Title:

Somatic genetics of CDR3 control H-bonds to MHC via V-domain rotational probability (*d*V) effecting germline CDR2 'scanning' (*d*θ) of polymorphic MHC

Running title:

Genetic calculus of TCR binding class-II pMHC

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) repertoire within a given *major histocompatibility complex* (MHC; HLA, in humans) genotype is essential for protection against rapidly dividing pathogens. Historically attributed to relative affinity, genetically vast TCRs are surprisingly "focused" towards a micromolar affinity for their respective pHLA 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 equation* (from fixed geometry) for any given V-domain in TCR bound to pMHC. We examined solved complexes involving HLA-DR and HLA-DQ, where all the available structures are still reasonably analysed. Certain V-beta domains displayed "rare" geometry within this panel—specifying a very low ("highly-restricted") rotational probability/volumetric density (*d*V). Remarkably, hydrogen (H)-bond *charge-relays* spanning CDR3-beta, the peptide, and polymorphic MHC distinguished these structures from the others; suggesting that CDR3 binding chemistry dictates CDR2 'scanning' on the respective MHC-II alpha-helix. Taken together, this analysis (*n* = 38 V-domains) supports a novel geometric theory for *MHC restriction*—one not constrained by a thymus "optimized" TCR-affinity.

Keywords:

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

Introduction:

For better and (for) worse all the biology of the adaptive immune system cascades from a stochastic genetic mechanism (*somatic DNA rearrangement*)—generating TCR and antibodies (sIg) "individualized" within each precursor of a given T-cell or B-cell clone [1–4]. TCR (here-in, refers to α/β-type only) are utilized against antigenic structures which are all partially similar—all are MHC (human, HLA) proteins transporting a single 8–25 amino acid (a.a.) peptide (*abbrv*., pMHC). The peptide component is extremely variable however; derived from intra-cellular, proteolytic 'processing' of virtually any *self*, or non-self protein [4, 5]. In-fact, TCR diversity exceeds Ig diversity, although antibodies historically defined a 'V-region,' *e.g.*, their elicitation against hitherto not-synthesised organic structures [6]. Classic MHC class-I (HLA-A, -B, -C) and -II (HLA-DR, -DQ, -DP) genes include the most polymorphic in the human genome, yet the germline limits diversity to the alleles on maternal and paternal sixth chromosomes [5, 7, 8]. Thus, despite extreme MHC polymorphism within the population, and co-dominance within the individual, overwhelmingly, most of the diversity in a patient's pMHC structures stem from the peptide possibilities for these ligands—indeed, the 'hyper-variant' CDR3 loops are positioned (conspicuously) atop the peptide component in solved TCR:pMHC complexes [9–15]; although recent investigations have challenged ubiquity of this structure/function [16-18].

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) [17]. *Burrows et al*., also found conserved CDR3 contacts directly with MHC-I [19]. More recently, *Sharon et al* formally linked germline TCR usage to HLA, particularly via the class-II alleles [15]. Here, the *STCR-Dab* (http://opig.stats.ox.ac.uk/webapps/stcrdab/) was used to identify all current TCR complexes involving HLA-DR and HLA-DQ (as of 07/06/19; 3.0 Å resolution cut-off). Note, also included are *PDB* 4H1L, at 3.3 Å, and 3T0E, at 4.0 Å (see below), and the 'unusual topology' complexes (*PDB* 1YMM, 4Y1A & 4Y19) are excluded based on this investigation targeting "characteristic" features.

Using *Euler*'s geometry, we established 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 [17]. Here, we found a linear relationship between calculated pitch and a new measure, "*d*V"—by multivariable calculus. As is discussed, *d*V interprets V-domain orientation into a rotational probability involving an apparent CDR2–MHC α-helix scanning function (*d*θ). While *d*V were 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; further suggesting the chemistry of "somatic-TCR" might dictate positioning of "germline-TCR" [20–22]. Within the seven pHLA-DQ structures, the two *highly-restricted d*V TCR (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 *d*V structure, 4OZF). Crucially, this involves an additional H-bond (MHC:MHC) in suitable 'charge-relay' mechanisms of 4OZG and 4OZH. We will discuss how these data suggest that TCR binding *specificity* could be controlled by *orientation* (dynamics) and *not* densities of pMHC, or relative forward and reverse rates of the reaction, *per se*.

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 [10]. As indicated in the summary tables (Tables 1A–1D) some of the structures share the same TCR, where different peptide (2IAM *vs.* 2IAN), or different DRB1 (1J8H *vs.* 1FYT; 6CQL *vs.* 6CQQ *vs.* 6CQR) were compared. Some structures share pMHC binding different TCR (2IAM *vs.* 4E41; 6CQL *vs.* 6CQN; 4OZF–I). All the DR structures share the DRA chain [8, 23] and in some cases, so, too, the DRB1 allotype (2IAM, 2IAN, 4E41, 1FYT, 6CQR; all DRB1-1), (1J8H, 3T0E; both DRB1-4), (6CQL, 6CQN; both DRB1-11); 6CQQ is DRB1-15, 4H1L is DRB3 and 1ZGL is DRB5 (isotypes). The DQ structures are DQ8 (5KSA, 4GG6), DQ2.5 (5SK9), or DQ2.2 (4OZF–I) (*NCBI* annotation and *BLAST*; www.ncbi.nlm.nih.gov). While bound peptides in the DR group are of diverse protein origins, all the DQ peptides are wheat *gliadins*; either the glia-gamma-1, glia-alpha-2, or glia-MM1. Interest in these particular DQ structures is due to *celiac disease*, an autoimmune-like condition known to be mediated by predominant T-cell

clonotypes against ingested wheat. Some 95% of *celiac* patients express HLA-DQ2.5 (DQA*05/DQB1*02), while the rest are either HLA-DQ8 (DQA*03/DQB1*03:02), or DQ2.2 (DQA*02:01/DQB1*02) [24]. For our purposes, the crystal structures of these nearly identical complexes facilitated isolating the chemistry associated with V-domain rotational probability/volumetric density (*d*V). Note that 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, -DQ 'grooves' (*e.g.*, *PDB* 1J8H). We originally used *Euler* angles (www.mathword.wolfram.com/EulerAngles.html) to establish the basic method (for a TCR-Vα cohort in pHLA-A2 complexes) [17]; here, we modified the method for the available solved TCR:pHLA-DR, -DQ 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.

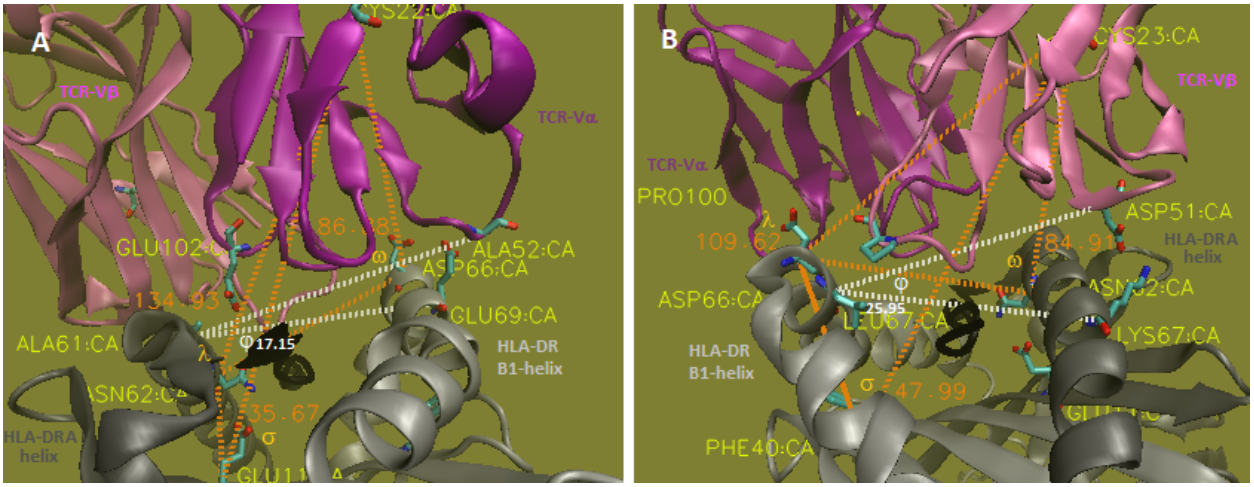

Figure 1 **The "twist/tilt/sway'" of TCR-Vα and -Vβ relative to pMHC-II (A & B, respectively).** Illustrated by the example of 1J8H. This same analysis was performed on all 19 structures (Tables 1A–1D). All angle measurements were from Cα in *VMD-1.9.1* (www.ks.uiuc.edu) used to examine the diversity 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°.

