Review

Endoplasmic reticulum aminopeptidases in the pathogenesis of ankylosing spondylitis

Tony J. Kenna¹, Philip C. Robinson² and Nigil Haroon³,⁴

Abstract
There has been significant progress in our understanding of the pathogenesis of AS. The advent of genome-wide association studies has increased the known loci associated with AS to more than 40. The endoplasmic reticulum resident aminopeptidases (ERAP) 1 and 2 were identified in this manner and are of particular interest. There appears to be a genetic as well as a functional interaction of ERAP1 and 2 with HLA-B27 based on the known functions of these molecules. Recent studies on the structure, immunological effects and the peptide-trimming properties of ERAP 1 and 2 have helped to provide insight into their pathogenic potential in AS. In this review, we explore the role of ERAP 1 and 2 in the pathogenesis of AS.

Key words: aminopeptidase, endoplasmic reticulum resident aminopeptidase 1, endoplasmic reticulum resident aminopeptidase 2, endoplasmic reticulum, unfolded protein response, T cell response, NK cell, peptide, misfolding.

Rheumatology key messages
- With recent technical advances and formation of large genetics consortia, multiple genetic loci associated with AS have been identified.
- There is a genetic and functional interaction between endoplasmic reticulum aminopeptidases and HLA-B27 in the pathogenesis of AS.
- Novel therapeutic agents for AS could be developed by functional studies and targeting molecular pathways identified by novel genetic associations.

Introduction
AS is the most common member of the SpA family of musculoskeletal diseases, affecting 0.1–1.4% of the global population [1]. AS is a chronic inflammatory arthritis predominantly affecting the spine and sacroiliac joints. AS causes joint pain and stiffness that can result in joint fusion and spinal deformities with a significant impact on patients’ quality of life. Disease onset is typically in the third or fourth decade of life and, thus, AS is considered to have a greater patient impact than inflammatory disorders with later development, such as RA. AS shares overlapping clinical manifestations with other SpAs, including acute anterior uveitis, reactive arthritis, psoriatic SpA, IBD-associated arthritis and Behçet’s disease [2]. Inflammation is a critical component of AS that precedes bone formation, but understanding of the inflammatory processes in AS is limited. Currently, there is no cure for AS and the available treatment modalities suppress pain and inflammation in a proportion of patients only.

The pathogenesis of AS is not well understood. Family and twin studies demonstrate a strong role for genes in the pathogenesis of AS. The strongest genetic association of AS is with HLA-B27, a MHC-I molecule involved in antigen presentation [3]. Compared with 4–7% prevalence of HLA-B27 in control subjects, AS patients have 80–90% prevalence of the gene [3, 4]. Despite >40 years of research since this discovery, we are not clear how HLA-B27 contributes to the pathogenesis of AS. Following the advent of genome-wide association studies, the number of known loci associated with AS is now in excess of 40. Endoplasmic reticulum aminopeptidase 1 (ERAP1) is second only to HLA-B27 in the strength of association with AS. This was a major breakthrough in AS genetics,
and recently, the structurally and functionally related aminopeptidase ERAP2 was also found to be associated with AS. Other major genes found to be associated with AS include IL23R, IL12R, I2, prostaglandin E receptor 4 (subtype EP4; PTGER4), leucyl-cystinyl aminopeptidase (LNPEP) and TNF-receptor superfamily member 1A. A more complete list of genes is available in the review by Tsui et al. [5].

The association of AS with ERAP1 and 2 is especially interesting because they can act in the same pathogenic pathway as HLA-B27. Several important and interesting studies on the functional role of ERAP1 and ERAP2 have been published recently. In this review, we provide a critical analysis of recent literature and propose a pathogenic model for AS that includes ERAP1 and 2 and HLA-B27.

Genetic associations of the aminopeptidases

The first described association of the aminopeptidase ERAP1 gene with disease was in cervical cancer [6]. Subsequently, haplotypes containing ERAP1 and ERAP2 have been associated with eight immune-mediated diseases: AS, non-AS SpA, type 1 diabetes mellitus, psoriasis, multiple sclerosis, IBD, JIA subtype enthesitis-related arthritis, anterior uveitis and birdshot chorioretinopathy [7–16]. In addition, ERAP1 has been associated in genetic studies with resistance to HIV and in functional experiments with resistance to the intracellular pathogen Chlamydia [17, 18]. This range of associated infections, immune-mediated diseases and cancer indicates the previously unappreciated potential importance of these molecules.

