Title Page

Full title:

CDR3 binding chemistry controls TCR V-domain rotational probability and germline CDR2 'scanning' of polymorphic MHC

Running title:

Genetic calculus of TCR binding class-II pMHC

Joseph S. Murray†

jsmurray@xenolausgenetics.com

†***Xenolaüs Genetics LLC***
Los Angeles, CA, 90024, USA

CDR3 binding chemistry controls TCR V-domain rotational probability and germline CDR2 scanning of polymorphic MHC

Joseph S. Murray†*

†***Xenolaüs Genetics LLC***
Los Angeles, CA UNITED STATES

**Email Author: Dr. J. S. Murray, Ph.D. (jsmurray@xenolausgenetics.com)*

Abstract:

The mechanism which adapts the T-cell antigen receptor (TCR) within a given major histocompatibility complex (*MHC/HLA*) genotype is essential for protection against pathogens. Historically attributed to relative affinity, genetically vast TCRs are surprisingly focused towards a micromolar affinity for their respective peptide (p) plus MHC (pMHC) ligands. Thus, the somatic diversity of the TCR with respect to MHC-restriction, and (ultimately) to pathogens, remains enigmatic. Here, we derive a triple integral solution (from fixed geometry) for any given V domain in TCR bound to pMHC. Solved complexes involving HLA-DR and HLA-DQ, where genetic linkage to the TCR is most profound, were examined in detail. Certain V domains displayed rare geometry within this panel—specifying a restricted rotational probability/volumetric density (*dV*). Remarkably, hydrogen (H) bond charge-relays distinguished these structures from the others; suggesting that CDR3 binding chemistry dictates CDR2 contacts on the opposite MHC-II alpha helix. Together, these data suggest that TCR recapitulate *dV* and specialise target pMHC recognition.

Keywords:

T-cell receptor (TCR); Major Histocompatibility Complex (MHC); Clonal Selection; MHC Restriction; Protein Chemistry; Mathematical Biology

Introduction:

T-cell antigen receptors (TCR) and antibodies (sIg) are individualized within each precursor of a given T-cell or B-cell clone. TCR (α/β type) have genetically variable (V) domains, wherein *complementary determining regions* (CDR1–CDR3) contain closest amino acid (a.a.) contacts with the peptide (p) plus *major histocompatibility complex* (MHC) protein (together abbrev., pMHC) composite ligand. CDR1 and CDR2 are encoded in the germline, via the particular V-region DNA segment involved in the *RAG1/RAG2* recombination mechanism responsible for somatic construction, together with the *D*- and/or *J*-segment(s), of the third, CDR3 [1–5]. This paper seeks to understand two puzzles of the TCR-pMHC interaction, wherein a novel examination of the first can be used to re-examine the second in the context of existing

evidence. Firstly, compared to antibodies and indeed other protein:protein binding reactions, TCRs display quite low (μM) binding affinities for pMHC ligands [6–8]. Nevertheless, like antibodies, TCRs have exquisite specificity, where single peptide or MHC a.a. changes can dictate T-cell reactivity or non-reactivity [9–17]. Second, the nature of T-cell selection in the thymus prior to seeding secondary lymphoid tissues is thought to operate on the basis of this same relative TCR:pMHC affinity, where overt 'self' peptide plus self-MHC recognition (*negative selection*) leads to deletion from the repertoire; as does a complete lack of recognition (*positive selection*) [4]. This intermediate TCR:pMHC binding in the fetal thymus forms the basis of 'simultaneous' self plus pathogen recognition in adulthood, *i.e.*, the *MHC-restricted* adaptive immune response [5, 18]. Somatic CDR3 loops are roughly positioned atop the peptide component in solved TCR:pMHC complexes [9–18]; although more recent investigations have challenged the notion that the most diverse component of the ligand (the peptide) is particularly interfaced with the most diverse regions of the receptor (CDR3α/β) [12, 19].

For example, we previously reported that one of the CDR3 loops (CDR3α) also makes consistently close contacts with a conserved HLA-A component, *i.e.*, the *R65-motif* (a.a., R65-X-X-K68-A69-X-S71-Q72) [20]. *Burrows et al.*, also found conserved CDR3 contacts directly with MHC-I [21]. More recently, *Sharon et al.,* formally linked germline TCR usage to HLA, particularly via the class-II alleles [18]. Here, the *STCR-Dab* (www.opig.stats.ox.ac.uk/webapps/stcrdab/) was used to identify all current TCR complexes involving HLA-DR and HLA-DQ (3.0 Å resolution cut-off). Note, also included are *PDB* 4H1L, at 3.3 Å, and 3T0E, at 4.0 Å (see below). A simple relationship between the observed (*i.e.*, measured) CDR3–CDR2 "pitch" angle for a given V domain and its predicted (*i.e.*, calculated) CDR3–CDR2 pitch angle was observed [20]. Here, a linear relationship between calculated pitch and a new measure, "*dV*"—by multivariable calculus, was found. As is discussed, the simple volume element, *dV*, interprets V-domain orientation into a rotational probability involving an apparent CDR2–MHC α-helix scanning function (*d*θ). While the *dV* was unique for each TCR on each pHLA-DR, one TCR displayed a dramatic restriction in *d*V for Vβ. This was isolated to a charge-relay H-bonding mechanism for CDR3β; hence, the chemistry of "somatic-TCR" dictates positioning of "germline-TCR" (see **Discussion**). Within the seven pHLA-DQ structures, the two *highly-restricted dV* TCR (*PDB* 4OZG & 4OZH) displayed distinct, yet functionally similar mechanisms that shifted the same TCR:MHC H-bond by one MHC a.a. position (relative to the nominal *dV* structure, 4OZF). This involves an additional H-bond (MHC:MHC) also found in suitable *charge-relay* mechanisms of 4OZG and 4OZH.

Results:

PDB files of solved TCR:pHLA-DR and TCR:pHLA-DQ structures were used to investigate V-domain geometry amongst the available complexes, *i.e.*, involving similar (but different) TCR, and/or similar (but different) pMHC. All of these structures share the canonical (diagonal) orientation of the TCR over pMHC, which was one of the earliest observed similarities between different complexes [13]. As indicated in the summary tables (Tables 1A–1D; separation into

different tables is simply for clarity) all sequences for each component of all of these structures are available under the appropriate *PDB* file name at the *NCBI* (www.ncbi.nlm.nih.gov).

TCR-pMHC Geometry

Shown in Figures 1A and 1B is an example of the geometry analysis based on fixed a.a. positions in the HLA-DR/HLA-DQ grooves (*e.g.*, *PDB* 1J8H). We originally used the concept of *Euler* angles (www.mathword.wolfram.com/EulerAngles.html) to establish the basic method (for a TCR-Vα cohort in pHLA-A2 complexes) [20]; here, the method was modified for the available solved TCR-pMHC-II structures (summarized in Tables 1A–1D). In brief, the analysis is based on measuring three angles corresponding to the *twist* (ω), *tilt* (λ), and *sway* (σ) of each V domain over the pMHC.

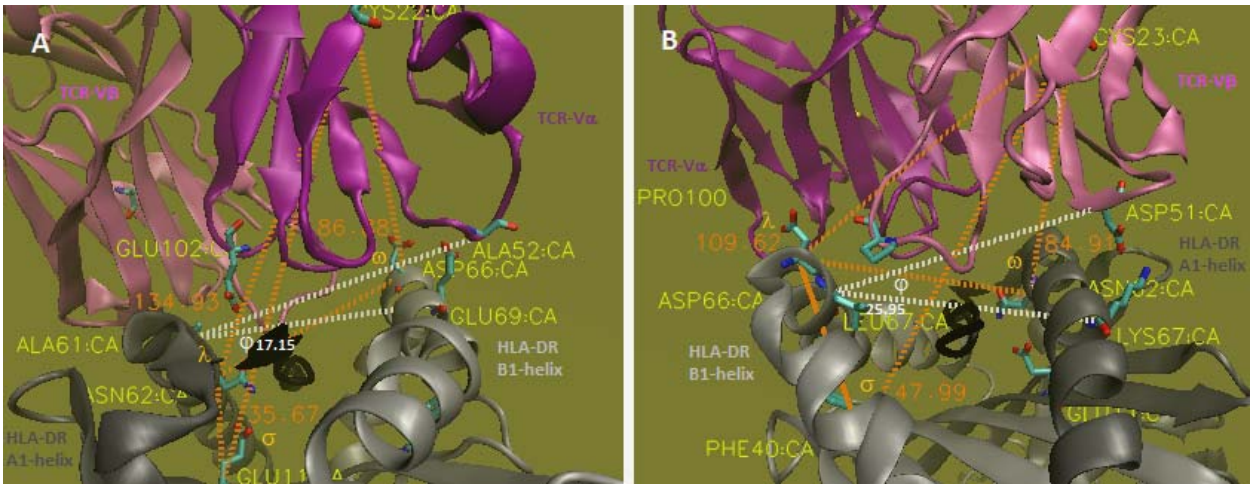

Figure 1 **The twist/tilt/sway of TCR-Vα and -Vβ relative to pMHC-II (A & B, respectively).** Illustrated by the example of 1J8H. This same analysis was performed on all 19 structures (Tables 1A–1D). All angle measurements were from Cα in *VMD-1.9.1* (www.ks.uiuc.edu) used to examine the diversity through the 38 V-domains: (i) in-plane to the MHC-groove {*twist* = ω}, with (ii) displacements perpendicular to the groove {*tilt* = λ}, including (iii) side-to-side variation {*sway* = σ}. The a.a. positions used as coordinates for angular measures (*dotted orange* vectors) across structures were fixed and are labelled (see text). The measured incline of a V domain, or "pitch" {*pitch* = φ_m}, is shown by *white dotted* vectors and could be approximated by the equation: $\varphi_c = [\sigma \div (\lambda + \sigma)]\omega$ (see text, Tables 1A–1D). Side chains of a.a. in measurements are shown by *CPK licorice*; Cα-backbones are in *new cartoon* and labelled; peptides are in *black*. Note from A to B the view of the structure rotates 180°. Figures are the original output of the *PDB* file as analysed in *VMD-1.9.1*. Best viewed at 100%.