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. Analogously 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 (not shown). 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, angles 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:

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

that predicts the pitch of a given V-domain (*pitch* = angle φ) from the ω, λ, and σ. By contrast, measuring the φ-angle is based on finding the closest contact between the domain CDR2 and the α-helix that is opposite 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., then back to the α-helix a.a. that is in contact with the 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 observed *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 [17]. Similarly, the CDR2α closest contacts range for the DR structures here (Tables 1A–1C) implicate an α2-helix range from a.a. E69 to a.a. T77; CDR2β α1-helix contacts range: a.a. Q57 to a.a. K67. For the DQ complexes: CDR2α contacts range is α2-helix a.a. E69 to a.a. D76; CDR2β: Q57 to H68 (Table 1D). Thus, we hypothesized that if these α-helical a.a. are indeed swept/scanned by CDR2 of the V-domain, that we should be able to model such with integration from *spherical coordinates* in each structure [25–28]. To normalize the equation across structures and use the geometry, the central cysteine (C22/23) was chosen as a point rotating from the fixed across-the-groove a.a. position that defines *twist*, *tilt*, and *sway* (Figs. 1 & 2). As shown in Figure 2B, the coordinates model a 'slice' of a cone [26], where for each of the 38 V-domains we measured the distance 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).

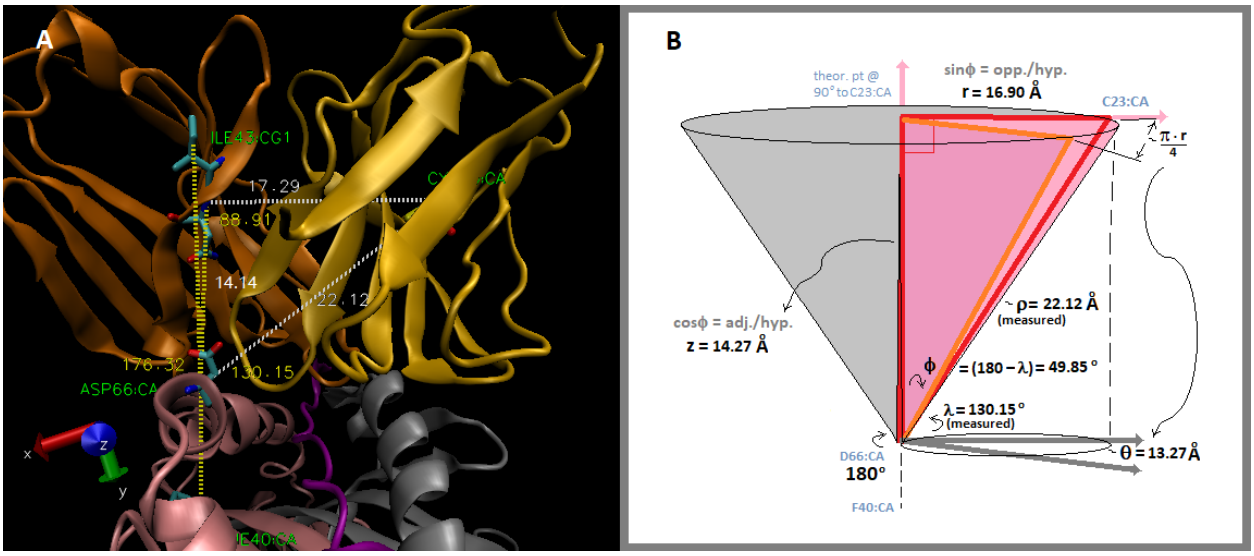

Figure 2 **Derivation of the V-domain *rotational* volumetric-density equation.** The same a.a. positions used for the tilt (λ) angles were used to define the phi (ϕ) angles (*e.g.*, 49.85° for this structure, 2IAM Vβ, *gold*) by subtraction of λ from a theoretical 180° vector. This is illustrated with 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* ($d\theta$) for CDR2.

This is integration of a 'volume element' in spherical coordinates (all *d*V values in Tables 1A–1D). For strictly compositional reasons throughout "*the triple integral of d*V" may be referred to as simply "*d*V." As discussed above it is not surprising that the ($\pi \cdot r / 4$) 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:

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

Since ρ and ϕ are measured for each V-domain, derivation 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:

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

¶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¥:

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

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***. 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 [26]. 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 [6] and each V-domain 'slice' is thus the volumetric-density through its CDR2 *scanning* (*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—suggesting here, a classic protein function, *allostery* [28, 29]; see also 'protein machinery' in **Discussion**.

Table 1B Summary Geometry TCR:pHLA-DR

PDB	CDR2[1]	CDR3	σ ¶	λ	ω	φ	k ‡	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 ‡	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[◊]	*d*V[#]
5KSAα[†]	Y57CE1:R70CG 3.71	Y113OH:D55OD2 2.64	36.16	134.95	94.04	15.90c 19.87m	*1.25*	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	465
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 5ksa (Bel502); MHC (DQ8.5) different from 5ksa (DQ8); peptide different [glia-gamma-1 (1-11)] from 5ksa [glia-gamma-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 5ksa.

[1] refer to Table 1A legend.

Accordingly, the *mean dV* of these TCR V-domains was 955 ± 507 Å^3, excluding the unusually large *d*V of the 1ZGL Vα (see below). Importantly, with the exclusion of 1ZGL (an apparent *outlier*) there is a *linear relationship* between calculated pitch (φ_c) and the *d*V triple integral (see plotted values from Tables 1A–1D in the **Suppl.2**)—corresponding to the equation: $y = 83.60x - 719.40$; where $R^2 = 0.900$, by *linear regression* analysis. The lowest *d*V 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 its capacity to "find" a compatible binding interface with side-chains involving the aforementioned α-helix regions.

CDR3:pMHC Contacts

Particularly close TCR contacts with an autologous MHC during thymic development could theoretically enhance 'positive selection' [30]. However, some of the CDR2/α-helix interface

measurements (Tables 1A–1D) could suggest such TCR are not selected. Alternatively, we know close CDR2β contacts with monomorphic DRA must contribute to linkages between DRB1 and autoimmunity [15], but considerations here cannot directly exclude 'negative selection' that deletes 'forbidden clones' specific for self-peptide:self-MHC complexes [4]. All these structures, even those with the highest resolution, are subject to limitations of crystallography [15]; and indeed, we compute the H-bonds based on these coordinates. Nevertheless, we sought a correlation between relative *d*V and contact distances for hydrogen bonds of the five components; see Tables 2A–2C. For example, asking if particularly strong (close) H-bonds are linked to a restricted *d*V?

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**.

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).

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	C35G: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.

We examined the structures 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; complexity of certain structures (2IAN, 1ZGL) precluded computing 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 extend to the Vβ directly binding with N62. 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 (recall 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 [28, 31]. 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" *d*V 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 *d*V for Vβ of 1747 $Å^3$ and 1024 $Å^3$, respectively (Table 1B). Shown is our analysis of the 2IAM-Vβ versus 1FYT-Vβ for contacts made by each CDR3β (Figs. 3A–D). 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, while the β-sheet a.a. E28 (with R71:NE and R71:NH2 bonds of precisely the same distance) cancel out. 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 we 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 [10], and suggests that alloreactivity of 1FYT/1J8H involves the capacity of the single G98:0 to K315:N bond maintaining the ≈ 1740 $Å^3$ *d*V of the two Vβ (Table 1B). Still, we had to examine the chemistry of each possible bond to arrive at feasible mechanisms; arguably, this is necessary to understand binding.

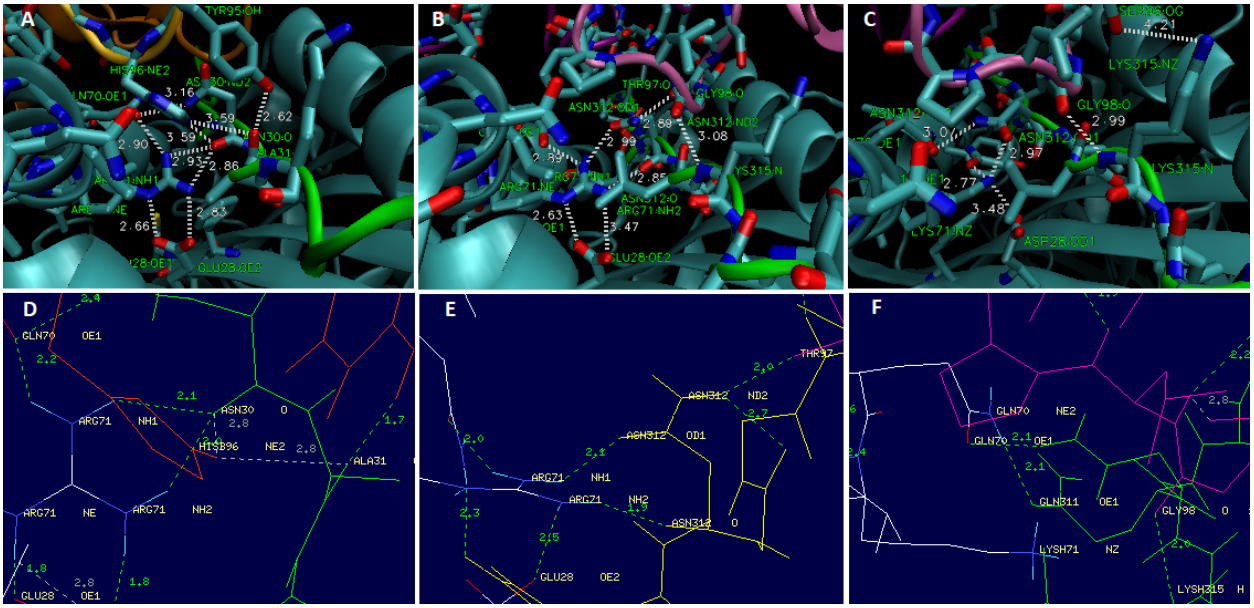

Figure 3 **CDR3 H-bonding networks of 2IAM-Vβ (A & D), 1FYT-Vβ (B & E) and 1J8H-Vβ (C & F).** *PDB* files analysed with *VMD* (top) and *Swiss-Deepview* (bottom). In *Swiss*, hydrogens added and H-bond distances computed; see *dotted* lines in *green*. Involved side chains shown by atom in *licorice* (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*.

