

A TCR α framework-centered codon shapes a biased T cell repertoire through direct MHC and CDR3 β interactions

Kristin Støen Gunnarsen,^{1,2} Lene Støkken Høydahl,^{1,2} Louise Fremgaard Risnes,¹ Shiva Dahal-Koirala,¹ Ralf Stefan Neumann,¹ Elin Bergseng,¹ Terje Frigstad,³ Rahel Frick,^{1,2} M. Fleur du Pré,¹ Bjørn Dalhus,^{4,5} Knut E.A. Lundin,^{1,6,7} Shuo-Wang Qiao,^{1,6} Ludvig M. Sollid,^{1,6} Inger Sandlie,^{1,2} and Geir Åge Løset^{1,2,3}

¹Centre for Immune Regulation and Department of Immunology, University of Oslo and Oslo University Hospital-Rikshospitalet, Oslo, Norway. ²Department of Biosciences, University of Oslo, Oslo, Norway. ³Nextera AS, Oslo, Norway. ⁴Department of Microbiology, Oslo University Hospital-Rikshospitalet, Oslo, Norway. ⁵Department of Medical Biochemistry, Institute for Clinical Medicine, University of Oslo, Oslo, Norway. ⁶KG Jebsen Coeliac Disease Research Centre and Department of Immunology, University of Oslo, Oslo, Norway. ⁷Department of Gastroenterology, Oslo University Hospital-Rikshospitalet, Oslo, Norway.

Selection of biased T cell receptor (TCR) repertoires across individuals is seen in both infectious diseases and autoimmunity, but the underlying molecular basis leading to these shared repertoires remains unclear. Celiac disease (CD) occurs primarily in HLA-DQ2.5⁺ individuals and is characterized by a CD4⁺ T cell response against gluten epitopes dominated by DQ2.5-glia- α 1a and DQ2.5-glia- α 2. The DQ2.5-glia- α 2 response recruits a highly biased TCR repertoire composed of TRAV26-1 paired with TRBV7-2 harboring a semipublic CDR3 β loop. We aimed to unravel the molecular basis for this signature. By variable gene segment exchange, directed mutagenesis, and cellular T cell activation studies, we found that TRBV7-3 can substitute for TRBV7-2, as both can contain the canonical CDR3 β loop. Furthermore, we identified a pivotal germline-encoded MHC recognition motif centered on framework residue Y40 in TRAV26-1 engaging both DQB1*02 and the canonical CDR3 β . This allowed prediction of expanded DQ2.5-glia- α 2-reactive TCR repertoires, which were confirmed by single-cell sorting and TCR sequencing from CD patient samples. Our data refine our understanding of how HLA-dependent biased TCR repertoires are selected in the periphery due to germline-encoded residues.

Introduction

The central role of the T cell receptor (TCR) in initiation and orchestration of adaptive immune responses underlies a long-standing focus on understanding the rules governing TCR-MHC recognition. Structural studies have revealed that the TCR engages peptide-MHC (pMHC) in a conserved binding topology where the germline-encoded CDR1 and CDR2 loops generally contact the MHC helices and the somatic, flexible CDR3 loops primarily contact the peptide (1). Moreover, biased and public TCR responses have been reported for infections, and malignant and autoimmune disorders (2). In general, the molecular basis for selection of biased TCR repertoires is not well understood, but has been suggested to involve a combination of preferences in the somatic rearrangement machinery (2) and germline-encoded features that influence thymic selection (3). A recent study showed a strong *trans* association between variation in the MHC locus and TCR variable (V) gene usage, further suggesting a codependent expression pattern (4). However, it is still unclear if and how certain TCR V genes are predisposed to interact with MHC and underlie the existence of biased repertoires.

To address these questions, we studied celiac disease (CD), a human T cell-mediated chronic inflammatory disorder with autoimmune features affecting the small intestine and which recruits biased TCR repertoires. CD is strongly HLA dependent, where 95% of CD patients express the HLA class II molecule HLA-DQ2.5 (*DQA1*05/DQB1*02*), while the remaining express HLA-DQ8 (*DQA1*03/DQB1*03:02*) or HLA-DQ2.2 (*DQA1*02:01/DQB1*02*) (5). There are 2 immunodominant HLA-DQ2.5-restricted T cell

Authorship note: K.S. Gunnarsen and L.S. Høydahl contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: May 17, 2017

Accepted: August 3, 2017

Published: September 7, 2017

Reference information:

JCI Insight. 2017;2(17):e95193.

<https://doi.org/10.1172/jci.insight.95193>.

insight.95193.

epitopes derived from wheat α -gliadin, DQ2.5-glia- α 1a (PFPQPELPY) and DQ2.5-glia- α 2 (PQPELPYPQ), both found within a naturally processed 33mer peptide fragment (6, 7). In addition, responses to DQ2.5-glia- ω 1 and DQ2.5-glia- ω 2 derived from ω -gliadin are commonly observed (8). Notably, efficient T cell recognition of the gluten-derived peptides requires posttranslational deamidation (glutamine to glutamate) by the enzyme transglutaminase 2 (9). In particular, the DQ2.5:DQ2.5-glia- α 2 response recruits a dominant and semipublic TCR repertoire characterized by *TRAV26-1* and *TRBV7-2*, coupled with a canonical CDR3 β loop (ASSxRxTDTQY) (10, 11). Central in this loop is a conserved non-germline-encoded R in position 5 (p5R) that makes important contacts with DQ2.5-glia- α 2 (12). Although *TRAV4* combined with *TRBV20-1* or *TRBV29-1* appears to be overrepresented in the DQ2.5:DQ2.5-glia- α 1a response, *TRBV7-2* is also used in some patients (10, 12).

To gain insight into the molecular details characterizing the dominant T cell responses in CD and the selection of biased repertoires, we first analyzed TCR fine-specificity. In the case of the signature DQ2.5-glia- α 2 specificity, we extended the molecular characterization to map the contribution of germline-encoded residues. We identified a single residue, Y40 of *TRAV26-1*, to be of critical importance for T cell activation. We propose that Y40, together with Y38 and H55 of *TRAV26-1*, influences TCR chain pairing and positioning of the canonical CDR3 β loop to interact with the peptide. Furthermore, this trio makes crucial contacts with the HLA-DQ2 molecule. Importantly, from our in vitro analyses we could prospectively identify *TRAV* and *TRBV* usage in the TCR repertoire. We show that germline-encoded TCR-MHC interactions underpin the recruited chronic and pathogenic TCR repertoire in CD, alluding to why TCR-biased germline repertoires are so frequently seen in HLA-associated diseases.

Results

Generation of α -gliadin-specific T cell clones and TCR cloning. The dominating HLA-DQ2.5-restricted CD4⁺ T cell response in CD patients is directed towards a proteolytically resistant α -gliadin 33mer peptide harboring the 3 physically coupled epitopes, DQ2.5-glia- α 1a, DQ2.5-glia- α 1b, and DQ2.5-glia- α 2. In particular, the response towards DQ2.5-glia- α 2 exhibits a prominent TCR V gene bias with public features (11). To better understand the underlying molecular details governing this response, we here studied 2 representative T cell clones (TCCs) derived from gut biopsies of 2 CD patients. The DQ2.5:DQ2.5-glia- α 2-specific TCC364 uses the signature V gene pair *TRAV26-1/TRBV7-2* in combination with the canonical CDR3 β and has been described previously (10). The DQ2.5:DQ2.5-glia- α 1a-specific TCR of TCC380 was cloned in this study and uses *TRAV9-2/TRBV7-2*. Upon restimulation, both TCCs responded specifically to their cognate minimal epitopes presented as exogenously loaded peptides on HLA-DQ2.5⁺ EBV-B cells, with TCC380 responding at a lower peptide concentration compared with TCC364 (Figure 1, A and B). Notably, the proliferative response of TCC364 increased markedly to the epitope in the form of the 33mer peptide, underscoring the potency of this naturally occurring and highly immunogenic multivalent peptide (Figure 1B) (6).

The soluble TCRs show binding characteristics previously seen for microbially derived TCRs. To elucidate the intrinsic pMHC binding properties of the TCRs 364 and 380, we first expressed and purified both TCRs in a single-chain TCR (scTCR) format containing stabilizing mutations (P50^{TRAV}, L50^{TRBV}, F103^{TRBV}, and P2^{linker}) (13, 14). We then performed surface plasmon resonance (SPR) studies employing recombinant, soluble TCRs and pMHCs. scTCR s380 bound specifically to DQ2.5:DQ2.5-glia- α 1a with an average K_D of 17 μ M (Figure 1C, Supplemental Figure 1, and Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.95193DS1>). Similarly, scTCR s364 bound specifically to DQ2.5:DQ2.5-glia- α 2 with an average affinity of 38 μ M (Figure 1D, Supplemental Figure 1, and Supplemental Table 1). Neither scTCR s380 nor s364 reacted with DQ2.5:CLIP2 (Figure 1, C and D). To gain insight into the TCR fine-specificity, we extended the SPR binding experiments to a panel of deamidated gliadin-pMHCs: HLA-DQ2.5 in complex with DQ2.5-glia- α 1a, DQ2.5-glia- α 2, DQ2.5-glia- ω 1, and DQ2.5-glia- ω 2 (Figure 1, E and F). While scTCR s380 showed a weak, concentration-dependent binding to the 3 gluten-pMHC complexes tested, scTCR s364 bound weakly to the highly similar DQ2.5-glia- ω 2 only (Figure 1, G and H). Thus, the affinity of both TCRs are in the 1–100 μ M range typically observed for microbially derived peptides (15, 16), as well as for other gluten-pMHC complexes (12, 17, 18), and reflected the EC50 values obtained in the T cell activation assays using the originating TCCs (Figure 1, A and B).

