Coeliac disease and rheumatoid arthritis: similar mechanisms, different antigens

Frits Koning, Ranjeny Thomas, Jamie Rossjohn and Rene E. Toes

Abstract | Rheumatoid arthritis (RA) and coeliac disease are inflammatory diseases that both have a strong association with class II HLAs: individuals carrying HLA-DQ2.5 and/or HLA-DQ8 alleles have an increased risk of developing coeliac disease, whereas those carrying HLA-DR shared epitope alleles exhibit an increased risk of developing RA. Although the molecular basis of the association with specific HLA molecules in RA remains poorly defined, an immune response against post-translationally modified protein antigens is a hallmark of each disease. In RA, understanding of the pathogenetic role of B-cell responses to citrullinated antigens, including vimentin, fibrinogen and α-enolase, is rapidly growing. Moreover, insight into the role of HLAs in the pathogenesis of coeliac disease has been considerably advanced by the identification of T-cell responses to deamidated gluten antigens presented in conjunction with predisposing HLA-DQ2.5 molecules. This article briefly reviews these advances and draws parallels between the immune mechanisms leading to RA and coeliac disease, which point to a crucial role for T-cell–B-cell cooperation in the development of full-blown disease. Finally, the ways in which these novel insights are being exploited therapeutically to re-establish tolerance in patients with RA and coeliac disease are described.

Koning, F. et al. Nat. Rev. Rheumatol. advance online publication 19 May 2015; doi:10.1038/nrrheum.2015.59

Introduction

Many autoimmune diseases are characterized by a strong association with variants in the HLA region. The precise mechanisms through which these variants contribute to disease are often unknown, although such associations are generally accepted to be best explained by a contribution of antigen presentation to T cells, which is a critical stage in the pathogenesis of autoimmune diseases. The HLA class II region encodes the HLA-DR, HLA-DQ and HLA-DP proteins (as well as other proteins, such as HLA-DM) involved in presentation of peptides to HLA-class II-restricted CD4+ T-helper cells. Therefore, association of the HLA class II region with an autoimmune disease directly implicates the action and participation of CD4+ T cells in the induction and/or perpetuation of the underlying autoimmune response. In several of these diseases, an association with the HLA class II region is linked to the presence of autoantibodies, which require B-cell interactions with T follicular helper cells for their generation. This Review outlines current knowledge of this field and draws parallels between the immune mechanisms leading to rheumatoid arthritis (RA) and coeliac disease, which point to a crucial role for T-cell–B-cell cooperation in the development of full-blown disease. Finally, we describe ways in which these novel insights could be exploited therapeutically to re-establish immune tolerance in patients with RA and coeliac disease.

HLA class II associations

Rheumatoid arthritis

RA was one of the first autoimmune disorders in which an association with the HLA class II region was noted. Analysis of mixed lymphocyte cultures from patients with RA revealed that these individuals shared an HLA determinant, now known as HLA-DR4. Since then, important progress has been made in understanding the basis of this association between the HLA region and RA. Not only are HLA-DR4 molecules associated with RA, but this association is accounted for by HLA haplotypes that encode HLA-DR molecules characterized by a common amino acid sequence in the HLA-DRB1 chain, termed the HLA shared epitope.

In the past 3 years, further refinement of these observations stemmed from finding that specific amino acids at positions 11, 13, 71 and 74 of the HLA-DRB1 chain, as well as position nine of HLA-B and position nine of HLA-DPB1, are associated with the greatest risk of RA. These positions are located within the antigen-binding groove of the HLA molecule, which makes a compelling case for the argument that residues at these positions influence the binding and presentation of specific peptide epitopes, including arthritisogenic epitopes involved in the pathogenesis of RA. Indeed, analyses of natural epitope

Competing interests

R.T. has filed patent applications (PCT/AU2007/001555: Compositions and methods for modulating immune responses, USA; PCT/AU2013/000303: Citrullinated aggrecan peptides for immunotherapy in rheumatoid arthritis, USA) related to technology for targeting dendritic cells to achieve antigen-specific tolerance, and is a director of Dendright, a company developing commercial vaccines that target dendritic cells to suppress autoimmune diseases. The other authors declare no competing interests.
Immune responses to enzymatically modified protein antigens are a hallmark of both these diseases. Environmental insults are probably key to breaking of immune tolerance to enzymatically modified protein antigens.

B-cell to T-cell presentation of target enzymatically modified protein antigens provides a powerful amplification loop sustaining the autoimmune disease process.

This paradigm provides multiple targets for specific interventions aimed at reinstating immune tolerance to enzymatically modified protein antigens.

Key points
- Coeliac disease and rheumatoid arthritis are multifactorial HLA-associated diseases
- Immune responses to enzymatically modified protein antigens are a hallmark of both these diseases
- Environmental insults are probably key to breaking of immune tolerance to enzymatically modified protein antigens
- B-cell to T-cell presentation of target enzymatically modified protein antigens provides a powerful amplification loop sustaining the autoimmune disease process
- This paradigm provides multiple targets for specific interventions aimed at reinstating immune tolerance to enzymatically modified protein antigens

To date, the nature of the association between non-citrullinated autoantigens and HLA shared epitope and non-shared-epitope allomorphs in RA has not been determined. HLA-DR3 has the strongest association with ACPA-negative disease. Conversely, other HLA haplotypes such as HLA-DR13 protect against the development of RA. As these protective effects are again only observed in relation to ACPA-positive disease, the predisposing HLA shared epitope alleles and protective HLA alleles are thought to act in the same biological pathway leading to RA. Uncovering the molecular details underlying these associations and determining how HLA–peptide interactions shape the T-cell repertoire will be an exciting avenue for future research, and has potential therapeutic implications.