H-Bonding Mechanisms

Shown in Figure 4A is a suitable organic reaction mechanism for the 2IAM CDR3β binding to the peptide (triosephosphate isomerse, 15mer)-HLA-DR complex (*PDB* ref., 12). When considering the probability of hydrogen bonds in a mechanism, we 'drew-out' all the possibilities using *ChemSketch* (ACD Labs; www.acdlabs.com/resources) and traced electron pathways using standard evaluations of electron configuration [31]. 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 at HLA-DM exchange, with foreign peptide (ref. 5). 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 [28]; here, a relay is utilized in a similar (albeit, noncovalent) role. Undoubtedly, this would still be a 'stabilised' transition-state [31], *i.e.*, given that these H-bonds are *de-localised* across 3/5 components of the structure [32, 33].

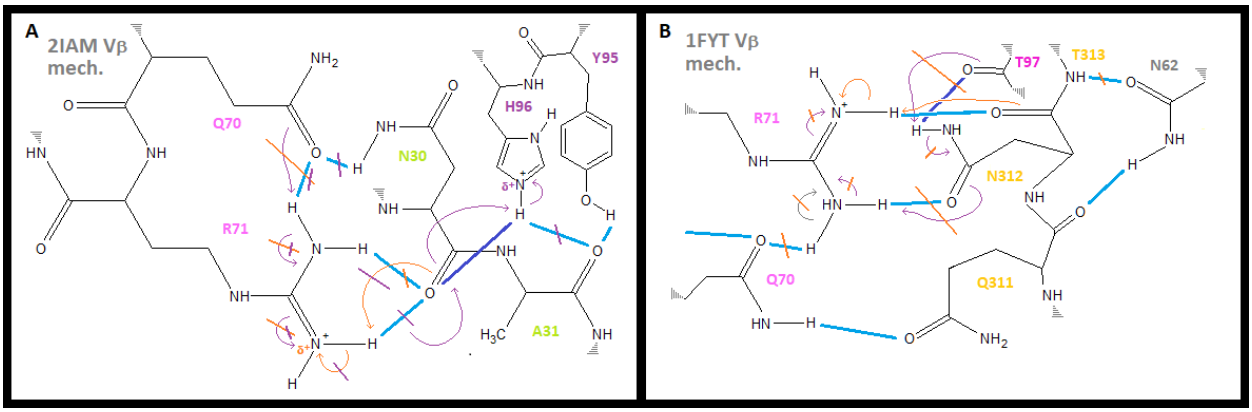

Figure 4 **H-bonding mechanism for CDR3β binding to peptide and MHC.** *ChemSketch* drawn mechanisms are shown for 2IAM (A) and 1FYT (B) and were based on analyses in *VMD* and *Swiss* (Fig. 3). *Not necessarily to scale*. Electrons shown with arrows (*purple*, involving the TCR; *orange* related to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

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 again 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, the chemistry driving restricted *d*V in 2IAM. Note that the 2IAN-Vβ mechanism appears to be similar to that of 2IAM (structures differ by a single, sufficiently "distant" a.a. of the peptide), at least as comparable with the *VMD* analysis (not shown). 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 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 [34], this conclusion holds as well, where B1D95:OD2 has a possible H-bond to R9:NH2 at 1.6 Å (not shown); again, not favouring a CDR3β N100:O bond to the peptide.

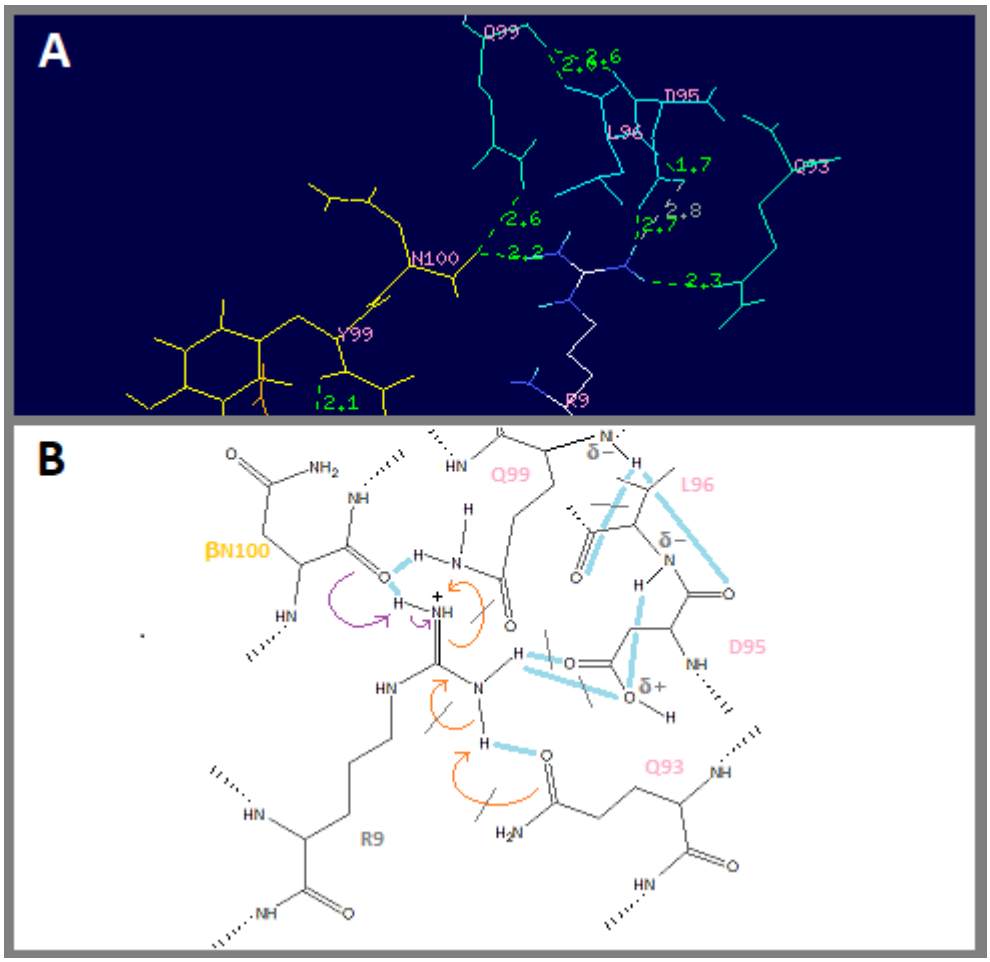

Figure 5 **H-bonds and H-bonding mechanism for 3T0E CDR3β.** *Swiss* was used as before (non-involved a.a. 'deselected' for simplicity) to compute possible H-bonds for 3T0E (A). Mechanism for 3T0E TCR binding via CDR3β to peptide and DRB1 (B) shows possible H-bonds in *blue*, with improbable bonds blocked 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).

DQ H-Bonding Mechanisms

As shown in Table 1D, we noted the "highly-restricted" *d*V in Vβ for two of the TCR:pHLA-DQ structures (4OZH & 4OZG; 267 Å^3 and 276 Å^3, respectively) [35]. Here, the control structure is 4OZF (Vβ *d*V, 554 Å^3), where the pMHC is the same; in-fact, these three structures only differ by their 'non-germline' TCR, *i.e.*, by their CDR3 regions (*all 38 IMGT-computed CDR3 junctions shown in* **Suppl.3**). Accordingly, we analysed H-bonding 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 (CDRβ, 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-L111-A112-A113 in the these TCR is ultimately the result of somatic genetic construction of the CDR3β loop [**Suppl.3**; ref. 4].

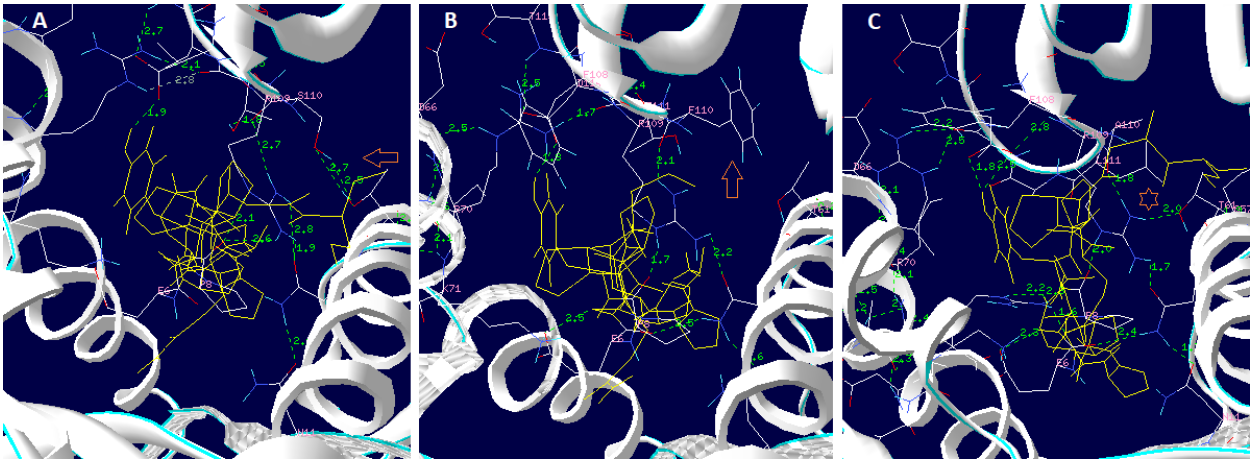

Figure 6 **H-bonds for three different TCR CDR3β, (from germline-identical TCR) on the same pHLA-DQ.** (A) 4OZH, showing S110:T61 H-bonds (*orange* arrow), (B) 4OZG, showing bulky F110 (orange arrow); both of which could preclude (C) , the conserved R109 H-bonds to T61 (*orange* star). *Swiss-Deepview* 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*.

