The role of MHC class II genes and post-translational modifications in celiac disease

Ludvig M. Sollid

Centre for Immune Regulation, KG Jebsen Coeliac Disease Research Centre, Department of Immunology, University of Oslo and Oslo University Hospital-Rikshospitalet, Oslo, Norway.

Correspondence: Ludvig M. Sollid E-mail: l.m.sollid@medisin.uio.no, Phone: +47 230273811

ORCID: 0000-0001-8860-704X

Acknowledgments: The work in the author's laboratory was supported by grants from the Research Council of Norway through its Centers of Excellence funding scheme (project number 179573/V40); by grants from the South-Eastern Norway Regional Health Authority; by the European Commission (grant ERC-2010-Ad-268541); and by Stiftelsen KG Jebsen. The author thanks former and present members of his laboratory as well as collaborators for invaluable contributions to the research. He also thanks numerous celiac disease patients for their donation of biological samples, thus making critical experiments possible.

Abstract

Our increasing understanding of the etiology of celiac disease, previously considered a simple food hypersensitivity disorder caused by an immune response to cereal gluten proteins, challenges established concepts of autoimmunity. HLA is a chief genetic determinant, and certain HLA-DQ allotypes predispose to the disease by presenting posttranslationally modified (deamidated) gluten peptides to $CD4^+$ T cells. The deamidation of gluten peptides is mediated by transglutaminase 2. Strikingly, celiac disease patients generate highly disease-specific autoantibodies to the transglutaminase 2 enzyme. The dual role of transglutaminase 2 in celiac disease is hardly coincidental. This paper reviews the genetic mapping and involvement of MHC class II genes in disease pathogenesis, and discusses the evidence that MHC class II genes, via the involvement of transglutaminase 2, influence the generation of celiac disease-specific autoantibodies.

Celiac disease is a gluten-sensitive enteropathy that is caused by a harmful immune response to cereal gluten proteins. While the condition was recognized already in ancient Greece, it was only after the Second World War that the role of gluten became clear. The Dutch pediatrician Willem K. Dicke established by elimination and provocation tests the culprit role of gluten proteins in the disorder (Dicke 1950). Gluten is the sticky protein part of wheat flour that remains after washing with water, and wheat gluten proteins can be divided into gliadins (α-, γ- and ω-types) and glutenins. Related proteins are found also in grains of barley (hordeins) and rye (secalins). Strictly speaking, gluten is a term referring to wheat proteins, yet this term is now also commonly used for proteins from barley and rye. Celiac disease can be treated effectively by a life-long, strict elimination of gluten from the diet. In accordance with this condition initially being classified as a food hypersensitivity disorder, elimination and provocation tests were central in the early diagnostic schemes (Meeuwisse 1970). Examination of proximal gut biopsies showing characteristic blunting of villi, crypt cell hyperplasia and infiltration of leukocytes both in the epithelium and the lamina propria was also an essential part of the diagnostic workup (Walker-Smith et al. 1990). However, a few years ago the European Society for Paediatric Gastroenterology and Hepatology (ESPGHAN) made the game-changing decision that the diagnosis could be made in some children without examination of gut biopsies (Husby et al. 2012). The reason for this development relates to the fact that celiac disease characteristic antibodies specific for the autoantigen transglutaminase 2 (TG2) and deamidated gluten peptides have extraordinarily high diagnostic sensitivities and specificities. The need for the biopsy examination is alleviated when such antibodies are present in serum at high levels. In fact, the diagnostic performance of the IgA anti-TG2 is greater than for any other human autoantibodies. This change in diagnostic procedures tellingly illustrates how celiac disease over the last fifty years has moved from the camp of food sensitivity disorders to the camp of autoimmune diseases. In addition to autoantibodies, celiac disease also shows other features of autoimmunity, such as the specific killing of enterocytes, similar to the killing of β cells in type 1 diabetes. Thus, today it seems obvious that lessons learned from the immunology of celiac disease have relevance for other autoimmune diseases.

Chief predisposing role of cis- or trans-encoded HLA-DQA1 and HLA-DQB1 alleles

Like most autoimmune diseases, celiac disease has a strong HLA association. The major role of HLA as a susceptibility factor became very clear by way of the first genome-wide association study of celiac disease (Hunt et al. 2008). However, the first observations that the MHC influences the risk for celiac disease came nearly 50 years earlier with the finding of a disease association with HLA-B8 (Falchuk et al. 1972; Stokes et al. 1972). Later it was discovered that there are disease associations with HLA-A1, HLA-DR3, and HLA-DR7 as well as with HLA-DQ2 (Betuel et al. 1980; DeMarchi et al. 1979; Keuning et al. 1976; Solheim et al. 1977; Tosi et al. 1983). The associations with HLA-A1, -B8, -DR3 and -DQ2 is due to linkage disequilibrium and to the fact that these HLA alleles are part of a very conserved HLA haplotype (A1-B8-DR-DQ2; AH8.1). The strong degree of linkage disequilibrium in the HLA region imposes a general problem for dissecting, which is the culprit locus/allele of a disease. In many HLA-associated diseases this challenge has not been fully resolved. This is not so in celiac disease. In Southern Europe it was observed that a substantial proportion of celiac disease patients are DR5/DR7 heterozygous (Mearin et al. 1983; Trabace et al. 1984). Furthermore, it was observed that individuals who carry the DR7 allotype are only at risk if they carry the DR3 or DR5 allotypes, and individuals who carry the DR5 allotype are only at risk if they carry the DR3 or DR7 allotypes. The explanation for these observations became evident when the *DQA1* and *DQB1* alleles of the DR3, DR7 and DR5 haplotypes were characterized. The *DQA1* allele of the DR3-DQ2 haplotype (*DQA1*05:01*) was almost identical to the *DQA1* allele of the DR5-DQ7 (*DQA1*05:05*) (only differing at one single residue in the leader sequence), and the *DQB1* allele of the DR3-DQ2 haplotype (*DQB1*02:01*) was almost identical to the *DQB1* allele of the DR7-DQ2 haplotype (only differing at one single residue at position 135 in the membrane proximal domain). A DR3-DQ2 individual and a DR5-DQ7/DR7-DQ2 individual thus express a nearly identical HLA-DQ molecule, either expressed in *cis* or *trans* configuration (Sollid et al. 1989). This molecule is nowadays often termed HLA-DQ2.5 for short (Table 1).

The DQ region of the DR3-DQ2 haplotype is likely evolutionarily connected to the DR5-DQ7 and DR7-DQ2 haplotypes, with a site of DNA recombination located between the *DQA1* and *DQB1* loci. Conceivably, susceptibility to celiac disease depends on an interaction between two or more genes of the DR3-DQ2 haplotype that are reestablished in DR5-DQ7/DR7-DQ2 heterozygous individuals. The *DQA1* and *DQB1* are prime suspects because their gene products interact by forming a protein heterodimer, and because they are located close to the haplotype recombination site (Sollid et al. 1989). This evolutionary-based argument, which is different from the statistical analysis of odds ratios usually employed in HLA and disease association studies, provides very strong genetic evidence linking *DQA1*05* and *DQB1*02* to the etiology of celiac disease.

Usually 90% or more celiac disease patients carry HLA-DQ2.5. The patients who do not express HLA-DQ2.5 most often either express HLA-DQ2.2 (encoded by *DQA1*02:01* and *DQB1*02:02* alleles of the DR7-DQ2 haplotype) or they express HLA-DQ8 (encoded by *DQA1*03* and *DQB1*03:02* alleles of the DR4-DQ8 haplotype) (Karell et al. 2003; Spurkland et al. 1992) (Table 1). Some very few patients express HLA-DQ7.5 (encoded by *DQA1*05:05* and *DQB1*03:01* alleles, usually of the DR5-DQ7 haplotype) (Karell et al. 2003). The distribution of DQ2.5, DQ2.2 and DQ8 among celiac disease patients differ to some degree between different populations. For instance, a particularly high fraction of patients carry DQ8 among Ashkenazi Jews (Tighe et al. 1993) and Amerindians of Chile (Araya et al. 2000).

Further insight from comparisons of conserved HLA haplotypes

Many different conserved HLA haplotypes exist. In Sardinia there is a predominant association of celiac disease with the B18-DR3-DQ2 (AH18.2) haplotype (Congia et al. 1992). The B18-DR3-DQ2 (AH18.2) haplotype differs from the B8-DR3-DQ2 (AH8.1) haplotype at most loci including the *DPB1* locus and also the *DRB3* locus, where AH18.2 carries *DRB3*02:02* and AH8.1 carries *DRB3*01:01*. As both the AH8.1 and AH18.2 haplotypes are associated with celiac disease, this indicates that a region where the two ancestral haplotypes are identical confers the risk. The obvious candidates are *DQA1*05:01* and

*DQB1*02:01,* as these alleles are shared between the two haplotypes. Similarly, in Northern India celiac disease shows a strong association with the A26-B8-DR3-DQ2 (AH8.2) and the Ax-B21-DR3-DQ2 haplotypes (Kaur et al. 2002). These haplotypes differ from the AH8.1 haplotype at many loci (AH8.2 for instance carries the *DRB3*02* allele), but not at the *DQA1* and *DQB1* loci, where both carry *DQA1*05:01* and *DQB1*02:01* (Witt et al. 2002). Again, this observation argues for the involvement of *DQA1*05:01* and *DQB1*02:01* in the etiology of celiac disease.