Coeliac disease
In contrast to RA, the antigen-specific T-cells and predisposing HLA molecules have been well defined in coeliac disease. Patients with coeliac disease are intolerant of gluten proteins, which are present in commonly consumed cereals including wheat, barley and rye. Roughly 1% of individuals of European, Indian and Middle Eastern ancestry develop coeliac disease, with a female: male ratio of 2:1. Virtually all such individuals carry either or both HLA-DQ2.5 (encoded by DQA1*05:01, DQB1*02:01 alleles) and HLA-DQ8 (encoded by DQA1*03, DQB1*03:02 alleles) molecules. HLA-DQ2.5 confers by far the highest risk of coeliac disease, as approximately 95% of affected patients carry this allele. Also, individuals homozygous for alleles encoding HLA-DQ2.5 molecules have an approximately fivefold increased risk of coeliac disease compared with heterozygous individuals, implying a strong gene-dose effect.

Nonetheless, most individuals who carry high-risk HLA alleles do not develop coeliac disease. Therefore, non-HLA susceptibility genes and environmental factors or precipitating events must converge to trigger the disease. Precipitating events might include incidental gastrointestinal infection(s) or purely stochastic processes that lead to the emergence and expansion of T cells with a disease-causative T-cell receptor (TCR) repertoire. Such expansion could also be associated with the adjuvant effect of invading pathogens, unfavourable changes in the microbiota, or innate events that promote the loss of tolerance in intestinal mucosa.

Generation of antigen-specific T cells

Coeliac disease
Many studies have demonstrated that HLA-DQ2.5-restricted or HLA-DQ8-restricted, gluten-specific T cells are common in patients with coeliac disease. Usually these T cells reside in the lamina propria of the small intestine, but they can migrate into peripheral blood upon gluten challenge. These cells do not secrete IL-17, but they do secrete IL-21 and copious amounts of IFN-γ. This cytokine profile suggests that these T cells have the capacity to induce local inflammation and provide B-cell help. Such gluten-specific T cells are

sequences eluted from shared epitope and non-shared-epitope HLA-DRB1 molecules demonstrated individual binding motifs and preferred anchor residues for each HLA-DRB1 allomorph. As shown in studies of the interactions between antigenic peptides and Lys 71 in the P4 pocket of the antigen-binding groove of HLA-DRB1*0401, epitopes containing negatively charged amino acids are preferred (and those containing arginine are disfavoured).

Another important advance came from the finding that the presence of HLA shared epitope alleles does not predispose individuals to RA as such, but only to autoantibody-positive RA. Furthermore, stratification of patients with RA for the presence of rheumatoid factor or anti-citrullinated-protein antibodies (ACPAs) revealed that the presence of HLA shared epitope alleles conferred a predisposition only to ACPA-positive RA, not to rheumatoid-factor-positive RA. These findings are important, as they indicate that the contribution of HLA shared epitope alleles to the risk of RA reflects the T-helper cell activity involved in formation of ACPA-producing B-cell responses and/or the initiation of T-cell responses to specific citrullinated proteins. ACPAs recognize post-translationally modified proteins in which an arginine residue has been converted to a citrulline residue through the action of protein arginine deiminase (PADI) enzymes. Importantly, genetic variants in loci encoding PADI enzymes also predispose to the development of RA. These genetic associations underscore the importance of protein citrullination by PADI enzymes in the development of RA.

In view of the high specificity of ACPAs for RA, and the observation that the predisposition to RA conferred by HLA shared epitope molecules is only apparent with ACPA-positive disease, an appealing hypothesis is that the susceptibility variants of these HLA molecules facilitate the activation of T cells involved in providing help to ACPA-producing B cells. However, which specific peptides are recognized by these T cells and how T-cells directed against citrullinated peptides contribute to the pathogenesis of RA remain poorly understood. For example, the commonly described epitopes containing citrullinated residues lie within B-cell epitopes, which might not be the dominant T-cell autoantigenic epitopes. Moreover, RA is heterogeneous, and the evidence suggests that multiple autoantigens are involved in its pathogenesis.
Together with the observed HLA-DQ2.5 gene-dose effect, this evidence led to the concept that exposure to immunogenic complexes of HLA-DQ2.5 and gluten peptides above a certain threshold level drives the development of coeliac disease.20,21 In high-risk individuals expressing HLA-DQ2.5 molecules, the maintenance of oral tolerance (the state of local and systemic immune unresponsiveness to antigens normally present in food) towards potentially immunogenic gluten proteins is hypothesized to be compromised by T-cell exposure to high numbers of immunogenic HLA-DQ2.5–gluten complexes in the intestinal mucosa or draining lymph nodes. Increased T-cell exposure to these complexes might occur through several mechanisms: increased expression of TGase-2 in response to a triggering event; abnormalities in the selection, proliferation or survival of regulatory T (TREG) cells; impaired induction of peripheral TREG cells in the gut; or an enhanced capacity of gut dendritic cells to stimulate effector T-cell responses to immunogenic sequences in an inflammatory milieu—for example, enhanced gluten presentation through IFN-γ-driven upregulation of HLA-DQ2.5 expression (Figure 1).

