# Structural mechanism underpinning peptide cross-reactivity of a human telomerase reverse transcriptase reactive CD8+ T-cell clone

**David K. Cole[1], Hugo A. van den Berg[2], Angharad Lloyd[1], Michael D. Crowther[1], Konrad Beck[3], Julia Ekeruche-Makinde[1], John J. Miles[1,4], Anna M. Bulek[1], Garry Dolton[1], Andrea J. Schauenburg[1], Aaron Wall[1], Anna Fuller[1], Mathew Clement[1], Bruno Laugel[1], Pierre J. Rizkallah[1], Linda Wooldridge[5*] and Andrew K. Sewell[1*]**

[1]Division of Infection and Immunity and Systems Immunity Research Institute, Cardiff University School of Medicine, Heath Park, Cardiff, CF14 4XN, UK; [2]Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK. [3]Cardiff University School of Dentistry, Heath Park, Cardiff, CF14 4XN, UK. [4]QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia, 4029. [5]Faculty of Health Sciences, University of Bristol, Bristol BS8 1TD, UK.
[*]These authors contributed equally to this study

Running title: Structural basis of T-cell cross-reactivity

**Correspondence:** Dr David K. Cole, Cardiff University School of Medicine, UK, E-mail: coledk@cf.ac.uk, or Prof Andrew Sewell, Cardiff University School of Medicine, UK, E-mail: sewellak@cf.ac.uk.

**Keywords:** $CD8^+$ T-cells, T-cell receptor (TCR), peptide-major histocompatibility complex (pMHC), T-cell degeneracy, telomerase, surface plasmon resonance, X-ray crystallography.

## ABSTRACT

**T-cell cross-reactivity is essential for effective immune surveillance, but has also been implicated as a pathway to autoimmunity. Previous studies have demonstrated that T-cell receptors (TCRs) that focus on a minimal motif within the peptide are able to facilitate a high level of T-cell cross-reactivity. However, the structural database shows that most TCRs do not utilize this mode of antigen binding. To further explore the structural features that allow the clonally expressed TCR to functionally engage with multiple peptide-major histocompatibility complexes (pMHCs), we examined the ILA1 $CD8^+$ T-cell clone that responds to a peptide sequence derived from human telomerase reverse transcriptase (hTERT). The ILA1 TCR contacted its pMHC with a broad peptide-binding footprint encompassing spatially distant peptide residues. Despite the lack of a focused TCR-peptide binding mode, the ILA1 T-cell clone was still cross-reactive. Overall, the mode of binding correlated well with the level of degeneracy at different peptide positions. Thus, the ILA1 TCR was less tolerant of changes at peptide residues that were at, or adjacent to, key contact sites. This study provides new insights into the molecular mechanisms that control T-cell cross-reactivity, with important implications for pathogen surveillance, autoimmunity and transplant rejection.**

Recognition of peptide-major histocompatibility complexes (pMHCs) by the clonally expressed αβ T-cell antigen receptor (TCR) mediates T-cell immunity. Although TCRs generally interact with pMHC *via* a conserved binding mode, with the TCRα chain positioned over the MHCα1 domain and the TCRβ chain over the MHCα2 domain, the TCR complementarity determining region (CDR) loops can use a variety of mechanisms to probe both the MHC surface and bound peptide (1). This flexible binding probably mediates the ability of a single TCR to interact productively with a large range of different epitopes (2–6). Thus, TCR degeneracy enables the circa 25 million distinct TCR clonotypes expressed by an individual host (7) to have the potential to recognise the entire theoretical peptide universe that could be presented by MHC (2, 8), minimising the likelihood of pathogens escaping immune surveillance. Given the highly diverse number of TCR-pMHC binding modes seen to

date, it is reasonable to predict that different TCRs will exhibit distinct levels of cross-reactivity, depending on the chemical characteristics of their CDR loops and how they interact with pMHC. Such distinctions could determine whether certain TCRs are more likely to offer sufficient protection against hyper-variable pathogens, such as HIV-1, HBV, HCV and influenza, or, conversely, to trigger autoimmune disease.

Although new quantitative information on the extent of T-cell cross-reactivity has recently come to light (3, 5, 9), the molecular rules that determine this important facet of cellular adaptive immunity remain unclear. Understanding TCR binding degeneracy and the ensuing T-cell cross-reactivity it enables is of emerging importance given the increasing use of T-cell therapies using modified TCRs, one of which has already demonstrated the dangers of unintentional T-cell cross-reactivity with self-ligands (10–12). Currently, there are few examples of TCR-pMHC complex structures for which the cross-reactivity profiles of the corresponding T-cell clone have also been determined (3, 13, 14). In this study, we demonstrated that an insulin-reactive human $CD8^+$ T-cell clone (1E6) could recognize upwards of 1 million unique peptide ligands. The structure of the 1E6 TCR with its cognate ligand revealed a focused TCR-peptide binding mode, with the interaction of only two TCR residues and two adjacent peptide residues accounting for the majority of the binding interface. We speculated that this focussed binding might enable the 1E6 TCR to tolerate changes outside of the core motif, mediating the high level of degeneracy. In support of this, another study recently demonstrated that a high level of cross-reactivity was mediated by a similar focused TCR-peptide binding mode by a MHC class II restricted TCR (13). However, whether TCRs must exhibit a focussed peptide mode in order to cross-react remains unclear. This is an important question because, unlike the two examples of focused TCR-peptide binding mentioned above, most TCRs that have been studied structurally to date make more comprehensive interactions with the pMHC surface. Thus, whether a 'typical' TCR binding mode can underpin T-cell cross-reactivity remains unknown.

Here, we used a well-characterized $CD8^+$ T-cell clone (ILA1) (15), that responds to residues 540 to 548 (sequence ILAKFLHWL) of hTERT to further investigate the structural basis of TCR degeneracy. We have previously characterized a limited number of altered peptide ligands (APLs) for the ILA1 T-cell clone that exhibit different potencies in terms of T-cell activation (9, 15), corresponding to a wide range of binding affinities with the ILA1 TCR (15–17). These previous findings clearly demonstrate that the ILA1 T-cell clone can recognize multiple different peptide ligands. Here, we solved the structure of the ILA1 TCR in complex with HLA-A*0201-ILAKFLHWL (A2-ILA) and several previously defined APLs. Combined with biophysical analysis, we demonstrate the molecular mechanism for antigen discrimination by the ILA1 TCR, and model the mode of cross-reactivity with these APLs. In addition, we used our previously published peptide sampling approach (3) to estimate the number of pMHCI molecules that could be recognized by the ILA1 TCR. These data offer novel insight into the molecular factors that determine T-cell cross-reactivity, extending our understanding of the nature of T-cell antigen discrimination.

## RESULTS

*The ILA1 TCR makes a broad contact network with A2-ILA* - We solved the structure of the ILA1 TCR-A2-ILA complex at 2.8 Å in space group P 1 21 1, with crystallographic $R_{work}/R_{free}$ ratios within accepted limits as shown by the theoretically expected distribution (18) (**Table 1**). The electron density was high quality throughout, represented by an omit map analysis of the ILA peptide (**Figure 1A**). The ILA1 TCR employed a canonical binding mode to engage A2-ILA (**Figure 1A**) with a buried surface area (2507.2$Å^2$) and surface complementarity (TCR-pMHC = 0.641) within the normal range (**Table 2**) (1). The TCR α-chain was orientated over the MHCI α1-helices and the TCR β-chain over the MHCI α2 helices (**Figure 1B**), positioning the CDR loops of the ILA1 TCR over the central portion of the peptide, enabling contacts with 4 of the 9 peptide residues (**Figure 1C**). Peptide residues Lys4 and Trp8 engaged in a complex network of contacts with the TCR CDR1/3α loops and CDR1/3β loops respectively, contributing 39 out of the 49 total peptide contacts (**Table 2**). Binding to Lys4 involved a tight ball-and-socket interaction with 7 TCR residues, whereas contacts with Trp8 were less restrictive, involving only TCRβ-chain residues Glu30 and Gln96 (**Table 3, Figure 1C**). The 49 TCR interactions with peptide were supported

by 58 contacts with MHC involving 12 TCR α-chain residues and 4 TCR β-chain residues (**Figure 1D**), contributing to the slightly TCR α-chain skewed binding mode (α- 58%, β- 42% of total contacts). Notably, the TCR α-chain residue Arg68 in the framework region loop (outside of the traditional CDR loops) formed two salt-bridges with MHC residue Glu166, and all of the MHC restriction triad residues (Arg65, Ala69 and Gln155) (19) interacted with the ILA1 TCR (**Table 3**).

We have previously shown that the ILA1 TCR binds with a moderate/weak affinity ($K_D$ = 34μM) to A2-ILA (20), consistent with the observation that TCRs specific for self pMHCIs usually bind at the lower end of the TCR-pMHC affinity scale (21, 22). We performed a thermodynamic analysis of the ILA1 TCR-A2-ILA interaction by measuring binding using SPR at a range of temperatures (5-37°C) (**Figure 2A**). These analyses demonstrated that the weak affinity was not temperature-dependent, ranging from $K_D$ = 34 μM (25°C) to 74 μM (5°C). At physiological temperature (37°C), the affinity was in the middle of this range ($K_D$ = 48 μM). The energetic analysis (**Figure 2B**) revealed that the ILA1 TCR-A2-ILA interaction was driven entropically (TΔS = 5.86 kcal/mol), with only a minor change in enthalpy (ΔH = –0.16 kcal/mol). These values indicate almost no net loss, or gain, in electrostatic interactions during complex formation, indicative of structural re-ordering of the TCR and/or pMHC when binding. The entropic contribution suggested that ordered water molecules are squeezed out at the interface as the TCR and pMHC engage. Overall, these analyses demonstrated that the ILA1 TCR utilizes a relatively broad binding mode, contacting spatially distant regions on the peptide and involving 21 different TCR residues contacting 17 A2-ILA residues (4 peptide, 13 MHC) contributing to an entropically driven, moderate-to-weak affinity interaction.

*Altered peptide ligands (APLs) guide ILA1 antigen recognition through structural alterations in peptide conformation* - We have previously characterized a number of APLs that alter the T-cell activation profile and TCR binding affinity of the ILA1 T-cell clone (9, 15, 16). To investigate how these ligands adjust TCR interactions to tune affinity, we solved the structure of four APLs (A2-ILA3G8R: IL**G**KFLH**R**L, A2-ILA3G: IL**G**KFLHWL, A2-ILA8T: ILAKFLH**T**L, A2-ILA8E: ILAKFLH**E**L), included our previously published APL structure (A2-ILA8R: ILAKFLH**R**L) (23), and used the ILA1-A2-ILA complex as a model. The electron density was high quality throughout, represented by an omit map analysis of the ILA peptide variants (**Figure 3A-D**) and B-factor analysis indicated that there were no major differences in peptide mobility across the peptide variants (**Figure 3E-J**). Thermal stability analysis demonstrated that most of the APLs had a similar apparent Tm value (the term 'apparent Tm' is used here because the protein irreversibly aggregates at high temperature) of around 55°C, with extremes in the range of 50°C to 61°C (**Figure 4**). These similar stabilities are consistent with our previous observation that all these APLs bind equally to HLA-A2 on the cell surface (15). As we have shown previously in other systems (24), apparent Tm values correlated poorly with antigen potency (for instance A2-ILA3G had the lowest apparent Tm value, but was a potent activator of the ILA1 T-cell clone), suggesting that different pMHC cell surface expression levels were a minor factor in T-cell recognition. The A2-ILA3G structure was determined at 2.7 Å resolution and the other APL structures were determined at resolutions between at 1.9 Å-1.8 Å, with crystallographic $R_{work}/R_{free}$ ratios within accepted limits as shown by the theoretically expected distribution (18) (**Table 1**). The overall conformation of A2-ILA3G8R, A2-ILA3G and A2-ILA8R were virtually identical to A2-ILA (**Figure 5A-D**), with Lys4, Leu6, and Trp8 pointing up and away from the MHC binding groove and Leu2, Phe5, His7, and Leu9 acting as primary and secondary anchors, indicating that a molecular mimicry mechanism underpins ILA1 TCR recognition of these APLs. In contrast, in the A2-ILA8T and A2-ILA8E structures, peptide residues 5-7 were flipped so that Leu6 acted as a secondary anchor and Phe5 and His7 were in more solvent-exposed positions (**Figure 5E-F**). In both peptide variants, the mutated residue was at position 8, distal from this structural rearrangement. Closer inspection of the structures did not reveal an obvious mechanism for this indirect effect on peptide conformation.