Again, we can think about the mechanism sequentially. 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), 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 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 DP) which displayed *highly-restricted d*V, also shared the likelihood of a *charge-relay* mechanism of CDR3β binding, where an MHC to MHC bond is involved. Interestingly, this finding may suggest selection in *celiac disease* for CDR3β joints that can disrupt the N11-N62 H-bond, which is an indirect result of T61 as the source of R109 neutralization. For example, via a Th1->Th2 (Th17)

functional (reduced inflammation) shift [21]; perhaps a change in TCR affinity between receptors is superfluous, *i.e.*, if one accepts thymic-selection based on a 'window' of TCR affinity [4, 21, 22, 30]. Yet, a TCR's *signature dV*, including the relationship between *pitch* and *dV* shown here, could be the first "phenotype" isolated to an alternative mechanism; importantly, no-less direct than affinity to H-bonding networks at the TCR:pMHC interface.

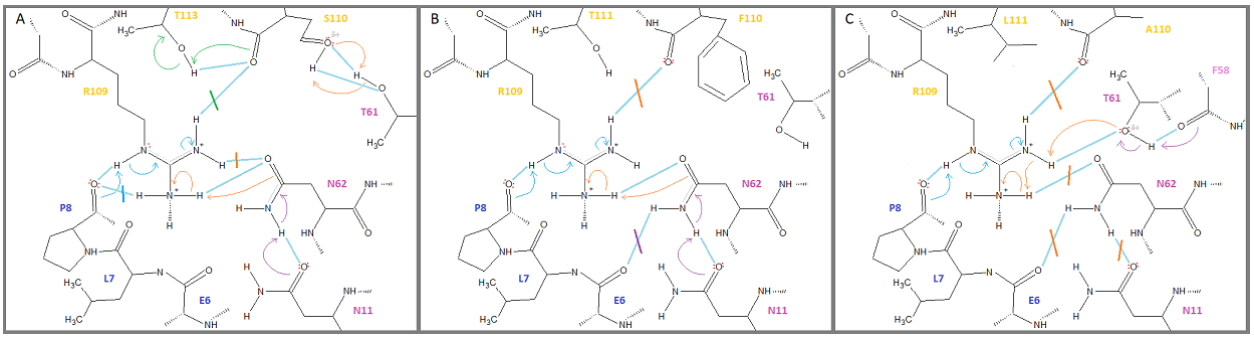

Figure 7 **H-bond mechanisms for three different CDR3β, (from germline-identical TCR) on the same pHLA-DQ.** (A) 4OZH, showing S110:T61 H-bonds (*orange* arrows), (B) 4OZG, showing bulky F110; both of which could preclude (C), the conserved R109 H-bonding to T61, as only observed for the nominal *dV* 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*.

Discussion:

We used spherical coordinates to derive a unique solved triple integral for each of the 38 V-domains (Figs. 1 & 2; **Suppl.1**). 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 of a cone with the vertex at the 'across-the-groove' MHC α-helix, where maximal CDR2 *scanning* ($d\theta$) 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 'across-the-groove' *tethering* via CDR3 (first formulated for TCR-Vα:pHLA-A2; ref. 17) was again, broadly apparent. Most importantly, here we found a linear relationship between V-domain *pitch*; calculated by:

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

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 very-low pitch may "restrict" $d\theta$, *i.e.*, the "CDR2-scanning" function. In retrospect, this confirms 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 [28, 29].

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 *d*V, *viz.*, $R^2 \approx 0.900$ (**Suppl.2**). Thus, *pitch* calculated from a given V-domain's *twist-tilt-sway* might be a kind of "hidden" correlate of TCR-selection; although there is not available crystallography on any thymic (selecting) ligands [36] to actually test this theory. 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 the H-bond shown (Fig. 3A; *PDB*, ref. 12); indeed, the key to understanding the TCR seems to be in this relationship between two 'induced-fit' TCR conformations [4, 28, 37]. What is suggested here is a conservation/approximation of *d*V for the V-domains when binding the *selecting* (thymus) and *activating* (peripheral) pMHC—despite that a given pitch on the two (potentially) different ligands *can* change dramatically, *usually* it does not.

Differences between MHC alleles with regard to transplantation, so-called 'permissive' versus 'non-permissive' haplotypes, could evoke an increased number of low-concentration peptides that a non-permissive graft might present with a given (selecting) *d*V [38, 39]. Alternatively/historically, *intrinsic* TCR affinity (*i.e.*, equilibrium and rate binding constants specified by the α/β TCR protein) has been used to explain thymic selection *and* MHC restriction [4, 40, 41]. However, as previously indicated, it is difficult to understand the intermediate ('window') nature of affinity that would result from selection in the context of a subsequent affinity-based MHC-restriction mechanism. Moreover, TCR affinity for pMHC is generally *micromolar* binding—weak reactions by antibody standards [42–44]. Related factors including the orientation of CDR2α, 3°-structure of α-chain constant domains and inter-chain disulfides, differences in CDR3α packing and the extra HV4, all likely contribute to pMHC-elicited monoclonal antibody Fab binding *not* like TCR [4, 45, 46]. Simply put, if positive/negative selection is based on TCR-affinity, it would be more logical that the selected repertoire reflects an *optimum* TCR-affinity; and thus, a qualitatively different mechanism might drive *MHC restriction*. In these and other regards, there is a *stereochemical* alternative to an affinity-limited (kinetics) binding reaction [31, 32].

Briefly, orientation is assumed to drive the TCR-pMHC reaction involving a high(er) energy "scanning" conformer, because scanning (leading to a suitable CDR2:MHC interface) has the effect of lowering the *transition-state* free energy, $\Delta G^{\ddagger}$ (*Curtin-Hammett* control) [31, 32, 47]. For example, in TCR like E8 (of 2IAM/2IAN) very little CDR2β scanning is apparently needed, due to the effect of the proposed *charge-relay* stabilised transition-state. Thus, *d*V is a universal consequence of the CDR3's chemistry—but usually the reaction requires more CDR2-scanning (*i.e.*, usually V-domains have larger *d*V values). Simply put, binding *specificity* could be controlled by *orientation* (*viz.*, $\Delta G^{\ddagger}$) and *not* densities of the reactants [40, 48, 49], or relative forward or reverse rates of the reaction, *per se* (ref. 31). One can envisage a free-energy 'trough' for successful induced-fit rates [37]; the possible corollary being its establishment during thymic selection [48]. Unfortunately, technology to directly investigate the issue is also limiting [reviewed in ref. 50]. Recently, 'ensemble refinement' of crystal data with MD simulations has suggested conformational diversity in the microsecond range, but the

conformational changes implicated here-in would be on the millisecond to minutes scale; in this regard, NMR of membrane-bound receptors may offer promise [50–54]

Finally, a dynamics mechanism driven by H-bonding networks is consistent with the knowledge that different T-cell dose-response curves are found at the same TCR affinity for the same pMHC [40, 55–58]. Even if we assume similar TCR-affinity for the three 4OZ structures, it appears that somatic selection directly effects H-bonds implicated in V-domain dynamics (*d*V), and (indirectly) germline:germline contacts between CDR2 and the MHC α-helices (*d*θ) [24, 44, 49, 59–61]. Stabilization of H-bonding via charge-relays is implicated for V-domains with very-low *pitch* angles, as specified by the *twist-tilt-sway* (*viz.*, orientation) and *highly-restricted d*V. What can change is the *measured* pitch angle. Thus, an orientation selected on a thymic pMHC-II could specify particular *d*V (characterized by the calculated pitch); and speculatively, this *d*V must be approximated on a peripheral pMHC-II, *i.e.*, leading to 'recognition' of this pMHC-II during an immune response. The 'plasticity' of pitch during recognition could further 'induced fit' and ostensibly allows some recapitulations on a pMHC-II with the different (peripheral) peptide; in effect, *MHC restriction* [4, 7, 15, 30]. Again, usually we see *k* values close to unity. While there are myriad 'downstream' implications, taken together, these data support that V-domain rotation and germline to germline contacts between TCR and pMHC *both* depend upon somatic TCR H-bonding with a highly-conserved, *across-the-groove* MHC motif [17, 19, 24, 30, 44]. In this case, the DRB1/DQB1 sequence, Y60-W-N-S-Q-K-D-L/I-L-E-Q/R70, which spans closest contacts across CDR3β of the DR and DQ TCR:pMHC-II structures (Tables 1A–1D), conserved throughout human and primate alleles (not shown; *NCBI BLAST* data, 09/03/2019).

