Published in final edited form as:

Immunol Rev. 2012 November ; 250(1): 82–101. doi:10.1111/imr.12006.

Diversity-oriented approaches for interrogating T-cell receptor repertoire, ligand recognition, and function

Michael E. Birnbaum, Shen Dong, and K. Christopher Garcia

Departments of Molecular and Cellular Physiology, and Structural Biology, Program in Immunology, and Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA

Summary

Molecular diversity lies at the heart of adaptive immunity. T-cell receptors and peptide-major histocompatibility complex molecules utilize and rely upon an enormous degree of diversity at the levels of genetics, chemistry, and structure to engage one another and carry out their functions. This high level of diversity complicates the systematic study of important aspects of T-cell biology, but recent technical advances have allowed for the ability to study diversity in a comprehensive manner. In this review, we assess insights gained into T-cell receptor function and biology from our increasingly precise ability to assess the T-cell repertoire as a whole or to perturb individual receptors with engineered reagents. We conclude with a perspective on a new class of high-affinity, non-stimulatory peptide ligands we have recently discovered using diversity-oriented techniques that challenges notions for how we think about T-cell receptor signaling.

Keywords

T cells; T-cell receptors; library; selection; evolution

Introduction

A functioning adaptive immune system is based on the generation and recognition of molecular and chemical diversity. The $\alpha\beta$ T-cell receptor (TCR)'s ability to combine fine specificity with sensitivity has made it a fascinating arena of study for the application of technical approaches embracing the concept of molecular diversity. The TCR repertoire is highly diverse, consisting of an estimated 10^{15} potential sequences (1). Yet, unlike other antigen receptors, the $\alpha\beta$ TCR is restricted to recognizing its antigen presented in the context of major histocompatibility complex (MHC) molecules, resulting in a significant degree of structural commonality between all peptide-MHC ligands that the TCR surveys (2). The germline-derived binding determinants of the TCR (CDR1 and CDR2 loops) primarily make contact with residues on the helices of the MHC while the somatically derived CDR3 loops primarily contact the displayed peptide (*Fig. 1*). Despite these commonalities, as few as 10 agonist peptide-MHC (pMHC) ligands can be discriminated from self-peptides, and during development, ligands with undetectably weak affinity are capable of inducing TCR signaling (3-5). The very sensitivity and complexity of the system that makes it so interesting also makes delimiting the rules of TCR ligand recognition and function difficult. As a field, we are still addressing questions such as (i) What is a good TCR agonist? (ii)

Correspondence to: K. Christopher Garcia Howard Hughes Medical Institute Depts. of Molecular and Cellular Physiology, and Structural Biology Stanford University School of Medicine Beckman B171B 279 Campus Drive Stanford, CA 94305-5345 Tel.: +1 650 498 7332 Fax: +1 650 725 8021 kcgarcia@stanford.edu.

How cross-reactive are TCRs? (iii) How many unique TCRs are in a given immune response? (iv) How does structure influence TCR signaling?

In this review, we focus on recent progress in answering some of these questions using methods that either allows for the study of or generation of molecular diversity. The ability to assess the T-cell repertoire and track the global effects of experimental manipulation has provided insight into T-cell function on a macroscopic level. Conversely, the ability to thoroughly interrogate a single pMHC-TCR interaction through the creation of molecular diversity has provided insight into the molecular underpinnings of TCR function on the microscopic level.

TCR repertoire analysis through next-generation sequencing

Technological advances in high-throughput sequencing have provided a method to take snapshots of the diversity of the entire TCR repertoire. 'Next generation' sequencing platforms such as Illumina and 454 have allowed for the rapid generation of millions of short, unique DNA reads (6, 7). The DNA reads are ideally sized to capture the identity of a V region and CDR3 sequence and have recently been used to sequence whole zebrafish antibody repertoires (8). Several studies have addressed the sequence diversity and makeup of the TCR β repertoire in humans. An attempt to exhaustively sequence the TCR β repertoire found ~1-5 million unique TCR β V region sequences, corresponding well to previous estimates derived before whole repertoire sequencing was feasible (9-11). This serves as a lower bound for diversity, because the sequence repertoire of any given blood sample may not completely represent the whole peripheral repertoire. It was also demonstrated that V region representation in the TCR β repertoire is highly variable, with ~2500 fold difference in usage between the most and least represented gene segment (12).

Sorting T cells based upon surface markers and then sequencing defined pools allows for the TCR sequence of specific cell types to be quantified. For example, when the memory TCR β sequence repertoire was assayed, it was only 5-10-fold less diverse than the naive repertoire. This observation is somewhat surprising, given the expectation that only a small subset of T cells would have been activated and converted to memory cells, leading to a correspondingly less diverse repertoire. Additionally, a large number of the TCR β sequences in the memory compartment were in low abundance, suggesting the memory repertoire contains many unexpanded clones (13, 14). While the implications of the finding are not yet clear, this is an excellent example for how snapshots of repertoire diversity can provide new hypotheses for study in T-cell biology.

Comparison of the TCR β sequence repertoire between two individuals showed between ~1% (14) and ~10% (9) of amino acid sequences are identical. Given the theoretical T-cell diversity of $\sim 10^{15}$ unique clones, fewer than 10 TCR β clones would be expected to be shared between any two individuals by chance. The large amount of sequence conservation is suggestive of the selective pressures of thymic selection and exposure to the environment (9), although it is also possibly a byproduct of V(D)J recombination (15). There are significant differences in the number of shared sequences between MHC-matched and -mismatched individuals, suggesting that allele-specific effects of MHC may also play a role in selection of the T-cell repertoire (9). Fewer common sequences between individuals are noted in the memory compartment of T cells, which is reasonable given a person's antigen experience is likely to be variable (14). This large public repertoire of shared TCR sequences has implications that reach beyond a common footprint imprinted onto the human immune system by thymic selection and environmental exposure. Study and manipulation of this shared pool has the potential to provide outsized benefits to human health given its mass applicability.