Gluten contains many potentially immunogenic amino acid sequences, and the T-cell response in children can be directed to epitopes that are rarely implicated in the immune responses observed in adults with coeliac disease.38 By contrast, in adults, T-cell responses to particular gluten peptides (especially the α-gliadins39,40) are almost invariably found, and the immune response therefore seems to be skewed towards these immunodominant epitopes. Furthermore, highly similar sequences are found in the ω-gliadins and gluten-like proteins in barley and rye.26,41 To reconcile these different observations in children and adults, our research group has previously proposed that gluten-specific immune responses might be initiated by any of a wide range of immunogenic sequences, but as coeliac disease progresses, the T cells with the highest affinity (namely those directed against the immunodominant peptides) outcompete those directed against other gluten epitopes. This process is termed epitope focusing.38

T cells from patients with coeliac disease that are specific for immunodominant gluten peptides express a strongly biased TCR repertoire dominated by a few TRAV (TCR α-chain V) and TRBV (TCR β-chain V) molecules.42–46 Also, a non-germline-encoded arginine residue is frequently found in complementarity-determining region (CDR) 3 of the majority of TCRs that are specific for immunodominant gluten peptides, and T-cell clones from unrelated patients with coeliac disease express very similar CDR3 regions.42–46 Direct measurements demonstrate a high level of affinity of the biased TCRs for their cognate ligand, similar to that typically observed with TCRs specific for microbial antigens.43,44,47,48 This high affinity contrasts with the moderate affinity of other autoimmune interactions between TCRs, self peptides and HLA molecules;49 this difference might relate to the non-self nature of gluten.

---

**Table 1 | Modification of gluten peptides by tissue transglutaminase**

<table>
<thead>
<tr>
<th>HLA-modified gluten complex</th>
<th>Original peptide sequence</th>
<th>HLA-DQ2-binding modified sequence (XXXEEXXX)*</th>
<th>HLA-DQ8-binding modified sequence (EXXXXXXX)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQ2.5–a1-gliadin</td>
<td>PPPOQPQPY</td>
<td>PPPOPELPY</td>
<td>–</td>
</tr>
<tr>
<td>DQ2.5–a2-gliadin</td>
<td>PPQPQVPYQ</td>
<td>PPPELPQPYQ</td>
<td>–</td>
</tr>
<tr>
<td>DQ2.5–ω1-gliadin</td>
<td>PPPOPQOPF</td>
<td>PPPOQEPQF</td>
<td>–</td>
</tr>
<tr>
<td>DQ2.5–ω1-gliadin</td>
<td>PPQPQOPPF</td>
<td>PPQEOQPPW</td>
<td>–</td>
</tr>
<tr>
<td>DQ2.5–a3-gliadin</td>
<td>FRPOQYPQY</td>
<td>FRPEQYPQY</td>
<td>–</td>
</tr>
<tr>
<td>DQ2.5–γ1-gliadin</td>
<td>PQSQPSQQQ</td>
<td>PQPSPEQQQ</td>
<td>PQPSPEQE</td>
</tr>
<tr>
<td>DQ8–a1-gliadin</td>
<td>QGSFQPSQQQ</td>
<td>–</td>
<td>EGFSQPSQE</td>
</tr>
<tr>
<td>DQ8–γ1a-gliadin</td>
<td>OQPQOPPPQ</td>
<td>–</td>
<td>EQPOQPPF</td>
</tr>
</tbody>
</table>

---

*HLA-DQ2.5 requires E at positions 4 and either 6 or 7 (or both), whereas HLA-DQ8 requires E at positions 1 and 9, or both. Conversion of neutral Q into negatively charged E enables high-affinity binding to HLA-DQ2.5, HLA-DQ8, or both. Abbreviations: E, glutamic acid; F, phenylalanine; G, glycine; L, leucine; P, proline; Q, glutamine; R, arginine; S, serine; W, tryptophan; Y, tyrosine.

---

not detectable in small intestine biopsy samples from HLA-DQ-matched healthy controls, and consequently expansion of this T-cell population seems to be unique to people who develop coeliac disease.20,21 The immunogenicity of gluten is linked to its unique glutamine-rich and proline-rich composition, which renders it highly resistant to pepsin-mediated degradation during digestion. Relatively long gluten fragments consequently reach the duodenal compartment, are transported across the epithelium, and can be taken up by dendritic cells in the lamina propria.30 These dendritic cells then migrate to the draining lymph nodes where priming of gluten-specific T-cells occurs.30 Gluten protein fragments frequently contain multiple immunogenic repetitive amino acid sequences that can induce vigorous T-cell responses.39 Such gluten fragments are an ideal substrate for protein-glutamine γ-glutamyltransferase 2 (also known as tissue transglutaminase [TGase-2]), an enzyme that selectively converts glutamine residues into glutamic acid.31 This change induces negative charges in the gluten fragments, facilitating their high-affinity binding to HLA-DQ2.5 and HLA-DQ8 molecules (Table 1).32,33 This precise but ultimately pathogenic match between post-translationally modified gluten fragments and disease-predisposing HLA-DQ2.5 and HLA-DQ8 molecules provides an explanation for the association of these genetic variants with coeliac disease.20,21,34–36 Strikingly, HLA-DQ2.2 molecules (encoded by DQA1*02:01, DQB1*02:01 alleles) are structurally very similar to HLA-DQ2.5, yet are not associated with coeliac disease. Several studies have indicated that this distinction is due to a subtle difference in the peptide-binding properties of HLA-DQ2.2 versus HLA-DQ2.5 that enables HLA-DQ2.5 molecules to bind to a substantially larger repertoire of immunogenic gluten peptides32 and to form more stable complexes37 than HLA-DQ2.2 does. Similarly, HLA-DQ8 molecules can only bind to a limited subset of gluten peptides, explaining why individuals with these molecules have a low risk of coeliac disease development versus those with HLA-DQ2.5.34