3D interactions do not predict the functional T cell reactivity. To investigate how the intrinsic binding characteristics of soluble molecules translate to the same TCR's ability to function as membrane-

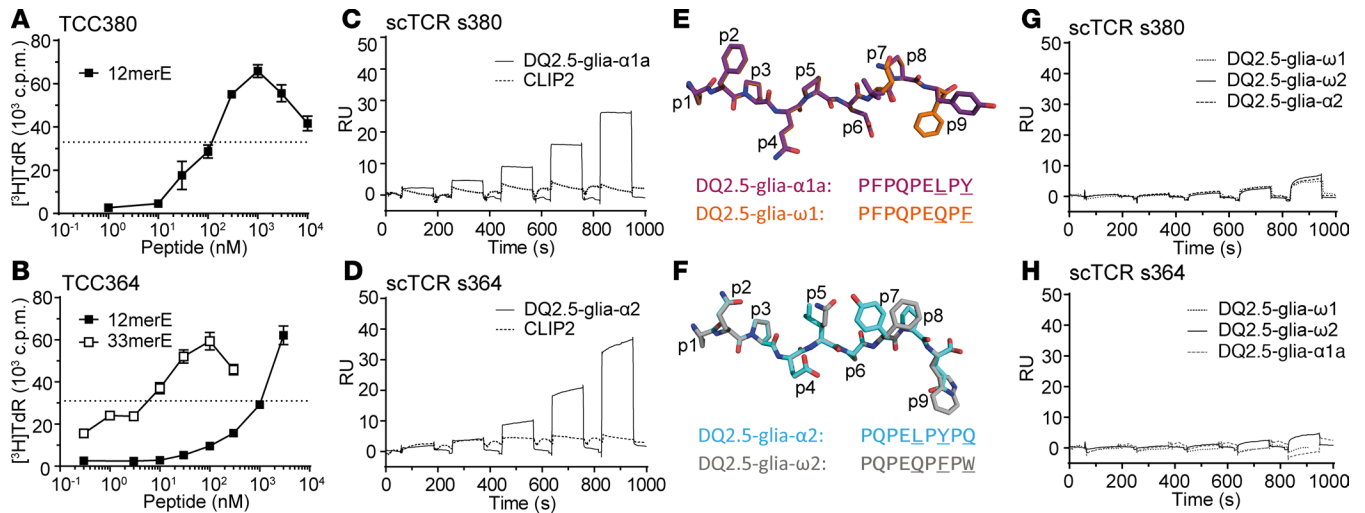


Figure 1. Mapping affinity and fine-specificity of T cell receptors (TCRs) 380 and 364. (A and B) Responsiveness of (A) TCC380 and (B) TCC364 to deamidated (E = glutamate) 12mer DQ2.5-glia- α 1a and DQ2.5-glia- α 2 peptides, respectively, presented by HLA-DQ2.5⁺ EBV-B cells was assessed by [³H]thymidine incorporation. For TCC364 the responsiveness to deamidated 33mer peptide was also assessed. The dotted line indicates half-maximum response to 12mer peptides and error bars indicate \pm SEM of triplicates. (C and D) Representative sensorgrams of (C) scTCR s380 binding to DQ2.5:DQ2.5-glia- α 1a ($n = 3$) and DQ2.5:CLIP2 ($n = 3$), and (D) binding of scTCR s364 to DQ2.5:DQ2.5-glia- α 2 ($n = 4$) and DQ2.5:CLIP2 ($n = 2$). n indicates the number of independent experiments performed. (E and F) Peptide overlays of (E) DQ2.5-glia- α 1a and DQ2.5-glia- ω 1, and (F) of DQ2.5-glia- α 2 and DQ2.5-glia- ω 2. Peptide sequences are indicated and the residues differing are underlined. Based on the crystal structures of DQ2.5:DQ2.5-glia- α 1a (PDB: 159V) and DQ2.5:DQ2.5-glia- α 2 (PDB: 40ZH) (12, 42). (G and H) Binding of scTCRs s380 (G) and s364 (H) to HLA-DQ2.5-gliadin complexes was assessed as indicated. The sensorgrams are representative of $n = 1$ –3 independent experiments.

bound receptors, we performed complementary cellular assays using HLA-DQ2.5⁺ EBV-B cells loaded with 12mer peptides as antigen-presenting cells (APCs) and retrovirally generated BW 380 and 364 T cells (Supplemental Figure 2, A and B). The BW 380 T cell was stimulated by both deamidated DQ2.5-glia- α 1a and DQ2.5-glia- ω 1 peptides, but not with DQ2.5-glia- α 2 or DQ2.5-glia- ω 2 peptides or native counterparts (Figure 2A). Thus, despite the specific, yet weak binding across epitopes observed in SPR, only the response against the highly similar DQ2.5-glia- ω 1 translated into functional activation of the BW 380 T cell. In contrast, the BW 364 T cell exclusively responded to deamidated DQ2.5-glia- α 2, even though a weak binding to DQ2.5-glia- ω 2 was observed in SPR (Figure 2B). The same was seen when using the longer, naturally occurring 33mer α -gliadin and ω -gliadin 17/19mer peptides (Figure 2, C and D). Both native and deamidated 33-mer α -gliadin peptide stimulated the T cells, but clearly favored the deamidated version.

To further study TCR fine-specificity, we engineered murine A20 B cells to express equal amounts of HLA-DQ2.5 harboring covalently attached peptide, thereby largely bypassing effects of differences in peptide off-rates (Supplemental Figure 2C). We constructed WT peptides, as well as versions of DQ2.5-glia- α 1a and DQ2.5-glia- α 2 to resemble the highly similar DQ2.5-glia- ω 1 and DQ2.5-glia- ω 2 epitopes. When assessed for their ability to activate the BW 380 T cells, DQ2.5-glia- α 1a pL7Q showed greatly reduced potency, while mutation DQ2.5-glia- α 1a pY9F slightly increased responses compared with the WT peptide (Figure 2E). Similarly, pQ9W mutation of DQ2.5-glia- α 2 increased the response of the BW 364 T cells, while mutations pL5Q and pP7F completely abolished T cell activation (Figure 2F). Neither the BW 380 nor the BW 364 T cells responded to CLIP2 (Figure 2, E and F). In summary, we observed a discrepancy between the binding characteristics using soluble molecules in SPR and functional T cell activation, illustrating the shortcomings of 3D interaction studies in predicting physiological TCR reactivity, as also observed by others (19, 20).

TRBV7-3 can substitute for TRBV7-2, while TRAV26-1 cannot be exchanged with its closest homolog TRAV26-2. While the DQ2.5:DQ2.5-glia- α 2 TCR repertoire is characterized by a highly biased *TRAV* and *TRBV* gene segment usage (*TRAV26-1/TRBV7-2*), such a bias is much less prominent in the DQ2.5:DQ2.5-glia- α 1a-reactive repertoire. To study the contribution of germline-encoded residues to pMHC recognition, we chose to focus on the DQ2.5:DQ2.5-glia- α 2 response and we used TCR 364 as a representative clone. We were intrigued by the fact that neither *TRAV26-2* nor *TRBV7-3*, the closest homologs of *TRAV26-1* and *TRBV7-2*, respectively,

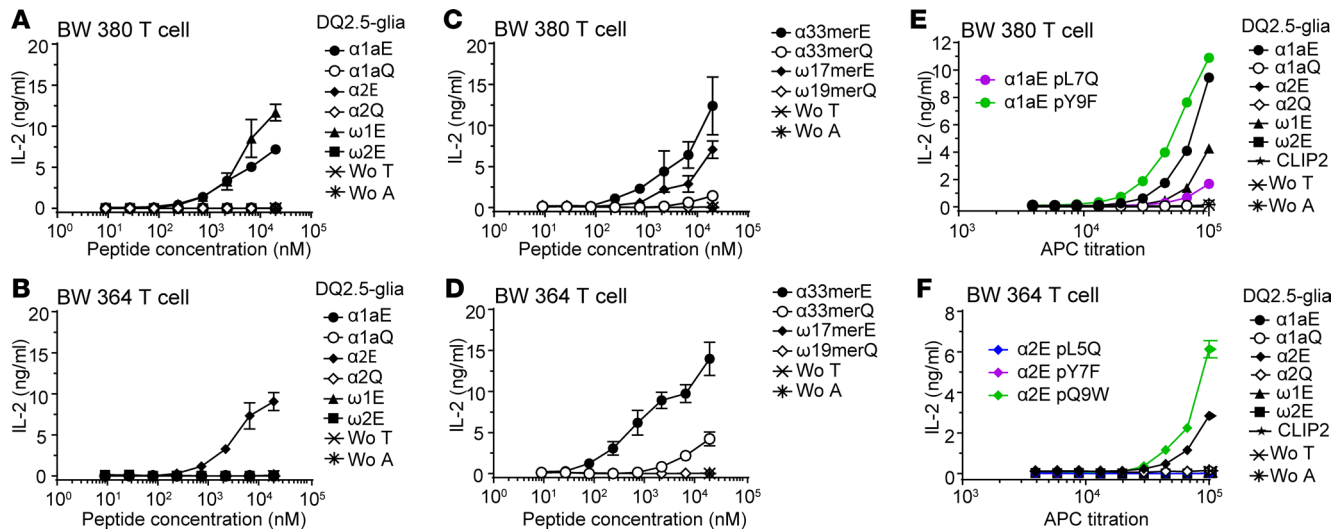


Figure 2. Responsiveness of T cell receptor (TCR) 380 and TCR 364 as reconstructed membrane-bound receptors. The ability of BW T cells retrovirally transduced with either TCR 380 or TCR 364 to respond to a panel of gluten peptides was monitored by measuring IL-2 secretion. (A and B) HLA-DQ2.5⁺ EBV-B cells were loaded with titrated amounts of 12mer peptides as indicated before addition of BW T cells. Both deamidated (E) and native (Q) peptides were used. Representative responses of (A) BW 380 T cells and (B) BW 364 T cells are shown. (C and D) HLA-DQ2.5⁺ EBV-B cells were loaded with titrated amounts of 33mer α -gliadin peptide (E or Q) containing both the DQ2.5-glia- α 1a and DQ2.5-glia- α 2 epitopes, or with 17mer/19mer ω -gliadin peptide (E and Q, respectively) containing both the DQ2.5-glia- ω 1 and DQ2.5-glia- ω 2 epitopes. Representative IL-2 responses of (C) BW 380 T cells and (D) BW 364 T cells are shown. (E and F) Titrated amounts of A20 B cells expressing HLA-DQ2.5 with covalently linked peptide were used as antigen-presenting cells (APCs) to stimulate (E) BW 380 T cells and (F) BW 364 T cells. Wo = without T and A are controls with T cells or APCs with peptide alone. Error bars indicate \pm SD of triplicates. (A–F) The experiment in each panel was performed 3 independent times. Figures were prepared using GraphPad Prism 7, and nonlinear regression analysis (3 parameters) was used to derive IL-2 concentrations from the standard curves.