In addition to sequence composition and conservation, deep sequencing of the immune repertoire is also providing insight into mechanisms of immune system function. A key place where this is playing out is through comparison of different T-cell functional compartments. When sequences are compared between T cells exhibiting a T-helper 1 (Th1), Th2, CD8⁺, or T-regulatory cell (Treg) phenotype, shared sequences are observed between essentially all compartments (16). This strongly suggests a model where undifferentiated progenitor cells polarize into various T-cell subsets through differentiation programs propagated by cytokines rather than an instructive signal from a TCR engaging a specific antigen. Interestingly, there is also a large pool of TCR β sequences shared between CD4⁺ and CD8⁺ cells (16). While there are known examples of individual receptors able to recognize both class I and class II MHC, the biological implications of this phenomenon is unknown (17-19).

The majority of sequencing analysis has been done on the TCR β chain, as it provides a larger source of diversity, and a single TCR β chain can be paired with multiple TCR α chains. Currently, the sequences of paired α and β chains cannot be determined, as the connectivity between TCR α and TCR β sequences is lost during sample preparation. Even if full receptor sequences could be easily recovered, there currently exists no plausible mechanism to intuit peptide or MHC specificity from sequence alone.

Analyzing the naive immune repertoire via direct counting

A second method to view the immune repertoire is through the number of specific ligands it is able to recognize. The connectivity between these two properties is not clear-cut: the sequence diversity of the immune repertoire does not accurately predict the number of antigens able to be recognized. The collection of receptors present in a given organism can likely recognize a larger pool of antigen than its sequence diversity would suggest due to TCR cross-reactivity. To obtain this type of information, work has been conducted to directly count the number of T cells in a given organism that recognize a pMHC of interest (20, 21). The small number of naive T cells specific for any given peptide ligand makes such a measurement extremely challenging. Recently, Moon *et al.* (22) devised a way to improve the signal to noise of quantitating low abundance T cells. Instead of a direct count of T cells stained with fluorescent pMHC tetramer, the T cells are first enriched for tetramer staining via magnetic activated cell sorting (MACS) based upon the tetramer's fluorophore and then counted via flow cytometry (22). The apparent frequency of fluorescent cells via flow cytometry is divided by the enrichment ratio achieved via MACS to achieve the organismal frequency for the T-cell specificity of interest.

Using this method, Moon *et al.* (22) found that when stained with pMHC tetramer, T cells that were specific for the tetramer of interest could be recovered and that these cells could be specifically re-stimulated by the same peptide that was used to isolate the cells in the context of the tetramer. Quantitation of three different foreign peptide specificities led to a naive frequency of ~20-200 CD4⁺ cells per mouse (22). They also found that the extent of T-cell expansion was closely related to the original number of naive T cells (e.g. ~20 to 200) able to recognize the pMHC ligand regardless of peptide dose. Finally, the V-region composition of T cells for a given pMHC ligand does not change between pre- and post-antigen challenge.

This approach has been adapted for studying the CD8⁺ T-cell repertoire, showing a substantially higher naive T-cell precursor frequency for the pMHCs studied, averaging 120-600 cells per mouse (23). The authors point out that at this frequency, the repertoire would rapidly be consumed by a small number of specificities, implying that TCR cross-reactivity is required for complete coverage of all possible pMHC epitopes. The reason for

the difference noted between number of pMHC-reactive cells between CD4⁺ and CD8⁺ T-cell subsets is not currently known. The methodology used for these studies also enabled direct assay of T-cell activation events on a single-cell level during early infection, before the onset of cell proliferation.

A limitation of this technique is that the study of multiple pMHCs at once is limited by the number of available fluorescently labeled streptavidin conjugates. A solution to this problem was demonstrated in a recently developed multiplexing system by Newell *et al.* (24). This system takes advantage of the fact that even though TCR cross-reactivity may be prevalent, cross-reactivity between any two distinct pMHC specificities is a very low probability event. Therefore, combinations of fluorophores can be used to identify a surprisingly large number of unique pMHC specificities. Although the mechanism is beyond the scope of this review, if six variants of labeled streptavidins are used, 63 distinct T-cell populations could be delineated. This technique could be further expanded by coupling to mass cytometry (i.e. CyTof), which allows for the simultaneous measurement of far more parameters than traditional flow cytometry (25, 26). Additionally, there has been some progress using microarray-based approaches, in which multiple pMHCs are immobilized in a spatially defined manner. With this, perhaps combined with other forms of enrichment, large numbers of specificities can be simultaneously observed (27-29).

There is also evidence that while pMHC tetramers have been an invaluable tool, they may underestimate the number of T-cell clones that signal in response to pMHC binding. In some cases, robustly stimulating peptide agonists can fail to stain their cognate T cells in the context of a tetramer (30). This phenomenon may be exacerbated for class II MHC, where issues such as longer peptides and multiple peptide binding registers can cause inefficient staining of T cells (31). The disparate effects of pMHC presentation geometry in streptavidin-arrayed tetramers cannot be discounted, and this is always an unknown variable. It is possible that reactivity could be missed using a streptavidin-based tetramer that would be detected using other arraying scaffolds, and vice versa.

Manipulation of the pMHC-TCR interaction through libraries, screens, and selections

The tripartite nature of the pMHC-TCR interaction, in which the MHC is highly polymorphic and the presented peptide and TCR are hugely diverse repertoires, has rendered systematic reductionist study of the system difficult. Recently, classical biochemical and biophysical approaches to find novel pMHC-TCR reagents have been augmented with new combinatorial biology techniques to create perturbations to the pMHC-TCR system. Much of this progress has been through the creation of libraries of peptides or proteins. In theory, a library can range in scale from a panel of a few dozen peptides to 10¹⁰ unique sequences via phage display. Sequences may be obtained from biologically relevant sources such as peptides verified to bind to MHC, a 'soft' randomization of a known molecule, or a complete randomization of a peptide or stretch of protein. Many library creation, screening, and selection approaches have been brought to bear to interrogate the pMHC-TCR system, allowing for an unparalleled array of reagents to advance our mechanistic understanding of antigen recognition.

