University of Massachusetts Medical School

eScholarship@UMMS

Program in Bioinformatics and Integrative Biology Publications and Presentations

Program in Bioinformatics and Integrative Biology

2016-03-01

How structural adaptability exists alongside HLA-A2 bias in the human alphabeta TCR repertoire

Sydney J. Blevins
University of Notre Dame

Et al.

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/bioinformatics_pubs

Part of the Amino Acids, Peptides, and Proteins Commons, Bioinformatics Commons, Cells Commons, Chemical Actions and Uses Commons, Computational Biology Commons, Immunology and Infectious Disease Commons, Investigative Techniques Commons, and the Structural Biology Commons

How structural adaptability exists alongside HLA-A2 bias in the human αβ TCR repertoire

Sydney J. Blevins[a,b], Brian G. Pierce[c], Nishant K. Singh[a,b], Timothy P. Riley[a,b], Yuan Wang[a,b], Timothy T. Spear[d,e], Michael I. Nishimura[d,e], Zhiping Weng[f], and Brian M. Baker[a,b,1]

[a]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556; [b]Harper Cancer Research Institute, University of Notre Dame, Notre Dame, IN 46556; [c]University of Maryland Institute for Bioscience and Biotechnology Research, Rockville, MD 20850; [d]Department of Surgery, Loyola University Chicago, Maywood, IL 60153; [e]Cardinal Bernardin Cancer Center, Loyola University Chicago, Maywood, IL 60153; and [f]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01605

How T-cell receptors (TCRs) can be intrinsically biased toward MHC proteins while simultaneously display the structural adaptability required to engage diverse ligands remains a controversial puzzle. We addressed this by examining αβ TCR sequences and structures for evidence of physicochemical compatibility with MHC proteins. We found that human TCRs are enriched in the capacity to engage a polymorphic, positively charged "hot-spot" region that is almost exclusive to the α1-helix of the common human class I MHC protein, HLA-A*0201 (HLA-A2). TCR binding necessitates hot-spot burial, yielding high energetic penalties that must be offset via complementary electrostatic interactions. Enrichment of negative charges in TCR binding loops, particularly the germ-line loops encoded by the TCR Vα and Vβ genes, provides this capacity and is correlated with restricted positioning of TCRs over HLA-A2. Notably, this enrichment is absent from antibody genes. The data suggest a built-in TCR compatibility with HLA-A2 that biases receptors toward, but does not compel, particular binding modes. Our findings provide an instructional example for how structurally pliant MHC biases can be encoded within TCRs.

T-cell receptor | peptide/MHC | structure | binding | MHC bias

MHC restriction is a hallmark of T-cell recognition (1). The first crystallographic structures of T-cell receptors (TCRs) in complex with antigen revealed the structural basis for MHC restriction, illustrating how TCRs bind a composite ligand consisting of the MHC protein and its presented peptide (pMHC) (2, 3). Around the same time, evidence began to mount that T cells, and by extension their TCRs, maintain a physical bias toward MHC proteins (4). The mechanism underlying this bias, and indeed its very existence, has remained controversial. On one hand, structural evidence has been presented that germ-line complementarity-determining region (CDR) loops incorporate evolved regions with intrinsic affinities toward MHC proteins (5). On the other hand, the requirement for coreceptor during thymic selection provides a mechanism for selecting TCRs that bind MHC (6–9).

Efforts to reconcile these positions have emphasized that they are not mutually exclusive (9–11). However, it is clear that any intrinsic bias TCRs have toward MHC proteins must exist alongside the need for TCRs to structurally adapt to different ligands or alterations in the TCR–pMHC interface (12). For example, changes to a peptide alone are sufficient to alter how the same TCR sits over the same MHC protein, altering TCR–MHC atomic contacts (13, 14). Similar results are seen with changes to TCR variable domains or hypervariable loops, or even MHC mutations (15–18). Perhaps not surprisingly then, there are no germ-line–based TCR–MHC contacts that are fully conserved at the atomic level when structures with TCRs sharing the same variable domain are compared. For example, TCRs sharing Vβ 8.2 make similar contacts with the helices of the class II MHC protein I-A, but actual contacts are modulated by different peptides, hypervariable loops, and Vα domains (15, 19–21). Although such variances have been noted and even predicted (5, 22), they nonetheless make it difficult to clearly identify interactions encoding a TCR–MHC bias (12).

Furthermore, in addition to class I and class II MHC proteins, TCRs can bind a variety of nonclassical MHC molecules, non-MHC ligands (6, 23, 24), and as demonstrated very recently can even bind MHC molecules in orientations reverse from the "canonical" diagonal binding mode (25). Thus, whatever MHC biases actually do exist do not exempt a receptor from identifying alternate binding solutions or targets.

A possibility often overlooked in discussions of TCR binding is that an evolutionarily encoded MHC bias does not need to specify attractive (i.e., energetically favorable) interactions. A more fundamental feature is structural or physicochemical complementarity, or simply "compatibility." A set of structurally/chemically compatible interactions whose formation is energetically weak, neutral, or even slightly unfavorable is better than a set of noncomplementary, unfavorable interactions. A well-understood example in protein chemistry is interactions between opposing charges in protein structures. Formation of such interactions require removal of the charges from bulk water, which comes with a substantial, unfavorable "desolvation" penalty (26, 27). The coulombic energy from the interaction between the charges is often insufficient to offset the desolvation penalty (28). However, a weak (or even unfavorable) charge–charge interaction is still energetically better than two uncompensated charges buried in a protein. In this way, charges can influence structural specificity even if they do not drive binding. Similar arguments can be made for shape complementarity: an energetically neutral but complementary set of interactions is better than a forced set of interactions that might result in steric hindrance or require costly conformational changes. This concept of compatibility is well understood in antibody–antigen interactions (29).

Significance

T-cell receptor (TCR) recognition of antigenic peptides presented by major histocompatibility complex (MHC) proteins defines specificity in cellular immunity. Evidence suggests that TCRs are intrinsically biased toward MHC proteins, yet how this bias coexists alongside the considerable structural variability that is necessary for TCRs to engage different ligands has been a longstanding puzzle. By examining structural and sequence data, we found evidence that human αβ TCRs have an inherent compatibility with structural and chemical properties of MHC proteins. This compatibility leads TCRs to an intrinsic MHC bias but does not compel the formation of particular modes of binding, providing a solution to how TCRs can be MHC-biased but still structurally adaptable.

Considering the above arguments, evolution could have specified a TCR bias toward MHC by facilitating structural/chemical compatibility. Because they need not be energetic drivers, regions of compatibility might be expected to be adaptable and even overridden when sufficient "glue" can be found elsewhere in an interface. To examine this possibility, we studied human αβ TCRs that recognize the highly prevalent human class I MHC protein, HLA-A*0201 (HLA-A2 or simply A2). Given the high occurrence of A2 in human populations (30), we reasoned that, should it exist, a signature of evolutionarily encoded compatibility should be present within the large amount of structural data available for A2-binding TCRs. We quantified sequence variability and used this to guide queries of electrostatic compatibility. In doing so, we identified a set of interactions between TCRs and a polymorphic region on A2 that are strongly conserved in TCR–A2 structures. These interactions, which involve a positively charged polymorphic "hot-spot" region that is almost exclusive to the A2 α1-helix (31–33), are correlated not only with A2 binding but also restricted positioning over the MHC. Different TCRs engage this hot spot differently, in some cases using hypervariable rather than (or in addition to) germ-line loops. Nonetheless, we show that TCR germ-line loops are enriched in the negatively charged amino acids needed to engage the hot spot and that this enrichment is absent from the corresponding loops in antibodies. Thus, the capacity for achieving structural and chemical complementarity with A2 is not only built-in to the human TCR repertoire but present in a way that facilitates the structural adaptability needed to engage diverse pMHC ligands. Although examples of other forms of evolved compatibility remain to be identified, our findings support the conclusion that TCRs have an evolutionarily encoded, inherent compatibility with MHC proteins that facilitates recognition but does not compel particular binding solutions. This can explain how TCRs can show MHC bias in functional experiments, while still displaying considerable structural adaptability, and sometimes violate expectations.

Results

TCR Germ-Line Loops Are Diverse. We began examining TCR germ-line loops by generating sequence logos for each CDR1 and CDR2 loop, focusing on those Vα/Vβ segments encoding the most common loop lengths (Fig. 1A; see Table S1 for loop statistics). Loop definitions and sequences were taken from the ImMunoGeneTics (IMGT) database, except that we used an extended definition for CDR2β that better accounts for loop structure in structures of TCRs bound to pMHC (Fig. S1A). We also separately calculated the Shannon entropy values for each position in each loop. Shannon entropies quantify diversity, with a value of zero indicating conservation of 1 aa and the maximum value of 4.32 indicating equal representation of all 20 aa. A previous analysis of TCR variable genes identified three general Shannon entropy categories: H_0 ($0 \leq H \leq 1$), H_1 ($1 < H \leq 2$), and H_2 ($H > 2$) (34). H_0 sites are highly conserved and associated with sites essential for fold or function. H_1 sites are semiconserved, and H_2 sites are highly variable with no obvious conservation pattern.