---

© 2015 Macmillan Publishers Limited. All rights reserved
Figure 1 | Binding of modified peptides to MHC class II molecules and their interactions with T cells in coeliac disease and RA. a | Deamidation of gluten peptides introduces a negatively charged E at p4, resulting in improved binding to HLA-DQ2.5, lowering of the threshold for T-cell activation, and increased T-cell proliferation and cytokine secretion. b | Interactions between the HLA-DQ2.5 α-chain (yellow), HLA-DQ2.5 β-chain (grey) and deamidated α2-gliadin epitope (green) are shown. Specific residues of the HLA-DQ2.5 β-chain interact with the deamidated E residue of α2-gliadin (blue) via hydrogen bonds (black dashes) and Van der Waals interactions (brown dots). c | Similarly, in RA, conversion of positively charged R into neutral Cit at p4 facilitates binding to a K residue in HLA-DRB1*04:01, which could influence thymic selection and result in autoreactivity. d | Interactions between the HLA-DRB1*04:01 α-chain (pale yellow), HLA-DRB1*04:01 β-chain (pale blue) and citrullinated vimentin (vimentin-71Cit67,71) epitope (purple) are shown. Specific residues of the HLA-DRB1*04:01 β-chain interact with Cit (bright blue) via hydrogen bonds (black dashes) and Van der Waals interactions (brown dots). Abbreviations: ↑, increased; →, strongly increased; ↓, decreased; Cit, citrulline; p, amino acid position; RA, rheumatoid arthritis; SNR, single-nucleotide polymorphism.

Whether this biased TCR repertoire is present in unaffected individuals who carry HLA-DQ2.5 alleles is not yet known. Answering this question will aid the understanding of how and when the biased TCR repertoire develops in relation to the putative triggering event in coeliac disease. Of potential benefit in both diagnosis and translational research, antigen-specific T cells can be readily visualized with HLA-DQ2.5–gluten tetramer reagents.44,50,51 These reagents facilitate the isolation and characterization of antigen-specific T cells (reviewed in detail elsewhere53).

Typically, up to 2% of the CD4+ T-cell population in the small intestinal lamina propria of patients with coeliac disease recognizes HLA-DQ2.5 complexed with immunodominant gluten peptides, when patients are on a gluten-containing diet.51 Structural studies have provided the first insight into the molecular basis for the TCR–HLA-DQ2.5–antigen and TCR–HLA-DQ8–antigen interactions central to coeliac disease. The structure of a biased TRAV26–TRBV9 TCR in complex with HLA-DQ8–α1-gliadin revealed that two germline-encoded residues in TRBV9 make crucial contacts with both the α1-gliadin peptide and HLA-DQ8, providing an explanation for the TRBV9 bias.44 In addition, a non-germline-encoded arginine in the α-chain of CDR3 (CDR3a) makes several contacts with both the α1-gliadin peptide and HLA-DQ8. Also, several ternary structures of the TRAV26-1–TRBV7.2 TCR and HLA-DQ2.5–α2-gliadin complexes have been elucidated (Figure 2). These studies revealed that germline-encoded residues within TRAV26-1 have a crucial role in the interaction with the HLA-DQ2.5–α2-gliadin complex, whereas non-germline-encoded amino acids (including an arginine) in the β-chain of CDR3 (CDR3β) are crucial in mediating contact between the TCR and HLA-DQ2.5–α2-gliadin.53 Thus, various combinations of non-germline-encoded and germline-encoded interactions mediate the biased TCR recognition of immunodominant gluten peptides—in which non-germline-encoded arginines in the CDR3 loop act as a lynchpin.

Rheumatoid arthritis

Analysis of the ACPA response in patients with RA provides compelling evidence for a role of T cells in shaping the B-cell response to citrullinated proteins. For example, ACPAs consist of all IgG subclasses, IgM and IgA, indicative of a T-cell-dependent B-cell response.52 Likewise, the variable region of these antibodies has undergone extensive somatic hypermutation, also pointing to a T-cell-driven B-cell response.54 Epidemiological evidence similarly indicates an important input of T cells to ACPA-producing B cells, as the RA-predisposing HLA shared epitope alleles influence ACPA levels as well as their fine specificity.55,56 These data recapitulate classic studies in mice, which showed that the MHC region controls the magnitude and specificity of antibody production.57