One of the challenges of going from associated single nucleotide polymorphisms (SNPs) to causal genes is the linkage disequilibrium in associated areas. For example, the associated region that spans the ERAP2 gene also spans the LNPEP gene that encodes the insulin-regulated aminopeptidase (IRAP). There are two strong functional candidates in ERAP2 (see below) and none identified to date in LNPEP [19]; therefore, the focus of attention has been on ERAP2, although without further evidence we cannot exclude the association of LNPEP [20, 21]. IRAP is involved in a process known as cross-presentation, whereby exogenous antigens are brought into the cell and presented on MHC class I molecules to the cells of the immune system [22]. This molecule and its important function in trimming peptides prior to presentation in this pathway could very plausibly be related to disease.

Specific haplotypes harbouring the same variants are consistently associated across different diseases. For example, the ERAP1 haplotype tagged by the rs30187 variant is associated with AS, non-AS SpA, type 1 diabetes mellitus, psoriasis, multiple sclerosis, anterior uveitis and enthesitis-related arthritis. In ERAP2, the haplotype tagged by rs2248374 is associated in AS, IBD and birdshot chorioretinopathy. Many associations have been demonstrated across multiple ethnic groups; for example, ERAP1 has been associated in Europeans, Canadians, Hungarians, Portuguese, French, Spanish, Belgians, Koreans and Han Chinese [7, 13, 16, 23–26]. This is consistent with a common founder population. There are examples in genetics where risk alleles are not shared across ethnic groups, such as the rs11209026 IL23R association that is a risk for AS in Europeans and not in Han Chinese, but this stems from the SNP not being polymorphic in the Chinese population [27]. It would be difficult to contemplate different underlying mechanisms of disease pathogenesis in different ethnic groups. These shared disease associations and implicated haplotypes in different ethnic groups along with gene–gene interactions (see below) strengthen the circumstantial evidence that these associations cause disease in the same or a similar way.

Genetic interaction with classical MHC alleles

Gene–gene interactions in human genetics are rare and have only been reported with ERAP1 and a number of classical MHC alleles. ERAP1 alleles in AS and psoriasis are only disease associated if a specific class I MHC allele is present, HLA-B*27 or HLA-Cw*6, respectively [8, 28]. In Behcet’s disease, a recessive interaction exists with the rs17462078 haplotype and HLA-B*51 [29]. This strongly implicates the MHC class I presentation pathway as the mechanism for disease association at these loci in these diseases. Initially, only HLA-B*27-positive patients were definitely described to have associations in ERAP2 in AS [30]. However, now in both HLA-B*27-positive and HLA-B*27-negative patients the ERAP2 region is associated with disease (P. Robinson, unpublished data).

ERAP1 and AS

Two haplotypes of ERAP1 have been associated with AS, tagged by rs10050860 and rs30187. Using a Bayesian approach, it was determined that these variants are also the most likely disease-causative variants. Each SNP has been mapped to an important part of the enzyme and has been demonstrated to have functional consequences for the enzyme (detailed below) [28, 31]. In ERAP1, the two independent associations are common, with the minor allele frequency of rs30187 being 0.30–0.45 and for rs10050860 being 0.10–0.20 depending on the population. This suggests that these variants either did not pose strong selective pressure in the past or were the subject of balancing selection, as has been suggested with ERAP2. SNPs in ERAP1 have been identified to be expression quantitative trait loci (eQTLs), which means that the SNPs control the level of protein expression. Furthermore, the eQTL SNPs are also associated with AS [32]. This eQTL association could result either from changes in the level of enzyme expression driving AS pathogenesis or from linkage with other causal SNPs. Experiments knocking down ERAP1 in different systems have led to both increased and decreased levels of MHC on the cell surface [33, 34]. Dendritic cells isolated from AS patients show elevated ERAP1 protein expression but no alteration to MHC class I surface expression compared with healthy individuals [35]. But the potential functional consequences of variability in cellular enzyme levels are not fully understood.
ERAP2 and AS

Variation at the rs2248374 SNP in ERAP2 (A/G) controls absolute expression of the ERAP2 protein. This SNP has been identified as influencing slicing of exon 10, and the G allele of this SNP causes elongation of exon 10 that subsequently includes two TAG stop codons [20]. The mRNA has been shown to undergo nonsense-mediated decay, a quality-control process that destroys incorrect mRNA. Therefore, haplotypes with the G allele produce no ERAP2 protein and homozygotes have absent ERAP2 protein. This absence of protein appears to be protective in AS, IBD and birdshot chorioretinopathy [10, 11, 30]. At ERAP2, the frequency of the two main haplotypes is equal and there is evidence of balancing selection on the resultant allele frequencies of ERAP2 [20]. For a gene that encodes an integral part of the MHC class I presentation pathway, this could suggest selective pressure from infectious agents in the past.