had appeared in the sequencing of the DQ2.5:DQ2.5-glia- α 2 TCR repertoire (10–12, 21). To understand this, we generated TCR variants where either *TRAV26-1* or *TRBV7-2* were exchanged with their closest homologs, while keeping the CDR3 α and CDR3 β as well as the J gene segments of TCR 364 unchanged (Table 1). We then introduced these modified TCRs into BW T cells (Supplemental Figure 3, A and B). When assessed for their ability to respond to the 33mer peptide, we observed a small decrease in reactivity when *TRBV7-2* was exchanged with *TRBV7-3* (Figure 3, A and B). In contrast, the exchange of *TRAV26-1* with *TRAV26-2* completely abolished activation of the T cells (Figure 3, A and B). As the CDR3 loops of the WT and the TCR 364 mutants were identical, the observed difference in DQ2.5:DQ2.5-glia- α 2 reactivity most likely arose from differences in germline-encoded V gene residues.

A single germline-encoded TRAV residue is a crucial determinant for T cell responsiveness. We then constructed and tested a panel of TCR variants (TCR1–13) where *TRAV26-1* or *TRBV7-2* residues were exchanged with the corresponding *TRAV26-2* or *TRBV7-3* residues (all mutated residues are specified in Table 1, and Figure 4, A and B, and clone validation shown in Supplemental Figure 4, A and B). Mutations within *TRAV26-1* resulted in more profound effects on T cell activation (Figure 4C and Supplemental Figure 4C) than the mutations

Table 1. Sequence alignment of TRAV26-1 and TRBV7-2 with their closest homologs

V-gene ^A	FR1	CDR1	FR2	CDR2	FR3
IMGT no.	1-----→26	27-----→38	39-----→55	56-----→65	66-----→104
TRAV26-1	DAKTTQ.PP SMDCAE GRAAN LPCNHS	TISG..... NEY	VY WYRQ I H S QGPQ Y I H	GLK..... NN	ET NE...MASL I ITEDR KSSTL L PH A T L RD T AVV Y CC
TRAV26-2 ^B	DAKTTQ.P NSMESNE EEPVH LPCNHS	TISG..... TDY	I H W YRQ L P S QGP E V I H	GL T SN	V NR...MASL A I A EDR KSSTL L HR A T L RD A AVV Y CC
TRBV7-2	GAGV S Q S PSNK V TEK G KD VELRCD P I	SGH..... TA	LYWYRQ S L G Q G L EFL I Y	FQ G NSA	P D K S G LP S D R F S A ERT.G G SV S T L I Q RT Q Q ED S AVV L CC
TRBV7-3 ^B	GAGV S Q T PSNK V TEK G KY VELRCD P I	SGH..... TA	LYWYRQ S L G Q G P EFL I Y	FQ G TGA	A DD S GL P ND R F FA V RP.E G SV S T L I Q RT E R T GD S AVV L CC

^ATRAV gene segment usage and numbering was defined by the IMGT Database. ^BClosest homolog gene fragments were identified using the IMGT database. TRAV26-2*01 and TRBV7-3*01 were chosen. Amino acids differing between TRAV26-1/TRAV26-2 and TRBV7-2/TRBV7-3, which were mutated in TCR 364, are illustrated with colors in the TRAV26-1/TRBV7-2 sequences corresponding to the different mutants in Figure 4 and in bold and underlined in the TRAV26-2/TRBV7-3 sequences.

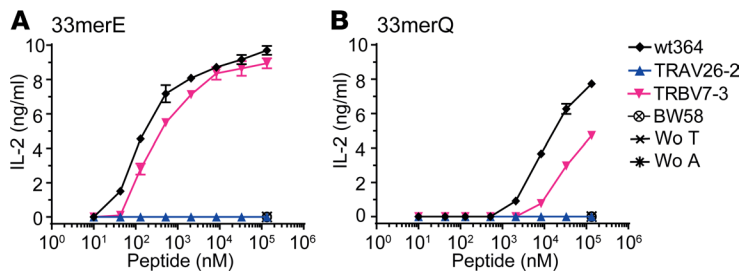


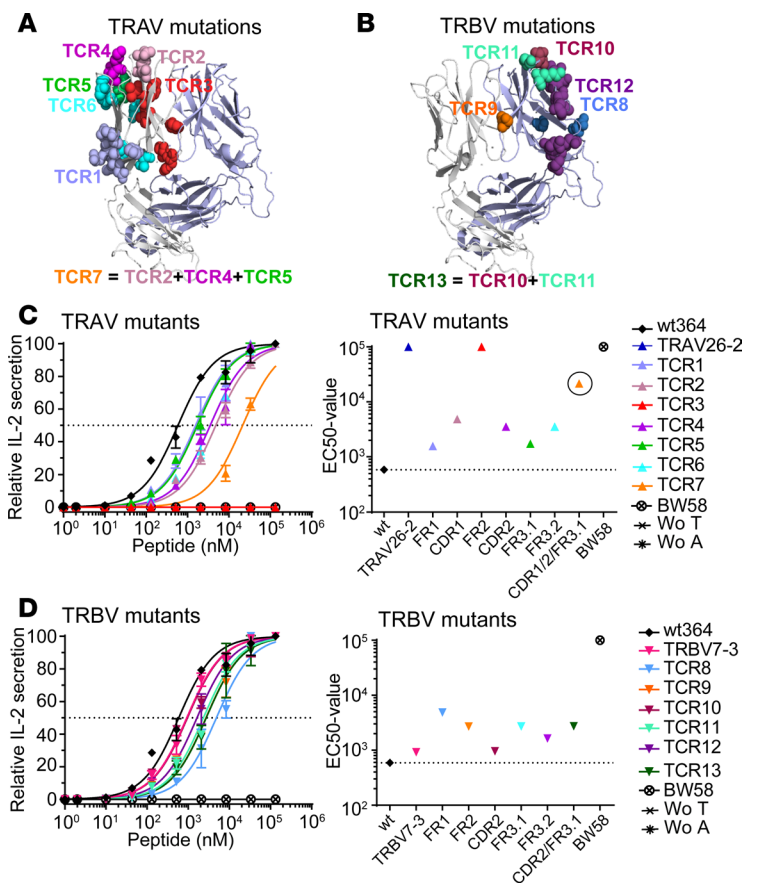
Figure 3. The effect of TRAV26-1 and TRBV7-2 germline segments on T cell responsiveness. (A and B) BW T cells were engineered to express T cell receptor (TCR) 364 variants where the entire TRAV26-1 or TRBV7-2 segments were exchanged with their closest homologs, TRAV26-2 and TRBV7-3. The effect of the germline-encoded residues on T cell responsiveness was assessed by measuring IL-2 secretion by ELISA. HLA-DQ2.5* EBV-B cells were loaded with titrated amounts of (A) deamidated 33merE peptide, and (B) native 33merQ peptide before addition of T cells. Four independent experiments were conducted and a representative experiment is shown. Untransduced BW T cells (BW58), and T cells (Wo A) or antigen-presenting cells (Wo T) with peptide alone were included as controls. Error bars indicate \pm SD of triplicates. Figures were prepared using GraphPad Prism 7, and nonlinear regression analysis (3 parameters) was used to derive IL-2 concentrations from the standard curves.

in TRBV7-2 (Figure 4D and Supplemental Figure 4C). TCR3 that contained 6 framework (FR) 2 mutations was completely nonresponsive (Figure 4C). The second most pronounced effect was seen in TCR7, where residues of CDR1, CDR2, and FR3 were exchanged for those of TRAV26-2 (Figure 4C). To dissect the individual effects of these regions, they were altered separately in TCR2, TCR4, and TCR5, respectively. Exchange of N36T^{TRAV} and E37D^{TRAV} within the CDR1 α of TCR2 led to a modest decrease in T cell responsiveness, as did K58T^{TRAV} and N64S^{TRAV} within the CDR3 α of TCR4. The FR3 α mutations of TCR5 had little impact. For the TRBV variants, only small changes in EC50 values relative to WT364 were observed, as could be expected from the exchange of TRAV7-2 with TRBV7-3 (Figure 3 and Figure 4D).

To map the structural basis for the observed effects of the mutations, we inspected the cocrystal structure of TCR S16 using TRAV26-1/TRBV7-2 and the canonical CDR3 β bound to DQ2.5:DQ2.5-glia- α 2 (12). Of the 6 FR2 residues mutated in TRAV26-1, residue Y40^{TRAV} participates in an extensive network of polar interactions and directly interacts with both R70^{MHC β} and T115^{TRBV} of the canonical CDR3 β .