Methods:

PDB Analysis

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* (ω), *tilt* (λ), and *sway* (σ) 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 V-domain's 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. Linear regression analysis by *MS-Excel* (**Suppl.2**); 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 footnotes and the text.

CDR3 Joint Analysis

Nucleotide sequences for all CDR3 of TCR were specified from *PDB* files (**Suppl.3**). 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. The pMHC contacts were computed by *bond-length* measurement from the *PDB* files with *VMD*, as previously described.

H-Bonding Mechanisms

Hydrogen (H) bonds were estimated either with *VMD*, or (more accurately) with *Swiss*. In *Swiss*, H-bonds are calculated 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 [31].

Declarations:

Ethics approval and consent to participate: yes; *Consent to publish:* yes; *Availability of data and materials:* All primary data are public and available at www.ncbi.nlm.nih.gov, and all analytic data are either in the paper, or supplementary materials; *Competing interests:* Dr. Murray and *Xenolaüs Genetics LLC* declare no competing interests regarding the research, or its publication; *Author Contributions:* Dr. Murray did the research and wrote the manuscript.

Acknowledgements:

This work was possible because of dedicated crystallographers, molecular biologists and computer scientists providing public availability of coordinates, sequences and software; thank you.

Supplementary material:

[Suppl.1.pdf]
[Suppl.2.pdf]
[Suppl.3.pdf]

References:

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

2. Tonegawa S. Somatic generation of antibody diversity. *Nature* (1983) **302**:575–581.

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

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

5. Wieczorek M, Abualrous ET, 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 (2017) **8**:292. doi: 10.3389/fimmu.2017.00292

6. Nisonoff A, Hopper JE, Spring SB. *The Antibody Molecule*, 1st ed. New York: Academic Press (1975).

7. McDevitt HO, and Bodmer WF. HL-A, immune response genes, and disease. *Lancet* (1974) **1**:1269–75. doi.org/10.1016/S0140-6736(74)90021-X

8. Brown JH, Jardetzky T, Saper MA, Samraoui B, Bjorkman PJ, Wiley DC. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature* (1988) **332**:845–50.

9. Davis MM, and Bjorkman PA. T-cell antigen receptor genes and T-cell recognition. *Nature* (1988) **334**:395–402. doi: 10.1038/334395a0

10. Hennecke J, Wiley DC. 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* (2001) **195**:571–81.

11. Li Y, Huang Y, Lue J, Quandt JA, Martin R, Mariuzza RA. 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* (2005) **24**:2968–79.

12. Deng L, Langley RJ, Brown PH, 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* (2007) **8**:398–408.

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

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

15. Sharon E, Sibener LV, Battle A, Fraser HB, Garcia KC, Pritchard JK. Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat Genet* (2016) **48**:995–1002. doi: 10.1038/ng.3625

16. Gras S, Chadderton J, Del Campo CM, Farenc C, Wiede F, Josephs TM, *et al*. Reversed T Cell Receptor Docking on a Major Histocompatibility Class I Complex Limits Involvement in the Immune Response. *Immunity* (2016) **45**:749–760. doi: 10.1016/j.immuni.2016.09.007

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

18. Beringer DX, Kleijwegt FS, Wiede F, van der Slik AR, Loh KL, Petersen J, *et al*. T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nat Immunol* (2015) **16**:1153–61. doi: 10.1038/ni.3271

19. Burrows SR, Chen Z, Archbold JK, Tynan FE, 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* (2010) **107**:10608–13. doi: 10.1073/pnas.1004926107

20. Ashwell JD, Chen C, Schwartz RH. High frequency and nonrandom distribution of alloreactivity in T cell clones selected for recognition of foreign antigen in association with self class II molecules. *J Immuno*l (1986) **136**:389–95.

21. Murray JS, Madri J, Tite J, Carding SR, Bottomly K. MHC control of CD4+ T cell subset activation. *J Exp Med* (1989) **170**:2135–40.

22. Gottschalk RA, Hathorn MM, Beuneu H, Corse E, Dustin ML, Altan-Bonnet G, Allison JP. Distinct influences of peptide-MHC quality and quantity on in vivo T-cell responses. *Proc Natl Acad Sci USA* (2012) **109**:881–6. doi: 10.1073/pnas.1119763109

23. Kasahara M, Vazquez M, Sato K, McKinney EC, Flajnik MF. Evolution of the major histocompatibility complex: class II A cDNA clones from the cartilaginous fish. *Proc Natl Acad Sci USA* (1992) **89**:6688–92.

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

25. Chen J, Thirumalai D. Interface Residues That Drive Allosteric Transitions Also Control the Assembly of l-Lactate Dehydrogenase. *J Phys Chem B* (2018) **122:**11195–11205. doi: 10.1021/acs.jpcb.8b06430

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

27. Das BB, Park SH, Opella SJ. Membrane protein structure from rotational diffusion. *Biochim Biophys Acta* (2015) **1848**:229–45. doi: 10.1016/j.bbamem.2014.04.002

28. Berg JM, Tymoczko JL, Stryer L. *Biochemistry*, 5th ed. New York: WH Freeman (2002).

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

30. Marrack P, Krovi SH, Silberman D, White J, Kushnir E, Nakayama M, *et al*. The somatically generated portion of T cell receptor CDR3α contributes to the MHC allele specificity of the T cell receptor. e*life* (2017) **6. pii**:e30918. doi: 10.7554/eLife

31. Carey FA, Sundberg RJ. *Advanced Organic Chemistry, pt. A. Structure and Mechanism*, 2nd ed. New York: Plenum Press (1984).

32. Hernández G, Anderson JS, LeMaster DM. Assessing the native state conformational distribution of ubiquitin by peptide acidity. *Biophys Chem* (2010) **153**:70–82. doi: 10.1016/j.bpc.2010.10.006

33. Koch M, Camp S, Collen T, Avila D, Salomonsen J, Wallny HJ, *et al*. Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* (2007) **27**:885-99.

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

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

36. Zerrahn J, Held W, Raulet DH. The MHC reactivity of the T cell repertoire prior to positive and negative selection. *Cell* (1997) **88**:627–36.

37. Pusuluk O, Farrow T, Cemsinan D, Burnett K, Vedral V. Proton tunneling in hydrogen bonds and its implications in an induced-fit model of enzyme catalysis. *Proc R Soc A* (2018) **474**:20180037. dx.doi.org/10.1098/rspa.2018.0037

38. Clement CC, Becerra A, Yin L, Zolla V, Huang L, Merlin S, *et al*. The dendritic cell major histocompatibility complex II (MHC II) peptidome derives from a variety of processing pathways and includes peptides with a broad spectrum of HLA-DM sensitivity. *J Biol Chem* (2016) **291**:5576–95. doi: 10.1074/jbc.M115.655738

39. Arrieta-Bolaños E, Crivello P, Metzing M, Meurer T, Ahci M, Rytlewski J, *et al*. Alloreactive T cell receptor diversity against structurally similar or dissimilar HLA-DP antigens assessed by deep sequencing. *Front Immunol* (2018) **9**:280. doi: 10.3389/fimmu.2018.00280

40. Murray JS. How the MHC selects Th1/Th2 immunity. *Immunol Today* (1998) **19**:157-63.

41. Faro J, Castro M, Molina-París C. A unifying mathematical framework for experimental TCR-pMHC kinetic constants. *Sci Rep* (2017) **7**:46741. doi: 10.1038/srep46741

42. Murray JS, Brown JC. Measurement of association constants in ELISA. Reactions between solid-phase antibody and fluid-phase biotinylated antigen. *J Immunol Methods* (1990) **127**:25-8.

43. Kamat V, Rafique A. Designing binding kinetic assay on the bio-layer interferometry (BLI) biosensor to characterize antibody-antigen interactions. *Anal Biochem* (2017) **536**:16–31. doi: 10.1016/j.ab.2017.08.002

44. Gee MH, Sibener LV, Birnbaum ME, Jude KM, Yang X, Fernandes RA, *et al*. Stress-testing the relationship between T cell receptor/peptide-MHC affinity and cross-reactivity using peptide velcro. *Proc Natl Acad Sci USA* (2018) **115**:E7369–E7378. doi: 10.1073/pnas.1802746115

45. Dunbar J, Knapp B, Fuchs A, Shi J, Deane CM. Examining variable domain orientations in antigen receptors gives insight into TCR-Like antibody design. *PLOS Comput Biol* (2014) **10**:e1003852. doi: 10.1371/journal.pcbi.1003852

46. Hülsmeyer M, Chames P, Hillig RC, Stanfield RL, Held G, Coulie PG, *et al*. A major histocompatibility complex-peptide-restricted antibody and T cell receptor molecules recognize their target by distinct binding modes. *J Biol Chem* (2005) **280**:2972–80. doi: 10.1074/jbc.M411323200

47. Lodola A, Mor M, Zurek J, Tarzia G, Piomelli D, Harvey JN, Mulholland AJ. Conformational effects in enzyme catalysis: reaction via high energy conformation in fatty acid amide hydrolase. *Biophys J* (2007) **92**:L20–22. doi: 10.1529/biophyj.106.098434

48. Pielak RM, O'Donoghue GP, Lin JJ, Alfieri KN, Fay NC, Low-Nam ST, Groves JT. Early T cell receptor signals globally modulate ligand:receptor affinities during antigen discrimination. *Proc Natl Acad Sci U S A* (2017) **114**:12190–12195. doi: 10.1073/pnas.1613140114