However, identification of the citrullinated epitopes recognized by autoreactive T cells in patients with RA has proven difficult. Evidence is emerging for T-cell recognition of several citrullinated epitopes, based on analyses using peptide–HLA tetramers and in vitro T-cell responses to candidate epitopes.27,58–60 Citrulline, in contrast to arginine, can be accommodated within the electrostatically charged pocket of two HLA shared epitope molecules: HLA-DRB1*04:01 and HLA-DRB1*04:04, which supports the argument that the change in characteristics afforded by conversion of arginine to citrulline (namely changing from positively charged to neutral) enables citrullinated autoantigenic peptide epitopes to bind where the native sequences cannot.7

Of note, a number of other self-peptide epitopes implicated in the pathogenesis of autoimmune diseases—such as insulin B chain (InsB)9,23 and myelin basic
protein acetylated N-terminal (MBP Ac)₁₋₉—also bind only weakly to MHC class II molecules, and such weak binding is proposed to be a mechanism by which potentially self-reactive T cells might escape thymic negative selection and enter the repertoire. In the same way, we speculate that the relatively permissive binding of the HLA shared epitope to citrullinated epitopes (that is, compared with its binding to non-citrullinated epitopes) might be most relevant to selection of a potentially self-reactive repertoire in the thymus. Whether or not autoimmunity subsequently develops depends on genetic and environmental factors that contribute to efficient TCR signalling in response to self antigens, such as a deficiency in Tₘₑ₂₅ cells, the presence of PTPN22 risk alleles, and production of inflammatory cytokines that promote the survival and expansion of memory T cells (Figure 1).

The molecular identity of the T-cell epitopes and TCRs that contribute to RA are not as well-defined as those involved in coeliac disease. In part, this lack of clarity exists because the pathogenetic mechanisms underlying HLA associations in RA are more complex than those in coeliac disease. For example, the HLA associations differ in ACPA-positive and ACPA-negative RA. Moreover, the HLA association with ACPA-positive RA is not with a single MHC class II molecule, but rather with a group of molecules containing a similar (but not identical) antigen-binding footprint, which the shared epitope concept assumes to be equivalent. In fact, a hierarchy of risk is associated with different HLA genotypes linked to RA. In coeliac disease, a small set of antigenic peptides seems to be presented by a very limited number of HLA-DQ allomorphs, leading to oligoclonal expansion of high-affinity, antigen-specific T cells. Although next-generation sequencing has provided evidence of an oligoclonal expansion of T-cell populations in HLA-DR4-positive patients with early RA, especially in the synovial compartment, the TCR antigen specificity of these cells is unknown. No public T-cell responses (in which T cells bearing identical TCRs dominate the response to the same antigenic epitope in multiple individuals) were identified among 12 patients with RA.

ACPAs are crossreactive with many different citrullinated proteins—in contrast to the antibodies associated with coeliac disease, which are highly specific for two well-defined antigens, namely gluten and TGase-2. ACPA crossreactivity is observed not only at the polyclonal level, but also at the monoclonal level. Therefore, although identifying the citrullinated proteins recognized by ACPAs might reveal the citrullinated T-cell epitopes involved in B cell help, isotype switching and somatic hypermutation, the same antibodies might also recognize other structurally (un)related proteins. An ACPA-producing B cell, because of the crossreactive nature of ACPAs, might conceivably recruit help from several different T cells specific for different citrullinated antigens, including a variety of autoantigens and foreign proteins. Binding of these various proteins to the immunoglobulin receptors of ACPA-producing B cells results in different T-cell epitopes being presented to T cells in the context of MHC class II molecules after antigen processing (Figures 3 and 4). In this context, the T cells that provide help to ACPA-producing B cells might not necessarily recognize citrullinated epitopes. Indeed, considerable evidence exists that non-citrullinated peptide epitopes (including collagen, aggrecan, glycosylated immunoglobulin, cartilage glycoprotein 39, and several heterogeneous nuclear ribonucleoproteins and immunoglobulin binding proteins) are also presented to T cells in RA.

Although the quest for the HLA-DR-restricted autoantigenic epitopes involved in RA is expected to continue, epidemiological studies have made considerable progress by identifying the stage at which HLA shared epitope alleles probably contribute to the pathogenesis of RA. ACPAs can be detected years before disease onset, and their presence is associated with a history of smoking. Shortly before clinical onset of disease, ACPAs show pronounced increases in titre, isotype range and antigen-recognition profiles. These observations strongly suggest that the ACPA-producing B cells acquire a ‘second hit’ immediately before disease onset, possibly representing the ability to recruit the T-helper activity required for maturation of the ACPA response. Indeed, in line with this idea, the results of a twin study showed that environmental (including smoking) or stochastic factors predominately contribute to ACPA positivity in healthy individuals, with a small contribution from RA-associated HLA shared epitope genotypes. Likewise, a study of almost 10,000 healthy individuals showed that ACPA positivity is not confined to patients with RA, as 1–2% of
these healthy people were ACPA-positive.81 Importantly, however, HLA shared epitope status did not correlate with ACPA positivity in this healthy population.81 By contrast, presence of the shared epitope is strongly associated with ACPA-positive RA.8

Furthermore, ACPA titres are lower (and the number of epitopes recognized fewer) in the absence than in the presence of RA.12 Together, these observations suggest that the B-cell response against citrullinated proteins might smoulder for years before the clinical symptoms of RA develop. Although smoking has a role in the first appearance of ACPA, in many patients T cells might not contribute to the initial breaking of B-cell tolerance to citrullinated self proteins. However, once the B cells acquire T-cell help, the ACPA response dramatically increases in range and intensity. Since clinical symptoms of RA begin shortly after this change, it is tempting to speculate that maturation of the autoimmune response might be ignited by an inflammatory ‘second hit’ that leads directly to the onset of arthritis, especially in the context of an ageing immune system, in which the T-cell repertoire is contracting (Figure 3).

