

# The molecular basis of T cell receptor recognition of citrullinated tenascin-C presented by HLA-DR4

Received for publication, April 3, 2025, and in revised form, May 16, 2025 Published, Papers in Press, June 2, 2025,  
<https://doi.org/10.1016/j.jbc.2025.110326>

Hien Thy Dao<sup>1</sup>, Tiing Jen Loh<sup>1</sup> , Ravi K. Sharma<sup>2,3,4</sup> , Lars Klareskog<sup>2,3</sup>, Vivianne Malmström<sup>2,3</sup> , Hugh H. Reid<sup>1</sup> ,  
Jamie Rossjohn<sup>1,5,\*</sup> , and Jia Jia Lim<sup>1,\*</sup> 

From the <sup>1</sup>Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Australia; <sup>2</sup>Division of Rheumatology, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden; <sup>3</sup>Center for Molecular Medicine, Karolinska Institutet, Solna, Sweden; <sup>4</sup>Department of Clinical Immunology and Rheumatology, All India Institute of Medical Sciences, Bilaspur (H.P), India; <sup>5</sup>Institute of Infection and Immunity, Cardiff University School of Medicine, Heath Park, Cardiff, United Kingdom

Reviewed by members of the JBC Editorial Board. Edited by Karen Fleming

CD4<sup>+</sup> T cell autoreactivity against citrullinated (cit) self-epitopes presented by HLA-DRB1 is associated with rheumatoid arthritis (RA) pathogenesis. We understand the molecular bases of T cell receptor (TCR) recognition of cit-fibrinogen, cit-vimentin, and cit- $\alpha$ -enolase epitopes, and the role of citrulline in shaping TCR repertoire usage. Nevertheless, how TCRs recognize other cit-epitopes, including tenascin-C (TNC) and how alternative citrullination positions may modulate the T cell recognition remains unclear. Here, we examined TNC<sup>1014,1016cit</sup> peptide, which contains citrullination at position P-1 and P2, to study the underlying TCR-HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> molecular interactions. Crystal structure of HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> at 2.4 Å resolution revealed a conserved peptide binding register to the established HLA-DRB1\*04:01-peptide structures, where both citrullines protruded upward. Next, we determined the crystal structure of a RA patient-derived TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> (PB) TCR in complex with HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> at 3.2 Å resolution. The CDR3 $\alpha$  loop (<sup>109</sup>VGNTN<sup>113</sup>) of PB TCR formed a secondary helical conformation at the N-terminus of the peptide binding cleft, allowing extensive interactions between the P-1 and P2 citrullines of TNC<sup>1014,1016cit</sup> peptide. Surface plasmon resonance, tetramer staining, and CD69 activation assays revealed that the PB TCR did not cross-react to other RA autoantigens, and the P-1-Cit, P2-Cit, and P5-Tyr of TNC<sup>1014,1016cit</sup> are the key determinants underlying the strict specificity of the PB TCR. Collectively, we provide molecular insight into citrullination in modulating TCR recognition.

Rheumatoid arthritis (RA) is a T cell-mediated autoimmune disease that affects primarily the joints, and its progression leads to significant morbidity and reduced life expectancy. The disease has a prevalence of 0.5% worldwide and more than 500,000 people in Australia (1, 2). A hallmark characteristic of RA is the presence of anti-citrullinated protein antibodies

(ACPA) in sera, for which >70% of the patients with RA are ACPA seropositive in established RA and before disease onset (3–9). ACPA recognizes proteins that have undergone citrullination, a post-translational modification (PTM), whereby a positively charged arginine is deiminated to the neutral amino acid citrulline, driven by peptidyl arginine deiminases (PAD), specifically PAD2 and PAD4 enzymes, leading to the emergence of neo-antigens and subsequently self-antigen immunogenicity (10, 11).

The genetic susceptibility of RA is strongly associated with the human leukocyte antigen (HLA) *loci*, specifically the HLA-DRB1 genes, which encode an HLA-DR  $\beta$ -chain possessing a common five-amino-acid sequence of either QKRAA, QRRAA, or RRRAA at positions 70 to 74 (12). This motif, known as the HLA-shared susceptibility epitope (SE), forms the P4 binding pocket of the HLA-DRB1 peptide binding cleft (12). With the presence of a positively charged arginine or lysine residue at position 71, a negatively charged or neutral polar residue is favored in the P4 pocket (13). This is consistent with the previous studies that the conversion of arginine into citrulline enhances the binding affinity of peptide antigen to HLA-DRB1 allomorphs bearing SE (13, 14). The prevalence and genetic susceptibility of RA can vary depending on ethnicity and racial stratification of SE-encoded HLA-DRB1 alleles, with the HLA-DRB1\*04:01 allele possessing the highest risk of RA development (odds ratio (OR) of 4.44), alongside with HLA-DRB1\*04:04 (odds ratio of 4.22), HLA-DRB1\*01:01 (OR of 2.17), and HLA-DRB1\*10:01 (OR of 4.22) are the predominant SE-encoding alleles in Europeans (15). In Japanese and East Asians, the most common SE-encoding allele is HLA-DRB1\*04:05 (OR of 4.22) (15, 16); Native Americans possess a common SE-coding allele of HLA-DRB1\*14:02 (17).

The inflamed and arthritic joint synovium of RA patients is characterized by CD4<sup>+</sup> T cell infiltration and the accumulation of neo-cit-epitopes originating from extracellular matrix proteins such as type II collagen (18), fibrinogen (19), tenascin-C (TNC) (20, 21), cartilage intermediate layer protein (CILP) (22), as well as cell-associated components including vimentin (22, 23) and  $\alpha$ -enolase (24). The cit-peptides reactive CD4<sup>+</sup> T

\* For correspondence: Jia Jia Lim, [Jiajia.lim@monash.edu](mailto:Jiajia.lim@monash.edu); Jamie Rossjohn, [Jamie.rossjohn@monash.edu](mailto:Jamie.rossjohn@monash.edu).

## T cell recognition of HLA-DR4 presenting cit-tenascin

cells reactive to cit-peptides derived from vimentin,  $\alpha$ -enolase, fibrinogen, and TNC were detected in patients with RA with the HLA-DRB1\*04:01 (*HLA-DRA1\*01:01/HLA-DRB1\*04:01*) as well as in HLA-DRB1\*04:01 transgenic mice immunized with cit-peptides (23, 25–29). Subsequent phenotypic characterization showed an increase of cit-peptide-specific T helper 1 (Th1) effector memory CD4<sup>+</sup> T cells in patients with SE<sup>+</sup> RA (30, 31). Consistent with their effector role, CD4<sup>+</sup> T cells produce proinflammatory cytokines in response to citrullinated antigens (21, 32), further implicating the effector function of CD4<sup>+</sup> T cells in RA pathogenesis.

We have previously reported the structural basis for the mouse CD4<sup>+</sup> T cell receptor (TCR) recognition of HLA-DRB1\*04:01 presenting cit-fibrinogen and revealed that citrullination at position P4 confers high affinity to HLA-DRB1\*04:01 allomorph, which in turns allows SE to co-recognize the cit-epitope and TCRs (26). In addition to P4 cit-fibrinogen, citrullination at position P2 impacted the responding TCR repertoire in immunized mice (26). Moreover, we have recently provided the key determinants underpinning TCR recognition of citrullinated vimentin and  $\alpha$ -enolase epitopes, both of which contain P4-citrulline (27). In this study, we use a citrullinated TNC<sup>1014,1016cit</sup> peptide, which contains two citrullination sites at positions P-1 and P2, to further investigate the direct impact of multi-citrullination on TCR recognition. The TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR was previously isolated from peripheral blood mononuclear cells (PBMC) of patient with RA (29). We show specific binding to double citrullinated TNC<sup>1014,1016cit</sup> peptide presented by HLA-DRB1\*04:01, and the citrullines play a critical role in determining PB TCR recognition. We provide insight into the key

determinants of PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> recognition and the cross-reactivity of TNC<sup>1014,1016cit</sup> peptide to another SE-encoded allomorphs.

## Results

### HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> presentation

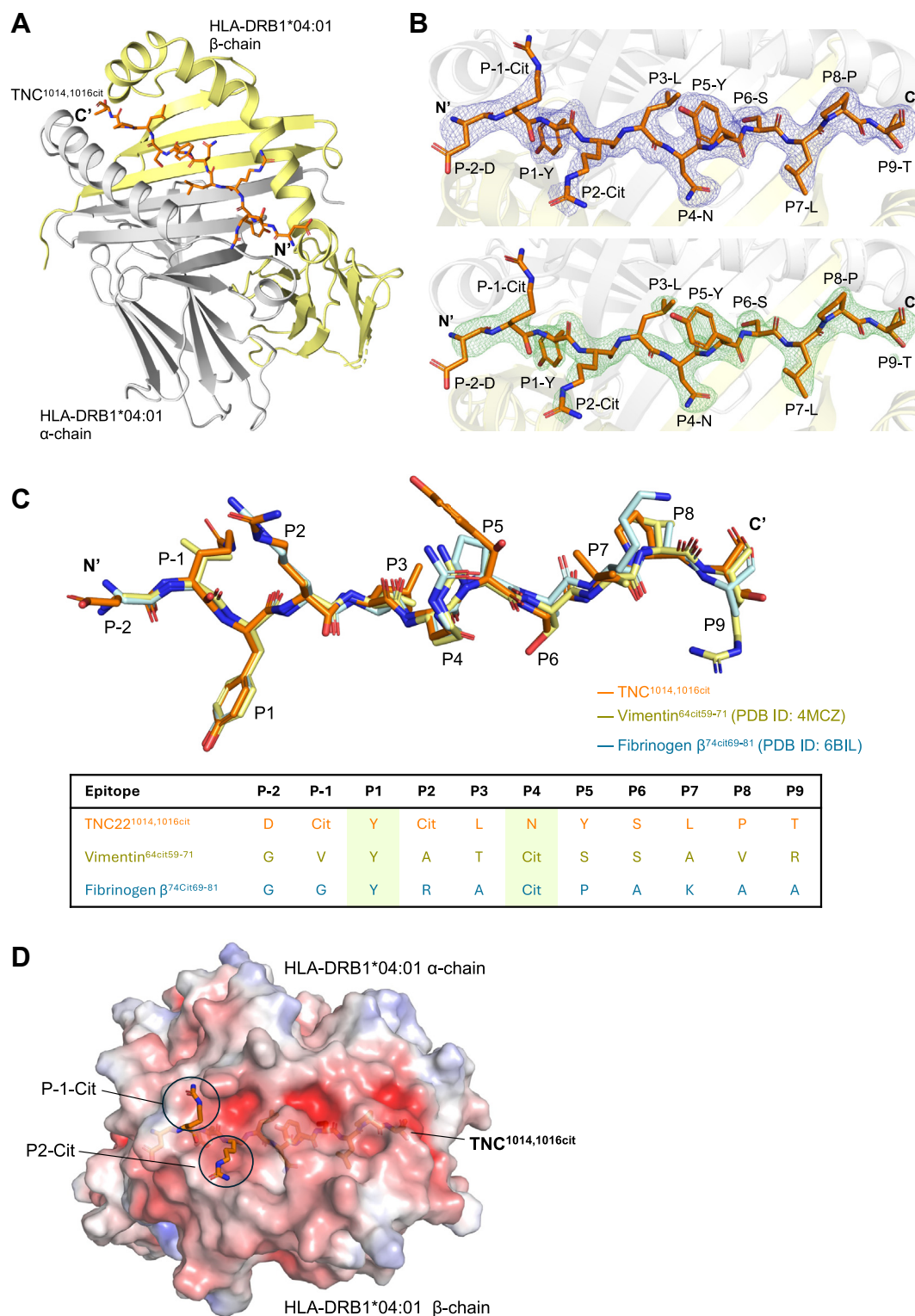
To understand the register of TNC<sup>1014,1016cit</sup> peptide (<sup>1013</sup>DcitYcitLNYSLPTG<sup>1024</sup>) presented by the HLA-DRB1\*04:01 molecule, we solved the crystal structure of HLA-DRB1\*04:01 presenting TNC<sup>1014,1016cit</sup> (<sup>1013</sup>DcitYcitLNYSLPTG<sup>1024</sup> KK, where two lysine residues were added at the C-terminus to improve the peptide solubility) at 2.4 Å resolution (Table 1, Fig. 1A). The TNC<sup>1014,1016cit</sup> peptide bound to HLA-DRB1\*04:01 in a canonical conformation, spanning from the N-terminus P-2 to C-terminus P9 residue of the peptide binding cleft, with a well-defined electron density map (Fig. 1B). As expected, the TNC<sup>1014,1016cit</sup> peptide was aligned with the established binding motif of previously reported HLA-DRB1\*04:01 presenting fibrinogen  $\beta$ <sup>74cit69–81</sup> and vimentin <sup>64cit59–71</sup> peptides, where the P1 pocket accommodated by a tyrosine residue, and the P4 pocket was occupied with a neutral asparagine or citrulline residue, respectively (13, 14) (Fig. 1, B and C). In this HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> structure, both citrullines at positions P-1 and P2 were protruding away from the binding cleft, indicating that the presence of citrulline at such positions did not impact HLA-DRB1\*04:01 binding (Fig. 1, B and C). This finding was supported by the similar relative binding strength reported for native and single/multiple citrullinated TNC<sup>1013–24</sup>