Only a single H_0 site in the TCR germ-line loops was identified, the histidine at position 29 in CDR1β, whose presence has been noted previously and is thought to be involved in loop organization (35). Five H_1 sites were identified, one each in CDR1α, CDR2α, and CDR1β, and two in CDR2β. The remaining sites are all H_2. This indicates that the majority of TCR germ-line loop positions are highly variable, with only a small fraction demonstrating sequence and presumably functional conservation.

There Is No Obvious Signal for Engaging HLA-A2 in Sequences of TCR Germ-Line Loops. We next asked whether binding a particular MHC protein dictates preferences in TCR germ-line loops. We assembled a set of structurally or biophysically characterized, natural human TCRs known to recognize peptides presented by the most common human class I MHC protein, HLA-A*0201 (A2) (Table S2, Complexes with HLA-A2). After accounting for duplicate genes, we used this set to compute CDR1/2 sequence logos and entropies (Fig. 1B). In all loops, diversity is reduced compared with the larger analysis of all sequences. To ask whether this result is attributable to
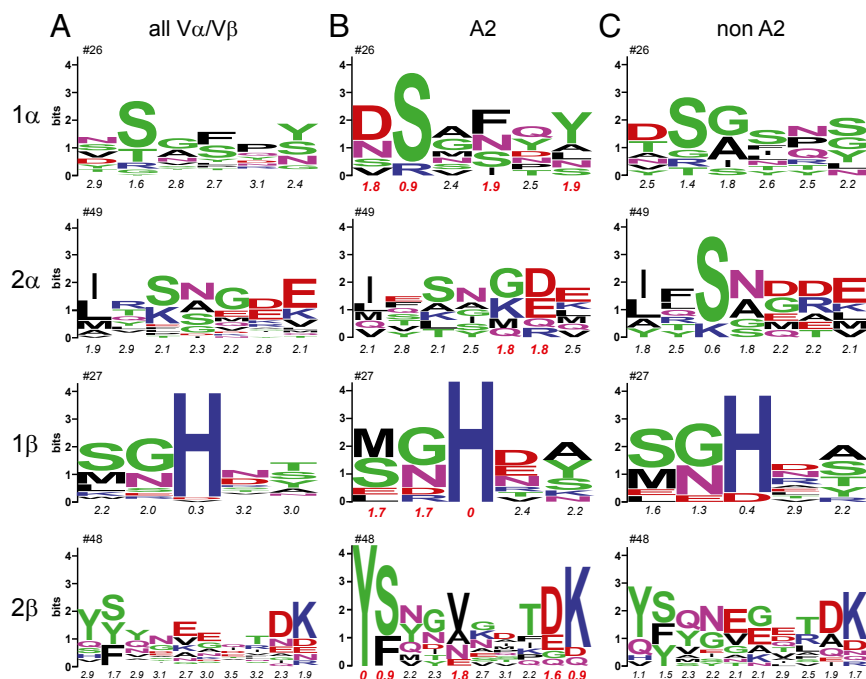
Fig. 1. There is no sequence signature for TCR binding HLA-A2. (A) Sequence logos for the most common germ-line loop lengths in human TCR Vα/Vβ genes. The numbers at the *Top* give the sequence number for the first amino acid in the loop. Numbers below give Shannon entropy values for each loop position. (B) Sequence logos for the set of structurally/physically characterized human TCRs that bind A2. Red entropy numbers indicate highly conserved H_0/H_1 sites. (C) Sequence logos for the set of structurally characterized human TCRs that bind human class I MHC proteins other than A2.

restriction toward A2 or just reflective of a smaller sample size, we randomly generated 10 identically sized datasets of germ-line loop sequences from the larger set of all human germ-line loop sequences. The average entropies of the loops of the A2-binding TCRs were all within one SD of the averages of the randomly generated datasets, indicating that reduced diversity in the data in Fig. 1B is not likely a function of engaging A2 (Table S3).

Examining the entropy data for the A2-binding TCRs in more detail, there are at least two sites per CDR loop that are H_0 or H_1 (highly conserved or semiconserved; Shannon entropy values highlighted in red in Fig. 1B). For comparison, we computed sequence logos and entropies for a set of structurally characterized, natural human TCRs shown to bind human class I MHC proteins other than A2 (Fig. 1C; TCRs in Table S2, Complexes with other class I HLA proteins). Although there are notable similarities and differences, in every case, the most common amino acids in the A2-binding H_0 or H_1 sites are also present in the equivalent position in the non–A2-binding sequences. χ^2 tests for independence indicated that the two sets of sequences are not statistically distinct. Thus, the available data indicate that there are no germ-line amino acids or amino acid sequences exclusively required for engaging A2.

TCRs That Bind HLA-A2 Show No Conserved Pairwise Amino Acid Contacts with A2 but a Propensity for Engaging a Key Molecular Contact on the MHC α1-Helix. We next asked whether TCRs use H_0 or H_1 sites to form preferred contacts with particular MHC proteins. Within the set of A2-binding TCRs, excluding redundant structures with altered peptides or mutants, there are 13 high-resolution crystal structures of naturally occurring human TCRs bound to peptide/A2 complexes (Table S2, Complexes with HLA-A2). We asked whether the most common amino acid in each H_0/H_1 position forms common interactions with A2. For example, does the strongly conserved serine at position 27 in CDR1α ($H = 0.9$) commonly contact a certain residue on A2?

Of a total of 249 pairwise amino acid interactions between the germ-line loops and A2 in the entire dataset, 44 are made using amino acids in the H_0/H_1 positions. Of these 44, 26 are made using the most frequently occurring residues. Of these 26, although some contacts are repeated within different complexes, there are none that were common to every complex (Table 1). In some instances, the most conserved residues at the H_0/H_1 sites do not contact A2 at all.

The only pattern observed in the contact analysis was between the conserved N-terminal tyrosine of CDR2β (Tyr48β) and Arg65 on the A2 α1-helix: these two residues contact each other in 6 of 13 complexes, although the numbers of contacts, their distances, and the relative positioning of the two side chains differ (Fig. 2A). This observation is consistent with earlier studies indicating a potential role for this tyrosine (22, 36). None of the interactions between Tyr48β and Arg65 involves a hydrogen bond or other electrostatic interaction, and the interaction is not restricted to a certain Vβ gene use. When in contact, the shape complementarity between Arg65 and Tyr48β ranged from moderate to high. This was quantified using the Sc statistic (37), yielding values from 0.64 for Mel5 to 0.86 for 1G4 (a value of 1.0 would indicate perfect complementarity between interacting residues).

A Correlation Between TCR Binding Position and Engagement of HLA-A2. The observation demonstrating a propensity for TCRs contacting Arg65 on A2 with a conserved tyrosine is notable, as position 65 is a polymorphic site that in part distinguishes HLA-A from HLA-B: position 65 is arginine in 85% of HLA-A alleles, whereas it is glutamine in 95% of HLA-B alleles and 100% of HLA-C alleles. To follow up, we examined the relationships between positions 48β and 65 in structures of TCRs bound to human class I MHC proteins other than A2, using the dataset in Table S2 (Complexes with other class I HLA proteins). Eleven of the TCRs in these 16 complexes have a tyrosine at position 48 in CDR2β, and only one of the MHC proteins has an arginine at position 65. Of the 11 non–A2-binding TCRs with a tyrosine at 48β, only one used it to contact position 65. This analysis revealed a stark difference between the A2 and the non-A2 complexes: in the non-A2 com-

Table 1. Contacts to HLA-A2 made by the most conserved residues in the germ-line loops of A2-binding human αβ TCRs

Loop	H_0/H_1 residue	HLA-A2 contact	TCR (PDB)
CDR1α	D26	W167	Mel5 (3HG1)
	D26	E58	DMF5 (3QDG)
	S27	W167	B7 (1BD2)
	S27	E166	ASO1 (3O4L)
	Y31	Q155	ASO1 (3O4L)
	Y31	L156	ASO1 (3O4L)
	Y31	Y159	ASO1 (3O4L)
	Y31	Q155	E16 (3UTT)
	Y31	Q155	B7 (1BD2)
CDR1β	N28	V76	1G4 (2BNR)
	H29	V76	DMF4[10] (3QDM)
	H29	V76	DMF4[9] (3QEQ)
CDR2β	Y48	R65	Mel5 (3HG1)
	Y48	R65	DMF5 (3QDG)
	Y48	R65	B7 (1BD2)
	Y48	R65	1G4 (2BNR)
	Y48	R65	E16 (3UTT)
	Y48	R65	DMF4[9] (3QEQ)
	Y48	Q72	RA14 (3GSN)
	V52	V56	1406 (4ZEZ)
	V52	Q72	JM22 (1OGA)
	V52	V76	JM22 (1OGA)
	D56	R65	1G4 (2BNR)
	D56	Q72	RA14 (3GSN)
	D56	R65	E16 (3UTT)
	D56	K68	DMF4[9] (3QEQ)

Contacts measured with a 4.5-Å interatomic cutoff as described in *Materials and Methods.*

plexes, position 48β is positioned across a broad stretch of the α1-helix (Fig. 2B). This differs from the complexes with A2, in which Tyr48β is restricted to a much tighter region, even for those cases in which Tyr48β and Arg65 are not in contact. The difference in positioning is not due to TCR crossing (docking) angles: the average and SDs of the crossing angles of the A2-binding TCRs is 46 ± 14°, whereas for the non–A2-binding TCRs it is 49 ± 13° (Table S2). The variation also cannot be attributed to different loop conformations, as the CDR2β loops for all of the TCRs in Fig. 2 adopt similar conformations (Fig. S1B).