Multiple non-HLA genes with small size effects also contribute to the risk

Almost no celiac disease patients exist that do not carry either one or more of the DQ2.5, DQ2.2 and DQ8 variants. However, most individuals carrying the celiac disease-associated HLA-DQ molecules will never develop celiac disease. Thus, certain HLA genes are thus necessary but not sufficient for the development of celiac disease. The explanation for this fact appears to be that particular non-HLA genes and possibly environmental factors other than gluten are mandatory for disease development. Celiac disease is thus a multifactorial and polygenic disorder. Over the last couple of decades considerable information has accumulated on the identity and nature of non-HLA genes in celiac disease, and the information mostly comes from genome-wide association studies (Dubois et al. 2010; Hunt et al. 2008; Trynka et al. 2011). As of 2016, 42 celiac-disease loci in addition to the HLA locus have been identified (Withoff et al. 2016). The effect size of each of these loci is very small, and compared to the effect size of HLA their individual effect sizes are almost negligible. It was estimated that whereas HLA contributed to 40% of the genetic risk, 37 non-HLA loci collectively contributed with 14% of the risk to the disease (Trynka et al. 2011). Issues remain on which of several single nucleotide polymorphisms (SNPs) in linkage disequilibrium at the disease susceptibility loci are responsible for the association, yet it appears clear that most of the primary candidates do not encode for change in the sequence of proteins (Withoff et al. 2016). In contrast, the SNPs are located in intergenic regions where they affect transcription, some of them by affecting the binding of transcription factors. It also appears that the polymorphisms primarily relate to regulation of the transcription of certain

genes in T cells and B cells. Thus, the unbiased insight coming from genetic studies supports the results of immunological studies of celiac disease patients, which I will return to, and clearly speak to an involvement of the adaptive immune system in the pathogenesis.

Gene-dosage effect of certain HLA-DQ alleles

The risk for celiac disease is markedly increased in individuals who are homozygous for HLA-DQ2.5 (*DQA1*05-DQB1*02*) (Ploski et al. 1993) and HLA-DQ8 (*DQA*03-DQB1*03:02*) (Karell et al. 2003). Individuals who are homozygous for *DQB1*02* but heterozygous for *DQA1*05* (i.e. *DQA1*05-DQB1*02:01/DQA1*0201-DQB1*02:02*) also have an increased risk for celiac disease (Ploski et al. 1993). A differential HLA risk profile was already noted, based on serological HLA-DR typing results, as DR3/DR3 and DR3/DR7 subjects were found to have increased risk compared to subjects who were DR3 heterozygous (Mearin et al. 1983; Trabace et al. 1984). With today's insight, we know that these features are explained by the extensive linkage disequilibrium between DR and DQ genes, and that the primary HLA-DQ associations dictate the biased distribution of DR allotypes among celiac disease patients.

No protective effect of HLA haplotypes or alleles

In type 1 diabetes there is a strong dominant protective effect of the DR2-DQ6 haplotype (Thomson et al. 1988). The effect of regulatory T cells could manifest as a dominant protective effect of HLA. Importantly, no protective effect of particular HLA haplotypes or alleles is observed in celiac disease. Looking at the frequency of *DQB1** alleles on the other chromosome in individuals who are *DQB1*02:01,* this does not follow the distribution of alleles in the control population. This could at first glance could be indicative of some protective effects. On closer inspection, however, it is apparent that this skewed distribution is due to a compensation of the excess of *DQB1*02:01* homozygous celiac disease patients (Ploski et al. 1993).

Is there any additional risk effect of non-DQ HLA genes?

A question has been whether any of the HLA region genes apart from HLA-DQ

contribute to celiac disease susceptibility. A slightly higher risk of the B8-DR3-DQ2 haplotype compared with the B18-DR3-DQ2 haplotype suggested that such risk genes do exist (Bolognesi et al. 2003; Louka et al. 2003). The analysis of 16 microsatellite markers and *DQA1* and *DQB1* alleles in 160 Dutch simplex *cis* DQ2.5-positive celiac disease families and 86 *cis* DQ2.5-positive control families revealed no significant differences, hence suggesting that DQ2.5 plays the major role (van Belzen et al. 2004). Further insight was obtained by an analysis of 12016 celiac disease cases and 11920 controls of European ancestry genotyped with the Illumina Immunochip array and imputing (Gutierrez-Achury et al. 2015). Stepwise logistic regression modeling assuming additive effects identified five independent variants with genome-wide significance beyond HLA-DQ: amino acid position 9 of HLA-DPβ1, *HLA-B*08:01* and *HLA-B*39:06* and two SNPs (rs1611710 and rs2301226). Knowing how difficult is is to fully control for linkage disequilibrium in the HLA region, some uncertainty probably remains with respect to localization of the identified risk factors. As for contribution to heritability, Gutierrez-Achury and coworkers (2015) estimated that 23% could be accounted for by HLA-DQ genes, 18% could be accounted for by non-DQ HLA genes and 7% could be accounted for by known non-HLA genes. Later it became clear that the estimate for non-DQ HLA genes was wrong, and the correct estimate is 2% and not 18% (Gutierrez-Achury 2015). Thus, within the HLA region the overriding risk effect for celiac disease is mediated by particular HLA-DQ allotypes with the modest contribution of certain non-DQ HLA genes.

HLA-DQ allotypes predispose to disease by presenting gluten peptide to CD4⁺ T cells

The role of DQ2.5 in predisposing for celiac disease became clear from experiments in which biopsies of celiac disease patients were stimulated with pepsin-trypsin digest of gluten and gluten-reactive $CD4^+$ T cells could be generated (Lundin et al. 1993). Strikingly, all the gluten-specific T cells were found to be restricted by the DQ2.5 molecule and not by any other HLA molecule expressed by the patients. This strongly suggested that the effect of HLA is mediated by the preferential presentation of gluten antigens to $CD4^+$ T cells in the periphery, although an effect of HLA molecules on T-cell selection in the thymus can not be excluded. Later it was found that when biopsies of DQ8 celiac

patients and DQ2.2 patients (who were not DQ2.5) were stimulated with gluten antigen, gluten-specific $CD4^+$ T cells restricted by these HLA molecules could be cultured (Bodd et al. 2012; Lundin et al. 1994). Thus, all celiac disease-associated HLA-DQ molecules act by their ability to preferentially present gluten antigens to $CD4^+$ T cells (Fig. 1).

The next challenge was to identify the epitope being recognized by these $CD4^+$ T cells. Gluten is an extremely complex antigen consisting of several hundred but slightly distinct proteins in a single wheat cultivar. This proved to be a formidable task. The work was greatly aided by mass spectrometry, which at the time had emerged as a powerful technology for protein and peptide analysis. A suspicion that the stimulatory gluten antigen was posttranslationally modified by deamidation became proven by this work. The first DQ2.5-restricted epitope identified was derived from a γ-gliadin protein, and in this epitope a glutamine-to-glutamate conversion (i.e. deamidation) at a particular position was crucial for T-cell recognition (Sjöström et al. 1998). A little later a DQ8-restricted epitope of α-gliadin was identified, but for this epitope deamidation was not found to be crucial for T-cell recognition (van de Wal et al. 1998b).

Gluten-specific T cells from the intestinal lesion were found to predominantly secrete interferon-(IFN)γ, but typically also many other cytokines including tumor necrosis factor, interleukin-(IL)4, IL-5, IL-10 and transforming growth factor (TGF)-β, hence having profiles compatible with TH1 or TH0 cells (Nilsen et al. 1995). Later it was found that gluten-specific T cells also produce IL-21, but not IL-17 (Bodd et al. 2010). The production of IFN-γ and IL-21 can thus be considered hallmarks of gluten-specific T cells in celiac disease.

TG2 generates post-translationally modified gluten peptides

The factor responsible for deamidation of gluten peptides remained for a short period unsolved. TG2 became a suspect because it was identified as the autoantigen in celiac disease (Dieterich et al. 1997) and because the enzyme, in addition to its main function of crosslinking polypeptides, was described to have the capacity to mediate deamidation. When non-deamidated (i.e. chymotrypsin

digested) gluten was treated with TG2 and then offered as antigen to biopsy-derived gluten-specific T cells, a vivid response could be recorded (Molberg et al. 1998). In the DQ2.5-restricted γ-gliadin epitope examined for TG2-mediated deamidation, not all glutamine residues were targeted by TG2, and strikingly the specificity of TG2 coincided with modifications that were important for T-cell recognition of the epitopes. Also for the DQ8-restricted α-gliadin epitope, TG2 treatment led to better T-cell recognition, further substantiating a role for TG2 in creating immunostimulatory gliadin peptides (van de Wal et al. 1998a). Thus, TG2 plays an essential role in making gluten peptides immunogenic for presentation by the HLA-DQ molecules DQ2 and DQ8 (Fig. 1).