Function and pathogenic role of ERAP1 in AS

It remains unclear how ERAP1 contributes to the pathogenesis of AS. Two functions have been proposed for ERAP1 [36]. *In vitro* studies reported that ERAP1 cleaves cytokine receptors from the cell surface, including IL-1RII, IL-6Ra and TNF-R1, thereby preventing intracellular cytokine signalling [37–39]; however, this so-called receptor sheddase function of ERAP1 was not supported by studies of cytokine receptor shedding in AS patients or ERAP1 knock-out mice [28, 40]. The second function is that of a molecular ruler, whereby ERAP1 is responsible for trimming of peptides that have been processed through the proteasome from 13 to 15 amino acids in length down to 8 or 9 amino acids, the optimal length for presentation on MHC-I [41].

ERAP1 displays trimming preferences depending on the composition of the peptide being trimmed. Changes at the N-terminus, the C-terminus and the internal sequence of the peptide influence its trimming ability [42–44]. ERAP1 exists in either a closed (enzymically active) or an open (enzymically inactive) state. X-ray crystallography revealed that one of the main AS-associated ERAP1 SNPs, rs30187, is located at the hinge region of the enzyme, which controls the switch between open and closed conformations and likely alters enzyme activity [31]. Indeed, the protective variant at rs30187 is associated with reduced enzyme activity *in vitro* [32].

ERAP1 is associated with AS only when HLA-B*27 is present; therefore, models for how ERAP1 operates in AS must align with the proposed roles for HLA-B*27 in AS pathogenesis. The following three classical models exist to explain how HLA-B*27 contributes to AS.

Arthritogenic peptide theory

Altered or inappropriate, MHC: self-peptide complexes are presented to the immune system and recognized as harmful or foreign, which elicits a self-reactive inflammatory response. The arthritogenic peptide theory suggests that altered properties of HLA-B*27 or the peptides presented by HLA-B*27 are necessary for initiation of inflammatory response. In support of this, some subtypes of HLA-B*27 are not associated with AS. These subtypes have amino acid changes in their binding groove, which means that the peptide repertoire they present is different from that of AS-associated HLA-B*27 subtypes [45].

Endoplasmic reticulum stress and the unfolded protein response hypothesis

The unfolded protein response (UPR) hypothesis is another theory of AS aetiology involving HLA-B*27. HLA-B*27 heavy chains have been observed to fold slowly and remain in the endoplasmic reticulum (ER) for extended periods of time compared with other class I MHC molecules [46]. This HLA-B*27 misfolding can precipitate a cellular response called the UPR. The UPR is a homeostatic mechanism initiated by the cell in an attempt to return itself to a normal state [47].

HLA-B*27 homodimer formation

HLA-B*27 can also form disulphide bonds with other proteins and homodimerize, and these homodimers have been observed on the cell surface [48]. Notably other MHC class I alleles have been shown to form homodimers on the cell surface, namely HLA-B07 and HLA-G, but not to the same degree as HLA-B*27 [49]. HLA-B*27 homodimers are recognized by killer-cell immunoglobulin-like receptors. Antigen presenting cells expressing HLA-B*27 homodimers can stimulate the proliferation of Th17 cells that carry a killer-cell immunoglobulin-like receptor that binds HLA-B*27 homodimers. These antigen presenting cells bearing HLA-B*27 homodimers also stimulate IL-17 production in the Th17 cells, and these Th17 cells account for most of the IL-23R expression peripherally. Their production of IL-17 was also increased by IL-23 [50].

ERAP1 polymorphisms may contribute to all three models for how HLA-B*27 operates in AS. Altered rates of peptide trimming mediated by ERAP1 could lead to cell surface presentation of aberrant peptides in the context of HLA-B*27 (Fig. 1A). Recently, ERAP1 variants have been reported to alter levels of HLA-B*27 cell surface free heavy chains [51]. Altered enzyme activity could affect the rate at which HLA-B*27:peptide complexes fold within the ER, thereby contributing to the UPR and ER stress (Fig. 1B). Recent data, however, cast some doubt over this model [52, 53]. Indeed, it has recently been suggested that autophagy rather than UPR regulates expression of inflammatory cytokines in AS patients [54]. The surface B27 homodimers, formed as a result of abnormal ERAP1 trimming and misfolding of HLA-B27, can interact with killer-cell immunoglobulin-like receptors of NK cells or Th17 cells, contributing to AS pathogenesis (Fig. 1B).