We next explored the binding affinity of the ILA1 TCR for the APLs included in this study using previously published (9, 15, 16) and new data (**Figure 6A-E**). Despite the relatively weak affinity between the ILA1 TCR and the natural A2-ILA ligand ($K_D$ = 34 μM), the ILA1 TCR could recognize A2-ILA3G8R and A2-

ILA3G with anti-viral-like affinities ($K_D$ = 1 μM and 3.7 μM respectively). Both of these ligands included substitution of peptide residue 3 from Ala to Gly, a substitution that was clearly indicated in our previously published unbiased combinatorial peptide library (CPL) screening using the ILA1 T-cell clone (9). Structural modelling of the ILA1 TCR with these ligands indicated that interaction with the N-terminal portion of the peptide was likely to be very similar for both ligands (**Figure 6F&G**). However, for A2-ILA3G8R, a major re-orientation of TCR β-chain residue Gln96 would be required in order to tolerate Arg at position 8 in the peptide. For both of these ligands, the extra flexibility, afforded at the N-terminus of the peptide by the substitution of Gly compared to Ala, may enable the ILA1 TCR to establish enhanced contacts with peptide residue Lys4. This represents a likely mechanism for the stronger affinity, supported further by the observation that all of the other APLs that did not include Gly at position 3 were bound by ILA1 with weaker affinity compared to the ILA1-A2-ILA interaction (**Figure 6C-E**). Structural modelling of ILA1 in complex with A2-ILA8R (**Figure 6H**) revealed the same potential for steric hindrance between the TCR β-chain residues Glu30 and Gln96, but without the compensatory substitution at position 3. The structural rearrangement that would be required for ILA1 to bind to A2-ILA8R ($K_D$ = 151 μM) was reflected by a much weaker affinity compared to ILA1-A2-ILA ($K_D$ = 34 μM). Both A2-ILA8T and A2-ILA8E ligands underwent a conformational transition compared to the other APLs (**Figure 5**). Modelling demonstrated that this alteration could lead to a steric conflict between TCR β-chain residue Gln96 and the C-terminal residues of the peptide (**Figure 6I&J**). However, the position of peptide residue Phe5 in A2-ILA8T may allow for compensatory interactions with TCR α-chain residues Asp91 and Tyr98, not present with A2-ILA8E. Furthermore, the smaller Thr8 side chain in A2-ILA8T, compared to Glu8 in A2-ILA-8E, would require smaller modifications in TCR docking. Taken together, these structural distinctions may explain the extremely weak affinity observed for ILA1 binding to A2-ILA8E compared to A2-ILA8T ($K_D$ >500 mM and 28 μM, respectively).

*Quantification of ILA1 TCR degeneracy* - Our previous investigations have demonstrated that even single residue substitutions outside of the two main peptide interaction zones (Lys4 and Trp8) could have a substantial impact on the ILA1 TCR-A2-ILA complex, reflected by the different binding affinities and antigen potencies shown here and published previously (9, 15, 16). This binding mode is in stark contrast to the focused TCR-peptide binding mode of the highly cross-reactive auto-reactive T-cell clone (1E6) (3, 6, 14), known to mediate β-cell destruction in patients with type 1 diabetes (25, 26). We next generated a degeneracy curve for the ILA1 TCR using our previously described approach that quantifies TCR cross-reactivity (3).

CPL scan data were used to design four different motif-restricted peptide sets (I: xLGxxxxRL (total set size $19^5$); II: xLxKFLxxL (total set size $19^4$); III: xLG{K/L}F{L/I}{M/F/Y/N/H}{R/T/Y/K/S/F/H/I/L/M/Q/V/G/N}{L/V} (total set size 10640); and, IV: {A/I/L/M/P/Q/W}L{G/A}{K/L}F{L/I}{N/H}{F/K/N/Q/T/Y}{L/V} (total set size 1344); where x denotes any of the 19 proteogenic amino acids excluding cysteine). Between 19 and 30 peptides were sampled at random from each of these motif-restricted peptides cohorts. In addition, we performed importance sampling where 20 peptides were sampled from an effective sample size of $1.5 \times 10^7$. The $pEC_{50}$ for each peptide was estimated by simultaneous curve fitting (**Figures 7**) and these values were used to construct a degeneracy curve for the ILA1 TCR (**Figure 8A**). These analyses indicated that ILA1 could recognise ~$2 \times 10^3$ peptides with a functional sensitivity at least as high as 1/10th the functional sensitivity of the optimal agonist. At two orders of magnitude from the optimum (i.e. peptides ranging from1/$100^{th}$ of the optimal agonist to the optimal agonist) ~$4 \times 10^4$ peptides could be recognised by the ILA1 T-cell clone.

This analysis suggests that the ILA1 TCR can cross-react with a diverse peptide universe. Although smaller than the estimated number of peptides recognised by the 1E6 T-cell clone (~$10^6$), it should be noted that 1E6 recognises a 10-mer peptide, whereas ILA1 recognises a 9-mer. Thus, the peptide universe under consideration for 1E6 is 20 times larger than that for ILA1. Although it is unknown how this difference in peptide length affects the comparison between the two degeneracy profiles, this difference may partly explain why ILA1 appears to be less cross-reactive than 1E6. This analysis also demonstrated a different pattern of cross-reactivity between the 1E6 and

ILA1 T-cell clones that was consistent with the respective binding modes of their TCRs. The 1E6 T-cell clone could recognise a large number of sub-libraries (46 in total) outside of the central binding zone (residues 4-6) (**Figure 8B**). In contrast, the ILA1 T-cell clone was generally more sensitive across the peptide backbone, reflecting a more globally coordinated interaction between the ILA1 TCR and the antigenic peptide (**Figure 8C**). Taken together, these results are broadly consistent with the idea that different modes of TCR binding (i.e. TCRs that focus on a minimal peptide motif compared to TCRs that make contacts across the peptide backbone) can enable T-cell cross-reactivity, adding further support to the general idea that T-cells must be cross-reactive to fully protect us against a highly variable pathogen universe (reviewed in (8)).

## DISCUSSION

In order to mount effective immune responses in the face of a diverse antigenic milieu, using a limited set of TCRs (estimated ~25 million distinct clonotypes in an individual), each T-cell must be able to interact productively with a vast array of different antigens (2, 8). Indeed, recent experimental evidence supports this notion, including our own study demonstrating that a single T-cell clone can recognize over 1 million different peptides at physiologically relevant concentrations (3). However, structural investigations of TCR-pMHC interactions have demonstrated that the TCR can establish a distinct and highly specific interaction with both peptide and MHC. In keeping with such specificity of binding requirements, even single mutations at key residues in the TCR, peptide, or MHC that are involved in the binding interface have been shown to abrogate antigen recognition (27–31). The importance of T-cell cross-reactivity is manifest not only in effective immune surveillance (2, 8), but also in auto-reactivity (32) and the design of therapeutics (10, 11). Structural rules must therefore exist that allow cross-reactivity to take place; here, we report early steps toward deconstructing these rules.

We investigated a TCR isolated from ILA1, a well-characterized HLA-A*0201-restricted, telomerase-specific, $CD8^+$ T-cell clone. Our previous work has shown that ILA1 cross-reacts with an array of APLs with different potencies, tuned by the CD8 co-receptor, and that antigen 'potency' generally correlates directly with the affinity of TCR binding (9, 15). Here, we solved the complex structure of the ILA1 TCR with the natural index ligand (A2-ILA) and used the structures of a number of unligated APLs to model the mode of APL recognition and determine the structural basis for ILA1 cross-reactivity.

We have previously demonstrated that the ILA1 T-cell clone is particularly sensitive to APLs with modifications at peptide residues 3 and 8. The unligated structures of five APLs with alterations in these positions demonstrated that the overall conformation of the Cα peptide backbone could be altered by introduction of Thr or Glu at position 8, possibly explaining the weaker binding affinity between ILA1 and A2-ILA8R and A2-ILA8E. In contrast, substitution of Ala to Gly at peptide residue 3 enhanced recognition, and Gly at this position was strongly recognised in combinatorial peptide library screens (9). Our structural analysis indicated that the surmised extra flexibility afforded to the N-terminus of the peptide mediated by substitution at position 3 with Gly would likely enable more favourable interactions with Lys4, which made a network of contacts with the ILA1 TCR through a 'ball-and-socket'-like interaction. In fact, substitution at position 3 with Gly could override the negative impact of modifications to peptide residue 8, revealed by the enhanced binding affinity of the ILA1-A2-ILA3G8R interaction ($K_D$ = 1 μM) compared to the index peptide ($K_D$ = 34 μM) and ILA8R ($K_D$ = 151 μM). Our structural analysis demonstrated that, again, even single peptide substitutions outside of the main interaction interface could have a substantial impact on TCR binding affinity and T-cell antigen potency, consistent with our previous findings (9, 15, 16). These observations add further evidence to our recent findings (19, 33–36) that peptide presentation by MHC-I can be dynamic and difficult to predict.

Recent reports have demonstrated that TCRs using a focused TCR-peptide binding mode can be highly cross-reactive (37, 38). However, this binding mode is not representative of most TCRs described in the literature. Indeed, on average TCR-peptide binding is spread out over ~60% of the peptide backbone for MHCI-restricted epitopes, often including contacts with both the N- and C-terminal regions of the peptide (1, 39). This interconnected binding network between the TCR and the peptide may not allow a high degree of cross-reactivity because most peptide modifications could

impact on binding. Unlike the focused TCR-peptide binding mode utilised by the 1E6 (37) and 42F3 (40) TCRs, the ILA1 TCR employed a more representative binding mode to A2-ILA. This was reflected by a larger BSA value (2540Å for ILA1 compared to 1640Å for 1E6) and a binding motif that included contacts spread out over peptide residues 4-8. Thus, we explored the consequences of the ILA1 TCR binding mode on ILA1 T-cell cross-reactivity using our previously published methodology (3). Despite the broader binding mode utilised by the ILA1 TCR, the ILA1 T-cell clone was still able to recognise ~4 × $10^4$ peptides with equal or greater sensitivity compared to the index peptide, and many more at lower potency. These data suggest that although the mode of TCR binding is very likely to tune T-cell cross-reactivity to some degree, the ability of T-cells to recognise a vast array of different peptides is likely to be commonplace.

In summary, we demonstrate that the interaction between a TCR from a human $CD8^+$ T-cell clone that recognises a peptide sequence from an important tumour antigen contacts the peptide at spatially distant sites along the peptide backbone. The affinity of this TCR can be tuned by various peptide modifications through both direct and indirect effects, demonstrating the dynamic nature of the interaction between TCR, peptide and MHC. Even though the ILA1 T-cell clone was sensitive to modifications along the peptide backbone, consistent with its broad binding interface, it was still able to cross-react with a vast array of different peptides. These data demonstrate that focused TCR-peptide binding is not a requirement for T-cell degeneracy. Indeed, a broader binding mode, as observed in most TCR-pMHC structures reported to date, is also likely to facilitate T-cell cross-reactivity. These results have important implications for immune surveillance, i.e. how a limited set of TCRs can recognize all potential antigens variants, and the complex mechanisms that may lead to auto-reactivity mediated by molecular mimicry.