A curiosity worth mentioning is that in the initial experiments the screening of the T-cell lines and T-cell clones was done with a peptic-tryptic digest of wheat gluten (Lundin et al. 1993). Pepsin digestion was done at low pH, which caused the spontaneous deamidation of the gluten antigen, thereby facilitating the detection of deamidation-dependent T cells. If a non-deamidated antigen (i.e. chymotrypsin digested gluten) had been used in these assays, the T-cell response would have gone unreported and no discovery would have been made. Luck definitely plays a role in science!

Over the years a number of DQ2.5- and DQ8-restricted gluten epitopes have been identified (Arentz-Hansen et al. 2000; Arentz-Hansen et al. 2002; Qiao et al. 2005; Stepniak et al. 2005; Tye-Din et al. 2010; Vader et al. 2002b). Some of the epitopes were identified by testing gluten-specific T cells from the peripheral blood of treated celiac-disease patients challenged with gluten (Tye-Din et al. 2010). Strikingly, the numbers of gluten-specific T cells recognizing deamidated epitopes increased dramatically on day six after initiation of a three-day oral gluten challenge in treated patients (Anderson et al. 2000; Anderson et al. 2005; Ráki et al. 2007). Generally, deamidation was found to be essential for all DQ2.5-restricted epitopes, and important for the majority of DQ8-restricted epitopes. A listing of known epitopes with their nomenclature has been published (Sollid et al. 2012).

Importance of stable peptide-HLA complexes for generating in vivo T-cell responses

What remained a conundrum was why DQ2.2 had a lower risk for celiac disease, yet most DQ2.5-restricted gluten epitopes could be presented by DQ2.2 molecules in a standard T-cell assay (Qiao et al. 2005; Vader et al. 2003). An analysis of peptide-HLA stability indicated that this could possibly be an underlying reason for the observed difference, as DQ2.5-restricted epitopes were only presented by DQ2.5 and not by DQ2.2 if antigen-presenting cells were pulsed with peptide antigen and incubated at extended time before T cells were added (i.e. a functional stability assay) (Fallang et al. 2009). The prediction from these studies was that DQ2.2 would present T-cell epitopes different from those presented by DQ2.5, and that these epitopes would bind stably to DQ2.2.

Binding preference of negatively charged peptides by disease-associated HLA-DQ molecules

In parallel with the identification of gluten T-cell epitopes, the binding motifs of the celiac disease associated HLA-DQ2 molecules were characterized. Both DQ2.5 and DQ8 were found to have a preference for binding negatively charged anchor residues albeit at different positions. Whereas DQ2.5 was found to prefer negatively charged anchors at positions P4, P6 and P7 (Johansen et al. 1996b; van de Wal et al. 1997; van de Wal et al. 1996; Vartdal et al. 1996), DQ8 was found to prefer negatively charged anchors at position P1 and P9 (Godkin et al. 1997; Kwok et al. 1996) (Fig. 2). As a consequence, the majority of gluten epitopes presented by DQ2.5 are distinct from those presented by DQ8 and vice versa (Tollefsen et al. 2006). Although DQ2.5 and DQ2.2 were found to have similar binding motifs with a preference for negatively charges anchors at P4, P6 and P7 (Johansen et al. 1996a; van de Wal et al. 1997), binding experiments suggested that DQ2.2, unlike DQ2.5, had an additional pocket at P3, with a preference for serine and threonine, and with proline being disfavored (van de Wal et al. 1997) (Fig. 2).

Gluten epitopes presented by DQ2.2 and DQ2.5 are distinct

The characterization of DQ2.2-restricted gluten epitopes revealed that these indeed are distinct from the DQ2.5-restricted epitopes (Bodd et al. 2012). Three

identified epitopes shared the feature of having a serine residue at the P3 position. When tested in the functional stability assay, it was found that the DQ2.2-restricted epitopes made stable complexes with DQ2.2 but not with DQ2.5. The serine residue was proven to be an important anchor for peptide binding to DQ2.2 but not to DQ2.5 (Bergseng et al. 2015). Serine and threonine are rarely distributed in gluten proteins. Thus the abundance of peptides in the TG2-modified gluten proteome that would be able to bind HLA would be less for DQ2.2 than for DQ2.5. This fact likely explains the different risks associated with the two DQ2 variants for celiac disease.

Gene-dosage effect of HLA-DQ relates to amount of peptide-HLA

The above-mentioned gene-dosage effect of HLA was – by testing in functional T-cell assays – ascribed to the enhanced presentation of gluten epitopes to $CD4^+$ T cells by HLA-DQ2.5 homozygous antigen-presenting cells (Vader et al. 2003). It was also suggested that there is a threshold of peptide-HLA complexes required for generating a pathogenic T-cell response in celiac disease. Recently, the notion of gene-dosage effect was contended, as similar levels of the cell-surface expression of DQ2.5 molecules and no difference in antigen presentation capability to gluten-specific $CD4^+$ T cells was seen when DQ2.5 homozygous and heterozygous celiac patients were compared (Pisapia et al. 2016). The absence of gene-dosage effects of MHC is surprising, given that this phenomenon has been demonstrated for I-E molecules in mice (Berg et al. 1990). What remains unclear is what explains the increased risk of individuals who are homozygous for *DQB1*02* but heterozygous for *DQA1*05* (i.e. *DQA1*05:01-DQB1*02:01*/*DQA1*02:01-DQB1*02:02*), and whether this only relates to an increased expression of DQ2.5 and an enhanced presentation of gluten epitopes to T cells in the periphery.

X-ray crystal structures relevant to gluten-specific T cells

Resolution of X-ray crystal structures have provided vital information on the molecular basis for the HLA association in celiac disease. The structure of the deamidated DQ2.5-glia-α1a epitope in complex with the DQ2.5 molecule revealed that Lys71β of DQ2.5 is instrumental in binding the P6 glutamate

anchor of this epitope (Kim et al. 2004). This same Lys71β residue is also instrumental in binding the P4 glutamate anchor in the DQ2.5-glia-α2 epitope (Jabri et al. 2014; Petersen et al. 2014). The latter structure, as it is a structure of the quaternary p-MHC-TCR complex, also revealed features of importance for T-cell recognition (Petersen et al. 2014). Quite surprisingly, an Arg109β of the TCR, which is part of a canonical motif in DQ2.5-glia-α2 epitope-specific TCRs carrying a semipublic TRBV7-2/TRAV26-1 receptor (Han et al. 2013; Qiao et al. 2014; Qiao et al. 2011), does not make direct contact with the P4 glutamate of the DQ2.5-glia-α2 epitope. Hence, the structural basis for TCRs being specific for deamidated gluten epitopes still remains unclear. A number of other X-ray crystal structures of quaternary complexes of p-MHC-TCR complexes relevant to celiac disease have been resolved, and they have also provided important information (Broughton et al. 2012; Petersen et al. 2015, 2016).

A role for regulatory T cells in celiac disease?

Conceivably, breakage of oral tolerance to gluten could be the underlying cause of celiac disease. This notion, together with the fact that the intestine is a hotbed for regulatory T cells, has fueled an interest in regulatory T cells in the pathogenesis of celiac disease. Several types of $CD4^+$ regulatory T cells have been described in the literature; $FOXP3^+$ T-regulatory (Tregs)cells that are either natural or induced (Sakaguchi et al. 2008) as well as Type 1 (Tr1) regulatory T cells that are hallmarked by the production of IL-10 and TGF-β (Roncarolo et al. 2006). In order to demonstrate that most subjects do not develop celiac disease, and that the majority of these subjects do not express the celiac disease-associated HLA-DQ allotypes, regulatory T cells controlling the anti-gluten response would need to recognize gluten peptides in the context of MHC class II molecules other than DQ2 or DQ8. Such cells have not been described, but a proper analysis is still lacking. However, the fact that there is no of protective effect of HLA alleles or haplotypes in celiac disease is a strong argument against the existence of regulatory T cells being restricted by protective alleles.