Precisely what event(s) drive the onset of RA is presently unclear, but delineating the HLA-associated pathways involved in maturation and evolution of the autoimmune response to citrullinated proteins is likely to prove a very rewarding field of research. Such studies could provide pivotal information required to halt this maturation, and thereby perhaps prevent the appearance of RA.

**Involvement of B cells**

**Coeliac disease**

Two types of antibodies are typically found in patients with coeliac disease but not healthy individuals: IgA autoantibodies specific for TGase-2 and antibodies specific for deamidated gliadin, a post-translationally modified, gluten-derived peptide.82–84 Assays for these antibodies are now commonly used in the diagnosis of coeliac disease, as they have high sensitivity and specificity. A surprisingly large percentage of intestinal plasma cells secrete antibodies specific for TGase-2 (on average 10%) and deamidated gliadin (on average 3%) in patients with active coeliac disease.85,86 The proportion of such cells correlates with exposure to dietary gluten, as TGase-2-specific and deamidated-gliadin-specific plasma cells decrease with a gluten-free diet.

Strikingly, and similarly to the TCR repertoire, the B-cell receptor (BCR) repertoires responsive to both TGase-2 and deamidated gliadin are biased, in that both are dominated by a few combinations of VH and VL chains (variable heavy and variable light regions of immunoglobulin molecules, respectively).85,86 Also similarly to the TCR repertoire, the IgA antibodies specific for TGase-2 and deamidated gliadin bind with high affinity, yet demonstrate little somatic hypermutation, indicating the presence of a relatively high affinity, germline-encoded immunoglobulin repertoire specific for TGase-2 and deamidated gliadin. The similarity of this lack of somatic hypermutation despite chronic exposure to both TGase-2 and deamidated gliadin suggests a common underlying mechanism, although its molecular basis remains unknown.

Most likely, these TGase-2-specific and deamidated-gliadin-specific IgA antibodies do not have a direct effector role in the pathogenesis of coeliac disease, as a gluten-free diet is highly effective in relieving symptoms despite the continued presence of both types of antibodies. However, IgA antibodies have a potential role in transport of gluten fragments across the epithelium, via a transferrin-receptor-mediated process.87 Moreover, these antibodies might amplify the gluten-specific T-cell response through B-cell to T-cell presentation, as discussed below. Finally, these antibodies could have a role in extragastrointestinal disorders such as dermatitis herpetiformis, a gluten-dependent disorder in which deposition of IgA antibodies specific for protein glutamine γ-glutamyltransferase E (also known as epidermal transglutaminase) in the skin has a crucial pathogenetic role.88

---

**Figure 3** | A two-hit model for the development of ACPA-positive RA. ACPAs are present long before onset of clinical symptoms of RA. Environmental factors and stochastic events probably drive the initial development of ACPAs, without a substantial contribution of HLA molecules. However, HLA risk alleles do contribute to the development of ACPA-positive disease, and so predisposing HLA molecules are thought to be involved in the ‘second hit’ that enables expansion of the ACPA response, which occurs shortly before disease precipitation. The CD4+ T-cell help required for isotype switching and maturation of the ACPA response might be generated by presentation of citrullinated (foreign or self) epitopes, or non-citrullinated epitopes derived from citrullinated foreign proteins, following infection. Abbreviations: ACPA, anti-citrullinated protein antibody; BCR, B-cell receptor; DC, dendritic cell; FcR, Fc receptor; RA, rheumatoid arthritis; TCR, T-cell receptor.
Rheumatoid arthritis

B cells are very likely to be involved in active RA, as B-cell depletion mediated by the CD20-specific monoclonal antibody rituximab is an effective treatment for this disease. However, which B cells contribute and how they contribute to the pathogenesis of RA is unclear. RA is characterized by the presence of several different autoantibodies, such as rheumatoid factor, ACPAs and anti-carbamylated-protein (anti-CarP) antibodies. Anti-CarP autoantibodies recognize proteins containing homocitrulline, and can also be found in patients with ACPA-negative RA.

All these antibodies (but especially ACPAs) are postulated to contribute directly to the pathogenesis of RA as they are highly specific for this disease and their presence is associated with its severity and persistence. Moreover, in vitro studies show that RA-associated autoantibodies recruit various immune effector mechanisms and can induce osteoclastogenesis and bone resorption in mice. Transfer of ACPAs to mice with
mild collagen-induced arthritis (CIA) leads to disease exacerbation, and inducing tolerance of citrullinated peptides in mice with CIA reduces the severity of this disease. Other studies suggest direct pathogenicity of ACPAs in RA. However, use of effective therapies in patients with RA does not necessarily induce reductions in levels or composition of the ACPAs present in serum. Therefore, whether the role of B cells in RA is derived from their ability to produce autoantibodies, and whether these autoantibodies are a main driver of the pathogenesis of RA, remains unclear. Likewise, ACPAs might not be directly pathogenic to joints in the absence of other immune abnormalities, as babies born to women with ACPA-positive RA are not affected by RA or RA-like symptoms. By contrast, babies born to women with systemic lupus erythematosus or Sjögren syndrome can (albeit in rare circumstances) develop syndromes such as neonatal lupus or congenital heart block that are mediated by maternal autoantibodies transferred across the placenta.

