

The Crystal Structure of a T Cell Receptor in Complex with Peptide and MHC Class II

Ellis L. Reinherz,^{1,2*} Kemin Tan,^{1,2†} Lei Tang,^{1,2†} Petra Kern,^{1,2‡}
 Jin-huan Liu,^{1,2} Yi Xiong,^{1,2} Rebecca E. Hussey,^{1,3}
 Alex Smolyar,^{1,2} Brian Hare,⁵ Rongguang Zhang,⁶
 Andrzej Joachimiak,⁶ Hsiu-Ching Chang,^{1,2} Gerhard Wagner,⁵
 Jia-huai Wang^{1,4*}

The crystal structure of a complex involving the D10 T cell receptor (TCR), 16-residue foreign peptide antigen, and the I-A^k self major histocompatibility complex (MHC) class II molecule is reported at 3.2 angstrom resolution. The D10 TCR is oriented in an orthogonal mode relative to its peptide-MHC (pMHC) ligand, necessitated by the amino-terminal extension of peptide residues projecting from the MHC class II antigen-binding groove as part of a mini β sheet. Consequently, the disposition of D10 complementarity-determining region loops is altered relative to that of most pMHC-specific TCRs; the latter TCRs assume a diagonal orientation, although with substantial variability. Peptide recognition, which involves P-1 to P8 residues, is dominated by the V α domain, which also binds to the class II MHC β_1 helix. That docking is limited to one segment of MHC-bound peptide offers an explanation for epitope recognition and altered peptide ligand effects, suggests a structural basis for alloreactivity, and illustrates how bacterial superantigens can span the TCR-pMHCII surface.

The adaptive immune response is dependent on the specific recognition function of $\alpha\beta$ T lymphocytes (1). Each T cell detects a protein fragment (that is, peptide) of a self protein or cell-associated pathogen derived from either viral, bacterial, fungal, parasitic, or tumor cell origin bound to an MHC molecule. The physical binding of the peptide-MHC (pMHC) complex to the TCR then initiates a series of signal transduction events. Once triggered, T lymphocytes release cytotoxic molecules or inflammatory cytokines (or both) that destroy the infected or otherwise altered cells through various effector mechanisms. For a given $\alpha\beta$ T lymphocyte, immune recognition is mediated by a clonotypic $\alpha\beta$ heterodimeric structure (Ti) noncovalently associated with the monomorphic CD3 signaling components.

Sequence analysis of TCR $\alpha\beta$ heterodimers first suggested that they would share with antibodies a common structure (2, 3). However, direct evidence supporting this

notion has been provided only in the last several years, initially from crystal structures of $\alpha\beta$ TCR components (4) and subsequently through analysis of intact $\alpha\beta$ TCR heterodimers alone (5) or in complex with pMHC (6, 7). As anticipated, aside from the C α domain, the three-dimensional (3D) structure of the TCR resembles an antibody Fab fragment such that each of the α and β chains consists of canonical immunoglobulin (Ig)-like variable and constant domains, with the hypervariable complementarity-determining regions (CDRs) from the two variable domains (V α and V β) forming the ligand binding site for pMHC within the immunorecognition module.

Class I and class II MHC molecules have evolved to facilitate T cell detection of pathogens residing in distinct intracellular compartments (8–10). Although the domain organization of the class I and class II MHC extracellular segments is different, these molecules possess a very similar overall antigen-presenting groove consisting of $\alpha 1$ plus $\alpha 2$ domains and $\alpha 1$ plus $\beta 1$ domains for class I and class II MHC, respectively (11–13). For both molecules, the α helices of these two domains form the sides of the antigen-binding groove with the floor created by an eight-stranded β sheet arising from both domains. However, unique structural features of the two MHC classes dictate the binding of peptides differing in length and composition (14). The bipartite nature of the immune recognition molecules expressed on antigen-

presenting cells is reflected at the level of $\alpha\beta$ T lymphocytes by the evolution of two subsets bearing specialized MHC binding structures, termed CD4 and CD8 (15, 16). CD8 cells are cytolytic precursor and effector cells, whereas CD4 cells comprise the helper T cell subset that initiates inflammatory responses. CD4 and CD8 molecules have been termed coreceptors because CD4 binds to the membrane-proximal $\beta 2$ domain of class II MHC whereas CD8 ($\alpha\alpha$ and $\alpha\beta$) isoforms bind to the corresponding $\alpha 3$ domain of class I MHC (17, 18).

At present, four distinct class I-restricted TCRs have been crystallized in complex with their specific pMHCII ligands (6, 7). Rather extensive interactions with the pMHC α helices have suggested a common “diagonal” docking mode, regardless of TCR specificity or species origin, in which the TCR V α domain overlies the class I MHC $\alpha 2$ helix and the V β domain overlies the MHC $\alpha 1$ helix. As a result, the CDR1 and CDR3 loops of the TCR V α and V β domains make the major contacts with the peptide, whereas the two CDR2 loops interact primarily with the MHC. Given the distinct nature of class II versus class I MHC expression, peptide binding, and the differential interactions with CD4 and CD8 T cell subsets, it was of interest to structurally define the TCR-pMHCII interaction. We now report the x-ray crystal structure of a TCR-pMHCII ternary complex. The complex contains the V module of the D10 TCR [single chain (sc) D10] derived from AKR/J (H-2^k) mouse T cell clone D10.G4 and a fragment of conalbumin (CA) bound to the self-I-A^k molecule (19, 20). A marked difference in TCR docking topology relative to TCR-pMHC I complex is noted.

Overview of the Complex Structure

The crystal structure of the scD10-CA/I-A^k complex was determined with molecular replacement and alternative cycles of model building and refinement (Table 1). In the asymmetric unit there are two complexes related to each other by a 115° rotation. The complex A and D10-B pack together to form layers perpendicular to the longest Y axis, whereas the I-A^k molecules B connect the layers, thus leaving spaces filled with large amounts of solvent. The structures of the two complexes are very similar. The root-mean-square deviation (rmsd) value of C α superposition is only 0.8 Å for the whole complex. Consequently, only complex A is discussed.

Figure 1 is a ribbon drawing of the scD10-CA/I-A^k complex. Figure 2 is an omit map around the CA peptide. In Fig. 1, the view is down the peptide-binding groove of the MHC molecule, hence the longer dimension of the MHC molecule with α and β chains on either

¹Laboratory of Immunobiology, Dana-Farber Cancer Institute, ²Department of Medicine, ³Department of Pathology, ⁴Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. ⁵Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA. ⁶Biosciences Division, Argonne National Laboratory, Argonne, IL 60439, USA.

*To whom correspondence should be addressed. E-mail: ellis_reinherz@dfci.harvard.edu; jwang@red.dfci.harvard.edu

†These authors contributed equally to this work.

‡Present address: N.V. Procter Gamble Eurocor, Temselaan 100, A128, B-1853 Strombeek-Bever, Belgium.

RESEARCH ARTICLES

side. The immediately striking observation is that the scD10 molecule sits on top of the MHC with its longer dimension crossing the bound peptide in an orthogonal manner, rather than the “diagonal” mode commonly recognized in structures of TCR-pMHCI complexes (6, 7). The V α domain of scD10 contacts the β 1 helical region of I-A^k, whereas the V β domain touches the α 1 helical region. Contacts between the D10 TCR and the CA/I-A^k pMHC ligand are listed in the supplementary table (21). In contrast to the class I pMHC-TCR ternary structures, the much

longer peptide stretches out both sides of the TCR-MHC complex. In particular, the COOH-terminal three residues have no interaction with either TCR or MHC. The orthogonal orientation for the TCR-pMHCI interaction noted herein excludes the possibility that direct TCR contact with COOH-terminal peptide flanking residues is the basis for any observed functional dependence on this peptide segment in T cell recognition (22).

Although there have been several structures of class I-restricted $\alpha\beta$ TCRs or derivative fragments (4–7), our scD10 represents

the crystal structure of a class II-restricted $\alpha\beta$ TCR V module in complex with its cognate pMHC partner. The structure of the V α -V β heterodimer is very similar to the recently published nuclear magnetic resonance (NMR) structure of an unligated scD10 (23). The rmsd's for all of the backbone atoms of residues in β strands between structures in the NMR ensemble and the crystal structure are 1.3 and 1.4 Å for V α and V β domains, respectively. Notably, there does not appear to be any significant 3D structural difference between TCRs that recognize peptides bound to class I versus class II MHC molecules. The human class I HLA-A2/Tax-specific B7 TCR is by far the most structurally similar to our murine class II-specific scD10. Virtually the entire V module of these two TCRs can be superimposed. The rmsd values of the superposition for the entire V α domain's 110 C α atoms (excluding the first residue, which is not seen in the density map of our scD10 structure) and 107 C α atoms of the V β domain (excluding part of the CDR3) are only 0.98 and 0.72 Å, respectively. Moreover, if