Both of these interactions are lost by the Y40H mutation in TCR3 (Figure 5A). Additional indirect effects on the H-bond network upon mutation, including the R70^{MHC β} -T115^{TRBV} interaction cannot be excluded. The reduced responsiveness of TCR2 and TCR4 could both be due to loss of direct interactions; a H-bond between N36^{TRAV} of CDR1 α and p2Q is lost by the N36T mutation of TCR2 (Figure 5B), and the K58T^{TRAV} mutation within CDR2 α of TCR4 abrogates the salt bridge between K58^{TRAV} and D76^{MHC α} (Figure 5C).

Figure 4. Fine-mapping the effect of germline-encoded TRAV26-1 and TRBV7-2 residues on T cell responsiveness. BW T cells engineered to express versions of TRAV26-1 and TRBV7-2 where either framework (FR) or germline CDR loops were mutated to the corresponding residues of TRAV26-2 and TRBV7-3 (exchanged residues are specified in Table 1) were evaluated for effect on T cell responsiveness. (A and B) Overview of the panel of T cell receptor (TCR) mutants (TCR 1–13) and the residues mutated in (A) TRAV26-1 and (B) TRBV7-2. TCR7 and TCR13 are combination mutants as indicated. TCR α -chain, light gray; TCR β -chain, pale blue. The crystal structure of TCR S16 containing TRAV26-1/TRBV7-2 was used (PDB: 4OZH) (12). (C and D) The left panels show the responsiveness of (C) TRAV and (D) TRBV variants to HLA-DQ2.5* EBV-B cells loaded with titrated amounts of 33merE peptide given as relative IL-2 secretion. Dotted line indicates half-maximum response. The right panels show the EC50 values derived from the responses to the 33merE peptide of the (C) TRAV and (D) TRBV variants. The EC50 value of TCR7 (marked with a circle) could not be reliably determined, as saturation was not reached. The dotted line indicates the EC50 value of WT364. Four independent experiments were conducted and a representative experiment is shown. Untransduced BW T cells and T cells or antigen-presenting cells (APCs) with peptide alone were included as controls. Error bars indicate \pm SD of triplicates. Figures were prepared using GraphPad Prism 7. Nonlinear regression analysis (3 parameters) was used to derive IL-2 concentrations from the standard curves. Graphs were normalized and EC50 values calculated using nonlinear regression analysis (log[agonist] vs. normalized response).



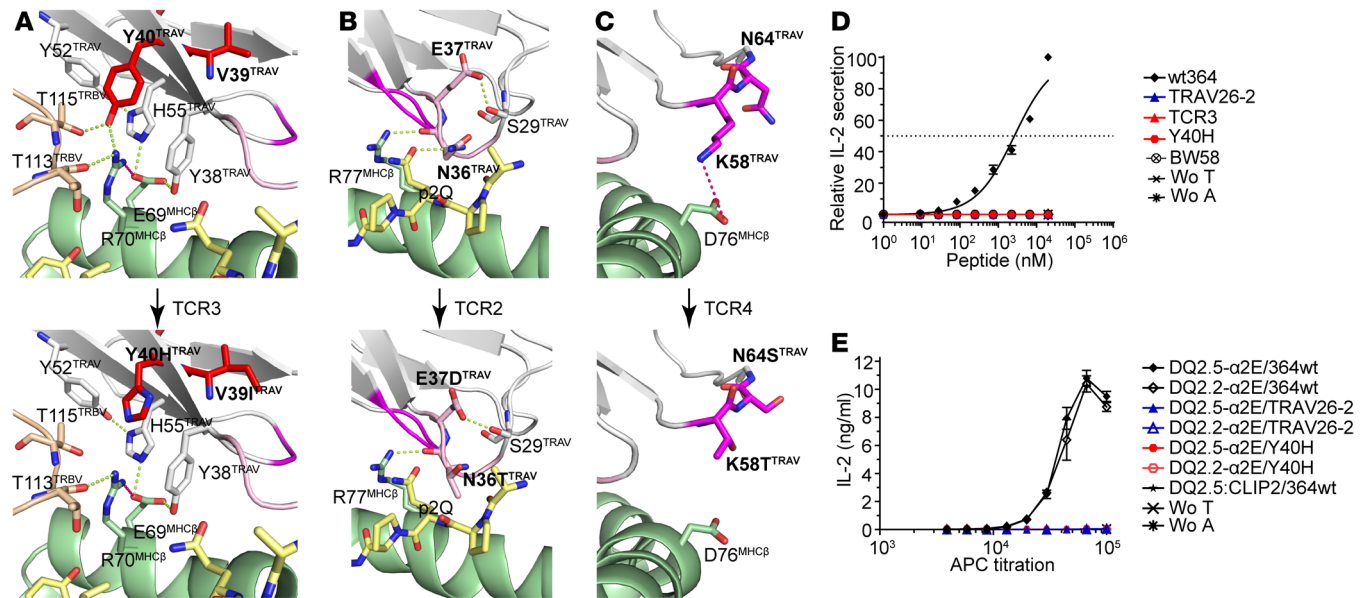


Figure 5. A structural basis for the reduced responsiveness of the TRAV26-1/TRAV26-2 T cell receptor (TCR) variants. TCR α -chain, light gray; MHC β -chain, pale green; DQ2.5-glia- α 2 peptide, yellow; canonical CDR3 β , wheat. Residues of TCR3 in red; TCR2 in light pink; TCR4 in magenta. All residues mutated in the TCRs are specified in Table 1. H-bonds, green dashes; salt bridges, pink dashes. Based on PDB: 4OZH (14). (A) Of the variant FR2 residues, only Y40^{TRAV} interacts directly with pMHC (R70^{MHC β}). Y40H of TCR3 partly disrupts the polar interactions network between TRAV, HLA-DQ2.5 β , and CDR3 β . Residue 39^{TRAV} is close to 40^{TRAV} and is illustrated. (B) The direct interaction of N36 with p2Q of DQ2.5-glia- α 2 is abrogated by the N36T mutation of TCR2. (C) The K58T alteration in TCR7 disrupts the salt bridge formed between K58 of TRAV26-1 and D76 of MHC α . (D) The effect of the single Y40H mutation on T cell responsiveness was assessed using HLA-DQ2.5^{*} EBV-B cells loaded with 33merE peptide. Dotted line indicates half-maximum response. (E) Titrated amounts of A20 B cells expressing HLA-DQ2.5 or HLA-DQ2.2 with covalently linked peptide were used as antigen-presenting cells (APCs) to stimulate BW T cells transduced with either 364WT, the TRAV26-2 exchange variant, or the 364 Y40H variant as indicated. (D and E) Error bars indicate \pm 5D of triplicates and the experiments were conducted 2 independent times. Figures were prepared using GraphPad Prism 7. Nonlinear regression analysis (3 parameters) was used to derive IL-2 concentrations from the standard curves.

Finally, to validate if the effect solely resided in Y40^{TRAV}, we constructed the single Y40H mutant. In line with predictions, the BW 364 Y40H mutant was completely nonresponsive to pMHC (Figure 5D). Taken together, these results show that the germline sequence of TRBV7-2 does not contain features explaining its selection over TRBV7-3. In contrast, the single residue, Y40^{TRAV}, which differs between TRAV26-1 and TRAV26-2, is crucial for dictating T cell responsiveness.

TRAV26-1/TRBV7-3 is used in the in vivo repertoire, whereas TRAV26-2 is not. Our data clearly show that TRBV7-2 and TRBV7-3 are interchangeable. Thus, we hypothesized that DQ2.5:DQ2.5-glia- α 2-reactive T cells using TRBV7-3 with the canonical CDR3 β paired with TRAV26-1, should be present in the repertoire of HLA-DQ2.5^{*} CD patients, and that TRAV26-2 should not be found paired with such TRBVs harboring the canonical CDR3 β .

To this end, we tetramer-sorted CD4⁺ T cells from blood of 9 gluten-challenged CD patients followed by single-cell coupled sequencing of TCR and TCR β genes (Louise Fremgaard Risnes, unpublished observations). Notably, these cells were sorted using a pool of HLA-DQ2.5:gluten tetramers consisting of 4 epitopes. After filtering the sequence data for TRAV26-1 paired with the canonical CDR3 β found in DQ2.5:DQ2.5-glia- α 2-reactive T cells only, we found that TRBV7-3 was in fact used, but 5 times less frequently than TRBV7-2 (Table 2). Interestingly, this difference corresponds well with the relative expression of TRBV7-2 and TRBV7-3 as reported in the naive repertoire (22, 23). We also looked for TRAV26-2 and found it in 6 clonotypes (of 536 clonotypes in the total pool). Importantly, we did not observe pairing with TRBV7-2 or TRBV7-3 (Table 2).

