The handle http://hdl.handle.net/1887/33221 holds various files of this Leiden University dissertation

**Author:** Heemst, Jurgen van  
**Title:** Towards a molecular basis for the association of HLA, rheumatoid arthritis and autoantibodies  
**Issue Date:** 2015-06-17
CHAPTER 1

Introduction

Adapted from:

Genetics of rheumatoid arthritis: What have we learned?
*Immunogenetics.* 2011 Aug;63(8):459-466

and

HLA and rheumatoid arthritis: How do they connect?
THE ADAPTIVE IMMUNE SYSTEM

During life, every day, we encounter many different micro-organisms including potentially pathogenic bacteria, viruses, fungi and parasites. The immune system is critically important in the defence against these pathogens. The immune system has many different components. These include a wide variety of immune cells, but also non-cellular effector mechanisms. The immune system can be roughly subdivided in two components: the innate and the adaptive immune system. Upon primary infection, the innate immune system is the first to act. The response is generally fast and unspecific. After a few days the adaptive immune system comes into play. This response is highly potent and extremely specific. In addition, after clearance of the pathogen, the adaptive immune system establishes memory resulting in a highly potent and fast specific immune response upon secondary infection.

The large specificity of the adaptive immune system is mainly established by two cell typed, the B and T cells. These lymphocytes express B or T cell receptors, highly specific receptors with a very large diversity. B and T cells differ in the way they recognize their antigen. B cells recognize their antigen in its native form. Antigen-recognition combined with different activation signals allows for B cell activation and their differentiation in antibody secreting plasma cells. T cells can provide these activation signals and are therefore critically important for the generation of potent antibody responses. In addition, T cells also have other effector functions including killing of infected cells and the production of inflammatory cytokines. In contrast to B cells that recognize native antigens, T cells only recognize processed antigens presented by antigen presenting molecules, predominantly by human leukocyte antigens (HLA) [1].

Due to the potency and specificity of adaptive immune response, B and T cells have to be strongly educated to distinguish self from non-self. Nevertheless, autoimmune diseases (AID), diseases characterized by a failure of the immune response to distinguish self from non-self, are highly prevalent in the western world. In the United States, the prevalence was estimated between 5-8% of the population [2]. B and T cells are implicated in the development of most of these AID.

HUMAN LEUKOCYTE ANTIGENS

HLA molecules present processed antigens to T lymphocytes. In the presence of foreign proteins, e.g. during infection, foreign peptides are presented by HLA molecules. Recognition of such HLA-peptide complexes by T cells results in T cell activation and its accompanying effector mechanisms. HLA molecules can be subdivided in two subclasses: HLA class I and class II molecules. HLA class I molecules present peptides derived from an intracellular source. T lymphocytes that are
specific for antigens presented by HLA class I molecules are distinguished by the expression of CD8, hence their name CD8+ T cells. HLA class II molecules present peptides derived from an extracellular source. T lymphocytes that are specific for HLA class II molecules express CD4, hence their name CD4+ T cells. The classical HLA class II molecules are named HLA-DQ, HLA-DR and HLA-DP and are composed of an alpha and a beta chain. Peptides are presented at a site of interaction between these two chains [3].

The genes encoding for HLA molecules are located in a region in chromosome 6 (Figure 1). The region has high linkage disequilibrium (LD) and the genes are highly polymorphic, resulting in a large variety of HLA molecules. Polymorphic residues are mainly located at sites important for antigen presentation resulting in large differences in peptide repertoire between different HLA molecules [4, 5].

Expression

HLA class I and HLA class II molecules differ in their expression pattern. HLA class I molecules are expressed by nearly all nucleated cells. In contrast, the expression of HLA class II molecules is more tightly regulated. Under physiological conditions, HLA class II molecules are predominantly expressed by professional antigen-presenting cells, including B lymphocytes, dendritic cells and macrophages. In inflammatory conditions, including in AID, other tissue cell types can also present HLA class II molecules [6].

Tolerance and autoimmunity

T cell are highly specific and potent. Therefore, it is very important to prevent their activation in response to processed self-antigens. In the thymus, precursor T cells are educated to distinguish self from non-self [7]. HLA molecules play a very important in role this process. Nevertheless, the HLA locus is the most important risk factor for the development of most AID [8]. Understanding the role of HLA molecules in the development of AID will likely offer very important insights in the AID pathogenesis.