However, in examining the structures, we observed a greater variation in the global positioning of the TCRs over the MHC in the non-A2 complexes than the A2 complexes. To quantify this, we computed the center of mass of the Vα/Vβ domains over the MHC. In the non-A2 complexes, the center of mass is broadly distributed over the pMHC (Fig. 2C). In contrast, the center of mass of the A2-binding TCRs is restricted to a tighter cluster. To quantify this difference, we computed the position of the average center of mass for the A2 and non-A2 datasets and tabulated its distance from each individual TCR's center of mass. For the A2 complexes, the average distance was 3.5 ± 1.3 Å. For the non-A2 complexes, the average distance was 6.3 ± 4.4 Å. A two-tailed Mann–Whitney test indicated these differences were significant at $P \leq 0.05$. [We note that the values for the non-A2 structures include the SB47, SB27, and CA5 TCRs, which engage a "superbulged" 13-mer presented by HLA-B*3508 (38). If these are excluded from the non-A2 data, the non-A2 average distance drops to a less significant 4.9 ± 2.5 Å, highlighting the need for more structural work on recognition of long peptides (39).]

The Interaction Between Tyr48β and Arg65 Does Not Drive TCR Binding to Peptide/HLA-A2 Complexes. To assess the importance of the interaction between Tyr48β and Arg65 in TCR recognition of peptide/A2 complexes, we examined the impact of mutations at these positions. These were made in the interface between the TCR DMF5 and the MART1$_{26(2L)-35}$ peptide presented by A2, a TCR–pMHC
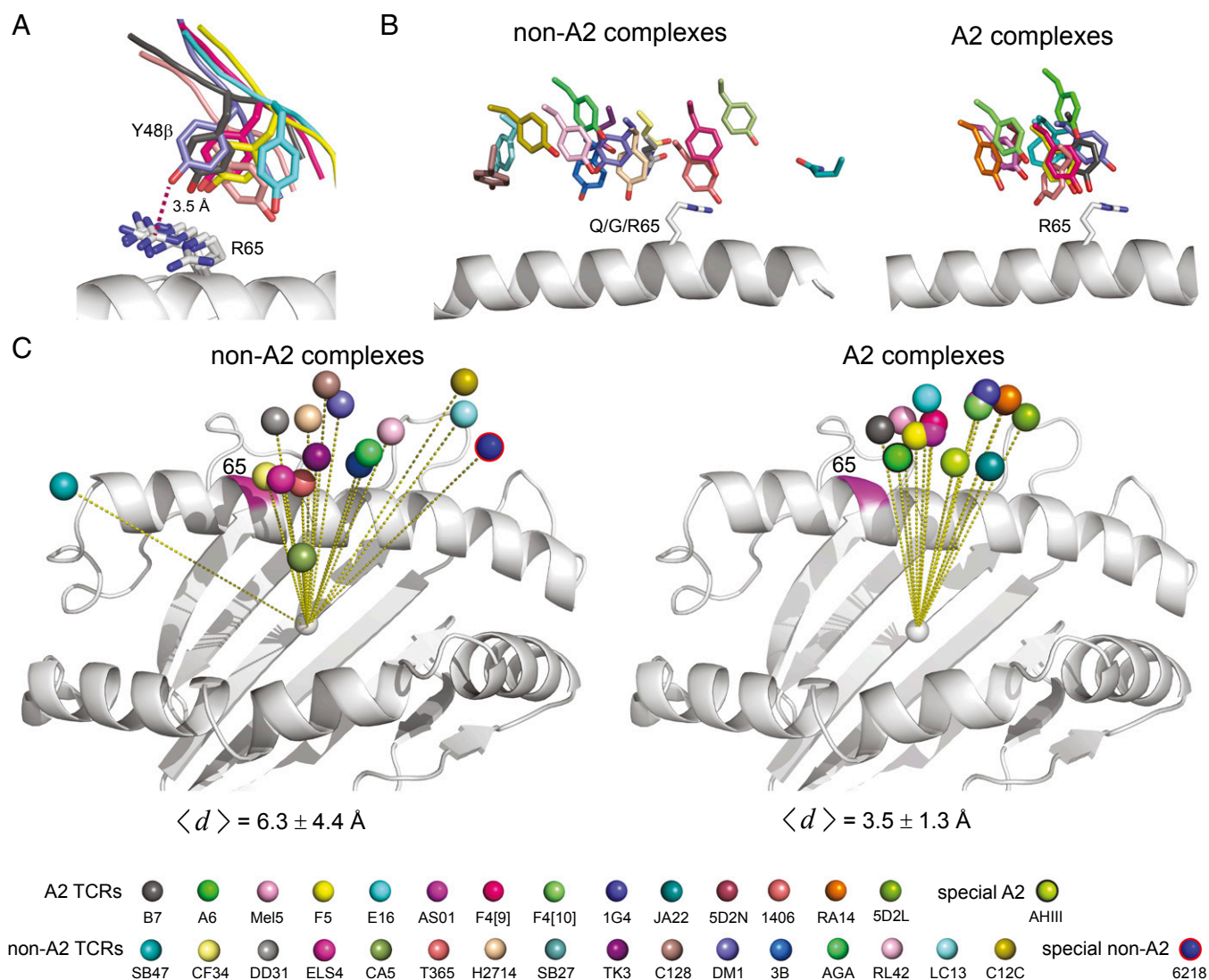
Fig. 2. TCRs that bind HLA-A2 are more restricted in their binding mode. (*A*) Alignment between Tyr48 of CDR2β and Arg65 on the A2 α1-helix in six TCR–pMHC complexes where Tyr48β and Arg65 are in contact. (*B*) TCRs that bind A2 place Tyr48β near position 65 on the α1 helix. The left image shows the position of residue 48β for TCRs that bind class I HLA proteins other than A2. The right image shows the Tyr48β position with A2-binding TCRs. (*C*) TCRs that bind A2 are more restricted in their position over the MHC, as shown by the TCR Vα/Vβ center of mass relative to the center of mass of the peptide-binding domain. The left image shows TCRs that bind class I HLA proteins other than A2; the right image shows A2 binding TCRs. The average distances between the Vα/Vβ centers of mass and the average center of mass are shown beneath each image. The color legend for TCRs is shown at the *Bottom* of the figure.

interaction we have previously studied in detail (13, 40). The shape complementarity between Tyr48β and Arg65 in the DMF5 interface is high, with an interresidue Sc statistic of 0.8. Mutation of Tyr48β to alanine reduced binding 40-fold ($\Delta\Delta G° = 2.1$ kcal/mol) (Fig. 3*A*). Mutation of Arg65 to alanine had a smaller but still significant effect, reducing binding eightfold ($\Delta\Delta G° = 1.2$ kcal/mol). Tyr48β and Arg65 are thus clearly important for DMF5 recognition. These results are consistent with alanine mutations at these positions in other TCR–pMHC interfaces (31, 40), as well as the impact of mutations to Tyr48β on T-cell development (36).

However, the two alanine measurements do not report on the strength of the Tyr48β–Arg65 interaction. To assess this, we measured the interaction between the two alanine mutants, which when combined with the measurements for the wild type and two single mutants allowed us to construct a double-mutant cycle targeting the Tyr48β–Arg65 interaction. In a double-mutant cycle, the interaction free energy between two side chains (ΔG_{int}) is equal to the difference between the $\Delta\Delta G°$ of the double mutant and the sum of the $\Delta\Delta G°$ values of the two single mutants (40, 41).

For the interaction between Tyr48 in DMF5 and Arg65 in A2, we measured a negligible ΔG_{int} of 0.3 kcal/mol (Fig. 3*A*). This indicates that, in the interface between DMF5 and A2, although both Tyr48β and Arg65 are important, the interaction between them does not provide a driving force for binding. This result is consistent with structure, as despite their high shape complementarity in the DMF5 complex, Tyr48β and Arg65 form only three van der Waals interactions. As Tyr48β and Arg65 interact similarly in all other TCR complexes with A2 (Fig. 2*A*; average number of contacts is 3 ± 1), this result is likely to be generally applicable.