49. Deng L, Langley RJ, Wang Q, Topalian SL, Mariuzza RA. Structural insights into the editing of germ-line-encoded interactions between T-cell receptor and MHC class II by Vα CDR3. *Proc Natl Acad Sci USA* (2012) **109**:14960–5. doi: 10.1073/pnas.1207186109

50. Natarajan K, Jiang J, May NA, Mage MG, Boyd LF, McShan AC, *et al*. The Role of Molecular Flexibility in Antigen Presentation and T Cell Receptor-Mediated Signaling. *Front Immunol* (2018) **9**:1657. doi: 10.3389/fimmu.2018.01657

51. Buckle AM, Borg NA. Integrating Experiment and Theory to Understand TCR-pMHC Dynamics. *Front Immunol* (2018) **9**:2898. doi: 10.3389/fimmu.2018.02898

52. Pfleger C, Minges A, Boehm M, McClendon CL, Torella R, Gohlke H. Ensemble- and Rigidity Theory-Based Perturbation Approach To Analyze Dynamic Allostery. *J Chem Theory Comput* (2017) **13**:6343–57. doi: 10.1021/acs.jctc.7b00529

53. Carlson GM, Fenton AW. What Mutagenesis Can and Cannot Reveal About Allostery. *Biophys J* (2016) **110**:1912–23. doi: 10.1016/j.bpj.2016.03.021

54. Sibener LV, Fernandes RA, Kolawole EM, Carbone CB, Liu F, McAffee D, *et al*. Isolation and visualization of a structural trigger that uncouples TCR signaling from pMHC binding. *Cell* (2018) **174**:672–687.e27. doi: 10.1016/j.cell.2018.06.017

55. De Oliveira DB, Harfouch-Hammoud E, Otto H, Papandreou NA, Stern LJ, Cohen H, *et al*. Structural analysis of two HLA-DR-presented autoantigenic epitopes: crucial role of peripheral but not central peptide residues for T-cell receptor recognition. *Mol Immunol* (2000) **37**:813–25.

56. Holzer U, Kwok WW, Nepom GT, Buckner JH. Differential antigen sensitivity and costimulatory requirements in human Th1 and Th2 antigen-specific CD4+ cells with similar TCR avidity. *J Immunol* (2003) **170**:1218–23.

57. Lin JJY, Low-Nam ST, Alfieri KN, McAffee DB, Fay NC, Groves JT. Mapping the stochastic sequence of individual ligand-receptor binding events to cellular activation: T cells act on the rare events. *Sci Signal* (2019) **12**:pii: eaat8715. doi: 10.1126/scisignal.aat8715

58. Galvez J, Galvez JJ, 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* (2019) **10**:349. doi: 10.3389/fimmu.2019.00349

59. Scharf L, West AP Jr, Sievers SA , Chen C, Jiang S, Han G, *et al*. Structural basis for germline antibody recognition of HIV-1 immunogens. *eLife* (2016) **5**:e13783. 10.7554/eLife.13783

60. Lu J, Van Laethem F, Bhattacharya A, Craveiro M, Saba I, Chu J, *et al*. Molecular constraints on CDR3 for thymic selection of MHC-restricted TCRs from a random pre-selection repertoire. *Nat Commun* (2019) **10**:1019. doi: 10.1038/s41467-019-08906-7

61. Khosravi-Maharlooei M, Obradovic A, Misra A, Motwani K, Holzl M, Seay HR, *et al*. Crossreactive public TCR sequences undergo positive selection in the human thymic repertoire. *J Clin Invest* (2019) **129**:2446–2462. doi: 10.1172/JCI124358

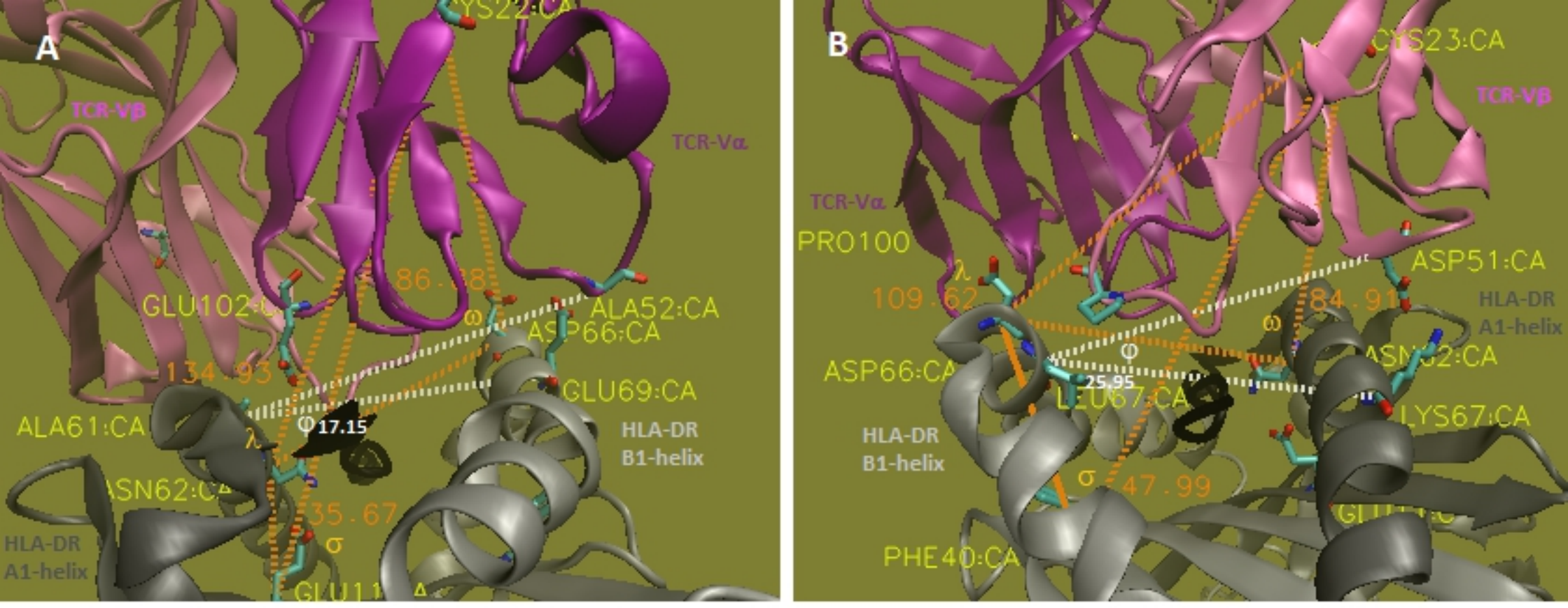
A
TCR-Vβ
TCR-Vα
GLU102:CA
ALA52:CA
ASP66:CA
GLU69:CA
ALA61:CA
φ 17.15
HLA-DR
B1-helix
ASN62:CA
HLA-DR
A1-helix
B
CYS23:CA
TCR-Vβ
TCR-Vα
PRO100
ASP51:CA
HLA-DR
A1-helix
ASP66:CA
φ
25.95
LYS67:CA
HLA-DR
B1-helix
PHE40:CA

Figure 1

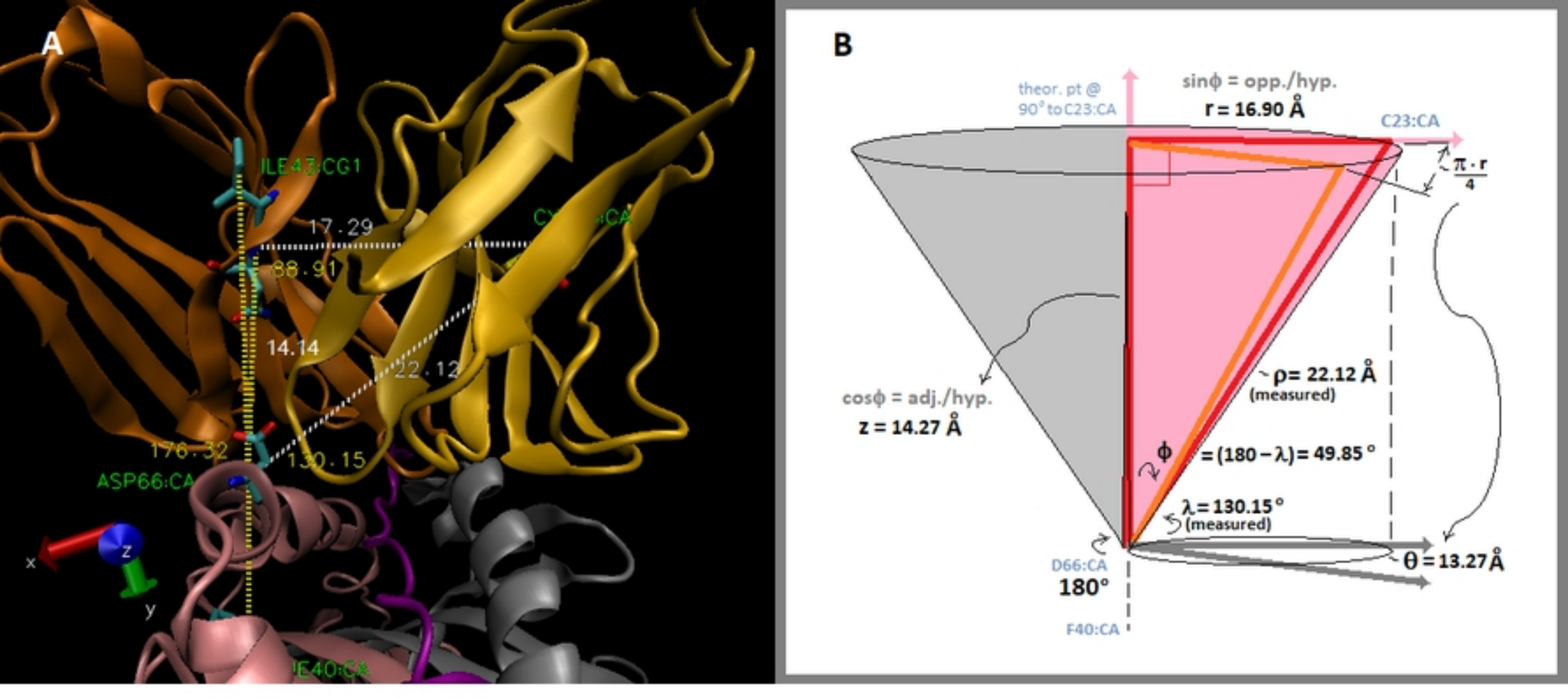
A
ILE43:CG1
17.29
88.91
14.14
22.12
176.32
130.15
ASP66:CA
x
z
y
B
theor. pt @
90° to C23:CA
sinϕ = opp./hyp.
r = 16.90 Å
C23:CA
π · r
4
cosϕ = adj./hyp.
z = 14.27 Å
ρ= 22.12 Å
(measured)
ϕ
= (180 − λ) = 49.85 °
λ = 130.15 °
(measured)
D66:CA
180°
θ = 13.27 Å
F40:CA