Evolution of high-affinity T-cell receptors

Engineering of T-cell receptors is now being carried out nearly with the same facility as the much earlier manipulation of antibodies. While antibodies and TCRs share antigen detection duties in the adaptive immune response, there are key structural and functional differences. Compared to antibodies, which can achieve quite high affinities for their ligands through

somatic hypermutation, TCRs have a restricted antigen set (consisting almost entirely of pMHC) and a much lower affinity range, generally 1-100 μ M (32).

There have been several hypotheses for the biological origin and importance of the low affinity of TCRs. It is possible that the lower affinity has structural origin – the generally poor shape complementarity between TCR and pMHC, and segregated nature of the composite pMHC surface, may result in a structural ‘ceiling’ to pMHC-TCR affinities (33). It is also possible there is a functional imperative to limit TCR affinity. Higher affinity TCR clones may be negatively selected, because the affinity is associated with a greater propensity for cross-reactivity. Several models for TCR signaling, such as the serial triggering of TCR by MHC and kinetic proofreading of activation, incorporate the reality of low affinity interactions with fast kinetics into integral features of robust T-cell activation (34-36).

Several groups have created libraries to isolate higher affinity TCR clones against pMHC of interest. Two techniques have been utilized to create higher affinity TCRs: yeast display and phage display. While using different techniques, different formats for expressing TCR, and different TCR model systems, both techniques have produced TCR variants with drastically higher affinity and provided insight into the TCR structure and function (*Table 1*).

Yeast display of TCRs

Affinity matured murine TCRs have been produced through the evolution of single-chain TCRs (scTCR) displayed on yeast (37). In yeast display, the pMHC of interest is fused to the highly expressed yeast surface protein Aga2p (38, 39). Expression and function of the protein can be tracked via included epitope tags and conformationally specific probes, respectively. Analysis is done using flow cytometry and fluorescently tagged probes. Yeast display combines the strengths of several other techniques. Relatively large library sizes are easily obtainable ($\sim 10^7$ - 10^9 unique clones), the reagent cost is limited, and the turnaround time for creating and selecting a library can be as little as 2-3 weeks. Analogous to scFv antibody libraries, the yeast displayed TCRs consist of the V α and V β domains connected via a GlySer linker (V β -linker-V α), fused to the C-terminus of the yeast protein Aga2p. The initial scTCR created was focused on the 2C TCR, a model system in wide use that has both syn- and allogeneic pMHC ligands. When the 2C scTCR was tested, it initially proved unable to bind either pMHC or a clonotypic antibody against 2C, 1B2 (37). Therefore, the scTCR needed to be evolved for correct folding before the affinity maturation could begin.

The optimization of 2C scTCR occurred in multiple steps. First, the construct was mutagenized through error-prone polymerase chain reaction (PCR) and selected for binding to the 1B2 antibody. Multiple mutations were found that restored 1B2 binding, clustered in the V α /V β interface and the portion of V β that contacts C β in the full length TCR (37). The ability to bind ligand on yeast tracked well with both the yield of secreted scTCR and its thermal stability (40). This property was cleverly applied to select for a more stable version of the scTCR by creating error-prone libraries of the initial scTCR hits and either inducing the cells at 37° instead of the usual 20° or including a high-temperature denaturation step before selections (41). Through these methods, a highly optimized version of 2C scTCR was obtained containing six mutations in the TCR V framework regions (37, 41).

The 2C TCR system is ideal for studying peptide specificity, because there are a range of known peptide ligands with different sequences, affinities, and activities. The display-optimized version of 2C was used to generate TCRs with widely different affinities for the various peptides. Using a relatively small ($\sim 10^5$ transformant) library randomizing only the TCR α CDR3 loop, the authors identified TCR variants with higher affinity for known 2C ligands. Multiple families of CDR3 α sequences were found that recognize QL9-L^d pMHC

with higher affinity than wildtype (WT) 2C (42). These sequence groups are highly diverse, with one containing a conserved Gly at the tip of the CDR loop and the other exhibiting a tandem Pro-Pro-Pro motif (42). This approach was expanded to recover peptide-specific 2C clones for the SIYR and dEV8 bound to K^b. TCR mutants with over 1,000-fold improvement in affinity were isolated, lowering the K_d from 30 μM to 20 nM for SIYR-K^b and 80 μM to 6.5 μM for dEV8-K^b (43). Despite affinity maturation for divergent pMHC and resulting diverging CDR3α sequences, many of the evolved clones retained cross-reactivity for several ligands (43). One potential implication of this unexpected finding is that the low affinity of WT TCRs for pMHC could be because higher affinity TCRs are more prone to cross-reactivity and subsequent deletion through negative selection (43). Nanomolar affinity TCR ligands were also obtained through mutagenesis of the other CDR loops of 2C; optimized CDR1α, CDR1β, and CDR2α each were able to significantly increase affinities while not sacrificing gross peptide specificity (44). However, mutations of CDR1/2 loops alone were not sufficient to recruit a previously nonbinding pMHC to 2C, suggesting that peptide recognition is required regardless of MHC-TCR affinity (44).