**Table 1**

**Data collection and refinement statistics of HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>, PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> and PB TCR structure**

Crystal structure and PDB entry	HLA-DRB1*04:01 <sup>TNC1014,1016cit</sup> cit (PDB ID: 9NIH)	PB TCR-HLA-DRB1*04:01 <sup>TNC1014,1016cit</sup> (PDB ID: 9NIG)	PB TCR (PDB ID: 9NII)
Data collection			
Space group	H3	P1	P22121
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	119.30 119.30 73.27	76.93 88.67 126.07	109.26 119.80 176.30
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00 90.00 120.00	81.47 76.75 65.45	90.00 90.00 90.00
Resolution (Å)	34.53–2.40 (2.49–2.40) <sup>a</sup>	47.39–3.20 (3.30–3.20) <sup>a</sup>	47.84–2.75 (2.82–2.75) <sup>a</sup>
R <sub>sym</sub> or R <sub>merge</sub>	0.092 (0.352) <sup>a</sup>	0.112 (0.311) <sup>a</sup>	0.128 (1.164) <sup>a</sup>
CC <sub>1/2</sub>	0.998 (0.970) <sup>a</sup>	0.924 (0.506) <sup>a</sup>	0.996 (0.704) <sup>a</sup>
I/ <i>s</i> (I)	16.4 (7.2) <sup>a</sup>	5 (1.3) <sup>a</sup>	8.3 (1.6) <sup>a</sup>
Completeness (%)	99.6 (99.3) <sup>a</sup>	98.9 (99.0) <sup>a</sup>	99.9 (99.9) <sup>a</sup>
Redundancy	10.2 (10.4) <sup>a</sup>	1.8 (1.8) <sup>a</sup>	6.1 (6.3) <sup>a</sup>
Refinement			
Resolution (Å)	34.53–2.40	44.320–3.2	47.84–2.75
No. reflections	15,153	48,029	60,604
R <sub>work</sub> /R <sub>free</sub>	0.1842/0.2186	0.2187/0.2604	0.2326/0.2644
No. atoms	3232	18,764	13,573
Protein	3076	18,651	13,498
Ligand/ion	46	113	8
Water	110	–	67
B-factors (Å <sup>2</sup> )	44.385	59.028	66.764
Protein	44.183	58.811	66.837
Ligand/ion	78.773	94.859	68.989
Water	35.682	–	51.873
R.m.s. deviations			
Bonds lengths (Å)	0.004	0.003	0.002
Bond angles (°)	0.691	0.537	0.475
Rama allowed (%)			
Rama favored (%)	97.29	96.55	95.89
Rama outlier (%)	0	0	0

<sup>a</sup> Values in parentheses refer to the highest resolution shell.



**Figure 1. HLA-DRB1\*04:01 in complex with TNC<sup>1014,1016cit</sup>.** A, cartoon representation of HLA-DRB1\*04:01 presenting TNC<sup>1014,1016cit</sup> peptide with HLA-DRB1\*04:01  $\alpha$ - and  $\beta$ -chains are colored in grey and yellow, respectively, whereas TNC<sup>1014,1016cit</sup> peptide is presented as orange stick. B, the refined 2mF<sub>o</sub> - DF<sub>c</sub> map (top) and SA omit map (bottom) of TNC<sup>1014,1016cit</sup> peptide are shown in blue and green, respectively. Both maps are contoured at 1 $\sigma$ . C, overlaid structure of TNC<sup>1014,1016cit</sup> (orange), vimentin<sup>64cit59-71</sup> (yellow; PDB ID: 4MCZ) and fibrinogen  $\beta$ <sup>74cit69-81</sup> peptides (pale cyan; PDB ID: 6BIL) from the binary complex with HLA-DRB1\*04:01. List of TNC<sup>1014,1016cit</sup>, vimentin<sup>64cit59-71</sup> and fibrinogen  $\beta$ <sup>74cit69-81</sup> peptide sequences from positions P-2 to P9. All amino acids are indicated in single-letter abbreviations, Cit = Citrulline. D, the Adaptive Poisson Boltzmann Solver-generated electrostatic surface of HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> binary structure. Two citrulline residues at positions P-1 and P-2 of TNC<sup>1014,1016cit</sup> are circled in black.



## T cell recognition of HLA-DR4 presenting cit-tenascin

peptide bound to HLA-DRB1\*04:01 ( $EC_{50} \sim 0.6\text{--}1\text{ }\mu\text{M}$ ) (21). Moreover, solvent-exposed residues of the  $\text{TNC}^{1014,1016\text{cit}}$  epitope, including P5-Tyr and P8-Pro, might also play a role in interacting with the TCR (Fig. 1, B and C). The electrostatic profile of HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$  binary complex revealed a distinct neutral charged feature at P-1 and P2 of citrullinated  $\text{TNC}^{1014,1016\text{cit}}$  peptide as opposed to its native form (positive charged arginine in both position), likely suggesting the preference of the overall surface charge of the contacting TCR CDR loops which would affect the TCR specificity (Fig. 1D). Overall, the highly conserved binding register of HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$  binary structure suggests the potential critical role of citrullines in modulating TCR recognition.

### Citrullinated $\text{TNC}^{1014,1016\text{cit}}$ is essential for TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR reactivity

A human TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> CD4<sup>+</sup> T cell clone was previously isolated from the PBMC of a HLA-DRB1\*04:01 patient with ACPA-positive RA, *via* HLA-DRB1\*04:01- $\text{TNC}^{1014,1016\text{cit}}$  tetramer (29) (Fig. 2A). To investigate the antigen specificity and impact of citrullination in TCR recognition, we transiently expressed TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR in the HEK293 T cell line and stained with individual HLA-DRB1\*04:01 tetramers presenting  $\text{TNC}^{1014,1016\text{cit}}$ , vimentin<sup>64cit59-71</sup>,  $\alpha$ -enolase<sup>15cit10-22</sup>, or fibrinogen  $\beta^{74\text{cit}69-81}$ , respectively (Figs. 2B and S1). As expected, TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR bound specifically to  $\text{TNC}^{1014,1016\text{cit}}$  peptide and did not cross-react to other RA autoantigens (Fig. 2B). Next, we expressed, refolded, and purified TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR and subsequently determined the steady state binding affinity of this TCR with four different variants of  $\text{TNC}^{1013-24}$  peptide, namely,  $\text{TNC}^{1014,1016\text{cit}}$  (P-1 and P2 citrullinated),  $\text{TNC}^{1014\text{cit}}$  (P-1 citrullinated),  $\text{TNC}^{1016\text{cit}}$  (P2 citrullinated), and native  $\text{TNC}^{1013-24}$  peptide *via* surface plasmon resonance (SPR) (Fig. 2C). The affinity of PB TCR-HLA-DRB1\*04:01- $\text{TNC}^{1014,1016\text{cit}}$  fell within the relative range of TCR-pMHC II interaction (33) as well as previously determined TCRs-HLA-DRB1\*04:01-cit-epitopes (26, 27). The PB TCR bound strongest in the presence of both citrullines to HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$ , with a  $K_D = 25.8\text{ }\mu\text{M}$  and did not recognize native  $\text{TNC}^{1013-24}$  peptide, highlighting the essential role of citrullination in PB TCR recognition (Fig. 2C). In particular,  $\text{TNC}^{1014\text{cit}}$  had a critical impact on PB TCR recognition, whereas  $\text{TNC}^{1016\text{cit}}$  displayed a twofold weaker affinity ( $K_D = 50\text{ }\mu\text{M}$ ) to PB TCR than the double citrullinated epitope (Fig. 2C).

Subsequently, we used the T cell activation assay to provide insight into the functionality of the PB TCR and HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$  interaction. Here, we generated a PB TCR transduced SKW-3 CD4<sup>+</sup> T cell line and measured the expression of CD69 and CD3 cell surface markers as an indicator of T cell activation. We observed an increase in CD69 expression and a concomitant down-regulation of CD3 expression, which corresponded with

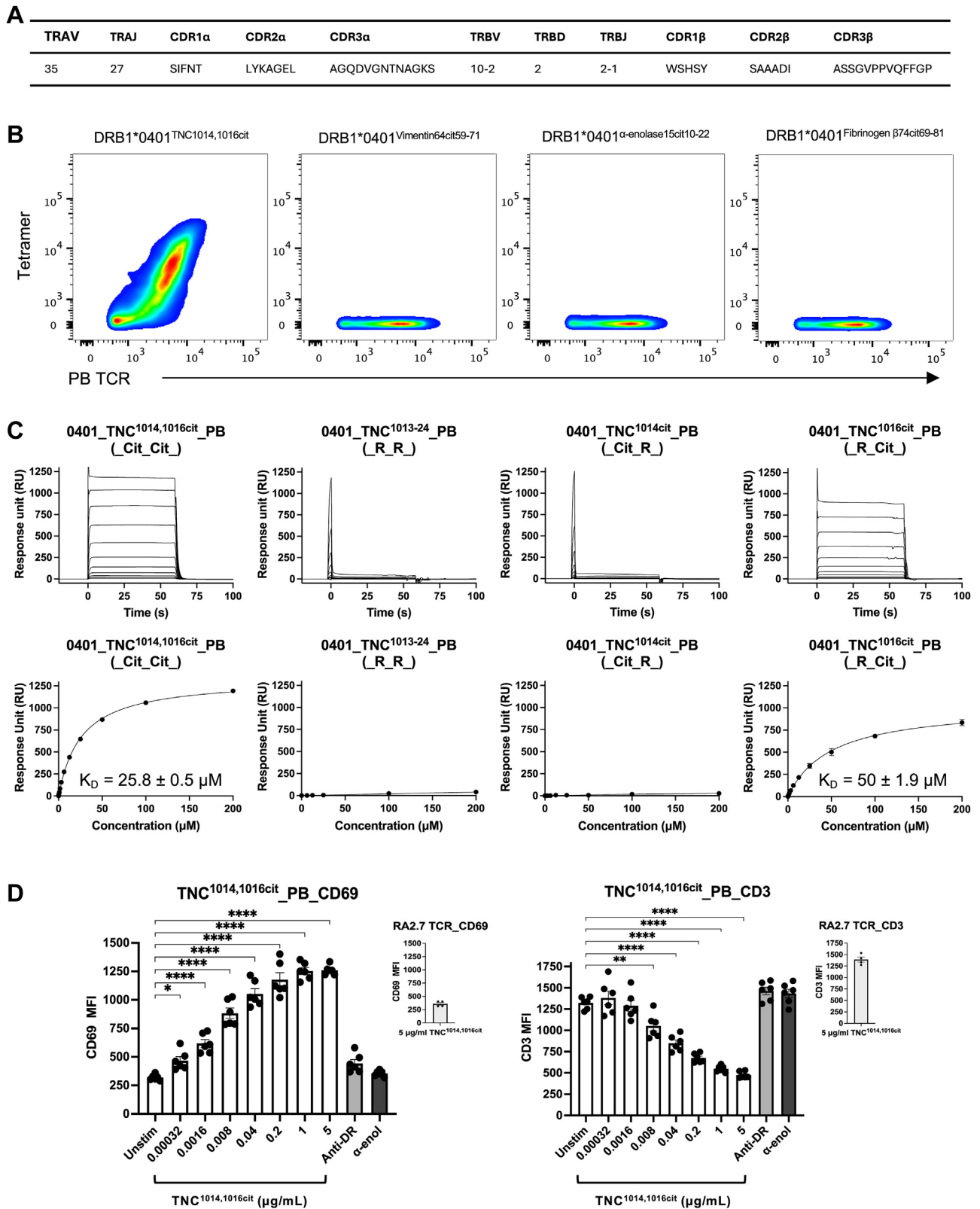
a dose-dependent response to the  $\text{TNC}^{1014,1016\text{cit}}$  peptide concentration, suggesting that PB T cell line activated the TCR signaling pathway in response to  $\text{TNC}^{1014,1016\text{cit}}$  peptide recognition (Fig. 2D). In particular, the PB TCR was highly reactive to  $\text{TNC}^{1014,1016\text{cit}}$  peptide, with approximately 0.32 ng/ml of peptide was sufficient to significantly activate PB TCR with a maximal response reached at  $\sim 1\text{ }\mu\text{g/ml}$  (Fig. 2D). Here, we used the HLA-DRB1\*04:01  $\alpha$ -enolase<sup>15cit10-22</sup> restricted TCR, RA2.7 (27) and an anti-HLA-DR blocking antibody (clone LB3.1) as controls, indicating that the activation of PB TCR is citrullinated  $\text{TNC}^{1014,1016\text{cit}}$ -specific and in a HLA-DRB1\*04:01-dependent manner (Fig. 2D). Taken together, double citrullination at  $\text{TNC}^{1014,1016\text{cit}}$  peptide is a key determinant underlying PB TCR recognition and reactivity, with citrulline at P2 being critical for recognition, while citrulline at P-1 enhances recognition.