RHEUMATOID ARTHRITIS

Rheumatoid arthritis (RA) is a chronic systemic AID characterized by extensive inflammation of the synovial joints. About 1% of the western population is affected by this disease resulting in a large disease burden. In the past decades seminal work was performed on the drivers of the inflammatory process, resulting in the development of a whole range of therapeutics that have greatly enhanced the quality of life of RA patients. Despite the identification of crucial pathways and the implementation of novel therapeutics, little is known about the factors driving the initiation of RA and about why the inflammation does not spontaneously resolve.
Introduction

RA is considered an autoimmune disease. Autoantibodies are an important hallmark of RA and several classes of autoantibodies have been described that precede the development of RA. These autoantibodies include rheumatoid factor, anti-citrullinated protein antibodies (ACPA) and the recently identified anti-carbamylated protein antibodies [10, 11].

ACPA are of particular interest as these autoantibodies are highly specific for RA and can be found in about 50% of early RA patients. This makes ACPA an important early clinical biomarker and ACPA-status was therefore recently added to the new criteria used for the classification of RA patients [12]. In addition, ACPA+ RA patients present with a faster rate of joint destruction. ACPA can therefore also be used as a biomarker for patients with a more severe disease phenotype and could be used to select those patients eligible for a more aggressive treatment [13].

Upon comparing ACPA+ with ACPA- RA patients, it was noted that these patient-groups do not only differ in clinical phenotype but also in genetic and environmental risk factors [14, 15]. It is therefore believed that these are two distinct disease subsets with a different underlying pathogenesis.

Citrullination

In 1964, it was first shown that RA patients harbour very specific antibodies, which were called anti-perinuclear factor or anti-keratin antibodies [16-18]. In a search for synovial targets to which these autoantibodies bind, fibrinogen and vimentin were identified [19, 20]. Proteins can undergo many types of post-translational modifications (PTMs). In 1998, it was demonstrated that the described autoantibodies only target citrullinated antigens, hence their name anti-citrullinated protein antibodies [21-23].

Citrullination is a physiological process catalysed by a family of enzymes called peptidylarginine deiminases (PADI1-4). These enzymes convert the positively charged amino acid (aa) arginine to a novel uncharged aa called citrulline. This process is linked to several aspects of cell biology including apoptosis, necrosis, netosis and histone modifications. PADI enzymes require relatively high concentrations of calcium for their activity allowing citrullination of the extracellular matrix when they are released from (dying) cells [24, 25].

Several data indicate an important role for citrullination in ACPA+ RA. Genome-wide
association studies identified single nucleotide polymorphisms (SNPs) in the region encoding the PADI4 gene that are associated with an increased risk to develop ACPA+ RA[26]. Also, smoking, the most important environmental risk factor, is described to associate with an increased PADI2 expression in bronchoalveolar lavage cells and bronchial mucosal biopsy sections [27]. Finally, it was recently shown that RA patients can carry autoantibodies specifically targeting PADI-enzymes, which can enhance the activity of these enzymes allowing them to function at lower calcium concentrations enabling enhanced citrullination of the extracellular matrix when they are released from (dying) cells [28].

Pathogenicity of ACPA

Because of the clear association between the rate of joint destruction and ACPA positivity, it has been suggested that ACPA may directly contribute to synovial inflammation. ACPA producing B cells are enriched in the synovial fluid, which suggests local production of ACPA in the synovial fluid and a direct role for ACPA and/or ACPA-producing B cells in synovial inflammation [29, 30]. The number of different ACPA isotypes is a predictor of radiographic damage [31]. As different isotypes can recruit different immune-effector mechanisms, this suggests that ACPA can contribute to damage using multiple different effector pathways. For example, stimulation of osteoclast precursors, cells involved in the degradation of bone, with ACPA has been reported to result in increased osteoclastogenesis and was suggested to explain the association between ACPA and more severe joint destruction [32, 33]. This was also confirmed in mice since injecting mice with experimental arthritis with ACPA aggravates arthritis severity [34]. Many antibody effector mechanisms are implicated in the recruitment and activation of inflammatory cells. For instance, ACPA can form immune complexes with citrullinated proteins, which can activate inflammatory cells and induce the production of TNF-alpha [35, 36]. Also, it was reported that ACPA are particularly good in activating the complement system and complement-mediated recruitment of inflammatory cells could therefore be an important effector mechanism of ACPA [37-39]. These data suggest that ACPA is not merely a biomarker for a more severe disease phenotype, but could be directly involved in the inflammatory process.