The creation of such a wide range of high affinity TCRs for a common set of pMHC provided a unique set of reagents for a surgical look into the biophysics and function of the pMHC-TCR interaction in a manner that was not previously possible. Despite their very slow off rates, even the highest affinity 2C TCR mutants were stimulated by peptide in a dose-dependent manner (43, 45, 46). This result suggested that fast kinetics was not strictly required for T-cell receptor signaling, which forced a serious reconsideration of the serial triggering model. The highest affinity interactions were CD8 independent, while lower affinity pMHC required CD8, which was consistent with previous ideas about co-receptor dependence (46). On a biophysical level, when the pMHC binding thermodynamics were measured for a panel of the high affinity TCRs, a wide range of entropic and enthalpic profiles were found, contrasting with previous notions that TCR recognition of pMHC was largely an enthalpy driven process (47, 48). Thus, collectively, the high affinity 2C series of TCRs were useful for showing us that TCR/pMHC recognition and signaling parameters can be highly diverse and not explainable by restrictive physicochemical rules of engagement.

A second study utilizing scTCR engineered for higher affinity was performed using the murine MHC class II-reactive TCR 3.L2 (49). An antibody specific to 3.L2 did not recognize the scTCR, and the mutations that rendered 2C functional were not transferrable to 3.L2, which has different V region usage (49). Therefore, an analogous error-prone strategy was used to identify a set of mutations for 3.L2 that enabled correct display. Mutations in CDR1α, CDR3α, and CDR3β improved the affinity for its hemoglobin (Hb) peptide complexed to I-E^k from ~15 μM to ~50 nM (49). The mutations grouped by CDR loop were essentially additive in their contribution to increased affinity. Peptide specificity and the ability to respond to antigen were retained for the higher affinity mutants, while there was also an increased tolerance for mutations at any given TCR contact position on the peptide (49, 50). Similar to 2C, increased degeneracy correlated well with increases in TCR affinity (50). The high affinity version of 3.L2 was also recently used in a biophysical study, showing the affinity was gained through more favorable entropy (51).

There have been several other TCRs that recently have been displayed on yeast in different formats, and these experiments demonstrate the highly empirical nature of this approach. The first human scTCRs have been successfully displayed, exhibiting a wide range of outcomes: the 868 TCR was functional without mutation, A6 and CE10 required optimization via error prone mutagenesis, and 1G4 did not gain function even after mutagenesis (52, 53). While it is possible that V-region specific or even pan-TCR mutations may be found in the future to facilitate scTCR production, currently creation of future scTCRs will require validation and optimization of each individual construct. A second

format of TCR displayed on yeast involves the whole TCR extracellular domain (52, 53). One chain of the TCR is tethered to Aga2p, while the second chain is encoded on a second plasmid and secreted. Functional TCR can also be obtained by this manner, but staining is generally less robust than optimized scTCR (52, 53). Additionally, a two-plasmid system would significantly complicate any potential libraries involving simultaneous changes to both TCR chains. As interest in yeast display of TCRs is growing rapidly, no doubt new and clever solutions will emerge for the aforementioned technical challenges.

Phage display of TCR

In addition to the studies using scTCR for yeast display, highly useful TCR libraries have also been produced using phage display. Human TCRs can be produced as a fusion to the gene III product on the surface of M13 phage (54). While multiple TCR formats have been verified for phage display (55, 56), the phage display approach that has gained the most traction adapts the widely used design of a TCR heterodimer linked with an engineered inter-domain disulfide between the C α and C β domains (57). The TCR β exists as the fusion protein while TCR α exists as an independently secreted protein on the vector, relying on pairing of the chains post-translationally. A strong factor in favor of phage display of TCR is that a wide range of model TCR sequences was shown to be both expressed and active in this format, with no apparent limitations with regards to V-region usage or target MHC molecule (54). There also is no apparent need to optimize the TCR through additional mutations to allow for efficient fold apart from the engineered disulfide (57).

Using this system, Li *et al.* (54) created libraries based upon the A6 and 1G4 TCRs. Initial libraries only mutating the CDR3 loops produced TCR variants that could recognize cognate antigen (HLA-A2-Tax for A6, HLA-A2-NY-ESO-1 for 1G4) at ~1-10nM. Further mutagenesis for 1G4 produced a mutant that bound to its pMHC at 26pM, an astounding million-fold increase over WT affinity (measured at 32 μ M). In spite of this increase in affinity, and even though there were significant mutations in the MHC-contacting residues in the CDR2 loops, the evolved TCRs retained the requirement for their cognate peptide for binding to MHC. Indeed, even when only the CDR2 loops are varied, TCR mutants that show significantly higher affinity yet still exhibit peptide specificity can be obtained (58). This presents an interesting contrast to the series of 2C studies that showed increased cross-reactivity of high affinity TCRs (43), again highlighting the recurrent theme that there are few, if any, generalizations that can be made about the physical chemistry of TCR binding, which appears to be as diverse as that seen between antibodies and antigens.

Given the CDR loops are heavily mutated in many of these TCR variants, how does the higher affinity manifest itself structurally? This is especially a matter of interest for the 1G4 mutants, since the highest affinity variants discovered significantly mutate three of the six CDR loops. Recently, crystal structures have been solved for higher affinity variants of both 2C (m6, m13, and m67 recognizing L^d-QL9) and 1G4 (c5c1, c49c50, c58c61, and c58c62, each recognizing HLA-A2-NY-ESO-1)(58-61). Despite increases in affinity of 3 to 6 orders of magnitude, the overall binding topology between each mutant and its parental TCR are essentially identical. Gains in affinity are realized through local rearrangements of the CDR loops, producing greater shape complementarity and buried surface area. This unexpected convergence to the WT sequence's binding footprint is suggestive of the evolutionary influence of the pMHC-TCR germline interactions (60, 62).