### Docking topology of TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR on HLA-DRB1\*04:01 $\text{TNC}^{1014,1016\text{cit}}$

To understand the molecular basis underpinning the specific recognition of PB TCR for  $\text{TNC}^{1014,1016\text{cit}}$  presented by HLA-DRB1\*04:01, we determined the ternary complex structure at 3.2 Å resolution (Table 1, Figs. 3A, and S2A). The electron density map at the interface of the PB TCR and HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$  was clearly defined (Fig. S2A). The PB TCR docked canonically at an angle of  $\sim 70^\circ$  across the central region of peptide binding cleft, with a total Buried surface area of  $\sim 1910\text{ Å}^2$  over HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$  (Fig. 3B and Table S1). The TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR chain usage was biased toward the  $\alpha$ -chain that constituted  $\sim 62\%$  of BSA as compared to  $\sim 38\%$  of the  $\beta$ -chain (Fig. 3B and Table S1). Particularly, the PB TCR CDR3 $\alpha$  and 2 $\alpha$  made major contributions to the HLA-DRB1\*04:01  $\text{TNC}^{1014,1016\text{cit}}$  interaction, by contributing 32.8% and 21.1% of total BSA, respectively, followed by CDR1 $\alpha$  (4.4%) and framework  $\alpha$  (FW $\alpha$ ) (3.7%) (Fig. 3B and Table S1). In contrast, CDR3 $\beta$ , 2 $\beta$ , 1 $\beta$  and FW $\beta$  contributed 17.1%, 9.6%, 7.6% and 3.7% of total BSA, respectively (Fig. 3B and Table S1). Intriguingly, the  $\text{TNC}^{1014,1016\text{cit}}$  peptide contacts were mainly derived from non-germline encoded CDR3 $\alpha$  and 3 $\beta$  loops, whereas germline encoded CDR1 $\beta$ , 2 $\beta$ , 1 $\alpha$  and 2 $\alpha$  loops contributed to the interaction with HLA-DRB1\*04:01 (Fig. 3B).

### Molecular basis of TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR recognition of HLA-DRB1\*04:01

The germline encoded TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR made extensive contacts with HLA-DRB1\*04:01, in which CDR- $\alpha$  and - $\beta$  loops mainly interacted with HLA-DRB1\*04:01  $\beta$ - and  $\alpha$ -chains, respectively, indicating a typical canonical docking mode of TCR-peptide-major histocompatibility complex II (pMHC II) (33) (Fig. 3, C–E and Table S2). Here, CDR1 $\alpha$  (Asn<sup>37</sup>) and FW $\alpha$  (Ala<sup>55</sup>) made contacts with Thr<sup>77</sup> and Asp<sup>66</sup> of the HLA-DRB1\*04:01  $\beta$ -chain, respectively, *via* van der Waals (VdW) interactions



**Figure 2. Identification of CD4<sup>+</sup> T cell restricted to HLA-DRB1\*04:01 presenting citrullinated TNC<sup>1014,1016cit</sup> peptide.** A, gene segment usage and CDR loops sequence of PB TCR (29). B, *in vitro* PB TCR expression and tetramer staining for individual HLA-DRB1\*04:01 tetramers presenting citrullinated RA autoantigens, including TNC<sup>1014,1016cit</sup>, Vimentin<sup>64cit59-71</sup>, α-enolase<sup>15cit10-22</sup>, and Fibrinogen β<sup>74cit69-81</sup> peptides. Gating strategy is shown in Fig.S1. C, binding affinity of PB TCR against HLA-DRB1\*04:01 presenting TNC<sup>1014,1016cit</sup>, native TNC<sup>1013-24</sup>, TNC<sup>1014cit</sup> and TNC<sup>1016cit</sup> peptides. HLA-DRB1\*04:01<sup>CLIP</sup> was immobilized in the reference flow cell to control non-specific binding. PB TCR equilibrium affinity constants ( $K_D$ ) values were determined from three independent experiments in duplicate and curve fitted using a 1:1 binding model. For each concentration, the points represent the mean values, and the error bars correspond to  $\pm$  s.e.m. from three independent experiments in duplicate. D, activation assay of PB TCR transduced SKW3 T cells against BLCL 9031 expressing HLA-DRB1\*04:01 stimulated with TNC<sup>1014,1016cit</sup> peptide. Upregulation CD69 expression (left) and down-regulation CD3 expression (right) of PB

(Fig. 3C and Table S2). Notably, germline-encoded CDR2 $\alpha$  residues (Tyr<sup>57</sup>, Lys<sup>58</sup>, and Glu<sup>64</sup>) formed multiple interactions with the DRB1\*04:01  $\beta$ -chain, including SE residues (Gln<sup>70</sup>, Ala<sup>73</sup>, and Arg<sup>72</sup>) and adjacent residues (Asp<sup>66</sup> and Glu<sup>69</sup>) *via* H-bonds, salt bridge, and VdWs, suggesting the importance of conserved TRAV gene usage in HLA-DRB1\*04:01  $\beta$ -chain recognition (Fig. 3C and Table S2). For the PB TCR  $\beta$ -chain, CDR1 $\beta$  (Ser<sup>37</sup> and Tyr<sup>38</sup>) and CDR2 $\beta$  (Ala<sup>57</sup>, Ala<sup>58</sup>, and Ile<sup>65</sup>) formed multiple VdWs with the HLA-DRB1\*04:01  $\alpha$ -chain (Gln<sup>57</sup>, Gly<sup>58</sup>, Leu<sup>60</sup>, Ala<sup>61</sup>, and Ala<sup>64</sup>) (Fig. 3D and Table S2). Tyr<sup>55</sup> and Asp<sup>67</sup> of PB TCR FW $\beta$  also interacted with Gln<sup>57</sup> of HLA-DRB1\*04:01  $\alpha$ -chain (Fig. 3D and Table S2). In the context of non-germline encoded interface, CDR3 $\alpha$  (Val<sup>109</sup>, Asn<sup>113</sup>, and Ala<sup>114</sup>) and CDR3 $\beta$  (Val<sup>109</sup>, Pro<sup>110</sup>, and Pro<sup>111</sup>) interacted with HLA-DRB1\*04:01  $\alpha$ - (Phe<sup>54</sup>, Glu<sup>55</sup>, Ala<sup>56</sup>, Gln<sup>57</sup>, and Gly<sup>58</sup>) and  $\beta$ -chains (Leu<sup>67</sup>, Gln<sup>70</sup>, and His<sup>81</sup>), respectively, *via* H-bonds and VdWs (Fig. 3E and Table S2). Overall, the substantial contribution of germline-encoded residues of TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR, alongside non-germline-encoded residues, highlights the potential TCR gene usage preference specific for HLA-DRB1\*04:01 engagement.

## Altered structural changes in the CDR3 $\alpha$ loop are critical for co-recognition of PB TCR-TNC<sup>1014,1016cit</sup>

The PB TCR and TNC<sup>1014,1016cit</sup> peptide interactions were primarily driven by the CDR3 $\alpha$  and CDR3 $\beta$  loops, followed by limited involvement of CDR1 $\alpha$  and CDR1 $\beta$  (Fig. 4A). Five residues in CDR3 $\alpha$  loop (<sup>109</sup>VGNTN<sup>113</sup>) formed a secondary helical conformation which in turn sat atop the N-terminus of TNC<sup>1014,1016cit</sup> peptide at position P-1 to P3 pocket (Fig. 4, A and B). To understand whether this secondary structure transition of the CDR3 $\alpha$  loop is ligand driven, we also determined the structure of the apo form of PB TCR at 2.75 Å resolution (Table 1 and Fig. S2B). Superposition of PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> holo and PB TCR apo form at the C $\alpha$  backbone of the TCR showed a very subtle change with a root mean square deviation (r.m.s.d) value of 0.57 Å (Fig. S2B). In the PB TCR apo form, the CDR3 $\alpha$  loop was unstructured, whereupon Asn<sup>111</sup> and Thr<sup>112</sup> were positioned downward (Fig. 4B). While in the PB TCR holo state, the <sup>109</sup>VGNTN<sup>113</sup> of the CDR3 $\alpha$  loop reoriented with Val<sup>109</sup> and Asn<sup>113</sup> to point downward facing HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>. Thus, the altered conformational transition of the CDR3 $\alpha$  loop in the PB TCR-TNC<sup>1014,1016cit</sup> holo state is ligand driven.

In the context of the TNC<sup>1014,1016cit</sup> peptide, there was a rearrangement of the citrulline residues upon CDR3 $\alpha$  loop docking. Namely, citrulline at P2 shifted ~40° toward the C-terminus, alongside with subtle rearrangement of the P-1 citrulline at the N-terminus, which allowed the docking of the conformationally rearranged CDR3 $\alpha$  loop (Fig. 4C).

## Detailed interactions of TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR and TNC<sup>1014,1016cit</sup> peptide

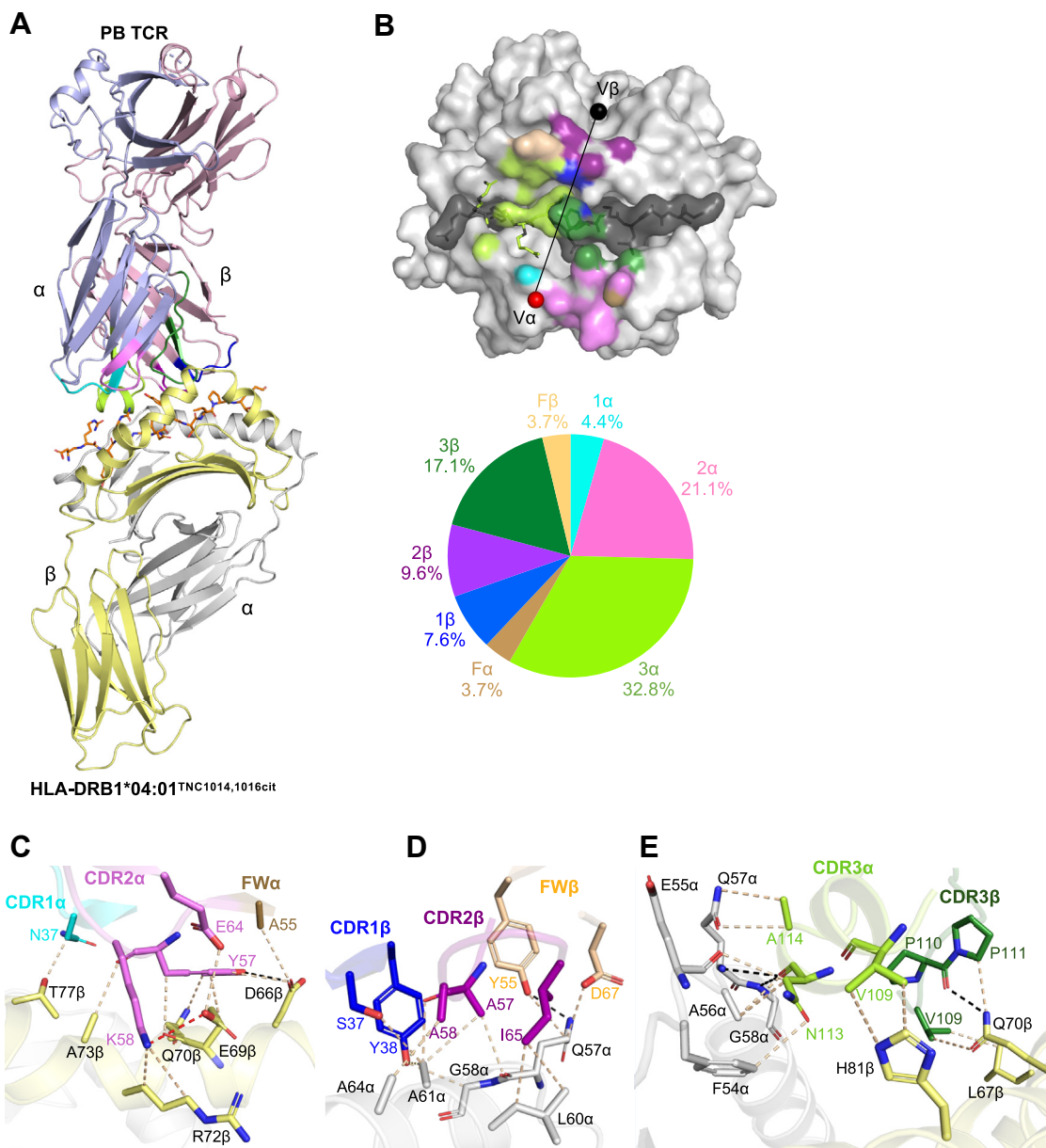
At the PB TCR-TNC<sup>1014,1016cit</sup> interface, the nature of the  $\alpha$ -helical turn in the CDR3 $\alpha$  loop enabled extensive interactions between P-1 to P3 residues of the TNC<sup>1014,1016cit</sup> peptide. Here, Val<sup>109</sup> projected downward and positioned at the center of the P-1 and P2 citrullines, forming multiple contacts with both citrullines, as well as the main chain interactions with P1-Tyr (Fig. 4D and Table S2). The adjacent Gly<sup>110</sup> and Asn<sup>113</sup> of CDR3 $\alpha$  also made extensive H-bonds and VdW contacts with P-1 citrulline, as well as main chain interactions with P1-Tyr, P2-Cit, and P3-Leu (Fig. 4D and Table S2). Moreover, Thr<sup>112</sup> and Asn<sup>37</sup> of CDR3 $\alpha$  and CDR1 $\alpha$  loops, respectively, contacted with P2 citrulline within the H-bond distance (Fig. 4D and Table S2). Another distinct feature of PB TCR recognition of TNC<sup>1014,1016cit</sup> peptide was observed at Position 5, where numerous interactions were made between P5-Tyr and CDR3 $\alpha$  (Thr<sup>112</sup>, Gln<sup>107</sup>, and Gly<sup>115</sup>), CDR3 $\beta$  (Pro<sup>110</sup> and Val<sup>109</sup>), and CDR1 $\beta$  (Tyr<sup>38</sup>) (Fig. 4D and Table S2). The neutral feature of citrullinated TNC<sup>1014,1016cit</sup> peptide in the presence of both P-1 and P2 citrullines, complemented with the hydrophobic Val<sup>109</sup> of CDR3 $\alpha$ , further explains the role of citrullination in recognizing PB TCR (Fig. 4D). Such hydrophobic feature of Val<sup>109</sup> in CDR3 $\alpha$  likely disfavored the positively charged arginine in the native TNC<sup>1013-24</sup> peptide or single citrullinated TNC<sup>1016cit</sup> peptide due to charge repulsion (Fig. 4D). The structural analyses were consistent with the SPR data of high affinity of PB TCR towards double citrullinated TNC<sup>1014,1016cit</sup> peptide as opposed to P2 citrullinated TNC<sup>1016cit</sup> peptide (Fig. 4D). Collectively, both citrullines at P-1 and P2, alongside with P5-Tyr, are the key determinants for TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR recognition.