We further observed that a substantial fraction of the T cells pairing TRAV26-1 with TRBV7-2 used a CDR3 β highly similar to the canonical semipublic sequence previously reported, also containing a p5R, emphasizing the importance of this CDR3 β feature in DQ2.5:DQ2.5-glia- α 2 recognition (Table 2) (10, 11). Further, of the 9 clonotypes using TRBV7-3, 8 had the canonical CDR3 β (Table 2). Both the canonical CDR3 β and the highly similar loop were formed by recombination of TRBV7-2 or TRBV7-3 with TRBJ2-3. Taken together, these results show that analysis of TCR V gene usage in CD patients sup-

Table 2. TRAV and TRBV7 pairing in CD patients

Chain pairing	No. of clonotypes
Total TRAV26-1/TRBVx	82
TRAV26-1/TRBV7-2 ^A	52
TRAV26-1/TRBV7-2 TRBJ2-3 (Can CDR3 β) ^B	39
TRAV26-1/TRBV7-2 TRBJ2-3 (P5R) ^C	6
TRAV26-1/TRBV7-3 ^A	9
TRAV26-1/TRBV7-3 TRBJ2-3 (Can CDR3 β) ^B	8
TRAV26-1/TRBV7-3 TRBJ2-3 (P5R) ^C	0
Total TRAV26-2/TRBVx	6
TRAV26-2/TRBV7-2	0
TRAV26-2/TRBV7-3	0
Total TRAVx/TRBV7-2/TRBV7-3 TRBJ2-3 (Can CDR3β)^B	7
TRAV5/TRBV7-2	1
TRAV8-6/TRBV7-2	1
TRAV12-3/TRBV7-2	1
TRAV13-2/TRBV7-2	1
TRAV29/DV5/TRBV7-2	1
TRAV38-1/TRBV7-2	1
TRAV39/TRBV7-3	1

Sequencing of CD4⁺ T cells from blood of 9 HLA-DQ2.5⁺ CD patients after gluten challenge identified 536 unique clonotypes. ^ATRAV26-1/TRBV7-2 or TRAV26-2/TRBV7-3 irrespective of CDR3 sequence. ^BSequences from ^A were filtered on the canonical (can) CDR3 β that is restricted to DQ2.5:DQ2.5-glia- α 2-reactive T cells. ^CSequences from ^A were filtered on the presence of a central P5R of CDR3 β , in the absence of the canonical CDR3 β . ^DClonotypes using TRBV7-2/7-3 with the canonical CDR3 β with other TRAVs than TRAV26-1.

ports the hypothesis, based on findings from the in vitro cellular assays, that TRAV26-1 is crucial for clonal expansion of the semipublic repertoire in the DQ2.5:DQ2.5-glia- α 2 response, and that frequent TRBV7-2 pairing is due to both the TRBV7-2 frequency in the naive repertoire and its ability to form the canonical CDR3 β loop.

Y38, Y40, and H55 form a recognition motif for HLA-DQ2.

The observed impact of the single Y40^{TRAV} position on T cell responsiveness was striking. Y40^{TRAV} aligns its aromatic side chain to interact with the side chains of Y38^{TRAV} and H55^{TRAV}, forming a trio interaction motif, which together coordinates an extensive network of polar interactions between TRAV, CDR3 β , and the MHC β (Figure 5A) and appears to act as a germline-encoded recognition motif suited to interact with DQB1*02. Such imprints may represent the coevolution of TCR and pMHC (3). Three other TRAV gene segments contain a Y40^{TRAV} residue (of 47 TRAVs in total in the IMGT database): TRAV5, TRAV34, and TRAV39. We then investigated whether or not we could find either of these gene segments represented among the DQ2.5:DQ2.5-glia- α 2-reactive T cells using TRBV7-2/7-3 with the canonical CDR3 β . Indeed, 7 such clonotypes were found that paired with other TRAVs than TRAV26-1; one used TRAV5 and one used TRAV39, while the others used different gene segments (Table 2, Supplemental Tables 2 and 3). Interestingly, all except 1 of the 7 also contained the Y38^{TRAV} characteristic of the trio of amino acids in TRAV26-1 (Supplemental Table 2).

A critical dependency on Y40^{TRAV} for binding to HLA-DQ2.2, which is highly similar to HLA-DQ2.5, would strongly indicate that the trio of amino acids engages HLA-DQ2.2 in a similar manner as to HLA-DQ2.5. With respect to the DQ2.5-glia- α 2 epitope presentation, HLA-DQ2.2 and HLA-DQ2.5 have been shown to be functionally identical, with the exception of kinetically segregating the T cell response due to pMHC stability differences (24). To experimentally test this, we used cells expressing HLA-DQ2.2 with covalently linked peptide. Indeed, the Y40H mutation completely abrogated recognition of HLA-DQ2.2 with the DQ2.5-glia- α 2 peptide (Figure 5E). Taking these results together, we suggest that this trio of amino acids (Y38^{TRAV}, Y40^{TRAV}, and H55^{TRAV}) constitutes a recognition motif shared by multiple DQ2.5:DQ2.5-glia- α 2-reactive TCRs. Importantly, this MHC recognition motif is engaged by virtue of also involving the public CDR3 β , thereby influencing chain pairing.

Discussion

The existence of biased TCR repertoires with preferential usage of certain TRAV and TRBV gene segments has emerged as a common phenomenon in adaptive immunity and has been described in the T cell response against gluten epitopes presented on both HLA-DQ2.5 and HLA-DQ8 in CD (10–12, 21, 25). To broaden our understanding of the molecular basis for the selection of biased TCR repertoires, we undertook a combination of affinity, directed mutagenesis, and cellular activation studies, complemented with structural analysis.

As a starting point, we mapped the intrinsic binding properties of 2 prototypic, patient-derived TCRs against the dominant α -gliadin epitopes, DQ2.5-glia- α 1a and DQ2.5-glia- α 2. There was a surprising discrepancy between specificity at the soluble protein level and the corresponding functional T cell reactivity. While TCR 364 preferentially bound DQ2.5-glia- α 2 in SPR, a weak reactivity was observed towards DQ2.5-glia- ω 2. Still, the BW 364 T cells responded solely to DQ2.5-glia- α 2 in cellular assays. In contrast, the BW 380 T cells responded to both DQ2.5-glia- α 1a and DQ2.5-glia- ω 1, even though DQ2.5-glia- ω 1 was only 1 of a set of gluten peptides with weak binding to TCR 380 in SPR. Consistent with this, a substantial fraction of DQ2.5:DQ2.5-glia- α 1a-sorted TCCs from CD patients cross-react

with DQ2.5:DQ2.5-glia- ω 1 (Shiva Dahal-Koirala, unpublished observations). To our surprise, pL7Q and pY9F mutations of DQ2.5-glia- α 1a to resemble DQ2.5-glia- ω 1 resulted in greatly reduced and increased T cell activation, respectively. A previous alanine scan showed a complete dependence on p7, whereas all other positions resulted in clone-dependent effects, even though the TCRs presumably do not directly interact with p7L (12). We speculate that the presence of a hydrophobic amino acid in p9 anchors the peptide more efficiently to the p9 pocket. Thus, recognition of DQ2.5-glia- ω 1 by TCR 380 seems to be due to a compensatory mechanism. Similarly, mutations of p5 and p7 completely abrogated activation of the BW 364 T cells in support of previous observations (12), whereas mutation of p9 to a hydrophobic anchor residue increased responses, albeit not compensating for the negative effects of the p5 and p7 mutations. Based on our data, TCR 380 may have been selected by either of the DQ2.5-glia- α 1a or DQ2.5-glia- ω 1 epitopes in vivo.

TRAV26-1 paired with *TRBV7-2* is characteristic of the DQ2.5:DQ2.5-glia- α 2 response, and the closest homologs, *TRAV26-2* and *TRBV7-3*, had not been observed using biased or unbiased approaches (10–12, 21). A combination of structural and mutagenesis data had suggested that a germline imprint of *TRAV26-1* was important for recognition of HLA-DQ2.5, as mutation of Y38, Y40, and L57 in *TRAV26-1* markedly affected affinity in SPR (12). Our data clearly show that SPR is not suitable to predict functional outcome. Thus, we here took on a cellular approach to map germline contributions. Exchange of *TRAV26-1* with *TRAV26-2* abolished T cell activation in our studies, presumably through a combination of direct disruption of MHC (R70^{MHC β}) and TCR β interactions (T115^{TRBV} of the canonical CDR3 β) mediated by the FR2 residue Y40. Y38 and L57 are both present in *TRAV26-2*, and therefore cannot explain the preference for *TRAV26-1* over *TRAV26-2* (12). In contrast to the strong dependence on *TRAV26-1*, exchange of *TRBV7-2* with *TRBV7-3* only weakly affected pMHC recognition. Dissection of the contribution of residues differing between *TRBV7-2/7-3* revealed only minor effects, in line with the small molecular imprint of *TRBV7-2* on MHC, where P66^{TRBV} is the only germline-encoded residue to interact with MHC through 2 H-bonds (12).

Based on these findings, we hypothesized that large-scale single-cell sequencing of the DQ2.5:DQ2.5-glia- α 2 TCR repertoire would reveal the presence of T cells using *TRBV7-3* and that *TRAV26-2* would be absent. We did indeed find *TRBV7-3* to be present, and the *TRAV26-1/TRBV7-3* T cells formed the canonical CDR3 β . Furthermore, the difference in the representation of *TRBV7-2* versus *TRBV7-3* mirrored the abundance of these gene segments in the naive repertoire (22, 23). Whether this is a general feature of antigen-expanded T cell populations remains unclear, but a previous study found that antigen-driven expansion of CD4⁺ T cells was reflected by the frequency of naive progenitors (26). Moreover, we found *TRAV26-2* to pair with neither *TRBV7-2* nor *TRBV7-3*. Interestingly, use of *TRBV7-2/7-3* and the canonical CDR3 β is almost always coupled with *TRAV26-1*, whereas use of *TRBV7-2* without the canonical CDR3 β show no pairing bias with *TRAV26-1* or any other particular *TRAVs* (11). This indicates that the preferential pairing is not due to features in the V $_{\alpha}$ /V $_{\beta}$ interphase, but rather influenced by the presence of the canonical CDR3 β , consistent with the direct CDR3 β T115 interaction with Y40 of *TRAV26-1* and the H-bond network further extending to T113^{TRBV} and R70^{MHC β} . We propose that the crucial interactions *TRAV26-1* Y40 makes with HLA-DQ2 and the canonical CDR3 β is the driver for the selection of this biased repertoire.