## EXPERIMENTAL PROCEEDURES

*T-Cells and target cells* - The ILA1 $CD8^+$ T-cell clone is specific for the HLA A*0201 restricted human telomerase reverse transcriptase-derived epitope ILAKFLHWL (residues 540–548) (41), and the 1E6 T-cell clone is specific for the human leukocyte antigen HLA A*0201 restricted autoantigen PPI (preproinsulin) epitope ALWGPDPAAA (residues 15 - 24) (25). $CD8^+$ T-cell clones were maintained in RPMI 1640 (Life Technologies, Rockville, MD) containing 100 U/ml penicillin (Life Technologies), 100 mg/ml streptomycin (Life Technologies), 2 mM L-glutamine (Life Technologies), and 10% heat inactivated FCS (Life Technologies) (R10) supplemented with 2.5% Cellkines (Helvetica Healthcare, Geneva, Switzerland), 200IU/ml IL-2 (PeproTech, Rocky Hill, NJ) and 25 ng/ml IL-15 (PeproTech). Hmy.2 C1R B-cells expressing full-length HLA A*0201 were generated as described previously (42).

*Protein expression, refolding and purification* - The ILA1 TCR, HLA-A*0201 α chain and human β2m chain sequences were generated as previously described (20), and cloned into separate pGMT7 expression plasmids under the control of the T7 promoter. The ILA1 TCR and HLA-A*0201 in complex with various different peptide variants (as indicated) were refolded and purified as described previously (14). Biotinylated pMHCI was prepared as previously described (43).

*pMHC stability assays* - Thermal stability of the HLA-A*0201-peptid complexes was assessed by circular dichroism spectroscopy monitoring the change in ellipticities at 218 nm upon heating as described (33). Briefly, samples were prepared in PBS at a concentration of ~3μM.and measured in 0.1-cm quartz cells. Melting curves were analyzed assuming a two-state trimer-to-monomer transition from the native to unfolded conformation and fitted as described (44). As all protein complexes aggregated to various degrees upon unfolding, the ellipticity of the unfolded state was set as a constant of -1.35 $M^{-1}$ $cm^{-1}$ (45).

*Surface Plasmon Resonance Analysis* - Binding analysis was performed in duplicate using a BIAcoreT200™ equipped with a CM5 sensor chip as previously described (46). Approximately 200-500 RU of HLA-A*0201-ILGKFLHRL or HLA-A*0201-ILAKFLHRLpeptide complex was attached to the CM5 sensor chip at a slow flow-rate of 10 μL/min to ensure uniform distribution on the chip surface. HLA-A*0201-ILAKFLHWL was used as a positive control as the binding affinity with the ILA1 TCR has been published previously (15, 20). The ILA1 TCR was purified and concentrated to ~300 μM on the same day of surface plasmon resonance (SPR) analysis. For equilibrium analysis, 10 serial dilutions were

prepared in duplicate for each sample and injected over the relevant sensor chips at 25°C. TCR was injected over the chip surface using kinetic injections at a flow rate of 45 µL/min using HLA-A*0201-ALWGPDPAAA as a negative control surface on flow cell 1. Results were analyzed using BIAevaluation 3.1, Excel and Origin 6.0 software. The equilibrium dissociation constant ($K_D$) values were calculated assuming a 1:1 interaction by plotting specific equilibrium-binding responses against protein concentrations followed by non-linear least squares fitting of the Langmuir binding equation. For kinetics analysis, the $k_{on}$ and $k_{off}$ values were calculated assuming 1:1 Langmuir binding and the data were analyzed using a global fit algorithm (BIAevaluation 3.1™).

*Crystal structure determination* - All protein crystals were grown at 18°C by vapour diffusion via the sitting drop technique. 200 nL of each pMHCI (10 mg/ml) in crystallization buffer (10 mM Tris pH 8.1 and 10 mM NaCl) was added to 200 nL of reservoir solution. ILA1-HLA-A*0201-ILAKFLHWL (ILA1-A2-ILA) and HLA-A*0201-ILAKFLHTL (A2-ILA8T) crystals were grown in 0.2 M ammonium sulphate, 0.1 M HEPES pH 7, 20% PEG 8000 (47); HLA-A*0201-ILGKFLHRL (A2-ILA3G8R) crystals were grown in 0.2 M ammonium sulphate, 0.1 M Tris pH 7.5, 25% PEG 8000 (47); HLA-A*0201-ILGKFLHWL (A2-ILA3G) crystals were grown in 0.2 M ammonium sulphate, 0.1 M MES pH 7, 15% PEG 8000; HLA-A*0201-ILAKFLHEL (A2-ILA8E) crystals were grown in 0.2 M ammonium sulphate, 0.1 M MES pH 7, 25% PEG 8000 (47). Crystallization screens were conducted using an Art-Robbins Phoenix dispensing robot (Alpha Biotech Ltd, UK) and data were collected at 100 K at the Diamond Light Source (DLS), Oxfordshire, UK at a wavelength of 0.98 Å using an ADSC Q315 CCD detector. Reflection intensities were estimated using XIA2 (48) and the data were analyzed with SCALA and the CCP4 package (49). Structures were solved with molecular replacement using PHASER (50). Sequences were adjusted with COOT (51) and the models refined with REFMAC5. Graphical representations were prepared with PYMOL (52). The reflection data and final model coordinates were deposited with the PDB database (ILA1-A2-ILA PDB: 5C60, A2-ILA3G8R PDB: 5C61, A2-ILA3G PDB: 5C62, A2-ILA8T PDB: 5C63, A2-ILA8E PDB: 5C64).

*$CD8^+$ T-cell effector function assays: MIP1β ELISA* - $6 \times 10^4$ C1R-A2 cells were incubated with peptide at various concentrations in duplicate for 2 hours at 37ºC. Subsequently, $3 \times 10^4$ ILA1 $CD8^+$ T-cells were added and the assay was incubated overnight at 37ºC. The supernatant was harvested and assayed for MIP1β by ELISA according to the manufacturer's instructions (R&D Systems). Functional sensitivity of individual peptides was expressed as the $pEC_{50}$ of each peptide, which is defined as $-1 \times$ the base 10 logarithm ($p$) of the 50% efficacy concentration (EC50).

*Quantification of ILA1 TCR degeneracy* - The degeneracy of the ILA1 TCR was estimated as described previously (3). Briefly, the degeneracy at ω, defined as the number of peptides whose functional sensitivity is at least as large as ω was estimated directly using importance sampling based on the combinatorial peptide library scan, and bounded below by sampling from motif-based subsets of the peptide universe. The degeneracy is reported by plotting this quantity as a function of ω, where the functional sensitivity ω was scaled relative to a clone-specific reference peptide (the "index").

**Acknowledgements**
The authors have declared that no competing interests exist. AKS is a Wellcome Trust Investigator. DKC is a Wellcome Trust Research Career Development Fellow (WT095767). This work was supported by the UK Biotechnology and Biological Sciences Research Council (Grant BB/ H001085/1) and a WT Intermediate Clinical Fellowship awarded to LW (WT079848MA). PJR was supported by a RCUK Fellowship. We thank the staff at Diamond Light Source for providing facilities and support.

**Conflict of interest**
The authors have no conflicts of interest or competing financial interests.

**Author Contributions**
H. A. B., A.L., M.D.C., K.B., J.E., J.J.M., A.M.B., G.D., A.J.S., A.W., A.F., M.C., B.L., P.J.R., L.W., and D.K.C., performed experiments and analysed the data. A.K.S., L.W., and D.K.C. wrote the manuscript. A.K.S., L.W., H. A. B., and D.K.C., conceived and directed the study. A.K.S., L.W., and D.K.C., funded the study. All authors contributed to discussions.

**REFERENCES**


1. Rossjohn, J., Gras, S., Miles, J. J., Turner, S. J., Godfrey, D. I., and McCluskey, J. (2015) T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200

2. Mason, D. (1998) A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol. Today*. **19**, 395–404

3. Wooldridge, L., Ekeruche-Makinde, J., Van Den Berg, H. A., Skowera, A., Miles, J. J., Tan, M. P., Dolton, G., Clement, M., Llewellyn-Lacey, S., Price, D. A., Peakman, M., and Sewell, A. K. (2012) A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177

4. Wilson, D. B., Wilson, D. H., Schroder, K., Pinilla, C., Blondelle, S., Houghten, R. A., and Garcia, K. C. (2004) Specificity and degeneracy of T cells. *Mol. Immunol.* **40**, 1047–1055

5. Birnbaum, M. E., Mendoza, J. L., Sethi, D. K., Dong, S., Glanville, J., Dobbins, J., ??zkan, E., Davis, M. M., Wucherpfennig, K. W., and Garcia, K. C. (2014) Deconstructing the peptide-MHC specificity of t cell recognition. *Cell*. **157**, 1073–1087

6. Cole, D. K., Bulek, A. M., Dolton, G., Schauenberg, A. J., Szomolay, B., Rittase, M., Trimby, A., Jothikumar, P., Fuller, A., Skowera, A., Rossjohn, J., Zhu, C., Miles, J. J., Peakman, M., Rizkallah, P. J., and Sewell, A. K. (2016) Hotspot autoimmune T-cell receptor binding to pathogen and insulin peptides e. *J. Clin. Invest.* 10.1172/JCI85679

7. Arstila, T. P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J., and Kourilsky, P. (1999) A direct estimate of the human alphabeta T cell receptor diversity. *Science*. **286**, 958–961

8. Sewell, A. K. (2012) Why must T cells be cross-reactive? *Nat. Publ. Gr.* **12**, 669–677

9. Wooldridge, L., Laugel, B., Ekeruche, J., Clement, M., van den Berg, H. A., Price, D. A., and Sewell, A. K. (2010) CD8 Controls T Cell Cross-Reactivity. *J. Immunol.* **185**, 4625–4632

10. Linette, G. P., Stadtmauer, E. A., Maus, M. V., Rapoport, A. P., Levine, B. L., Emery, L., Litzky, L., Bagg, A., Carreno, B. M., Cimino, P. J., Binder-Scholl, G. K., Smethurst, D. P., Gerry, A. B., Pumphrey, N. J., Bennett, A. D., Brewer, J. E., Dukes, J., Harper, J., Tayton-Martin, H. K., Jakobsen, B. K., Hassan, N. J., Kalos, M., and June, C. H. (2013) Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood*. **122**, 863–871

11. Cameron, B. J., Gerry, A. B., Dukes, J., Harper, J. V, Kannan, V., Bianchi, F. C., Grand, F., Brewer, J. E., Gupta, M., Plesa, G., Bossi, G., Vuidepot, A., Powlesland, A. S., Legg, A., Adams, K. J., Bennett, A. D., Pumphrey, N. J., Williams, D. D., Binder-Scholl, G., Kulikovskaya, I., Levine, B. L., Riley, J. L., Varela-Rohena, A., Stadtmauer, E. A., Rapoport, A. P., Linette, G. P., June, C. H., Hassan, N. J., Kalos, M., and Jakobsen, B. K. (2013) Identification of a Titin-derived HLA-A1-presented peptide as a cross-reactive target for engineered MAGE A3-directed T cells. *Sci. Transl. Med.* **5**, 197ra103

12. Raman, M. C. C., Rizkallah, P. J., Simmons, R., Donnellan, Z., Dukes, J., Bossi, G., Le Provost, G. S., Todorov, P., Baston, E., Hickman, E., Mahon, T., Hassan, N., Vuidepot, A., Sami, M., Cole, D. K., and Jakobsen, B. K. (2016) Direct molecular mimicry enables off-target cardiovascular toxicity by an enhanced affinity TCR designed for cancer immunotherapy. *Sci. Rep.* **6**, 18851