While all three models may, at least in part, contribute to AS pathogenesis, the arthritogenic peptide theory is probably the one that lends itself most readily to coordinated effects of HLA-B*27 and ERAP1. Altered rates of peptide trimming mediated by AS-associated ERAP1 variants likely alter the HLA-B*27:peptide complex presented to...
immune cells, facilitating recognition of peptidomes only presented in genetically predisposed individuals.

Peptide presentation and infection

**ERAP1** is associated with AS only when **HLA-B*27** is present. Changes in ERAP1 function influenced by the associated variants rs30187 and rs10050860 influence the trimming rate of the enzyme [28]. Studies on the peptidomes of cell lines expressing different ERAP1 variants demonstrate that these variants influence the molecular weight, length and amount of peptide presented [55]. There were also observed differences in the thermostability of the HLA-B*2704 molecule dependent on the ERAP1 variant expressed in the cell. As shown previously, reduced ERAP1-mediated trimming leads to increased HLA-I misfolding and decreased thermostability of HLA-B*2704 [51, 55]. These observations need further development, but these data certainly suggest that the repertoire of antigens being presented is significantly different between the ERAP1 variants.

The change in peptide repertoire that ERAP1 variants impose could therefore be the driving force behind immune dysfunction. *In vivo* studies have revealed how ERAP1 shapes the antigenic peptide repertoire against microbial and viral infections. This is most clearly demonstrated in ERAP1-deficient mice infected with the pathogen *Toxoplasma gondii*. Mice lacking ERAP1 are unable to trim and present immunodominant and appropriate-length peptides derived from *T. gondii*, leading to aberrant peptide presentation, an inability to mount an effective immune response to infection and lethality due to infection [17]. ERAPs also play an important role in response to viral infections. ERAP1 knock-out mice infected with lymphocytic choriomeningitis virus present dramatically different immunodominant epitopes derived from lymphocytic choriomeningitis virus for presentation compared with wild-type mice [56]. Human cytomegalovirus has evolved an immune evasion strategy whereby a micro-RNA, miR-US4-1, is expressed and specifically targets ERAP1, down-regulating expression and altering peptide trimming and presentation [57]. Interestingly, it was recently shown by comparing HLA-B*27 and B7 mice that ERAP1 could influence the host immune response in an HLA-allele-specific manner [58].

Reduced ERAP1 expression not only affects peptide presentation during infection, but also significantly alters the composition of the endogenous MHC-I peptidome. Murine studies have shown that lack of ERAP1 led to unstable and structurally unique peptide-MHC complexes, and that these complexes elicited a potent CD8+ T cell response [34, 59]. Taken together, this suggests that ERAP1 plays a critical role in overall immune system modulation, with systemic effects. Looking at individual polymorphisms in **ERAP1**, loss-of-function variants seem to be protective. However, there are multiple variations of ERAP1 in the same AS patient. It may not be appropriate to look at individual ERAP1 variants and extrapolate the functional effects, because the net result of multiple SNPs may vary. This was recently demonstrated in a study
looking at ERAP1 allotype pairs that occur in AS patients compared with healthy and non-AS disease controls [60]. The ERAP1 allotypes seen in AS patients compared with controls had decreased peptide-trimming activity. This is in sharp contrast to the previously held notion that decreased ERAP1 activity is protective in AS. It is possible that any altered ERAP1 activity (higher or lower) can result in changes in peptide processing, with pathological consequences [60].

The theory that altered ERAP1 activity contributes to systemic immunodeficiency certainly requires additional experimental validation. How the changes in peptide presentation may influence the immune control of gut microflora is one theory that deserves attention, especially in view of the strong association between AS and bowel inflammation. Whatever the mechanisms of action of ERAP1 in AS, a better understanding of them may provide therapeutic target(s) in AS.