Recently, an association was found between expression of *TRAV26-1* and *DQB1*02:01* (β -chain of HLA-DQ2.5), whereas this was not the case with *TRAV26-2* (4). We therefore hypothesize that the trio of amino acids, Y38, Y40, and H55 of *TRAV26-1*, constitute a germline-encoded recognition motif with specificity towards HLA-DQ2 (3). This same notion has been extensively investigated for TRBV-centric TCRs using V β 8.2 (*TRBV13-2*), where a trio of amino acids in the CDR1 β and CDR2 β were suggested to represent an evolutionary signature of germline specificity, forming close to the same contacts with distinct MHC alleles (IA^k, IA^u, and IA^b) irrespective of peptide and TRAV (27–31). As in our case, TRAV-centric germline codons have also been previously suggested for the V α 3-containing (*TRAV9D-3/TRAV9-4*) TCRs 2C and 42F3 specific for H2-L^d (32). It appears that amino acids in CDR1 and CDR2 can contribute to different MHC interaction modes influenced by the CDR3 regions (30). Our data further expand our understanding of the motifs involved and how the tuning of the recognition might occur, as we here report a TRAV FR residue (Y40) that coordinates the interaction network. The results explain the selection and expansion of the public T cell repertoire shared across patients. Of the 7 clonotypes we identified that harbored *TRBV7-2/7-3* with the canonical CDR3 β , but did not pair with *TRAV26-1*, two used *TRAV5* and *TRAV39* that both encode Y40. Y38 of the interaction trio is present in 6 out of 7 clonotypes, and it is not unlikely

Table 3. Oligonucleotides

TRAC_rv	5'-AGTCAGATTGTTGCTCCAGGCC-3'
TRBC_rv	5'-TTCACCCACCAGCTCAGCTCC-3'
poly-G-NotI_fw	5'-ATATGCGGCCGCGGGGGGGGGGGGGG-3'
TRAC_MluI_rv	5'-ATACCGTTCTCTCAGCTGGTACACGG-3'
TRBC_MluI_rv	5'-ATACCGTAGATCTCTGCTTCTGATGGC-3'
TRAV_NcoI_fw	5'-ATCCATGGCCGGAGATTCACTGACCCAGATGG-3'
TRAV_HindII_rv	5'-ATATATAAGCTTAGGTAACAGTCAATTGTGTCCC-3'
TRBV_MluI_fw	5'-ATATACGCGTAGGAGCTGGAGTCTCCAG-3'
TRBV_NotI_rv	5'-ATATGCGGCCGCTGTGACCGTGAGCCCTGG-3'
TRAV_L50P_fw	5'-AATATCCTGGAGAAGTCCGCAGCTCCTCTCAAAGCC-3'
TRAV_L50P_rv	5'-GGCTTTCAGGAGGAGCTGCGGACCTTCCAGGATATT-3'
Linker_L2P_fw	5'-TGACTGTTTTACTTAAGCCGTGAGGAGTGCATCCGCC-3'
Linker_L2P_rv	5'-GGCGGATGCACTCCCTGACGGCTTAGGTAACAGTCA-3'
TRBV_L103F_fw	5'-CAGGAGGACTCGGCCGTGATTTTTGTGCCAGCAG-3'
TRBV_L103F_rv	5'-CTGCTGGCACAAAATACACGGCCGAGTCTCTCTG-3'

that Y38 can rearrange and substitute for Y40 in these other *TRAVs*. Similarly, the DQ8:DQ8-glia- α 1 and DQ8:DQ8-glia- γ 1 responses select biased repertoires with R in CDR1 α or CDR3, which by structural rearrangement adjusts the TCR docking mode to position the R similarly (17).

We propose that the immunodominant DQ2.5-glia- α 2-reactive TCR repertoire in CD is positively selected in the thymus through conserved TRAV26-1 interactions with DQB1*02. A prerequisite may be a small TRBV footprint on DQA1 as well as ignorance to CLIP2 as seen for the prototypic TCR studied here. Since HLA-DQ2.5 is largely inert to HLA-DM-mediated peptide exchange, a high proportion of HLA-DQ2.5 molecules present CLIP peptides (24, 33–35). The consequence likely is a narrow peptide repertoire available for T cell interactions during thymic selection. Still, T cells using TRAV26-1 combined with TRBV7-2/TRBV7-3 encoding a p5R in CDR3 β escape negative selection, presumably rescued from

death by neglect by the critical Y40^{TRAV-R70}^{MHC β} interaction. Upon disease onset, the T cells initiate the inflammatory autoimmune response. Knowledge of the molecular basis for TCR-pMHC recognition opens the door for repertoire-based patient stratification and therapeutic intervention.

Methods

Isolation of T cells from CD patients and TCR cloning and expression. TCCs were established from intestinal biopsies of CD patients (CD364 and CD380) as previously described (36). Briefly, T cells were expanded with anti-CD3 and anti-CD28 beads (Invitrogen) for a minimum of 7 days prior to mRNA extraction and TCR cloning. mRNA was isolated from TCC380.E48 (herein denoted TCR 380) using an Absolutely RNA miniprep kit (Stratagene), followed by a seminested PCR approach. Briefly, first-strand cDNA synthesis was performed using gene-specific TRAC_rv and TRBC_rv primers (Table 3) and SuperScript II reverse transcriptase (Invitrogen). cDNA was precipitated using seeDNA (GE Healthcare) after RNase H digestion (NEB), followed by poly dCTP-tailing of the cDNA at the 3'-end with rTerminal transferase (Roche) and reprecipitation. The V α and V β were PCR amplified using Phusion HotStart DNA polymerase (Finnzymes) and the poly-G-NotI_fw primer and either the TRAC_MluI_rv or TRBC_MluI_rv primer (Table 3) annealing in the TCR constant domains. The PCR products were restriction enzyme (RE) digested and ligated into the RE-digested pFKPEN vector (37) and verified by sequencing. Gene-specific primers were designed (TRAV_NcoI_fw and TRAV_HindII_rv for V α ; TRBV_MluI_fw and TRBV_NotI_rv for V β ; Table 3) and the exact V domains were re-amplified, followed by ligation into RE-digested pFKPEN, generating the scTCR-WT380 (V α -linker-V β) construct followed by c-Myc and His6 tags. To generate pFKPEN encoding the stabilized scTCR s380 (P50^{TRAV}, L50^{TRBV}, F103^{TRBV}, and P2^{linker}), the scTCR-WT380 vector was used as template for QuikChange site-directed mutagenesis using TRAV_L50P_fw/TRAV_L50P_rv, Linker_L2P_fw/Linker_L2P_rv, and TRBV_L103F_fw/TRBV_L103F_rv primer pairs (Table 3). TCR 364.1.0.14 (herein denoted 364) and scTCR s364 have been described before (10). All constructs were verified by sequencing.

Soluble periplasmic TCR expression and purification. For soluble periplasmic expression of TCR, cells were inoculated from glycerol stocks into 1 liter of LB-TAG medium (1 \times LB medium supplemented with 30 μ g/ml tetracycline, 100 μ g/ml ampicillin, and 0.1 M glucose) and incubated overnight at 37°C/220 rpm. After reinoculation to an OD_{600nm} of 0.025, the cultures were further grown until an OD_{600nm} of 0.6–0.8. Cells were harvested by centrifugation (3,220 g for 10 minutes) and resuspended in 1 \times LB-TA. Notably, we abstained from IPTG induction of the LacPO promoter. After overnight incubation at 30°C/220 rpm, cells were pelleted at 3,220 g for 30 minutes and the resulting cell pellets were resuspended in 80 ml ice-cold periplasmic extraction solution (50 mM Tris-HCl, 20% sucrose, 1 mM EDTA) supplemented with 0.1 mg/ml RNase A and 1 mg/ml hen egg lysozyme and incubated 1 hour at 4°C with rotation. After centrifugation, the supernatant containing the periplasmic content was filtered (0.22 μ m) before purification.

TCRs were purified by IMAC purification according to the manufacturer's protocol (HiTrap HP, GE Healthcare) followed by size-exclusion chromatography using either a HiLoad 26/600 Superdex 200 column or a Superdex 200 10/30 GL (both GE Healthcare). A final polishing step of the samples on a ResourceQ or Superdex 75 10/300 GL (both GE Healthcare) were performed when needed. Columns were run in PBS supplemented with 150 mM NaCl and protein concentration was determined using the MW and extinction coefficient of each individual protein. Purified samples were verified by analytical gel filtration using either a Superdex 200 10/300 GL column or a Superdex 200 PC3.2/30 column (both GE Healthcare).

Expression and purification of soluble pMHC. Recombinant HLA-DQ2.5 with the gluten-derived peptides containing the deamidated T cell epitopes DQ2.5-glia- α 1a (QLQPFPOPELPY, underlined 9mer core sequence), DQ2.5-glia- α 2 (POPELPYPOPE), DQ2.5-glia- ω 1 (QQPFPOPEQPF), DQ2.5-glia- ω 2 (FPOPEQFPWQP), and DQ2.5-CLIP2 (MATPLLMQALPMGAL) coupled to the N-terminus of the DQ2.5 β -chain via a thrombin-cleavable linker peptide was expressed in insect cells using a baculovirus expression vector system as previously described (33, 38). Soluble, recombinant pMHC was affinity purified using mAb 2.12.E11 specific for the HLA-DQ2 β -chain (39). After site-specific biotinylation using BirA (Avidity), pMHCs were purified using a Superdex 200 GL10/300 run in PBS.