13. Adams, J. J., Narayanan, S., Birnbaum, M. E., Sidhu, S. S., Blevins, S. J., Gee, M. H., Sibener, L. V, Baker, B. M., Kranz, D. M., and Garcia, K. C. (2015) Structural interplay between germline interactions and adaptive recognition determines the bandwidth of TCR-peptide-MHC cross-reactivity. *Nat. Immunol.* **17**, 87–94

14. Bulek, A. M., Cole, D. K., Skowera, A., Dolton, G., Gras, S., Madura, F., Fuller, A., Miles, J. J., Gostick, E., Price, D. a, Drijfhout, J. W., Knight, R. R., Huang, G. C., Lissin, N., Molloy, P. E., Wooldridge, L., Jakobsen, B. K., Rossjohn, J., Peakman, M., Rizkallah, P. J., and Sewell, A. K. (2012) Structural basis for the killing of human beta cells by CD8(+) T cells in type 1 diabetes. *Nat. Immunol.* **13**, 283–289

15. Laugel, B., Van Den Berg, H. A., Gostick, E., Cole, D. K., Wooldridge, L., Boulter, J., Milicic, A., Price, D. A., and Sewell, A. K. (2007) Different T cell receptor affinity thresholds and CD8 coreceptor dependence govern cytotoxic T lymphocyte activation and tetramer binding properties. *J. Biol. Chem.* **282**, 23799–23810

16. Melenhorst, J. J., Scheinberg, P., Chattopadhyay, P. K., Lissina, A., Gostick, E., Cole, D. K., Wooldridge, L., van den Berg, H. A., Bornstein, E., Hensel, N. F., Douek, D. C., Roederer, M., Sewell, A. K., Barrett, A. J., and Price, D. A. (2008) Detection of low avidity CD8+ T cell populations with coreceptor-enhanced peptide-major histocompatibility complex class I tetramers. *J. Immunol. Methods*. **338**, 31–39

17. van den Berg, H. A., Ladell, K., Miners, K., Laugel, B., Llewellyn-Lacey, S., Clement, M., Cole, D. K., Gostick, E., Wooldridge, L., Sewell, A. K., Bridgeman, J. S., and Price, D. A. (2013) Cellular-level versus receptor-level response threshold hierarchies in T-cell activation. *Front. Immunol.* **4**, 250

18. Tickle, I. J., Laskowski, R. A., and Moss, D. S. (2000) R(free) and the R(free) ratio. II. Calculation of the expected values and variances of cross-validation statistics in macromolecular least-squares refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 442–450

19. Tynan, F. E., Burrows, S. R., Buckle, A. M., Clements, C. S., Borg, N. a, Miles, J. J., Beddoe, T., Whisstock, J. C., Wilce, M. C., Silins, S. L., Burrows, J. M., Kjer-Nielsen, L., Kostenko, L., Purcell, A. W., McCluskey, J., and Rossjohn, J. (2005) T cell receptor recognition of a "super-bulged" major histocompatibility complex class I-bound peptide. *Nat Immunol*. **6**, 1114–1122

20. Cole, D. K., Pumphrey, N. J., Boulter, J. M., Sami, M., Bell, J. I., Gostick, E., Price, D. A., Gao, G. F., Sewell, A. K., and Jakobsen, B. K. (2007) Human TCR-binding affinity is governed by MHC class restriction. *J. Immunol.* **178**, 5727–34

21. Bridgeman, J. S., Sewell, A. K., Miles, J. J., Price, D. A., and Cole, D. K. (2012) Structural and biophysical determinants of αβ T-cell antigen recognition. *Immunology*. **135**, 9–18

22. Aleksic, M., Liddy, N., Molloy, P. E., Pumphrey, N., Vuidepot, A., Chang, K.-M., and Jakobsen, B. K. (2012) Different affinity windows for virus and cancer-specific T-cell receptors: implications for therapeutic strategies. *Eur. J. Immunol.* **42**, 3174–9

23. Ekeruche-Makinde, J., Miles, J. J., Van Den Berg, H. A., Skowera, A., Cole, D. K., Dolton, G., Schauenburg, A. J. A., Tan, M. P., Pentier, J. M., Llewellyn-Lacey, S., Miles, K. M., Bulek, A. M., Clement, M., Williams, T., Trimby, A., Bailey, M., Rizkallah, P., Rossjohn, J., Peakman, M., Price, D. A., Burrows, S. R., Sewell, A. K., and Wooldridge, L. (2013) Peptide length determines the outcome of TCR/peptide-MHCI engagement. *Blood*. **121**, 1112–1123

24. Miles, K. M., Miles, J. J., Madura, F., Sewell, A. K., and Cole, D. K. (2011) Real time detection of peptide-MHC dissociation reveals that improvement of primary MHC-binding residues can have a minimal, or no, effect on stability. *Mol. Immunol.* **48**,

728–32

25. Skowera, A., Ellis, R. J., Varela-Calvi??o, R., Arif, S., Huang, G. C., Van-Krinks, C., Zaremba, A., Rackham, C., Allen, J. S., Tree, T. I. M., Zhao, M., Dayan, C. M., Sewell, A. K., Unger, W., Drijfhout, J. W., Ossendorp, F., Roep, B. O., and Peakman, M. (2008) CTLs are targeted to kill□?? cells in patients with type 1 diabetes through recognition of a glucose-regulated preproinsulin epitope. *J. Clin. Invest.* **118**, 3390–3402

26. Knight, R. R., Kronenberg, D., Zhao, M., Huang, G. C., Eichmann, M., Bulek, A., Wooldridge, L., Cole, D. K., Sewell, A. K., Peakman, M., and Skowera, A. (2013) Human□??-cell killing by autoreactive preproinsulin-specific CD8 T cells is predominantly granule-mediated with the potency dependent upon T-cell receptor avidity. *Diabetes*. **62**, 205–213

27. Cole, D. K., Miles, K. M., Madura, F., Holland, C. J., Schauenburg, A. J. A., Godkin, A. J., Bulek, A. M., Fuller, A., Akpovwa, H. J. E., Pymm, P. G., Liddy, N., Sami, M., Li, Y., Rizkallah, P. J., Jakobsen, B. K., and Sewell, A. K. (2014) T-cell Receptor (TCR)-peptide specificity overrides affinity-enhancing TCR-major histocompatibility complex interactions. *J. Biol. Chem.* **289**, 628–638

28. Borg, N. a, Ely, L. K., Beddoe, T., Macdonald, W. a, Reid, H. H., Clements, C. S., Purcell, A. W., Kjer-Nielsen, L., Miles, J. J., Burrows, S. R., McCluskey, J., and Rossjohn, J. (2005) The CDR3 regions of an immunodominant T cell receptor dictate the "energetic landscape" of peptide-MHC recognition. *Nat. Immunol.* **6**, 171–180

29. Madura, F., Rizkallah, P. J., Miles, K. M., Holland, C. J., Bulek, A. M., Fuller, A., Schauenburg, A. J. A., Miles, J. J., Liddy, N., Sami, M., Li, Y., Hossain, M., Baker, B. M., Jakobsen, B. K., Sewell, A. K., and Cole, D. K. (2013) T-cell receptor specificity maintained by altered thermodynamics. *J. Biol. Chem.* **288**, 18766–18775

30. Gras, S., Chen, Z., Miles, J. J., Liu, Y. C., Bell, M. J., Sullivan, L. C., Kjer-Nielsen, L., Brennan, R. M., Burrows, J. M., Neller, M. A., Khanna, R., Purcell, A. W., Brooks, A. G., McCluskey, J., Rossjohn, J., and Burrows, S. R. (2010) Allelic polymorphism in the T cell receptor and its impact on immune responses. *J. Exp. Med.* **207**, 1555–67

31. Burrows, S. R., Chen, Z., Archbold, J. K., Tynan, F. E., Beddoe, T., Kjer-Nielsen, L., Miles, J. J., Khanna, R., Moss, D. J., Liu, Y. C., Gras, S., Kostenko, L., Brennan, R. M., Clements, C. S., Brooks, A. G., Purcell, A. W., McCluskey, J., and Rossjohn, J. (2010) Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 10608–10613

32. Sethi, D. K., Gordo, S., Schubert, D. a, and Wucherpfennig, K. W. (2013) Crossreactivity of a human autoimmune TCR is dominated by a single TCR loop. *Nat. Commun.* **4**, 2623

33. Kløverpris, H. N., Cole, D. K., Fuller, A., Carlson, J., Beck, K., Schauenburg, A. J., Rizkallah, P. J., Buus, S., Sewell, A. K., and Goulder, P. (2015) A molecular switch in immunodominant HIV-1-specific CD8 T-cell epitopes shapes differential HLA-restricted escape. *Retrovirology*. **12**, 20

34. Madura, F., Rizkallah, P. J., Holland, C. J., Fuller, A., Bulek, A., Godkin, A. J., Schauenburg, A. J., Cole, D. K., and Sewell, A. K. (2015) Structural basis for ineffective T-cell responses to MHC anchor residue-improved "heteroclitic" peptides. *Eur. J. Immunol.* **45**, 584–91

35. Motozono, C., Pearson, J. A., De Leenheer, E., Rizkallah, P. J., Beck, K., Trimby, A., Sewell, A. K., Wong, F. S., and Cole, D. K. (2015) Distortion of the major histocompatibility complex class i binding groove to accommodate an insulin-derived

10-mer peptide. *J. Biol. Chem.* **290**, 18924–18933

36. Miles, J. J., Elhassen, D., Borg, N. A., Silins, S. L., Tynan, F. E., Burrows, J. M., Purcell, A. W., Kjer-Nielsen, L., Rossjohn, J., Burrows, S. R., and McCluskey, J. (2005) CTL Recognition of a Bulged Viral Peptide Involves Biased TCR Selection. *J. Immunol.* **175**, 3826–3834

37. Cole, D. K., Bulek, A. M., Dolton, G., Schauenberg, A. J., Szomolay, B., Rittase, W., Trimby, A., Jothikumar, P., Fuller, A., Skowera, A., Rossjohn, J., Zhu, C., Miles, J. J., Peakman, M., Wooldridge, L., Rizkallah, P. J., and Sewell, A. K. (2016) Hotspot autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. *J. Clin. Invest.* **126**, 2191–204

38. Degano, M., Garcia, K. C., Apostolopoulos, V., Rudolph, M. G., Teyton, L., and Wilson, I. a (2000) A Functional Hot Spot for Antigen Recognition in a Superagonist TCR/MHC Complex. *Immunity*. **12**, 251–261

39. Rudolph, M. G., Stanfield, R. L., and Wilson, I. A. (2006) How Tcrs Bind Mhcs, Peptides, and Coreceptors. *Annu. Rev. Immunol.* **24**, 419–466

40. Adams, J. J., Narayanan, S., Liu, B., Birnbaum, M. E., Kruse, A. C., Bowerman, N. A., Chen, W., Levin, A. M., Connolly, J. M., Zhu, C., Kranz, D. M., and Garcia, K. C. (2011) T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity*. **35**, 681–93

41. Laugel, B., Price, D. a, Milicic, A., and Sewell, A. K. (2007) CD8 exerts differential effects on the deployment of cytotoxic T lymphocyte effector functions. *Eur. J. Immunol.* **37**, 905–13

42. Purbhoo, M. A., Boulter, J. M., Price, D. A., Vuidepot, A. L., Hourigan, C. S., Rod Dunbar, P., Olson, K., Dawson, S. J., Phillips, R. E., Jakobsen, B. K., Bell, J. I., and Sewell, A. K. (2001) The Human CD8 Coreceptor Effects Cytotoxic T Cell Activation and Antigen Sensitivity Primarily by Mediating Complete Phosphorylation of the T Cell Receptor Chain. *J. Biol. Chem.* **276**, 32786–32792

