptide complexes, but also the expression of the free heavy chains of HLA class I (including HLA-B27) on the surface of monocytes (45). This can lead to an abnormal immune response following the binding of these free heavy chains to the killer cell immunoglobulin-like receptors (KIR) belonging to the T lymphocytes and natural killer.

Also, ERAP1 can remove peptides that are not well loaded onto MHC I molecules in the endoplasmic reticulum (46) which can be a potential mechanism through which ERAP1 is involved in AS.

Without a doubt there is a long road ahead until the complete understanding of the AS pathogenesis, but the discovery of new associations with genes outside the major histocompatibility complex, such as ERAP1 or IL-23R, may open not only new perspectives in understanding the intimate mechanisms of the disease, but it may also suggest new therapeutic approaches.

Acknowledgement

This paper is supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number SOP HRD/159/1.5/S/135760.
REFERENCES