SPR. SPR was performed on either a Biacore T100 (T200 sensitivity enhanced) or Biacore T200 (GE Healthcare) by immobilizing Neutravidin (10 μ g/ml in acetate buffer pH 4.5) on a CM3 Series S sensor chip by amine coupling to 1,000 resonance units (RU). One hundred or 300 RU of pMHC was captured on the chip surface, followed by injections of dilution series of TCR using a single-cycle kinetics or multi-cycle kinetics method (data collection rate 10 Hz). For experiments determining kinetics, 4-fold dilutions from 50 μ M or 3-fold dilutions from 55 μ M were used, and for cross-reactivity experiments, 3-fold dilutions from 30 μ M were used. All experiments were performed at 25°C with a flow rate of 30 μ l/min and using 1 \times PBS supplemented with 150 mM NaCl and 0.05% surfactant P20. Glycine-HCl pH 2.7 was used to regenerate the sensor chip surface. All data were zero-adjusted and the Neutravidin reference flow cell value was subtracted before evaluation with T200 Evaluation Software, version 1.0 and RI set to constant. A 1:1 Langmuir binding model was used for determination of K_D . Figures were prepared using GraphPad Prism 7 and graphs were smoothed by averaging 8 neighbors.

Retroviral TCR transduction of hybridoma T cells and FACS sorting of transduced cells. The BW58 $\alpha\beta$ human CD4⁺ T cell hybridoma line devoid of endogenous TCR (herein denoted BW T cells, obtained from Bernard Malissen, Centre d'Immunologie de Marseille-Luminy, Marseille, France) (40) was retrovirally transduced as described before to express TCR 380 and the mutant TCR 364 versions (10). Synthetic, codon-optimized DNA encoding the TCR α and TCR β V domains (Genscript) fused to murine TCR constant domains was cloned into pMIG-II-TCR α -P2A-TCR β -eGFP retroviral plasmid (obtained from Dario Vignali, University of Pittsburgh, Pittsburgh, Pennsylvania, USA) (41) between the EcoRI/BamHI site or BamHI/XhoI sites, respectively. The assembly connects the TCR α - and β -chains through a P2A peptide linker. The pMIG-II-TCR plasmids were cotransfected (Lipofectamine 2000, Invitrogen) together with the pCL-Eco plasmid into GP2-293 packaging cells (Clontech). Virus supernatants were collected 48 and 72 hours after transfection, followed by centrifugation and filtration (0.45 μ m) to remove cell debris. BW T (50,000 cells) were incubated with 1 ml virus-containing supernatant supplemented with 10 μ g/ml polybrene and subjected to centrifugation (3,000 g) at 32°C for 90 minutes. Following removal of the supernatant, the BW T cells were cultured in RPMI with 10% FCS. To ensure equal TCR expression levels, cells were stained with hamster anti-mouse TCR α/β Alexa-647 mAb (1:20, clone H57/597, Life Technologies) and sorted using a FACS Aria II (BD Biosciences) based on eGFP and TCR expression level. The BW T cells were cultivated in RPMI/10% FCS. Data were analyzed using FlowJo software V10. The BW 364WT transductant has been described before (10).

Retroviral pMHC transduction of A20 B cells and FACS sorting of transduced cells. The murine A20 B cell line was retrovirally transduced to express HLA-DQ2.5 or HLA-DQ2.2 with different covalently coupled peptides. Notably, the ectodomains were made to be identical to the soluble, recombinant molecules. Epitope mutants were generated by mutating single positions in the DQ2.5-glia- α 1a or DQ2.5-glia- α 2 constructs. Codon-optimized synthetic DNA (Genscript) encoding the different peptides coupled to the N-terminal end of the HLA-DQ2.5 β -chain (DQB1*02:01) through a peptide linker was cloned as BamHI/XhoI fragment into the pMIG-II-DQ2.5-eGFP retroviral plasmid already encoding the HLA-DQ2.5 α -chain (DQA1*05:01). HLA-DQ2.2 with DQ2.5-glia- α 2 was constructed by exchanging DQA1*05:01 with DQA1*02:01 clones as a BglII/BamHI fragment (Genscript). After assembly, the HLA-DQ2.5/2.2 α - and β -chains are connected through a P2A peptide linker. The pMIG-II-eGFP-peptide-DQ2.5/2.2 plasmids and the pAmpho plasmid

were cotransfected into GP2-293 packaging cells as for the TCR transductions. A20 (50,000 cells) were incubated with 1.3 ml of virus-containing supernatant. After 5 days of culturing in RPMI/10% FCS, A20 cells expressing peptide-HLA-DQ2.5/2.2 were sorted (FACSARIA II) based on eGFP expression level. The A20 B cells were cultivated in RPMI/10% FCS. Data were analyzed using FlowJo software V10.

Evaluation of TCR and HLA-DQ2-peptide transductants by flow cytometry. To assess TCR and CD4 expression levels on the BW T cell variants, 3×10^5 cells were stained with hamster anti-mouse TCR α/β Alexa-647 mAb (1:20, clone H57/597, Life Technologies), anti-human CD4-PerCP-Cy5.5 mAb (1:20, clone RPA-T4, eBioscience), or isotype control mAbs (1:20, catalog number MCA2357A647, AbD Serotec; 1:160, clone P3.6.2.8.1, eBioscience). For validation of pMHC expression level on the A20 B cell variants, 2×10^5 cells were stained with biotinylated mAb 2.12.E11 (5 $\mu\text{g}/\text{ml}$; see ref. 39) specific for the HLA-DQ2.5/2.2 β -chain (in-house biotinylated), followed by streptavidin Alexa-647 (1:200, Invitrogen). Nonspecific binding to Fc γ receptors was blocked with anti-mouse CD16/32 (1:200, BD Pharmingen). Samples were fixed using a 2% paraformaldehyde solution and acquired on a FACSCalibur (BD) and analyzed by FlowJo software V10.

T cell activation assays. For T cell activation assays using peptide-loaded APCs, 60,000 HLA-DQ2.5⁺ EBV-transduced human B cells (CD114 derived from a DR3-DQ2.5 (*DQA1*05:01/DQB1*02:01*) homozygous CD patient) were incubated in RPMI/10% FCS at 37°C for 1–4 hours with titrated amounts of peptide (as indicated in the figures) before addition of 30,000 BW T cells. The following native (Q) or deamidated (E) gliadin peptides were used (epitopes are underlined, except in the long peptides that contain several overlapping epitopes): DQ2.5-glia- α 1aQ (QLQPFPPQQLPY), DQ2.5-glia- α 1aE (QLQPFPPQPELPY), DQ2.5-glia- α 2Q (PQPQLPYQPQL), DQ2.5-glia- α 2E (PQPELPYQPQL), DQ2.5-glia- ω 1E (PQQPFPPQEPFP), DQ2.5-glia- ω 2E (FPQPEQPFPWQP), α 33merQ (LQLQPFPPQQLPYQPQLPYQPQLPYQPQPF), α 33merE (LQLQPFPPQPELPYQPPELPYQPPELPYQPQPF), ω 19merQ (FPQPFPPQFPWQPEQPFPQ), and ω 17merE (QPQPFPPQPEQPFPWQP). In assays using TCC380 and TCC364, the 12mer DQ2.5-glia- α 1aE and DQ2.5-glia- α 2E peptides and the 33merE peptide were used. In T cell activation assays using murine A20 cells transduced with HLA-DQ2.5 with covalently linked peptide, titrated amounts of APCs were incubated with 30,000 BW T cells as indicated in the figures. As a control, Cell Stimulation Cocktail containing PMA and ionomycin (eBioscience, 1:500) was added to wells containing BW T cells only. Culture supernatants (25 μl) were assayed for murine IL-2 secretion after overnight incubation by ELISA (1:2 dilutions with PBST–2% FCS). Wells were coated with 50 μl of 2 $\mu\text{g}/\text{ml}$ rat anti-mouse IL-2 (clone JES6-1A12, Pharmingen) and detected with 50 μl of 5 $\mu\text{g}/\text{ml}$ biotin rat anti-mouse IL-2 (clone JES6-5H4, Pharmingen) followed by 80 μl of 1:2,000 diluted Streptavidin-ALP (Amersham). Plates were blocked with PBS–2% FCS and recombinant mouse IL-2 (Biolegend) was used as standard. T cell activation assays with thymidine incorporation as readout was performed essentially as described above, except for the following: APCs (CD114) were irradiated with 75 Gy before overnight incubation with titrated amounts of peptide at 37°C, before addition of 50,000 T cells (TCC380 and TCC364) in medium containing 10% heat-inactivated, pooled human serum. [³H]Thymidine (1 $\mu\text{Ci}/\text{well}$) was added after 20 hours and cells were harvested after an additional 72 hours using an automated harvester, and [³H]thymidine incorporation was measured by liquid scintillation counting. Figures were prepared using GraphPad Prism 7. Nonlinear regression analysis (3 parameters) was used to derive IL-2 concentrations from the standard curves. Graphs were normalized and EC50 values calculated using nonlinear regression analysis (log[agonist] vs. normalized response).

Analysis of single-cell TCR sequencing data. The single-cell TCR sequencing data generated in a gluten challenge study (Louise Fremgaard Risnes, unpublished observations) was used for analysis of V gene usage and pairing in 9 CD patients. The cells were isolated using a pool of DQ2.5:DQ2.5-glia- α 1a, DQ2.5:DQ2.5-glia- ω 1, DQ2.5:DQ2.5-glia- α 2, and DQ2.5:DQ2.5-glia- ω 2 tetramers (10 $\mu\text{g}/\text{ml}$ each), directly ex vivo from blood samples. The IMGT/HighV-QUEST online tool (imgt.org/HighV-QUEST/) was used for identifying V, D, and J genes and CDR3 junction sequences. T cells with at least 1 productive TRA and TRB pair (allowing dual on either TRA or TRB, maximum 3 chains) with dupcount greater than 100 were defined as valid cells. The valid cells with identical V and J gene (subgroup level) together with identical CDR3 region (allowing for 1 nucleotide mismatch) were defined as clonotypes. These unique clonotypes were used for comparative analysis of *TRBV7-2*, *TRBV7-3*, *TRAV26-1*, and *TRAV26-2* pairing and usage in the entire TCR repertoire. All sequencing data are deposited in the NCBI Sequence Read Archive under accession number SRP102402.