## Energetic determinants underlying PB TCR recognition and HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>

To define the energetically important residues of the TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR attributing to HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> recognition, we conducted an alanine-scanning mutagenesis of a panel of 11 residues on PB TCR that are involved in contacts with HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>, and analyzed their impact on TCR-pMHC II binding using SPR. The impact of each mutation was classified into four categories: no effect (<2-fold reduced affinity compared to wildtype), moderate (2-5-fold reduced affinity), severe (5-10-fold reduced affinity), and deleterious (>10-fold reduced affinity). Alanine substitution of Asn<sup>113</sup> (CDR3 $\alpha$ ), Tyr<sup>38</sup> (CDR1 $\beta$ ), Val<sup>109</sup>, and Pro<sup>110</sup> (CDR3 $\beta$ ) residues which co-contacted both HLA-DRB1\*04:01 and TNC<sup>1014,1016cit</sup> peptide had a deleterious effect on PB TCR recognition (Figs. 5A, S3 and Table S3). CDR2 $\alpha$  (Tyr<sup>57</sup>) and CDR3 $\beta$  (Pro<sup>111</sup>) which contacted HLA-DRB1\*04:01 particularly at the shared epitope also revealed severe impact on

TCR upon serial dilution of TNC<sup>1014,1016cit</sup> peptide are shown in the bar chart. Three independent experiments in duplicate were performed. *p*-values were determined by one-way ANOVA with Dunnett's multiple comparison testing, \**p* ≤ 0.05, \*\**p* ≤ 0.01, \*\*\*\**p* ≤ 0.0001, and error bars correspond to ± s.e.m. from three independent experiments in duplicate.





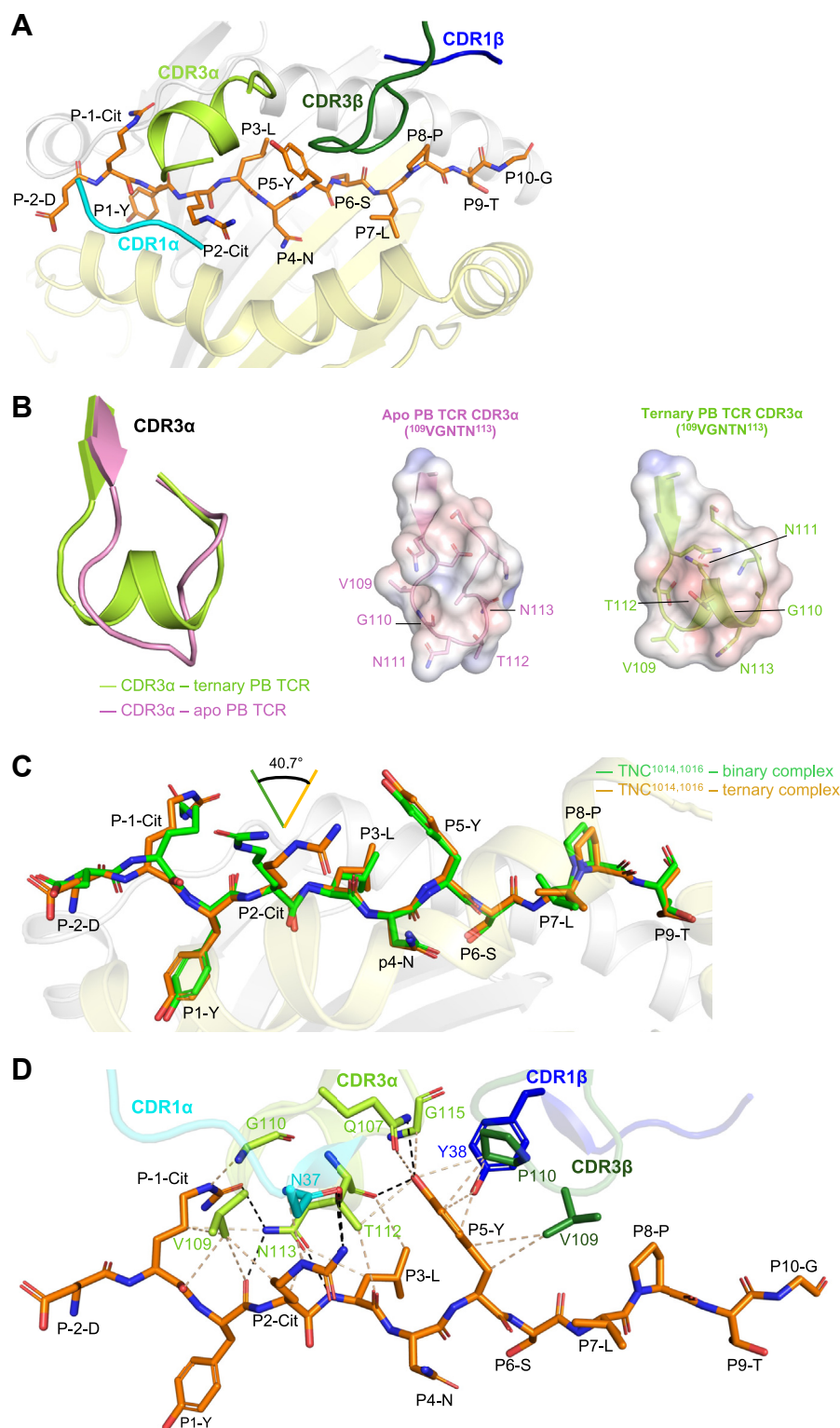
**Figure 3.** PB TCR recognition of HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>. **A**, overall cartoon representation of PB TCR complexed to HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>. The HLA-DRB1\*04:01  $\alpha$ - and  $\beta$ -chains are highlighted in grey and yellow, whereas PB TCR  $\alpha$ - and  $\beta$ -chains are presented in light purple and light pink, respectively. The TNC<sup>1014,1016cit</sup> peptide is presented as orange sticks. The CDR1 $\alpha$ , 2 $\alpha$  and 3 $\alpha$  loops are colored in cyan, violet and light green, while the CDR1 $\beta$ , 2 $\beta$  and 3 $\beta$  are highlighted in blue, purple and dark green, respectively. The FW $\alpha$  residues are colored in sand and FW $\beta$  residues are presented in beige. **B**, (top) surface representation of PB TCR footprint on the HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>. The atoms from HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> interacting with PB TCR are colored according to the nearest CDR loops that they are interacting with. The V $\alpha$  and V $\beta$  center of mass positions are presented as spheres in red and black, respectively, connecting via a black line. (Bottom) pie chart highlights the relative contribution of CDR loops to the interface of HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>. Detailed interactions of PB TCR between (C) germline encoded CDR1 $\alpha$ , CDR2 $\alpha$  and FW $\alpha$ , (D) CDR1 $\beta$ , CDR2 $\beta$  and FW $\beta$  and (E) non-germline encoded CDR3 $\alpha$  and CDR3 $\beta$  with HLA-DRB1\*04:01 are shown. Black dashes represent H-bond within 3.5 Å, beige dashes correspond to VdW interaction within 4 Å and red dashes denote disulfide bond within 4.5 Å distance. All amino acids are indicated in single-letter abbreviations.

HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> recognition with >5-fold reduced in affinity when compared to wildtype TCR (Figs. 5A, S3 and Table S3). Other residues, including CDR1 $\alpha$  (Asn<sup>37</sup>), CDR3 $\alpha$  (Val<sup>109</sup> and Thr<sup>112</sup>), and CDR1 $\beta$  (Ser<sup>37</sup>), which interacted with either HLA-DRB1\*04:01 or TNC<sup>1014,1016cit</sup> peptide, showed a moderate impact on PB TCR recognition. In contrast, FW $\beta$  (Tyr<sup>55</sup>) had no impact on pHLA recognition (Figs. 5A, S3 and Table S3). Collectively, the impact of germline encoded and variable residues in CDR

loops formed an energetic hotspot on HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> at the N-terminus region P1-P5 across the HLA-DRB1\*04:01 peptide-binding cleft (Fig. 5B).

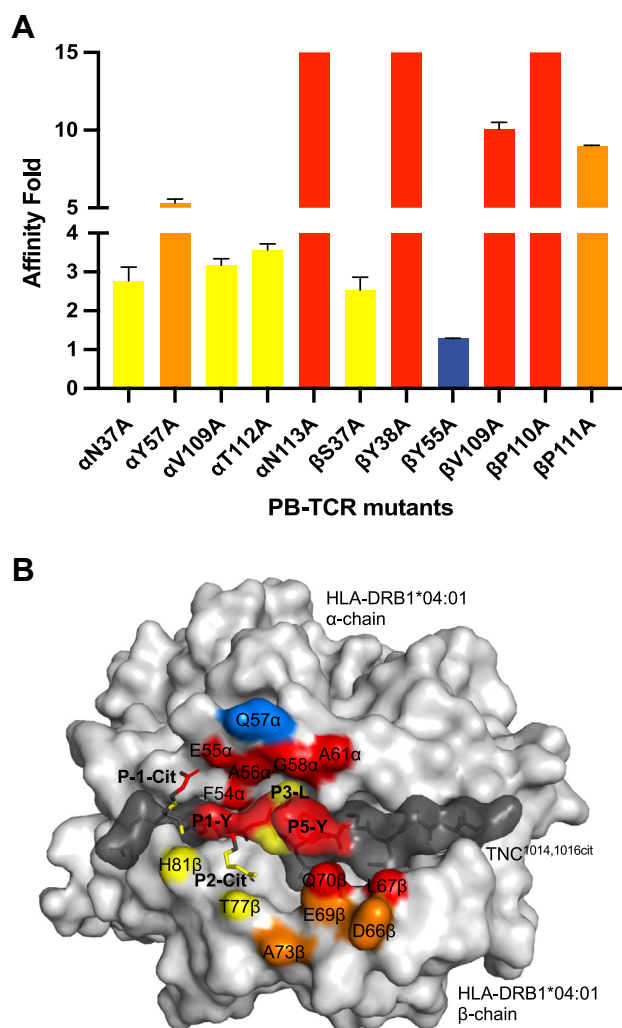
#### Role of HLA-DRB1\*04:01 SE in TCR recognition

Given that the TNC<sup>1014,1016cit</sup> peptide contains P4-Asn instead of P4-Cit (26, 27), we then analyzed the SE (<sup>70</sup>QKRAA<sup>74</sup>) interaction with P4-Asn of TNC<sup>1014,1016cit</sup>



**Figure 4.** TNC<sup>1014,1016cit</sup> peptide-mediated PB TCR interactions. *A*, schematic representation depicting the docking topology of CDR1 $\alpha$ , 3 $\alpha$ , 1 $\beta$  and 3 $\beta$  loops atop of HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup>. HLA-DRB1\*04:01  $\alpha$ - and  $\beta$ -chains are colored in grey and yellow, while the CDR1 $\alpha$ , 3 $\alpha$ , 1 $\beta$  and 3 $\beta$  loops are highlighted in cyan, light green, blue and dark green, respectively. TNC<sup>1014,1016cit</sup> peptide is presented as orange stick. *B*, (left) overlaid CDR3 $\alpha$  loop from the unliganded PB TCR apo structure, in pink, and from the PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> ternary complex, in light green. The Adaptive Poisson Boltzmann Solver-generated electrostatic surface of (middle) apo PB TCR-CDR3 $\alpha$  loop and (right) holo PB TCR-CDR3 $\alpha$  loop, displaying the arrangement of <sup>109</sup>VGNTN<sup>113</sup> sequence. *C*, overlaid P-2 to P-9 residues of TNC<sup>1014,1016cit</sup> peptide from the HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> binary complex (green) and from the PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> ternary complex (orange). *D*, detailed interactions of PB TCR in contact with the TNC<sup>1014,1016cit</sup> peptide are shown in sticks. Black dashes represent H-bond within 3.5 Å distance and beige dashes correspond to VdW interaction within 4 Å distance. All amino acids are indicated in single-letter abbreviations. Citrulline denotes as Cit.