Figure 2

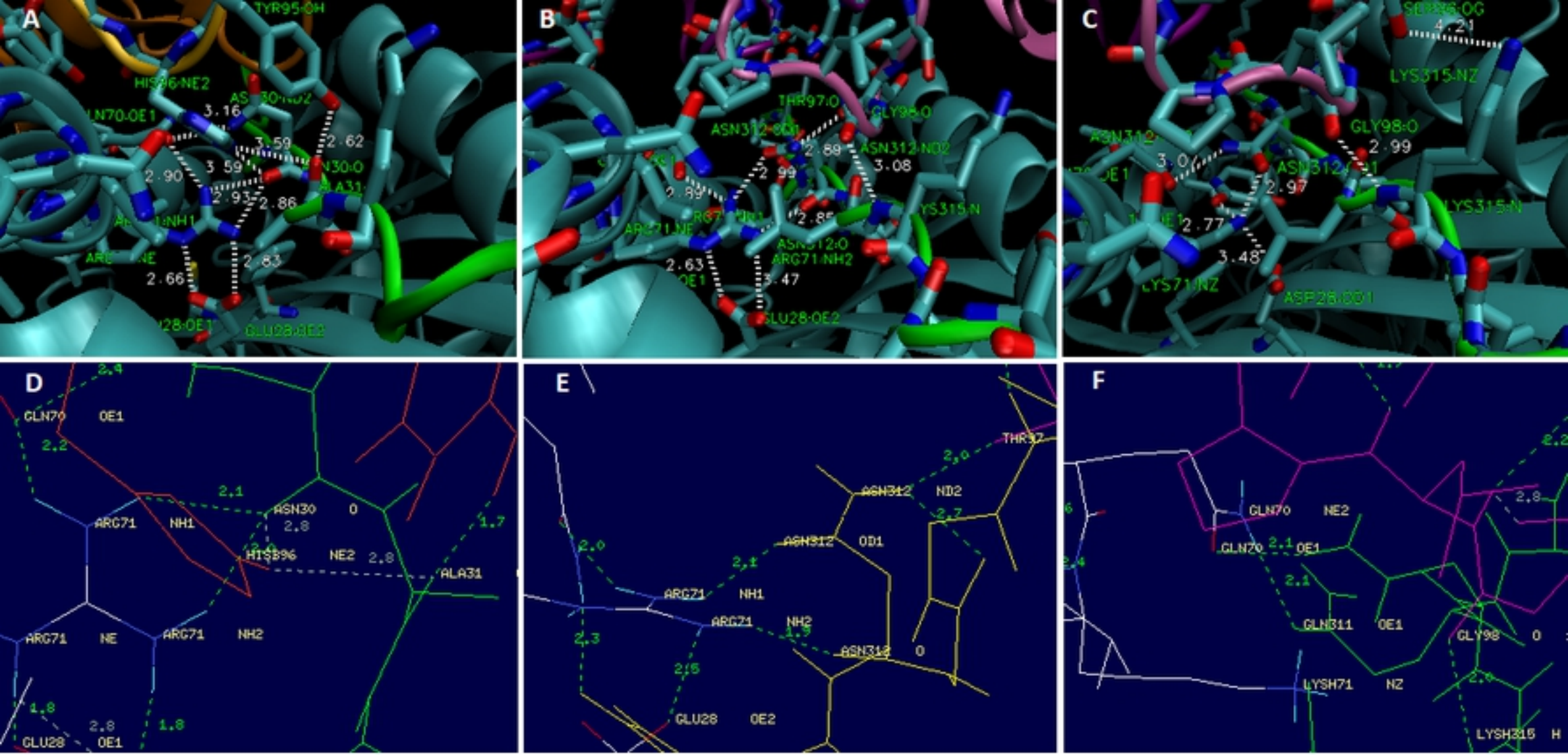
A
B
C
D
E
F
TYR95:OH
HIS96:NE2
GLN70:OE1
ASN30:ND2
ASN30:O
ALA31
ARG71:NH1
ARG71:NE
GLU28:OE2
THR97:O
GLY98:O
ASN312:OD1
ASN312:ND2
LYS315:N
ARG71:NE
ARG71:NH2
ASN312:O
LYS315:NZ
GLY98:O
ASN312:OD1
LYS71:NZ
ASP28:OD1
GLN70 OE1
ARG71 NH1
ASN30 O
HIS96 NE2
ALA31
ARG71 NE
ARG71 NH2
GLU28 OE1
THR97
ASN312 ND2
ASN312 OD1
ARG71 NH1
ARG71 NH2
ASN312 O
GLU28 OE2
GLN70 NE2
GLN70 OE1
GLN311 OE1
GLY98 O
LYSH71 NZ
LYSH315 H

Figure 3

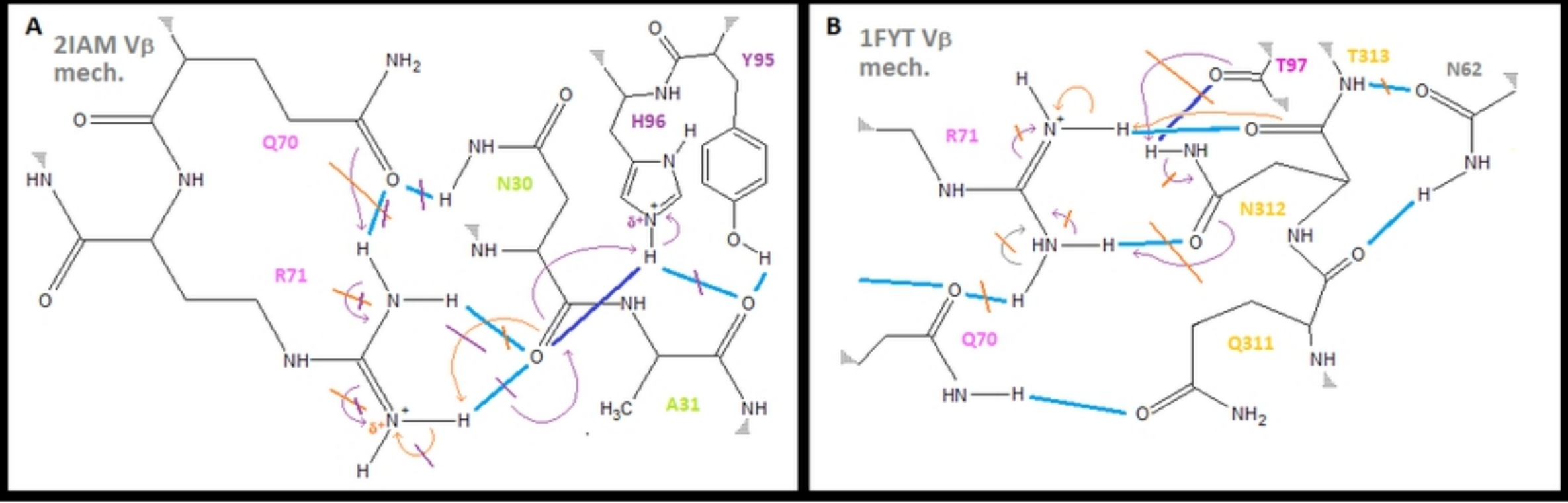
A
2IAM Vβ
mech.
Q70
R71
N30
H96
Y95
A31
B
1FYT Vβ
mech.
R71
T97
T313
N62
N312
Q70
Q311