For HLA-DR a vector from the DRA a.a. N62 alpha-carbon (N62:Cα) to DRB1/3/5 a.a. D66 (D66:Cα; Cα used for measurements unless otherwise noted) bisects the MHC-groove from the DRA α-helix to the DRB1/3/5 α-helix, then the angle at D66 to the Vα central cysteine (C22) is computed with the *VMD angle-label* tool. This ω-angle can be seen in Figure 1A (Vα) and 1B (Vβ; where the vectors run D66:N62:C23) as dotted *orange* lines near the *melon*-coloured "ω" symbol, 86.88° (Vα), 84.91° (Vβ). The E11 (beta-sheet floor) to N62 to C22 (*tilt*-angle) is similarly shown (1A) near the "λ" symbol, 134.93°; for Vβ, D66 is used (1B), F40:D66:C23, at

109.62°. Finally the σ-angle is measured from N62:E11:C22 (Vα) or D66:F40:C23 (Vβ), here at 35.67° and 47.99°, respectively. For HLA-DQ TCR:pMHC structures, Vα *twist* is based on a DQA a.a. N62:Cα to DQB D66:Cα vector bisecting the groove, where the angle vertex is at D66 to Vα C23:Cα. Here, N11:Cα (not E11 found in DR) to the N62 vertex to C23 defines *tilt*, and the reverse N62:N11:C23 is *sway*. For the DQ Vβ's, angle a.a. are the same as in DR structures. Specifically, *twist* is defined by D66:N62:C23, *tilt* by F40:D66:C23, and *sway* as D66:F40:C23 (Table 1D). We formulated the equation:

$$\varphi_c = \mathbf{k}\varphi_m = [\sigma \div (\lambda + \sigma)]\ \omega$$

that predicts the pitch of a given V domain (*pitch* = angle φ) from the ω, λ, and σ (orientation). By contrast, measuring the φ-angle is based on finding the closest contact between the domain CDR2 and the α-helix that is opposite CDR3 across-the-groove, N62 (for Vα), or D66 (for Vβ) (*yellow* highlighted in Tables 1A–1D). The vectors are directed from that CDR2 a.a. to the across-the-groove α-helix a.a. in closest contact with CDR3, then back to the α-helix a.a. that is in contact with said CDR2 a.a. (*white* dotted lines, Figs. 1A & 1B). In Tables 1A–1D the calculated (φ_c) and measured (φ_m) pitch angles are shown along with the **k** values. Thus, to the extent **k** approaches 1.00 indicates correlation between the measured pitch and calculated pitch values. ***Statistics***: HLA-DR group: *mean k* = 0.99 ± 0.36 (*s*), *n* = 24, *t* = 0.915, μ_0 = 1.00, *p* = 0.376, 95% CI, 0.880–1.10; (paired *Student*'s *t*-test; www.graphpad.com/quickcalcs; www.select-statistics.co.uk). HLA-DQ group: *mean k* = 0.88 ± 0.35 (s), *n* = 14, *t* = 1.24, μ_0 = 1.00, *p* = 0.234, 95% CI, 0.750–1.01. Overall (DR & DQ): *mean k* = 0.95 ± 0.35 (s), *n* = 38, *t* = 0.85, μ_0 = 1.00, *p* = 0.402, 95% CI, 0.870–1.03.

Triple Integrals

We observed that the closest CDR2α contacts with the α2-helix 'across-the-groove' from the conserved R65 contact with CDR3α in HLA-A2 structures specified a polymorphic region from a.a. H/R151 to a.a. A158 [20]. Similarly, the CDR2α closest contacts range for the DR structures here (Tables 1A–1C) implicate a polymorphic DRB α-helix range from a.a. E69 to a.a. T77; for CDR2β, a DRA α-helix contacts range: a.a. Q57 to a.a. K67. For the DQ complexes: CDR2α contacts range is the polymorphic DQB α-helix a.a. E69 to a.a. D76; and for CDR2β, DQA: Q57 to H68, is also polymorphic (Table 1D). Thus, the hypothesis that if these α-helical a.a. are indeed swept/scanned by CDR2 of the V-domain, that one should be able to model such with integration from spherical coordinates in each structure [22, 23]. To normalize the equation across structures and use defined geometry, the central cysteine (C22/23) was chosen as a point rotating from the fixed a.a. position that defines *twist*, *tilt*, and *sway* (Figs. 1 & 2). As shown in Figure 2A & 2B, the coordinates model a 'slice' of a cone, where for each of the 38 V-domains a measured distance was taken between the C22/23 and said α-helix a.a. (*e.g.*, 2IAM Vβ measures 22.12 Å)—this is the *rho* (ρ) distance in the equation. The other measure was the angle ϕ, which is simply the difference from 180° of the previously determined *tilt* angle (λ). The other variables were derived by trigonometry (see Fig. 2B).

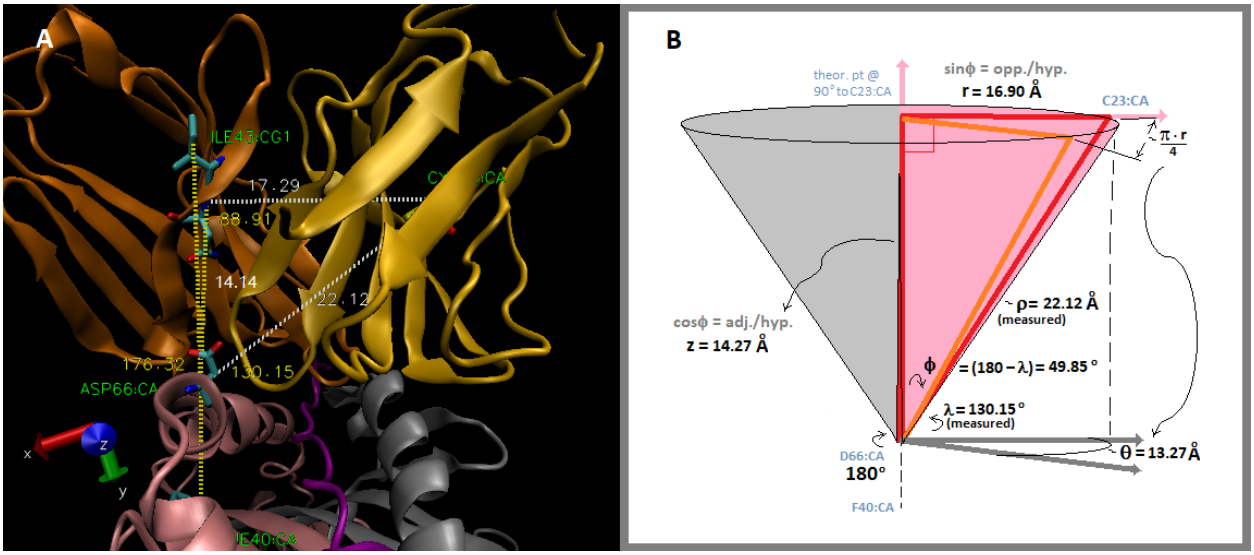

Figure 2 **Derivation of the V-domain *rotational* volumetric-density equation.** The same a.a. positions used for the tilt (λ) angles were used to define the phi (ϕ) angles (*e.g.*, 49.85° for this structure, 2IAM Vβ, *gold*) by subtraction of λ from a theoretical 180° vector. This is illustrated with actual positions in 2IAM (A), showing the 180° approximating, 176.32° vector; the ~ 90°, 88.91° angle; the ~ 14.27 Å, 14.14 Å line segment; and the ~ 16.90 Å, 17.29 Å line segment, *i.e.,* to show the essential geometry within a structure; these are more accurately calculated by trigonometry in the derivation (B). Additionally, the rho (ρ) segment from the appropriate α-helix position (see λ definitions) to a given V-domain's C22/23 was measured (here, 22.12 Å). The cone (full rotation) includes a probability (slice volume) through a *scanning path* (*d*θ) for CDR2. Figure (A) is the original output of the *PDB* file as analysed in *VMD-1.9.1.* Figure (B) is a diagram constructed in *MS-3D-paint*. Best viewed at 100%.

This is integration of a volume element in spherical coordinates (all *dV* values in Tables 1A–1D). As discussed above it is not surprising that the (π · r / 4) values were close to the actual distances between CDR2 contacts in the aforementioned MHC α-helix ranges using this approach, *e.g.*, 13.27 Å, Figure 2B. Thus, the volume of the cone slice for each V-domain was determined upon integration for the three spherical coordinates, rho (ρ), theta (θ) and phi (ϕ), where the upper limit of each integral is derived from the measured ρ and ϕ values of each V domain as shown:

$$\iiint_V dV = \int_0^{\sin^{-1}(r/\rho)} \int_0^{\pi \cdot r/4} \int_0^{z/\cos\phi} \rho^2 \sin\phi \, d\rho \, d\theta \, d\phi$$

Since ρ and ϕ are measured for each V domain, derivations of upper limits for the first and third integrals simply provide trigonometric relationships.

Table 1A Summary Geometry TCR:pHLA-DR

PDB	CDR2[1]	CDR3	σ¶	λ	ω	φ	k‡	ρ (Å)◊	*dV*¥
2IAMα*	K53NZ:E69OE2 2.77	G92O:G58N 3.34	50.75	116.86	101.05	30.73c 18.13m	*1.69*	29.55	1703
2IAMβ	V50CG2:Q57CB 3.96	T98OG1:D66OD2 2.53	30.97	130.15	66.21	12.73c 17.69m	*0.719*	22.12	297
2IANα	K53NZ:E69OE1 2.95	A93N:E55O 3.51	50.99	116.16	99.72	30.42c 19.32m	*1.57*	29.39	1710
2IANβ	Y48OH:Q57NE2 3.30	T98OG1:D66OD1 2.74	30.27	130.76	64.98	12.21c 29.69m	*0.411*	21.78	270
4E41α	K53NZ:E69OE2 2.63	T93OG1:E55OE1 3.17	48.39	119.79	100.80	29.00c 16.93m	*1.71*	29.51	1512
4E41β	N54ND2:E55OE2 3.24	T97OG1:D66OD2 2.51	32.08	129.37	65.97	13.11c 26.55m	*0.494*	23.67	406

*Same TCR (E8) and MHC (DRA1, DRB1*0104; pep I28) as 2ian (pep T28); same pMHC as 4e41 (TCR, G4).

[1]Contacts in angstroms (Å) measured with VMD software (www.ks.uiuc.edu); *yellow* highlighted contacts of the measured *pitch* angle (φ_m), see text. Colour-coded possible H-bonds, pH 7: red (strong), magenta (medium), green (weak); hydrophobic (blue). The calculated *pitch* angle (φ_c) of Vα and Vβ determined as before, using the formula:

$$^{\ddagger}\ \varphi_c = [\sigma \div (\sigma + \lambda)]\ \omega$$

¶Measured angles with VMD in degrees; ◊The value for ρ as measured with VMD in Å; volume of the rotational density distribution (about N-term Cys of the intra-domain disulfide) for Vα and Vβ was modeled as a cone (see text) where¥:

$$\iiint_V dV = \int_0^{\sin^{-1}(r/\rho)} \int_0^{\pi \cdot r/4} \int_0^{z/\cos\phi} \rho^2 \sin\phi \, d\rho \, d\theta \, d\phi$$

And where, $\pi/180$ is the conversion factor for the angular component of the integrand to yield dV in Å^3.

Two example solutions are given below (parameters as defined, Figs. 1 & 2; text)—*for all 38 triple integral solutions*, see **Suppl.1.I***. The ($\pi \cdot r / 4$) circumference segment is used as the upper limit of the $d\theta$ integral because it is accurate to a *path*, *i.e.*, a distance; formally (for θ-angle in degrees): $\pi \cdot r\ (\theta) / 180$ = *arc length* in Å [22]. Hence, $d\phi$ is the only integrand for conversion (multiplying by $\pi \div 180$) to yield the *cubic angstrom* (Å^3) unit of volume. The C22/23 cysteine is historically used as the center of any given V domain [10] and each V-domain cone-slice is thus the volumetric-density through its CDR2 scanning path (*e.g.*, Fig. 2B). The calculus interprets a comprehensive geometry of the V domain into a probability of scanning using only the ground-state structure, *i.e.*, without crystallographic data on theoretical conformers [24].