**Figure 5. Effect of PB TCR point mutations at the HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> interface.** A, the affinity in fold of PB TCR mutants as compared to the native PB TCR was calculated corresponding to equilibrium affinity constants ( $K_D$ ) values in Figure S3 and Table S3. The impact of each mutation was categorized as no effect (<2-fold reduced affinity compared to wildtype, blue), moderate (2-5-fold reduced affinity, yellow), severe (5-10-fold reduced affinity, orange), and deleterious (>10-fold reduced affinity, red). All data were derived from two independent measurements in duplicate and the error bars correspond to  $\pm$  s.e.m. B, energetic footprint of PB TCR on HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> complex. The impact of each mutation is colored according to (A). Surface representation of HLA-DRB1\*04:01 and the TNC<sup>1014,1016cit</sup> peptide are colored in light grey and dark grey, respectively.

peptide and PB TCR to understand the role of HLA-DRB1\*04:01 SE in PB TCR recognition (Fig. 6A). Here, we showed that the P4-Asn of TNC<sup>1014,1016cit</sup> peptide is anchored with Lys<sup>71</sup> of SE, and multiple contacts are formed between other SE residues (Gln<sup>70</sup>, Arg<sup>72</sup>, and Ala<sup>73</sup>) and CD2 $\alpha$  (Tyr<sup>57</sup> and Lys<sup>58</sup>) and CDR3 $\beta$  (Val<sup>109</sup>, Pro<sup>110</sup>, and Pro<sup>111</sup>) (Fig. 6A). A constant pattern of dual-recognition between Lys<sup>71</sup> of SE and P4-Cit/P4-Asn, as well as Gln<sup>70</sup> of SE and TCR CDR loops, was observed in this study, alongside three previously reported citrullinated epitopes, including vimentin<sup>64cit59-71</sup>,  $\alpha$ -enolase<sup>15cit10-22</sup>, and fibrinogen  $\beta$ <sup>74cit69-81</sup>, albeit with distinct CDR loops or residues involved, suggesting this pattern is a hallmark of SE-peptide-TCR recognition (Fig. 6, A–D).

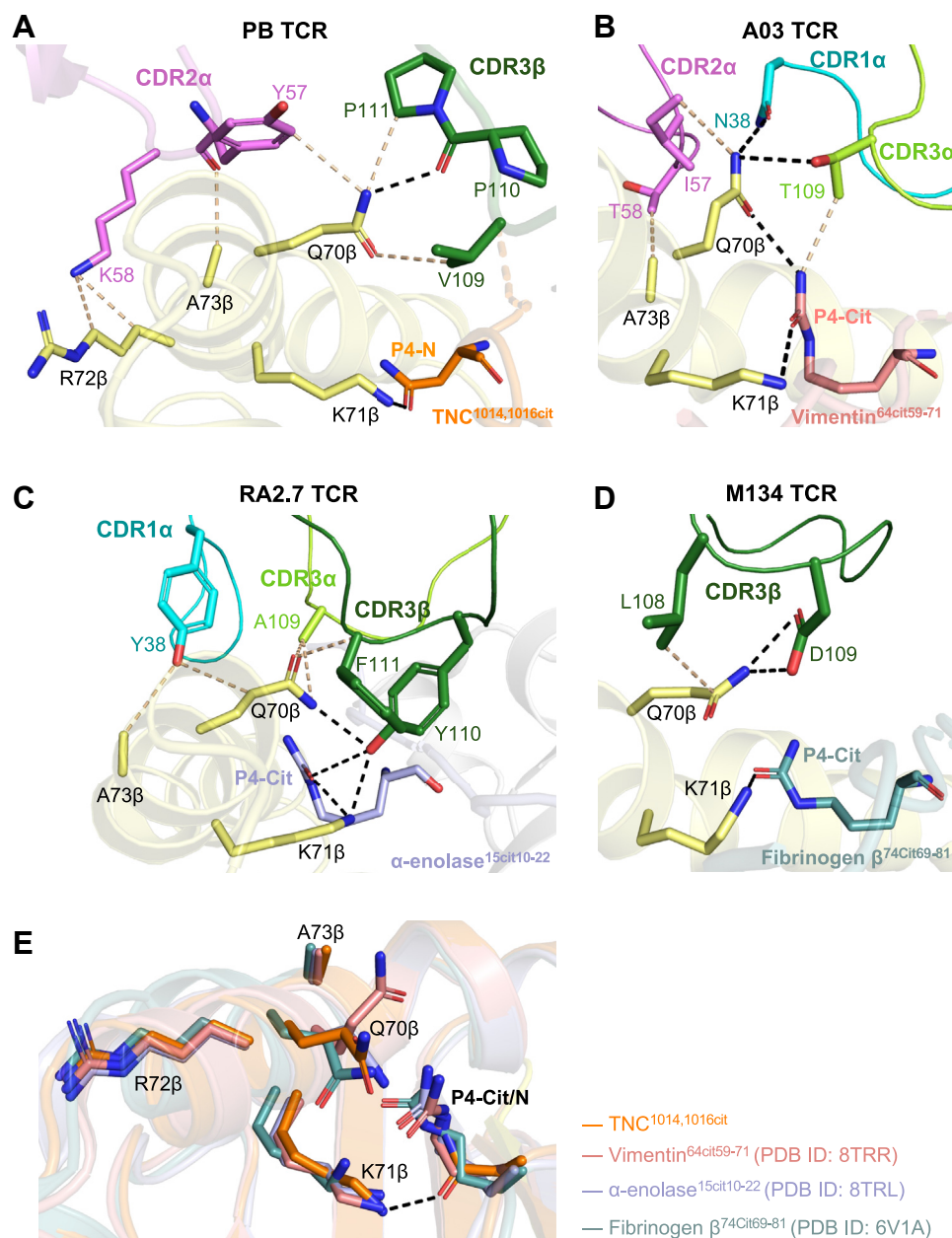
Superposition of TCR-pHLA complexes of TNC<sup>1014,1016cit</sup>, vimentin<sup>64cit59-71</sup>,  $\alpha$ -enolase<sup>15cit10-22</sup>, and fibrinogen  $\beta$ <sup>74cit69-81</sup> at the SE region revealed a highly similar pattern between P4-Cit/P4-Asn and SE residues, with certain flexibility arises from Gln<sup>70</sup> and Lys<sup>71</sup> due to rearrangement of the alpha helix of the HLA-DRB1\*04:01  $\beta$ -chain (Fig. 6E). Accordingly, the HLA-DRB1\*04:01 SE plays a consistent role in shaping PB TCR recognition of TNC<sup>1014,1016cit</sup> despite asparagine at position P4.

#### Presentation of TNC<sup>1014,1016cit</sup> peptide by other SE<sup>+</sup> HLA-DRB1 allomorphs

Considering the high homology between SE<sup>+</sup> HLA-DRB1 allomorphs, we next investigated if the TNC<sup>1014,1016cit</sup> peptide can be presented by other SE<sup>+</sup> HLA-DRB1 allomorphs and impacted in cross-reactivity to PB TCR. We performed a peptide competition assay using the fluorescence polarization technique to measure the relative binding strength of TNC<sup>1014,1016cit</sup> peptide for other HLA-DRB1 allomorphs (Fig. 7A). HA peptide that binds relatively strong to all HLA-DRB1 allomorphs, albeit weaker binding to DRB1\*04:04 was used as a control peptide (Fig. S4). As expected, HLA-DRB1\*01:01, \*04:01, and \*14:02 exhibited comparable strong binding to TNC<sup>1014,1016cit</sup> peptide, with a  $IC_{50}$  of 0.8  $\mu$ M, 1.1  $\mu$ M, and 1.3  $\mu$ M, respectively (Fig. 7A). Moreover, HLA-DRB1\*04:05 revealed more than fourfold weaker affinity when compared to HLA-DRB1\*01:01, \*04:01, and \*14:02. The relative binding strength of HLA-DRB1\*01:01, \*04:01 and \*14:02 fell within the “strong binding” range as previously established HLA-DRB1\*04:01 and cit-epitopes including cit-fibrinogen and cit-vimentin (14). In contrast, HLA-DRB1\*04:04 allomorph had limited binding to the TNC<sup>1014,1016cit</sup> peptide, with a  $IC_{50}$  of over 300  $\mu$ M and did not reach 100% inhibition (Fig. 7A). The sequence alignment of these SE<sup>+</sup> HLA-DRB1 allomorphs at the peptide binding cleft revealed high sequence identity of the anchoring residues at the P1, P4, P6, and P9 pockets for HLA-DRB1\*04:05, \*01:01, and \*14:02 alleles (Fig. 7B), consistent with established MHC II binding motif (34). Superposed crystal structures of HLA-DRB1\*04:01-TNC<sup>1014,1016cit</sup> and HLA-DRB1\*04:04 revealed that the HLA-DRB1\*04:04 allomorph did not bind the TNC<sup>1014,1016cit</sup> peptide due to the hydrophobic Val<sup>86</sup> in P1, which might inhibit the bulky aromatic tyrosine residue being accommodated in the P1 pocket due to steric clashes (Fig. 7, B and D).

#### Cross-reactivity of PB TCR toward HLA-DRB1\*04:05 presenting TNC<sup>1014,1016cit</sup> peptide

Next, we characterized the PB TCR cross-reactivity to other SE<sup>+</sup> HLA-DRB1 allomorphs by transient expression and staining of PB TCR transfectants with individual DRB1\*04:05, \*01:01, and \*14:02 tetramers loaded with TNC<sup>1014,1016cit</sup> peptide. Despite the observation that TNC<sup>1014,1016cit</sup> peptide can bind to these three HLA-DRB1 allomorphs, PB TCR showed only some cross-reactivity to the HLA-DRB1\*04:05 allomorph presenting TNC<sup>1014,1016cit</sup> peptide (Fig. 7C). This result