Applications for high affinity TCRs

Combinatorial manipulation of TCRs is not only for academic purposes. High-affinity TCRs are being actively studied as a method to control disease in an antigen-specific manner. Promising results have been found for cancer and human immunodeficiency virus (HIV)

(63-65). For cancer, the 1G4 CDR2 and CDR3 mutants developed via phage display were assessed for their ability to specifically respond to cells expressing the NY-ESO-1 peptide (63). Even though all clones only stain with the cognate pMHC tetramer, it was found that the highest affinity variants responded to HLA-A2 positive cells independently of peptide (63). It appears that even though in solution the pMHC-TCR interaction seems specific, HLA-A2 in the context of a cell surface and/or the presence of coreceptors is sufficient to cause peptide-independent activity (63). When CD8 is blocked or when the high affinity TCRs are expressed in CD4⁺ cells, peptide specificity is recovered. Thus, a future strategy may be to focus on the peptide-specific CDR3 loops even at the cost of more modest gains in affinity.

A similar approach was taken in an HIV system. T-cell clones against the HIV gag-derived peptide SL9 were engineered to achieve picomolar affinity (65). At such high affinities, these clones are similar enough to antibodies that they may prove useful as a method to target a toxic payload to virally infected cells. T-cell clones expressing the high-affinity TCRs also showed the ability to cross-react with common SL9 escape mutations, suggesting the possibility that these ultra high affinity TCRs could be particularly useful in cases of a highly variable virus such as HIV (65).

Soluble high affinity receptors have also been useful as reagents for specific targeting to a pMHC of interest. They can serve as imaging probes to find cancer antigens with high specificity and sensitivity, able to detect as low as ~10 peptides presented per cell (66, 67). TCRs engineered for preferential binding to bacterial superantigens rather than to a pMHC provide a potential therapeutic to prevent undue T-cell activation (68-71). There has also been a recent paper in which high affinity soluble TCR is fused to an anti-CD3 scFv, producing a soluble molecule that can potentially repurpose the extant immune response to target tumor antigen-expressing cells (64). This approach, termed immune-mobilizing monoclonal TCRs against cancer (ImmTACs), is especially promising, as there is no need for manipulating T cells to express a different receptor.

There has been work showing that high affinity TCRs can be repurposed as a method to specifically target viruses to cells displaying a pMHC of interest (72). TCR-directed infection of viruses potentially could be used as a method to target gene therapies to only cells expressing a protein of interest. Clearly, as we gain a better handle on the technical solutions, TCR engineering is poised to break through as a viable strategy for developing human therapeutics, it appears only a matter of time.

Discovery of novel peptide ligands

While it is possible to find many T-cell clones that recognize a given pMHC antigen through pMHC tetramer staining, there lacks a corresponding way to find multiple pMHC ligands for a given TCR. Collections of peptides with varying sequence and affinity would be a powerful means of systematically testing pMHC-TCR structure/function relationships. The ability to find pMHC ligands for an arbitrary TCR also may prove profoundly useful in 'de-orphanizing' T cells that are resident in cancer, autoimmune diseases, infection, or even TCRs from T-cell subsets where we do not know the role of specific antigen recognition, such as regulatory T cells (73). Knowledge of a given T cell's ligand could be useful as a biomarker or potential avenue to a useful therapeutic, especially combined with the high affinity TCR approaches discussed above.

Several methods have been applied to discover peptide ligands reactive with TCRs when presented by MHC. Due to the extreme number of possible peptides that can be presented by MHC (a nonameric peptide, the typical peptide displayed on a class I MHC, has 209 or 5.1×10^{11} potential sequences; class II peptides tend to be longer and therefore contain even

more diversity), no system has the capability to systematically look at all peptides. Instead, tradeoffs must be made between comprehensively looking at a limited sequence space and looking at a diverse array of sequences without complete coverage of the set. Both of these approaches have been applied to generate peptides that provide new information about TCR recognition and function.

Altered peptide ligands

The most conservative approach to identify new peptide ligands for a given TCR is to generate a panel of substituted peptides based upon a known ligand and assay them for differences in affinity or activity. These peptides, called altered peptide ligands (APLs) (74, 75), have been extensively used to tease out potential binding 'hotspot' residues (*Fig. 1A*). Researchers have been able to determine information about regions of a peptide where degeneracy is tolerated compared to residues that are immutable while retaining binding. Changes in the potency or composition of the T-cell response can also be measured.

Such directed mutagenesis studies, conceptually similar to alanine scanning, have over the past two decades produced useful data for model systems too numerous to count (76). As just one of many examples, an APL was described using the Hb peptide bound to I-E^k, where a single amino acid substitution was able to produce a pMHC-TCR interaction capable of inducing a cytokine response while not causing T-cell proliferation, showing a decoupling of canonical T-cell effector function (77). This change in APL function was later linked with directly measurable differences in TCR signaling proximal to the membrane, such as differences in CD3 ζ phosphorylation ratios (78, 79). APLs were also used to show that a peptide can specifically induce positive selection of a T-cell clone (80, 81).

APLs have also been useful for studying the structure/function relationships for pMHC-TCR interactions. In a study conducted using the A6 TCR binding to HLA-A2 by Wiley and colleagues (82), APLs are identified that show a vast decoupling between 3D affinity as measured by Surface Plasmon Resonance (SPR), effector functions, and structure. The Tax peptide was compared to an APL, V7R. Tax and V7R only differed by approximately sevenfold in KD and fivefold in off rate. Through this modest difference in binding (both peptides certainly fall within the agonist range of pMHC-TCR interactions), V7R exhibits a 2-log difference in effector functions such as cytokine production and cell lysis. The observed deviation in function cannot be explained by a wholesale difference in structure, as the complexes solved by X-ray crystallography are essentially identical.

