

# A biased view toward celiac disease

J Rossjohn<sup>1,2,3</sup> and F Koning<sup>4</sup>

**Celiac disease is an autoimmune-like disorder that is triggered by dietary gluten and has a strong genetic association with the human leukocyte antigen locus, specifically, HLA-DQ2.5/DQ8. Here, Dahai-Koirala *et al.* apply *ex vivo* single-cell sequencing of TCRs from celiac disease patients, and show that biased T-cell receptor usage underpins the response to two gluten epitopes, which has implications for disease pathogenesis, diagnosis, and treatment.**

Celiac disease (CD) patients are intolerant to gluten, proteins found in wheat and the closely related cereals barley and rye. Gluten-specific T cells and autoimmune antibodies against the enzyme tissue transglutaminase (TG2) are typically found in patients, whereas both are absent in healthy controls.<sup>1–3</sup> CD shares an important feature with autoimmune-mediated diseases such as type 1 diabetes, rheumatoid arthritis, and narcolepsy: a strong HLA-association. In CD this association is explained as the gluten-specific T cells almost invariably respond to gluten peptides bound to the disease predisposing HLA-DQ2.5 and/or HLA-DQ8 molecules. Most of these T-cell stimulatory gluten peptides have been modified by the TG2, a modification that strongly enhances the binding of such gluten epitopes to either HLA-DQ2.5 or HLA-DQ8. In return, the gluten-specific

T cells likely drive the production of TG2-specific autoantibodies by giving help to TG2-specific B cells that can bind to TG2 crosslinked to gluten fragments. This creates a powerful amplification loop; as such, B cells efficiently internalize and process these TG2-gluten complexes, which would lead to strongly enhanced presentation of gluten peptides bound to HLA-DQ on TG2-specific B cells (**Figure 1**).

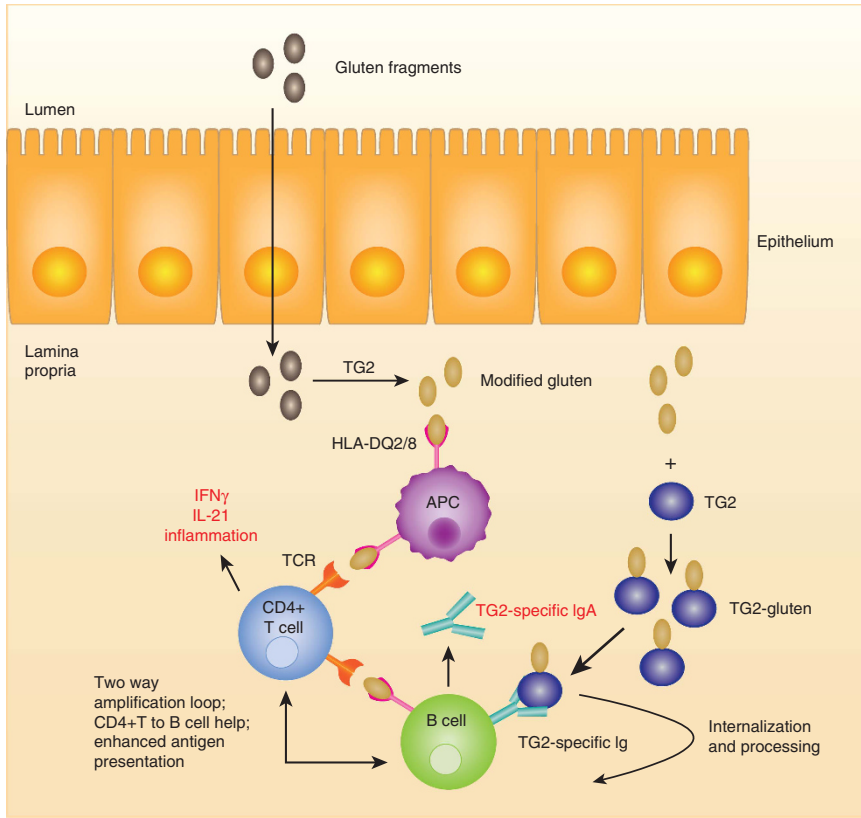
It is well-known that compared with HLA-DQ8, HLA-DQ2.5 is more strongly associated with the disease as some 95% of patients are HLA-DQ2.5 positive.<sup>1–3</sup> This is not related to the gene frequency as both HLA-alleles have rather similar frequencies in the western hemisphere. Also, although 17 HLA-DQ2.5-restricted gluten epitopes have been identified, only 4 epitopes exist for HLA-DQ8.<sup>4</sup> Finally, it is striking that