Table 1B Summary Geometry TCR:pHLA-DR

PDB	CDR2[1]	CDR3	σ ¶	λ	ω	φ	k ǂ	ρ (Å)◊	*dV* ¥
1J8Hα†	A52CB:E69CD 3.54	E102CD:A61CB 4.07	35.67	134.93	86.88	18.17c 17.15m	*1.06*	29.77	740
1J8Hβ	D51OD1:K67NZ 2.54	P100CG:L67CD1 3.43	47.99	109.62	84.91	25.85c 25.95m	*0.996*	27.91	1734
1FYTα	A52CB:E69CD 4.42	K103NZ:Q57OE1 4.36	37.11	133.00	86.29	18.81c 16.90m	*1.11*	29.70	826
1FYTβ	D51OD1:K67NZ 2.67	P100CB:D66CB 4.96	47.87	109.97	85.82	26.03c 23.44m	*1.11*	28.04	1747
3T0Eα	L50CD1:T106CG2 3.65	K96NZ:Q57OE1 4.06	38.06	132.31	88.95	19.87c 21.36m	*0.930*	29.99	855
3T0Eβ‡	E50OE1:K67NZ 3.77	N100CB:Q99CD 4.08	44.41	110.49	76.75	22.00c 20.98m	*1.05*	24.63	1024
4H1Lα	F50CD2:Q70CG 3.17	T98OG1:E55OE2 4.32	44.40	125.30	94.81	24.75c 19.78m	*1.25*	32.62	1782
4H1Lβ	I49CD1:L60CB 3.30	Y97CZ:D66CB 3.77	41.83	117.27	75.97	19.98c 18.84m	*1.06*	27.16	1198
1ZGLα	K51NZ:D76OD1 6.91	S91OG:E55OE2 2.90	51.82	117.60	85.36	26.11c 22.11m	*1.18*	33.78	2831
1ZGLβ	N51CG:A64CB 3.70	R99NH2:D66OD2 2.57	46.60	108.58	69.79	20.96c 14.06m	*1.49*	25.67	1281

† Same TCR, peptide and DRA (DRB1-4) as 1fyt (DRB1*0104).

‡ DRB1 numbering is +29 from conventional due to attached peptide; Q93=Q64; L96=L67; D95=D66; Q99=Q70; T106=T77.

[1] Same as Table 1A legend

Table 1C Summary Geometry TCR:pHLA-DR

PDB	CDR2[1]	CDR3	σ ¶	λ	ω	φ	k ǂ	ρ (Å)◊	*dV* ¥
6CQLα¶	L58CD1:A73CB 3.85	A105N:E55OE1 3.00	37.14	133.17	93.59	20.41c 23.24m	*0.878*	29.01	746
6CQLβ	S66OG:Q57NE2 3.01	R108NH2:Y60OH 2.93	32.68	129.70	70.67	14.22c 18.95m	*0.750*	25.56	542
6CQNα	L58CD2:T77CG2 3.73	A105N:E55OE1 2.89	36.31	134.45	92.61	19.69c 22.34m	*0.881*	28.98	690
6CQNβ	S66OG:Q57NE2 3.17	L109CD1:D66CB 3.73	33.40	128.57	70.09	14.45c 23.56m	*0.613*	25.68	586
6CQQα	L58CD2:A73CB 3.78	A105N:E55OE1 2.88	35.95	134.32	89.29	18.85c 24.02m	*0.785*	28.97	693
6CQQβ	S66OG:Q57OE1 3.64	R108NH2:Y60OH 2.97	34.88	127.12	72.65	15.64c 22.76m	*0.687*	26.34	800
6CQRα	T57OG1:Q70NE2 3.17	A105N:E55OE1 3.03	36.52	133.75	87.44	18.75c 23.44m	*0.800*	28.99	718
6CQRβ	E60OE2:K67NZ 3.07	R108NH2:Y60OH 3.11	36.34	124.36	69.70	15.76c 29.07m	*0.542*	25.95	746

¶ Same TCR (F24) and peptide and DRA (DRB1-11) as 6cqq (DRB1-15) and 6cqr (DRB1-1); same pMHC as 6cqn (TCR, F5; CDR3β G108R).

[1] Same as Table 1A legend

Table 1D Summary of TCR:pHLA-DQ Geometry

PDB	CDR2[1]	CDR3	σ[¶]	λ	ω	φ	k[‡]	ρ[◊]	*dV*[#]
5KSAα[†]	Y57CE1:R70CG 3.71	Y113OH:D55OD2 2.64	36.16	134.95	94.04	15.90c 19.87m	*0.800*	28.76	651
5KSAβ	Y57CE2:A64CB 4.13	Y114CD1:E66CG 3.59	40.26	119.15	71.38	18.03c 23.32m	*0.773*	26.02	939
5KS9α	S58OG:D76OD2 2.65	N114ND2:T61OG1 4.24	41.63	128.32	92.51	22.66c 16.96m	*1.34*	28.65	918
5KS9β	R66NE:Q57OE1 3.14	D114OD2:E66OE2 3.67	33.80	126.25	67.09	14.17c 18.03m	*0.786*	23.57	698
4OZFα[‡]	H55NE2:R70NH2 2.89	A113CB:F58C2 5.05	40.39	131.50	111.69	26.24c 29.44m	*0.891*	33.17	1397
4OZFβ	N63OD1:K67NZ 3.50	D114OD1:R70NE 3.73	37.47	120.69	66.87	15.84c 25.68m	*0.617*	23.16	554
4OZGα	H55NE2:E69OE1 2.93	A115CB:Q57CB 4.11	45.55	124.99	100.43	26.82c 26.64m	*1.01*	31.91	1657
4OZGβ	Y55OH:Q57OE1 3.35	Y117OH:D66OD2 2.72	30.93	128.29	61.35	11.92c 31.10m	*0.383*	21.22	276
4OZHα	H55NE2:E69OE1 2.79	A110CB:D55CB 3.83	48.60	121.57	102.66	29.32c 26.64m	*1.10*	32.35	2030
4OZHβ	Y55OH:Q57OE1 2.86	Y117OH:D66OD2 3.21	29.94	130.19	61.67	11.53c 32.11m	*0.359*	21.47	267
4OZIα	Q55OE1:R70NH1 2.78	S111OG:D55OD1 3.21	44.59	125.38	110.06	28.87c 19.13m	*1.51*	31.72	1588
4OZIβ	T66OG1:Q57OE1 4.08	Y115OH:D66OD1 2.56	38.48	116.06	66.45	16.55c 23.76m	*0.700*	22.10	549
4GG6α[¶]	L57CD2:A73CB 3.98	R110CD:F58CD1 3.71	38.77	132.70	98.93	22.37c 25.38m	*0.881*	30.87	982
4GG6β	N58ND2:H68ND1 3.94	T113OG1:R70NE 2.95	35.96	124.32	68.59	15.39c 22.73m	*0.677*	25.02	645

[†] Bel602 TCR same germline (TRAV 20*01/TRBV 9*01) as 5ks9 (Bel502); MHC (DQ8.5) different from 5ks9 (DQ8); peptide different [glia-gamma-1 (1-11)] from 5ks9 [glia-alpha-1 (1-16)].

[‡] Same MHC (DQ2) and peptide [glia-alpha-2 (1-13)] as 4ozg and 4ozh; 4ozi has same MHC (DQ2) with different peptide [glia-alpha-1 (1-13)]; all have different TCR.

[¶] Same MHC (DQ8) as 5ks9.

[1] refer to Table 1A legend.

Accordingly, the *mean dV* of these TCR V-domains was 955 Å^3, excluding the unusually large *dV* of the 1ZGL Vα. Importantly, with the exclusion of 1ZGL (an apparent *outlier*) there is a linear relationship between calculated pitch (φ_c) and the *dV* triple integral (see plotted values from Tables 1A–1D in the **Suppl.1.II**)—corresponding to the equation: $y = 83.60x - 719.40$; where $R^2 \approx 0.900$, by *linear regression* analysis. The lowest *dV* structures have the lowest calculated *pitch* values. Overall, it says something about TCR:pMHC that may not be intuitive—that a "flush" V-domain geometry could limit CDR2 "finding" an optimal binding interface with side-chains involving the aforementioned α-helix regions. This is the clonotypic nature of TCR function.

CDR3:pMHC Contacts

All these structures, even those with the highest resolution, are subject to limitations of crystallography; and indeed, computed H-bonds are based on these coordinates. Nevertheless, a correlation between relative *dV* and contact distances for hydrogen bonds of the five components could be investigated; see Tables 2A–2C. For example, asking if particularly strong (close) H-bonds are linked to a restricted *dV*?

2IAM-Vα

$\lambda = 116.86°$ $\quad \rho = 29.55$ Å

$\phi = 180 - 116.86 = 63.14°$ $\quad r = \sin\phi\,(\rho) = 0.892\,(29.55) = 26.36$ Å

$\pi \cdot r / 4 = 3.1\bar{4}\,(26.36) / 4 = 20.70$ Å

$$\iiint_V dV = \int_0^{63.14}\int_0^{20.70}\int_0^{29.55} \rho^2 \sin\phi\, d\rho\, d\theta\, d\phi = \int_0^{63.14} 20.70\left[\frac{(29.55)^3}{3}\sin\phi\right] d\phi \quad \text{where, } \int \sin\phi = -\cos\phi$$

$$= \int_0^{63.14}\int_0^{20.70} \frac{1}{3}\rho^3 \sin\phi \Big|_{\rho=0}^{\rho=29.55} d\theta\, d\phi = -20.70\left[\frac{(29.55)^3}{3}\right]\cos\phi \Big|_{\phi=0}^{\phi=63.14} \quad \text{where, } \cos 0 = 1$$

$$= \int_0^{63.14}\int_0^{20.70} \frac{(29.55)^3}{3}\sin\phi\, d\theta\, d\phi = .452\left[-20.70\left[\frac{(29.55)^3}{3}\right]\right] - 1\left[-20.70\left[\frac{(29.55)^3}{3}\right]\right]$$

$$= \int_0^{63.14}\left[\frac{(29.55)^3}{3}\sin\phi\right]\theta \Big|_{\theta=0}^{\theta=20.70} d\phi = -80{,}475 - \left[-178{,}042\right] = 97{,}567 \quad \text{where, multipl. x } \frac{\pi}{180}$$

corrects for angular integrand of ϕ

$$\therefore \iiint_V dV = 1{,}702.87\ \text{Å}^3$$

2IAM-Vβ

$\lambda = 130.15°$ $\quad \rho = 22.12$ Å

$\phi = 180 - 130.15 = 49.85°$ $\quad r = \sin\phi\,(\rho) = 0.764\,(22.12) = 16.90$ Å

$\pi \cdot r / 4 = 3.1\bar{4}\,(16.90) / 4 = 13.27$ Å

$$\iiint_V dV = \int_0^{49.85}\int_0^{13.27}\int_0^{22.12} \rho^2 \sin\phi\, d\rho\, d\theta\, d\phi = \int_0^{49.85} 13.27\left[\frac{(22.12)^3}{3}\sin\phi\right] d\phi \quad \text{where, } \int \sin\phi = -\cos\phi$$

$$= \int_0^{49.85}\int_0^{13.27} \frac{1}{3}\rho^3 \sin\phi \Big|_{\rho=0}^{\rho=22.12} d\theta\, d\phi = -13.27\left[\frac{(22.12)^3}{3}\right]\cos\phi \Big|_{\phi=0}^{\phi=49.85} \quad \text{where, } \cos 0 = 1$$

$$= \int_0^{49.85}\int_0^{13.27} \frac{(22.12)^3}{3}\sin\phi\, d\theta\, d\phi = .645\left[-13.27\left[\frac{(22.12)^3}{3}\right]\right] - 1\left[-13.27\left[\frac{(22.12)^3}{3}\right]\right]$$

$$= \int_0^{49.85}\left[\frac{(22.12)^3}{3}\sin\phi\right]\theta \Big|_{\theta=0}^{\theta=13.27} d\phi = -30{,}879 - \left[-47{,}875\right] = 16{,}996 \quad \text{where, multipl. x } \frac{\pi}{180}$$

corrects for angular integrand of ϕ

$$\therefore \iiint_V dV = 296.64\ \text{Å}^3$$

*for all (*n* = 38) triple-integral solutions, see **Suppl.1.I**.