Regulatory cells that have been described in the literature in relation to celiac disease are all restricted by the disease-associated HLA-DQ allotypes. Gluten-

specific and HLA-DQ2.5-restricted Tr1 cells were described in celiac disease patients (Gianfrani et al. 2006). Clones of such Tr1 cells were found to suppress the proliferation of pathogenic Th0 cells. Of note, however, is that the Tr1 cells generally had concomitant production of IFN-γ along with IL-10, so whether they are really a separate cell type from the presumed pathogenic T cells is questionable. Moreover, circulating $FOXP3^{+}CD39^{+}$ Tregs reactive with gluten and also restricted by DQ2.5 were described recently in subjects with celiac disease (Cook et al. 2017). Cells of this phenotype surprisingly constituted 80% of gluten-specific cells in peripheral blood upon an oral gluten challenge involving treated celiac disease patients. The T cells appeared to use the same type of TCRs previously described among presumed pathogenic T cells, and they could be isolated with HLA-DQ2.5 tetramers representing a couple of immunodominant gluten epitopes. While peripheral polyclonal Tregs from celiac disease patients were found to have a normal suppressive function, the gluten-specific $FOXP3^{+}CD39^{+}$ Tregs after a short in vitro expansion exhibited a significantly reduced antigen-unspecific suppressive function compared to polyclonal Tregs. The authors interpreted this to indicate that the dysfunction of Tregs may contribute to the pathogenesis of celiac disease. Again, a pertinent question is whether these cells truly represent cells different from the cells considered to be pathogenic T cells. In humans, FOXP3 is a marker expressed by activated T cells, not only by Tregs, hence obscuring FOXP3 as a precise marker in clinical studies. Moreover, the finding of $FOXP3^{+}CD39^{+}$ Tregs in celiac disease patients does not explain why most DQ2.5 subjects never develop celiac disease. A relevant observation in this regard is that no gluten-specific and DQ2.5-restricted Tregs specific for immunodominant gluten epitopes could be identified in healthy subjects using HLA-DQ2.5 tetramers (Christophersen et al. 2016). Thus, taken together there is as yet no immunogenetic or functional evidence that regulatory T cells control the immune response to gluten in healthy subjects, and that a breach of this control results in celiac disease.

Antibodies to deamidated gluten peptides

Not only T-cell epitopes but also antibodies to gluten in celiac disease are preferentially directed against deamidated gluten epitopes (Osman et al. 2000).

Several epitopes have been identified. Particularly high serum antibody activity was observed with regard to peptides having the amino acid sequences QPEQPF (Osman et al. 2000), QPEQAFPE and PFPEQxFP (Ballew et al. 2013) and QPEQPF[PS]E (Pantazes et al. 2016). Plasma cells of the celiac lesion reactive with deamidated gluten peptides have been isolated and used to generate recombinant monoclonal antibodies (Steinsbø et al. 2014). These monoclonal antibodies had a biased usage of the *IGHV3-23/IGLV4-69* and *IGHV3-15/IGKV4-1* gene segments, and many of the antibodies were reactive with the peptide PLQPEQPFP. When the monoclonal antibodies were used to pull down peptides from complex proteolytic digests of deamidated gluten, it was striking that they typically brought down long peptide fragments with a representation of multiple copies of the antibody epitopes (Dørum et al. 2016). Recognition of multivalent gluten epitopes by B cells thus seems be important for generation of the antibody response. Interestingly, the long proteolytically stable gluten fragments harboring B-cell epitopes typically also harbor multiple distinct T-cell epitopes, speaking to T-cell and B-cell co-operation as an essential part of the celiac disease pathogenesis (Dørum et al. 2016).

Autoantibodies to TG2

TG2 being the target for autoantibodies in celiac disease as well as being instrumental in creating potent T-cell epitopes is hardly coincidental. A model that incorporates this dual role of TG2 is one in which complexes of TG2 and gluten play a key role (Sollid et al. 1997) (Fig. 3). In incubating TG2 and gluten, two types of covalent complexes are formed: one with an isopeptide linkage of TG2 surface lysine residues with gluten peptide glutamine residues and another in which the gliadin glutamine forms a thioester with TG2's active site cysteine (Fleckenstein et al. 2004). Such complexes may bind the surface B-cell receptor of TG2-specific B cells, and, upon internalization and processing, HLA-DQ molecules may bind deamidated gluten peptides for T-cell presentation. The B cells can in this way get the required T-cell help for plasma cell differentiation, and, importantly, the B cells will serve as antigen-presenting cells for T cells, thereby leading to amplification of the anti-gluten T-cell response. The engagement of TG2-specific B cells, but also B-cells specific for deamidated

gluten peptides, may thus play essential roles in mounting a pathogenic T-cell response in celiac disease. This may be a reason that anti-TG2 and anti-deamidated gliadin antibodies are excellent proxies for celiac disease in the clinic. This model regarding recognition of hapten-carrier-like complexes will also explain why TG2 antibodies are formed only in DQ2- and DQ8-expressing individuals when they are eating gluten.

Two modifications of the original hapten-carrier model have been suggested. One is based on the observation that IgD immunoglobulin is a substrate for TG2, and that the enzyme can crosslink gluten peptide to anti-TG2 IgD molecules on the surface of B cells, thereby facilitating the uptake of gluten peptides (Iversen et al. 2015). The other is based on the observation that TG2 is a good substrate for itself, and that multivalent complexes of TG2 can be generated by TG2 autocatalysis (Stamnaes et al. 2015). In the presence of gluten peptides, gluten peptides will be incorporated into these multimeric TG2 complexes – antigenic complexes that will be very potent B-cell antigens both for anti-TG2 and anti-deamidated gluten peptide-specific B cells. Multivalent antigen may be a reason that both anti-TG2-specific (Di Niro et al. 2012) and anti-deamidated gluten peptide-specific (Steinsbø et al. 2014) plasma cells have fewer mutations in their immunoglobulin genes than other plasma cells in the gut.

Rules governing the selection of gluten T-cell epitopes

Given the diversity of gluten proteins, usage of the same gluten epitopes across individuals suggests clear rules for epitope selection. The epitopes are confined to regions of gluten proteins rich in proline residues (Arentz-Hansen et al. 2002). The reason for this is twofold. First, proline-rich sequences are hard to digest by gut proteases, and such peptides will survive better with a minimal length required for HLA binding and T-cell receptor recognition (Shan et al. 2002). Second, proline is involved in substrate recognition by TG2 (Fleckenstein et al. 2002; Vader et al. 2002a). The role of TG2 in selecting T-cell epitopes is signified by the fact that the known T-cell epitopes are among the preferred substrates in very complex proteolytic digests of gluten (Dørum et al. 2010). The role of TG2 in epitope selection might well be mediated via an uptake of gluten-TG2 complexes

in B cells. Finally, binding to HLA molecules obviously are also important in the selection of T-cell epitopes. Epitopes able to generate stable HLA binding will survive and thereby prime T cells in vivo.

HLA-DQ as a drug target in celiac disease

Given its central role in the pathogenesis of celiac disease, HLA presents as an attractive drug target. The blocking of peptide binding to DQ2.5 with high-affinity HLA ligands could be observed (Xia et al. 2007; Xia et al. 2006). However, the blocking of T-cell activation was incomplete, particularly when live, non-fixed antigen-presenting cells were tested (Xia et al. 2007). Thus transforming this approach to an efficient treatment in vivo appears not to be straightforward. Another approach that also targets the presentation of gluten epitopes to T cells is the inhibition of cathepsin S. This inhibition aims to prevent degradation of the invariant chain in endosomes, and therefore exposure of the MHC II peptide-binding site. A phase 1 clinical trial (ClinicalTrials.gov Identifier: NCT02679014) exploring this approach in celiac disease with the drug RG7625 (Roche) is ongoing. Alternative targeting of gluten-specific T cells is an approach that is also pursued in celiac disease: for instance, by antigen-specific immunotherapy (ClinicalTrials.gov Identifier: NCT02528799).

Concluding remarks

Celiac disease is a disorder whose HLA association is now remarkably well understood. The perception of the disease over the last decades has changed from it being only a food hypersensitivity disorder to becoming a bona fide autoimmune disease. The fact that the immune reaction to the exogenous antigen gluten is driving autoimmune reactions and the immunopathology is raising the pertinent question as to whether yet undefined exogenous agents can also drive autoimmunity in other conditions such as, for instance, rheumatoid arthritis (Sollid and Jabri 2013). Thus, celiac disease, this food-induced condition, continues to give us food for thought.