Table 1. Crystallographic analysis. Crystals of the ternary complex were grown by the conventional hanging-droplet vapor diffusion method at room temperature. The *Escherichia coli*-expressed scD10 (57) and glycosylated CA/I-A^k from CHO Lec3.2.8.1 cells (58) were mixed at a 1:1 molar ratio to a final concentration of 23 mg/ml in 0.1 M tris-HCl buffer (pH 8.5). The protein solution was further mixed with a crystallization buffer of 8% PEG 8000, 0.1 M tris (pH 8.5), 0.01 M KCl and then sealed against a reservoir with the same buffer. These crystals belong to the space group *P*2₁2₁2 with unit cell parameters *a* = 97.6 Å, *b* = 345.3 Å, and *c* = 97.7 Å. There are two complexes in the asymmetric unit with 78% solvent. Crystals were stepwise transferred to cryoprotectant solution that contains 30% glycerol in addition to the crystallization buffer before freezing. One data set was collected at the SBC-CAT of Advanced Photon Source (APS) at the Argonne National Laboratory with an APS1 mosaic 3 × 3 CCD detector at 100 K. The wavelength used was 1.069 Å. Data were processed with programs DENZO and SCALEPACK (59). The structure was solved by molecular replacement with AMoRe (60). The refined structure of CA/I-A^k (61) was taken as the search model. At the beginning, only one of the CA/I-A^k pMHC molecules (molecule A) was identified. The CA/I-A^k molecule B was located only after the first one was rigid body refined and fixed. We then carried out the rigid body refinement of the two I-A^k molecules, each of the domains and the bound peptide being treated as one rigid body. A few degrees of rotations were seen for the α 2 and β 2 domains. After positional and individual *B*-factor refinement, the *R*_{free} dropped, and the Ig-like domains of scD10, especially the one in complex A, were already visible in the calculated 2*F*_o - *F*_c difference map. Cycles of model building and refinement gradually improved the density, allowing the correct side-chain assignment and eventually the completion of the model building and refinement. All the refinement was done with the program X-PLOR (62) and model building with program O (55). Ten percent of reflections were set aside for *R*_{free} calculation. Noncrystallographic symmetry (NCS) refinement was not used because the two complexes are related by an improper rotation and the two I-A^k molecules have very different temperature factors. Using NCS at an early stage of the refinement led to serious intermolecular collision. In the current model, each of the complexes contains residues 1 to 182 and 2 to 190 of I-A^k α and β chains, respectively, all 16 residues of the bound peptide (three leader-derived and 13 CA-derived), as well as residues 2 to 117 and 3 to 116A of D10 V α and V β domains, respectively. Ten carbohydrate moieties were modeled in three potential glycosylation sites in I-A^k molecules. At this resolution no water molecules were included. The final 2*F*_o - *F*_c map is of excellent quality, particularly in the TCR regions and the interface between TCR and pMHC. There are very few density breaks, mainly in the BC loops of the I-A^k β 2 domains. Figure 2 is an omit map in the bound peptide region.

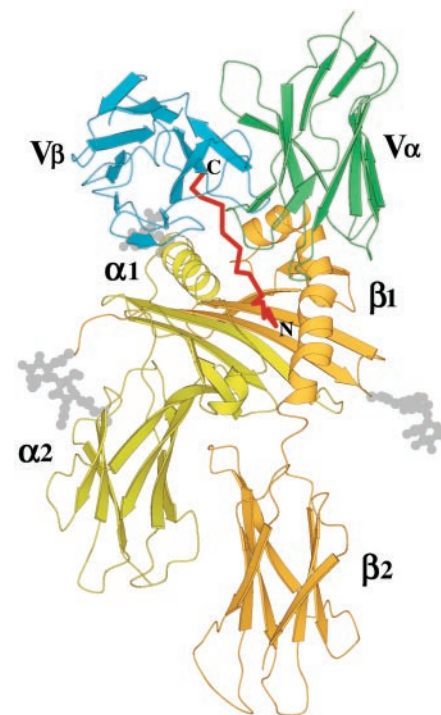


Fig. 1. Structure of the scD10-CA/I-A^k complex. Ribbon diagram showing the overall orientation of the scD10 V module to CA/I-A^k (pMHC). Domains are color-coded and labeled as follows: TCR V α (green), V β (blue), MHC α -chain (light green), and MHC β chain (orange). The secondary structures, β strands, and α helices of all component domains are defined by the program DSSP (63). The peptide is drawn in red tracing of the C α atoms with NH₂- and COOH-terminals labeled. The three glycans (GlcNAc) (gray) are drawn in ball-and-stick format. This figure was generated with the program MOLSCRIPT (64).

Data collection	
Resolution limit (last shell)	30.0 to 3.2 Å (3.31 to 3.20 Å)
Reflections	
Total number	501,406
Unique	52,592 (<i>I</i> > 0)
<i>I</i> / σ (<i>I</i>)	10.0 (2.2)
Completeness	95.2% (87.1%)
<i>R</i> _{merge} *	7.0% (29.2%)
Refinement statistics	
Resolution range	15.0 to 3.20 Å
Number of reflections†	46,332 (<i>F</i> > 0)
<i>R</i> _{work}	24.7%
<i>R</i> _{free}	29.3%
rms deviations	
Bonds	0.007 Å
Angles	1.4°
Dihedrals	30.2°
Improper	0.7°
Ramachandran plot	
Favored	71.4%
Allowed	21.7%
Generous	6.9%
Unfavored	0%

**R*_{merge} = $\sum(|i(hkl) - \langle i(hkl) \rangle|) / \sum i(hkl)$. †ln working set.

the two $V\alpha$ domains are superimposed, then the orientations of the two $V\beta$ domains differ only by a 3.7° rotation, indicating that $V\alpha$ - $V\beta$ dimerization is very similar for these two TCRs as well.

The Orthogonal Binding Mode

The orthogonal docking mode was not correctly predicted by either extensive mutagenesis studies (19) or modeling with a scD10 NMR structure in conjunction with a CA/I-A^k crystal structure as a starting point (23). To establish a quantitative and comparative measurement of binding orientation among TCRs and their pMHC ligands, we have defined an angle between two vectors. One vector passes through the mass centers of the $V\alpha$ and $V\beta$ domains of the TCR, while the other is drawn from the NH_2 -terminal $C\alpha$ atom at the P1 position (the first residue bound in the P1 pocket of the MHC) to the $COOH$ -terminal $C\alpha$ atom (the P9 position, the last buried residue) of the bound peptide. The angle for scD10-CA/I-A^k complex is 80° , very close to a right angle. Table 2 lists the orientation angle calculated for all known TCR-pMHC complex structures. For class I complexes, the peptide vector is defined between the anchoring residues at the two termini. The angles for TCR-pMHCI complexes span a broad range, from diagonal (45°) to close to orthogonal (70°). The difference between orthogonal and diagonal docking is illustrated by comparing the scD10-CA/I-A^k structure to the 2C-dEV8/H-2K^b (pMHCI) complex (Fig. 3, A and B).

Garboczi *et al.* (6) argue that in pMHC structures, there are two high "peaks" near the NH_2 -termini of the α -helical regions forming the side wall of the peptide binding groove. These two "peaks" limit the TCR-pMHC class I binding to a diagonal mode such that the TCR can fit at a low enough point on the MHC surface to contact the entire complexed antigenic peptide. Teng *et al.* (7) have compared three TCR-pMHC class I complex structures and identified a common docking mode of the TCR relative to the MHC with substantial variation of twist, tilt, and shift, however (Table 2 legend). Furthermore, we have noticed that the inherent left-handed twist of the eight-stranded β sheet that forms the platform of the binding groove is the structural basis for the breaks in the two helical regions, resulting in the formation of high "peaks." In this context, an MHC class II molecule is similar to an MHC class I molecule because all MHC molecules have the same platform. However, there are distinct features to the mode of peptide binding between the two classes. In the class I system, the 8- to 10-residue peptide has its termini anchored into two binding pockets whose unique chemical environments determine the polarity of the bound

peptide. In addition, the bulky side chains of the conserved Trp^{167 α} and Tyr^{84 α} from the MHC molecule occlude the peptide-binding groove at both ends. In class II MHC, these blocking side chains are replaced by smaller ones or are reoriented (or both); the open ends eliminate the peptide length restriction. Moreover, the peptide (15 to 20 amino acids long) binds to the class II MHC molecule with hydrogen bonds not only at the termini, but throughout the entire peptide through main-chain atoms (14).

Figure 4A shows the hydrogen-bonding pattern between the CA peptide and the I-A^k molecule that is conserved in other pMHC class II structures. Compared with the class I system, the P-3 to P-1 segment is an extension. This extension plays a unique role in the orthogonal docking mode. The peptide binding groove is much wider in the middle relative to its tapered ends so that the MHC class II molecule needs to use side chains of multiple conserved residues from $\alpha 1$ and $\beta 1$ helical regions to reach the peptide mainchain atoms. The residues include asparagine and glutamine, which form bidentate hydrogen

bonds to the peptide backbone. This hydrogen bonding pattern determines the peptide binding polarity in the class II MHC system (11-14). An important characterization of class II MHC molecules is that the $\alpha 1$ helix is two turns shorter in the NH_2 -terminus than the corresponding class I MHC molecule $\alpha 1$ helix (Fig. 4A). In particular, from Arg^{52 α} to Glu^{55 α} , the helix is replaced by an extended strand that reaches close enough to the NH_2 -terminal extension segment of the bound peptide to form a mini parallel β sheet through the use of main-chain atoms. The pair of main-chain—main-chain hydrogen bonds between Arg^{53 α} of the MHC class II molecule and the P-2 and P1 residues at the NH_2 -terminal part of the peptide are conserved among all known pMHCI structures. The beginning of the $\alpha 1$ helix, Gln^{57 α} , is at the high "peak," so from Glu^{55 α} to Arg^{52 α} toward the NH_2 terminus the chain runs down, away from the TCR binding surface. However, the left-handed twist of the mini β sheet then forces the NH_2 -terminus of the peptide to point in the opposite direction, curving up toward the TCR binding surface. Together,

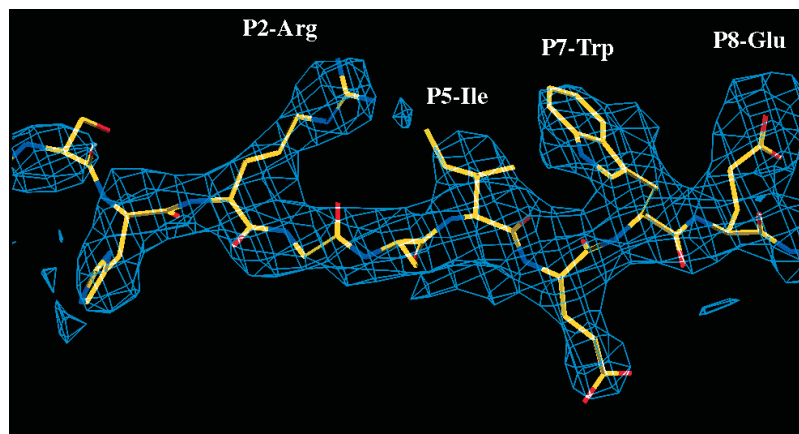


Fig. 2. Conalbumin peptide omit map. The core of CA (P-1 to P8) involved in TCR-based immune recognition is shown. Its σ_A weighted $2F_o - F_c$ omit electron density map (in blue) contoured at 1.0σ was generated with the program O (55) and prepared with a cover radius around the atoms shown. The omit map was generated by omitting the CA peptide entirely and after a round of torsion angle dynamics calculation. Carbon, oxygen, and nitrogen atoms are depicted in yellow, red, and blue, respectively. Those key peptide residues interacting with the D10 CDR loops are labeled.