There are two major limitations to using APLs for the study of pMHC-TCR structure/function relationships. First, APLs require the directed creation and testing of single site mutants of a known agonist peptide. Thus, it is very difficult to stray far from the properties of the cognate peptide in a rational manner. Changes to residues unimportant to pMHC-TCR binding or mutations that completely ablate the interaction are possible. Even if the design of APLs is based upon known pMHC-TCR complex structure, there is no method available to lead an investigator to a peptide sequence or mutant that will possess a particular property of interest. This can be somewhat mitigated through a systematic point mutation scan of the entire peptide, but the process of individually producing and screening ~200 peptides is laborious (83).

APLs have distinct limitations in that the focus is on a known agonist sequence and APL sequences generally closely resemble that of the original agonist, so it is very much 'searching for keys under the lamp-post'. While APLs have allowed for the discovery of novel pMHC-TCR functions, it would be quite surprising to find a point mutant capable of dramatically altering the mechanism of recognition of an APL by a TCR. If one wants to truly stress the hypotheses of TCR cross-reactivity or the functional implications of altered

pMHC-TCR docking geometries, finding peptide sequences unrelated to the known ligand is crucial.

Synthetic peptide and 'split pool' libraries

An earlier approach that expanded upon the identification and characterization of peptides for a given MHC-TCR system was to create libraries of synthetically derived peptides and screen them for a desired activity. Key to this method is that individual peptides are not initially screened for activity. Instead, the inherent sensitivity of T cells for activating ligands is used to detect low abundance agonist sequences among a mixture of peptides. Subsets of the peptide library as a whole are split into pools containing a defined, limited subset of the library. These so-called 'split-pool' libraries are then assayed for activity with increasingly finer granularity until individual sequences or convergent families of sequences can be extracted (84, 85) (*Fig. 2A*).

Two general methods have been applied for the creation of split-pool libraries. In the most classical approach that was key to establishing paradigms for peptide binding to MHC, pools of peptides derived from a known source were eluted from MHC. In this fashion, peptides from MHC were sequenced as a pool as a method to determine the MHC binding epitopes for such alleles as HLA-A2, -B8, -B53, and -DR1 (86-89). This approach succeeded in defining the preferred anchor residues for many different allotypes of MHC. Positions that have a clear preference for a given amino acid are presumably important for MHC binding, as this measurement is independent of any TCR binding. Such studies have been conducted with naturally processed peptides derived from the proteome and synthetic peptide libraries (86).

To extend beyond pools of endogenous peptides eluted from MHC, a limited library approach can be combined and used to identify a peptide whose existence is known based on the ability of an HPLC fraction to activate a T cell, but exact sequence identity is not. Tallquist and colleagues (90, 91) used a combination of Edman degradation of eluted peptides, peptide libraries, and assay of HPLC-fractionated peptides to identify a peptide agonist for the 2C TCR, dEV8. The libraries were crucial for resolving sequencing ambiguities from the active HPLC fraction eluted from H2-K^b expressing cells and revealed a peptide derived from the mouse proteome that showed markedly different activity than other studied 2C TCR ligands (91).

Extending beyond the delineation of MHC binding motifs, a TCR agonistic peptide sequence approach can be taken that relies on split pool synthesis. One method is to utilize bead-based solid phase peptide synthesis to create a library where each bead only contains one peptide sequence. Through T-cell activation assays, it is possible to isolate a pool of beads to find active peptide ligands and iterate the process until a single peptide can be identified. This has been used to successfully find peptide agonists for a TCR with no known ligand (92, 93). A drawback of this technique is that the number of peptides tested was only 8×10^6 , a relatively small number for naive peptide agonist scouting.

In a second technique, each position is kept constant for a given pool with all other peptide positions being fully randomized (94). For example, to find the amino acid specificity at P1, 20 peptide pools could be created in which each P1 has a known identity with P2-P9 a random mixture of all other amino acids. With this approach, each pool is quite large (1.7×10^{10} for a 9mer peptide library) so recovery of individual sequences of interest for any given pool is not possible. Instead, preferences for each position are then combinatorially combined to create potential consensus sequences to be tested. Unlike point mutation scans, even those that may be entirely comprehensive (83), the simultaneous variance of every

position in theory allows for new consensus sequences entirely unrelated to any known agonist peptide.

Split pool libraries have been used to successfully find new sequence information on multiple occasions, several of which are highlighted here. The peptide SIYR, discussed extensively above in the study of the 2C TCR, was isolated through the creation of a split-pool library on K^b (95). There are multiple additional cases where split pool libraries were utilized to determine the sequence of a peptide agonist for TCRs with unknown ligands (96-98). A similar approach was applied to a class I system using HLA-A2 with HIV Gag epitopes, exploring both the cross-reactivity of the T-cell clone studied and the ability to find stronger agonist peptides related to Gag to invoke a potentially stronger response for vaccination against HIV (99, 100). Use of a split-pool library with HLA-A2 was also used in a study for cross-reactivity of autoimmune TCRs. Combined with mathematical modeling, the authors extrapolated that the 1E6 TCR was capable of reacting to $\sim 1 \times 10^6$ sequences with EC50 comparable to its known insulin peptide. Further, they found that the highest activity ligands were a full 100-fold more potent than the natural ligand (101). Class II libraries were used for the I-E^k-MCC and HLA-DR2-MBP TCR systems to identify the optimal TCR and MHC binding epitopes, as well as show the interplay of tolerance for mutations at different positions and cross-reactivity between peptides (102). Additionally, split pool libraries were used to determine the degeneracy and optimal binding motif for the 172.10 TCR, which recognizes I-A^u-MBP, finding peptides able to activate 172.10 with a six log difference in potency (103).