Table 2A Possible H-bonding networks in TCR:pHLA-DR complexes

PDB	CDR3α:pep	Nn MHC:pep[§]	CDR3β:pep	pep:pep	α:α, α:β, β:β	MHC:MHC	α/β:MHC
2IAM*	none	A1N62ND2:L29O 2.0 A31N:A1N62OD1 2.2 A1N69ND2:A32O 2.1 V34N:A1N69OD1 1.8 B1R71NH1:N30O 2.1 B1R71NH2:N30O 2.0 N30ND2:B1Q70OE1 2.4	T95OH:A31O 1.7 H96NE2:N30O 2.8 H96NE2:A31O 2.8	none	αA93N:αQ91O 2.3 αQ94NE2:αA93O 2.1 αQ94NE2:αI90O 2.8	B1R71NH1:Q70OE1 2.2	αQ91NE2:B1T77O 2.2 A1G58N:αG92O 2.4 βT98N:B1D66OD1 2.0
2IAN (no H's)	none	A1N62ND2:L29O 3.0 A1N69ND2:A32O 3.1 V34N:A1N69OD1 3.0 B1R71NH1:N30O 2.9 B1R71NH2:N30O 2.8	T95OH:A31O 2.5 H96NE2:N30O 3.8 H96NE2:A31O 3.8	none	αG92N:αI90O 3.3	B1R71NH1:Q70OE1 3.2	αQ91NE2:B1T77O 2.9 βT98OG1:D66OD1 2.7
4E41	I28N:G91O 1.9	A31N:A1N62OD1 1.7 A1N62ND2:L29O 2.0 A1N69ND2:A32O 2.1 V34N:A1N69OD1 2.0 B1R71NH1:N30OD1 2.1 B1R71NH2:N30O 1.8 B1Q70NE2:N30O 2.8	R95NH1:I28O 1.9	none	βR95NH2:αG91O 2.0 βR95NH2:αT93O 2.4 βR95NE:αD89OD1 2.5	none	βT97N:D66OD2 1.9

[§]Nn, nearest neighbour; shown possible H-bonds, donor → acceptor; colour-code: ≤ 2.5, red; 2.6 to 2.9, mauve; ≥ 3.0, green.

*Same TCR (E8) and MHC (DRA , DRB1*0104; pep I28) as 2ian (pep T28); same pMHC as 4e41 (TCR, G4).

249
250

Table 2B Possible H-bonding networks in TCR:pHLA-DR complexes

PDB	CDR3α:pep	Nn MHC:pep[§]	CDR3β:pep	pep:pep	α:α, α:β, β:β	MHC:MHC	Nn α/β:MHC
1J8H[†]	none	A1N62ND2:Q311O 1.9 A1N69ND2:L314O 1.8 L316N:A1N69OD1 1.7 B1Q70NE2:Q311OE1 2.1 N312ND2:B1Q70OE1 2.1	K315N:G98O 2.0	none	αN99ND2:βL99O 2.2 βY104N:βL99O 1.9	none	none
1FYT	none	A1N62ND2:Q311O 2.0 T313N:A1N62OD1 1.9 A1N69ND2:L314O 2.1 L316N:A1N69OD1 1.7 B1R71NH1:N312OD1 2.1 B1R71NH2:N312O 1.9 B1Q70NE2:Q311OE1 2.3	N312ND2:T97O 2.0 K315N:G98O 2.1	N312ND2:T313O 2.7 Q311NE2:K310O 2.8	αN99ND2:βL99O 2.0 βY104N:βL99O 2.1	B1R71NH1:Q70OE1 2.0	none
3T0E[‡]	A5N:G92O 2.4	A1N62ND2:E6O 2.4 R9NH1:B1Q93OE1 2.3 R9NH1:B1D95OD1 2.8 R9NH1:B1D95OD2 2.7	S98N:Q8O 2.5 R9NH2:N100O 2.2	G7N:E6OE1 2.6	βR95NH2:αN95OD1 1.8 βY99N:βS98OG 2.1	B1L96N:D95OD1 1.7	βS98OG:A1N62OD1 1.4 B1Q99NE2:βN100O 2.6
4H1L	none	A1N62ND2:N4O 1.9	none	C3SG:N4O 2.8	αN97ND2:βT98OG1 2.8	none	βR94NH2:A1N62OD1 1.6
1ZGL (no H's)	V1N:D99OD1 3.0 H2N:D99OD2 3.3	A1N62ND2:I7O 3.2 N6ND2:A162OD1 3.0	none	none	αK103N:βT102OG1 3.3 αS93OG:βT102O 3.2 βR99NH1:βE103OE1 2.8	none	αS101OG:A1E55OE2 2.9 αS101N:A1E55OE2 3.1 βR99NH2:B5D66OD2 2.6 βR99NH2:B5D66O 2.8 βV100N:B5D70OD1 3.3

[§] Nn, nearest neighbour; shown possible H-bonds, donor → acceptor; colour-code: ≤ 2.5, red; 2.6 to 2.9, mauve; ≥ 3.0, green.

[†] Same TCR, peptide and DRA (DRB1-4) as 1fyt (DRB1*0104).

[‡] DRB1 numbering is +29 from conventional due to attached peptide; Q93=Q64; L96=L67; D95=D66; Q99=Q70.

The structures were examined in *Swiss-PDB Viewer/Deepview-v4.1* (www.wpdbv.vital-it.ch); in *Swiss* H-bonds can be computed after hydrogens are added based on coordinates; two PDB files, (2IAN, 1ZGL) precluded adding hydrogens, and the shown H-bonding distances are thus ≈ 1 Å larger (Tables 2A, 2B). Here, an obvious feature of all the complexes is that there are two principal foci of H-bonding; one involving N62 of DRA (A1) (shaded *orange* in the tables), which may include a separate secondary grouping at N69 (*light-orange* shade), and the DRB1 (B1) centre involving R71 (*green* shades). Note, R71 is polymorphic, so structures involving allotypes/isotypes that have a different a.a. at pos. 71 (1J8H, 3T0E, 1ZGL, 6CQQ & 4H1L) involve a different a.a. (1J8H, 3T0E, 6CQQ), or simply do not have a corresponding beta-chain H-bonding centre (4H1L & 1ZGL). Superimposed upon the H-bonding centres are the H-bonds

between the CDR3α and peptide, the CDR3β and peptide, any peptide to peptide H-bonding, any intra- or inter- bonds between the TCR V-domains, any MHC to MHC H-bonding, and direct H-bonds between TCR V-domains and the MHC (Tables 2A–2C, *columns*).

Table 2C Possible H-bonding networks in TCR:pHLA-DR complexes

PDB	CDR3α:pep	Nn MHC:pep[§]	CDR3β:pep	pep:pep	α:α, α:β, β:β	MHC:MHC	Nn α/β:MHC
6CQL[¶]	R95NH1:A105O 2.0 R95NH1:G106O 2.6 R95NH1:N107OD1 2.3 R95NH2:N107OD1 2.0	A1N69ND2:E97O 2.2 A99N:A1N69OD1 1.8 B1R71NH1:R95O 2.2 B1R71NH2:R95O 2.0	A110N:E97OE1 1.9 G111N:E97OE2 1.9	Q98N:E97OE1 2.0	βG112N:βL109O 2.7	B1R71NH2: D70OD2 2.4	αA105N:A1E55OE1 2.1
6CQN	R95NH1:A105O 2.1 R95NH1:G106O 2.6 R95NH1:N107OD1 2.2 R95NH2:N107OD1 1.9	A96N:A1N62OD1 1.9 A1N62ND2:L94O 1.9 A1N69ND2:E97O 2.2 A99N:A1N69OD1 1.8 B1R71NH1:R95O 2.0 B1R71NH2:R95O 1.9	A110N:E97OE1 1.7 G111N:E97OE2 2.0	Q98N:E97OE1 2.2	βG112N:βL109O 2.5	B1R71NH2: D70OD2 1.9	αA105N:A1E55OE1 2.0
6CQQ	R95NH1:G106O 2.4 R95NH1:N107OD1 2.3 R95NH2:N107OD1 1.8	A96N:A1N62OD1 1.8 A1N62ND2:L94O 1.8 A1N69ND2:E97O 2.1 A99N:A1N69OD1 2.0 S100N:B1D57OD1 2.6	R95NH2:A110O 2.0 A110N:E97OE1 2.1 G111N:E97OE2 2.3	Q98N:E97OE1 2.2	βQ115NE2:αN114O 2.1	none	αA105N:A1E55OE1 2.0 βR108NH2:B1D66OD1 2.2 βR108NH2:B1D66OD2 2.8
6CQR	R95NH1:A105O 2.2 R95NH1:N107OD1 2.1 R95NH2:N107OD1 2.0	A96N:A1N62OD1 1.9 A1N62ND2:L94O 2.1 A1N69ND2:E97O 2.2 A99N:A1N69OD1 1.8 B1R71NH1:R95O 2.4 B1R71NH2:R95O 1.8	R95NH2:A110O 2.2 A110N:E97OE1 1.8 G111N:E97OE2 2.2	Q98N:E97OE1 2.4	βQ115NE2:αN114O 2.1	none	αA105N:A1E55OE1 2.0

[§] Nn, nearest neighbour; shown possible H-bonds, donor → acceptor; colour-code: ≤ 2.5, red; 2.6 to 2.9, mauve; ≥ 3.0, green.

[¶] Same TCR (F24) and peptide and DRA (DRB1-11) as 6cqq (DRB1-15) and 6cqr (DRB1-1); same pMHC as 6cqn (TCR, F5; CDR3β G108R).

In all cases, the CDR3α:peptide bonds are either non-existent, or isolated from the MHC:peptide centres. In 4E41, 6CQQ and 6CQR, CDR3α may link to CDR3β through both bonding nearby with peptide (*blue* shading). For 2IAM, 1J8H, and 1FYT there are bonds between intra-V (2IAM), or intra- and inter-V (1J8H, 1FYT), but these are not linked to the peptide (different *blue* shade). N62 is connected to a peptide:peptide H-bond (4H1L), or to an inter-Vβ bond (3T0E); in both cases these potential networks can include Vβ directly binding with N62 of DRA. 1ZGL is the only structure with an inter-V bond potentially linked to a direct V-domain interaction with MHC, while 4H1L is unique in not showing any DRB H-bonding involving the CDR3 (these two structures involve non-DRB1 isotypes). In 6CQL and 6CQN there is a potential DRB1:peptide link with an MHC:MHC bond, but this does not involve the TCR. The

6CQ series all have CDR3β:peptide bonds also involving a peptide:peptide bond; this is not found in any of the other complexes (*mauve* shading). Finally, and most significantly here, in 2IAM/2IAN, 1FYT and 3T0E the DRB1:peptide centre is potentially connected through a CDR3β:peptide bond with an MHC:MHC H-bond (Table 2A & 2B, *green* shades).