TCRs That Bind HLA-A2 Form Electrostatic Interactions with Arg65. We next asked whether other conservations exist between TCRs and A2 that could account for the more restricted positioning of TCRs over A2. For example, residues on the α1- or α2-helices may interact with different residues in TCR loops that nonetheless provide similar chemical complementarity. Using the set of TCR–peptide/A2 structures, we examined pairwise TCR–A2 interactions from the perspective of the MHC protein, asking which TCR amino acids
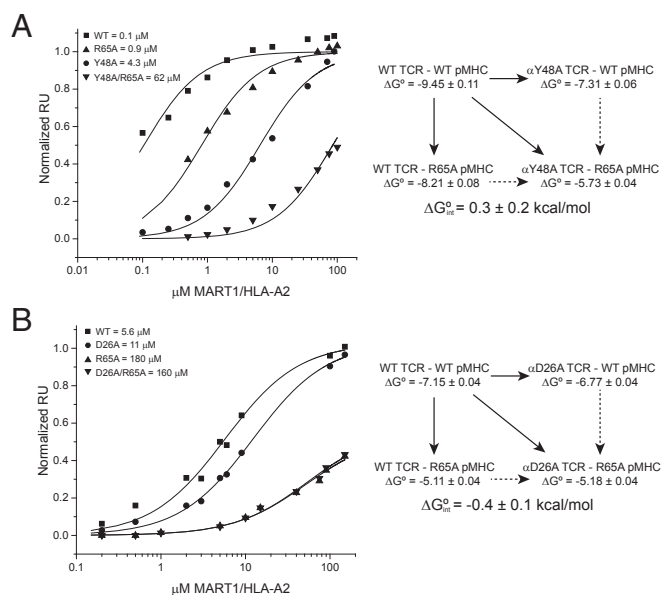
Fig. 3. Double-mutant cycles in the DMF5-MART1/HLA-A2 interface. (*A*) Double-mutant cycle between Arg65 and Tyr48β. Mutation of either Arg65 or Tyr48β to alanine weakens TCR binding, although a double-mutant cycle shows that the interaction between the two side chains is negligible. This experiment was done in the background of the affinity-enhancing D26αY/L98βW mutations in DMF5 (65). (*B*) Double-mutant cycle between Arg65 and Asp26α. Both mutations again have an impact on binding, although the double-mutant cycle analysis reveals a weak electrostatic interaction between the two side chains. This experiment was done in the background of the affinity-enhancing L98βW mutation in DMF5 (65).

interacted with each MHC amino acid on the α1- and α2-helices. We chose to focus on electrostatic interactions as these are more directional and of longer range than van der Waals interactions.

The strongest pattern that emerged from this analysis again involved Arg65 on the A2 α1-helix (Fig. 4). Arg65 forms a charge–charge interaction with a Glu or Asp near the end of CDR2β in 4 of the 13 complexes, as well as a fifth crystallized with an engineered TCR (42). Three other TCRs use a Glu or Asp at the start of CDR1α to interact with Arg65. Overall, 8 of 13 TCRs interact with Arg65 using a germ-line–encoded, negatively charged amino acid (Fig. S2 *A* and *B*).

Is this propensity for an electrostatic interaction with Arg65 reflected in germ-line loop sequences? An enrichment in negative charges in TCR CDR2β loops has been noted (22). Indeed, in the sequence logos for the A2-binding TCRs, position 56 in CDR2β and position 26 in CDR1α are low-diversity, H_1 sites for which Asp and Glu are the most common amino acids (Fig. 1*B*). These charges are also present in the non–A2-binding TCRs, although the frequencies are lower and the positions more diverse, with $H > 2$ (Fig. 1*C*). The pattern of interacting with Arg65 is not associated with variable gene use: only two of five Vβ 6-5 TCRs use an Asp in CDR2β to interact with Arg65, and two of three Vα 12-2 TCRs use an Asp in CDR1α (Table S2, Complexes with HLA-A2). Thus, the capacity to engage Arg65 electrostatically is represented in TCRs that bind A2 and, to a lesser extent, non-A2, and it is not restricted to nor is it obligated by particular gene segments.

Expanding the search to include hypervariable loops, we observed that, in addition to germ-line loops, some TCRs use hypervariable loops to interact with Arg65. Two use only an Asp or Glu in CDR3α, and two use only noncharged CDR3α loop residues to form hydrogen bonds with Arg65 (Fig. 4 and Fig. S2 *C* and *D*). Several receptors use both germ-line and hypervariable residues. Overall, 11 of 13 TCRs directly engage Arg65 using a germ-line CDR1α/CDR2β loop or a hypervariable

CDR3α loop. In the two complexes where Arg65 is not engaged directly, Arg65 interacts indirectly with a TCR polar or charged side chain via a bridging water or ion (Fig. S2*E*). Thus, in every structure examined, Arg65 is involved in a direct or indirect electrostatic interaction with the TCR.

Burial of Arg65 and the Adjacent Lys66 on the HLA-A2 α1-Helix Requires Overcoming a Large Desolvation Penalty. Position 65 is among the three most commonly contacted MHC residues in TCR interfaces with class I MHC proteins (18). This is attributable to its location near the center of the MHC α1-helix, which makes TCR burial of position 65 a near necessity (Fig. S1*C*). Indeed, for the A2- and non–A2-binding TCRs studied here, the average solvent accessibility of the position 65 side chain in the TCR–pMHC complex is 30%, compared with 82% in the free pMHC (Table S2, Complexes with HLA-A2). Notably, in A2, Arg65 is adjacent to another positively charged polymorphic amino acid, Lys66. In A2, the solvent accessibility of Lys66 is reduced upon TCR binding from an average of 23% to 6%. Only HLA-A alleles have the Arg65/Lys66 combination, with A2 and its suballeles accounting for 93% of the sequences. The combination of these two amino acids results in a significant and distinctive positive charge density on the surface of A2 that is absent from other MHC proteins. This is shown in Fig. 5*A*, which compares the electrostatic surface potentials of A2 (Arg65/Lys66) and HLA-B51 (Gln65/Ile66).

Burial of charged amino acids in a protein interface is always associated with a penalty resulting from the removal of water molecules interacting with the charge (26, 27). We computed this desolvation penalty for Arg65 and Lys66 in each TCR interface with A2, using a continuum electrostatics approach that yields the desolvation and coulombic energies associated with charge burial (27, 43). We used this approach previously in studying TCR recognition of pMHC, validating it with mutational and binding studies (32). Across the 13 TCR–A2 interfaces, the average desolvation penalty for Arg65 was an unfavorable +2.0 kcal/mol, ranging from +0.5 to +3.8 kcal/mol (Fig. 5*B*). The desolvation penalty was equally compensated by favorable coulombic interactions, which averaged −2.0 kcal/mol. We performed the same calculation for Lys66, which showed an average desolvation penalty of +1.7 kcal/mol (Fig. 5*B*).

These computed results are in good agreement with experimental measurements. For electrostatic interactions, the ΔG_{int} from a double-mutant cycle reflects the interresidue coulombic interactions (28). With the A6 TCR binding Tax/A2, we previously used double-mutant cycles to measure the strength of the interaction between Arg65 and Asp99 of CDR3α (40). The experimental ΔG_{int} was a highly favorable value of −2.8 kcal/mol. For comparison, in the A6-Tax/A2 interface, the continuum electrostatic calculations yielded a value of −3.2 kcal/mol for the intermolecular coulombic interactions with Arg65. As a second check, we performed a double-mutant cycle in the interface between Arg65 and Asp26 of CDR1α in the DMF5–MART1/A2 interface. Here, the ΔG_{int} was −0.4 kcal/mol (Fig. 3*B*). For comparison, the computed intermolecular coulombic interactions of Arg65 with DMF5 amounted to 0.5 kcal/mol.

The Xenoreactive TCR AHIII Demonstrates Electrostatic Interactions with Arg65 and Lys66 Are Not Strictly Encoded. Our analysis so far indicates that the need to interact with and counteract a region of positive charge on the A2 α1-helix is reflected in the composition of TCR germ-line sequences, and doing so is correlated with more restricted TCR positioning over pMHC. However, the utilization of both hypervariable and germ-line loops in this capacity, even with TCRs sharing variable gene segments, indicates that the germ-line sequences do not obligate particular interactions with Arg65. In other words, although TCR germ-line loops are enriched in the capacity to interact with Arg65/Lys66, inherent structural adaptability and sequence variations allow charge complementarity to occur opportunistically via a variety of means.

As a test of this hypothesis, we examined the structure of the xenoreactive mouse TCR AHIII bound to the p1049 peptide

HLA-A2 α1 helix

TCR	Glu58	Asp61	Arg65	Lys66	Lys68	Ala69	Gln72	Thr73	Arg75
B7	Q31α		E94α	D30α	D56β				
JM22			Q96α						
A6			D26α,D99α, T98α						
1G4			D56β,G98α				T72β,A52β	E30β	
RA14			R92α			N96α	Y48β,D56β, N96α		
Mel5			E59β,A94α						
ASO1			E58β						E50β
DMF5		K1α	D26α,F92α				N50β		
DMF4 [10]			T92α						D54β
DMF4 [9]			G93α		D56β	Y31β	D54β		D54β
E16			D56β,S95α	D94α					
C25			D95α				N50β,Q55β		E52β
1406			E28α,E101α	E28α					E56β
α1β1ILA1			D56β						
AHIII			E56β	S99α	E56β		D30β		

HLA-A2 α2 helix

TCR	Lys146	Ala149	Ala150	His151	Glu154	Gln155	Ala158	Thr163	Glu166
B7						S51α			
JM22	D30β		R98β	R98β		G94α, R98b			
A6					R102β	G101β		K68α	K68α
1G4			Q50α						
RA14	E30β	Y101β							K68α
Mel5						T100β			R27α
ASO1			N103β		Y48α	N103β		T30α	K68α
DMF5								K68α	K68α,N52α
DMF4 [10]	E95β					Y50α,Q100β			
DMF4 [9]						Q100β		N29α,R68α	R68α
E16									
C25	E103β		Y32α		T51α				
1406	D30β		Y98β		K68α			D30α	
α1β1ILA1	E30β		T98β		Q55α	Y48α			R68α
AHIII					K54α,K68α			S98α	