although many of the DQ2.5 epitopes contain multiple proline residues, rendering them highly resistant to degradation,<sup>5,6</sup> the most immunodominant DQ8 epitope contains only one proline and is thus likely more sensitive to enzymatic degradation in the gastrointestinal tract.<sup>7</sup> Altogether, these observations indicate that the existence of a large repertoire of gluten epitopes that are resistant to degradation leads to a high chance of disease development. This goes hand-in-hand with the well-established HLA-DQ2.5 gene dose effect where homozygosity for HLA-DQ2.5 leads to an at least fivefold higher chance of disease compared with HLA-DQ2/X heterozygosity and this has been linked to the strength of the gluten-specific T-cell response.<sup>8</sup> Nevertheless, even though gluten contains many immunogenic sequences, and T-cell responses to multiple gluten peptides are commonly observed in patients, particularly sequences in two gluten protein families, the  $\alpha$ - and the  $\omega$ -gliadins, are clearly immunodominant as T-cell responses to these peptides are observed in the large majority of patients, implying a crucial role for such responses in the disease process.<sup>6,9–11</sup>

In a number of recent studies, the T-cell receptor (TCR) repertoire expressed by T cells specific for these immunodominant gliadin epitopes has been investigated<sup>12–16</sup> with Dahai-Koirala *et al.*<sup>17</sup> being the latest addition. Notably, the authors have for the first time performed *ex vivo* single-cell sequencing of TCRs reactive to the immunodominant and highly homologous DQ2.5-glia- $\alpha$ 2 and DQ2.5-glia- $\omega$ 2 peptides (**Table 1**).

<sup>1</sup>Australian Research Council Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton, Victoria, Australia. <sup>2</sup>Institute of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, UK. <sup>3</sup>Department of Biochemistry and Molecular Biology, Infection and Immunity Program, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia and <sup>4</sup>Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands. Correspondence: F Koning (F.Koning@lumc.nl)

Published online 23 March 2016. doi:10.1038/mi.2016.17



**Figure 1** CD4<sup>+</sup> T- and B-cell responses. For details see first paragraph of main text.

**Table 1** Amino acid sequences of four immunodominant gliadin epitopes

	P1	P2	P3	P4	P5	P6	P7	P8	P9
DQ2.5-glia- $\alpha$ 1	P	F	<b>P</b>	Q	<b>P</b>	E	<b>L</b>	<b>P</b>	Y
DQ2.5-glia- $\alpha$ 2	P	Q	P	E	<b>L</b>	P	<b>Y</b>	P	Q
DQ2.5-glia- $\omega$ 2	P	Q	<b>P</b>	E	<b>Q</b>	P	<b>F</b>	P	W
DQ8-glia- $\alpha$ 1	E	G	<b>S</b>	F	<b>Q</b>	P	S	<b>Q</b>	E

Shown is the p1–p9 core sequence which occupies the peptide-binding groove of HLA-DQ2.5. The DQ2.5-glia- $\alpha$ 2 and DQ2.5-glia- $\omega$ 2 epitopes are the subject of the current study by Dohai-Koirala *et al.* Bold font indicates T-cell receptor contact residue, whereas bold and italics font indicates major T-cell receptor contact residue. Amino acid code: P = proline; F = phenylalanine; Q = glutamine; E = glutamic acid; L = leucine; Y = tyrosine; W = tryptophane; G = glycine; S = serine.

Importantly, these cells were isolated directly from peripheral blood of patients with corresponding HLA-DQ-gliadin tetramers and thus representative of the actual TCR repertoire specific for the gluten epitopes investigated. They observed clonal dominance in the T-cell response to both gliadin epitopes: the response is dominated by a few clones in all three patients examined. Also, the most used TRAV and TRBV genes accounted for roughly 50% of the T cells analyzed and there is a bias in the

TCR gene usage as similar TRAV and TRBV usage was observed in these patients. Although the DQ2.5-glia- $\alpha$ 2 and DQ2.5-glia- $\omega$ 2 peptides are highly related (**Table 1**), it was notable that the extent of TCR crossreactivity toward these epitopes was limited. As the current data confirm the previously reported preferential usage of TRAV26.1 in combination with TRBV7-2 by DQ2.5-glia- $\alpha$ 2 specific T cells,<sup>12,14,15</sup> there was no preferential pairing observed in DQ2.5-glia- $\omega$ 2