References

Anderson RP, Degano P, Godkin AJ, Jewell DP, Hill AV (2000) In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. Nat Med 6*:*337-342

Anderson RP, van Heel DA, Tye-Din JA, Barnardo M, Salio M, Jewell DP, Hill AV (2005) T cells in peripheral blood after gluten challenge in coeliac disease. Gut 54*:*1217-23

Araya M, Mondragon A, Perez-Bravo F, Roessler JL, Alarcon T, Rios G, Bergenfreid C (2000) Celiac disease in a Chilean population carrying Amerindian traits. J Pediatr Gastroenterol Nutr 31*:*381-6

Arentz-Hansen H, Körner R, Molberg Ø, Quarsten H, Vader W, Kooy YM, Lundin KEA, Koning F, Roepstorff P, Sollid LM, McAdam SN (2000) The intestinal T cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. J Exp Med 191*:*603-612

Arentz-Hansen H, McAdam SN, Molberg Ø, Fleckenstein B, Lundin KE, Jorgensen TJ, Jung G, Roepstorff P, Sollid LM (2002) Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. Gastroenterology 123*:*803-809

Ballew JT, Murray JA, Collin P, Maki M, Kagnoff MF, Kaukinen K, Daugherty PS (2013) Antibody biomarker discovery through in vitro directed evolution of consensus recognition epitopes. Proc Natl Acad Sci USA 110*:*19330-5

Berg LJ, Frank GD, Davis MM (1990) The effects of MHC gene dosage and allelic variation on T cell receptor selection. Cell 60*:*1043-53

Bergseng E, Dørum S, Arntzen MO, Nielsen M, Nygard S, Buus S, de Souza GA, Sollid LM (2015) Different binding motifs of the celiac disease-associated HLA molecules DQ2.5, DQ2.2, and DQ7.5 revealed by relative quantitative proteomics of endogenous peptide repertoires. Immunogenetics 67*:*73-84

Betuel H, Gebuhrer L, Descos L, Percebois H, Minaire Y, Bertrand J (1980) Adult celiac disease associated with HLA-DRw3 and -DRw7. Tissue Antigens 15*:*231-238

Bodd M, Kim CY, Lundin KE, Sollid LM (2012) T-cell response to gluten in patients with HLA-DQ2.2 reveals requirement of peptide-MHC stability in celiac disease. Gastroenterology 142*:*552-61

Bodd M, Ráki M, Tollefsen S, Fallang LE, Bergseng E, Lundin KE, Sollid LM (2010) HLA-DQ2-restricted gluten-reactive T cells produce IL-21 but not IL-17 or IL-22. Mucosal Immunol 3*:*594-601

Bolognesi E, Karell K, Percopo S, Coto I, Greco L, Mantovani V, Suoraniemi E, Partanen J, Mustalahti K, Maki M, Momigliano-Richiardi P (2003) Additional factor in some HLA DR3/DQ2 haplotypes confers a fourfold increased genetic risk of celiac disease. Tissue Antigens 61*:*308-316

Broughton SE, Petersen J, Theodossis A, Scally SW, Loh KL, Thompson A, van Bergen J, Kooy-Winkelaar Y, Henderson KN, Beddoe T, Tye-Din JA, Mannering SI, Purcell AW, McCluskey J, Anderson RP, Koning F, Reid HH, Rossjohn J (2012) Biased T cell receptor usage directed against human leukocyte antigen DQ8-restricted gliadin peptides is associated with celiac disease. Immunity 37*:*611-21

Christophersen A, Risnes LF, Bergseng E, Lundin KE, Sollid LM, Qiao SW (2016) Healthy HLA-DQ2.5+ subjects lack regulatory and memory T cells specific for immunodominant gluten epitopes of celiac disease. J Immunol 196*:*2819-26

Congia M, Frau F, Lampis R, Frau R, Mele R, Cucca F, Muntoni F, Porcu S, Boi F, Contu L, et al. (1992) A high frequency of the A30, B18, DR3, DRw52, DQw2 extended haplotype in Sardinian celiac disease patients: further evidence that disease susceptibility is conferred by DQ A1*0501, B1*0201. Tissue Antigens 39*:*78-83

Cook L, Munier CM, Seddiki N, van Bockel D, Ontiveros N, Hardy MY, Gillies JK, Levings MK, Reid H, Peterson J, Rossjohn J, Anderson RP, Zaunders J, Tye-Din JA, Kelleher AD (2017) Circulating gluten-specific FOXP3+CD39+ regulatory T cells have impaired suppressive function in celiac disease. J Allergy Clin Immunol

DeMarchi M, Borelli I, Olivetti E, Richiardi P, Wright P, Ansaldi N, Barbera C, Santini B (1979) Two HLA-D and DR alleles are associated with coeliac disease. Tissue Antigens 14*:*309-316

Di Niro R, Mesin L, Zheng NY, Stamnaes J, Morrissey M, Lee JH, Huang M, Iversen R, du Pre MF, Qiao SW, Lundin KE, Wilson PC, Sollid LM (2012) High abundance of plasma cells secreting transglutaminase 2-specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions. Nat Med 18*:*441-5

Dicke WK (1950) Coeliac disease. Investigations of the harmful effects of certain types of cereal on patients suffering from coeliac disease. Thesis. University of Utrecht

Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, Schuppan D (1997) Identification of tissue transglutaminase as the autoantigen of celiac disease. Nat Med 3*:*797-801

Dørum S, Arntzen MO, Qiao SW, Holm A, Koehler CJ, Thiede B, Sollid LM, Fleckenstein B (2010) The preferred substrates for transglutaminase 2 in a complex wheat gluten digest are peptide fragments harboring celiac disease T-cell epitopes. PLoS One 5*:*e14056

Dørum S, Steinsbø Ø, Bergseng E, Arntzen MO, de Souza GA, Sollid LM (2016) Gluten-specific antibodies of celiac disease gut plasma cells recognize long proteolytic fragments that typically harbor T-cell epitopes. Sci Rep 6*:*25565

Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, Zhernakova A, Heap GA, Adany R, Aromaa A, Bardella MT, van den Berg LH, Bockett NA, de la Concha EG, Dema B, Fehrmann RS, Fernandez-Arquero M, Fiatal S, Grandone E, Green PM, Groen HJ, Gwilliam R, Houwen RH, Hunt SE, Kaukinen K, Kelleher D, Korponay-Szabo I, Kurppa K, MacMathuna P, Maki M, Mazzilli MC, McCann OT, Mearin ML, Mein CA, Mirza MM, Mistry V, Mora B, Morley KI, Mulder CJ, Murray JA, Nunez C, Oosterom E, Ophoff RA, Polanco I, Peltonen L, Platteel M, Rybak A, Salomaa V, Schweizer JJ, Sperandeo MP, Tack GJ, Turner G, Veldink JH, Verbeek WH, Weersma RK, Wolters VM, Urcelay E, Cukrowska B, Greco L, Neuhausen SL, McManus R, Barisani D, Deloukas P, Barrett JC, Saavalainen P, Wijmenga C, van Heel DA (2010) Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42*:*295-302

Falchuk ZM, Rogentine GN, Strober W (1972) Predominance of histocompatibility antigen HL-A8 in patients with gluten-sensitive enteropathy. J Clin Invest 51*:*1602-1605

Fallang LE, Bergseng E, Hotta K, Berg-Larsen A, Kim CY, Sollid LM (2009) Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation. Nat Immunol 10*:*1096-101

Fleckenstein B, Molberg Ø, Qiao SW, Schmid DG, Von Der MF, Elgstoen K, Jung G, Sollid LM (2002) Gliadin T cell epitope selection by tissue transglutaminase in celiac disease. Role of enzyme specificity and pH influence on the transamidation versus deamidation process. J Biol Chem 277*:*34109-34116

Fleckenstein B, Qiao SW, Larsen MR, Jung G, Roepstorff P, Sollid LM (2004) Molecular characterization of covalent complexes between tissue transglutaminase and gliadin peptides. J Biol Chem 279*:*17607-17616

Gianfrani C, Levings MK, Sartirana C, Mazzarella G, Barba G, Zanzi D, Camarca A, Iaquinto G, Giardullo N, Auricchio S, Troncone R, Roncarolo MG (2006) Gliadin-specific type 1 regulatory T cells from the intestinal mucosa of treated celiac patients inhibit pathogenic T cells. J Immunol 177*:*4178-86

Godkin A, Friede T, Davenport M, Stevanovic S, Willis A, Jewell D, Hill A, Rammensee HG (1997) Use of eluted peptide sequence data to identify the binding characteristics of peptides to the insulin-dependent diabetes susceptibility allele HLA-DQ8 (DQ 3.2). Int Immunol 9*:*905-911

Gutierrez-Achury HJ (2015) HLA and other tales: The different perspectives of celiac disease. Thesis. University of Groningen

Gutierrez-Achury J, Zhernakova A, Pulit SL, Trynka G, Hunt KA, Romanos J, Raychaudhuri S, van Heel DA, Wijmenga C, de Bakker PI (2015) Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. Nat Genet 47*:*577-8

Han A, Newell EW, Glanville J, Fernandez-Becker N, Khosla C, Chien YH, Davis MM (2013) Dietary gluten triggers concomitant activation of CD4+ and CD8+ αβ T cells and γδ T cells in celiac disease. Proc Natl Acad Sci USA 110*:*13073-8

Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M, Romanos J, Dinesen LC, Ryan AW, Panesar D, Gwilliam R, Takeuchi F, McLaren WM, Holmes GK, Howdle PD, Walters JR, Sanders DS, Playford RJ, Trynka G, Mulder CJ, Mearin ML, Verbeek WH, Trimble V, Stevens FM, O'Morain C, Kennedy NP, Kelleher D, Pennington DJ, Strachan DP, McArdle WL, Mein CA, Wapenaar MC, Deloukas P, McGinnis R, McManus R, Wijmenga C, van Heel DA (2008) Newly identified genetic risk variants for celiac disease related to the immune response. Nat Genet 40*:*395-402

Husby S, Koletzko S, Korponay-Szabo IR, Mearin ML, Phillips A, Shamir R, Troncone R, Giersiepen K, Branski D, Catassi C, Lelgeman M, Maki M, Ribes-Koninckx C, Ventura A, Zimmer KP, Diagnosis EWGoCD, Committee EG, European Society for Pediatric Gastroenterology H, Nutrition (2012) European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 54*:*136-60