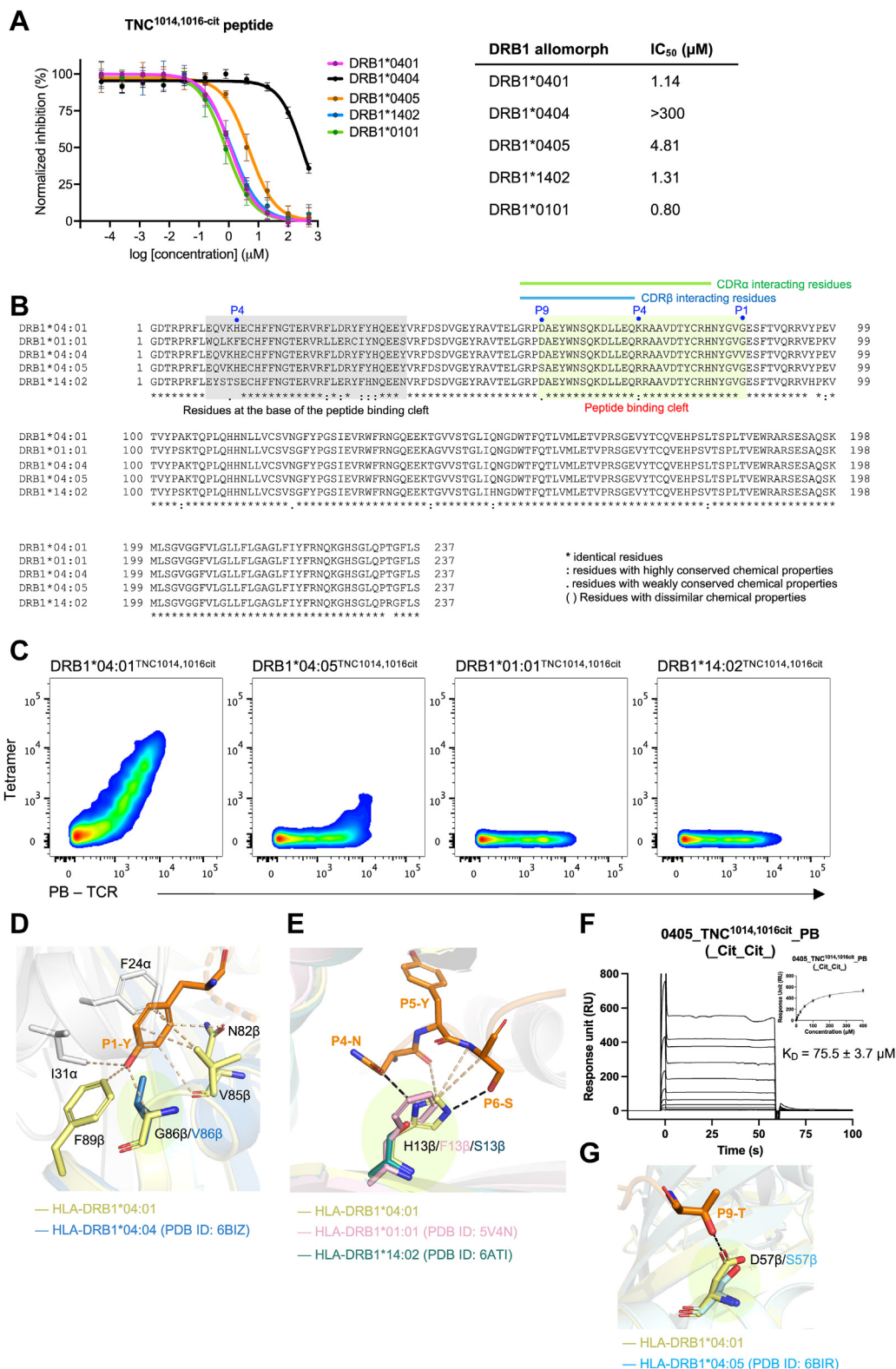


**Figure 6. Detailed SE interactions of HLA-DRB1\*04:01 with cit-epitopes and restricted TCRs.** Detailed SE interactions for (A) TNC<sup>1014,1016cit</sup> restricted PB TCR, (B) vimentin<sup>64cit59-71</sup> restricted A03 TCR (PDB ID: 8TRR), (C)  $\alpha$ -enolase<sup>15cit10-22</sup> restricted RA2.7 TCR (PDB ID: 8TRL) and (D) fibrinogen  $\beta$ 74Cit69-81 restricted M134 TCR (PDB ID: 6V1A) are shown in sticks. P4 residue of cit-epitopes for (A) TNC<sup>1014,1016cit</sup>, (B) vimentin<sup>64cit59-71</sup>, (C)  $\alpha$ -enolase<sup>15cit10-22</sup>, and (D) fibrinogen  $\beta$ 74Cit69-81 are colored in orange, pink, light purple, and teal, respectively. The SE residues of HLA-DRB1\*04:01  $\beta$ -chain are presented in yellow sticks, whereas the CDR1 $\alpha$ , 2 $\alpha$ , 3 $\alpha$ , and 3 $\beta$  loops are colored in cyan, violet, light green and dark green, respectively. Black dashes represent H-bond within 3.5 Å distance, beige dashes correspond to VdW interaction within 4 Å distance. E, overlaid SE residues of above mentioned four different cit-epitopes ternary complexes and P4-Asn/Cit. All amino acids are indicated in single-letter abbreviations. Citrulline denotes as Cit.

suggests the polymorphism within the HLA-DRB1 allomorphs plays a role in TCR recognition. Structural alignment of HLA-DRB1\*04:01, \*01:01, and \*14:02 revealed a deviation at residue 13, which is located on the  $\beta$ -sheet floor with their side chains oriented into the peptide-binding groove. This feature is important for P4 and P6 anchoring. The substitution of His<sup>13</sup> in HLA-DRB1\*04:01 to either Phe<sup>13</sup> (DRB1\*01:01) or Ser<sup>13</sup> (DRB1\*14:02), likely affects the rearrangement of the C-terminal region of the TNC<sup>1014,1016cit</sup> peptide to avoid clashes or to improve binding to the base of the peptide-binding groove.

This substitution will thus influence TCR recognition (Fig. 7, B and E).

Subsequent SPR analyses of PB TCR affinity for HLA-DRB1\*04:05<sup>TNC1014,1016cit</sup> revealed a  $K_D$  of 75.5  $\mu$ M, indicating the potential cross-reactivity of PB TCR to the DRB1\*04:05 allomorph (15) (Fig. 7F). The approximately threefold reduced in affinity of PB TCR towards HLA-DRB1\*04:05<sup>TNC1014,1016cit</sup>, as compared to HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> was likely due to the polymorphisms embedded in the peptide binding cleft, which are located at the P9 and P4 pocket, where Asp<sup>57</sup> and



**Figure 7. PB TCR cross-reactivity towards other SE<sup>+</sup> HLA-DRB1 allomorphs presenting TNC<sup>1014,1016</sup>cit.** A, titration curves of competitive binding of TNC<sup>1014,1016</sup>cit peptide to DRB1\*04:01, \*01:01, \*04:04, \*04:05, and \*14:02 allomorphs. Each data point represents normalized relative binding (in percentage) for two independent experiments in triplicate, and the binding affinity at 50% inhibition of total binding was calculated as IC<sub>50</sub> (μM) and the error bars correspond to ± s.e.m. B, multiple sequence alignment for HLA-DRB1\*04:01, \*01:01, \*04:04, \*04:05, and \*14:02 alleles. Residues at the peptide binding cleft and form the base of the peptide binding cleft are highlighted in light green and grey, respectively. The conservation of the residues are denoted as '\*' identical, '.' highly conserved, and ' ' low similarity, and ' ' distinct. C, *in vitro* PB TCR expression and tetramer staining analysis for TNC<sup>1014,1016</sup>cit peptide presented by HLA-DRB1\*04:01, \*04:05, \*01:01, and \*14:02 allomorphs. Corresponding gating strategy is shown in Fig.S1. D, superposed HLA-DRB1\*04:01 (yellow) and \*04:04 (blue) at P1-Tyr of TNC<sup>1014,1016</sup>cit peptide. E, overlaid polymorphism at residue 13 of HLA-DRB1\*04:01 (yellow), \*01:01 (light pink), and \*14:02 (teal) and impact of interactions with P4-P6 of TNC<sup>1014,1016</sup>cit peptide. F, binding affinity of PB TCR for HLA-DRB1\*04:05-TNC<sup>1014,1016</sup>cit. HLA-



## T cell recognition of HLA-DR4 presenting cit-tenascin

Lys<sup>71</sup> in HLA-DRB1\*04:01 are substituted by Ser<sup>57</sup> and Arg<sup>71</sup> in DRB1\*04:05 allele, respectively (Fig. 7B). The Ser<sup>57</sup> substitution in HLA-DRB1\*04:05 causes the loss of contact with P9-Thr of TNC<sup>1014,1016cit</sup> peptide, likely affects the P9 anchoring, which is implicated in PB TCR docking (Fig. 7G). This is consistent with our FP assay result of reduced binding strength (IC<sub>50</sub>) for DRB1\*04:05<sup>TNC1014,1016cit</sup> as compared to DRB1\*04:01<sup>TNC1014,1016cit</sup>. Collectively, the PB TCR restricted to the HLA-DRB1\*04:01, the strongest genetic RA risk factor in Europeans, can cross-react to highly conserved and susceptible RA allomorphs, namely, HLA-DRB1\*04:05 in Asians, albeit with weaker affinity.

### Discussion

Post-translational modification (PTM) of peptide antigens can confer the ability to bind to MHC II and have been implicated in immune disorders, as exemplified in citrullination in RA (14), deamidation of glutamine in coeliac disease (35), and peptide trans-splicing in type I diabetes (36, 37). In the context of RA, we have previously described the molecular basis for the specific TCR recognition of cit-fibrinogen (26), cit-vimentin (27), and cit- $\alpha$ -enolase (27) epitopes. We showed that citrullination at position P4 is not only critical in conferring the ability to occupy the P4 pocket of the SE in the HLA-DRB1\*04:01 but also has direct contact in TCR recognition (26, 27). In contrast, an additional citrullination at P2 of fibrinogen  $\beta^{72,74cit69-81}$  peptide had weaker binding to fibrinogen  $\beta^{74cit69-81}$ -restricted TCRs, resulting in an altered TCR repertoire in immunized mice (26). In the present study, we focus on cit-tenascin-C peptide (TNC<sup>1014,1016cit</sup>), which contains two citrullination sites at positions P-1 and P2 beyond the P4-SE anchor, to further understand the impact of citrullination in TCR recognition. The crystal structure of PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> revealed that while P5-Tyr is a distinct feature for TNC<sup>1013-24</sup> peptide, citrullines at P-1 and P2 are the key determinants for PB TCR recognition. The synergistic effect of P-1 and P2 citrullines of TNC<sup>1014,1016cit</sup> peptide confers high-binding affinity to PB TCR, suggesting the complementary role of multi-citrullination in TCR recognition. This is consistent with the SPR result, where single citrullination at the P-1 or P2 position of TNC<sup>1013-24</sup> peptide either led to reduced affinity or completely abolished the interaction with PB TCR, respectively.

The lack of cross-reactivity observed for PB TCR towards other cit-RA autoantigens, including fibrinogen  $\beta^{74cit69-81}$ , vimentin<sup>64cit59-71</sup>, and  $\alpha$ -enolase<sup>15cit10-22</sup>, highlighted the strict specificity of PB TCR in TNC<sup>1014,1016cit</sup> recognition, consistent with previously reported analysis (26, 27). Despite a highly conserved peptide binding register at P1, P4, and P6 across four different peptides, the distinct characteristics at positions P2 and P5 of cit-epitopes will discriminate against another cit-

epitope restricted TCRs. For instance, in P5, the bulky aromatic feature of P5-Tyr in the TNC<sup>1014,1016cit</sup> peptide, in comparison to small polar residues of P5-Ser of vimentin<sup>64cit59-71</sup>, P5-Pro of fibrinogen  $\beta^{74cit69-81}$ , and P5-Gly of the  $\alpha$ -enolase<sup>15cit10-22</sup> peptide. In P2, there is a neutral P2-Cit in TNC<sup>1014,1016cit</sup> peptide, as opposed to small residue P2-Ala, positively charged P2-Arg, and negatively charged P2-Asp in vimentin<sup>64cit59-71</sup>, fibrinogen  $\beta^{74cit69-81}$ , and  $\alpha$ -enolase<sup>15cit10-22</sup> epitopes, respectively. These features indicate that the likelihood of cross-reactivity between cit-epitope-restricted TCRs is unlikely. Although TCR cross-reactivity has been reported in other autoimmune disorders such as coeliac disease (38) and Type 1 diabetes (37); nevertheless, these antigens involved shared high sequence homology. In contrast, the immunodominant epitopes recognized by TCRs in RA are peptide antigens with very diverse and distinct features that originate from different tissues. It is therefore not surprising that little or no cross-reactivity was observed between different cit-epitopes. Nevertheless, we observed constant duality recognition patterns of SE-TNC<sup>1014,1016cit</sup> peptide and PB TCR interaction, consistent with our previously described fibrinogen  $\beta^{74cit69-81}$ , vimentin<sup>64cit59-71</sup>, and  $\alpha$ -enolase<sup>15cit10-22</sup> TCR-pMHC II complexes. This was the case with either a neutral asparagine or a neutral citrulline at P4, thereby affirming the role of SE<sup>+</sup>-HLA in TCR recognition (26, 27).

Moreover, we demonstrated the capability of the TRAV35<sup>+</sup>/TRBV10-2<sup>+</sup> PB TCR to cross-react with the highly homologous HLA-DRB1\*04:05 allomorph, albeit with weaker affinity. The distinct polymorphism underlying the peptide-binding groove is located at P9 pocket, with a Ser<sup>57</sup> residue in HLA-DRB1\*04:05, as opposed to Asp<sup>57</sup> in HLA-DRB1\*04:01. This residue at position 57 was reported to have a susceptibility effect which accounts for the detrimental association between the SE and joint destruction in Japanese patients with ACPA-positive RA (16). Furthermore, although the PB TCR did not recognize citrullinated TNC<sup>1014,1016cit</sup> peptide presented by HLA-DRB1\*01:01 or \*14:02, there is certainly a potential that additional TCRs may have the capacity to recognize this peptide also in the context of those alleles. The deviation at residue 13 of HLA-DRB1\*04:01, \*01:01, and \*14:02 that affects the binding preferences of PB TCR, likely contributes to odds ratio of RA pathogenesis, consistent with reported study of residue 13 is as important as the SE in associated with seropositive RA (15). Overall, our study has provided a structural insight into how citrullination shapes specific CD4<sup>+</sup> T cell recognition in RA.

### Experimental procedures

#### Peptide

Peptides including TNC<sup>1014,1016cit</sup> (<sup>1013</sup>DcitY-citLNYS<sup>1024</sup>PTG<sup>1024</sup>, where cit represents citrulline residue),

DRB1\*04:01<sup>CLIP</sup> was immobilized in the reference flow cell to control the non-specific binding. PB TCR equilibrium affinity constants (K<sub>D</sub>) value was determined from three independent experiments in duplicate. For each concentration, the point represents the mean value, and the error bar corresponds to  $\pm$  s.e.m. G, the D57S polymorphism at the P9 binding pocket of HLA-DRB1\*04:01 and \*04:05 (pale cyan) and the interaction with P9-Thr of TNC<sup>1014,1016cit</sup>. Black dashes represent H-bond within 3.5 Å, beige dashes correspond to VdW interaction within 4 Å. All amino acids are indicated in single-letter abbreviations.

native TNC<sup>1013-24</sup> (<sup>1013</sup>DRYRLNYSLPTG<sup>1024</sup>), TNC<sup>1014cit</sup> (<sup>1013</sup>DcitYRLNYSLPTG<sup>1024</sup>), TNC<sup>1016cit</sup> (<sup>1013</sup>DRY-citLNYSLPTG<sup>1024</sup>), Vimentin<sup>64cit59-71</sup> (<sup>59</sup>GVYATcitS-SAVRLR<sup>71</sup>),  $\alpha$ -enolase<sup>15cit10-22V20G</sup> (<sup>10</sup>EIFDScitGNPTGEV<sup>22</sup>), Fibrinogen  $\beta$ <sup>74cit69-81</sup> (<sup>69</sup>GGYRACitPAKAAAT<sup>81</sup>), and HA<sup>306-318</sup> (<sup>306</sup>PKYVKQNTLKLAT<sup>318</sup>) were synthesized by GL Biochem. The integrity of the peptides was verified by reverse-phase high-performance liquid chromatography and mass spectrometry.