43. Cole, D. K., Rizkallah, P. J., Boulter, J. M., Sami, M., Vuidepot, A. L., Glick, M., Gao, F., Bell, J. I., Jakobsen, B. K., and Gao, G. F. (2007) Computational design and crystal structure of an enhanced affinity mutant human CD8 coreceptor. *Proteins Struct. Funct. Genet.* **67**, 65–74

44. Greenfield, N. J. (2004) Analysis of Circular Dichroism Data. *Methods Enzymol.* **383**, 282–317

45. Venyaminov, S. Y., Baikalov, I. A., Shen, Z. M., Wu, C. S., and Yang, J. T. (1993) Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. *Anal. Biochem.* **214**, 17–24

46. Cole, D. K., Dunn, S. M., Sami, M., Boulter, J. M., Jakobsen, B. K., and Sewell, A. K. (2008) T cell receptor engagement of peptide-major histocompatibility complex class I does not modify CD8 binding. *Mol. Immunol.* **45**, 2700–2709

47. Bulek, A. M., Madura, F., Fuller, A., Holland, C. J., Schauenburg, A. J. A., Sewell, A. K., Rizkallah, P. J., and Cole, D. K. (2012) TCR/pMHC Optimized Protein crystallization Screen. *J. Immunol. Methods*. **382**, 203–210

48. Winter, G., Lobley, C. M. C., and Prince, S. M. (2013) Decision making in xia2. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1260–1273

49. Bailey, S., and Collaborative Computational Project, N. 4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763

50. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674

51. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132

52. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. *Schrödinger LLC wwwpymolorg*. **Version 1.**, http://www.pymol.org

**Figure Legends**

**Figure 1. The ILA1 TCR uses a broad binding mode to engage A2-ILA.** (**A**) The overall binding mode of ILA1 TCR (blue and cyan cartoon – CDR loops shown in multi-colored cartoon) in complex with A2-ILA (grey cartoon and orange sticks). The box shows the observed map (top) at 1.0 σ and an omit maps (below) in which the model was refined in the absence of the ILA peptide with difference density is contoured at 3.0 σ; positive contours are shown in green, and negative contours are red. (**B**) Position of the ILA1 TCR CDR loops (multi-coloured sticks) with the ILA peptide (orange sticks) is shown in the HLA-A*0201 binding groove (grey surface). The crossing angle of the ILA1 TCR (black line) was calculated using previously published parameters (39). Briefly, this crossing angle represents the angle between a best- fit straight line through the Cα atoms from the two MHC helices, and a line that links the disulphide bond in the TCR α chain variable region to the disulphide bond in the TCR β chain variable region. (**C**) Interaction between residues in the ILA1 TCR CDR loops (multi-colored sticks) and the ILA1 peptide (orange sticks) with the MHCα1 helix shown as grey cartoon. (**D**) The ILA1 TCR residues in the CDR loops that contact the MHC surface are shown in multicolored cartoon and surface with the MHC binding groove shown in grey cartoon and surface. CDR loops are colored as follows throughout: CDR1α – red, CDR2α – green, CDR3α – blue, CDR1β – yellow, CDR2β

**Figure 2. Thermodynamic analysis of the ILA1 TCR with A2-ILA.** (**A**) Eight serial dilutions of the ILA1 TCR were injected, in duplicate, over A2-ILA at 5°C, 13°C, 21°C, 25°C, 29°C, and 37°C. The equilibrium binding constants ($K_D$) were calculated using a nonlinear curve fit ($y= (P_1x)/P_2 + x)$) and kinetic and affinity parameters are shown in the table. (**B**) The binding free energies, ΔG° (ΔG° = $RTlnK_D$), were plotted against temperature (K) using non-linear regression to fit the three-parameters van 't Hoff equation, ($RT \ln K_D = \Delta H° - T\Delta S° + \Delta Cp°(T-T_0) - T\Delta Cp° \ln (T/T_0)$ with $T_0$=298 K).
– purple, andCDR3β – cyan.

**Figure 3. Density plot, omit map and B-factor analysis of ILA1 peptide variants.** Above: The observed map at 1.0 σ is shown. Below: omit maps are shown, in which the model was refined in the absence of the ILA peptide variants, with difference density contoured at 3.0 σ; positive contours are shown in green, and negative contours are red. (**A**) A2-ILGKFLHRL (purple sticks). (**B**) A2-ILGKFLHWL (green sticks). (**C**) A2-ILAKFLHTL (sand sticks). (**D**) A2-ILAKFLHEL (black sticks). A2-ILAKFLHRL was solved previously (23). (**E-J**) Each APL is coloured by B-factor, with light blue representing a low B-factor, and red representing a high B-factor. The conformation of each APL (sticks) with arrows indicating the direction of each residue in the peptide (solvent exposed, MHC anchor, or in between) with the MHCα1 helix shown as grey cartoon. Residues in red indicate differences from the index sequence. Up arrow indicates solvent exposed, down arrow indicates anchor position, no arrow indicates an intermediate position. (**E**) A2-ILAKFLHWL. (**F**) A2-ILGKFLHRL. (**G**) A2-ILGKFLHWL. (**H**) A2-ILAKFLHRL. (**I**) A2-ILAKFLHTL. (**J**) A2-ILAKFLHEL.

**Figure 4*:* Stability of HLA-A2-ILA variants using circular dichroism.** (**A**). CD thermal denaturation curves recorded at 218nm are shown for selected pHLAI samples. Dots represent measured values fitted assuming a 2-state trimer-to-monomer transition (dashed lines) as described in Materials and Methods. (**B**). Bar graphs of the thermal stability with respect to melting temperature (upper) and van 't Hoff's enthalpy of unfolding (lower panel).

**Figure 5. Structural analysis of ILA1 TCR ligands.** Conformation of each APL (sticks) demonstrating the direction of each residue in the peptide (solvent exposed, MHC anchor, or in between) with the MHCα1 helix shown as grey cartoon. Residues in red indicate differences from the index sequence. Up arrow indicates solvent exposed, down arrow indicates anchor position, no arrow indicates an intermediate position. (**A**) A2-ILAKFLHWL (orange sticks). (**B**) A2-ILGKFLHRL (purple sticks). (**C**) A2-ILGKFLHWL (green sticks). (**D**) A2-ILAKFLHRL (brown sticks), reproduced from (23). (**E**) A2-ILAKFLHTL (sand

sticks). (**F**) A2-ILAKFLHEL (black sticks). The circled residues in (**E**) and (**F**) face in different directions as compared to the index telomerase sequence (ILAKFLHWL) in (**A**).

**Figure 6. Equilibrium binding analysis and structural modelling of the ILA1 TCR interaction with the APLs.** Binding affinity of the ILA1 TCR interaction with different APLs at 25°C. Eight serial dilutions of the ILA1TCR were injected over A2-ILA8R and A2-ILA3G8R and representative data from 3 independent experiments are plotted after deducting binding to a control sample (HLA-A*0201-ALWGPDPAAA). The equilibrium binding constants ($K_D$) were calculated using a nonlinear curve fit ($y= (P_1x)/P_2 + x)$). (**A**) ILA1-A2-ILA3G8R (**B**) ILA1-A2-ILA3G (reproduced from (15)), (**C**) ILA1-A2-ILA8R, (**D**) ILA1-A2-ILA8T (reproduced from (15)), (**E**) ILA1-A2-ILA8E (reproduced from (15)). The ILA1 TCR was modelled with each of APL ligands by aligning the uncomplexed APLs with A2-ILA in the ILA1-A2-ILA complex structure (HLA-A*0101 α1 helix shown in gray cartoon in panels F-J). Potential steric clashes are highlighted in red circles. (**F**) A2-ILA3G8R (peptide in purple) showing modelled positions of the TCR CDR3α and β loops (blue and cyan, respectively) (**G**) A2-ILA3G (peptide in green) superposed with the A2-ILA peptide (orange), (**H**) A2-ILA8R (peptide in brown) showing modelled positions of the TCR CDR1, 2 and 3 α β loops (yellow, pink and cyan, respectively), (**I**) A2-ILA8T (peptide in sand) showing modelled positions of the TCR CDR3α and β loops (blue and cyan, respectively) and (**J**) ILA1-A2-ILA8E (peptide in black) showing modelled positions of the TCR CDR3α and β loops (blue and cyan, respectively).

**Figure 7: pEC50 values for all peptide ligands tested**. Simultaneous curve fitting was used to estimate functional sensitivity measured as pEC50 for peptides sampled from: I. xLGxxxxRL (set size $19^5$; 30 peptides sampled at random); II. xLxKFLxxL (set size $19^4$; 30 peptides sampled at random); III. xLG{K/L}F{L/I} {M/F/Y/N/H}{R/T/Y/K/S/F/H/I/L/M/Q/V/G/N }{L/V} (set size 10640; 30 peptides sampled at random); IV. {A/I/L/M/P/Q/W}L{G/A}{K/L}F{L/I}{N/H} {F/K/N/Q/T/Y}{L/V} and V. replicate of a biased sampling each set (20 peptides sampled from an effective sample size of $1.5 \times 10^7$).

**Figure 8. TCR binding mode contributes to T-cell cross-reactivity.** (**A**) Peptide recognition degeneracy for ILA1. The degeneracy curve plots the estimated number of peptides that have a functional sensitivity at least as strong as abscissa. The highest value of abscissa corresponds to inferred functional sensitivity of optimal peptide, whereas the lowest value of abscissa lies 10 orders of magnitude below this optimum. Bias: degeneracy curve based on agonist-biased importance sampling. Curves I–IV are motif-based, i.e., sampled from subsets of the entire peptide universe, and therefore lie below the degeneracy curve. I: xLGxxxxRL (set size $19^5$; 30 peptides sampled at random); II: xLxKFLxxL (set size 194; 30 peptides sampled at random); III:xLG{K/L}F{L/I}{M/F/Y/N/H}{R/T/Y/K/S/F/H/I/L/M/Q/V/G/N}{L/V} (set size 10640;30 peptides sampled at random); IV: {A/I/L/M/P/Q/W}L{G/A}{K/L}F{L/I}{N/H} {F/K/N/Q/T/Y}{L/V} (set size 1344; 19peptides sampled at random); x denotes any of the 19 amino acids excluding cysteine and y denotes any of the 20 amino acids. Cross-reactivity of the 1E6 (**B**) and ILA1 (**C**) $CD8^+$ T-cell clones estimated by the number of residues recognized in a combinatorial peptide library screen generating responses over >0.5ng/mL MIP-1β. Bars that represent positions in the peptide that are anchor residues are colored grey. Bars that represent residues that are main contact sites for each respective TCR are colored red.