Table 2. The orientation angle of a TCR onto a pMHC ligand.

TCR-peptide/MHC complex	Orientation angle* ($^\circ$)	MHC class
D10-CA/I-A ^k	80	II
2C-dEV8/H-2K ^b	45	I
N15-VS8/H-2K ^b	54	I
A6-Tax//HLA-A2	56	I
B7-Tax//HLA-A2	70	I

*The orientation angle of a TCR on MHC is defined as the angle between two vectors determined for the orientation of the TCR and pMHC, respectively. The vector representing the TCR direction is drawn from the mass center of $V\alpha$ to the mass center of $V\beta$. The vector representing the pMHC complex direction is drawn from the NH_2 -terminal $C\alpha$ atom to the $COOH$ -terminal $C\alpha$ atom of the peptide in the case of MHC class I. In the case of MHC class II, the vector is drawn from the P1 residue to the P9 residue of the peptide. Note that in Teng *et al.* (7), twist and tilt were used for semi-quantitative comparison among different TCR-MHC complexes. Essentially, the twist and tilt angles are two projections of the orientation angle more accurately defined here. Whereas the twist is a top view from the TCR toward the MHC, the tilt is a side view, perpendicular to the bound peptide.

the extended NH₂-terminus of the bound peptide and the MHC molecule now form a broader high “peak,” or a small protruding

“ridge” (Fig. 4B). For comparison, Fig. 4C is the same view of pMHC I taken from the 2C-dEV8/H-2K^b structure. Our scD10-CA/I-

A^k structure shows that a diagonal TCR docking would result in a collision between the V α domain of TCR and the pMHCII on the “left” side as viewed in Fig. 4B. Moreover, the tilt angle of a TCR relative to an MHC molecule (see Table 2 legend for the definition of tilt angle) exacerbates this potential clash by maintaining the V α domain in close proximity to MHC. We propose that while the TCR-pMHC class I docking may have more variation in terms of the orientation angle as demonstrated in Table 2, the topology of TCR binding to pMHC class II may be more closely restricted to an orthogonal mode due to the “ridge” described above. It is interesting that the protrusion of the peptide’s NH₂-terminus has been suggested as a site for disruption by DM in the process of exchanging CLIP for an antigenic peptide in the MHC class II molecule (24).

The Interface

The interaction between D10 and CA/I-A^k buries 1718 Å² of surface area, 861 Å² from the pMHC and 857 Å² from the TCR as determined by a 1.7 Å probe (25). Twenty-three percent of the pMHC buried surface involves the peptide. In general, the size of the buried surface is comparable to that previously reported for three class I ternary structures (1700 to 1880 Å²). However, the Sc value (the shape correlation statistic, a measurement of the degree of geometric match between two juxtaposed surfaces, where interfaces with Sc = 1 fit perfectly, whereas interfaces with Sc = 0 effectively define topologically uncorrelated surfaces) (26) of the interface between the scD10 V α V β module and the CA/I-A^k ligand is 0.70, higher than for class I TCR-pMHC interfaces whose Sc values range from 0.45 to 0.64 (27, 28). Moreover, the number of atomic contacts (29) in our class II complex structure is about twice as many as those for the class I complexes. For I-A^k, 68 atomic contacts exist with D10. By contrast, there are just 27 H-2K^b contacts with the 2C TCR, and 27 and 34 HLA-A2 contacts for the A6 and B7 TCRs, respectively. These results suggest a much better shape complementarity of the scD10-CA/I-A^k interface, and agree well with the higher affinity of D10 for its pMHC ligand relative to that of 2C, for example (1 to 2 μM versus 100 μM) (6, 7, 30). Of particular relevance is the finding that the additional interface atomic contacts can largely be ascribed to contacts between the TCR and the I-A^k molecule rather than between the TCR and the CA peptide (21). Assuming that these results are representative for other class II MHC-specific TCRs, the dominance of this TCR-MHC class II contact may explain why expression of a single pMHC class II complex in the thymus can select many different TCRs (31). Our data also show that despite having roughly the same buried surface area, the complementarity of TCR-pMHC recogni-

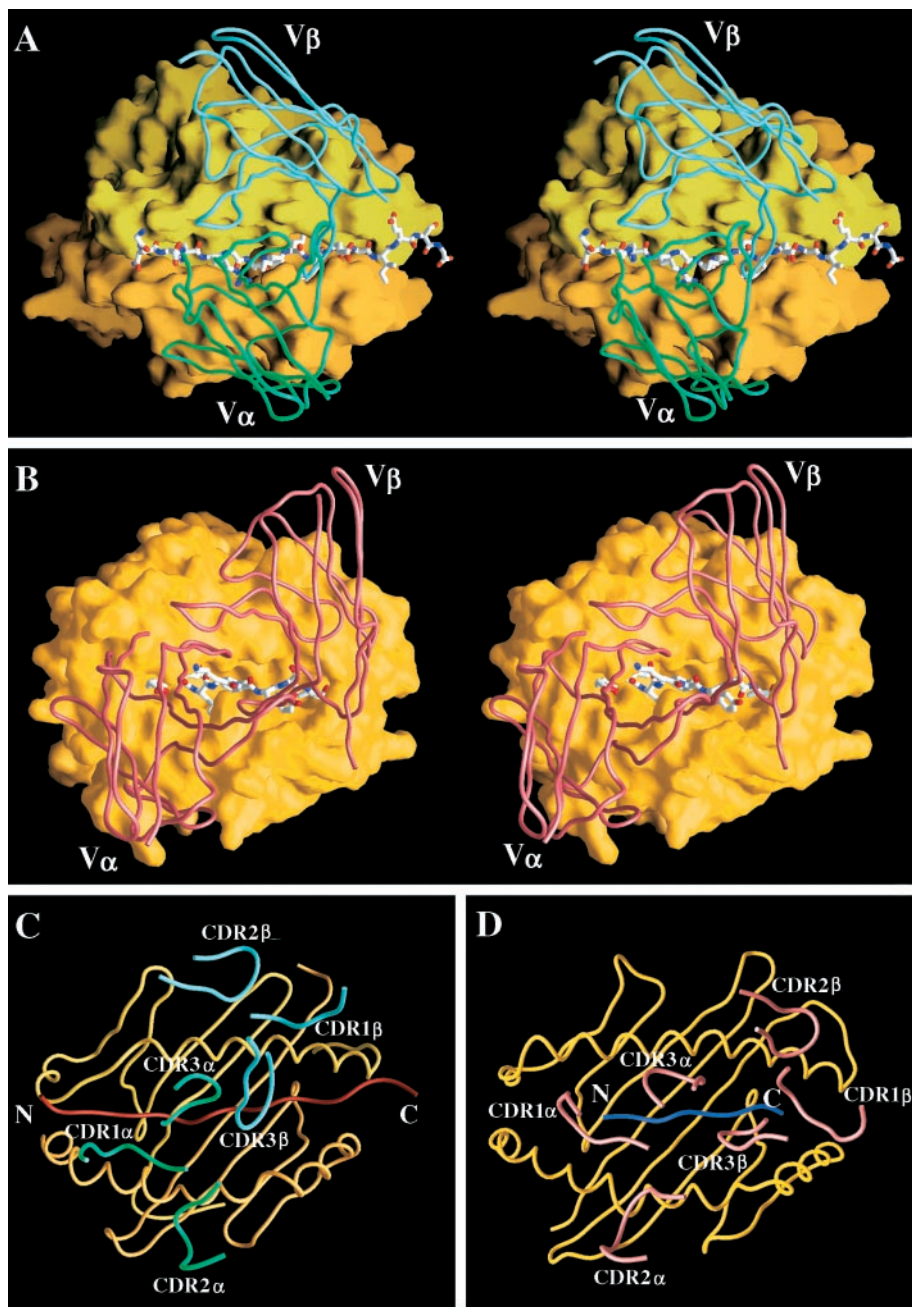


Fig. 3. Comparison of docking modes of TCRs on pMHC I versus pMHC II ligands. (A) Stereo view of scD10 (V α V β) on CA/I-A^k. The MHC molecule I-A^k is shown in molecular surface representation with the α chain in light green and the β chain in orange. The peptide is drawn in ball-and-stick format with NH₂- and COOH-termini at the left and right, respectively, in this view. The TCR V domains are shown as a backbone worm diagram, with V α in green and V β in blue. (B) Stereo view of 2C (V α V β) on dEV8/H-2K^b. The MHC molecule, H-2K^b, is shown in yellow in molecular surface representation. The peptide is drawn in ball-and-stick format with NH₂- and COOH-termini oriented as in (A). The TCR molecule is shown in brown as a backbone worm diagram. (C) Projection of the D10 TCR CDRs on the pMHC of CA/I-A^k. The CDR loops from V α and V β , as labeled, are in green and blue, respectively, shown as a backbone worm diagram. The peptide is in red with NH₂- and COOH-termini labeled. The α 1 and β 1 domains of I-A^k are in light green and orange, respectively. (D) Projection of the CDRs of TCR 2C on the pMHC of dEV8/H-2K^b. The CDR loops from V domains are drawn in brown. The peptide is in blue with NH₂- and COOH-termini labeled. The α 1 and α 2 H-K^b domains are in yellow. Figures 3 to 6 were prepared with the program GRASP (56).

tion surfaces can vary substantially from a low extreme to one even better than that of an antigen-Fab complex, as is the case for the scD10-CA/I-A^k complex. In complexes like 2C-dEV8/H-2K^b, a few large cavities (6, 7) contribute to poor shape complementarity. The presence or absence of such cavities may vary for different TCR-pMHC complexes, thereby influencing the shape of complementarity.

