FEBS Letters 585 (2011) 3619–3626

journal homepage: www.FEBSLetters.org

Review

New insights into the functional role of the rheumatoid arthritis shared epitope

Denise E. de Almeida, Song Ling, Joseph Holoshitz *

Department of Internal Medicine, University of Michigan Medical Center, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 28 February 2011
Revised 14 March 2011
Accepted 15 March 2011
Available online 22 March 2011

Edited by Richard Williams, Alexander Flügel and Wilhelm Just

Keywords:
Calreticulin
Nitric oxide
Oxidative stress
Rheumatoid arthritis
Shared epitope
Th17

ABSTRACT

The shared epitope (SE) – an *HLA-DRB1*-encoded 5-amino acid sequence motif carried by the vast majority of rheumatoid arthritis (RA) patients – is a risk factor for severe disease. The mechanistic basis of RA-SE association is unknown. This group has previously demonstrated that the SE acts as a signal transduction ligand that activates nitric oxide and reactive oxygen species production. SE-activated signaling depends on cell surface calreticulin, a known innate immunity receptor previously implicated in immune regulation, autoimmunity and angiogenesis. Recent evidence that the SE enhances the polarization of Th17 cells, which is a key mechanism in autoimmunity, is discussed highlighting one of several potential functional effects of the SE in RA.

1. Introduction

Rheumatoid arthritis (RA) affects 0.5–1.0% of the population [1]. The disease is characterized by chronic inflammatory changes in both articular and extra-articular tissues. Due to its high prevalence and debilitating nature, RA inflicts a major economic burden on society. In recent years it has been realized that in addition to causing pain and disability, the disease significantly shortens life expectancy due to accelerated atherosclerosis [2].

Although genes play a major role in RA risk, the disease appears to have low sibling occurrence with a concordance rate of 12–15% in monozygotic twins [3]. Overall, the contribution of genetic factors to RA risk is calculated at approximately 60%, while the remaining 40% are believed to be contributed by environmental factors. The observations that RA is more common in urban versus rural populations, a recent decline in the incidence of the disease in high-incidence of populations, and the effect of birth cohort on disease incidence are all indirectly supporting environmental influences. Importantly, over the past few years it has been conclusively shown that the disease is strongly associated with environmental pollutants, such as cigarette smoking [4].

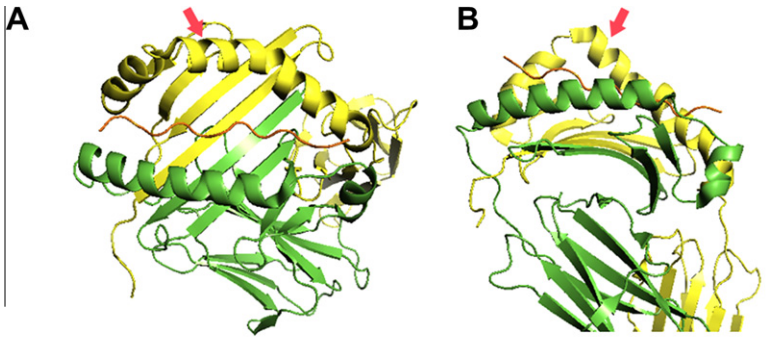
Among the genetic risk factors, the *HLA-DRB1* locus is the most significant one. RA has long been shown to associate with human leukocyte antigen (HLA) genes. The pioneering studies on HLA-RA association were carried out in the late 1970s by Peter Stastny [5] and Robert Winchester's group [6], who independently concluded that HLA-DR4 is significantly more common among patients with RA. It was subsequently found that other HLA-DR serotypes, for example, HLA-DR1 in Mediterranean, or HLA-DR14 in Native Americans, are also associated with the disease. With the advent of modern DNA sequencing techniques it had become apparent in the 1980's that there is no RA-specific HLA-DR sequence. Instead, it was found that the majority of RA patients share a short sequence motif coded by several *HLA-DRB1* alleles. This revelation had prompted the *Shared Epitope Hypothesis* [7].

2. The RA shared epitope

The term "shared epitope" (SE) most commonly refers to a five amino acid sequence motif in residues 70–74 of the DRβ chain coded by several *HLA-DRB1* alleles that are over-represented among RA patients (Fig. 1). The SE motif consists of three homologous amino acid sequence variants: (1) QKRAA, the SE variant that is the most common motif among Caucasian, is coded primarily by the *HLA-DRB1*∗0401* allele; (2) The second most common motif, QRRAA, is coded by several alleles, among them *HLA-DRB1*∗0404*, *HLA-DRB1*∗0101*, and *HLA-DRB1*∗0405*; (3) The third motif, RRRAA, coded by allele *HLA-DRB1*∗1001*, is the rarest. In addition to increasing RA risk, SE-coding *HLA-DRB1* alleles have been shown to associate with more severe disease [8] and to exhibit allele-dose

* Corresponding author. Address: University of Michigan School of Medicine, 5520D MSRB1, 1150 W Medical Center Drive, Ann Arbor, MI 48109-5680, USA. Fax: +1 734 763 4151.
E-mail address: jholo@umich.edu (J. Holoshitz).