A modification of this approach is to use iterated rounds of peptide libraries to construct a sequence for a strong agonist. This method sequentially locks in peptide positions that show an increase in activity over baseline and then proceeds to randomize other positions. By this technique, a peptide nine orders of magnitude more potent than the initial library was isolated for a T-cell reactive against an unknown peptide from a mouse lymphoma line (104). A combination of this method and more traditional split pool screening was used to find the optimal peptide size and composition of peptides for a T-cell clone that recognizes MBP peptide in the context of HLA-DR15. This study then used the idealized recognition motif to identify both self and foreign peptides present in the proteome capable of activating the T cell more potently than MBP, suggesting the role cross-reactivity may have in inciting an autoimmune response (105).

Selections of novel peptide ligands utilizing cell-based selection systems While there has been useful information obtained through structure-based mutagenesis and large-scale screening of peptides, these synthetic techniques have fundamental limitations. Synthetic approaches rely upon sufficient peptide solubility and efficient exogenous loading of peptide into MHC on the APC, potentially leading to many false positives and false negatives. Thus, the mathematically derived estimates of library diversity almost certainly, in practice, vastly overestimate the number of peptides that load and are presented by the MHC on the APC. Also, screening of synthetic libraries is dependent upon the peptide of interest having signal detectable above baseline. If the expected frequency of a hit is low, low-potency ligands may not be efficiently observed. Screening of split-pool libraries also relies upon the readout to be activity based, in which a pMHC can induce signal in a large number of cells even if it is not heavily represented in a library. Detection of direct binding between pMHC and TCR is not tenable in this approach. For these cases, it is advisable to make use of selections that rely on direct TCR/pMHC binding (*Fig. 2B*).

Three systems have been utilized to create pMHC libraries for selections, based upon baculovirus, yeast, and phage display. While the large library sizes obtainable through phage display make it a tempting system to use for pMHC libraries, there are relatively few

examples of robust and functional pMHC displayed on phage (106-110). There is one example of novel peptides being isolated from a phage display library, but this library only expresses peptide and relies upon the addition of exogenous MHC (106). The relatively complex fold of pMHC likely results in inefficient secretion on phage. The vast majority of published results using pMHC display rely on baculovirus and yeast systems.

Discovery of novel peptide ligands using baculovirus display

Baculovirus display of MHC is an elegant methodology developed by John Kappler and colleagues (111) that couples the ability to generate diversity to selections using direct TCR binding. Insect cell culture has long been a favored tool for protein production, because of its balance of ease of use and eukaryotic protein folding machinery. Baculovirus display leverages this through the selection of viral particles or virally infected insect cells for a desired property such as binding to a protein of interest (112, 113).

In baculovirus display, the MHC is modified for retention in the insect cell membrane. It is also necessary for the peptide of interest to be covalently linked to the cell that is expressing it, so as to retrieve the enriched genetic information. To accomplish these steps, the MHC is fused to a transmembrane domain or created as an N-terminal fusion to a VSVG anchor protein (111, 114, 115). Peptides are covalently linked via a Gly-Ser linker to either a soluble $\beta 2m$ molecule in the case of class I MHC or to the MHC β chain for class II (111, 114). There have been two class II designs reported in the literature: either the peptide and both MHC chains in a large single-chain construct, or a bicistronic vector that produces both chains, with leucine zippers to aid in MHC chain pairing (111, 115).

The major benefit of baculovirus display is that the native MHC usually folds into functional molecules without need for mutagenesis. While only a few MHC alleles have been reported as a baculovirus display construct, the fact that the reported functional constructs are relatively disparate (mouse and human, class I and class II) combined with how readily most MHC molecules are produced via baculovirus suggests that baculovirus display can be easily adapted to other MHC (*Table 2*).

Baculovirus display has produced impressive results in several TCR systems. Using the murine class II I-A^b, peptide mimotopes were found for two different TCRs recognizing the p3K peptide, showing a difference in the two TCRs' cross reactivity (111). A D^d baculovirus display library was used to identify a peptide capable of activating a T-cell clone that was known to be D^d reactive without any previously known ligand (114). A mimotope of this peptide was then found in the mouse proteome that was able to also activate the T-cell clone, albeit with lower efficiency (114). For both of these cases, limited diversity was ameliorated through the fixation of residues that contact the MHC to known efficient anchors. This is a potential limitation especially for class I MHC selections, as differential peptide topologies and MHC contact registers are possible based upon altered peptide sequence. Baculovirus display has also been shown to be able to 'optimize' a peptide for a given property of interest. A library can discriminate peptide positions that influence superantigen binding or fine-tune affinity of a TCR with a known ligand through optimization of MHC contacts of the biological ligand (116).

Recently, MacDonald *et al.* (117) utilized baculovirus display to provide insight into T-cell cross-reactivity and alloreactivity, directly plugging the results of their selections into a structure-function study. The authors used a fully randomized peptide library for HLA-B*4405 to find a peptide allotope for the well-studied TCR LC13, which recognizes an epitope from Epstein-Barr virus. The allotope peptide is notable for differing from the EBV 'virotope' peptide in 8/9 positions. The allotope peptide also led to a 'mimotope' peptide encoded in the human genome that also activated LC13. The crystal structures of all three

pMHC-TCR complexes were remarkably similar. The overall docking topology was retained with conserved TCR-MHC germline contacts. The TCR CDR3 loops accommodated the different TCRs through molecular mimicry: each peptide contained a P6-P8 motif of a bulky aromatic residue flanked by two small residues, allowing the TCR to engage each peptide in a highly similar manner.

An additional benefit of baculovirus display of pMHC constructs is that the insect cells themselves are capable of serving as synthetic antigen-presenting cells by inducing a T-cell response. Coexpression of T-cell costimulatory molecules such as ICAM or B7 enhances this effect (111). Indeed, experiments have been conducted using insect cells expressing pMHC as a method to manipulate the immune response (118).

Discovery of novel peptide ligands using yeast display

Yeast surface display is an alternative technique that has recently gained significant traction in finding novel peptides for a given MHC-TCR system (*Table 2*). While each approach may have suitability to specific applications, baculovirus and yeast pMHC display are highly complementary. Yeast may be faster and produce larger libraries, but baculovirus display has the advantage of expressing most MHC without the need to optimize expression.