CDR3:pMHC Chemistry

H-bonds computed in *Swiss* are potential H-bonds, and to assess a given network the chemistry of each bond must be examined [24, 25]. The analyses of Tables 2A–2C indicated that a potential H-bonding network where CDR3β is connected through peptide-bonding to DRB1:peptide and DRB1:DRB1 H-bonds might correlate with the *highly-restricted dV* of Vβ in 2IAM (297 $Å^3$) and 2IAN (270 $Å^3$). However, these networks would need to be distinct from this same type of potential network observed in 1FYT and 3T0E (Table 2B), which have *dV* for Vβ of 1747 $Å^3$ and 1024 $Å^3$, respectively (Table 1B). Shown is the analysis of the 2IAM-Vβ versus 1FYT-Vβ for contacts made by each CDR3β (Fig. 3). For 2IAM-Vβ a potential H-bonding network involving CDR3β a.a. Y95 and H96 focused on central peptide a.a. N30 and A31 is apparent; and DRB1 a.a. Q70 and R71 appear in this same network. More specifically, there are three *potential* H-bonds with the N30:O (at R71:NH1, R71:NH2, and H96:NE2); Y95 appears to H-bond to A31:O, and there is also a potential bond between Q70:OE1 and R71:NH1 (Figs. 3A & D). By contrast, 1FYT-Vβ displays single H-bonds between T97:O and N312:ND2 (peptide), and G98:O and K315:N (peptide). The DRB1, R71:NH2 is involved with H-bonding to N312:O, and R71:NH1 has possible H-bonds with N312:OD1 and with Q70:OE1 (Figs. 3B & E). Since 1FYT and 1J8H only differ via DRB1 alleles, we compared H-bonding for the two (Fig. 3B *vs*. 3C; 3E *vs*. 3F). Illustrating the importance of the *Swiss* computation, in Figure 3C it was anticipated that the R71K polymorphism would effect three H-bonds—involving K71 in a manner similar to R71. However, as is shown in Figure 3F, R71K actually shifted peptide contacts to Q70; in-fact, K71 did not show any H-bonds. Importantly, this confirms the change in peptide conformation between the two structures as originally reported [13], and suggests that alloreactivity of 1FYT/1J8H involves the single G98:O to K315:N bond in maintaining the ≈ 1740 $Å^3$ *d*V of the two Vβ (Table 1B). Still, the chemistry of each possible bond must be examined to arrive at feasible mechanisms; arguably, this is necessary to understand binding.

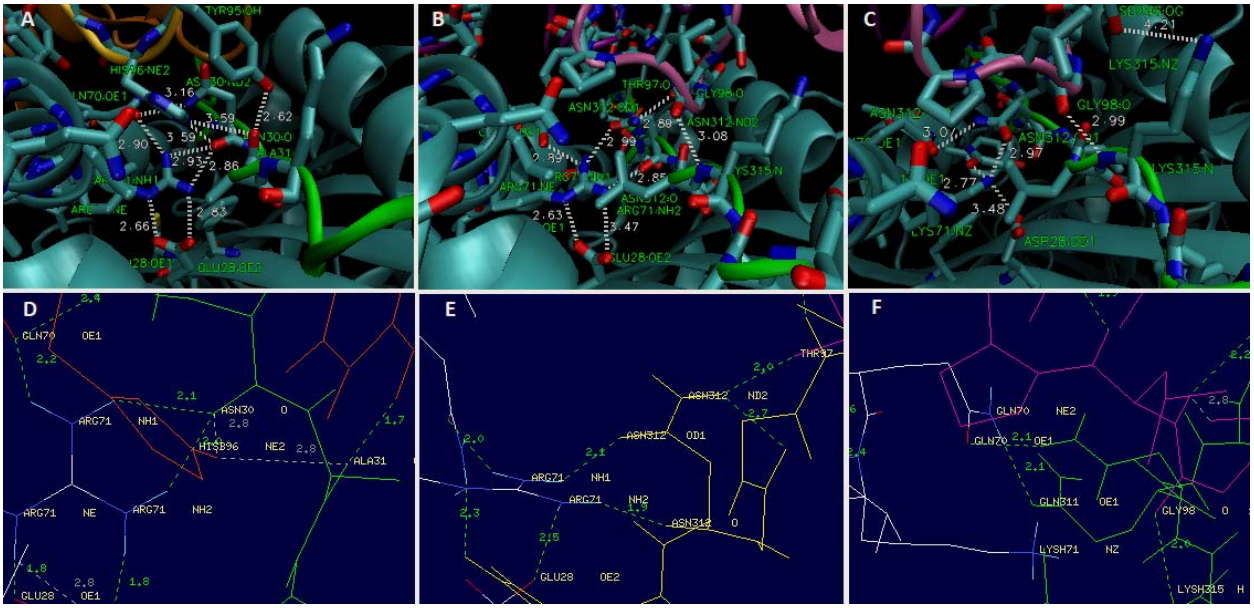

Figure 3 **CDR3 H-bonding networks of 2IAM-Vβ (A & D), 1FYT-Vβ (B & E) and 1J8H-Vβ (C & F).** *PDB* files analysed with *VMD* (top) and *Swiss-Deepview* (bottom). In *Swiss*, hydrogens added and H-bond distances computed; see *dotted* lines in *green*. Involved side chains shown by atom in *licorice* (top) or *stick* (bottom); The 2IAM CDR3β in *gold* (A) and *red* (D); 1FYT/1J8H CDR3β in *mauve* (B, C, E, F); and peptides in *green*. Figure (A⍰C) are the original output of the indicated *PDB* file as analysed in *VMD-1.9.1*. Figure (D⍰F) are the original *PDB* outputs as analysed in *Swiss-Deepview-v4.1.* Best viewed at 125%.

H-Bonding Mechanisms

Shown in Figure 4A is a suitable reaction mechanism for the 2IAM CDR3β binding to the peptide (trios phosphate isomerse, 15mer)-HLA-DR complex (*PDB* ref., 15). When considering the probability of hydrogen bonds in a mechanism, all possibilities were drawn using *ChemSketch* (ACD Labs; www.acdlabs.com/resources) and electron pathways traced using standard evaluations of electron configuration [24, 25]. Thus DRB1 (B1) Q70:OE1 would be in an H-bond with B1R71:NH1 in the absence of the peptide, with the downstream effect of relieving the charge on the B1R71:NH2. Bound peptide has the same effect on R71:NH2, where the N30:O attacks the R71:NH2 (*orange* arrows). Here, the peptide would indirectly break the H-bond between Q70 and R71 (*orange* blocks). When the TCR binds, the βH96:NE2 charge is preferentially attacked by the N30:O (*purple* arrows); this reverses the previous bond between the peptide and DRB1 (*purple* block) and has the downstream effect of re-forming the intra-MHC bond between Q70 and R71 (*purple* blocks on the *orange* blocks). Note, this is the only obvious mechanism that relieves both the H96 and R71 charges, and would be favoured over an A31:O attack on H96:NE2 by the neighbouring βY95 H-bond with A31:O (Fig. 4A). Therefore, *just three* of the possible H-bonds are predicted; and the overall effect is a charge-relay [28, 32, 33] between the MHC, the peptide, and the TCR.

Physiologically, the charge on R71:NH2 could be initially neutralized via the neighbour Q70:OE1—subsequently replaced during processing by CLIP, and/or at HLA-DM exchange, with

antigenic peptide (ref. 9). During T-cell conjugation with the antigen presenting cell, the TCR would replace the A30:O to R71:NH2 attack with an A30:O to H96:NE2 attack, which *relays* the charge back to R71:NH2—subsequently, neutralized again by Q70:OE1. While the hallmark charge-relay network of the 'catalytic triad' within serine proteases stabilizes formation of a covalent acyl-enzyme intermediate [24]; here, a relay is utilized in a similar (albeit, noncovalent) role. Undoubtedly, this would still be a particularly stabilised transition-state, *i.e.*, given that these H-bonds are de-localised across 3/5 components of the structure [24–27].

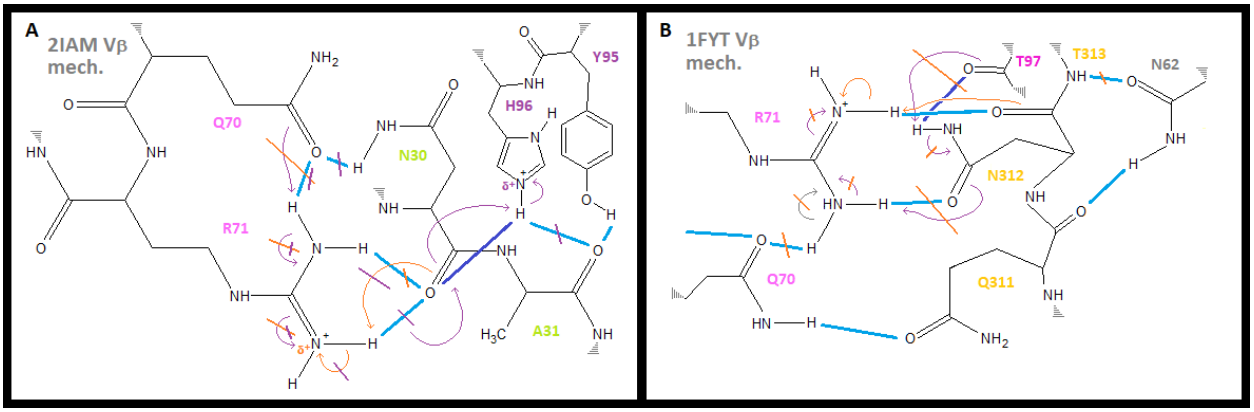

Figure 4 **H-bonding mechanism for CDR3β binding to peptide and MHC.** *ChemSketch* drawn mechanisms are shown for 2IAM (A) and 1FYT (B) and were based on analyses in *VMD* and *Swiss* (Fig. 3). *Not necessarily to scale*. Electrons shown with arrows (*purple*, involving the TCR; *orange*, related to the peptide). Possible H-bonds are shown in *blue* (darker shade for key TCR bond) and are crossed with the color-coded *block* line if not probable (see text). Best viewed at 125%.