0014-5793/$36.00
doi:10.1016/j.febslet.2011.03.035



C RA Association with *HLA-DRB1* Alleles in Different Populations

Population	*HLA-DRB1*	AA Sequence (70-74)
North European, North American, Caucasoid	**0401*	QKRAA
Caucasoid	**0404*	QRRAA
Japanese East Asian	**0405*	QRRAA
Mediterranean, Caucasoid	**0408*	QRRAA
West European, Asian, Caucasoid	**0101*	QRRAA
Native American	**1402*	QRRAA
Asian, Mediterranean, African	**1001*	RRRAA

Fig. 1. Structure and epidemiology of the RA SE. (A) 'Top' view of a SE-expressing (HLA-DR1) molecule. The α chain is shown in green, the β chain is in yellow and the groove peptide is in brown. The red arrow points at the SE-containing α helical loop. (B) 'Side' view of the same molecule. Note the localization of the SE near the β chain 'kink'. (C) Ethnic and geographic distribution of RA-associated *HLA-DRB1* alleles and their SE products.

effect, i.e., patients with 2 SE-coding alleles tend to experience more severe disease than patients with 1 allele, who, in turn, have more severe RA than SE-negative patients.

The mechanism underlying the effect of the SE is unclear. Based on the known role of MHC class II molecules in antigen presentation, the prevailing paradigms postulate that presentation of arthritogenic self-peptides [9], molecular mimicry with foreign antigens [10], or T cell repertoire selection [11] are involved. While these hypotheses are all plausible, they are difficult to reconcile with the fact that data supporting antigen-specific responses as the primary event in RA are inconclusive. Additionally, several other human diseases have also been shown to be associated with SE-encoding *DRB1* alleles, including polymyalgia rheumatica [12], giant cell arteritis [12], Type I diabetes [13], erosive bone changes in psoriatic arthritis [14] and lupus [15], autoimmune hepatitis [16] and early-onset chronic lymphoid leukemia [17], among other conditions. The SE is also associated with spontaneous arthritis in dogs [18] and, in *HLA-DRB1*0401* transgenic mice it increases the incidence of spontaneous diabetes [19] and the severity of both collagen-induced arthritis (CIA) [20] and experimental autoimmune encephalomyelitis (EAE) [21]. Thus, although it is best known for its involvement in RA, the SE associates with several pathogenically unrelated diseases and experimental disease models, and its effect seems to lack antigen- or species-specificity. These promiscuities are incongruent with fundamental tenets of MHC-restricted antigen presentation theory.

3. Activation of innate signaling by the SE: a new paradigm

Given the inconsistencies of SE-RA association with antigen presentation-based theories, over the past few years, our laboratory has examined an alternative hypothesis concerning the role of the SE in RA [22–28]. Based the known tri-dimensional homology among products of the MHC gene family, we postulated that similar to class I MHC-coded molecules [29], the SE may be acting as a ligand that can trigger innate immune signaling. The rationale of this antithetic hypothesis relates to the fact that the SE is located near the apex of α helical tri-dimensional structural motif that has been preserved throughout the entire MHC gene family and seems to be enriched in signal transduction ligands.

The first crystal structure of a class II MHC molecule, published in 1993 by Don Wiley's group [30], revealed a remarkable tridimensional similarity to a previously reported class I MHC molecule. The degree of the similarity was surprising, given a substantial evolutionary distance between the two molecules and the fact that the peptide-binding groove in class I molecules is coded by a single gene, while in class II it is formed jointly by the products of two distinct genes. The extent of evolutionary 'choreography' required to bring these two disparate MHC molecules to form a near-identical tri-dimensional structure, is staggering. One of the notable features of the similarity is a 'kink' in the α2 domain of the class I MHC molecule, which could be almost perfectly superimposed on a similar structure in the β1 domain of the class II molecule. The 'kink' region in both molecules involves allele-diversity regions. Subsequent crystal analyses have shown very similar tridimensional structures in the entire MHC gene product family, irrespective of whether or not they can present antigens [31]. In all cases, this region forms a sharp protrusion 'above' the MHC groove plane (Fig. 1).

The remarkable conservation of a similarly-shaped 'kink' in the midst of allele diversity regions in MHC molecules independent of their antigen presentation capabilities suggests that this region may possess important allele-specific, conformationally-dependent, non-antigen presentation functions. Indeed, there are some indications that this region performs such functions. For example, in both classical and non-classical (HLA-E) class I MHC molecules, this region contains ligands for natural killer (NK) cell receptors [32]; in HFE (an empty-grooved human class I-like molecule), it interacts with transferrin receptor [33]; In M10 (a mouse class I-like molecule), the same region has been proposed as an interaction site with a pheromone receptor [34].

These considerations have led us to pursue a novel hypothesis which postulates that similar to its structural homologue in the class I MHC molecule, the SE functions as a signal transduction ligand that interacts with an evolutionarily-conserved receptor. Our