**Tables:**

**Table 1. Data collection and structure refinement statistics**

| | ILA1-A2-ILA | A2-ILA3G8R | A2-ILA3G | A2-ILA8T | A2-ILA8E |
|---|---|---|---|---|---|
| **Data collection** | | | | | |
| PDB code | 5C60 | 5C61 | 5C62 | 5C63 | 5C64 |
| Space group | P 1 21 1 | P 21 21 21 | P 21 21 21 | P 1 21 1 | P 21 21 21 |
| Beamline | DLS I24 | DLS I04 | DLS I02 | DLS I04 | DLS I02 |
| **Cell dimensions** | | | | | |
| *a*, *b*, *c* (Å) | 93.2, 48.7, 118.1 | 49.2, 74.9, 125.8 | 119.4, 169.6, 47.1 | 53.34, 81.44, 56.77 | 45.7, 119.0, 170.3 |
| α, β, γ (°) | 90, 108.2, 90 | 90, 90, 90 | 90, 90, 90 | 90, 113.5, 90 | 90, 90, 90 |
| Resolution Maximum (Å) | 2.81 (2.88 - 2.81) | 1.77 (1.82 - 1.77) | 2.71 (2.78 - 2.71) | 2.27 (2.33 – 2.27) | 1.88 (1.93 -1.88) |
| $R_{merge}$ (%) | 0.100 (0.73) | 0.105 (0.718) | 0.105 (.831) | 0.086 (0.632) | 0.098 (0.621) |
| Total Measurements | 91,454 (6,913) | 332,927 (22,417) | 195,005 (13,976) | 74,664 (5,741) | 545,975 (42,439) |
| Unique reflections | 25,031 (1,865) | 46,011 (3,362) | 26,894 (1,974) | 20,523 (1,527) | 76,630 5,603 |
| $I / \sigma I$ | 9.2 (1.9) | 10.5 (2.5) | 14.1 (2.3) | 12.8 (1.9) | 10.8 (3.2) |
| CC1/2 | n/a | n/a | 0.997 (0.714) | 0.997 (0.842) | 0.997 (0.913) |
| Completeness (%) | 99.7 (99.8) | 100 (100) | 100 (100) | 99.3 (98.9) | 99.9 (100) |
| Multiplicity | 3.7 (3.7) | 7.2 (6.7) | 7.3(7.1) | 3.5 (3.8) | 7.1 (7.6) |
| **Refinement** | | | | | |
| Resolution (Å) | 56.11 - 2.81 | 48.16 - 1.77 | 42.39 - 2.71 | 48.90 – 2.27 | 34.06 - 1.9 |
| No. reflections in Work set | 23,666 | 43,626 | 25,511 | 19,484 | 72,747 |
| No reflections in $R_{free}$ set | 1,267 | 2,319 | 1,333 | 1,023 | 3,797 |
| $R_{work}/R_{free}$ (%) | 18.9/27.2 | 17.1/21.4 | 18.2/23.9 | 20.3/26.4 | 18.4/22.4 |
| Mean B value (Å$^2$) | 57.1 | 27.7 | 52.2 | 37.7 | 33.0 |
| Wilson B factor (Å$^2$) | 70.5 | 21.9 | 39.6 | 32.97 | 28.6 |
| Overall coord error (Å) | 0.377 | 0.082 | 0.252 | 0.232 | 0.098 |
| **R.m.s. deviations** | | | | | |
| Bond lengths (Å) | 0.017 | 0.019 | 0.013 | 0.019 | 0.019 |
| Bond Angles (°) | 2.025 | 1.983 | 1.658 | 1.976 | 1.967 |
| **Ramachandran Plot statistics** | | | | | |
| Most Favoured region (%) | 92.27 | 97.85 | 97.23 | 96.54 | 98.12 |
| Allowed Region (%) | 6.38 | 2.15 | 2.77 | 3.19 | 1.88 |
| Outliers (%) | 1.35 | 0 | 0 | 0.27 | 0 |

N.B. Figures in brackets refer to the highest resolution bin

*** One crystal was used for solving each structure.**

**Table 2:** ILA1-A2-ILA Contact Summary

| | vdW (≤4 Å) | H-bonds (≤3.4 Å) | salt bridges (≤3.4 Å) |
|---|---|---|---|
| MHC | 58 | 6 | 2 |
| Peptide | 49 | 4 | 1 |
| Peptide Lys4 | 20 | 2 | 1 |
| Peptide Trp8 | 19 | 2 | 0 |
| TCRα | 62 | 7 | 3 |
| CDR1α | 16 | 2 | 0 |
| CDR2α/FW | 9 | 2 | 2 |
| CDR3α | 37 | 3 | 1 |
| TCRα Asp97 | 6 | 2 | 0 |
| TCRα Arg68 | 4 | 0 | 2 |
| TCRβ | 45 | 3 | 0 |
| CDR1β | 13 | 0 | 0 |
| CDR2β | 7 | 0 | 0 |
| CDR3β | 25 | 3 | 0 |
| TCRβ Gln96 | 18 | 3 | 0 |
| **Total contacts** | **107** | **10** | **3** |

| | |
|---|---|
| BSA(Å$^2$) | 2507.2 |
| SC (TCR-MHC) | 0.635 |
| SC (TCR-p) | 0.707 |
| SC (TCR-pMHC) | 0.641 |
| Crossing angle (°)* | 51.8 |

BSA – buried surface area
SC – surface complementarity
*calculated as previously described (39)

**Table 3:** ILA1-A2-ILA contacts

| CDR loop | Gene Usage | TCR residue | Peptide residue | MHC residue | vdW (≤4 Å) | H-bonds (≤3.4 Å) |
|---|---|---|---|---|---|---|
| CDR1α | *TRAV22* | Asp27$^{O\delta2}$ | | Thr163$^{O\gamma1}$ | 4 | 1 |
| | *TRAV22* | Asp27$^{O\delta1}$ | | Glu166$^{O\varepsilon2}$ | 2 | 1 |
| | *TRAV22* | Ser28 | | Ala158 | 1 | |
| | *TRAV22* | Val29 | | Gln155 | 4 | |
| | *TRAV22* | Val29 | Lys4 | | 4 | |
| | *TRAV22* | Asn30 | | Gln155 | 1 | |
| CDR2α | *TRAV22* | Tyr48 | | Glu154 | 1 | |
| | *TRAV22* | Tyr48 | | Gln155 | 2 | |
| | *TRAV22* | Ser51$^{O}$ | | Glu154$^{O\delta}$ | 1 | 1 |
| | *TRAV22* | Ser51 | | Arg157 | 1 | |
| | *TRAV22* | Gln55$^{O\varepsilon1}$ | | Glu154$^{O\varepsilon2}$ | | 1 |
| FWα | *TRAV22* | Arg68$^{NH1/NH2}$ | | Glu166$^{O\varepsilon1/O\varepsilon2}$ | 4 | 2 salt bridges |
| CDR3α | *TRAJ40* | Asp91$^{O\delta1}$ | Lys4$^{N\zeta}$ | | 3 | 1 salt bridge |
| | *TRAJ40* | Ser92$^{O}$ | Lys4$^{N\zeta}$ | | 2 | 1 |
| | *TRAJ40* | Ala93 | | Lys66 | 1 | |
| | *TRAJ40* | Ala93 | | Thr163 | 2 | |
| | *TRAJ40* | Ala93 | Lys4 | | 2 | |
| | *TRAJ40* | Thr94$^{O}$ | | Lys66$^{N\zeta}$ | 3 | 1 |
| | *TRAJ40* | Thr94 | | Trp167 | 3 | |
| | *TRAJ40* | Ser95 | Lys4 | | 4 | |
| | *TRAJ40* | Gly96 | | Arg65 | 3 | |
| | *TRAJ40* | Thr97$^{O}$ | Lys4$^{N\zeta}$ | | 3 | 1 |
| | *TRAJ40* | Tyr98 | | Lys66 | 3 | |
| | *TRAJ40* | Tyr98 | | Ala69 | 2 | |
| | *TRAJ40* | Tyr98 | Lys4 | | 2 | |
| | *TRAJ40* | Tyr98 | Leu6 | | 4 | |
| CDR1β | *TRBV6* | Glu30 | Trp8 | | 13 | |
| CDR2β | *TRBV6* | Val50 | | Val152 | 1 | |
| | *TRBV6* | Val50 | Leu6 | | 1 | |
| | *TRBV6* | Val50 | Trp8 | | 1 | |
| | *TRBV6* | Ile54 | | Gln72 | 4 | |
| CDR3β | *TRBJ1-1* | Tyr95 | | Lys146 | 5 | |
| | *TRBJ1-1* | Tyr95 | | Ala150 | 2 | |
| | *TRBJ1-1* | Gln96 | | Lys146 | 2 | |
| | *TRBJ1-1* | Gln96$^{O\varepsilon1}$ | | Trp147$^{N\varepsilon1}$ | 4 | 1 |
| | *TRBJ1-1* | Gln96 | | Ala150 | 1 | |
| | *TRBJ1-1* | Gln96 | | Val152 | 1 | |
| | *TRBJ1-1* | Gln96 | Leu6 | | 2 | |
| | *TRBJ1-1* | Gln96 | His7 | | 3 | |
| | *TRBJ1-1* | Gln96$^{O\varepsilon1}$ | Trp8$^{N/O}$ | | 5 | 2 |

*N* = N-nucleotide insertion

*A 3.4Å cut-off was used for H-bonds and salt bridges and a 4Å cut-off was used for vdW.

**Table 4:** ILA1 TCR binding affinity to peptide variants

| HLA-A*0201 ILA variant | $K_{on}$ $M^{-1}s^{-1}$ | $K_{off}$ $s^{-1}$ | Chi² for $K_{on}$ & $K_{off}$ | $K_D$ µM |
|---|---|---|---|---|
| HLA-A*0201-IL<u>G</u>KFLH<u>R</u>L | $3x10^4$ | 0.16 | 4.2 | 1±0.1 |
| HLA-A*0201-IL<u>G</u>KFLHWL (41) | $1.6x10^4$ | 0.05 | 3.6 | 3.7±0.2 |
| HLA-A*0201-IL<u>G</u>KFLH<u>T</u>L (17) | $1.95x10^4$ | 0.05 | 2.5 | 2.5±0.5 |
| HLA-A*0201-ILAKFLH<u>Y</u>L (41) | n/m | n/m | n/m | 22.6±2.1 |
| HLA-A*0201-ILAKFLHWL (20) | $4.5x10^3$ | 0.15 | 4.3 | 34±2 |
| HLA-A*0201-ILAKFLH<u>T</u>L (41) | $2.2x10^3$ | 0.08 | 1.9 | 28±5 |
| HLA-A*0201-ILAKFL<u>Y</u>WL (41) | n/m | n/m | n/m | 82±8 |
| HLA-A*0201-ILA<u>L</u>FLHWL (16) | $1.7x10^3$ | 0.2 | 4.2 | 117±6 |
| HLA-A*0201-ILAKFLH<u>R</u>L | n/m | n/m | n/m | 151±8 |
| HLA-A*0201-ILAK<u>Y</u>LHWL (17) | $1.3x10^3$ | 0.32 | 3.5 | 242±20 |
| HLA-A*0201-ILAKFLH<u>E</u>L (41) | n/m | n/m | n/m | >500 |

n/m = kinetics were too fast to accurately measure
$K_D$ calculated from equilibrium binding experiments