By contrast, Figure 4B shows a suitable mechanism for the analogous CDR3β binding reaction with the influenza HA peptide-HLA-DR complex of 1FYT. The differences are the peptide and TCR, as both 2IAM and 1FYT are DRB1*0104 structures. Similarly the peptide would replace the internal B1Q70:OE1 to B1R71:NH1 with an H-bond directly with the charged R71:NH2 via the peptide N312:O (*orange* arrows). Again, the TCR could replace this bond with an attack by the CDR3β (here, T97:O on N312:ND2; *purple* arrows). However, the downstream effect of such an attack in this case is a partial charge on N312:OD1 and an attack on R71:NH1. Note this would restore the charge on R71:NH2, effectively breaking the peptide-MHC bond. Also, despite what looks like proximity in the *ChemSketch* diagram, T97 is actually far removed from R71 (Fig. 3E) such that an attack directly on R71:NH2 is not possible. Thus, because of a remaining charge on R71:NH2, the T97:O attack on N312:ND2 (with its downstream effects) would *not* be favoured (*orange* blocks). While similarly to 2IAM there is a neighbouring H-bond (in this case to Q311), the net loss of a CDR3β H-bond clearly distinguishes the two mechanisms, and shows that the 1FYT CDR3β does not favour a charge-relay mechanism; *i.e.*, would not have the stabilized CDR3β tether at B1R71—ostensibly, essential chemistry driving restricted *dV* in 2IAM. Note that the 2IAN-Vβ mechanism appears to be identical to that of 2IAM (structures differ by a single, sufficiently distant, a.a. of the peptide). Next, 3T0E also showed the TCR:peptide:MHC:MHC type of potential network by contacts analysis (Table 2B). Figure 5A shows the a.a. involved in this potential network using *Swiss*, and the proposed mechanism is shown in Figure 5B. Note that the DRB1-4 allotype of 3T0E has the R71K polymorphism, and

that the involved peptide a.a. is R9. In terms of distinguishing the mechanism from 2IAM, the question is whether the CDR3β N100 H-bond to R9 could exist with a B1:peptide H-bond. In fact, if the Vβ closest contact N100:O is to attack the positive charge of R9:NH2 (purple arrows), it would have to compete (indirectly) with the H-bond of B1Q93:OE1 to R9:NH1 (orange arrows). Together with the B1Q99:NE2 possible bond to the same N100:O, the B1Q93 bond would be favoured; thus, the 3T0E mechanism is not similar to 2IAM, most notably because CDR3β is not likely to form an H-bond with the peptide. Note that in the related structure, *PDB* 3O6F, this conclusion holds as well, where B1D95:OD2 has a possible H-bond to R9:NH2 at 1.6 Å; again, not favouring a CDR3β N100:O bond to the peptide [28].

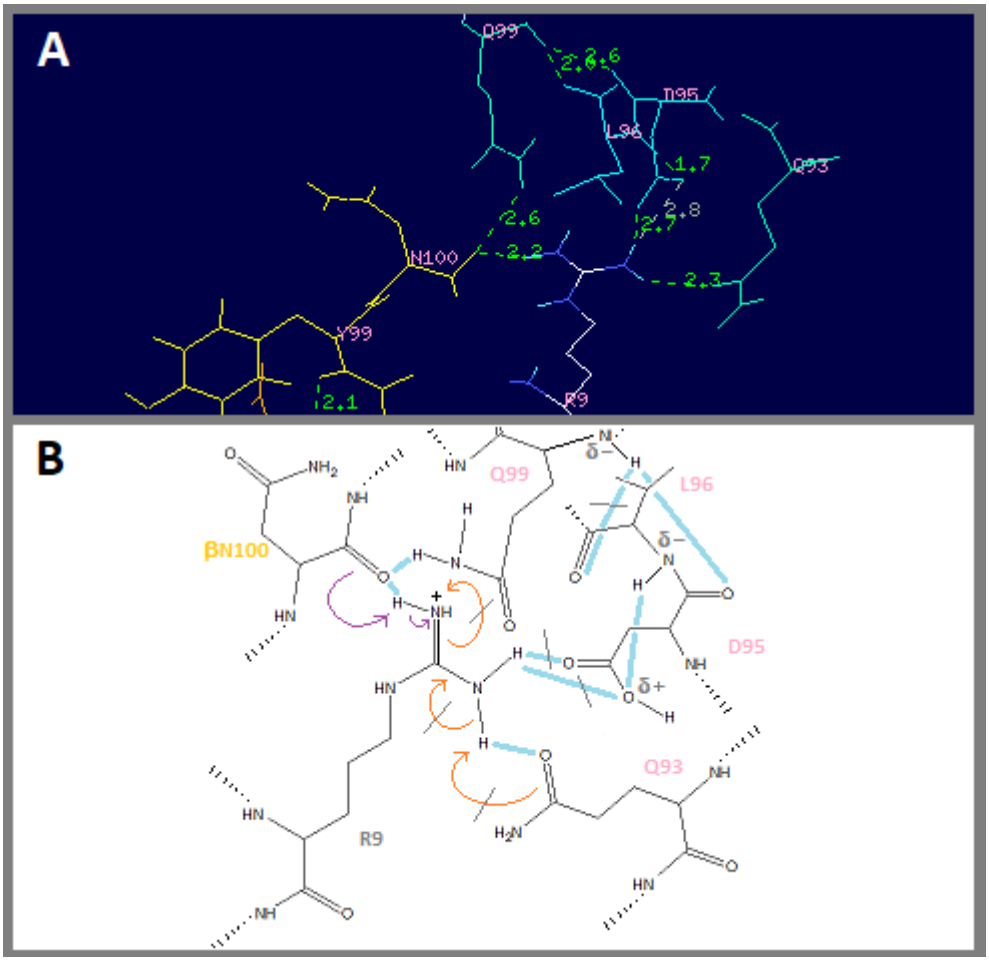

Figure 5 **H-bonds and H-bonding mechanism for 3T0E CDR3β.** *Swiss* was used as before (non-involved a.a. deselected for simplicity) to compute possible H-bonds for 3T0E (A). Mechanism for 3T0E TCR binding via CDR3β to peptide and DRB1 (B) shows possible H-bonds in *blue*, with improbable bonds blocked with *grey* lines (*ChemSketch*). Electron flow shown with *purple* arrows (TCR), or orange (MHC). CDR3β in *yellow* (top) or a.a. labelled *yellow* (bottom); DRB1 in *mauve* (top) or labelled a.a. in *pink* (bottom); peptide by atom (top) or a.a. labelled *grey* (bottom). Figure (A) is the original output of the *PDB* file as analysed in *Swiss-Deepview-v4.1.* Figure (B) is a diagram constructed in *ChemSketch*. Best viewed at 100%.

DQ H-Bonding Mechanisms

As shown in Table 1D, we noted the highly-restricted *dV* in Vβ for two of the TCR:pHLA-DQ structures (4OZH & 4OZG; 267 Å^3 and 276 Å^3, respectively) [*PDB* ref. 29]. Here, the control structure is 4OZF (Vβ *dV*, 554 Å^3), where the pMHC is the same; in-fact, these three structures only differ by their CDR3 regions (*all 38 IMGT-computed CDR3 junctions shown in* **Suppl.1.III**). Accordingly, H-bonding was analysed with *Swiss-Deepview* as before (Fig. 6). While very similar, we noted the absence (Fig. 6A & B) of the DQA1, T61:OG1 role in H-bonding to Vβ, R109:NH1 (Fig. 6C). This is accomplished in 4OZH by preferential H-bonding to Vβ a.a. S110 (Fig. 6A), or sterically in 4OZG by Vβ a.a. F110 (Fig. 6B). Thus, only in 4OZF is T61:OG1 free to H-bond to the pivotal arginine (CDR3β, R109); note, this shift from R109 contacting MHC at a.a. N62 to T61 indirectly disrupts the MHC:MHC component of the charge-relay mechanism (Fig. 7C *vs*. 7A & B). Crucially, the chemistry of S110 versus F110 versus A110 in these TCR is ultimately the result of somatic genetic construction of the CDR3β loop [**Suppl.1.III**; ref. 4].

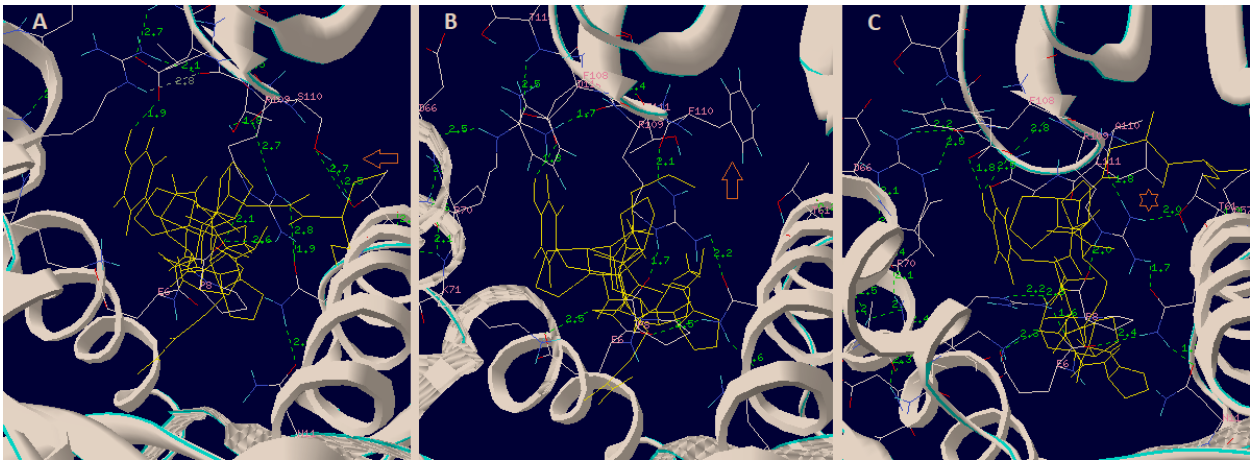

Figure 6 **H-bonds for three different TCR CDR3β, (from germline-identical TCR) on the same pHLA-DQ.** (A) 4OZH, showing S110:T61 H-bonds (*orange* arrow), (B) 4OZG, showing bulky F110 (orange arrow); both of which could preclude (C) 4OZF, the conserved R109 H-bonds to T61 (*orange* star). *Swiss-Deepview* analysis as previously; DQ and CDR3β backbones in *white* ribbon; peptides in *yellow* wire; involved side chains in *CPK*; computed possible H-bonds and distances in *green*. Figures are the original output of the indicated *PDB* files as analysed in *Swiss-Deepview-v4.1.* Best viewed at 150%.

A sequentially mechanism can be envisioned. First, there is internal MHC:MHC H-bonding between N11 and N62 (*purple* arrows, Fig. 7A & B). This facilitates the TCR binding to peptide via a P8:O attack on the R109:NE proton (*blue* arrows), and this leads to neutralizing the charge on R109:NH2. Uniquely here (*purple* arrows, Fig. 7C), when the CDR3β is permissible for the T61:OG1 to R109:NH1 H-bond, *i.e.*, the loop a.a., A110-L111-A112-A113 (4OZF), this appears to preclude both N62:OD1 attack on the charged R109:NH2 and the MHC-MHC H-bond (N11:OD1 to N62:ND1); compare in Fig. 7C to 7A & B). Thus, all three TCR (across DR and DQ) which displayed highly-restricted *dV*, also shared the likelihood of a charge-relay mechanism of CDR3β binding, where an MHC to MHC bond is involved.