Molecular interaction studies. Molecular structures of TCR and pMHC were prepared using PyMol v1.8.4.2 using the crystal structure of TCR S16 bound to DQ2.5:DQ2.5-glia- α 2 (PDB: 4OZH) (12). Sequence alignments and residue nomenclature were assigned by use of the IMGT database (<http://www.imgt.org>). Modeling of the DQ2.5-glia- ω 1 and DQ2.5-glia- ω 2 epitopes was done using SPDV based on DQ2.5:DQ2.5-glia- α 1a (PDB: 1S9V) and DQ2.5:DQ2.5-glia- α 2 (PDB: 4OZH), respectively (12, 42). The 4OZH structure contains 2 asymmetric units. We used chains A, B, G, H, and J for structural analysis. Molecular visualization was done in PyMol.

Statistics. Statistical analyses were performed using GraphPad Prism 7.01 software. Values are expressed as mean and standard deviation or standard error of the mean as indicated in the figure legends. Nonlinear regression analysis (3 parameters) was used to derive IL-2 concentrations from the standard curves. Graphs were normalized and EC50 values calculated using nonlinear regression analysis (log[agonist] vs. normalized response).

Study approval. Biological material was obtained from CD patients according to protocols approved by the Regional Ethics Committee of South-Eastern Norway (2010/2720 and 2013/1237), and informed written consent was given by all subjects.

Author contributions

KSG and LSH contributed equally to the manuscript and are co-first authors. KSG, LSH, LFR, SDK, SWQ, and GÅL designed and performed experiments and analyzed data. RSN performed experiments. EB provided some of the recombinant, soluble pMHC. RF, TF, and MFP provided some of the BW or A20 transduced cells. BD contributed to structural analyses. KEAL provided biopsies from CD patients. SWQ, IS, LMS, and GÅL designed experiments, analyzed data, and supervised the study. KSG, LSH, IS and GÅL wrote the manuscript. All authors critically reviewed the manuscript.

Acknowledgments

The authors would like to thank Sivaganesh Sathiaruby for excellent technical assistance, Kjetil Taskén (Biotechnology Center of Oslo) for access to the Biacore T100 instrument, The Flow Cytometry Core Facility (FCCF) at Oslo University Hospital for cell sorting, and the HSØ Regional Core Facility for Structural Biology (grant 2015095). This work received funding from the South-Eastern Norway Regional Health Authority (grant 2012046) and from the Research Council of Norway through its Centers of Excellence funding scheme, project number 179573/V40.

Address correspondence to: Geir Åge Løset, Department of Biosciences, University of Oslo, N-0316 Oslo, Norway. Phone: 47.45025421; Email: g.a.loset@ibv.uio.no.

1. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol.* 2006;24:419–466.
2. Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses? *Nat Rev Immunol.* 2008;8(3):231–238.
3. Garcia KC, Adams JJ, Feng D, Ely LK. The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nat Immunol.* 2009;10(2):143–147.
4. Sharon E, Sibener LV, Battle A, Fraser HB, Garcia KC, Pritchard JK. Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat Genet.* 2016;48(9):995–1002.
5. Sollid LM, Lie BA. Celiac disease genetics: current concepts and practical applications. *Clin Gastroenterol Hepatol.* 2005;3(9):843–851.
6. Shan L, et al. Structural basis for gluten intolerance in celiac sprue. *Science.* 2002;297(5590):2275–2279.
7. 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.* 2000;191(4):603–612.
8. Tye-Din JA, et al. Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci Transl Med.* 2010;2(41):41ra51.
9. Molberg O, et al. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med.* 1998;4(6):713–717.
10. Qiao SW, et al. Posttranslational modification of gluten shapes TCR usage in celiac disease. *J Immunol.* 2011;187(6):3064–3071.
11. Qiao SW, Christophersen A, Lundin KE, Sollid LM. Biased usage and preferred pairing of α - and β -chains of TCRs specific for an immunodominant gluten epitope in coeliac disease. *Int Immunol.* 2014;26(1):13–19.
12. Petersen J, et al. T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. *Nat Struct Mol Biol.* 2014;21(5):480–488.
13. Richman SA, et al. Structural features of T cell receptor variable regions that enhance domain stability and enable expression as single-chain ValphaVbeta fragments. *Mol Immunol.* 2009;46(5):902–916.
14. Gunnarsen KS, et al. Chaperone-assisted thermostability engineering of a soluble T cell receptor using phage display. *Sci Rep.*

- 2013;3:1162.
15. Cole DK, et al. Human TCR-binding affinity is governed by MHC class restriction. *J Immunol.* 2007;178(9):5727–5734.
 16. Davis MM, et al. Ligand recognition by alpha beta T cell receptors. *Annu Rev Immunol.* 1998;16:523–544.
 17. Petersen J, et al. Diverse T cell receptor gene usage in HLA-DQ8-associated celiac disease converges into a consensus binding solution. *Structure.* 2016;24(10):1643–1657.
 18. Broughton SE, et al. Biased T cell receptor usage directed against human leukocyte antigen DQ8-restricted gliadin peptides is associated with celiac disease. *Immunity.* 2012;37(4):611–621.
 19. Huppa JB, et al. TCR-peptide-MHC interactions in situ show accelerated kinetics and increased affinity. *Nature.* 2010;463(7283):963–967.
 20. Liu B, Chen W, Natarajan K, Li Z, Margulies DH, Zhu C. The cellular environment regulates in situ kinetics of T-cell receptor interaction with peptide major histocompatibility complex. *Eur J Immunol.* 2015;45(7):2099–2110.
 21. Dahal-Koirala S, et al. TCR sequencing of single cells reactive to DQ2.5-glia- α 2 and DQ2.5-glia- ω 2 reveals clonal expansion and epitope-specific V-gene usage. *Mucosal Immunol.* 2016;9(3):587–596.
 22. Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol.* 2014;32(7):684–692.
 23. Rubelt F, et al. Individual heritable differences result in unique cell lymphocyte receptor repertoires of naïve and antigen-experienced cells. *Nat Commun.* 2016;7:11112.
 24. Fallang LE, Bergseng E, Hotta K, Berg-Larsen A, Kim CY, Sollid LM. 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.* 2009;10(10):1096–1101.
 25. Petersen J, et al. Determinants of gliadin-specific T cell selection in celiac disease. *J Immunol.* 2015;194(12):6112–6122.
 26. Moon JJ, et al. Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity.* 2007;27(2):203–213.
 27. Feng D, Bond CJ, Ely LK, Maynard J, Garcia KC. Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction ‘codon’. *Nat Immunol.* 2007;8(9):975–983.
 28. Dai S, et al. Crossreactive T cells spotlight the germline rules for alphabeta T cell-receptor interactions with MHC molecules. *Immunity.* 2008;28(3):324–334.
 29. Garcia KC, et al. A closer look at TCR germline recognition. *Immunity.* 2012;36(6):887–888.
 30. Stadinski BD, et al. A role for differential variable gene pairing in creating T cell receptors specific for unique major histocompatibility ligands. *Immunity.* 2011;35(5):694–704.
 31. Scott-Browne JP, White J, Kappler JW, Gapin L, Marrack P. Germline-encoded amino acids in the alphabeta T-cell receptor control thymic selection. *Nature.* 2009;458(7241):1043–1046.
 32. Adams JJ, et al. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity.* 2011;35(5):681–693.
 33. Fallang LE, et al. Complexes of two cohorts of CLIP peptides and HLA-DQ2 of the autoimmune DR3-DQ2 haplotype are poor substrates for HLA-DM. *J Immunol.* 2008;181(8):5451–5461.
 34. Hou T, et al. An insertion mutant in DQA1*0501 restores susceptibility to HLA-DM: implications for disease associations. *J Immunol.* 2011;187(5):2442–2452.
 35. Nguyen TB, Jayaraman P, Bergseng E, Madhusudhan MS, Kim CY, Sollid LM. Unraveling the structural basis for the unusually rich association of human leukocyte antigen DQ2.5 with class-II-associated invariant chain peptides. *J Biol Chem.* 2017;292(22):9218–9228.
 36. Molberg O, McAdam SN, Lundin KE, Sollid LM. Studies of gliadin-specific T-cells in celiac disease. *Methods Mol Med.* 2000;41:105–124.
 37. Gunnarsen KS, Lunde E, Kristiansen PE, Bogen B, Sandlie I, Løset GÅ. Periplasmic expression of soluble single chain T cell receptors is rescued by the chaperone FkpA. *BMC Biotechnol.* 2010;10:8.
 38. Quarsten H, et al. Staining of celiac disease-relevant T cells by peptide-DQ2 multimers. *J Immunol.* 2001;167(9):4861–4868.
 39. Viken HD, et al. Characterization of an HLA-DQ2-specific monoclonal antibody. Influence of amino acid substitutions in DQ beta 1*0202. *Hum Immunol.* 1995;42(4):319–327.
 40. Letourneur F, Malissen B. Derivation of a T cell hybridoma variant deprived of functional T cell receptor alpha and beta chain transcripts reveals a nonfunctional alpha-mRNA of BW5147 origin. *Eur J Immunol.* 1989;19(12):2269–2274.
 41. Holst J, Vignali KM, Burton AR, Vignali DA. Rapid analysis of T-cell selection in vivo using T cell-receptor retrogenic mice. *Nat Methods.* 2006;3(3):191–197.
 42. Kim CY, Quarsten H, Bergseng E, Khosla C, Sollid LM. Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proc Natl Acad Sci USA.* 2004;101(12):4175–4179.