T cells. However, frequent usage of TRAV4 by both DQ2.5-glia- $\alpha$ 2- and DQ2.5-glia- $\omega$ 2-specific T cells was found (**Table 2**). Thus, there are similarities and differences between the expressed TCR repertoires, where the observation that such T cells express a biased and public TCR repertoire stands out. This gives further weight to previous observations that have reported the presence of a biased gluten-specific TCR repertoire against the immunodominant gliadin peptides DQ2.5-glia- $\alpha$ 1 and DQ2.5-glia- $\omega$ 2 in HLA-DQ2.5-positive patients<sup>12,14,15</sup> and against DQ8-glia-1 in HLA-DQ8-positive patients.<sup>13,16</sup> Although the TRAV26/TRBV7-2 combination is dominant in the response to DQ2.5-glia-2, DQ8-glia-1-specific T cells preferentially express TRAV26/TRBV9. The available ternary crystal structures between TRAV26<sup>+</sup> TCRs and DQ2.5-glia-2<sup>14</sup> and DQ8-glia-1<sup>13</sup> demonstrate that such TCRs are positioned very similarly above the HLA-DQ2 and HLA-DQ8  $\beta$ -chains, likely facilitated by the highly similar amino acid sequences of the HLA-DQ2.5 and -DQ8  $\beta$ -chains that interact with these TCRs. Similar mechanisms may underlie the observed shared usage of TRAV4, TRAV7, and TRBV4 by DQ2.5-glia- $\alpha$ 2- and DQ2.5-glia- $\omega$ 2-specific T cells and that of TRBV20 by DQ2.5-glia- $\alpha$ 1- and DQ2.5-glia- $\alpha$ 2-specific T cells. As noted by Dahai-Koirala *et al.*,<sup>17</sup> it is well possible that thymic selection of an HLA-DQ2.5-restricted TCR repertoire may be partially responsible for the observed bias. But next to the biased TCR repertoire there are features of the CDR3 regions that underlie the specific interaction between these TCR and the DQ-gliadin peptide complexes. In particular, a non-germline encoded CDR3 Arg residue is found in virtually all TCRs specific for DQ2.5-glia- $\alpha$ 2 and DQ8-glia-1. The availability of a number of structures between these TCRs and their ligands has revealed the molecular basis for the requirement of this Arg as it interacts with the p5-leucine residue in the DQ2.5-glia- $\alpha$ 2 epitope. In addition, a CDR3 encoded Asp-residue forms a hydrogen bond with the p7-tyrosine residue in DQ2.5-glia- $\alpha$ 2. With this in

**Table 2 Shared and unique TRAV and TRBV gene usage by T cells specific for immunodominant gliadin epitopes**

	DQ2.5-glia- $\alpha$ 1	DQ2.5-glia- $\alpha$ 2	DQ2.5-glia- $\omega$ 2	DQ8-glia- $\alpha$ 1
TRAV	–	26		26
		4	4	
			12	8
TRBV		7	7	9
		4	4	
	20	20		(20)
				6

Bold values indicate strong correlation between the recognition of the gliadin epitope and the TRAV and/or TRBV gene segment indicated.

mind it comes as no surprise that the CDR3 regions of the DQ2.5-glia- $\omega$ 2-specific TCRs are quite distinct from those specific for the highly homologous DQ2.5-glia- $\alpha$ 2 epitope as the two critical amino acid differences between these peptides concern the p5 and p7 TCR contact residues (**Table 1**). The rules that apply to recognition of the  $\alpha$ - and  $\omega$ -epitopes are thus different. Similarly, the TCR repertoire specific for the largely overlapping DQ2.5-glia- $\alpha$ 1 and DQ2.5-glia- $\alpha$ 2 are distinct as the different binding registers of these peptides (**Table 1**) imposes different requirements for a productive interaction between the TCR and its ligand.

#### IMPLICATIONS FOR CD

Although it is obvious that gluten is immunogenic, particularly after modification by TG2, the large majority of HLA-predisposed individuals will never develop CD. Also, the disease can essentially manifest at any age. This is most commonly explained by the multiple hit model: a combination of intrinsic factors and (multiple) external stimuli are required to precipitate disease development. Although HLA-DQ is clearly the most important intrinsic factor known, the nature of other contributing factors remains largely undetermined. In this respect it worthwhile to consider the fact that gluten-specific HLA-DQ2.5 and/or -DQ8-restricted T cells are commonly detected in patients with

CD but never in healthy controls. In combination with the recent TCR repertoire studies this implies that what distinguishes patients from controls is the presence of T cells specific for immunodominant gliadin peptides, some of which express a highly similar TCR. The latter is particularly the case for the DQ2.5-glia- $\alpha$ 2 and the DQ8-glia-1 epitopes where particular rules apply as these T cells share both highly conserved TRAV/TRBV pairing and a critical non-germline encoded Arg residue in either CDR3 $\alpha$  or CDR3 $\beta$ . As the generation of the TCR repertoire is based on a stochastic process, randomness enters the disease arena: could it be that the development of CD depends on the generation of high-affinity TCRs specific for particular gliadin peptides? It may seem farfetched but there are some issues that speak in favor. First, the potential TCR repertoire is vast and only a limited number of the potential TCRs will ever be made in a lifetime, indicating that not each and every individual will necessarily generate TCR that have high affinity for immunodominant gluten epitopes. Second, and in agreement with the above, so far no T cells specific for gliadin have been found in controls, indicating that even if they exist their numbers are very low. Third, affinity measurements demonstrate that DQ2.5-glia- $\alpha$ 2- and the DQ8-glia-1-specific TCR indeed have relatively high affinity for their ligands<sup>13,14,16</sup> and it is tempting to

speculate that this may be essential for the breaking of tolerance.