### Protein expression and purification

The TCR  $\alpha$ - and  $\beta$ -chains were designed and expressed as previously described (26, 39). In brief, the extracellular domains of TCR  $\alpha$ - and  $\beta$ -chains were engineered with a disulfide linkage in the constant domains to stabilize the heterodimer. TCR  $\alpha$ - and  $\beta$ -chains were then expressed independently as inclusion bodies in *Escherichia coli* BL21 (DE3) and, subsequently, refolded in a buffer containing 5 M Urea, 100 mM Tris pH 8.0, 0.4 M L-Arginine, 2 mM EDTA, 0.2 mM phenylmethylsulfonyl fluoride, 0.5 mM oxidized glutathione, and 5 mM reduced glutathione for 72 h at 4 °C with rapid stirring. The refolded samples were dialyzed with 10 mM Tris pH 8.0 and purified on a DEAE (Cytiva) anion exchange column, followed by size exclusion (HiLoad 16/600 Superdex 200pg column; Cytiva), hydrophobic interaction (HiTrap Phenyl HP column; Cytiva), and anion exchange (HiTrap TM Q HP column; Cytiva) chromatography.

The expression of HLA-DRB1\*04:01 was performed as described in previous article (27). Briefly, the extracellular domains of the  $\alpha$ -chain and  $\beta$ -chain of HLA-DRB1\*04:01 (*HLA-DRA*\*01:01 and *HLA-DRB1*\*04:01) were covalently linked to invariant chain (CLIP) peptide, and cloned into the lentiviral vectors, namely, pLV-EF1 $\alpha$ -MCS-IRES-GFP and pLV-EF1 $\alpha$ -MCS-IRES-RFP (Biosettia), respectively. The HLA-DRB1\*04:01 lentivirus was produced by co-transfection of these vectors, along with viral packaging plasmids (pMD2.G, pMDLg/pRRE, pRSV-REV; Addgene), in the human embryonic kidney (HEK) 293 T cells. The HLA-DRB1\*04:01 lentivirus was harvested and transduced into glycosylation deficiency HEK293S (GnT1<sup>-</sup>) (CRL-3022, ATCC) cells and subsequently sorted by single-cell FACS (Becton Dickinson) to generate a cell line that stably expresses HLA-DRB1\*04:01. To produce HLA-DRB1\*04:01 protein, stably expressed clones were cultured in Expi293 Expression Medium (serum free media; Gibco, Thermo Fisher Scientific) in shaking incubator at 37 °C in 5% CO<sub>2</sub>. HLA-DRB1\*04:01 protein was then harvested and purified as previous described (13). Briefly, the supernatant containing soluble HLA-DRB1\*04:01 protein was concentrated and dialysed to 10 mM Tris pH 8.0 and 150 mM NaCl using tangential flow filtration (TFF) on a Cogent M1 TFF system (Merck Millipore), followed by subsequent purification *via* immobilized metal ion affinity (Nickel-Sepharose 6 Fast Flow; Cytiva), and size exclusion (Superdex 200, 16/600; Cytiva) chromatography.

The construct design and expression of other HLA-DRB1 proteins (*HLA-DRA*1\*01:01, \*04:04, \*04:05, \*01:01 or \*14:02)

were as previously described (14). The C-terminus of the DRB1  $\alpha$ -chain had a Fos leucine zipper, and the  $\beta$ -chain had a Jun leucine zipper, followed by a BirA biotin ligase biotinylation recognition sequence and a polyhistidine tag. The N-terminus of the  $\beta$ -chain was covalently linked to a factor Xa-cleavable Strep-tag invariant chain (CLIP) peptide. The extracellular domains of  $\alpha$ - and  $\beta$ -chains were independently cloned into the pHLsec vector, transfected using polyethyleneimine (PEI) (BioScientific) at a ratio of 1:3 of DNA to PEI. The transfected cells were incubated at 37 °C with 5% CO<sub>2</sub> in a 120-rpm shaker incubator for a week. The soluble recombinant HLA-DRB1 proteins were then purified from the cell culture supernatant as stated above. Purified monomeric peptide-HLA-DRB1 was biotinylated using biotin protein ligase (BirA) in buffer containing 0.05 M bicine pH 8.3, 0.01 mM ATP, 0.01 mM MgOAc, 50  $\mu$ M d-biotin, and 2.5  $\mu$ g BirA. BirA was made according to protocols outlined in O'Callaghan C *et al.* (40).

### Peptide loading of HLA-DRB1

The HLA-DRB1 proteins presenting CLIP peptide were digested with Factor Xa (New England Biolabs) to cleave the covalently linked Strep-CLIP in TBS150 buffer (10 mM Tris pH 8.0, 150 mM NaCl) containing 2 mM CaCl<sub>2</sub> for 6 h at room temperature. 5 mM EDTA was used to stop the enzymatic reaction. The cleaved HLA-DRB1 was subsequently loaded with 20 M excess of peptide in 50 mM trisodium citrate pH 5.4 in the presence of HLA-DM at a molar ratio of 5:1 and incubated for 72 h at 37 °C. The peptide-loaded HLA-DRB1 was passed through a Strep-Tactin Sepharose (IBA) column to remove the partially digested or unloaded HLA-DRB1-Strep-CLIP.

### In vitro TCR expression and tetramer staining

HEK 293T cells (ATCC, #CRL-3216) were plated at 3.5 x 10<sup>5</sup> cells/well of a six well plate in 3 ml RF10 media containing RPMI-1640, 10% fetal bovine serum (FBS, Sigma), glutamax (Gibco, #35050061), Non-essential amino acid (Gibco, #11140050), HEPES (Gibco, #15630130), sodium pyruvate (Gibco, #11360070), penicillin-streptomycin (Gibco, #15070063), 50  $\mu$ M 2-mercaptoethanol (Merck), for 24h at 37 °C, 5% CO<sub>2</sub>. 420 ng of individual lentiviral vector pLV-EF1 $\alpha$ -MCS-IRES-GFP encoding TCR  $\alpha$ -chain and pLV-EF1 $\alpha$ -MCS-IRES-RFP (Biosettia) encoding TCR  $\beta$ -chain were transiently expressed together with the pLV encoding CD3 $\gamma\delta\epsilon\zeta$  subunits in HEK 293T cells using FuGene 6 HD (Promega, #E2691). On the following day, transfected HEK 293T cells were detached and repeatedly washed with FACS buffer (PBS + 2% FBS) by centrifugation at 350g for 5 min, prior being labelled with 0.5  $\mu$ g of individual peptide-loaded HLA-DRB1 tetramer for 1 h in dark at room temperature. Cells were then stained with 1:100 diluted BUV395 mouse anti-human CD3 antibody (clone UCHT1, BD Biosciences) for 1 h in dark at 4 °C, washed three times with FACS buffer, followed by live/dead cell staining with 1:10,000 diluted 4',6-diamidino-2-phenylindole (DAPI; BD Biosciences) viability stain for 15 min before

## T cell recognition of HLA-DR4 presenting cit-tenascin

being analyzed on a BD LSRFortessa X-20 with FACSDiva 8.0.1 software (BD Immunocytometry Systems). Three independent experiments were conducted for all tetramer staining analysis. Collected data were analyzed using FlowJo v10.9.0 (FlowJo).

### Surface plasmon resonance

The affinity measurements were performed using surface plasmon resonance on a Biacore T200 instrument (Cytiva). Approximately 3000 response units (RU) of biotinylated peptide-loaded HLA-DRB1 were immobilized on a streptavidin (SA) sensor chip (Cytiva). HLA-DRB1\*04:01<sup>CLIP</sup> was immobilized in the reference flow cell and acted as negative control. Serial dilutions of TCRs were passed over the flow cell's surface at the rate of 10  $\mu$ l/min in 20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, and 0.005% v/v surfactant P20 (Cytiva). Three independent experiments in duplicate were performed for PB TCR, and two independent experiments in duplicate were performed for PB TCR mutants. Collected data were analyzed on Prism 10 (GraphPad Software, version 10.2.0) using a one-site specific binding model and plotted as sensorgrams and equilibrium response curves. M134 TCR (26) and A03 TCR (27) were used as negative control TCRs.

### T cell stimulation assay

The PB TCR  $\alpha$ - and  $\beta$ -chains were cloned into pLV-EF1 $\alpha$ -MCS-IRES-GFP and pLV-EF1 $\alpha$ -MCS-IRES-RFP, respectively, and subsequently transduced into the SKW3 T cell line (TCR deficient; German Collection of Microorganisms and Cell Cultures) for stable expression using the lentiviral transduction system as previously described (41). PB TCR-transduced SKW3 T cells were cultured in RF10 media at 37 °C in 5% CO<sub>2</sub>. Briefly, approximately 0.1  $\times$  10<sup>6</sup> BLCL 9031 cells (HLA-DRA1\*01:01, HLA-DRB1\*04:01; sourced from The International Histocompatibility Working Group (IHWG) Cell and DNA Bank) acting as antigen presenting cells, were incubated with serial dilutions of TNC<sup>1014,1016cit</sup> peptide (5:1 dilution starting at 5  $\mu$ g/ml), 50  $\mu$ g of  $\alpha$ -enolase<sup>15cit10-22</sup> peptide (negative control), or 35  $\mu$ g/ml of anti-HLA-DR monoclonal antibody (clone LB3.1, blocking antibody; negative control) in 96-well round-bottom plate (Corning) for 4 h at 37 °C, 5% CO<sub>2</sub>. Subsequently, 1  $\times$  10<sup>5</sup> PB TCR transduced SKW3 T-cells, RA2.7 TCR transduced SKW3 T-cells (control) or untransduced SKW3 parental cells were added to the wells accordingly and incubated overnight at 37 °C, 5% CO<sub>2</sub>. On the following day, the cells were then washed twice with FACS buffer, then stained with a mixture of 1:100 diluted BUV395-mouse anti-human CD3 (clone UCHT1, BD Biosciences), and APC mouse anti-human CD69 (clone FN50, BD Biosciences) for 1 h at 4 °C in dark. Cells were then washed 6 times with FACS buffer to remove excess antibodies, followed by live/dead staining with DAPI (BD Biosciences) at 1:10,000 ratio for 15 min. Subsequently, cells were analyzed *via* flow cytometry using FACSDiva 8.0.1 software on the BD LSRFortessa X-20 (BD Immunocytometry System). Three independent experiments were conducted, and all samples were performed in

duplicate. Collected data were analyzed using FlowJo v10.9.0 and plotted with Prism 10 (GraphPad Software, version 10.2.0). One-way ANOVA multiple comparison with Dunnett's multiple comparison testing was used to determine the statistical significance between the MFI values of unstimulated SKW3 PB T-cells *versus* the peptide-stimulated SKW3 PB T-cells.

### Crystallization, data collection, and processing

For crystallization, the monomeric TNC<sup>1014,1016cit</sup> loaded HLA-DRB1\*04:01 was subjected to HRV 3C protease to remove C-terminal Fos/Jun leucine zipper tagging. For the ternary complex, HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> was mixed with PB TCR at 1:1 M ratio and incubated for 6 h at room temperature. Proteins were concentrated up to 10 mg/ml and undertook high-throughput crystallization screening at the Monash Molecular Crystallization Platform (MMCP) using an automated robotic NT8 system. The HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> binary complex was crystallized in reservoir solution containing 0.1 M Bis-Tris pH 5.5, 0.2 M NaCl and 29% w/v PEG3350; the PB TCR – HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> ternary complex was crystallized in 0.1 M Tris pH 7.5, 0.3 M NaCl, 0.05 M Glutamic Acid, 0.05 M Arginine and 20% w/v PEG3350; the apo PB TCR was crystallized in 0.1 M Na Acetate pH 7.8 and 8% w/v PEG 4K. Single crystals were treated with mother liquor containing cryoprotectant (15–25% glycerol or ethylene glycol) prior to flash freezing in liquid nitrogen. Diffraction data were collected at the Australian Synchrotron's MX2 beamline, auto-processed and scaled with XDS and CCP4 Software Suite version 8.0.

### Structure determination, refinement, and validation

Crystal structures of PB TCR-HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> ternary complex, PB TCR apo form, and HLA-DRB1\*04:01<sup>TNC1014,1016cit</sup> binary were solved by molecular replacement in PHASER (CCP4 Software Suite, version 8.0) using a separate search model for HLA-DRB1\*04:01 and TCR (PDB ID: 6V1A) (26). Multiple rounds of model building in Coot (42) and automated refinement using Phenix.refine (PHENIX) (43). The quality of the structures was validated at the Protein Data Bank (PDB) validation and deposition server. The PB TCR structure was numbered according to the IMGT unique numbering system (44). Data processing and refinement statistics were summarized in Table 1. Ramachandran statistics of final models revealed ~ 95%-97% of residues were in favored regions, with no outlier residue. Buried surface area and TCR-pHLA contact analyses were determined using the program Areaimol and Contact in CCP4 Program Suite, respectively. PyMOL (version 2.5.2) was used to generate all structural figures.