The antigens targeted by ACPA

Since the identification of the molecular nature of the antigens targeted, much work has been performed on identifying potential citrullinated proteins recognized by ACPA. Unfortunately, this is hampered by the characteristics of the ACPA response. The ACPA repertoire is highly diverse as it displays reactivity against many different citrullinated proteins. This is explained both by the polyclonality of the ACPA response and by several reports that showed that ACPA molecules can directly cross-react between different citrullinated antigens [30, 40-42]. An increasing list of
citrullinated proteins identified within the synovial compartment and recognized by ACPA is now known including proteins like collagen, vimentin, fibrinogen, enolase, fibronectin, vinculin and histones [19, 43-46]. It is therefore likely that a plethora of citrullinated proteins are recognized by ACPA in the synovial compartment that could all conceivably contribute to synovial inflammation. However, due to the high degree of ACPA cross-reactivity, it is difficult to identify those citrullinated antigens that play a critical role in the initiation or development of ACPA+ RA.

Characteristics of the ACPA response

The presence of ACPA can precede RA onset by up to ten years, without any clinical signs of arthritis [47, 48]. Interestingly, by comparing the characteristics of ACPA before and after disease onset, it has become clear that the ACPA response is, on average, not the same between these subjects. In RA patients, ACPA are increased in level, use more isotypes, display a different glycosylation pattern and are more cross-reactive towards different citrullinated epitopes. Interestingly, the “maturation” of the ACPA response takes place before disease onset. Prior to the start of clinical symptoms a sharp rise in ACPA levels, isotype-usage and epitope recognition-profile is observed [49]. These findings are in line with the notion that before disease onset the ACPA-response undergoes a “second hit”, which suggests a role for a matured ACPA response in the onset and the pathogenesis of ACPA+ RA. As an “immature” ACPA response can be identified in serum up to ten years before disease onset, this provides a potentially interesting therapeutic window for interventions preventing maturation of the ACPA response. In a search for such interventions, it is important to understand the progression from ACPA positivity without disease in healthy subjects to ACPA+ RA.

HLA MOLECULES IN RA

Genetic risk

The most important genetic risk factor for RA is the HLA class II locus (Figure 1). Importantly, it was shown that this locus is strongly associated with ACPA+ RA, not with ACPA- RA. It was estimated that about 20% of the genetic variance in ACPA+ RA patients is explained by SNPs in the HLA locus. All the other known loci together account for less than 5% of the genetic variance (unpublished data).

Susceptibility

In 1976, it was first shown that the presence of HLA-DRB1*04 strongly predisposes to RA [50]. The HLA-DRB1 gene is in LD with genes encoding the alpha and the beta chain of HLA-DQ (HLA-DQA1 and HLA-DQB1) and these genes are inherited together
in haplotypes. In 1986, the shared epitope (SE) hypothesis was postulated [51]. This hypothesis assumes that the HLA association is explained by polymorphisms in the HLA-DRB1 chain. Upon comparing different HLA-DRB1*04 alleles it was observed that these alleles differ mostly in aa position 70-74 [52]. Haplotypes encoding predisposing HLA-DRB1 alleles share the sequences QRRAA, QKRAA and RRRAA at these positions and hence are collectively called HLA-SE alleles (Figure 2). It was postulated that this sequence directly influences peptide presentation or T cell recognition [51, 53]. Recently, the association of individual amino acids with ACPA+ RA was revisited by using a novel statistical approach leading to the formation of an “HLA-SE version 2.0” hypothesis. Variations encoding amino-acids 11 and 13 of the HLA-DRB1 chain explain, statistically, most of the genetic risk followed by “SE aa” 71 and 74. Based on these four amino acid positions, RA patients can be subdivided in 16 groups, each with their own genetic risk-profile [54]. The three groups with the highest odds ratios still harbor HLA-DR-DQ haplotypes encoding the SE alleles derived from HLA-DRB1*01, DRB1*04 and HLA-DRB1*10 alleles.

Notably, HLA-DRB1*04 positive RA patients are also HLA-DQ3 positive and it has also been proposed that the predisposing effect of HLA-DR4 can be explained by its LD with genes encoding for HLA-DQ3 [55]. The HLA-DR/HLA-DQ linkage makes it difficult to identify which of the two molecules is responsible for the association with RA. Therefore, it is still unclear whether the HLA-DQ-genes in linkage with the HLA-DR encoding SE genes can be excluded from the association with ACPA+ RA.

**Protection**

In addition to HLA-DRB1 alleles that contribute to RA susceptibility, other HLA-DRB1 alleles confer protection against the disease. HLA-DRB1*13 alleles are most strongly associated with protection, but only from ACPA+ RA. Interestingly, these alleles can also protect in the presence of SE alleles [56]. The dominant role of these protective alleles could suggest that these alleles work on a similar pathway. Likewise, protection can be transferred from a mother carrying a protective allele to a child lacking...
such allele, indicating that protection is an active and dominant process [57].