Fig. 4. Hydrogen bonds and charge interactions between αβ TCRs and the HLA-A2 α1/α2 helices in TCR–peptide/HLA-A2 structures. Red indicates charge–charge interactions, black indicates hydrogen bonds, and green indicates a water or ion bridge. Hydrogen bonds and charge interactions were identified as described in *Materials and Methods*. HLA-A2 amino acids not involved in electrostatic interactions are not shown. AHIII is a murine TCR bound to HLA-A2.

presented by A2 (44), reasoning that strict genetic programming between germ-line loops and A2 would have diverged between mice and humans, altering how Arg65 is accommodated. In this structure, however, Arg65 interacts with Glu56 in CDR2β (Fig. 4). Tyr48 of the AHIII CDR2β loop also contacts Arg65, and the TCR binds with a center of mass that falls within the range seen for other A2-binding TCRs (Fig. 2*C*). The "cognate" syngeneic ligand for AHIII is a peptide presented by the mouse class I MHC protein H-2D[b], which notably has glutamine rather than arginine at position 65 on the α1-helix. Although efforts to crystallize the AHIII-peptide/D[b] complex failed, in the complex between the mouse 6218 TCR and an unrelated viral peptide presented by D[b] (45), the 6218 TCR does not contact position 65, and the Vα/Vβ center of mass is skewed away from the cluster associated with TCR recognition of A2 (Fig. 2*C*). We conclude from this analysis that neutralization of the charges on the A2 α1-helix and the impact on TCR binding modes results from opportunistic rather than "hardwired" interactions, but the opportunistic interactions are facilitated by negative charges present in TCR germ-line loops.

Evolution Has Enhanced the Ability of TCR Germ-Line Loops to Engage HLA-A2. Although our analysis shows that TCRs are not hardwired to engage Arg65/Lys66 in predetermined ways, the ability to

counteract these charges is nonetheless reflected in germ-line loop sequences. Has evolution imparted this as a way to promote—if not compel—MHC binding? We addressed this by comparing the sequences of human TCR and antibody germ-line CDR loops, using data for all functional genes from the IMGT database. Antibody and TCR genes share a common ancestral history, diverging ~500 Mya (46). Unlike TCRs, antibodies are unrestricted in the ligands they bind, and we presumed that an imprint of evolution would be visible in the extent that TCR and antibody germ-line sequences differ, particularly in the CDR1/CDR2 germ-line loops that in TCRs complement positive charges on the A2 α1-helix.

We first looked at the overall frequency of amino acids that comprise all four TCR and antibody germ-line loops. As shown in Fig. 6*A*, the top five amino acids for antibodies are serine, glycine, tyrosine, threonine, and asparagine. The physicochemical basis for antibody enrichment in these amino acids is well understood (47, 48). The distribution in TCR germ-line loops resembles that in antibodies, but with some distinctions. The top five amino acids are serine, tyrosine, asparagine, glycine, and notably, glutamate. Indeed, TCR germ-line loops are twice as negatively charged than those of antibodies (14% Asp/Glu in TCRs, compared with 7% in antibodies). The negative charge difference between TCRs and antibodies is statistically significant as indicated by a χ^2 test (29).
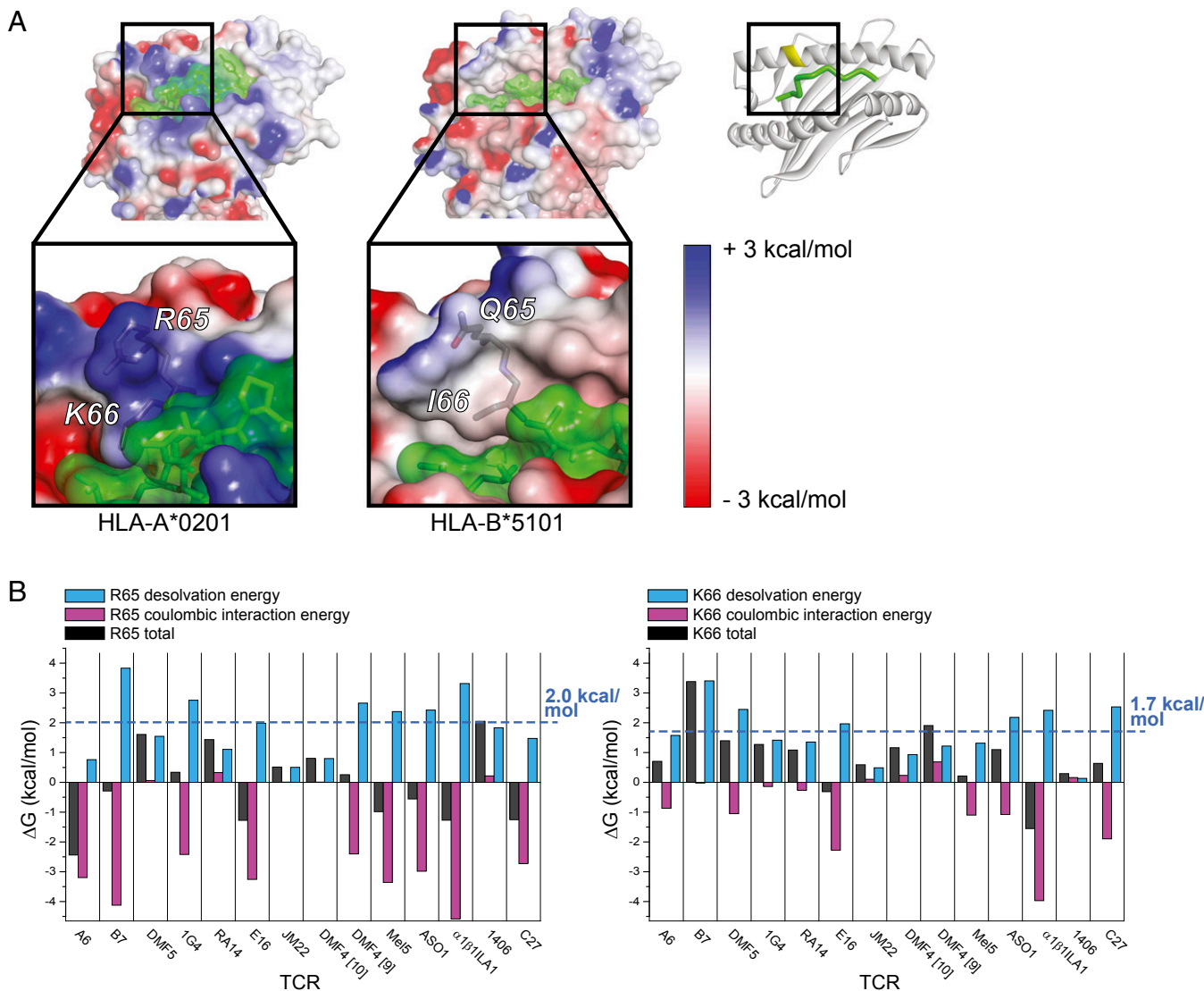
Fig. 5. TCR burial of Arg65 and Lys66 of HLA-A2 requires overcoming a large desolvation penalty. (*A*) Electrostatic surface potentials mapped to the surface of HLA-A2 (*Left*; PDB 1DUZ) and HLA-B*5101 (*Right*; PDB 4MJI). Surface potentials are colored from blue to red using the indicated scale. Peptide surface is green. The ribbon diagram on the *Upper Right* gives the orientation of the pMHC. The regions in the square highlight the surface potentials of positions 65/66. (*B*) Desolvation energies, coulombic interaction energies, and total interaction energies for burial of Arg65 (*Left*) or Lys66 (*Right*) in TCR complexes with A2. The average desolvation penalty is +2.0 kcal/mol for Arg65 and +1.7 kcal/mol for Lys66 as indicated by the blue lines. Coulombic interactions shown are the sum of the intermolecular and intramolecular terms.

Studies have shown that differences between TCR and antibody gene sequences are not due to genetic drift (34).

As we have found, when using germ-line loops, TCRs that bind A2 use negative charges in CDR2β and CDR1α, in that order of preference, to offset the charge on the A2 α1-helix. We thus compared TCR and antibody germ-line CDR loops. Previous antibody/TCR sequence comparisons indicate a VH = Vβ/VL = Vα equivalence (49). The antibody equivalents to CDR2β and CDR1α are accordingly CDR2-H and CDR1-L. The amino acid distributions for these TCR and antibody loops are noticeably different, particularly in their incorporation of negative charges (Fig. 6 *B* and *C*). Negative charges account for 19% and 8% of the composition of TCR CDR2β and CDR1α loops, respectively, compared with 8% and 5% in antibody CDR2-H and CDR1-L loops. These differences remain if the domain equivalencies are switched (i.e., VH = Vα and VL = Vβ), if alternative definitions are chosen for antibody CDR loops, or whether we use the classical IMGT definition of TCR CDR2β. The strategic enrichment in negative charges in TCR CDR loops compared with antibodies suggests evolution has imprinted A2 compatibility onto TCR germ-line loops.