Iversen R, du Pré MF, Di Niro R, Sollid LM (2015) Igs as substrates for transglutaminase 2: Implications for autoantibody production in celiac disease. J Immunol 195*:*5159-68

Jabri B, Chen X, Sollid LM (2014) How T cells taste gluten in celiac disease. Nat Struct Mol Biol 21*:*429-31

Johansen BH, Jensen T, Thorpe CJ, Vartdal F, Thorsby E, Sollid LM (1996a) Both α and β chain polymorphisms determine the specificity of the disease-associated HLA-DQ2 molecules, with β chain residues being most influential. Immunogenetics 45*:*142-150

Johansen BH, Vartdal F, Eriksen JA, Thorsby E, Sollid LM (1996b) Identification of a putative motif for binding of peptides to HLA-DQ2. Int Immunol 8*:*177-182

Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, Greco L, Ciclitira PJ, Sollid LM, Partanen J (2003) HLA types in celiac disease patients not carrying the DQA1*05-DQB1*02 (DQ2) heterodimer: results from the European Genetics Cluster on Celiac Disease. Hum Immunol 64*:*469-477

Kaur G, Sarkar N, Bhatnagar S, Kumar S, Rapthap CC, Bhan MK, Mehra NK (2002) Pediatric celiac disease in India is associated with multiple DR3-DQ2 haplotypes. HumImmunol 63*:*677-682

Keuning JJ, Pena AS, van Leeuwen A, van Hooff JP, van Rood JJ (1976) HLA-DW3 associated with coeliac disease. Lancet 1*:*506-508

Kim CY, Quarsten H, Bergseng E, Khosla C, Sollid LM (2004) Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. Proc Natl Acad Sci USA 101*:*4175-4179

Kwok WW, Domeier ML, Raymond FC, Byers P, Nepom GT (1996) Allele-specific motifs characterize HLA-DQ interactions with a diabetes-associated peptide derived from glutamic acid decarboxylase. J Immunol 156*:*2171-2177

Louka AS, Lie BA, Talseth B, Ascher H, Ek J, Gudjonsdottir AH, Sollid LM (2003) Coeliac disease patients carry conserved HLA-DR3-DQ2 haplotypes revealed by association of TNF alleles. Immunogenetics 55*:*339-343

Lundin KEA, Scott H, Fausa O, Thorsby E, Sollid LM (1994) T cells from the small intestinal mucosa of a DR4, DQ7/DR4, DQ8 celiac disease patient preferentially recognize gliadin when presented by DQ8. Hum Immunol 41*:*285-291

Lundin KEA, Scott H, Hansen T, Paulsen G, Halstensen TS, Fausa O, Thorsby E, Sollid LM (1993) Gliadin-specific, HLA-DQ(α1*0501,β1*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients. J Exp Med 178*:*187-196

Mearin ML, Biemond I, Pena AS, Polanco I, Vazquez C, Schreuder GT, de Vries RR, van Rood JJ (1983) HLA-DR phenotypes in Spanish coeliac children: their contribution to the understanding of the genetics of the disease. Gut 24*:*532-537

Meeuwisse GW (1970) European society for paediatric gastroenterology meeting in Interlaken September 18, 1969. Acta Paediat Scand 59*:*461-463

Molberg Ø, McAdam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, Fugger L, Scott H, Noren O, Roepstorff P, Lundin KE, Sjostrom H, Sollid LM (1998) Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. Nat Med 4*:*713-7

Nilsen EM, Lundin KEA, Krajci P, Scott H, Sollid LM, Brandtzaeg P (1995) Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma. Gut 37*:*766-776

Osman AA, Gunnel T, Dietl A, Uhlig HH, Amin M, Fleckenstein B, Richter T, Mothes T (2000) B cell epitopes of gliadin. Clin Exp Immunol 121*:*248-254

Pantazes RJ, Reifert J, Bozekowski J, Ibsen KN, Murray JA, Daugherty PS (2016) Identification of disease-specific motifs in the antibody specificity repertoire via next-generation sequencing. Sci Rep 6*:*30312

Petersen J, Kooy-Winkelaar Y, Loh KL, Tran M, van Bergen J, Koning F, Rossjohn J, Reid HH (2016) Diverse T cell receptor gene usage in HLA-DQ8-associated celiac disease converges into a consensus binding solution. Structure 24*:*1643-1657

Petersen J, Montserrat V, Mujico JR, Loh KL, Beringer DX, van Lummel M, Thompson A, Mearin ML, Schweizer J, Kooy-Winkelaar Y, van Bergen J, Drijfhout JW, Kan WT, La Gruta NL, Anderson RP, Reid HH, Koning F, Rossjohn J (2014) T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. Nat Struct Mol Biol 21*:*480-8

Petersen J, van Bergen J, Loh KL, Kooy-Winkelaar Y, Beringer DX, Thompson A, Bakker SF, Mulder CJ, Ladell K, McLaren JE, Price DA, Rossjohn J, Reid HH, Koning F (2015) Determinants of gliadin-specific T cell selection in celiac disease. J Immunol 194*:*6112-22

Pisapia L, Camarca A, Picascia S, Bassi V, Barba P, Del Pozzo G, Gianfrani C (2016) HLA-DQ2.5 genes associated with celiac disease risk are preferentially expressed with respect to non-predisposing HLA genes: Implication for anti-gluten T cell response. J Autoimmun 70*:*63-72

Ploski R, Ek J, Thorsby E, Sollid LM (1993) On the HLA-DQ(α1*0501, β1*0201)-associated susceptibility in celiac disease: a possible gene dosage effect of DQB1*0201 Tissue Antigens 41*:*173-177

Qiao SW, Bergseng E, Molberg Ø, Jung G, Fleckenstein B, Sollid LM (2005) Refining the rules of gliadin T cell epitope binding to the disease-associated DQ2 molecule in celiac disease: importance of proline spacing and glutamine deamidation. J Immunol 175*:*254-261

Qiao SW, Christophersen A, Lundin KE, Sollid LM (2014) Biased usage and preferred pairing of α- and β-chains of TCRs specific for an immunodominant gluten epitope in coeliac disease. Int Immunol 26*:*13-9

Qiao SW, Ráki M, Gunnarsen KS, Løset GA, Lundin KE, Sandlie I, Sollid LM (2011) Posttranslational modification of gluten shapes TCR usage in celiac disease. J Immunol 187*:*3064-71

Ráki M, Fallang LE, Brottveit M, Bergseng E, Quarsten H, Lundin KE, Sollid LM (2007) Tetramer visualization of gut-homing gluten-specific T cells in the peripheral blood of celiac disease patients. Proc Natl Acad Sci USA 104*:*2831-2836

Roncarolo MG, Gregori S, Battaglia M, Bacchetta R, Fleischhauer K, Levings MK (2006) Interleukin-10-secreting type 1 regulatory T cells in rodents and humans. Immunol Rev 212*:*28-50

Sakaguchi S, Yamaguchi T, Nomura T, Ono M (2008) Regulatory T cells and immune tolerance. Cell 133*:*775-87

Shan L, Molberg Ø, Parrot I, Hausch F, Filiz F, Gray GM, Sollid LM, Khosla C (2002) Structural basis for gluten intolerance in celiac sprue. Science 297*:*2275-2279

Sjöström H, Lundin KEA, Molberg Ø, Körner R, McAdam SN, Anthonsen D, Quarsten H, Noren O, Roepstorff P, Thorsby E, Sollid LM (1998) Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition. Scand J Immunol 48*:*111-115

Solheim BG, Albrechtsen D, Thorsby E, Thune P (1977) Strong association between an HLA-Dw3 associated B cell alloantigen and dermatitis herpetiformis. Tissue Antigens 10*:*114-118

Sollid LM, Jabri B (2013) Triggers and drivers of autoimmunity: lessons from coeliac disease. Nat Rev Immunol 13*:*294-302

Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E (1989) Evidence for a primary association of celiac disease to a particular HLA-DQ α/β heterodimer. JExpMed 169*:*345-350

Sollid LM, Molberg Ø, McAdam S, Lundin KEA (1997) Autoantibodies in coeliac disease: tissue transglutaminase--guilt by association? Gut 41*:*851-2

Sollid LM, Qiao SW, Anderson RP, Gianfrani C, Koning F (2012) Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. Immunogenetics 64*:*455-60

Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E (1992) HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. Hum Immunol 35*:*188-192

Stamnaes J, Iversen R, du Pre MF, Chen X, Sollid LM (2015) Enhanced B-cell receptor recognition of the autoantigen transglutaminase 2 by efficient catalytic self-multimerization. PLoS One 10*:*e0134922

Steinsbø Ø, Henry Dunand CJ, Huang M, Mesin L, Salgado-Ferrer M, Lundin KE, Jahnsen J, Wilson PC, Sollid LM (2014) Restricted VH/VL usage and limited mutations in gluten-specific IgA of coeliac disease lesion plasma cells. Nat Comm 5*:*4041