**HLA involvement in ACPA positivity or in ACPA positive disease?**

As mentioned above, the presence of ACPA can be found up to ten years before disease onset. This suggests that the progression to ACPA+ RA involves two “hits”. Initially individuals are healthy and ACPA negative. After the “first hit” they become positive for ACPA. Then, presumably after a “second hit”, individuals progress from ACPA positivity to ACPA+ RA, which is accompanied by a maturation of the ACPA response. The HLA class II locus is strongly associated with ACPA+ RA and understanding this association would provide important insights in disease pathogenesis. The association suggests an involvement of antigen-specific CD4+ T cells. This is strengthened by genome-wide association studies that identified many genes involved in the regulation of adaptive immune responses (Figure 3) [58]. In understanding the HLA class II association it is important to understand if this locus contributes to either the first or the second “hit” thought to accompany ACPA-positivity or ACPA-positive disease, respectively. As antigen-experienced CD4+ T cells are critically important for the generation of mature antibody responses, it seems likely that activated antigen-experienced (self-reactive) CD4+ T cells are involved in the maturation of the ACPA response. Importantly, recent evidence suggests that the HLA class II locus is not associated with the risk to become ACPA+, but with the risk to progress from ACPA+ to ACPA+ RA [59, 60]. This supports the hypothesis that the HLA class II locus is not directly involved in the formation of ACPA responses, but rather in its maturation, possibly via T cells providing help to ACPA producing B cells before the precipitation of disease (Figure 4).

**THESIS OUTLINE**

The general aim of this thesis was to provide a better understanding of the role of HLA class II dependent antigen presentation in both physiological conditions as well
as in rheumatoid arthritis. To achieve this aim, our studies were separated in two parts.

In **PART I**, we focused on the expression of HLA class II molecules and the characteristics of the ligands presented by HLA class II molecules, starting with the observation that human mast cells can present HLA class II molecules on their surface (**Chapter 2**). Potentially this expression could play a role in rheumatoid arthritis as mast cells are frequent cells in the synovium [61]. The nature of peptides presented by HLA class I molecules have recently been extensively characterized [62]. In **Chapter 3**, we have now performed highly similar analyses in the context of HLA class II with a focus on HLA-DR3. HLA-DR3 is one of the most common HLA-DR molecules and associated with different autoimmune diseases including thyroid autoimmunity and ACPA negative RA [63, 64]. In **Chapter 4**, we have characterized peptides eluted from HLA-DQ6.3. This particular HLA-DQ molecule is in tight linkage disequilibrium with HLA-DRB1*13:01 and is strongly associated with protection from autoimmune diseases including RA and narcolepsy [56, 65].

In **PART II**, we studied the association between HLA class II, RA and citrullination. As discussed, the HLA locus is primarily a risk factor for ACPA+ RA. Several tests are now available to determine ACPA-positivity using model antigens (e.g. anti-CCP-2/3) [11]. In **Chapter 5**, we have examined a novel ACPA multiplex-test in comparison to the conventional tests [66]. In **Chapter 6**, we provide an opinion on recent genetic advances in identifying causal variants within the HLA locus associated with ACPA-positive rheumatoid arthritis. To study the contribution of protective and predisposing alleles in the ACPA+ RA it is important to determine in time where these alleles modulate risk. This has previously been performed for the predisposing alleles and in **Chapter 7** we have expanded this to protective HLA-DRB1*13 alleles. In **Chapter 8**, we provide a detailed molecular basis for the contribution of protective and predisposing HLA class II molecules in the HLA-RA connection. We have also studied an alternative hypothesis that assumes that
the HLA-RA connection is explained by the presentation of shared (citrullinated) antigens by HLA-SE alleles [67]. In Chapter 9, we studied the ability of different SE-positive HLA class II molecules to present citrullinated antigens. In Chapter 10, we expanded these analyses to SE-negative HLA-DR and HLA-DQ molecules. Finally, in Chapter 11 we discuss the implications of the results and future directions.

REFERENCES

18 Chapter 1


36. van Steendam K, et al. Citrullinated vimentin as an important antigen in immune


42. van de Stadt LA, et al. Monoclonal anti-citrullinated protein antibodies selected on citrullinated fibrinogen have distinct targets with different cross-reactivity patterns. Rheumatology (Oxford) 2013;52:631-5.