Discussion

TCRs have long been hypothesized to be biased toward MHC proteins, but how such a bias can exist alongside TCR structural adaptability has remained a controversial puzzle. From a structural and energetic perspective, one challenge is that atomic contacts between germ-line regions of TCRs and MHC proteins are not conserved even when the same TCR variable regions and MHC proteins are compared (22). We expanded on this here by demonstrating that, in the structural database of TCRs that bind HLA-A2, not only are there no conserved TCR–MHC contacts, even those amino acids in TCR germ-line loops that are the least variable make different contacts or do not contact A2 at all. These results are consistent with findings that diversified TCRs can still bind MHC (35), and prompted us to search for a more fundamental basis for MHC bias based upon compatibility rather than simply contacts.
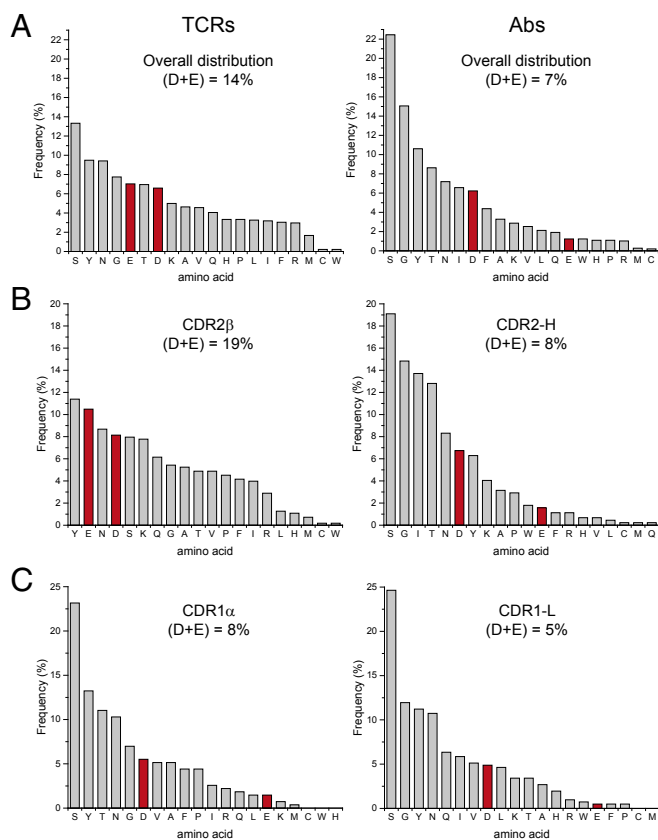
Fig. 6. TCR germ-line loops are enriched in negative charges compared with the corresponding antibody loops. (*A*) Distribution of all 20 aa in TCR (*Left*) and antibody (*Right*) germ-line loops. Total percent negative charges are shown. (*B*) Amino acid distributions in TCR CDR2β and antibody CDR2-H loops. (*C*) Amino acid distributions in TCR CDR1α and antibody CDR1-L loops.

Through a combination of sequence, structural, and biophysical analyses, we demonstrated that TCR germ-line loops are enriched in the means to complement a polymorphic hot-spot region on the α1-helix of the class I MHC protein HLA-A2, consisting of positively charged residues that are absent from most other HLA-A alleles, as well as those of HLA-B and HLA-C. Burial of these charges is required for TCRs to sample peptides effectively, and without the ability to complement them, their burial would be highly unfavorable due to the energetic cost of desolvation (26, 27).

Given the high frequency of A2 in human populations (30) and the recognized role of evolution in shaping TCR genes (34, 50), it is perhaps not surprising that evolution has enriched for charge complementarity with A2. Importantly, however, evolution has not selected for an exact structural solution in how to complement charges on A2. Negatively charged amino acids are found in appropriate locations in both the germ-line CDR1α and CDR2β loops, and structures show utilization of both, with some bias toward CDR2β. Moreover, hypervariable CDR3α loops can also perform this role. We suggest this reflects a built-in "permissiveness" in TCR binding. From the perspective of the TCR, the different solutions that are possible permit structural adjustments when optimizing binding with a particular peptide/A2 complex. From the perspective of the MHC, a large array of receptors will be available to engage A2, including those that lack the necessary structure and chemistry in the hypervariable CDR3α loop. Thus, evolution seems to not only have imparted MHC compatibility into the TCR repertoire, it has done so in a manner that allows structural adaptability and facilitates the cross-reactivity needed for immune function.

Does charge complementarity reflect an inherent MHC bias within TCRs? Interactions between charges in proteins are not

always favorable, and indeed, our results show a range of free energies for complementing the positively charged Arg65/Lys66 pair on A2. Thus, charge complementarity will not always drive binding. However, because uncompensated buried charges will always be energetically worse than even an unfavorable charge–charge interaction, charge complementarity will influence what structures are adopted, i.e., the structural specificity of a TCR–pMHC complex. This is apparent when the structures for TCRs in complex with A2 are compared with TCRs in complex with other human class I MHC proteins: although there is structural variation in how TCRs bind A2, it is less than seen with TCRs that bind other human class I MHCs. However, does this imply a bias? Consider the outcome if negative charges in CDR2β and CDR1α were absent from the TCR repertoire: due to increased reliance on CDR3α, fewer TCRs would be compatible with A2. The greater restriction on the sequence and structure of the CDR3α loop would in turn have a negative impact on the number and type of peptides that could be recognized when presented by A2. The presence of charge complementarity between TCR germ-line loops and the A2 α1-helix therefore increases the number of TCRs available to bind, translating into an intrinsic bias.

As noted above, the distinctive charged surface of A2 and its high prevalence in human populations make it an ideal candidate for identifying contributors to MHC bias. Other features likely exist in the TCR genes that promote compatibility with MHC proteins (our analysis of hydrogen bonds and salt bridges hints at this; for example, interactions with Lys146 and Glu166 on the A2 α2-helix may deserve closer scrutiny). As demonstrated here, however, these other features need not drive binding, and each one alone need not strictly specify a binding orientation. We suggest, however, that they do allow structural and chemical complementarity to occur when thermodynamic drivers exist elsewhere in the interface, and acting together they could strongly influence binding modes. An example may be the interaction of the strongly conserved Tyr48 in CDR2β with Arg65 on A2. This interaction is energetically neutral, and indeed is not present in all structures with A2. However, in cases where it is present, it may act as a "check" on the TCR binding mode. The removal of such checks or other sites encoding compatibility from the repertoire would be expected to impact TCR recognition, explaining the negative impact on T-cell development seen with mutation of Tyr48β in mice (36).

Because our findings suggest a bias in the true sense of the word, they are broadly compatible with other results. For example, they do not exclude the possibility of TCRs binding structurally/chemically compatible non-MHC ligands, or instances in which an evolved bias is "overruled" due to unique features in a ligand or TCR. They do not rule out a role for coreceptor or thymic selection in shaping the repertoire to help select for TCR–MHC compatibility. Following this logic, it remains possible that the more restricted positioning of TCRs over A2 that emerges from charge complementarity may be functionally significant.

Materials and Methods

TCR and Antibody Sequence Analysis. The sequences of all functional, naturally occurring, nonredundant human TCR and antibody genes were taken from the IMGT database (51). Germ-line CDR loops were selected based on the IMGT CDR definitions. For TCRs, TCR–pMHC crystallographic structures were used ensure CDR loop definitions adequately matched structural observations. For the TCR CDR2β loop, one N-terminal residue and three C-terminal residues were added to the loop definition to better fit structural observations as shown in Fig. S1A. Due to variations in the numbering of TCR residues, a uniform numbering was assigned to all HLA-A2–binding TCRs, using the B7 TCR [Protein Data Bank (PDB) ID code 1BD2] as a template (52). The numbering scheme also applies to CDR loops of the most common length. The CDR1α loop is comprised of residues 26–31, CDR2α loop is residues 49–55, CDR1β is residues 27–31, and CDR2β is residues 48–57. TCRs recognizing human class I MHC proteins were identified via PDB and the IMGT database. For sequence logos, unique sequences matching the most commonly used loop length were used as shown in Table S1. Sequence logos were created using the online WebLogo server at weblogo.berkeley.edu (53). Loop compositions in Fig. 6 used IMGT sequences for all functional TCR and antibody genes.

TCR–pMHC Structure Analysis. Structures of human αβ TCRs in complex with human class I MHC proteins were identified and downloaded from PDB. Contacts between TCR and MHC were identified using ContPro (54), the PyMOL script DistancesRH, and Accelrys Discovery Studio 4.1, using a 4.5-Å distance cutoff. We used this more liberal cutoff distance (compared with the more common 4 Å) to better account for coordinate error and the nonstatic nature of the protein interfaces (55, 56). TCR–pMHC complexes were superimposed by the PyMOL pair fit function, fitting the α carbon of residues 1–180 of the MHC protein. The two orientations of the DMF4 TCR over the nonamer and decamer MART-1 peptides bound to HLA-A2 were analyzed separately (referred to as DMF4[9] and DMF4[10]). Hydrogen bonds were identified using Discovery Studio, using a 3.4-Å cutoff and a DHA angle range of 90–180°. Charge–charge interactions were identified using Discovery Studio with a maximum distance of 5.6 Å. This distance is greater than the 4.0 Å sometimes used for formal salt bridges, but allows for a better accounting of long-distance interactions and helps account for coordinate error and protein dynamics. In the TCR contact analysis, the structure with the C7 TCR (53) was omitted as electron density for many of the key side chains was missing. Center of mass calculations were performed for the TCR variable domains and MHC peptide-binding domains using the PyMOL center of mass script. Shape complementarity values (Sc statistic) were computed with the Sc program (37). Solvent-accessible surface areas were calculated using Discovery Studio with a 1.4-Å radius probe. Percent accessible surface areas in Table S2 are relative to the surface area of an arginine or lysine in an extended Ala-X-Ala tripeptide. TCR crossing and incident angles were calculated as described (58, 59). Literature references for each TCR–pMHC complex analyzed are listed in Table S4.