Stepniak D, Vader LW, Kooy Y, van Veelen PA, Moustakas A, Papandreou NA, Eliopoulos E, Drijfhout JW, Papadopoulos GK, Koning F (2005) T-cell recognition of HLA-DQ2-bound gluten peptides can be influenced by an N-terminal proline at p-1. Immunogenetics 57*:*8-15

Stokes PL, Asquith P, Holmes GK, Mackintosh P, Cooke WT (1972) Histocompatibility antigens associated with adult coeliac disease. Lancet 2*:*162-164

Thomson G, Robinson WP, Kuhner MK, Joe S, MacDonald MJ, Gottschall JL, Barbosa J, Rich SS, Bertrams J, Baur MP, et al. (1988) Genetic heterogeneity, modes of inheritance, and risk estimates for a joint study of Caucasians with insulin-dependent diabetes mellitus. Am J Hum Genet 43*:*799-816

Tighe MR, Hall MA, Ashkenazi A, Siegler E, Lanchbury JS, Ciclitira PJ (1993) Celiac disease among Ashkenazi Jews from Israel. A study of the HLA class II alleles and their associations with disease susceptibility. HumImmunol 38*:*270-276

Tollefsen S, Arentz-Hansen H, Fleckenstein B, Molberg Ø, Ráki M, Kwok WW, Jung G, Lundin KE, Sollid LM (2006) HLA-DQ2 and -DQ8 signatures of gluten T cell epitopes in celiac disease. J Clin Invest 116*:*2226-2236

Tosi R, Vismara D, Tanigaki N, Ferrara GB, Cicimarra F, Buffolano W, Follo D, Auricchio S (1983) Evidence that celiac disease is primarily associated with a DC locus allelic specificity. ClinImmunolImmunopathol 28*:*395-404

Trabace S, Giunta A, Rosso M, Marzorati D, Cascino I, Tettamanti A, Mazzilli MC, Gandini E (1984) HLA-ABC and DR antigens in celiac disease. A study in a pediatric Italian population. Vox Sang 46*:*102-106

Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, Bakker SF, Bardella MT, Bhaw-Rosun L, Castillejo G, de la Concha EG, de Almeida RC, Dias KR, van Diemen CC, Dubois PC, Duerr RH, Edkins S, Franke L, Fransen K, Gutierrez J, Heap GA, Hrdlickova B, Hunt S, Izurieta LP, Izzo V, Joosten LA, Langford C, Mazzilli MC, Mein CA, Midah V, Mitrovic M, Mora B, Morelli M, Nutland S, Nunez C, Onengut-Gumuscu S, Pearce K, Platteel M, Polanco I, Potter S, Ribes-Koninckx C, Ricano-Ponce I, Rich SS, Rybak A, Santiago JL, Senapati S, Sood A, Szajewska H, Troncone R, Varade J, Wallace C, Wolters VM, Zhernakova A, Thelma BK, Cukrowska B, Urcelay E, Bilbao JR, Mearin ML, Barisani D, Barrett JC, Plagnol V, Deloukas P, Wijmenga C, van Heel DA (2011) Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet 43*:*1193-201

Tye-Din JA, Stewart JA, Dromey JA, Beissbarth T, van Heel DA, Tatham A, Henderson K, Mannering SI, Gianfrani C, Jewell DP, Hill AV, McCluskey J, Rossjohn J, Anderson RP (2010) Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. Sci Transl Med 2*:*41ra51

Vader LW, de Ru A, van Der WY, Kooy YM, Benckhuijsen W, Mearin ML, Drijfhout JW, van Veelen P, Koning F (2002a) Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. J Exp Med 195*:*643-649

Vader W, Kooy Y, van Veelen P, de Ru A, Harris D, Benckhuijsen W, Pena S, Mearin L, Drijfhout JW, Koning F (2002b) The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. Gastroenterology 122*:*1729-1737

Vader W, Stepniak D, Kooy Y, Mearin L, Thompson A, van Rood JJ, Spaenij L, Koning F (2003) The HLA-DQ2 gene dose effect in celiac disease is

directly related to the magnitude and breadth of gluten-specific T cell responses. Proc Natl Acad Sci USA 100*:*12390-12395

van Belzen MJ, Koeleman BP, Crusius JB, Meijer JW, Bardoel AF, Pearson PL, Sandkuijl LA, Houwen RH, Wijmenga C (2004) Defining the contribution of the HLA region to cis DQ2-positive coeliac disease patients. Genes Immun 5*:*215-220

van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, Papadopoulos G, Koning F (1998a) Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. J Immunol 161*:*1585-1588

van de Wal Y, Kooy YC, Drijfhout JW, Amons R, Papadopoulos GK, Koning F (1997) Unique peptide binding characteristics of the disease-associated DQ(α1*0501, β1*0201) vs the non-disease-associated DQ(α1*0201, β1*0202) molecule. Immunogenetics 46*:*484-492

van de Wal Y, Kooy YM, van Veelen PA, Pena SA, Mearin LM, Molberg Ø, Lundin KEA, Sollid LM, Mutis T, Benckhuijsen WE, Drijfhout JW, Koning F (1998b) Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. Proc Natl Acad Sci USA 95*:*10050-10054

van de Wal Y, Kooy YMC, Drijfhout JW, Amons R, Koning F (1996) Peptide binding characteristics of the coeliac disease-associated DQ(α1*0501, β1*0201) molecule. Immunogenetics 44*:*246-253

Vartdal F, Johansen BH, Friede T, Thorpe CJ, Stevanovic S, Eriksen JA, Sletten K, Thorsby E, Rammensee HG, Sollid LM (1996) The peptide binding motif of the disease associated HLA-DQ(α1*0501, β1*0201) molecule. Eur J Immunol 26*:*2764-2772

Walker-Smith JA, Guandalini S, Schmitz J, Shmerling DH, Visakorpi JK (1990) Revised criteria for diagnosis of coeliac disease. Arch Dis Child 65*:*909-911

Withoff S, Li Y, Jonkers I, Wijmenga C (2016) Understanding celiac disease by genomics. Trends Genet 32*:*295-308

Witt CS, Price P, Kaur G, Cheong K, Kanga U, Sayer D, Christiansen F, Mehra NK (2002) Common HLA-B8-DR3 haplotype in Northern India is different from that found in Europe. Tissue Antigens 60*:*474-480

Xia J, Bergseng E, Fleckenstein B, Siegel M, Kim CY, Khosla C, Sollid LM (2007) Cyclic and dimeric gluten peptide analogues inhibiting DQ2-mediated antigen presentation in celiac disease. Bioorg Med Chem 15*:*6565-6573

Xia J, Siegel M, Bergseng E, Sollid LM, Khosla C (2006) Inhibition of HLA-DQ2-mediated antigen presentation by analogues of a high affinity 33-residue peptide from α2-gliadin. J Am Chem Soc 128*:*1859-1867

Figure legends

Fig. 1. Depiction of key events in the activation of tissue-resident $CD4^+$ T cells recognizing deamidated gluten peptides in the context of the celiac disease-associated HLA molecules DQ2 (DQ2.5 and DQ2.2) and DQ8. Ingested gluten proteins are hard to process for gastrointestinal proteases due to their high content of proline residues, and long peptide fragments survive the digestion. Somehow these long peptide fragments manage to cross the single layer of enterocytes traversing from the lumen to the lamina propria. Some glutamine residues in a sequence-specific manner become targeted by the enzyme transglutaminase 2 (TG2) and become deamidated (i.e. glutamine-to-glutamate conversion). Exactly where this deamidation takes place is not known. Deamidated peptides will be picked up by antigen-presenting cells (APC) expressing the HLA molecules DQ2 and DQ8. The deamidated gluten peptides bind better to these HLA-DQ molecules than their native counterparts, and thereby the deamidated peptides can be presented to the gluten-specific $CD4^+$ T cells. The gluten-specific $CD4^+$ T cells secrete many cytokines including interferon (IFN)-γ and interleukin (IL)-21. These $CD4^+$ T cells serve as director of the immunological orchestra, directing the immune reactions that eventually lead to formation of the celiac lesion. The killing of enterocytes is likely mediated by intraepithelial $CD8^+$ T cells.

Fig. 2. The celiac disease-predisposing HLA molecules DQ2.5, DQ2.2 and DQ8 all have a preference for binding peptides with negatively charged anchor residues. In DQ2.5 and DQ2.2 it is the pockets P4, P6 and P7 that have a preference for negatively charged anchor residues, whereas in DQ8 it is the pockets P1 and P9 that have this preference. A major pocket for peptides binding to DQ2.2 is located at P3, and this pocket has a preference for binding serine (S) and threonine residues. Native gluten peptides are rich in proline (P) and glutamine (Q) residues, but scarce in negatively charged residues. Upon deamidation by transglutaminase 2 (TG2), glutamine residues in a sequence- dependent manner are converted to glutamate (E). The deamidated gluten peptides with the introduced negatively charged glutamate residues bind better to celiac disease-

associated HLA-DQ molecules. Schematically depicted are the native and deamidated variants of the immunodominant epitopes DQ2.5-glia-α2, DQ2.2-glut-L1 and DQ8-glia-α1.