### Fluorescence polarization assay

The relative binding strength of TNC<sup>1014,1016cit</sup> peptide for HLA-DRB1\*04:01, \*01:01, \*04:04, \*04:05 and \*14:02 was determined through the fluorescence polarization assay, as described previously (14, 45). In brief, serial dilution of peptide, starting from 500  $\mu$ M was incubated in competition with



20 nM TAMRA-HA fluorescent labeling peptide, to bind with 100 nM HLA-DRB1 protein in the presence of 20 nM HLA-DM, in the buffer comprising of 100 mM trisodium citrate pH 5.4, 50 mM NaCl and 5 mM EDTA. The fluorescent polarization was measured by PHERAstar microplate reader (BMG LABTECH) after 24 h, 48 h and 72 h incubation at 37 °C. The peptide binding curves were plotted by non-linear regression in Prism 10 (GraphPad Software, version 10.2.0) using a sigmoidal dose-response curve. IC<sub>50</sub> values were calculated as the peptide concentration required for 50% inhibition of TAMRA-HA fluorescent labeling peptide binding to HLA-DRB1 protein. All data were derived from two independent experiments in triplicate.

## Data availability

The X-ray crystal structures were deposited in the Protein Data Bank (PDB) with the following accession codes: DRB1\*04:01<sup>TNC1014</sup>, <sup>1016cit</sup>, 9NIH; PB TCR-DRB1\*04:01<sup>TNC1014,1016cit</sup>, 9NIG; PB TCR, 9NIL.

**Supporting information**—This article contains supporting information.

**Acknowledgments**—We thank the staff at the Australian Synchrotron MX beamline for assistance with data collection, Monash Macromolecular crystallization facility and Monash FlowCore. J. R. supported by an NHMRC Leadership Investigator Fellowship (2008981).

**Author contributions**—J. R., H. H. R., V. M., L. K., T. J. L., H. T. D., and J. J. L. writing—review & editing; J. R. validation; J. R., H. H. R., V. M., and L. K. supervision; J. R., V. M., and L. K. project administration; J. R., V. M., and L. K. funding acquisition; J. R. and J. J. L. conceptualization; L. K. resources; R. K. S. and H. T. D. investigation; T. J. L. and H. T. D. formal analysis. H. T. D. and J. J. L. writing—original draft.

**Conflict of interests**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: CILP, cartilage intermediate layer protein; cit, citrullinated; HLA, human leukocyte antigen; PAD, peptidyl arginine deiminases; PTM, post-translational modification; RA, rheumatoid arthritis; SE, susceptibility epitope; TCR, T cell receptor; TFF, tangential flow filtration.

## References

- Almutairi, K., Nossent, J., Preen, D., Keen, H., and Inderjeeth, C. (2021) The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatol. Int.* **41**, 863–877
- Health, A. I. o., and Welfare (2024) *Rheumatoid Arthritis*, AIHW, Canberra
- Berglin, E., Padyukov, L., Sundin, U., Hallmans, G., Stenlund, H., Van Venrooij, W. J., et al. (2004) A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis. *Arthritis Res. Ther.* **6**, 1–6
- Kroot, E. J. J., De Jong, B. A., Van Leeuwen, M. A., Swinkels, H., Van Den Hoogen, F. H., Van't Hof, M., et al. (2000) The prognostic value of anti-

- cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. *Arthritis. Rheum.* **43**, 1831–1835
- Huizinga, T. W., Amos, C. I., van der Helm-van Mil, A. H., Chen, W., Van Gaalen, F. A., Jawaheer, D., et al. (2005) Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. *Arthritis. Rheum.* **52**, 3433–3438
- Derksen, V., Huizinga, T., and Van Der Woude, D. (2017) The role of autoantibodies in the pathophysiology of rheumatoid arthritis. In *Seminars in Immunopathology*. Springer, Berlin, Germany
- Darrah, E., and Andrade, F. (2018) Rheumatoid arthritis and citrullination. *Curr. Opin. Rheumatol.* **30**, 72–78
- Kurowska, W., Kuca-Warnawin, E. H., Radzikowska, A., and Maslinski, W. (2017) The role of anti-citrullinated protein antibodies (ACPA) in the pathogenesis of rheumatoid arthritis. *Cent. Eur. J. Immunol.* **42**, 390–398
- Klareskog, L., Stolt, P., Lundberg, K., Kallberg, H., Bengtsson, C., Grunewald, J., et al. (2006) A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis. Rheum.* **54**, 38–46
- Foulquier, C., Sebbag, M., Clavel, C., Chapuy-Regaud, S., Al Badine, R., Méchin, M. C., et al. (2007) Peptidyl arginine deiminase type 2 (PAD-2) and PAD-4 but not PAD-1, PAD-3, and PAD-6 are expressed in rheumatoid arthritis synovium in close association with tissue inflammation. *Arthritis. Rheum.* **56**, 3541–3553
- Trouw, L. A., Rispens, T., and Toes, R. E. (2017) Beyond citrullination: other post-translational protein modifications in rheumatoid arthritis. *Nat. Rev. Rheumatol.* **13**, 331–339
- Gregersen, P. K., Silver, J., and Winchester, R. J. (1987) The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis. Rheum.* **30**, 1205–1213
- Scally, S. W., Petersen, J., Law, S. C., Dudek, N. L., Nel, H. J., Loh, K. L., et al. (2013) A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J. Exp. Med.* **210**, 2569–2582
- Ting, Y. T., Petersen, J., Ramarathinam, S. H., Scally, S. W., Loh, K. L., Thomas, R., et al. (2018) The interplay between citrullination and HLA-DRB1 polymorphism in shaping peptide binding hierarchies in rheumatoid arthritis. *J. Biol. Chem.* **293**, 3236–3251
- Raychaudhuri, S., Sandor, C., Stahl, E. A., Freudenberg, J., Lee, H.-S., Jia, X., et al. (2012) Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat. Genet.* **44**, 291–296
- Terao, C., Yano, K., Ikari, K., Furu, M., Yamakawa, N., Yoshida, S., et al. (2015) Brief report: main contribution of DRB1\*04:05 among the shared epitope alleles and involvement of DRB1 amino acid position 57 in association with joint destruction in anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis. Rheumatol.* **67**, 1744–1750
- Ferucci, E. D., Templin, D. W., and Lanier, A. P. (2005) Rheumatoid arthritis in American Indians and Alaska Natives: a review of the literature. *Semin. Arthritis. Rheum.* **34**, 662–667
- Burkhardt, H., Sehnert, B., Bockermann, R., Engström, Å., Kalden, J. R., and Holmdahl, R. (2005) Humoral immune response to citrullinated collagen type II determinants in early rheumatoid arthritis. *Eur. J. Immunol.* **35**, 1643–1652
- Masson-Bessiere, C., Sebbag, M., Girbal-Neuhausser, E., Nogueira, L., Vincent, C., Senshu, T., et al. (2001) The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the  $\alpha$ - and  $\beta$ -chains of fibrin. *J. Immunol.* **166**, 4177–4184
- Raza, K., Schwenzer, A., Juarez, M., Venables, P., Filer, A., Buckley, C. D., et al. (2016) Detection of antibodies to citrullinated tenascin-C in patients with early synovitis is associated with the development of rheumatoid arthritis. *RMD Open* **2**, e000318
- Song, J., Schwenzer, A., Wong, A., Turcinov, S., Rims, C., Martinez, L. R., et al. (2021) Shared recognition of citrullinated tenascin-C peptides by T and B cells in rheumatoid arthritis. *JCI Insight* **6**
- der Heijden, V. E. N. E., and ALora, M. (2004) Rheumatoid arthritis specific anti-Sa antibodies target citrullinated vimentin. *Arthritis. Res. Ther.* **6**, R142150

23. Snir, O., Rieck, M., Gebe, J. A., Yue, B. B., Rawlings, C. A., Nepom, G., *et al.* (2011) Identification and functional characterization of T cells reactive to citrullinated vimentin in HLA-DRB1\* 0401-positive humanized mice and rheumatoid arthritis patients. *Arthritis. Rheum.* **63**, 2873–2883
24. Kinloch, A., Tatzert, V., Wait, R., Peston, D., Lundberg, K., Donatien, P., *et al.* (2005) Identification of citrullinated  $\alpha$ -enolase as a candidate autoantigen in rheumatoid arthritis. *Arthritis. Res. Ther.* **7**, 1–9
25. Hill, J. A., Bell, D. A., Brintnell, W., Yue, D., Wehrli, B., Jevnikar, A. M., *et al.* (2008) Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. *J. Exp. Med.* **205**, 967–979
26. Lim, J. J., Jones, C. M., Loh, T. J., Ting, Y. T., Zareie, P., Loh, K. L., *et al.* (2021) The shared susceptibility epitope of HLA-DR4 binds citrullinated self-antigens and the TCR. *Sci. Immunol.* **6**, eabe0896
27. Loh, T. J., Lim, J. J., Jones, C. M., Dao, H. T., Tran, M. T., Baker, D. G., *et al.* (2024) The molecular basis underlying T cell specificity towards citrullinated epitopes presented by HLA-DR4. *Nat. Commun.* **15**, 6201
28. Feitsma, A. L., van der Voort, E. I., Franken, K. L., el Bannoudi, H., Elferink, B. G., Drijfhout, J. W., *et al.* (2010) Identification of citrullinated vimentin peptides as T cell epitopes in HLA-DR4-positive patients with rheumatoid arthritis. *Arthritis. Rheum.* **62**, 117–125
29. Sharma, R. K., Boddul, S. V., Yoosuf, N., Turcinov, S., Dubnovitsky, A., Kozhukh, G., *et al.* (2021) Biased TCR gene usage in citrullinated Tenascin C specific T-cells in rheumatoid arthritis. *Sci. Rep.* **11**, 24512
30. James, E. A., Rieck, M., Pieper, J., Gebe, J. A., Yue, B. B., Tatum, M., *et al.* (2014) Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis. Rheumatol.* **66**, 1712–1722
31. Pieper, J., Dubnovitsky, A., Gerstner, C., James, E. A., Rieck, M., Kozhukh, G., *et al.* (2018) Memory T cells specific to citrullinated alpha-enolase are enriched in the rheumatic joint. *J. Autoimmun.* **92**, 47–56
32. Law, S. C., Street, S., Yu, C.-H. A., Capini, C., Ramnouruth, S., Nel, H. J., *et al.* (2012) T-cell autoreactivity to citrullinated autoantigenic peptides in rheumatoid arthritis patients carrying HLA-DRB1 shared epitope alleles. *Arthritis. Res. Ther.* **14**, 1–12
33. Rossjohn, J., Gras, S., Miles, J. J., Turner, S. J., Godfrey, D. I., and McCluskey, J. (2015) T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200
34. Tadros, D. M., Eggenschwiler, S., Racle, J., and Gfeller, D. (2023) The MHC Motif Atlas: a database of MHC binding specificities and ligands. *Nucleic. Acids. Res.* **51**, D428–D437
35. Petersen, J., Montserrat, V., Mujico, J. R., Loh, K. L., Beringer, D. X., van Lummel, M., *et al.* (2014) T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. *Nat. Struct. Mol. Biol.* **21**, 480–488
36. Tran, M. T., Faridi, P., Lim, J. J., Ting, Y. T., Onwukwe, G., Bhattacharjee, P., *et al.* (2021) T cell receptor recognition of hybrid insulin peptides bound to HLA-DQ8. *Nat. Commun.* **12**, 5110
37. Tran, M. T., Lim, J. J., Loh, T. J., Mannering, S. I., Rossjohn, J., and Reid, H. H. (2024) A structural basis of T cell cross-reactivity to native and spliced self-antigens presented by HLA-DQ8. *J. Biol. Chem.* **300**, 107612
38. Ciacchi, L., Farenc, C., Dahal-Koirala, S., Petersen, J., Sollid, L. M., Reid, H. H., *et al.* (2022) Structural basis of T cell receptor specificity and cross-reactivity of two HLA-DQ2.5-restricted gluten epitopes in celiac disease. *J. Biol. Chem.* **298**, 101619
39. Boulter, J. M., Glick, M., Todorov, P. T., Baston, E., Sami, M., Rizkallah, P., *et al.* (2003) Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein. Eng.* **16**, 707–711
40. O'Callaghan, C., A., Byford, M. F., Wyer, J. R., Willcox, B. E., Jakobsen, B. K., McMichael, A. J., *et al.* (1999) BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation. *Anal. Biochem.* **266**, 9–15
41. Lim, J. J., Jones, C. M., Loh, T. J., Dao, H. T., Tran, M. T., Tye-Din, J. A., *et al.* (2025) A naturally selected alphabeta T cell receptor binds HLA-DQ2 molecules without co-contacting the presented peptide. *Nat. Commun.* **16**, 3330
42. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta. Crystallogr. D. Biol. Crystallogr.* **66**, 486–501
43. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta. Crystallogr. D. Biol. Crystallogr.* **66**, 213–221
44. Lefranc, M. P., Pommie, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., *et al.* (2003) IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* **27**, 55–77
45. Yin, L., and Stern, L. J. (2014) Measurement of peptide binding to MHC class II molecules by fluorescence polarization. *Curr. Protoc. Immunol.* **106**, 5