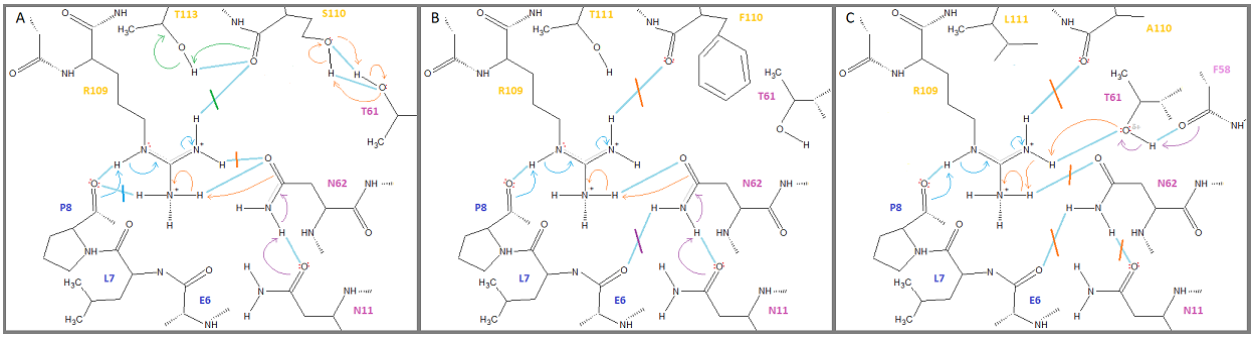

Figure 7 **H-bond mechanisms for three different CDR3β, (from germline-identical TCR) on the same pHLA-DQ.** (A) 4OZH, showing S110:T61 H-bonds (*orange* arrows), (B) 4OZG, showing bulky F110; both of which could preclude (C), the conserved R109 H-bonding to T61, as only observed for the nominal *d*V Vβ of 4OZF. As previously, possible H-bonds in *blue*, with improbable bonds blocked with *colour-matched* lines (*ChemSketch*). Electron flow shown with *blue* arrows (peptide:TCR), or *orange* arrows (MHC:TCR), or *purple* arrows (MHC:MHC). CDR3β a.a. labels in *yellow*, peptide a.a. in *blue*, DRB1 a.a. in *mauve*. Best viewed at 150%.

Discussion:

We used spherical coordinates to derive a unique solved triple integral for each of the 38 V-domains (Figs. 1 & 2; **Suppl.1.I**). Using this approach the mean volumetric-density of a TCR V-domain through the putative CDR2 scanning path was 1060 Å^3. As indicated, $\iiint_V$ *dV* is a 'slice' volume element of a cone with the vertex at the 'across-the-groove' MHC α-helix, where maximal CDR2 scanning (*d*θ) calculates close to the actual range of α-helix distances between the most N-term. and most C-term. CDR2 contacts (Tables 1A–1D). Also, overall geometry consistent with MHC tethering via CDR3 (first formulated for TCR-Vα:pHLA-A2; ref. 20) was again, broadly apparent. Most importantly, here a linear relationship was found between V-domain *pitch*; calculated by:

$$\varphi_c = \mathbf{k}\varphi_m = [\sigma \div (\lambda + \sigma)]\ \omega$$

and V-domain "*dV*" calculated by:

$$\iiint_V dV = \int_0^{\sin^{-1}(r/\rho)} \int_0^{\pi - r/4} \int_0^{z/\cos\phi} \rho^2 \sin\phi \, d\rho \, d\theta \, d\phi$$

This suggests that reduced pitch limits *d*θ, *i.e.*, the CDR2-scanning function. In retrospect, this might confirm intuition (excluding the broader *TCR-CD3 complex,* of course) on possibilities for a mechanism involving just the five components. Indeed, H-bonding chemistry which effects relatively simple physics is a hallmark of quite diverse protein machinery [24, 30]. Also, TCR had **k** values that varied in either direction—indicating that conformation adjustments might moderate flush and open pitch *without* much increasing or decreasing of the *dV*, *viz.*, $R^2 \approx 0.900$ (**Suppl.1.II**). Thus, pitch calculated from a given V-domain's *twist-tilt-sway* (orientation) might be a kind of "hidden" correlate of TCR-selection; although there is not available crystallography on any thymic (selecting) ligands [31]. For 2IAM, there are data for the uncomplexed TCR, and the CDR3β backbone is displaced ≈ 3.4 Å upon binding pMHC, while βY95 moves ≈ 9.0 Å to form

466 the H-bond shown (Fig. 3A); indeed, the key to understanding the TCR may be in this
467 relationship between two induced-fit TCR conformations [32–35]. Suggested here is a
468 conservation/approximation of *dV* for V-domains in TCR binding the *selecting* (thymus) and
469 *activating* (peripheral) pMHC. As shown (Fig. 3E *vs*. 3F), alloreactivity displays shared *dV*
470 accomplished by quite different CDR3 binding chemistry. Alternatively/historically, *intrinsic* TCR
471 affinity (*i.e.*, equilibrium and rate binding constants specified by the α/β TCR protein) has been
472 used to explain both thymic TCR selection and peripheral TCR recognition [4, 36, 37]. Clearly
473 TCR distinguish rare target (agonist) pMHC from thousands of nonagonist pMHC on the APC
474 surface; and this is not mutually exclusive of whether or not a physical force (TCR loading) is an
475 obligate component of TCR function [8, 37]. In these and other regards, there is a
476 stereochemical alternative to an affinity-limited binding reaction [25–27].
477
478 Briefly, we assume TCR:pMHC reactions involve a high(er) energy "scanning" conformer—
479 because, scanning (leading to a suitable CDR2:MHC interface) has the effect of lowering the
480 transition-state free energy, $\Delta\Delta G^{\ddagger}$ (*i.e.*, *Curtin-Hammett* control) [25, 27, 38]. For example, in
481 TCR like E8 (of 2IAM/2IAN) very little CDR2β scanning is apparently needed, due to the effect of
482 the proposed charge-relay stabilised CDR3β:pHLA-DR transition-state. Thus, *dV* is a universal
483 consequence of CDR3 chemistry; but usually the reaction requires more CDR2-scanning (*i.e.*,
484 usually V-domains have larger *dV* values; Tables 1A–1D). Recently, ensemble refinement of
485 crystal data with MD simulations has suggested conformational diversity in the microsecond
486 range, but the conformational changes implicated here-in would be on the millisecond to
487 minutes scale; in this regard, NMR of membrane-bound receptors offers promise [39], and may
488 elucidate the potential of internal water(s) in these H-bonding networks [32]. Finally, spherical
489 coordinates have been used by others to calculate TCR *center-of-mass* variation in TCR:pMHC
490 complexes [40]. These investigators did not consider Vα and Vβ independently, although
491 *Hoffmann et al.*, had shown the angle between Vα and Vβ characterized different TCR in pMHC
492 complexes; allowing these researchers to group a panel of TCR into six different clusters linked
493 to clonotype [41]. While there are myriad downstream implications, taken together, these data
494 support that V-domain rotation and germline to germline contacts between TCR and pMHC
495 *both* depend upon CDR3 H-bonding with highly-conserved MHC α-helix motifs. Thus, unlike
496 TCR-affinity, V-domain *dV* is clonotypic. In this regard, similar to the class-I, R65-motif [20, 42],
497 the class-II motifs presented here are found in *Galago sp.* (NCBI Acc. No. AAA96291) as well as
498 in both *tarsiers* and *lemurs*; suggesting conservation for at least 63 million years [43, 44].
499

500 # Methods:

501 *Availability of data and materials:* *PDB* files are public and available at www.ncbi.nlm.nih.gov; all
502 analytics data are either in the paper, or supplementary materials; *Competing interests:* Dr. Murray and
503 *Xenolaüs Genetics LLC* declare no competing interests regarding the research, or its publication; *Author*
504 *Contributions:* J.S.M. did the research and wrote the manuscript.
505

506 ## *PDB* Analysis

507

VMD-1.9.1 software (www.ks.uiuc.edu) was used for *PDB* files downloaded via *NCBI* (www.ncbi.nlm.nih.gov) from the *RSCB-PDB* (www.rscb.org); views normalized with the *VMD xyz-axis* tool; alpha carbon (Cα) main-chains in *new cartoon*; all alleles named per *NCBI* annotation. *Euler*'s methods (www.mathword.wolfram.com/EulerAngles.html) were the basis for the specific angle analyses, as previously reported. Briefly, three angles corresponding to the *twist*, *tilt*, and *sway* of each domain over the pMHC were measured from fixed Cα through the 19 structures: (i) in-plane to the MHC-groove {*twist* = ω}, with (ii) displacements perpendicular to the groove {*tilt* = λ}, including (iii) side-to-side variation {*sway* = σ}. The a.a. positions used as coordinates for angular measures across structures were *fixed*; see previous. The incline of a V-domain, {*pitch* = φ_m}, was approximated (calculated) by the equation: $\varphi_c = [\sigma \div (\lambda + \sigma)]\ \omega$ (see also **Results**, Tables 1A–1D). Pitch was also measured by using the closest determined CDR2 contact Cα for an angle across-the-groove to the closest CDR3 contact with an α-helix side-chain (vertex), then back to said CDR2 closest contact within the opposite α-helix (≈ 2-fold symmetry); angular value in degrees via the *VMD angle-label* tool. Triple integrals for all 38 V-domains based on the fixed geometry were solved as described in **Results** (**Suppl.1.I**). Linear regression analysis by *MS-Excel* (**Suppl.1.II**); and statistics by paired two-tailed *Student*'s t-test (www.insilico-net/tools/statistics/ttest).

Contacts Analysis

All measures were performed with *Swiss* or *VMD-1.9.1* as is specified in **Results**. Closest contacts in *angstroms* (Å) were determined by examining appropriate coordinates between structures (computed in *Swiss*, or measured/computed in *VMD*). Individual atomic contacts are named per software annotation. DRA chain contacts are *abbrv*. "A1" to avoid confusion with single-letter a.a. code for alanine (A); potential H-bond networks are colour-coded as is described in table captions and the text.

CDR3 Joint Analysis

Nucleotide sequences for all CDR3 of TCR were specified from *PDB* files (**Suppl.1.III**). TCR a.a. sequences were reverse translated using the *SMS* tool at www.bioinformatics.org. These were then imported into *IMGT* algorithms for joint analysis (www.imgt.org). Amino acid sequences of resulting CDR3 joints were determined by the *IMGT* algorithm; consensus *IMGT* numbering.

H-Bonding Mechanisms

Hydrogen (H) bonds were estimated either with *VMD*, or (more accurately) with *Swiss*. In *Swiss*, H-bonds are computed after computing hydrogens to the structures. H-bond distances were used as a factor in determining suitable organic reaction mechanisms, where relevant side-chains were reproduced with *Chemsketch* (www.acdlabs.com/resources). Standard evaluations of electron configuration per relevant atoms were used to predict electron flows [25].

Acknowledgements:

Crystallographers, molecular biologists and computer scientists providing public availability of coordinates, sequences and software made this work possible. Thank you.

Supplementary details:

[Suppl.1.pdf] **I.** Triple Integral Solutions for all V-domains
II. Linear Regression Analysis of Calculated Pitch versus *dV* (all V-domains)
III. *IMTG* CDR3 Joint Analysis (all V-domains)

References:

1. Allison, J. P., McIntyre, B. W. & Bloch, D. Tumor-specific antigen of murine T-lymphoma defined with monoclonal antibody. *J. Immunol.* **129**, 2293–300 (1982).