Alternatively, microbial proteins can be citrullinated by microbial PADI, and ACPAs might, therefore, have a functional role as neutralizing antibodies in antimicrobial immunity. Thus, the presence of RA-associated autoantibodies might conceivably reflect both autoimmune and antimicrobial responses, and these autoantibodies might be neither direct participants in the pathogenesis of RA nor capable of its perpetuation. In this scenario, the autoantibody-producing B cells, rather than the antibodies they produce, would be most likely to be responsible for disease perpetuation. B cells secrete substantial amounts of cytokines, including TNF and IL-6, upon BCR-mediated activation. Importantly, increased numbers of newly activated ACPA-producing B cells (plasmablasts) are present in the synovial compartment of patients with RA, and the numbers of these cells in blood correlates with ACPA levels. Besides their ability to produce autoantibodies, cytokines and other proinflammatory molecules, B cells are likely to be important synovial antigen-presenting cells, given their large numbers relative to dendritic cells.

Antibodies to TGase-2 are typically found in patients with coeliac disease, and anti-PADI autoantibodies can, similarly, be detected in patients with RA. Strikingly, anti-PADI antibodies from patients with RA increase the activity of their target enzymes, which amplifies the generation of citrullinated autoantigens. By contrast, the anti-TGase-2 antibodies in patients with coeliac disease do not influence enzyme activity, indicating differing roles for autoantibodies in the pathogenesis of coeliac disease and RA.

**B-cell to T-cell antigen presentation**

The presence of an amplification loop involving B-cell to T-cell antigen presentation is strongly implicated in both coeliac disease and RA (Figure 4). The unique presence of antibodies targeting TGase-2 and deamidated gliadin in patients with coeliac disease, as well as the sharp decline in these antibody titres after the introduction of a gluten-free diet, suggests that secretion of these antibodies is driven by gluten-specific T cells. In the case of antibodies specific for deamidated gliadin, we can assume that B cells expressing immunoglobulins specific for deamidated gliadin would internalize and process these peptides for binding to HLA-DQ2.5 or HLA-DQ8. The subsequent activation of gluten-specific T cells would result in effective T-cell to B-cell help, leading to immunoglobulin class switching and further antibody secretion by B cells (Figure 4a). Conversely, antibody-mediated antigen uptake by B cells results in highly efficient antigen presentation to T cells, amplifying the gluten-specific T-cell response. Gluten-specific T cells are also likely to initiate and maintain the secretion of TGase-2-specific antibodies. TGase-2 can crosslink itself to gluten, and uptake of the resulting TGase-2–gluten complexes by TGase-2–specific B cells would likewise result in efficient presentation of gluten peptides to gluten-specific T cells by TGase-2–specific B cells, leading to further antibody secretion and a second amplification loop—a ‘hapten-like’ effect (Figure 4a).

Although direct evidence for the existence of such antibody-mediated amplification loops has not yet been provided, a few case reports suggest that B-cell-depleting therapies might be effective in coeliac disease. These observations provide some evidence that a B-cell to T-cell amplification loop is present. Likewise, whether cooperation between B-cells and T-cells contributes to disease progression in RA is unclear, as antigen–specific T-cell responses have not been evaluated after B-cell-depleting therapies in patients with RA. However, rituximab treatment is associated with a reduction in numbers of immune cells in the RA synovium, including a decrease in T-cell numbers, and these observations suggest a crucial role of the B-cell compartment in RA, similar to that in coeliac disease (Figure 4).

**Therapeutic implications**

The available evidence surrounding the pathogenetic roles of HLA class II molecules, peptide autoantigens, antigen–specific T cells, antigen–specific B cells and their interactions in coeliac disease and RA opens up exciting opportunities for the development of antigen–specific therapy in these diseases. Clearly, if disease is perpetuated through the presentation of antigens to T cells whose effector function includes help for B-cell autoantibody production, then interruption or regulation of this process in an antigen–specific manner has the potential to block disease perpetuation without interfering with antimicrobial and antitumour immune responses. Indeed, the signs and symptoms of coeliac disease can be inhibited by a gluten–free diet, indicating that avoidance of antigen exposure effectively prevents the symptoms. By analogy, inhibition of PADI activity might offer a similar opportunity to prevent the citrullination of antigens, including self antigens; a proof–of–concept study has been conducted in mice.