Figure 3 This model explains how anti-transglutaminase 2 (TG2) autoantibodies can be formed in HLA-DQ2- or HLA-DQ8-positive celiac disease patients when they eat gluten. The incubation of TG2 with gluten peptides led to the formation of covalently linked TG2-gluten peptide complexes. Such complexes can be bound by the B-cell receptor of TG2-specific B cells. The complexes are internalized and processed in endosomes of the B cells. Released gluten peptides, which will need to become deamidated, will inside the B cells bind to the HLA-DQ molecules so that they can subsequently be displayed on the surface of the cells for recognition by T cells. Celiac disease patients have gluten-specific $CD4^+$ T, which − with the involvement of accessory molecules and T-cell-derived cytokines − provides help to B cells so that they can differentiate to plasma cells and produce antibodies specific for TG2. In this process, B cells serve as antigen-presenting cells for T cells, and the gluten-specific $CD4^+$ T cells will become activated and clonally expand.

Table 1

HLA-DQ molecules with corresponding alleles and haplotypes that confer the risk of celiac disease

HLA molecule	HLA alleles	HLA haplotype	Risk
DQ2.5 *cis*	*DQA1*05:01-DQB1*02:01*	DR3-DQ2	High
DQ2.5 *trans*	*DQA1*05:05-DQB1*03:01*	DR5-DQ7	High
	*DQA1*02:01-DQB1*02:02*	DR7-DQ2	
DQ2.2	*DQA1*02:01-DQB1*02:02*	DR7-DQ2	Low
DQ7.5	*DQA1*05:05-DQB1*03:01*	DR5-DQ7	Very low
DQ8	*DQA1*03:01-DQB1*03:02*	DR4-DQ8	Low

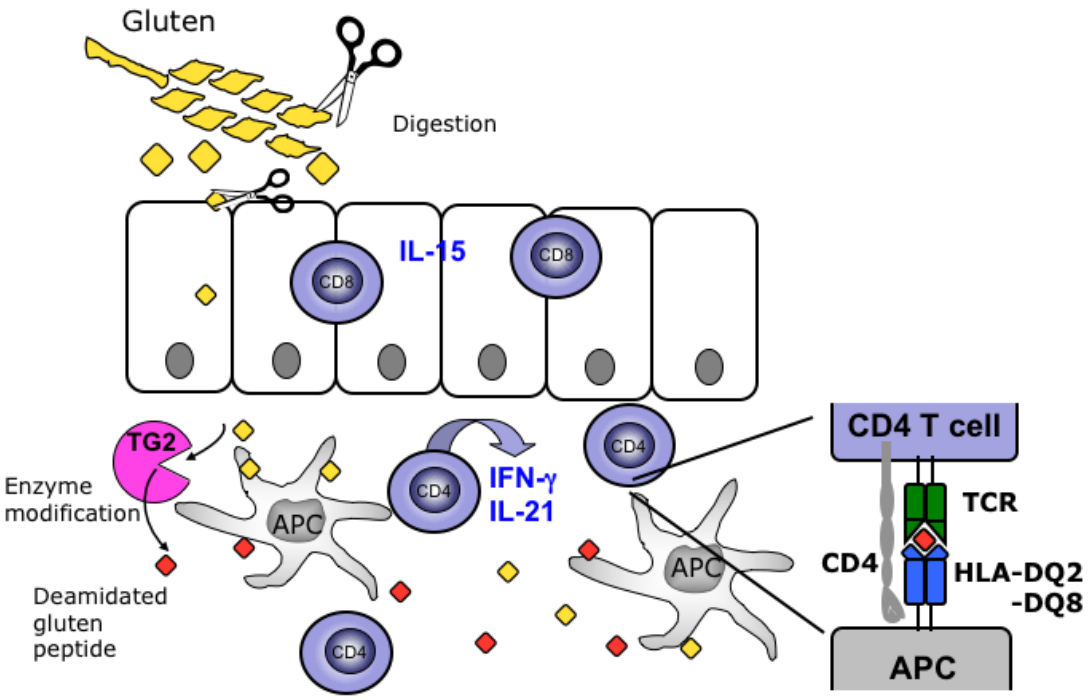
Gluten
Digestion
IL-15
CD8
CD8
TG2
Enzyme
modification
APC
CD4
IFN-γ
IL-21
CD4
Deamidated
gluten
peptide
CD4
APC
CD4 T cell
TCR
CD4
HLA-DQ2
-DQ8
APC

Figure 1

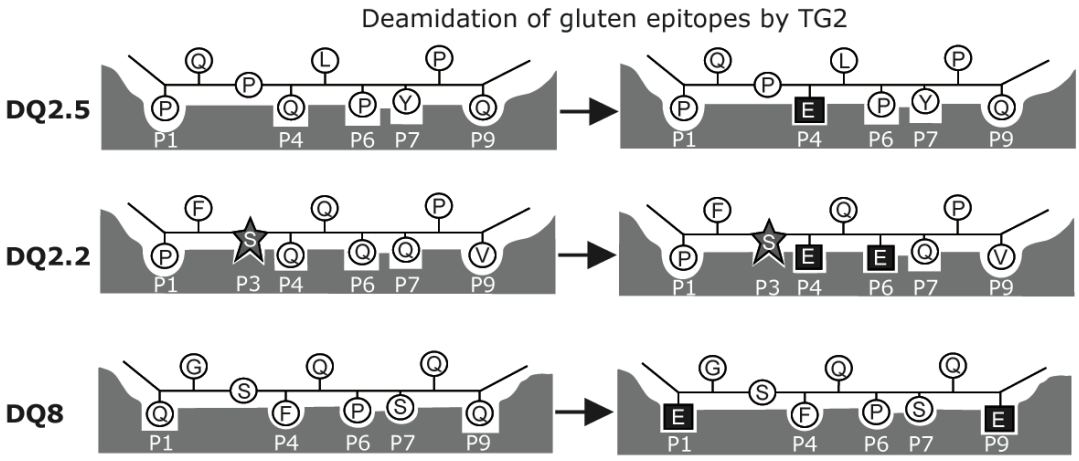
Deamidation of gluten epitopes by TG2
DQ2.5
DQ2.2
DQ8
P1
P3
P4
P6
P7
P9

Figure 2

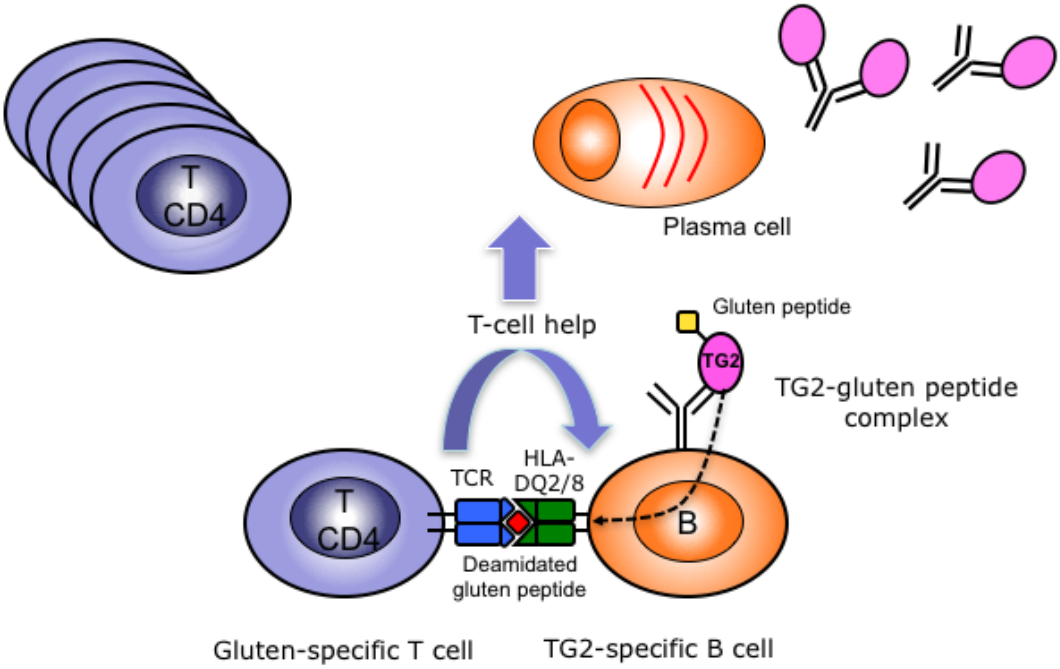
T
CD4
Plasma cell
T-cell help
Gluten peptide
TG2
TG2-gluten peptide
complex
TCR
HLA-
DQ2/8
T
CD4
B
Deamidated
gluten peptide
Gluten-specific T cell
TG2-specific B cell

Figure 3