Does this explain everything? Certainly not. It is well-known that first-degree relatives of CD patients have 10% chance of developing the disease.<sup>1-3</sup> Thus at least 10% of the population should generate gluten-specific T cells but only 1% develops CD. So there are other factors as well, like environmental insults including, but not limited to, gastrointestinal infections that lead to a (temporary) break of tolerance in which gluten-specific T cells, which normally are kept in check because of the powerful tolerogenic milieu in the intestine, may suddenly have a chance to develop into effector memory T cells. To shed further light on the issue of randomness, it may be very rewarding to determine the frequency of “normal” T cells that share features with high-affinity gliadin-specific T cells, like the observed preferential TRAV/TRBV pairing in combination with the required CDR Arg residue. In any case, the realization that high-affinity gliadin-specific T cells are shared among patients may provide a novel tool to monitor disease development in individuals at risk as the appearance and expansion of such cells with a memory phenotype should have the alarm bell ringing. Finally, as noted by Dahai-Koirala *et al.*,<sup>17</sup> it is tempting to speculate that T cells expressing a biased TCR repertoire may have a role in other HLA-linked diseases as well, like in type 1 diabetes, rheumatoid arthritis, and narcolepsy. Searching for such cells may thus provide valuable tools for the identification of the target antigens of such autoreactive T cells, which would provide novel opportunities for prevention and therapies in those diseases as well.

#### ACKNOWLEDGMENTS

JR is supported by the National Health and Medical Research Council of Australia and the Australian Research Council.

#### DISCLOSURE

The authors declare no conflict of interest.

## REFERENCES

1. Tjon, J.M., van Bergen, J. & Koning, F. Celiac disease: how complicated can it get? *Immunogenetics* **62**, 641–651 (2010).
2. Abadie, V., Sollid, L.M., Barreiro, L.B. & Jabri, B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu. Rev. Immunol.* **29**, 493–525 (2011).
3. Koning, F., Thomas, R., Rossjohn, J. & Toes, R.E. Coeliac disease and rheumatoid arthritis: similar mechanisms, different antigens. *Nat. Rev. Rheumatol.* **11**, 450–461 (2015).
4. Sollid, L.M., Qiao, S.W., Anderson, R.P., Gianfrani, C. & Koning, F. Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* **64**, 455–460 (2012).
5. Arentz-Hansen, H. *et al.* Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology* **123**, 803–809 (2002).
6. van de Wal, Y. *et al.* Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J. Immunol.* **161**, 1585–1588 (1998).
7. Vader, W. *et al.* 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 (2003).
8. Arentz-Hansen, H. *et al.* The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J. Exp. Med.* **191**, 603–612 (2000).
9. Anderson, R.P., Degano, P., Godkin, A.J., Jewell, D.P. & Hill, A.V. 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 (2000).
10. Shan, L. *et al.* Structural basis for gluten intolerance in celiac sprue. *Science* **297**, 2275–2279 (2002).
11. Tye-Din, J.A. *et al.* Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* **2**, 41ra51 (2010).
12. Qiao, S.W., Christophersen, A., Lundin, K.E. Sollid, L.M. Biased usage and preferred pairing of alpha- and beta-chains of TCRs specific for an immunodominant gluten epitope in coeliac disease. *Int. Immunol.* **26**, 13–19 (2014).
13. Qiao, S.W. *et al.* Posttranslational modification of gluten shapes TCR usage in celiac disease. *J. Immunol.* **187**, 3064–3071 (2011).
14. Broughton, S.E. *et al.* Biased T cell receptor usage directed against human leukocyte antigen DQ8-restricted gliadin peptides is associated with celiac disease. *Immunity* **37**, 611–621 (2012).
15. Petersen, J. *et al.* T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. *Nat. Struct. Mol. Biol.* **21**, 480–488 (2014).
16. Petersen, J. *et al.* Determinants of Gliadin-specific T cell selection in celiac disease. *J. Immunol.* **194**, 6112–6122 (2015).
17. Dahal-Koirala, S. *et al.* TCR sequencing of single cells reactive to DQ2.5-glia- $\alpha$ 2 and DQ2.5-glia- $\omega$ 2 reveals clonal expansion and epitope-specific V-gene usage. *Mucosal Immunol.* **9**, 587–596 (2016).