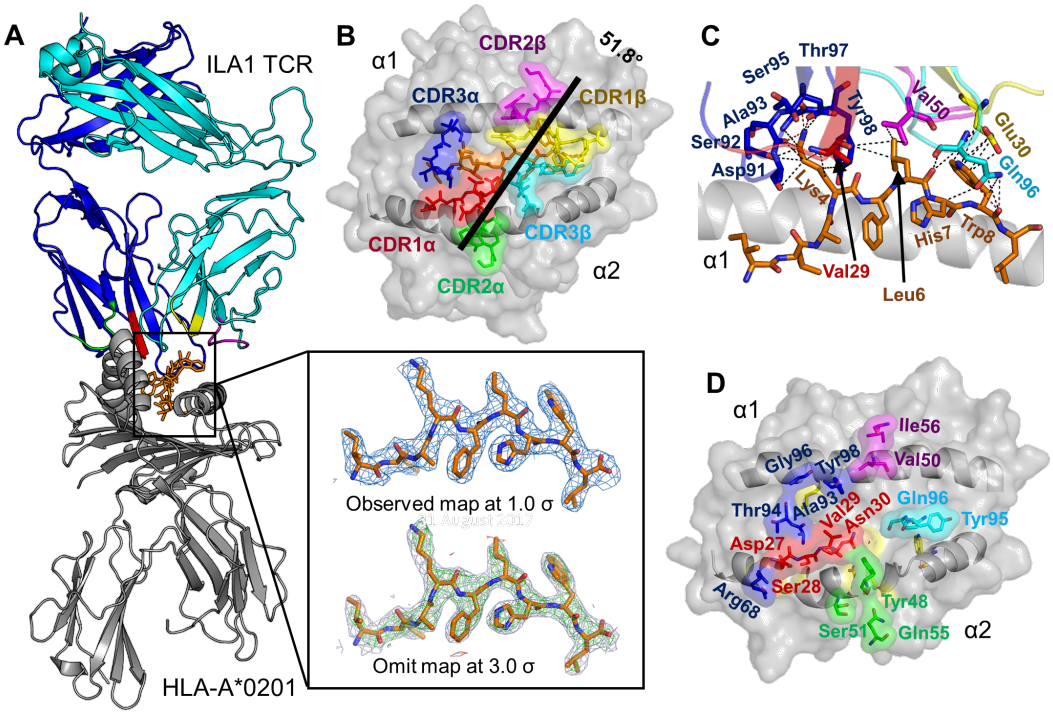
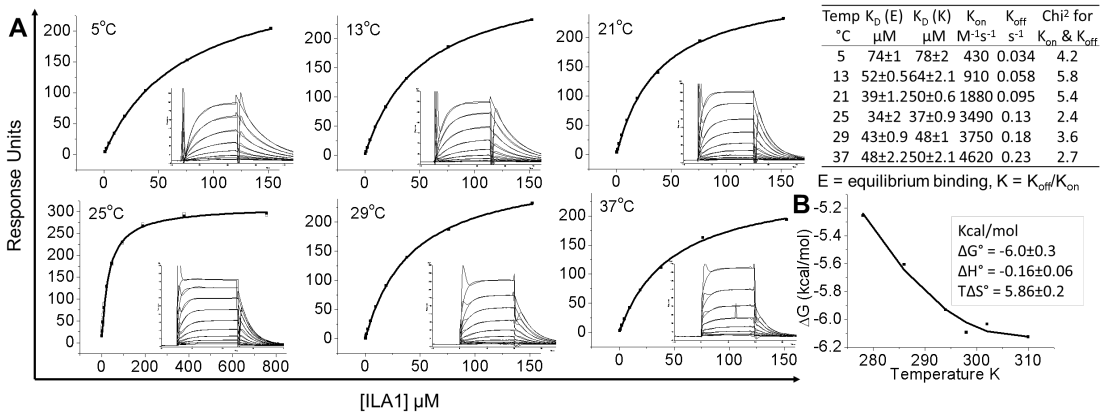

A
ILA1 TCR
HLA-A*0201
B
α1
CDR2β
51.8°
CDR3α
CDR1β
CDR1α
CDR3β
CDR2α
α2
Observed map at 1.0 σ
Omit map at 3.0 σ
C
Ser95
Thr97
Ala93
Tyr98
Val50
Glu30
Ser92
Gln96
Asp91
Lys4
Trp8
His7
α1
Val29
Leu6
D
α1
Ile56
Gly96
Tyr98
Val50
Thr94
Ala93
Val29
Asn30
Gln96
Tyr95
Asp27
Ser28
Tyr48
Arg68
Ser51
Gln55
α2

*Figure 1*




| Temp °C | $K_D$ (E) µM | $K_D$ (K) µM | $K_{on}$ $M^{-1}s^{-1}$ | $K_{off}$ $s^{-1}$ | Chi² for $K_{on}$ & $K_{off}$ |
|---|---|---|---|---|---|
| 5 | 74±1 | 78±2 | 430 | 0.034 | 4.2 |
| 13 | 52±0.5 | 64±2.1 | 910 | 0.058 | 5.8 |
| 21 | 39±1.2 | 50±0.6 | 1880 | 0.095 | 5.4 |
| 25 | 34±2 | 37±0.9 | 3490 | 0.13 | 2.4 |
| 29 | 43±0.9 | 48±1 | 3750 | 0.18 | 3.6 |
| 37 | 48±2.2 | 50±2.1 | 4620 | 0.23 | 2.7 |

E = equilibrium binding, K = $K_{off}/K_{on}$

*Figure 2*

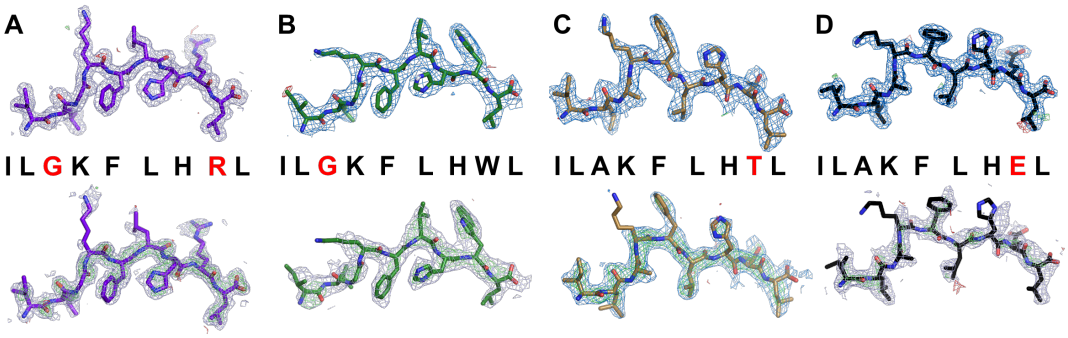



*Figure 3*

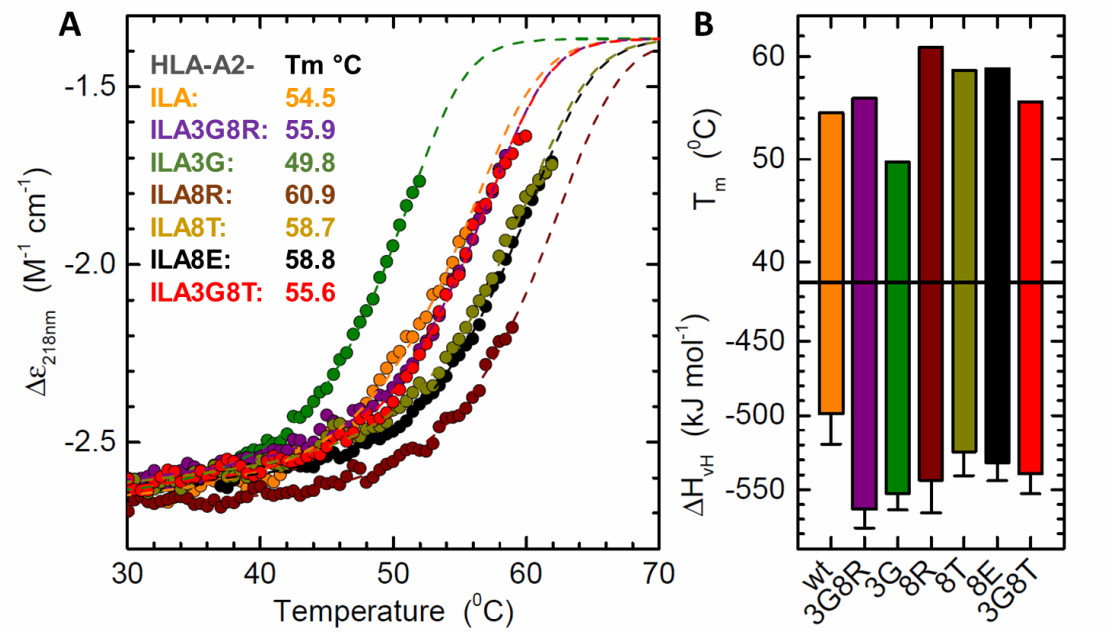

A
HLA-A2- Tm °C
ILA: 54.5
ILA3G8R: 55.9
ILA3G: 49.8
ILA8R: 60.9
ILA8T: 58.7
ILA8E: 58.8
ILA3G8T: 55.6
Δε218nm (M-1 cm-1)
-1.5
-2.0
-2.5
30
40
50
60
70
Temperature (0C)
B
Tm (0C)
60
50
40
ΔHvH (kJ mol-1)
-450
-500
-550
wt
3G8R
3G
8R
8T
8E
3G8T