Of the total buried surface area, V α accounts for 519 Å², whereas V β accounts for 338 Å² of the TCR buried surfaces. This result is consistent with the notion that V α dominates in the interaction, which is generally true for the class I system as well. Our calculations show that the buried surface areas of V α and V β are 480 and 430 Å² for the 2C-dEV8/H-2K^b complex, 576 and 319 Å² for A6-Tax/HLA-A2, and 555 and 260 Å² for B7-Tax/HLA-A2, respectively. Perhaps more importantly, amongst different TCR-pMHC complexes, the variation in buried surfaces is significantly smaller for V α than for V β . Given that the rotation angle of known TCRs relative to their MHC ligands varies by as much as 35° (Table 2), these data suggest that the pivot point is closer to V α , so that the V α domain location on the pMHC will not change as much as the V β domain, which can alter dramatically (compare Fig. 3, A and B). Differences in the disposition of CDR loops reflect this pivot point (Fig. 3, C and D). Variability in TCR docking also arises from differences in the tilt angle as described by Teng *et al.* (7) and noted in the Table 2 legend. The extreme is the A6-Tax/HLA-A2 structure, where the large tilt essentially precludes CDR1 β and CDR2 β from making contact with the MHC molecule (6, 7). Given that V α is critical for TCR selection in thymic development as well as mature T cell activation (32), this V α dominance in immune recognition is not unexpected.

Comparison between the scD10 TCR interaction with CA/I-A^k analyzed here and the 2C TCR interaction with dEV8/K^b (6, 7) shows how a single TCR V β 8.2 domain can bind in distinct orientations to class I and class II pMHC ligands (Fig. 3, C and D). In the 2C-dEV8/K^b complex, the germ line V β 8.2 segment recognizes the K^b α 1 helical MHC residues Gln⁷², Val⁷⁶, and Arg⁷⁹ through CDR2, and the K^b α 2 helical residues Lys¹⁴⁶, Gln¹⁴⁹, and Ala¹⁵⁰ through CDR1 (6, 7). In the scD10-CA/I-A^k complex, the identical germ line V β 8.2 segment recognizes the I-A^k α 1 helical residues Lys³⁹, Gln⁵⁷, and Leu⁶⁰ through CDR2 and Gln⁶¹ through CDR1. Given that these two docking interactions are to highly conserved MHC class I and to highly conserved class II residues, respectively, it is tempting to speculate that the V β domain plays a major role in MHC recognition by both classes of TCRs and perhaps in pre-TCRs as well (33).

Antigenic Peptide Recognition

Although the peptide in our ternary structure is 16 residues long, designated as from P-3 to P13, the TCR interaction is restricted to the P-1 to P8 segment. The supplementary table (21) lists all the contacts to the peptide. It is noteworthy that of 27 atomic contacts with the peptide, 23 involve V α and only 4 involve V β . This dominance of the V α domain

in peptide recognition was not appreciated previously, although early molecular modeling efforts correctly suggested that an orthogonal TCR docking mode was possible (3). The spiral conformation of bound peptide (Fig. 4) (14) dictates that of the deeply buried peptide residues, only those at positions P2, P5, and P8 are accessible to the TCR molecule. The Trp at the P7 position is an excep-

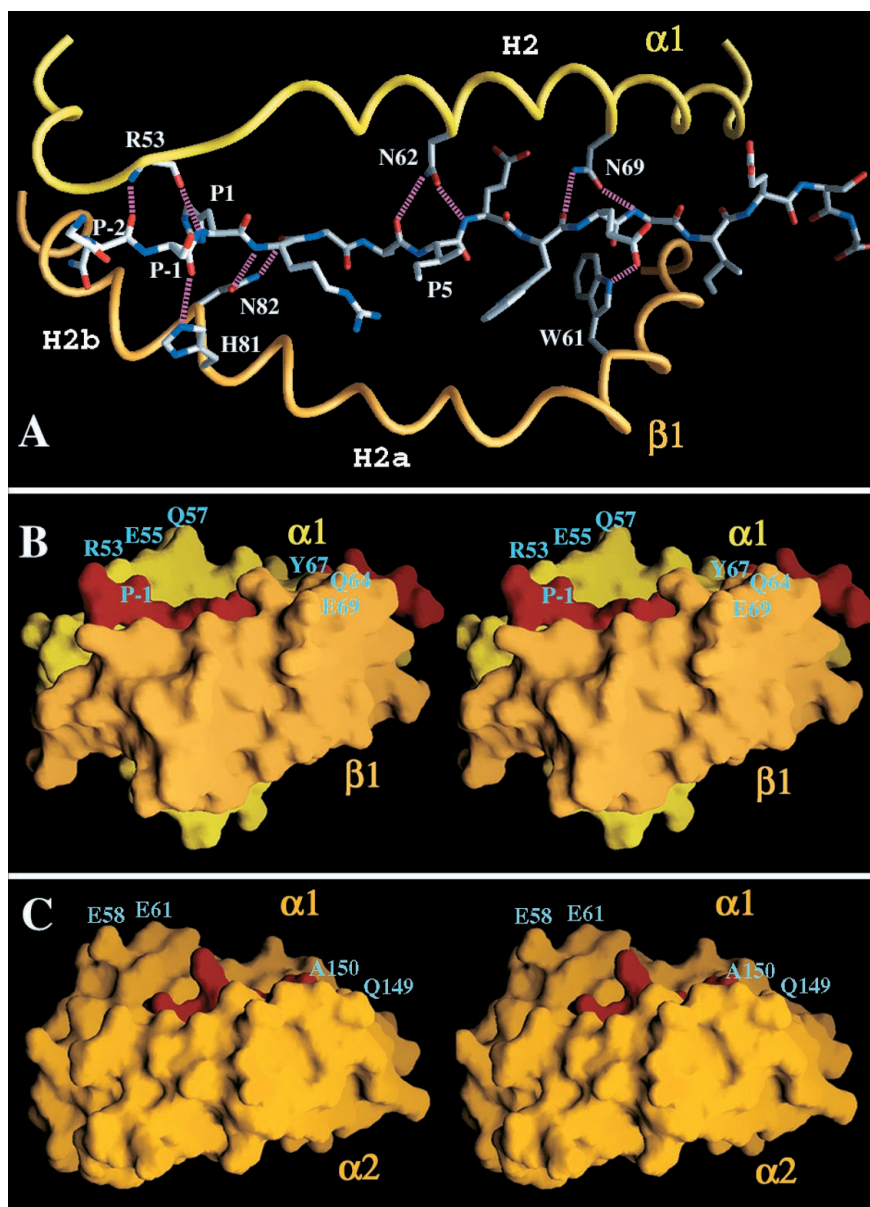


Fig. 4. The high point "ridge" in pMHCII ligands is created, in part, by the peptide. (A) Hydrogen bond network between the CA peptide and I-A^k. The 10 hydrogen bonds between the CA and I-A^k are shown as magenta dashed lines. These bonds are conserved in known pMHCII structures. The helical regions from I-A^k (colors as in Fig. 3) are shown as a backbone worm diagram with those side chains and main-chain atoms involved in the interactions displayed. The H2 α 1 helix and H2a and H2b β 1 helices are labeled. (B) Stereo view of the molecular surface of I-A^k (α 1 β 1) together with the CA peptide (red), showing the surface topology of the D10 docking platform on the CA/I-A^k ligand. (C) Stereo view of the molecular surface of H-2K^b (α 1/ α 2) together with the dEV8 peptide (red) (6, 7) in the same view as in (B), showing the smaller high point on the left side of the docking platform for the MHC class I-restricted TCR molecule. The peptide is mostly buried and makes little, if any, contribution to the elevated points. Those residues contributing to the high points of the platform are labeled in cyan in (B) and (C).

tion due to its bulky indole ring, which is partially exposed on the TCR binding surface. As for the rest of the peptide, the backbone of the P-2 residue is engaged in a mini parallel β sheet with the MHC molecule as discussed above, whereas the P-3 and the COOH-terminal three residues (P11 to P13) have no contacts with MHC or TCR whatsoever, although well defined by unambiguous densities.

The P2 residue is an Arg. It forms multiple salt bridges with both Asp^{30 α} in the D10 CDR1 α and I-A^k Glu^{74 β} , respectively. The same TCR Asp^{30 α} also interacts with I-A^k Arg^{70 β} . Moreover, the upward-pointing P2-Arg is within van der Waals contacts to backbone of CDR3 α Gly^{99 α} and CDR1 α Thr^{28 α} (21). This knitted local structure packs closely onto the side chain of Ile at the P5 position from the NH₂-terminal side of the peptide. The P5 residue is important structurally and biologically. Alteration of this residue adversely affects D10-TCR recognition of CA/I-A^k (19). The side chain of Ile at P5 fits extremely well into a hydrophobic pocket. Apart from the neutralized network discussed above, on the COOH-terminal peptide side is the indole ring of the Trp at P7 stacking onto the isobutyl group of the P5-Ile. On the top, from the TCR direction, the P5-Ile contacts the backbone of the tip of CDR3 α , which consists of Gly^{99 α} -Ser^{100 α} -Phe^{101 α} . The phenolic ring of Phe^{101 α} bends toward the P7-Trp position. The exposed tip of the indole ring of P7-Trp makes contacts with the Phe^{101 α} aromatic ring. The other peptide residue engaged in recognition is the P8-Glu residue. P8-Glu forms bifurcated hydrogen bonds to side chains of Tyr^{60 β} and Tyr^{67 β} of the I-A^k molecule. Only the aliphatic portion of the P8-Glu side chain makes van der Waals interactions with CDR3 α Phe^{101 α} and the aliphatic part of CDR3 β Gln^{97 β} . In addition, there is one hydrogen bond between the carbonyl oxygen of Gly^{96 β} and the side chain of P8-Glu. Together, it appears that the TCR recognition of the particular antigenic peptide in question is largely hydrophobic and involves a number of backbone associations with the TCR molecule. Although scD10 TCR recognition of CA/I-A^k is centered at the P5 position, it is also coordinated with interactions to peptide residues at P2, P7, and P8. The observed contacts are consistent with studies mapping the D10 footprint onto CA/I-A^k (20, 34).