2. Tonegawa, S. Somatic generation of antibody diversity. *Nature.* **302**, 575–81, doi: 10.1038/302575a0 (1983).

3. Agrawal. A., Eastman, Q. M. & Schatz D. G. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature.* **394**, 744–51, doi: 10.1038/29457 (1998).

4. Janeway, C. A., Jr. *Immunobiology*. 5th ed. (Garland, New York, 2001).

5. Zinkernagel, R. M. & Doherty, P. C. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature.* **248**, 701–2, doi: 10.1038/248701a0 (1974).

6. Galvez, J., Galvez, J. J. & Garcia-Penarrubia, P. Is TCR/pMHC Affinity a Good Estimate of the T-cell Response? An Answer Based on Predictions From 12 Phenotypic Models. *Front. Immunol.* **10**, 349, doi: 10.3389/fimmu.2019.00349 (2019).

7. Gee, M. H., Sibener, L. V., Birnbaum, M. E., Jude, K. M., Yang, X., Fernandes, R. A., *et al*. Stress-testing the relationship between T cell receptor/peptide-MHC affinity and cross-reactivity using peptide velcro. *Proc. Natl. Acad. Sci. USA.* **115**, E7369–78, doi: 10.1073/pnas.1802746115 (2018).

8. Hwang, W., Mallis, R. J., Lang, M. J. & Reinherz E. L. The αβTCR mechanosensor exploits dynamic ectodomain allostery to optimize its ligand recognition site. *Proc. Natl. Acad. Sci. USA.* **117**, 21336–45, doi: 10.1073/pnas.2005899117 (2020).

9. Wieczorek, M., Abualrous, E. T., Sticht, J., Álvaro-Benito, M., Stolzenberg, S., Noé, F. & Freund, C. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. *Front. Immuno*l. **8**, 292, doi: 10.3389/fimmu.2017.00292 (2017).

10. Nisonoff, A., Hopper, J. E. & Spring, S. B. *The Antibody Molecule.* 1st ed. (Academic Press, New York, 1975).

11. Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J. & Wiley, D. C. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature.* **332**, 845–50, doi: 10.1038/332845a0 (1988).

12. Davis, M. M. & Bjorkman, P. J. T-cell antigen receptor genes and T-cell recognition. *Nature.* **334**, 395–402, doi: 10.1038/334395a0 (1988).

13. Hennecke, J. & Wiley, D. C. Structure of a complex of the human alpha/beta T cell receptor (TCR) HA1.7, influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): insight into TCR cross-restriction and alloreactivity. *J. Exp. Med.* **195**, 571–81, doi: 10.1084/jem.20011194 (2001).

14. Li, Y., Huang, Y., Lue, J., Quandt, J. A., Martin, R. & Mariuzza, R. A. Structure of a human autoimmune TCR bound to a myelin basic protein self-peptide and a multiple sclerosis-associated MHC class II molecule. *EMBO J.* **24**, 2968–79, doi: 10.1038/sj.emboj.7600771 (2005).

15. Deng, L., Langley, R. J., Brown, P. H., Xu, G., Teng, L., Wang, Q., *et al*. Structural basis for the recognition of mutant self by a tumor-specific, MHC class II-restricted T cell receptor. *Nat. Immunol.* **8**, 398–408, doi: 10.1038/ni1447 (2007).

16. Yin, L., Crawford, F., Marrack, P., Kappler, J. W. & Dai, S. T-cell receptor (TCR) interaction with peptides that mimic nickel offers insight into nickel contact allergy. *Proc. Natl. Acad. Sci. USA*. **109**, 18517–22, doi: 10.1073/pnas.1215928109 (2012).

17. Yin, Y., Wang, X. X. & Mariuzza, R. A. Crystal structure of a complete ternary complex of T-cell receptor, peptide-MHC, and CD4. *Proc. Natl. Acad. Sci. USA.* **109**, 5405–10, doi: 10.1073/pnas.1118801109 (2012).

18. Sharon, E., Sibener, L. V., Battle, A., Fraser, H. B., Garcia, K. C. & Pritchard, J. K. Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat. Genet.* **48**, 995–1002, doi: 10.1038/ng.3625 (2016).

19. Gunnarsen, K. S., Hoydahl, L. S., Risnes, L. F., Dahal-Koirala, S., Neumann, R. S., Bergseng, E., *et al*. A TCRα frame-work centered codon shapes a biased T cell repertoire through direct MHC and CDR3β interactions. *JCI Insight.* **2**, 95193, doi:10.1172jci.insight.95193 (2017).

20. Murray, J. S. An old Twist in HLA-A: CDR3α Hook up at an R65-joint. *Front. Immunol.* **6**, 268, doi: 10.3389/fimmu.2015.00268 (2015).

21. Burrows, S. R., Chen, Z., Archbold, J. K., Tynan, F. E., Beddoe, T., Kjer-Nielsen, L., *et al*. Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. *Proc. Natl. Acad. Sci. USA.* **107**, 10608–13, doi: 10.1073/pnas.1004926107 (2010).

22. Stewart, J. *Calculus.* 6th ed. (Brooks/Cole Thomson, Belmont, CA, 2007).

23. Das, B. B., Park, S. H. & Opella, S. J. Membrane protein structure from rotational diffusion. *Biochim. Biophys. Acta.* **1848**, 229–45, doi: 10.1016/j.bbamem.2014.04.002 (2015).

24. Berg, J. M., Tymoczko, J. L. & Stryer, L. *Biochemistry.* 5th ed. (W. H. Freeman, New York, 2002).

25. Carey, F. A. & Sundberg, R. J. *Advanced Organic Chemistry, pt. A. Structure and Mechanism.* 2nd ed. (Plenum Press, New York, 1984).

26. Chen, J. & Thirumalai, D. Interface residues that drive allosteric transitions also control the assembly of l-lactate dehydrogenase. *J. Phys. Chem. B.* **122**, 11195–205, doi: 10.1021/acs.jpcb.8b06430 (2018).

27. Hernández, G., Anderson, J. S. & LeMaster, D. M. Assessing the native state conformational distribution of ubiquitin by peptide acidity. *Biophys. Chem.* **153**, 70–82, doi: 10.1016/j.bpc.2010.10.006 (2010).

28. Yin, Y., Li, Y., Kerzic, M. C., Martin, R. & Mariuzza, R. A. Structure of a TCR with high affinity for self-antigen reveals basis for escape from negative selection. *EMBO J.* **30**, 1137–48, doi: 10.1038/emboj.2011.21 (2011).

29. Petersen, J., Montserrat, V., Mujico, J. R., Loh, K. L., Beringer, D. X., van Lummel, M., *et al*. T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. *Nat. Struct. Mol. Biol.* **21**, 480–8, doi: 10.1038/nsmb.2817 (2014).

30. Brunori, M. The Bohr effect before Perutz. *Biochem. Mol. Biol. Educ.* **40**, 297–9, doi: 10.1002/bmb.20629 (2012).

31. Zerrahn, J., Held, W. & Raulet, D. H. The MHC reactivity of the T cell repertoire prior to positive and negative selection. *Cell.* **88,** 627–36, doi: 10.1016/s0092-8674(00)81905-4 (1997).

679 32. Pusuluk, O., Farrow, T., Cemsinan, D., Burnett, K. & Vedral, V. Proton tunneling in hydrogen
680 bonds and its implications in an induced-fit model of enzyme catalysis. *Proc. R. Soc. A.* **474**,
681 20180037, dx.doi.org/10.1098/rspa.2018.0037 (2018).
682
683 33. Rudolph, M. G. & Wilson, I. A. The specificity of TCR/pMHC interaction. *Curr. Opin.*
684 *Immunol.* **14**, 52–65, doi: 10.1016/s0952-7915(01)00298-9 (2002).
685
686 34. Deng, L., Langley, R. J., Wang, Q., Topalian, S. L. & Mariuzza, R. A. Structural insights into
687 the editing of germ-line-encoded interactions between T-cell receptor and MHC class II by Vα
688 CDR3. *Proc. Natl. Acad. Sci. USA.* **109**, 14960–5, doi: 10.1073/pnas.1207186109 (2012).
689
690 35. Buckle, A. M. & Borg, N. A. Integrating Experiment and Theory to Understand TCR-pMHC
691 Dynamics. *Front. Immunol.* **9**, 2898, doi: 10.3389/fimmu.2018.02898 (2018).
692
693 36. Faro, J., Castro, M. & Molina-París, C. A unifying mathematical framework for experimental
694 TCR-pMHC kinetic constants. *Sci. Rep.* **7**, 46741, doi: 10.1038/srep46741 (2017).
695
696 37. Limozin, L., Bridge, M., Bongrand, P., Dushek, O., van der Merwe, P. A. & Robert, P. TCR-
697 pMHC kinetics under force in a cell-free system show no intrinsic catch bond, but a minimal
698 encounter duration before binding. *Proc. Natl. Acad. Sci. USA.* **116**, 16943–48, doi:
699 10.1073/pnas.1902141116 (2019).
700
701 38. Lodola, A., Mor, M., Zurek, J., Tarzia, G., Piomelli, D., Harvey, J. N. & Mulholland, A. J.
702 Conformational effects in enzyme catalysis: reaction via high energy conformation in fatty acid
703 amide hydrolase. *Biophys J.* **92**, L20–2, doi: 10.1529/biophyj.106.098434 (2007).
704
705 39. Mallis, R. J., Brazin, K. N., Duke-Cohan, J. S., Hwang, W., Wang, J. H., Wagner, G., *et al.*
706 NMR: an essential structural tool for integrative studies of T cell development, pMHC ligand
707 recognition and TCR mechanobiology. *J. Biomol. NMR.* **73**, 319–32, doi: 10.1007/s10858-019-
708 00234-8 (2019).
709
710 40. Singh, N. K., Abualrous, E. T., Ayres, C. M., Noé, F., Gowthaman, R., Pierce, B. G., *et al.*
711 Geometrical characterization of T cell receptor binding modes reveals class-specific binding to
712 maximize access to antigen. *Proteins.* **88**, 503–13, doi: 10.1002/prot.25829 (2020).
713
714 41. Hoffmann, T., Krackhardt, A. M. & Antes, I. Quantitative Analysis of the Association Angle
715 between T-cell Receptor Vα/Vβ Domains Reveals Important Features for Epitope Recognition.
716 *PLoS Comput. Biol.* e1004244, doi: 10.1371/journal.pcbi.1004244 (2015).
717
718 42. Murray, J. S. Transposon-mediated death of an ancestral A-23-like allele: evolution of TCR-
719 positioning motifs in the HLA-A lineage. *Immunogenetics.* **67**, 473–6, doi: 10.1007/s00251-015-
720 0852-3 (2015).
721

722 43. Steiper, M. E. & Seiffert, E. R. Evidence for a convergent slowdown in primate molecular
723 rates and its implications for the timing of early primate evolution. *Proc. Natl. Acad. Sci. USA.*
724 **109**, 6006–11, doi: 10.1073/pnas.1119506109 (2012).
725
726 44. Kasahara, M., Vazquez, M., Sato, K., McKinney, E. C. & Flajnik, M. F. Evolution of the major
727 histocompatibility complex: class II A cDNA clones from the cartilaginous fish. *Proc. Natl. Acad.*
728 *Sci. USA.* **89**, 6688–92, doi: 10.1073/pnas.89.15.6688 (1992).