*Figure 4*

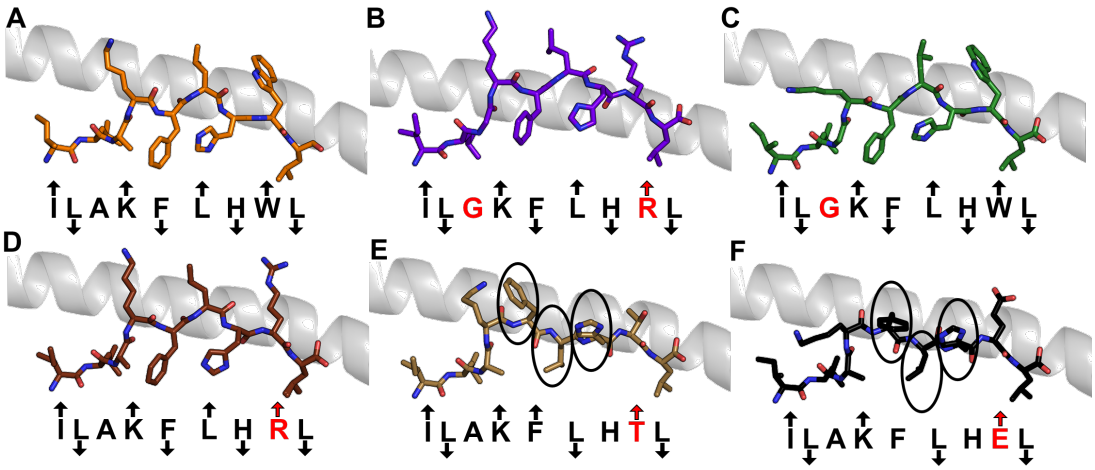

A
ILAKFLHWL
B
ILGKFLHRL
C
ILGKFLHWL
D
ILAKFLHRL
E
ILAKFLHTL
F
ILAKFLHEL

*Figure 5*

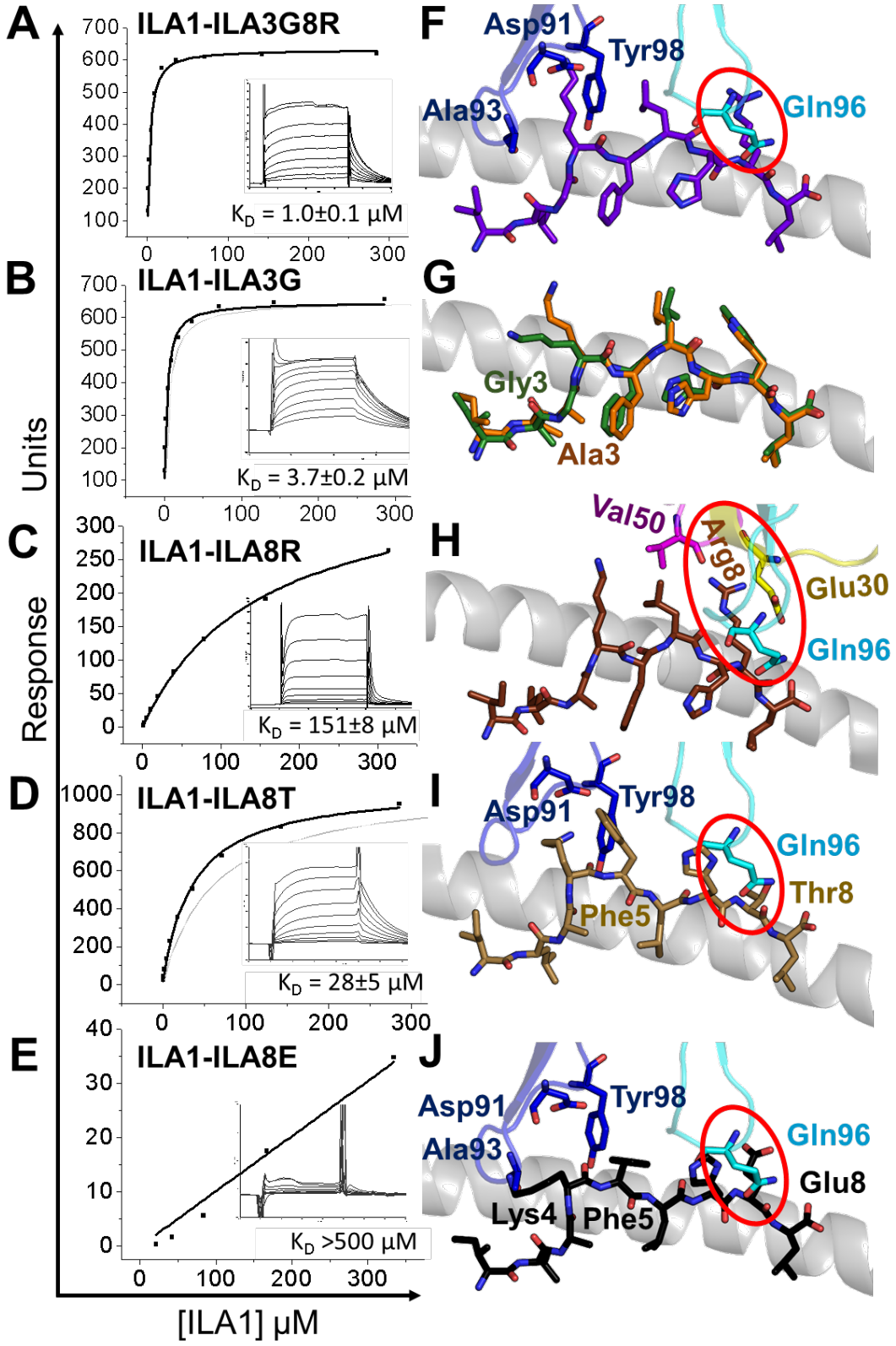

A
ILA1-ILA3G8R
K_D = 1.0±0.1 µM
B
ILA1-ILA3G
K_D = 3.7±0.2 µM
C
ILA1-ILA8R
K_D = 151±8 µM
D
ILA1-ILA8T
K_D = 28±5 µM
E
ILA1-ILA8E
K_D >500 µM
Response Units
[ILA1] µM
F
Asp91
Tyr98
Ala93
Gln96
G
Gly3
Ala3
H
Val50
Arg8
Glu30
Gln96
I
Asp91
Tyr98
Gln96
Thr8
Phe5
J
Asp91
Tyr98
Ala93
Gln96
Glu8
Lys4
Phe5

*Figure 6*

**I xLGxxxxRL motif**

| u | Peptide Sequence | $pEC_{50}$ |
|---|---|---|
| 1 | MLGLIQPRL | 3.918 |
| 2 | PLGYSEGRL | 4.932 |
| 3 | DLGYWGTRL | 5.016 |
| 4 | NLGAPLVRL | 5.062 |
| 5 | GLGWTPARL | 5.069 |
| 6 | FLGKQSHRL | 5.081 |
| 7 | WLGSSDORL | 5.815 |
| 8 | RLGGYYARL | 5.856 |
| 9 | QLGPMLRRL | 5.864 |
| 10 | VLGMRTMRL | 5.866 |
| 11 | PLGNWKPRL | 5.883 |
| 12 | PLGRIIHRL | 5.897 |
| 13 | SLGMTYRRL | 5.909 |
| 14 | ALGIIAARL | 5.926 |
| 15 | ALGMQFDRL | 5.928 |
| 16 | LLGDMSTRL | 5.937 |
| 17 | LLGQYGSRL | 5.943 |
| 18 | QLGQLVMRL | 5.951 |
| 19 | GLGKQQIRL | 5.971 |
| 20 | KLGPFLNRL | 5.987 |
| 21 | ELGVYQKRL | 5.990 |
| 22 | GLGLHYRRL | 6.007 |
| 23 | SLGLYYVRL | 6.010 |
| 24 | WLGELEVRL | 6.037 |
| 25 | KLGVREVRL | 6.045 |
| 26 | ILGPKNYRL | 6.045 |
| 27 | ALGNIMYRL | 6.058 |
| 28 | GLGLWKWRL | 6.071 |
| 29 | ALGLNYWRL | 6.094 |
| 30 | NLGEFYARL | 6.990 |
| Index | ILAKFLHWL | 8.983 |

**II xLxKFLxxL motif**

| Number | Peptide Sequence | $pEC_{50}$ |
|---|---|---|
| 1 | YLMKFLAEL | 2.860 |
| 2 | KLEKFLWKL | 3.045 |
| 3 | KLPKFLIEL | 3.062 |
| 4 | LLIKFLDQL | 3.526 |
| 5 | DLTKFLESL | 3.557 |
| 6 | ILAKFLYKL | 3.639 |
| 7 | ALVKFLNWL | 3.649 |
| 8 | GLKKFLMPL | 4.252 |
| 9 | LLPKFLRLL | 4.269 |
| 10 | VLLKFLKKL | 4.291 |
| 11 | WLHKFLQHL | 4.379 |
| 12 | ILNKFLPWL | 4.415 |
| 13 | DLTKFLVQL | 4.539 |
| 14 | NLDKFLQKL | 4.837 |
| 15 | PLSKFLYIL | 4.906 |
| 16 | PLQKFLPKL | 4.931 |
| 17 | FLKKFLGDL | 4.935 |
| 18 | KLQKFLVQL | 4.976 |
| 19 | MLVKFLMNL | 5.086 |
| 20 | ALRKFLRQL | 5.089 |
| 21 | WLRKFLVKL | 5.277 |
| 22 | MLSKFLMKL | 5.702 |
| 23 | HLLKFLKQL | 5.708 |
| 24 | SLGKFLQGL | 5.718 |
| 25 | HLGKFLIIL | 5.778 |
| 26 | FLWKFLIEL | 5.821 |
| 27 | VLGKFLTML | 6.101 |
| 28 | PLRKFLYSL | 6.374 |
| 29 | ELGKFLLSL | 6.493 |
| 30 | NLPKFLHQL | 6.614 |
| Index | ILAKFLHWL | 8.455 |

**III xLG{K/L}F{L/I}{M/F/Y/N/H} {R/T/Y/K/S/F/H/I/L/M/Q/V/G/N} {L/V} motif**

| Number | Peptide Sequence | $pEC_{50}$ |
|---|---|---|
| 1 | NLGLFIFGL | null |
| 2 | VLGLFINNL | null |
| 3 | HLGKFIYHL | null |
| 4 | ALGLFLYSV | null |
| 5 | RLGLFIFNL | null |
| 6 | SLGLFIYSL | null |
| 7 | HLGLFIMTV | 3.384 |
| 8 | YLGKFLYIV | 3.385 |
| 9 | KLGKFIFVL | 3.524 |
| 10 | WLGLFIFFL | 3.578 |
| 11 | FLGLFLFTV | 3.611 |
| 12 | MLGLFIMIL | 3.619 |
| 13 | HLGKFIMML | 3.721 |
| 14 | ALGKFIHSV | 3.849 |
| 15 | RLGLFIHGV | 3.890 |
| 16 | QLGKFIHSV | 5.015 |
| 17 | ELGLFLFSL | 5.026 |
| 18 | DLGKFIHLV | 5.192 |
| 19 | NLGLFLFRL | 5.231 |
| 20 | WLGKFIYKL | 5.372 |
| 21 | WLGLFLNQL | 5.442 |
| 22 | PLGKFLNQV | 5.805 |
| 23 | SLGKFLFMV | 5.885 |
| 24 | PLGKFIFQL | 6.111 |
| 25 | WLGKFLHFL | 6.189 |
| 26 | QLGLFINKL | 6.306 |
| 27 | ALGKFLHYV | 7.159 |
| 28 | LLGLFLNNL | 7.201 |
| 29 | ILGKFIHYV | 8.495 |
| 30 | MLGKFLNTV | 8.833 |
| Index | ILAKFLHWL | 8.889 |

**IV {A/I/L/M/P/Q/W}L {G/A}{K/L}F{L/I}{N/H} {F/K/N/Q/T/Y}{L/V} motif**

| Number | Peptide Sequence | $pEC_{50}$ |
|---|---|---|
| 1 | QLAKFLNTV | 3.705 |
| 2 | ILALFLNYV | 3.926 |
| 3 | WLAKFLNFV | 4.717 |
| 4 | MLAKFLNKL | 4.742 |
| 5 | WLAKFIHFV | 5.140 |
| 6 | QLALFLHQV | 5.339 |
| 7 | QLAKFLHNV | 5.746 |
| 8 | PLALFLHFV | 5.967 |
| 9 | QLGKFINTV | 6.023 |
| 10 | QLAKFINFV | 6.054 |
| 11 | LLALFLNYV | 6.611 |
| 12 | ILGKFINQL | 6.987 |
| 13 | MLGLFINKL | 7.276 |
| 14 | QLALFLHQL | 7.400 |
| 15 | MLALFIHFL | 7.516 |
| 16 | MLGKFLNFL | 9.172 |
| 17 | MLGLFINTL | 9.299 |
| Index | ILAKFLHWL | 9.381 |
| 18 | QLGKFLHTV | 10.352 |
| 19 | MLGKFIHTV | 10.572 |

**V Biased sampling set**

| Number | Peptide Sequence | $pEC_{50}$ |
|---|---|---|
| 1 | VSHKFLART | null |
| 2 | FTGKVWHKL | null |
| 3 | MQKKFLRRV | null |
| 4 | LVGLFDFTT | null |
| 5 | RTGKFIIQL | null |
| 6 | IYGKFLPRL | null |
| 7 | WVGLWLFSL | null |
| 8 | ITGKSPWQF | 2.519 |
| 9 | MSGKFGNID | 2.626 |
| 10 | MSGKSLAYT | 3.030 |
| 11 | LNYKFLHFL | 3.288 |
| 12 | MLRKFLMYL | 3.637 |
| 13 | VTGKFPFKI | 3.863 |
| 14 | ITGLWLFTL | 3.962 |
| 15 | VIQKFLFTI | 4.010 |
| 16 | IILKWLYRI | 4.017 |
| 17 | IVWKHWFRM | 4.310 |
| 18 | YIGKFLFSQ | 4.395 |
| 19 | MQKLIWHYT | 4.571 |
| 20 | MVGKVIHTI | 4.993 |
| Index | ILAKFLHWL | 8.337 |

*Figure 7*

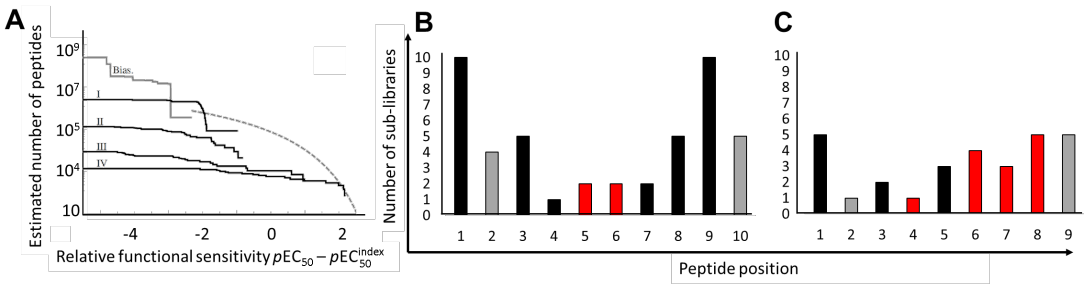

A
B
C
Estimated number of peptides
Bias.
I
II
III
IV
Relative functional sensitivity $pEC_{50} - pEC_{50}^{Index}$
Number of sub-libraries
Peptide position

*Figure 8*