Implications for Class II MHC-based Immune T Cell Recognition

The current scD10-CA/I-A^k complex offers several insights into immune recognition of other pMHC class II ligands by other TCRs. First, the size of a TCR footprint on the MHC covers maximally nine peptide residues (~25 Å). Hence, while MHC class II molecules capture peptides of substantially larger length, only a subset of residues is "read out"

by the bound TCR. Second, the P5 residue of the MHC-bound peptide occupies the central position [corresponding to the P4 position of the MHC class I-bound peptide (32)]. As such, this central, solvent-exposed residue is critically important for the TCR binding process. Therefore, even a minor conservative substitution at this residue can destroy binding (that is, null ligand) or lead to altered peptide ligands with very weak agonist or in fact, antagonist activity (35–37). Third, for all class II molecules examined, there appear to be three to four pMHC binding pockets (at P1, P4, P6, and P9 for I-A^k, I-E^k, and DR and P1, P4, and P9 for I-A^d). Only several upward-pointing peptide residues can serve as direct TCR contacts. On the basis of the observed molecular envelope of the TCR and the observed orthogonal orientation for class II MHC-restricted TCR interaction, it is likely that these basic principles apply in a general way to recognition of multiple pMHCII ligands including HEL_{48–62}/I-A^k (12, 38), Hb_{64–76}/I-E^k (39), moth cytochrome C (MCC)_{93–103}/I-E^k (35), and DR2-restricted myelin basic protein (MBP)_{85–99} (40, 41). Moreover, it has been suggested that a single TCR can recognize multiple pMHCII ligands (40, 41). As the class II-specific TCR focuses on the central P5 residue, mutations that affect non-P5 positions may be less detrimental to the recognition process.

Despite overall structural similarity, CDR3 conformations appear to differ between free and complexed scD10. In the x-ray structure of the complex, the CDR3 loops are close to one another. Packing among the side chain of Gln^{106 β} , Ala^{104 β} , and Leu^{104 α} forms a hydrophobic core between the CDR3 loops. In contrast, there is no evidence that CDR3 α packs with CDR3 β in unligated scD10. Numerous nuclear Overhauser enhancements (NOEs) are observed between the methyl group of Ala^{104 β} and other CDR3 β residues (Gly⁹⁶, Gln⁹⁷, Arg⁹⁹, Glu¹⁰⁵), but contacts to Leu^{104 α} or other CDR3 α residues are not observed. In all of the calculated NMR structures, the CDR3 α and CDR3 β loops are well separated. The backbones of both CDR3s are also highly mobile on the picosecond time scale in the free protein (23), suggesting that they are not tightly packed. Hence, during immune recognition, the mobile CDR3 loops of scD10 assume their pMHCII binding conformation, clamping down on the central peptide region.

Structural Basis of Alloreactivity

About 1 to 10% of peripheral T cells are able to recognize allogeneic MHC molecules to which they were never exposed (42). The precise molecular basis of alloreactivity is yet to be fully defined. In this regard, the complex of scD10-CA/I-A^k is informative because the D10 TCR not only recognizes the antigenic CA peptide

bound to I-A^k but also responds to all MHC class II molecules whose I-A β chain contains the sequence PEI at positions 65 to 67, including I-A^b, v, p, q, d (19). In comparison, MHC class II molecules having a Tyr at this position such as I-A^f, k, r, s, u, g7, cannot stimulate D10 cells in the absence of the CA peptide. Various mutagenesis studies conducted on D10 showed that a hybrid I-A^k α /I-A^b β MHCII molecule can stimulate D10 cells in the absence of exogenous antigen, suggesting that polymorphic residues critical for alloreactivity are located on the I-A β chain.

In order to elucidate this source of alloreactivity from the structural perspective, we compared the CA/I-A^k ligand and the alloreactive I-A^d molecule. The latter was taken from the recently solved x-ray structure of

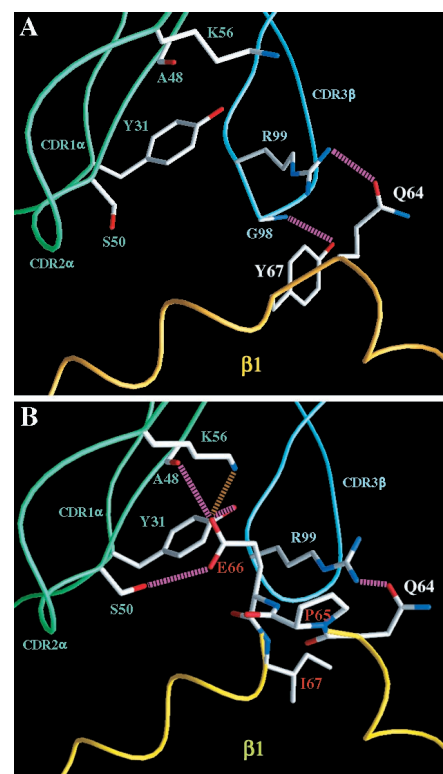


Fig. 5. Structural basis of alloreactivity. (A) V β CDR3 interaction with self MHC. The interactions between the side chains of Tyr⁶⁷ and Gln⁶⁴ on the I-A^k β 1 helix (orange) and the D10 V β CDR3 loop (blue) residues Gly⁹⁸ and Arg⁹⁹ are shown. The side chains and main-chain atoms displayed on the CDR1 and CDR2 loops of D10 V α make no hydrogen bonds or salt bridges to I-A^k β 1 helix residues. (B) Model of scD10 bound to I-A^k based on superposition of I-A^d and I-A^k. The α 1 H2 helix and the β 1 region between and including H2a-H2b of the two class II MHC molecules were superimposed (46 C α atoms, rmsd = 0.55 Å). The potential interactions between Glu⁶⁶ from the PEI_{65–67} motif of I-A^d to Tyr³¹ of CDR1 and Ala⁴⁸, Ser⁵⁰, and Lys⁵⁶ of CDR2 of D10 V α are indicated. The hydrogen bonds are drawn as magenta dashed lines. The salt bridge is drawn as a brown dashed line.

ovalbumin (323–339) complexed with I-A^d (13). Because all residues from I-A^k involved in the interaction with D10 as listed in the supplementary table are conserved in I-A^d, with the exception of β residues Tyr⁶⁷ in I-A^k and Pro⁶⁵-Glu⁶⁶-Ile⁶⁷ (PEI) in I-A^d, it is likely that D10 docks onto I-A^d in the same way as onto I-A^k. A model was constructed with the I-A^d superimposed onto the scD10-CA/I-A^k complex (Fig. 5). The major structural difference involves the β -chain residues Pro⁶⁵, Glu⁶⁶, and Ile⁶⁷ in I-A^d, which form a protrusion interrupting the β -chain α helix. As a consequence, in I-A^d, residue Ile⁶⁷ assumes a similar position to residue Tyr⁶⁷ in I-A^k (Fig. 5, A and B). The aliphatic side chain of the Ile⁶⁷ in I-A^d can replace the aromatic ring on the side chain of Tyr⁶⁷ in I-A^k, forming van der Waal's contacts with the V β CDR3 loop. To avoid steric clashes, side chains from residue Arg⁹⁹ of D10 V β and residue Glu⁶⁶ of I-A^d β 1 are rotated and the main-chain conformation around PEI on I-A^d is slightly modified. The backbone NH vectors of residues directly adjacent to Arg⁹⁹ are among the most mobile in scD10 (23).

Glu⁶⁶ can form multiple potential interactions with CDR1 and CDR2 of D10 V α (Fig. 5B). Additionally, the hydrogen bond between Gln⁶⁴ and Arg⁹⁹ from CDR3 of V β is preserved. Therefore, despite loss of one hydrogen bond of Tyr⁶⁷ to D10 V β Gly⁹⁸ (Fig. 5A), these potential additional contacts between the CDR loops of D10 V α and the inserted PEI residues can enhance the affinity between MHC and D10. Other TCR allo-pMHCII interactions cannot be excluded. Consistent with this view, it has also been suggested that in the case of the 2C allo-

MHCI response (L^d), allorecognition results from increased interaction between the 2C TCR V β domain and the allostimulus (43). The ability of exposed MHC helical polymorphic residues to permute the number and nature of contacts with the TCR is a feature of other class II MHC-restricted allogeneic responses (44, 45). For example, the naturally occurring I-A^b mutant H-2^{bm12} generates a strong alloresponse in H-2^b mice. This molecule differs from I-A^b at only three positions: 67 β , 70 β , and 71 β .

While self peptides bound to MHC have been shown to play a critical role in alloreactivity against MHC class I molecules (45), less is known about the nature of peptide ligands in class II MHC-based alloreactivity (46). Given the additional contacts between the D10 TCR and I-A^d, it is possible that there are fewer interactions required between the peptide (or peptides) associated with I-A^d molecules and the D10 TCR. However, a peptide (or peptides) must also be involved because replacement of the PEI sequence in lieu of Tyr⁶⁷ in the I-A^k β chain is not sufficient to create an allostimulatory molecule for the D10 T cell clone (19).