Continuum Electrostatics Calculations. Desolvation and electrostatic interaction energies for Arg65 and Lys66 of HLA-A2 were performed using finite-difference Poisson–Boltzmann methods as previously described (32), with minor modifications as listed below. Methods of computing energies followed those of Sheinerman and Honig (43). Briefly, TCR structures were imported into Accelrys Discovery Studio. Hydrogen atoms were added and the structure's energy minimized to optimize hydrogen placement, using 1,000 steps of steepest descent followed by 300 steps of conjugate gradient minimization using the CHARMM22 force field. Electrostatic potentials were calculated using Delphi 4.11 (26) as implemented in Discovery Studio. Solvent dielectric was set to 80 and the interior dielectric set to 10. The value of 10 for an interior dielectric was chosen based on our previous work (32) and is consistent with values used to reproduce experimental observations in other systems (60). We note, however, that although absolute values will change, our overall conclusions will be unaffected by the use of other interior dielectrics (see ref. 32, table 4). Ionic strength was set at 0.15 M. Grid spacing was 0.5 Å per point. Focusing was used, moving from 50% to 90% grid occupancy in two steps. Calculations performed on TCR–pMHC complexes and free pMHC molecules used the same grid mapping.

Desolvation penalties, intermolecular coulombic interactions, and intramolecular coulombic interactions were calculated from electrostatic potentials computed with partial charges present only on the atoms of Arg65 or Lys66 (Fig.

5B). Desolvation penalties were computed as the difference between the charging free energies of the TCR–pMHC complexes and the unbound pMHC. Intermolecular coulombic interaction energies were calculated as the interaction of the TCR–pMHC potentials with the full set of TCR partial charges. Intramolecular coulombic interaction energies were calculated similarly, except the calculation involved the full set of pMHC partial charges except those for Arg65 or Lys66, and is reported as the difference between calculations with pMHC potentials calculated for the free and TCR-bound states (43). Coulombic interaction energies reported in Fig. 5B are the sum of the intermolecular and intramolecular terms. Energies of kT reported by Delphi were converted to free energy values at 298 K. Structures for the pMHC molecules in Fig. 5A were of the Tax peptide bound to HLA-A2 (61) and an HIV 8-mer bound to HLA-B*5101 (62), and electrostatic surface potentials in this figure were calculated using all formal charges.

Protein Expression and Double-Mutant Cycles. Soluble constructs for the A6 and DMF5 TCRs and HLA-A2 were generated from bacterially expressed inclusion bodies as previously described (63). TCRs used an engineered disulfide bridge connecting the constant domains to improve stability (64). DMF5 mutants were produced by site-directed mutagenesis and confirmed by sequencing. Peptides were purchased from Chi Scientific. Steady-state binding was measured with a Biacore 3000 instrument as previously described (63). Experiments were performed at 25 °C in 10 mM Hepes, 3 mM EDTA, 150 mM NaCl, and 0.005% surfactant P20 at pH 7.4. Refolded and purified TCR were coupled to the surface of a CM5 sensor chip via amine linkage to a density of ~1,500 RU.

Double-mutant cycles were performed and analyzed as described previously (40). Briefly, soluble wild-type followed by mutant pMHC complexes were injected over wild-type or mutant TCR surfaces in a series of increasing concentration points until steady-state binding was attained. Each injection series was repeated twice. In any one experiment, three flow cells were used: a blank, wild-type TCR, and mutant TCR. The output of one single experiment thus consisted of two sets of all four measurements that comprise a double-mutant cycle as shown in Fig. 3. Data were processed in BiaEvaluation 4.1, and all eight datasets were simultaneously analyzed in OriginPro 7 with a custom global fitting function. The wild-type and mutant TCR surface densities and the three $\Delta\Delta G$ values that make up the cycle were global fitting parameters. Errors were propagated using standard statistical error propagation methods (63). High-affinity variants of the DMF5 TCR were used to facilitate the analysis (65). For analysis of Tyr48β–Arg65 pair, the D26αY/L98βW double mutant was used. For analysis of D26α–Arg65 pair, the L98βW single mutant was used. We note that, excluding the sites of the two mutations, the structures of the high-affinity and wild-type DMF5 TCR complexes are essentially identical (13, 65). The utilization of affinity-enhancing mutations is unlikely to impact the double-mutant cycle experiments, owing to the structural similarities and the robust nature of double-mutant cycle analysis (40, 41).

1. Zinkernagel RM, Doherty PC (1974) Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytic choriomeningitis. *Nature* 251(5475):547–548.
2. Garboczi DN, et al. (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384(6605):134–141.
3. Garcia KC, et al. (1996) An alphabeta T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex. *Science* 274(5285):209–219.
4. Zerrahn J, Held W, Raulet DH (1997) The MHC reactivity of the T cell repertoire prior to positive and negative selection. *Cell* 88(5):627–636.
5. Garcia KC, Adams JJ, Feng D, Ely LK (2009) The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nat Immunol* 10(2):143–147.
6. Van Laethem F, et al. (2007) Deletion of CD4 and CD8 coreceptors permits generation of alphabetaT cells that recognize antigens independently of the MHC. *Immunity* 27(5):735–750.
7. Tikhonova AN, et al. (2012) αβ T cell receptors that do not undergo major histocompatibility complex-specific thymic selection possess antibody-like recognition specificities. *Immunity* 36(1):79–91.
8. Mazza C, Malissen B (2007) What guides MHC-restricted TCR recognition? *Semin Immunol* 19(4):225–235.
9. Rangarajan S, Mariuzza RA (2014) T cell receptor bias for MHC: Co-evolution or co-receptors? *Cell Mol Life Sci* 71(16):3059–3068.
10. Van Laethem F, Tikhonova AN, Singer A (2012) MHC restriction is imposed on a diverse T cell receptor repertoire by CD4 and CD8 co-receptors during thymic selection. *Trends Immunol* 33(9):437–441.
11. Garcia KC (2012) Reconciling views on T cell receptor germline bias for MHC. *Trends Immunol* 33(9):429–436.
12. Miles JJ, McCluskey J, Rossjohn J, Gras S (2015) Understanding the complexity and malleability of T-cell recognition. *Immunol Cell Biol* 93(5):433–441.
13. Borbulevych OY, Santhanagopolan SM, Hossain M, Baker BM (2011) TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. *J Immunol* 187(5):2453–2463.
14. Adams JJ, et al. (2011) T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* 35(5):681–693.
15. Stadinski BD, et al. (2011) A role for differential variable gene pairing in creating T cell receptors specific for unique major histocompatibility ligands. *Immunity* 35(5): 694–704.
16. Stadinski BD, et al. (2014) Effect of CDR3 sequences and distal V gene residues in regulating TCR-MHC contacts and ligand specificity. *J Immunol* 192(12): 6071–6082.
17. Deng L, Langley RJ, Wang Q, Topalian SL, Mariuzza RA (2012) 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* 109(37):14960–14965.
18. Burrows SR, et al. (2010) Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. *Proc Natl Acad Sci USA* 107(23):10608–10613.
19. Dai S, et al. (2008) Crossreactive T cells spotlight the germline rules for alphabeta T cell-receptor interactions with MHC molecules. *Immunity* 28(3):324–334.
20. Reinherz EL, et al. (1999) The crystal structure of a T cell receptor in complex with peptide and MHC class II. *Science* 286(5446):1913–1921.
21. Feng D, Bond CJ, Ely LK, Maynard J, Garcia KC (2007) Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction "codon." *Nat Immunol* 8(9):975–983.

22. Marrack P, Scott-Browne JP, Dai S, Gapin L, Kappler JW (2008) Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annu Rev Immunol* 26:171–203.

23. Magarian-Blander J, Ciborowski P, Hsia S, Watkins SC, Finn OJ (1998) Intercellular and intracellular events following the MHC-unrestricted TCR recognition of a tumor-specific peptide epitope on the epithelial antigen MUC1. *J Immunol* 160(7):3111–3120.

24. Castro CD, Luoma AM, Adams EJ (2015) Coevolution of T-cell receptors with MHC and non-MHC ligands. *Immunol Rev* 267(1):30–55.

25. Beringer DX, et al. (2015) T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nat Immunol* 16(11):1153–1161.

26. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. *Science* 268(5214):1144–1149.

27. Hendsch ZS, Tidor B (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci* 3(2):211–226.

28. Bosshard HR, Marti DN, Jelesarov I (2004) Protein stabilization by salt bridges: Concepts, experimental approaches and clarification of some misunderstandings. *J Mol Recognit* 17(1):1–16.

29. Collis AVJ, Brouwer AP, Martin ACR (2003) Analysis of the antigen combining site: Correlations between length and sequence composition of the hypervariable loops and the nature of the antigen. *J Mol Biol* 325(2):337–354.