Figure 4

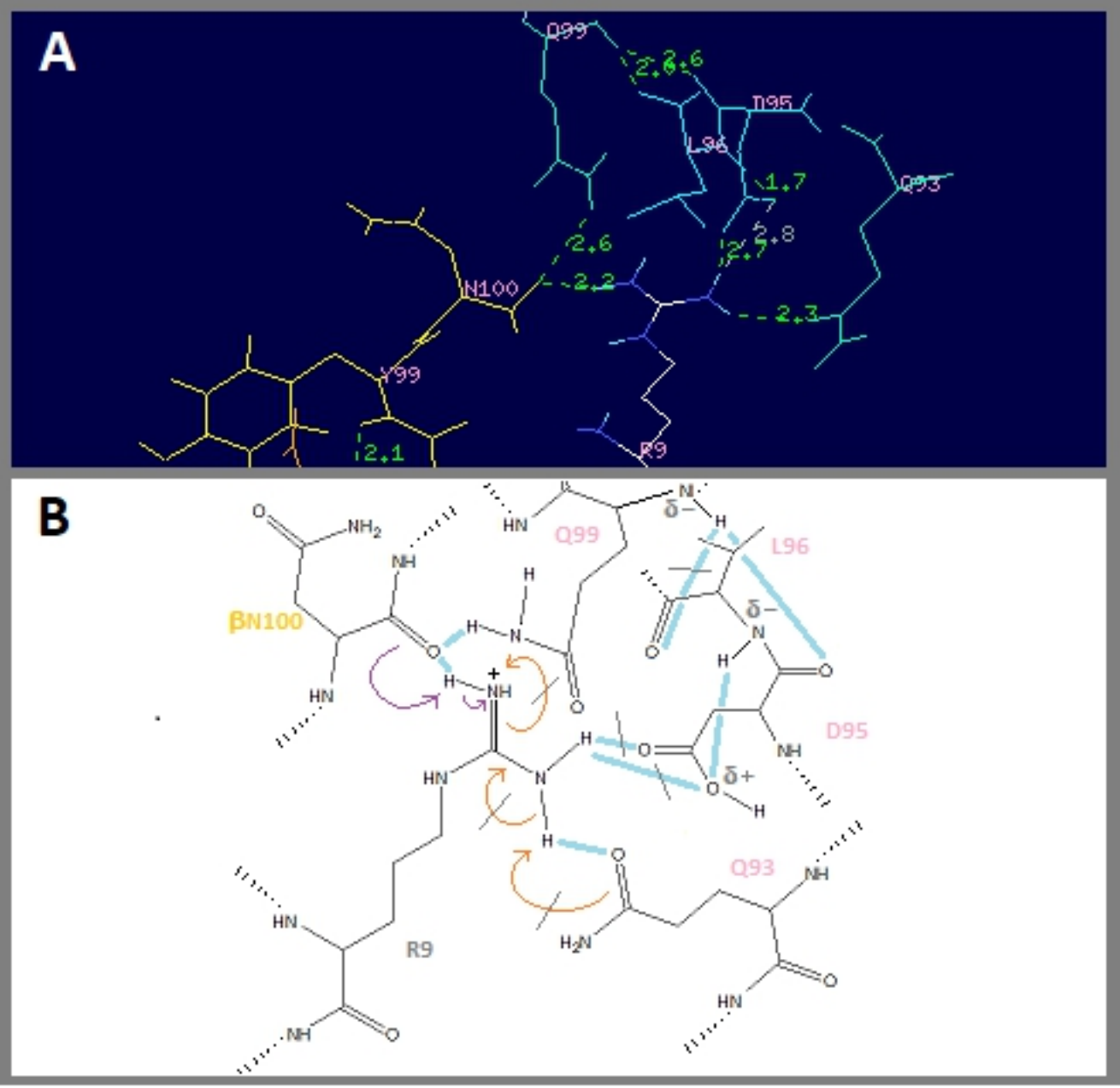
A
Q99
2.6
2.6
D95
L96
1.7
Q93
2.8
2.6
2.7
N100
2.2
2.3
Y99
2.1
R9
B
βN100
Q99
L96
D95
Q93
R9

Figure 5

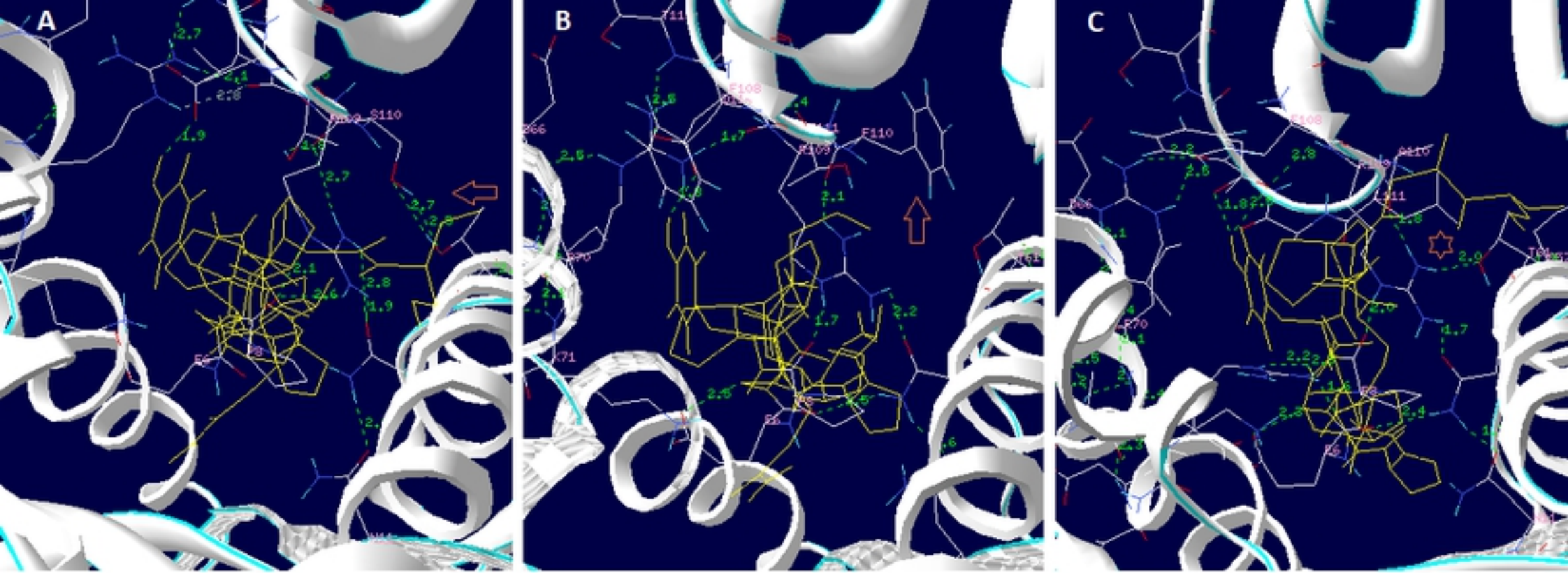

Figure 6

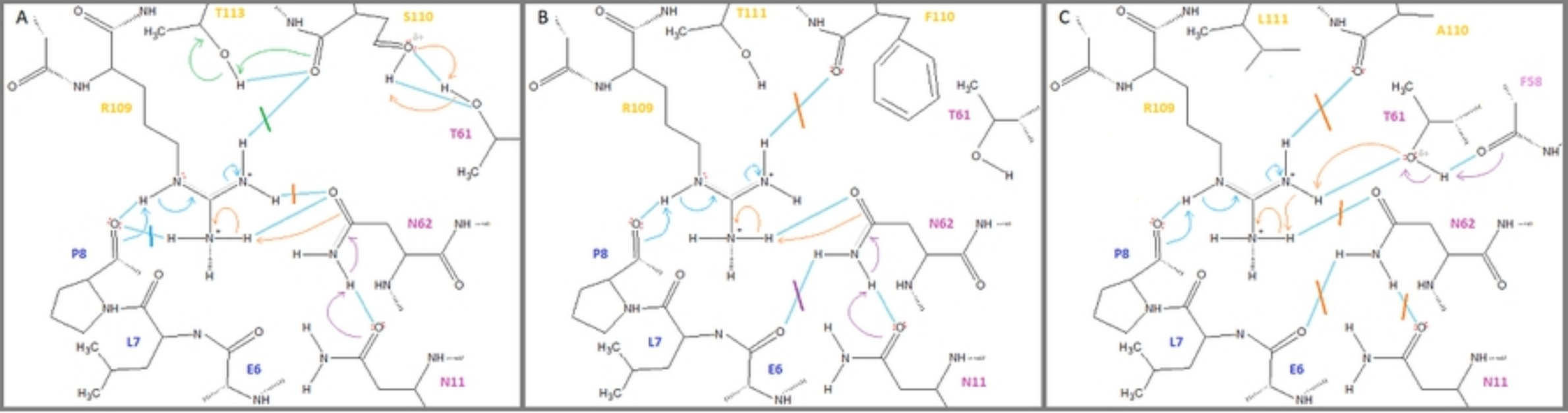
A
T113
S110
R109
T61
N62
P8
L7
E6
N11
B
T111
F110
R109
T61
N62
P8
L7
E6
N11
C
L111
A110
R109
T61
F58
N62
P8
L7
E6
N11

Figure 7