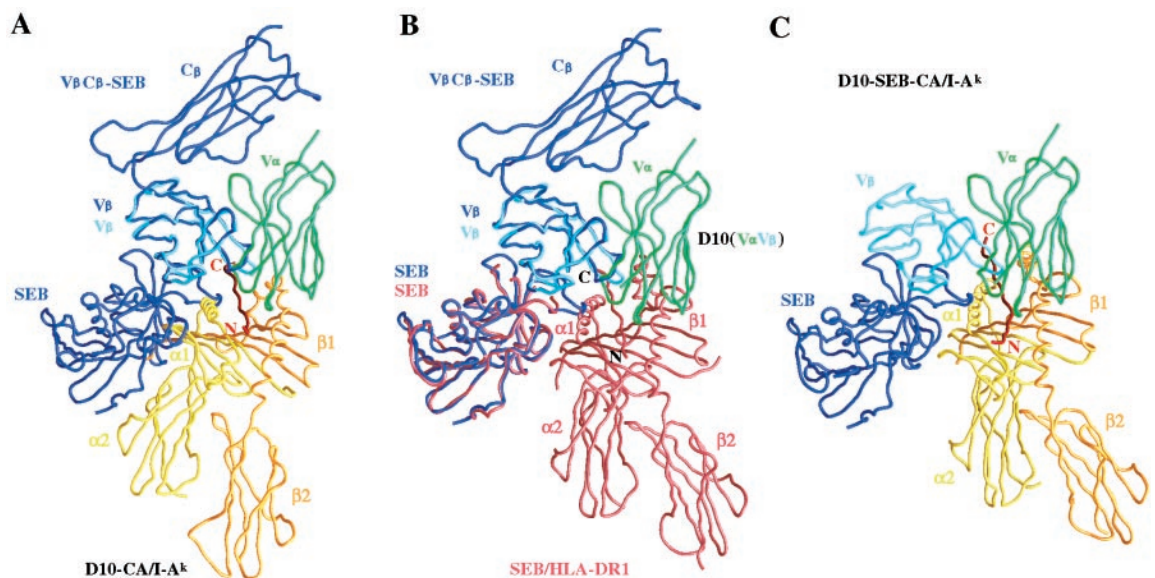
Superantigen Binding

Superantigens (SAGs) are a family of immunostimulatory and disease-causing proteins derived from bacterial or endogenous retroviral genes that are capable of activating a large fraction of the T cell population (47). In general, the activation appears to require a bridging interaction between the V β domain of the TCR and an MHC class II molecule. Although crystal structures (48, 49) showing the detailed interactions between SEB, a rep-

resentative bacterial SAG, and a TCR V β 8.2 chain or SEB and the HLA-DR1 class II MHC molecule have been determined, the physiologically relevant tripartite TCR-SAG-pMHC complex has not yet been characterized. A structural model of TCR-SAG-pMHC complex was previously generated (48, 49) based on least-squares superposition of (i) the 14.3.d V β C β -SEB complex, (ii) the SEB-HLA-DR1 complex, and (iii) the 2C TCR $\alpha\beta$ heterodimer. However, because the docking mode of TCR on the class II MHC was structurally unknown and presumed to be similar to the observed diagonal mode of TCR on class I MHC, it was noted that the rotational orientation of the TCR and MHC molecules in the predicted TCR-SEB-pMHC complex was substantially different ($\sim 40^\circ$) from the 2C-dEV8/K^b complex. The structural determination of the D10-CA/I-A^k complex reported here enables us to offer additional insight into the nature of the SAG binding to TCR and pMHC.

Fig. 6A shows the 14.3.d V β 8.2 C β -SEB complex superimposed onto the D10-CA/I-A^k complex. Because the TCR docks on the MHC molecule in a nearly perpendicular manner, the SEB directly interacts with the MHC α 1 helix without any requirement for TCR rotation. However, certain segments of SEB and the α 1 helix collide. Because there is no significant conformational change observed for either of the component domains involved in this interaction, we reasoned that a relative domain movement could alleviate any steric clash. To test this idea, we removed I-A^k from the complex and then superimposed the SEB HLA-DR1 complex onto the 14.3.d V β C β -SEB complex (Fig. 6B). From

Fig. 6. Model of the scD10-SEB-pMHCII interaction complex. **(A)** Superposition of the V β C β -SEB complex (in dark blue) and the scD10-CA/I-A^k complex. The V β domains from each complex were used for least-square fitting (92 C α atoms from residues Val³-Gly⁹⁴ of V β , rmsd = 0.67 Å). The color scheme for the complex of D10-CA/I-A^k is as labeled. **(B)** The superposition of the SEB/HLA-DR1 complex (in brown) to the already superimposed V β C β -SEB complex (in dark blue) and scD10 (V α V β) module derived from (A). The two SEB superantigen molecules were used for least-square fitting (83 C α atoms, rmsd = 0.63 Å). **(C)** scD10-SEB-CA/I-A^k interaction complex. DR1 in Fig. 6B has been replaced with I-A^k on the basis of structural alignment of residues of the two helices of each MHC molecule (43 C α atoms, rmsd = 1.02 Å).



this second model, we observed that the direct interaction between V β and the MHC α 1 helix is disrupted by SAG. The key interaction site for SEB involves CDR2 (Tyr⁵⁰, Ala⁵², Gly⁵³, Ser⁵⁴, Thr⁵⁵) and certain other V β residues (Glu⁵⁶, Lys⁵⁷, Tyr⁶⁵, Lys⁶⁶, Ala⁶⁷) as reported by Li *et al.* (49). In this way, the superantigen wedges itself between V β and the MHC class II α 1 helix, forcing the MHC to swing away from V β and toward V α while preserving the direct interaction between the V α domain of the TCR and MHC class II β 1 helix. This latter interaction has been proposed to be critical in stabilizing the TCR-SAG-pMHC complex. In fact, T cell activation by SAG is believed to be dependent on the interaction between a given TCR V α domain and the MHC class II β 1 helix (50). In Fig. 6C, we have replaced the HLA-DR with the I-A^k molecule on the basis of the structural alignment, thereby creating a third model. From the latter, we estimate that the relative swing angle between TCR and MHC in the TCR-SAG-pMHCII complex compared with the TCR-pMHCII complex is $\sim 17^\circ$.

Differential TCR Binding and Coreceptor Selection in the Thymus

Given that there are no intrinsic structural differences between class I versus class II MHC-restricted TCR V modules as shown above, what directs expression of a TCR to the proper CD4 or CD8 subset? During thymocyte development, progenitor cells transit from a CD4⁻CD8⁻ double negative (DN) stage through a CD4⁺CD8⁺ double positive (DP) stage and then into a CD4⁺CD8⁻ or CD4⁻CD8⁺ single positive (SP) stage (51). Selection for maturation occurs upon the interaction of thymocytes with stromal cells expressing self-pMHC I or self-pMHC II ligands within the thymus, beginning at the DP stage where the TCR first appears. Differentiation to the SP thymocyte stage, however, requires a match between the MHC class specificity of the TCR that a thymocyte bears and the CD4 or CD8 coreceptor it expresses. To explain how a thymocyte precisely coordinates coreceptor expression and TCR specificity, two models have been proposed (52). The "instruction model" argues that coengagement of TCR and CD4 or CD8 on a DP thymocyte specifically signals the cell to move down one pathway while extinguishing the expression of the inappropriate coreceptor. On the other hand, the "selection model" postulates that cells initiate stochastically or otherwise a process that terminates expression of one of the two coreceptors. If the correct match was chosen, then the cell further differentiates, but if not, differentiation is stalled.

Distinctions between class I versus II pMHC complexes and variation in TCR docking observed here offer strong backing

for the notion of the "instruction model." We suggest that depending on the degree of complementarity of a given TCR recognition surface and a self-pMHC I or self-pMHC II complex, binding occurs and a diagonal docking mode with substantial variability onto pMHC I or a preferred orthogonal docking mode onto pMHC II is established. Subsequently, CD8 preferentially coengages with the former and CD4 with the latter. Expression of the irrelevant coreceptor is then extinguished. On the basis of CD8-MHC class I crystal structures and on mapping of MHC class II residues involved in CD4 binding, the two coreceptors likely occupy an "homologous" orientation relative to the TCR (17, 18). Thus, we postulate that the differential TCR docking to self-pMHC I versus self-pMHC II contributes specificity for coordination of appropriate coreceptor selection.