30. Chen KY, Liu J, Ren EC (2012) Structural and functional distinctiveness of HLA-A2 allelic variants. *Immunol Res* 53(1-3):182–190.

31. Baker BM, Turner RV, Gagnon SJ, Biddison WE (2001) Identification of a crucial energetic footprint on the α1 helix of human histocompatibility leukocyte antigen (HLA)-A2 that provides functional interactions for recognition by tax peptide/HLA-A2-specific T cell receptors. *J Exp Med* 193(5):551–562.

32. Gagnon SJ, et al. (2005) Unraveling a hotspot for TCR recognition on HLA-A2: Evidence against the existence of peptide-independent TCR binding determinants. *J Mol Biol* 353(3):556–573.

33. Baxter TK, et al. (2004) Strategic mutations in the class I major histocompatibility complex HLA-A2 independently affect both peptide binding and T cell receptor recognition. *J Biol Chem* 279(28):29175–29184.

34. Stewart JJ, et al. (1997) A Shannon entropy analysis of immunoglobulin and T cell receptor. *Mol Immunol* 34(15):1067–1082.

35. Holland SJ, et al. (2012) The T-cell receptor is not hardwired to engage MHC ligands. *Proc Natl Acad Sci USA* 109(45):E3111–E3118.

36. Scott-Browne JP, White J, Kappler JW, Gapin L, Marrack P (2009) Germline-encoded amino acids in the alphabeta T-cell receptor control thymic selection. *Nature* 458(7241):1043–1046.

37. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234(4):946–950.

38. Liu YC, et al. (2013) Highly divergent T-cell receptor binding modes underlie specific recognition of a bulged viral peptide bound to a human leukocyte antigen class I molecule. *J Biol Chem* 288(22):15442–15454.

39. Hassan C, et al. (2015) Naturally processed non-canonical HLA-A*02:01 presented peptides. *J Biol Chem* 290(5):2593–2603.

40. Piepenbrink KH, Blevins SJ, Scott DR, Baker BM (2013) The basis for limited specificity and MHC restriction in a T cell receptor interface. *Nat Commun* 4:1948.

41. Horovitz A (1996) Double-mutant cycles: A powerful tool for analyzing protein structure and function. *Fold Des* 1(6):R121–R126.

42. Cole DK, et al. (2014) T-cell receptor (TCR)-peptide specificity overrides affinity-enhancing TCR-major histocompatibility complex interactions. *J Biol Chem* 289(2):628–638.

43. Sheinerman FB, Honig B (2002) On the role of electrostatic interactions in the design of protein-protein interfaces. *J Mol Biol* 318(1):161–177.

44. Buslepp J, Wang H, Biddison WE, Appella E, Collins EJ (2003) A correlation between TCR Valpha docking on MHC and CD8 dependence: Implications for T cell selection. *Immunity* 19(4):595–606.

45. Day EB, et al. (2011) Structural basis for enabling T-cell receptor diversity within biased virus-specific CD8+ T-cell responses. *Proc Natl Acad Sci USA* 108(23):9536–9541.

46. Flajnik MF, Kasahara M (2010) Origin and evolution of the adaptive immune system: Genetic events and selective pressures. *Nat Rev Genet* 11(1):47–59.

47. Birtalan S, et al. (2008) The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. *J Mol Biol* 377(5):1518–1528.

48. Shiroishi M, et al. (2007) Structural consequences of mutations in interfacial Tyr residues of a protein antigen-antibody complex. The case of HyHEL-10-HEL. *J Biol Chem* 282(9):6783–6791.

49. Dunbar J, Knapp B, Fuchs A, Shi J, Deane CM (2014) Examining variable domain orientations in antigen receptors gives insight into TCR-like antibody design. *PLoS Comput Biol* 10(9):e1003852.

50. Bontrop RE, Otting N, Slierendregt BL, Lanchbury JS (1995) Evolution of major histocompatibility complex polymorphisms and T-cell receptor diversity in primates. *Immunol Rev* 143(1):33–62.

51. Robinson J, et al. (2011) The IMGT/HLA database. *Nucleic Acids Res* 39(Database issue):D1171–D1176.

52. Ding YH, et al. (1998) Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/Tax peptide complex using different TCR amino acids. *Immunity* 8(4):403–411.

53. Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14(6):1188–1190.

54. Firoz A, Malik A, Afzal O, Jha V (2010) ContPro: A Web tool for calculating amino acid contact distances in protein from 3D-structures at different distance threshold. *Bioinformation* 5(2):55–57.

55. Kass I, Buckle AM, Borg NA (2014) Understanding the structural dynamics of TCR-pMHC complex interactions. *Trends Immunol* 35(12):604–612.

56. Hawse WF, et al. (2014) TCR scanning of peptide/MHC through complementary matching of receptor and ligand molecular flexibility. *J Immunol* 192(6):2885–2891.

57. Yang X, et al. (2015) Structural basis for clonal diversity of the public T cell response to a dominant human cytomegalovirus epitope. *J Biol Chem* 290(48):29106–29119.

58. Rudolph MG, Stanfield RL, Wilson IA (2006) How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* 24:419–466.

59. Pierce BG, Weng Z (2013) A flexible docking approach for prediction of T cell receptor-peptide-MHC complexes. *Protein Sci* 22(1):35–46.

60. Fitch CA, et al. (2002) Experimental pK(a) values of buried residues: Analysis with continuum methods and role of water penetration. *Biophys J* 82(6):3289–3304.

61. Khan AR, Baker BM, Ghosh P, Biddison WE, Wiley DC (2000) The structure and stability of an HLA-A*0201/octameric tax peptide complex with an empty conserved peptide-N-terminal binding site. *J Immunol* 164(12):6398–6405.

62. Motozono C, et al. (2014) Molecular basis of a dominant T cell response to an HIV reverse transcriptase 8-mer epitope presented by the protective allele HLA-B*51:01. *J Immunol* 192(7):3428–3434.

63. Davis-Harrison RL, Armstrong KM, Baker BM (2005) Two different T cell receptors use different thermodynamic strategies to recognize the same peptide/MHC ligand. *J Mol Biol* 346(2):533–550.

64. Boulter JM, et al. (2003) Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Eng* 16(9):707–711.

65. Pierce BG, et al. (2014) Computational design of the affinity and specificity of a therapeutic T cell receptor. *PLoS Comput Biol* 10(2):e1003478.

66. Ding YH, Baker BM, Garboczi DN, Biddison WE, Wiley DC (1999) Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. *Immunity* 11(1):45–56.

67. Chen J-L, et al. (2005) Structural and kinetic basis for heightened immunogenicity of T cell vaccines. *J Exp Med* 201(8):1243–1255.

68. Gras S, et al. (2009) Structural bases for the affinity-driven selection of a public TCR against a dominant human cytomegalovirus epitope. *J Immunol* 183(1):430–437.

69. Bulek AM, et al. (2012) Structural basis for the killing of human beta cells by CD8+ T cells in type 1 diabetes. *Nat Immunol* 13(3):283–289.

70. Stewart-Jones GB, McMichael AJ, Bell JI, Stuart DI, Jones EY (2003) A structural basis for immunodominant human T cell receptor recognition. *Nat Immunol* 4(7):657–663.

71. Cole DK, et al. (2009) Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J Biol Chem* 284(40):27281–27289.

72. Miles JJ, et al. (2010) Genetic and structural basis for selection of a ubiquitous T cell receptor deployed in Epstein-Barr virus infection. *PLoS Pathog* 6(11):e1001198.

73. Liu YC, et al. (2014) A molecular basis for the interplay between T cells, viral mutants, and human leukocyte antigen micropolymorphism. *J Biol Chem* 289(24):16688–16698.

74. Nivarthi UK, et al. (2014) An extensive antigenic footprint underpins immunodominant TCR adaptability against a hypervariable viral determinant. *J Immunol* 193(11):5402–5413.

75. Ladell K, et al. (2013) A molecular basis for the control of preimmune escape variants by HIV-specific CD8+ T cells. *Immunity* 38(3):425–436.

76. Shimizu A, et al. (2013) Structure of TCR and antigen complexes at an immunodominant CTL epitope in HIV-1 infection. *Sci Rep* 3:3097.

77. Gras S, et al. (2012) A structural basis for varied αβ TCR usage against an immunodominant EBV antigen restricted to a HLA-B8 molecule. *J Immunol* 188(1):311–321.

78. Macdonald WA, et al. (2009) T cell allorecognition via molecular mimicry. *Immunity* 31(6):897–908.

79. Gras S, et al. (2009) The shaping of T cell receptor recognition by self-tolerance. *Immunity* 30(2):193–203.

80. Stewart-Jones GB, et al. (2012) Structural features underlying T-cell receptor sensitivity to concealed MHC class I micropolymorphisms. *Proc Natl Acad Sci USA* 109(50):E3483–E3492.

81. Archbold JK, et al. (2009) Natural micropolymorphism in human leukocyte antigens provides a basis for genetic control of antigen recognition. *J Exp Med* 206(1):209–219.

82. Tynan FE, et al. (2007) A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat Immunol* 8(3):268–276.

83. Tynan FE, et al. (2005) T cell receptor recognition of a "super-bulged" major histocompatibility complex class I-bound peptide. *Nat Immunol* 6(11):1114–1122.