**ERAP2 in health and disease**

ERAP2 is a 110 kDa glycoprotein resident in the ER that helps to trim peptides for MHC-I presentation. There is 50% sequence identity between ERAP1 and ERAP2 [61, 62]. ERAP2 exists in homodimers or can form heterodimers with ERAP1 [61, 62]. ERAP2 potentially works in concert with ERAP1 in trimming peptides and, although the location and functions of ERAP1 and ERAP2 overlap, the substrate specificities of the two enzymes seem to differ [61]. While ERAP1 preferentially trims peptides with N-terminal hydrophobic amino acids, ERAP2 prefers positively charged amino acids [63]. Suppression of ERAP1 or ERAP2 alone reduces the MHC-I surface expression by 10% each, while combined suppression of ERAP1 and ERAP2 leads to a 20% drop in the level of surface MHC-I expression [61].

Balancing selection potentially resulted in maintenance of a shorter form of ERAP2 that undergoes nonsense-mediated decay, resulting in the absence of protein expression [20]. There are two main haplotypes of ERAP2, with haplotype A resulting in ERAP2 expression while haplotype B results in the absence of ERAP2 [20]. The frequency of these haplotypes is almost equal (50%), and therefore 25% of the population that is homozygous for haplotype B has no ERAP2 expression. Likewise, the non-synonymous SNP rs2549782 leads to a lysine/arginine switch at position 392 (K392N). ERAP2 (K392N) can trim both hydrophobic and positively charged N-terminal amino acid-containing peptides, while K392 preferentially trims peptides with positively charged N-terminal amino acids [21].

The ERAP2 SNP rs2549782 was initially reported to be associated with AS in a haplotype including ERAP1 [16]. Subsequently, genome-wide association studies have established the ERAP2-AS link [30]. The region of ERAP2-LNPEP is in strong linkage disequilibrium, and the ERAP2 rs2549782 risk allele T tags haplotype B of ERAP2 [64]. Another study looking exclusively (without correcting for ERAP1 polymorphisms) at the ERAP2 SNP rs2248374 (which leads to the ERAP2-null phenotype) did not find any association with AS [65].

ERAP2 has been linked to pre-eclampsia, and this is thought to be due to altered enzymic processing of vasoactive intermediates, such as vasopressin [66]. Interestingly a 6-fold decreased expression of ERAP2 was reported from chorionic villus samples of women who went on to develop pre-eclampsia on follow-up [67]. It was recently shown that placental tissue from patients with pre-eclampsia has higher gene expression of ERAP1, ERAP2 and LNPEP [68].

ERAP2 is also known to be associated with psoriasis, Crohn’s disease and paediatric IBD. ERAP2 (haplotype A) presence and HLA-B57 have been shown to be protective against HIV infection [64]. The lack of ERAP2 expression as seen in patients with haplotype B leads to decreased antigen presentation and HIV infection [64]. There are parallels here with the ERAP1–MHC class I association reported in other conditions, such as AS, Behc¸et’s disease and psoriasis [29]. Thus, both ERAP1 and ERAP2 seem to be pathogenic because of altered interactions with MHC-I. Functional studies on ERAP2 have not been widely published. This could be due to the absence of ERAP2 in mice. The available data point towards abnormal HLA-B27 expression, with an increase in MHC-I free heavy chains and activation of the UPR response [69].

While there is no doubt that further progress in elucidating the role of ERAP1 and ERAP2 will come from in vitro and in vivo functional investigations, the role of other types of variation is unknown. Additional classes of variation, such as epigenetic, large insertions and deletions and copy number variation, may also play a role in modifying the effect of these disease-associated SNP variants or inducing other disease associations at these loci. Additional investigations are required to explore this further, along with the continuing growth of disease cohorts included in studies.

**Summary and conclusions**

ERAP1 and ERAP2 are aminopeptidase enzymes residing in the ER that shape the MHC–peptide complex. Genetic and functional interactions of these genes with MHC-I molecules are pathogenic in several diseases, including AS. Concerted efforts to understand this interaction in greater detail could help to unravel the pathogenesis of AS.

**Funding:** No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

**Disclosure statement:** The authors have declared no conflicts of interest.

**References**

Tony J. Kenna et al.


33 York IA, Chang SC, Saric T et al. The ER aminopeptidase ERAP1 enhances or limits antigen presentation by...
49 Dangoria NS, DeLay ML, Kingsbury DJ et al. HLA-B27 misfolding is associated with aberrant intermolecular disulfide bond formation (dimerization) in the endoplasmic reticulum. J Biol Chem 2002;277:23459–68.
62 Birtle JR, Saridakis E, Stratikos E, Mavridis IM. The crystal structure of human endoplasmic reticulum aminopeptidase 2 reveals the atomic basis for distinct roles in antigen processing. Biochemistry 2012;51:286–95.