References and Notes

1. E. L. Reinherz, S. C. Meuer, S. F. Schlossman, *Immunol. Rev.* **74**, 83 (1983); P. Marrack and J. Kappler, *Adv. Immunol.* **38**, 1 (1986); H. Clevers, B. Alarcon, T. Wileman, C. Terhorst, *Ann. Rev. Immunol.* **6**, 629 (1988); J. D. Ashwell, R. D. Klausner, *Ann. Rev. Immunol.* **8**, 139 (1990).
2. O. Acuto *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3851 (1984); M. Fabbri, O. Acuto, J. E. Smart, E. L. Reinherz, *Nature* **312**, 269 (1984); G. K. Sim *et al.*, *Nature* **312**, 771 (1984); Y. Chien *et al.*, *Nature* **312**, 31 (1984); N. R. Gascoigne, Y. Chien, D. M. Becker, J. Kvalner, M. M. Davis, *Nature* **310**, 387 (1984); S. M. Hedrick, E. A. Nielsen, J. Kvalner, D. I. Cohen, M. M. Davis, *Nature* **308**, 153 (1984); S. M. Hedrick, D. I. Cohen, E. A. Nielsen, M. M. Davis, *Nature* **308**, 149 (1984); P. Patten *et al.*, *Nature* **312**, 40 (1984); H. Saito *et al.*, *Nature* **312**, 36 (1984); Y. Yanagi *et al.*, *Nature* **308**, 145 (1984); A. C. Hayday *et al.*, *Nature* **316**, 828 (1985); J. Novotny, S. Tonegawa, H. Saito, D. M. Kranz, H. N. Eisen, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 742 (1986); C. Chothia, D. R. Boswell, A. M. Lesk, *EMBO J.* **7**, 3745 (1988).
3. M. M. Davis and P. J. Bjorkman, *Nature* **334**, 395 (1988).
4. G. A. Bentley, G. Boulton, K. Karjalainen, R. A. Mariuzza, *Science* **267**, 1984 (1995); B. A. Fields *et al.*, *Science* **270**, 1821 (1995); D. Housset *et al.*, *EMBO J.* **16**, 4205 (1997).
5. K. C. Garcia *et al.*, *Science* **274**, 209 (1996); J. Wang *et al.*, *EMBO J.* **17**, 10 (1998).
6. D. N. Garboczi *et al.*, *Nature* **384**, 134 (1996); K. C. Garcia *et al.*, *Science* **279**, 1166 (1998); Y. H. Ding *et al.*, *Immunity* **8**, 403-411 (1998); Y. H. Ding, B. M. Baker, D. N. Garboczi, W. E. Biddison, D. C. Wiley, *Immunity* **11**, 45 (1999).
7. M.-K. Teng *et al.*, *Curr. Biol.* **8**, 409 (1998).
8. B. Benacerraf, *Harvey Lect.* **67**, 109 (1973); B. Benacerraf, *J. Immunol.* **120**, 1809 (1978); M. J. Bevan, *Nature* **269**, 417 (1977); R. M. Zinkernagel and P. C. Doherty, *Nature* **248**, 701 (1974); R. M. Zinkernagel *et al.*, *J. Exp. Med.* **147**, 882 (1978); R. N. Germain, *Cell* **76**, 287 (1994).
9. F. Momburg, J. Roelse, G. J. Hämmerling, J. J. Neefjes, *J. Exp. Med.* **179**, 1613 (1994); M. J. Androlewicz and P. Cresswell, *Immunity* **1** (1994).
10. A. Rudenski, P. Preston-Hurlburt, S. C. Hong, A. Barlow, C. A. J. Janeway, *Nature* **353**, 622 (1991); R. M. Chicz *et al.*, *Nature* **358**, 764 (1992); D. F. Hunt *et al.*, *Science* **256**, 1817 (1992).
11. D. H. Fremont, M. Matsumura, E. A. Stura, P. A. Peterson, I. A. Wilson, *Science* **257**, 919 (1992); D. R. Madden, J. C. Gorga, J. L. Strominger, D. C. Wiley, *Cell* **70**, 1035 (1992); M. Matsumura, D. H. Fremont, P. A. Peterson, I. A. Wilson, *Science* **257**, 927 (1992); J. H. Brown *et al.*, *Nature* **364**, 33 (1993); L. J. Stern *et al.*, *Nature* **368**, 215 (1994); D. H. Fremont, W. A. Hendrickson, P. Marrack, J. W. Kappler, *Science* **272**, 1001 (1996); A. Dessan, C. M. Lawrence, S. Cupo, D. M. Zaller, D. C. Wiley, *Immunity* **7**, 473 (1997); V. L. Murthy and L. J. Stern, *Structure* **5**, 1385 (1997).
12. D. H. Fremont, D. Monnaie, C. A. Nelson, W. A. Hendrickson, E. R. Unanue, *Immunity* **8**, 305 (1998).
13. C. A. Scott, P. A. Peterson, L. Teyton, I. A. Wilson, *Immunity* **8**, 319 (1998).
14. Reviewed in L. J. Stern and D. C. Wiley, *Structure* **2**, 245 (1994); D. R. Madden, *Annu. Rev. Immunol.* **13**, 587 (1995).
15. E. L. Reinherz and S. F. Schlossman, *Cell* **19**, 821 (1980).
16. S. Meuer, S. F. Schlossman, E. L. Reinherz, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4395 (1982); W. E. Biddison, P. E. Rao, M. A. Talle, G. Goldstein, S. Shaw, *J. Exp. Med.* **156**, 1065 (1982); A. M. Krensky, C. S. Reiss, J. W. Mier, J. L. Strominger, S. J. Burakoff, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2365 (1982).
17. C. A. Doyle and J. L. Strominger, *Nature* **330**, 256 (1987); R. König, L. Y. Huang, R. Germain, *Nature* **356**, 796 (1992); U. Moebius, P. Pallai, S. C. Harrison, E. L. Reinherz, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8259 (1993); J. M. Connolly, T. H. Hansen, A. L. Ingold, T. A. Potter, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2137 (1990); R. D. Salter *et al.*, *Nature* **345**, 41 (1990).
18. G. F. Gao *et al.*, *Nature* **387**, 630 (1997); P. Kern *et al.*, *Immunity* **9**, 519 (1998).
19. J. Kaye, S. Porcelli, J. Tite, B. Jones, C. A. J. Janeway, *J. Exp. Med.* **158**, 836 (1983); P. Portales, J. M. Rojo, C. A. J. Janeway, *J. Mol. Cell. Immunol.* **4**, 129 (1989); S. C. Hong *et al.*, *Cell* **69**, 999 (1992); S.-C. Hong *et al.*, *J. Immunol.* **159**, 4395 (1997).
20. D. B. Sant'Angelo *et al.*, *Immunity* **4**, 367 (1996).
21. The supplemental Web table is available at www.sciencemag.org/feature/data/1045623.shl.
22. R. T. Carson, K. M. Vignali, D. L. Woodland, D. A. Vignali, *Immunity* **7**, 387 (1997).
23. B. J. Hare *et al.*, *Nature Struct. Biol.* **6**, 574 (1999).
24. L. Mosyak, D. M. Zaller, D. C. Wiley, *Immunity* **9**, 377 (1998).
25. C. Chothia and J. Janin, *Nature* **256**, 705 (1975); M. L. Connolly, *J. Appl. Crystallogr.* **16**, 548 (1983).
26. M. C. Lawrence and P. M. Coleman, *J. Mol. Biol.* **234**, 946 (1993).
27. X. Ysern, H. Li, R. A. Mariuzza, *Nature Struct. Biol.* **5**, 412 (1998).
28. K. C. Garcia, L. Teyton, I. A. Wilson, *Annu. Rev. Immunol.* **17**, 369 (1999).
29. S. Sheriff, W. A. Hendrickson, J. L. Smith, *J. Mol. Biol.* **197**, 273 (1987).
30. E. L. Reinherz *et al.*, unpublished results.
31. C. P. Liu, D. Parker, J. Kappler, P. Marrack, *J. Exp. Med.* **186**, 1441 (1997); L. Ignatowicz *et al.*, *Immunity* **7**, 179 (1997); S. Tourne *et al.*, *Immunity* **7**, 187 (1997); C. D. Surh, D. S. Lee, W. Fung Leung, L. Karlsson, J. Sprent, *Immunity* **7**, 209 (1997).
32. Y. Ghendler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10061 (1998); B. T. Backström, U. Müller, B. Hausmann, E. Palmer, *Science* **281**, 835 (1998).
33. Y. Shinkai *et al.*, *Science* **259**, 822 (1993); H. J. Fehling, A. Krotkova, C. Ruf-Saint, H. von Boehmer, *Nature* **375**, 795 (1995); C. N. Levelt, B. Wang, A. Ehrfeld, C. Terhorst, K. Eichmann, *Eur. J. Immunol.* **25**, 1257 (1995); Y. Xu, L. Davidson, F. W. Alt, D. Baltimore, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2169 (1996); S. Koyasu *et al.*, *Int. Immunol.* **9**, 1475 (1997).
34. Results from specificity and sequence analysis of T cell hybridomas derived from mice carrying a given V α 2 or V β 8.2 D10 TCR transgene immunized with altered peptide ligands of CA (20) are consistent with the structural information reported here. For example, immunization of D10 TCR V α 2 transgenic (tg) mice with the Glu⁸→Ala CA peptide variant gives rise to T cell hybridomas, all of which use V β 8.2 but with variation in CDR3 β , consistent with the view that CDR3 β is involved in recognition of the P8 position. Immunization of D10 TCR α tg mice or D10 TCR β tg mice with the Ile⁵→Lys CA variant both failed to generate specific hybridomas, implying that CDR3 α or CDR3 β (or both) may be important for Ile⁵ recognition. Immunization of D10 TCR β tg mice with the Arg²→Asp CA variant resulted in a switch in V α usage from V α 2 to V α 8, suggesting that the germ