Intriguingly, however, the available evidence in coeliac disease suggests that a broad range of highly specific intervention strategies is possible. The presence of a highly biased TCR repertoire in all patients with coeliac...
Block antigen processing
Block post-translational protein modification
Block B-cell to T-cell antigen presentation
Reinstate tolerance using peptide-based therapy

Box 1 | Potential intervention strategies

- Block antigen processing
- Block antigen presentation
- Block post-translational protein modification
- Block B-cell to T-cell antigen presentation
- Eliminate pathogenic antigen-specific T cells
- Induce T regulatory cell responses
- Reinstate tolerance by targeting dendritic cells
- Reinstate tolerance using peptide-based therapy

disease examined to date points to a crucial involvement of antigen-specific T cells in its pathogenesis. Elimination or suppression of such cells could, therefore, be an effective immunotherapy for coeliac disease. The available structural data indicate that these highly biased TCRs engage HLA-DQ2.5–gliadin complexes in a highly conserved fashion, in which a relatively small number of germline-encoded and non-germline-encoded residues in the TCR make critical interactions with a HLA-DQ2.5–gliadin complex (Figure 2). This information can be used to build a structure-based pharmacophore model for virtual screening of compound libraries to identify lead molecules that will specifically bind to these biased TCRs, an approach that has already been applied to identify immunomodulatory small molecules capable of influencing presentation of InsB to T cells. After optimization, such compounds could be coupled to fluorochromes (for ex vivo detection of the specific T cells that express these biased TCRs), or to toxins (for in vivo elimination of such T cells). Currently, the knowledge base in RA lags behind that in coeliac disease, in that no HLA–self-peptide–TCR structures have been solved, and our understanding of the autoantigen-specific T-cell repertoire is rudimentary. Furthermore, given the diversity of autoantigens in RA, the choice of antigen for immunotherapy is less obvious than in coeliac disease.

A number of strategies have been developed to induce antigen-specific tolerance in patients affected by autoimmune diseases and allergies. Many of these strategies are in various stages of translation to the clinic, including administration of low or escalating doses of relevant antigenic peptides, exposure of tolerogenic dendritic cells to relevant peptides, and other strategies targeting dendritic cells (reviewed elsewhere). However, understanding which autoantigens are involved and characterizing their responding T cells is insufficient to cure any disease, just yet. The current challenge is to design clinical trials that include appropriate choices of peptide antigen(s), dosing and delivery strategy, as well as outcomes reflecting immune response parameters that are (at least initially) surrogates of treatment efficacy. Although preliminary clinical data are starting to emerge into the public domain from tolerance-induction trials—especially evidence of the safety of this approach—progress will be erratic without well-designed accompanying immune biomarker studies that can pick apart the effects of immunotherapies on the numbers and phenotypes of antigen-specific T cells (including Treg cells), antigen-specific B cells and their effector functions. The need for appropriate, well-conducted immune biomarker studies to be included by design in clinical trials of antigen-specific immunotherapies for rheumatic diseases is, therefore, greater than ever before.

In this regard, peptide–HLA multimer technologies have great potential utility for the enumeration and phenotyping of antigen-specific CD4+ T cells, and for extensive single-cell analysis using available genomic tools. Nonetheless, the high-affinity binding of peptide–HLA tetramers and high frequency of antigen-specific cells in the peripheral blood of patients with coeliac disease seems to be exceptional. Identification of rare antigen-specific cells with low-affinity binding, as occur in RA, is much more challenging. In addition, mass cytometry is emerging as a novel approach for multi-parameter analysis of biological samples. Currently, this technique can analyse up to 36 markers simultaneously, a number that is expected to rise to 100 in the near future. Mass cytometry is, therefore, a considerable advance over conventional flow cytometry, which is limited by fluorochrome availability, spectral overlap and compensation, biological background and auto-fluorescence. Moreover, because mass cytometry data are analysed by groupings of markers rather than individual markers, they can reveal relationships between cell subtypes that have previously not been recognized. We expect that use of this technique will lead to substantial advances in the discovery of disease-specific biomarkers, and that ongoing methodological development will see such technologies become more widespread.

Despite all these challenges, the exciting data emerging from the convergence of structural, biochemical and cellular immunology described herein are beginning to lead towards antigen-specific therapy in diseases such as coeliac disease and RA.

Conclusions

The strong HLA associations observed in coeliac disease and RA point to the involvement of detrimental immune responses in both diseases. In coeliac disease, T cells specific for post-translationally modified gluten peptides bound to predisposing HLA-DQ molecules are typically present in patients but not healthy individuals. These T cells drive the production of (auto)antibodies specific for modified gluten and TGase-2, leading to B-cell to T-cell presentation of immunogenic gluten epitopes, and forming an immune-response amplification loop. Similarly, antibodies to post-translationally modified antigens are found in RA, and the nature of the HLA association likewise suggests the presence of (auto)antigen-specific T cell responses and B-cell to T-cell presentation. However, the exact nature of the T-cell response in RA is still rather poorly understood.

In a broad sense, several therapeutic approaches to these diseases can be envisioned (Box 1), and the well-defined role of the immune system in coeliac disease provides several possibilities for therapeutic interventions based on these strategies. The results of such approaches could pave the way for similar attempts to halt or cure RA once sufficiently detailed knowledge on the nature of the underlying immune response has been gained.
REVIEWS


Acknowledgements
The authors’ research work is supported by the Australian Research Council, the National Health and Medical Research Council (NHMRC) of Australia and the Dutch Arthritis foundation. R.T is supported by an NHMRC Research fellowship and Arthritis Queensland, J.R. by an NHMRC Australia Fellowship, and R.E.T by a Vici fellowship from Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). The authors thank Jan Petersen for help with the drafting of the original figures.

Author contributions
The authors contributed equally to all aspects of the manuscript: researching data for the article, discussions of its content, writing and review or editing of the manuscript before submission.