- line CDR1 or CDR2 loops (or both) interact with this peptide residue.
35. P. A. Reay, R. M. Kantor, M. M. Davis, *J. Immunol.* **152**, 3946 (1994).
 36. G. J. Kersh and P. M. Allen, *J. Exp. Med.* **184**, 1259 (1996).
 37. K. Matsui, J. J. Boniface, P. Steffner, P. A. Reay, M. M. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12862 (1994); D. S. Lyons *et al.*, *Immunity* **5**, 53 (1996); G. J. Kersh, E. N. Kersh, D. H. Fremont, P. M. Allen, *Immunity* **9**, 817 (1998).
 38. P. M. Allen *et al.*, *Nature* **327**, 713 (1987).
 39. B. D. Evavold, J. Sloan-Lancaster, K. J. Wilson, J. B. Rothbard, P. M. Allen, *Immunity* **2**, 655 (1995).
 40. K. J. Smith, J. Pyrdol, L. Gauthier, D. C. Wiley, K. W. Wucherpfennig, *J. Exp. Med.* **188**, 1511 (1998).
 41. K. W. Wucherpfennig and J. L. Strominger, *Cell* **80**, 695 (1995).
 42. L. A. Sherman and S. Chattopadhyay, *Annu. Rev. Immunol.* **11**, 385 (1993).
 43. J. A. Speir *et al.*, *Immunity* **8**, 553 (1998).
 44. I. F. McKenzie *et al.*, *J. Exp. Med.* **150**, 1323 (1979); J. Bill *et al.*, *J. Exp. Med.* **169**, 115 (1989).
 45. Reviewed in J. A. Frelinger and M. McMillan, *Immunol. Rev.* **154**, 45 (1996).
 46. C. Daniel, S. Horvath, P. M. Allen, *Immunity* **8**, 543 (1998).
 47. M. T. Scherer, L. Ignatowicz, G. M. Winslow, J. W. Kappler, P. Marrack, *Annu. Rev. Immunol.* **9**, 101 (1993); S. R. Webb, N. R. J. Gascoigne, *Curr. Opin. Immunol.* **6**, 467 (1994).
 48. T. S. Jardetzky *et al.*, *Nature* **368**, 711 (1994).
 49. H. Li *et al.*, *Immunity* **9**, 807 (1998).
 50. D. L. Woodland and M. A. Blackman, *Immunol. Today* **14**, 208 (1993); K. Daly *et al.*, *J. Immunol.* **155**, 27 (1995); M. A. Blackman and D. L. Woodland, *Immunol. Res.* **15**, 98 (1996); D. Donson *et al.*, *J. Immunol.* **158**, 5229 (1997); A. M. Deckhut, Y. Chien, M. A. Blackman, D. L. Woodland, *J. Exp. Med.* **180**, 1931 (1994); N. Labrecque, J. Thibodeau, W. Mourad, R. P. Sekaly, *J. Exp. Med.* **180**, 1921 (1994).
 51. E. L. Reinherz, P. C. Kung, G. Goldstein, R. H. Levey, S. F. Schlossman, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1588 (1980); B. J. Fowlkes and D. M. Pardoll, *Adv. Immunol.* **44**, 207 (1989); A. Kruisbeek, *Curr. Opin. Immunol.* **5**, 227 (1993).
 52. S. H. Chan, D. Cosgrove, C. Waltzinger, C. Benoist, D. Mathis, *Cell* **73**, 225 (1993); S. H. Chan, C. Benoist, D. Mathis, *Immunol. Rev.* **135**, 119 (1993); C. B. Davis, N. Kileen, M. E. Casey Crooks, D. Raulet, D. R. Littman, *Cell* **73**, 237 (1993); H. S. Teh *et al.*, *Nature* **335**, 229 (1988); J. P. M. van Meerwijk and R. N. Germain, *Science* **261**, 911 (1993); J. P. M. van Meerwijk, E. M. O'Connell, R. N. Germain, *Semin. Immunol.* **6**, 231 (1994).
 53. H. Kozono, J. White, J. Clements, P. Marrack, J. W. Kappler, *Nature* **369**, 151 (1994).
 54. H.-C. Chang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11408 (1994); J. Liu *et al.*, *J. Biol. Chem.* **271**, 33639 (1996).
 55. T. A. Jones, J.-Y. Zou, S. W. Cowtan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
 56. A. Nicholls, K. A. Sharp, B. Honig, *Proteins Struct. Funct. Genet.* **11**, 281 (1991).
 57. Production and purification of scD10 TCR: The scD10 TCR, constructed by polymerase chain reaction, consists of 237 residues and is organized from NH₂- to COOH-terminus as follows: V β 8.2 (residues 3 to 110)-linker (GSADDAKKDAKKDG)-V α 2 (AV225) (residues 1 to 112) with a Cys²³⁵→Ser mutation. This linker was modified from that previously used for NMR studies (23) because the longer linker failed to give rise to I-A^k cocrystals of diffraction quality. For bacterial expression, we used the T7 promoter expression vector pET-11a. An overnight preculture of transformant (20 ml) was inoculated into 1 liter of fresh Luria broth supplemented with carbenicillin (50 μ g/ml) at 37°C and grown until the optical density at 600 nm reached 0.6. Isopropyl- β -D-thiogalactopyranoside was added to a final concentration of 1 mM, and the culture was incubated for a further 4 hours. Cells were harvested by centrifugation and lysed by sonication. The inclusion bodies were washed and dissolved in 100 mM tris-HCl (pH 8.0) containing 6 M guanidine chloride, 10 mM EDTA, and 10 mM dithiothreitol. An efficient refolding was achieved by diluting rapidly into a refolding buffer [50 mM tris-HCl (pH 8.0), 400 mM arginine, 2 M urea, 2 mM EDTA, 4 mM reduced glutathione, and 0.4 mM oxidized glutathione]. The refolded material was then applied to a 3D3 affinity column [5 mg/ml mAb coupled to Gamma Bind Plus Sepharose (Pharmacia, Uppsala, Sweden)] followed by gel filtration on Superdex 75 and exchanged to crystallization buffer [20 mM sodium acetate (pH 5.0) with 0.025% sodium azide].
 58. Production and purification of MHC class II I-A^k: A 13-residue hen egg CA peptide (residues 134 to 146) that is recognized by D10 TCR was fused (53) to the NH₂-terminus of the mature I-A^k β chain through a flexible linker. The 37-residue leucine zipper (LZ) sequences (54) were attached to both the α and β chains, with ACID-p1 to the β chain and BASE-p1 to the α chain through flexible thrombin-cleavable linkers. The cDNA constructions were subcloned into the pEE14 vector and expressed in Lec3.2.8.1 Chinese hamster ovary (CHO) cells. The screening of secreted recombinant protein in the culture supernatant was performed by both Sandwich enzyme-linked immunosorbent assay and BIAcore with antibodies specific for I-A^k (10.2.16) and the LZ epitope (2H11 or 13A12). The yield was ~0.7 mg of I-A^k/l supernatant. The production supernatant was applied to 2H11 affinity column [5 mg/ml mAb coupled to Gamma Bind Plus Sepharose (Pharmacia, Uppsala, Sweden)] and I-A^k protein was eluted by 50 mM citrate, 20 mM tris, 0.5 M NaCl, and 10% glycerol (pH 4.0). The eluted protein was then exchanged to 50 mM tris-HCl (pH 8.0) and cleaved by thrombin (2 U per 50 μ g of I-A^k) at 4°C for 4 hours. Thrombin was then removed by passage through benzamide Sepharose 6B beads and gel filtration on Superdex 75. Subsequently, the purified I-A^k was exchanged to 10 mM Hepes (pH 7.0), 0.025% sodium azide for crystallization.
 59. Z. Ortwinowski and W. Minor, *Methods Enzymol.* **276**, 307 (1997).
 60. J. Navaza, *Acta Crystallogr.* **A50**, 157 (1994).
 61. CA/I-A^k was crystallized by the hanging-droplet vapor diffusion method by equilibrating protein solution (7.2 mg/ml) against a buffer of 16% isopropanol, 18% polyethylene glycol (PEG) 4000, 0.1 M sodium citrate (pH 5.6) in reservoir. The crystals belong to space group P2₁2₁2₁ with unit cell parameters a = 99.1 Å, b = 122.4 Å, and c = 68.4 Å. A single crystal was transferred to the cryoprotectant solution of 30% glycerol, 18% PEG 4000, 0.1 M sodium citrate for half an hour before dipping into liquid nitrogen for freezing. A 95% complete data set was collected at Brookhaven National Synchrotron Light Source X12C beamline by using wavelength 1.072 Å with Brandeis charge-coupled device (CCD) detector. The data were integrated and scaled by DENZO and SCALPACK (55). The R_{merge} of the data was 9.2% to 3 Å resolution for 19,854 unique reflections. The structure was solved by molecular replacement method with AMoRe (56), by using a structure of I-A^k complexed with a peptide from hen egg lysozyme residues 50 to 62 [Protein Data Bank (PDB) code 1iak] as a search model. The structure was refined with X-PLOR. The final R_{free} is 29.5%, and R_{work} is 23.1%. Details will be published elsewhere.
 62. A. T. Brünger, Yale Univ. Press, New Haven, CT (1992).
 63. W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
 64. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).
 65. Supported by NIH grant AI19807 (E.L.R.), GM56008 (J.-H.W.), AI/CA37581 (G.W.), and U.S. Department of Energy, Office of Health and Environmental Research, contract W-31-109-Eng-38 (A.J. and R.Z.). We thank B. Benacerraf and M. Eck for careful review of the manuscript. We thank C. Janeway for providing the D10 TCR α and β cDNA clones and the 3D3 hybridoma. The coordinates of the D10 TCR-pMHCII complex have been deposited in the PDB (accession number 1D9K).

21 September 1999; accepted 26 October 1999

Predicting the Evolution of Human Influenza A

Robin M. Bush,^{1*} Catherine A. Bender,² Kanta Subbarao,² Nancy J. Cox,² Walter M. Fitch¹

Eighteen codons in the HA1 domain of the hemagglutinin genes of human influenza A subtype H3 appear to be under positive selection to change the amino acid they encode. Retrospective tests show that viral lineages undergoing the greatest number of mutations in the positively selected codons were the progenitors of future H3 lineages in 9 of 11 recent influenza seasons. Codons under positive selection were associated with antibody combining site A or B or the sialic acid receptor binding site. However, not all codons in these sites had predictive value. Monitoring new H3 isolates for additional changes in positively selected codons might help identify the most fit extant viral strains that arise during antigenic drift.

Antigenic drift due to mutations in the hemagglutinin gene necessitates frequent replacement of influenza A strains in the human vaccine. Antibodies against hemagglutinin are a primary determinant of susceptibility to

infection. However, the effects on antigenicity of specific mutations in the hemagglutinin gene are not well understood because multiple amino acid changes often occur together in important antigenic variants. Changes in antigenicity depend not only on the nature and position of the amino acid replacements but also on the amino acids currently encoded at other key positions in the HA1. We have developed a method for predicting the evolution of the virus that makes use of genetic data in the absence of specific knowledge

¹Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA. ²Influenza Branch, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA.

*To whom correspondence should be addressed. E-mail: rmbush@uci.edu