In the past few years, there have been several MHC molecules successfully displayed on the surface of yeast: principally class II MHC. Murine I-A^{g7} and human HLA-DR1 and HLA-DR15 have been expressed as full-length single-chain molecules with their peptides covalently linked (115, 119, 120). I-A^{g7} and HLA-DR15 required mutations for correct folding of the MHC, while HLA-DR1 was expressed and folded as a native sequence, with level of surface display dependent upon the peptide's affinity for MHC. Yeast-displayed HLA-DR1 exhibited the ability to stimulate T cells expressing a reactive TCR, and *bona fide* HLA-DR1 binding peptides could be isolated from a library created from influenza hemagglutinin protein (120). However, yeast displayed HLA-DR15 and I-A^{g7} did not have the ability or were not tested for the ability to bind to TCR or activate T cells, respectively (115, 119). HLA-DR1 has also been expressed as a single-chain molecule on the surface of yeast without peptide covalently linked (121). Without peptide, HLA-DR1 was not stable until the acquisition of stability mutations. These mutations allowed the stable fold of HLA-DR1 in the absence of peptides for timescales on the order of weeks (121).

Alternative formats for class II yeast display have also been developed: for HLA-DR4, instead of having the peptide and MHC α and β chains expressed as a single chain trimer, involving a lengthy linker between the MHC α and β chains, the MHC α chain is secreted as a soluble protein, free to pair with the Aga2p linked peptide-MHC β single chain fusion post-translationally (122, 123). This technique was furthered through the creation of a peptide-only linkage to Aga2p, secreting HLA-DR1 α and β chains (123). This system was shown to be able to efficiently screen peptides for anchor residue specificity by measuring for the presence of soluble HLA-DR1 relative to fused peptide, and is shown to be a method to create MHC variants with altered anchor specificity (123). When tested, MHC-expressing yeast were unable to stimulate T cells (122). Again, no data were presented regarding the ability of the yeast-displayed molecules to bind TCR (122, 123).

There have been fewer forays into yeast display for class I MHC molecules. Class I MHC has the complication of variable affinity to its accessory molecule β 2m and a 'closed' peptide-binding groove that constrains peptide length, providing no path for a linker between peptide and MHC. There have been two reported class I MHC molecules displayed on the yeast, the murine MHCs K^b and L^d. K^b was expressed as a single chain trimer, with peptide, β 2m, and the α 1 α 3 domains of the class I MHC all linked (124). The protein was active via staining of anti-MHC antibodies, and the yeast could be used to activate T cells in

an antigen-specific manner (124). However, only yeast pulsed with exogenous peptide were able to be stained with TCR; even though the single-chain trimer yeast could be stained via antibody, no TCR staining was seen (124). For L^d, the authors linked the QL9 peptide and used error prone mutagenesis as a method to take a modestly staining initial construct and markedly improve its stability and staining with its cognate TCR (125). Surprisingly, the most robustly staining clone was a truncation mutant of L^d, eliminating the $\alpha 3$ and $\beta 2m$ domains and leaving the $\alpha 1\alpha 2$ MHC 'platform'. For the selection of peptides, the L^d platform was further optimized: through rational mutagenesis, additional mutations were found that increased the stability of the platform (126). We also included a Trp167Ala mutation to open a 'notch' at one end of the peptide-binding groove for the linker. This is a necessary step for covalently linking the peptide to the MHC to minimize the chance the linker would interfere with TCR binding (127).

The fact that the L^d platform was both a very robust reagent and a ligand for the 2C family of TCRs led us to use it for as the basis for our peptide libraries. We were interested in finding as many unique peptides as possible to ask questions about TCR recognition, germline bias, and signaling (62). For example, would starkly different peptides reorient the TCR docking footprint on the MHC, and could we use different peptides to assess the linkage between TCR docking angle, 2D and 3D affinity in TCR signaling? Starting with a robustly displayed MHC on yeast was crucial.

An inescapable reality of any of these library selections is that the achievable library diversity is a vanishingly small fraction of the possible sequence space (in this case, for nonameric peptides). We addressed this potential problem in two ways. First, in each library, the primary MHC anchor residues are allowed very limited sequence diversity via a codon which encodes a small degree of degeneracy (e.g. VTC to only produce the hydrophobic residues Ile, Val, and Leu). Without this step, a large portion of the library would be rendered useless, as the peptides would not bind to MHC let alone serve as potential TCR ligands. Secondly, a series of libraries were developed: one allowing diversity at every non-anchor position, one only mutating the 'TCR contact' residues based upon the 2C-L^d-QL9 crystal structure, and one only mutating the non-anchor 'MHC contact' residues.

The three libraries were selected with the 42F3 TCR (a TCR that natively binds to the QL9 peptide bound to L^d)(128), and each produced a unique result. The TCR contact library produced a new P4-P5 consensus of Asp-Arg as compared to Pro-Phe in QL9. The MHC contact library showed a preference for a Trp at P6, a surprising occurrence given the limited amount of space for such a bulky side chain for a MHC contact. Finally, the totally random library of 10⁸ peptides produced a single peptide sequence, differing from QL9 at all nine positions.

For their differences in sequence, the library-derived peptides all showed affinities via SPR within the range of what is expected for bona fide TCR ligands, from 4 to 50 μ M. When measured for activity, differences became readily apparent. The peptides obtained from the MHC contact library showed a dependence upon CD8 on the T-cell surface for signaling, while the TCR contact peptides robustly signaled in both the presence and absence of coreceptor. This difference cannot be explained through affinity, as both libraries produced relatively high affinity sequences.

Strikingly, a peptide from the completely random library, p3A1, showed no activity via IL-2 release regardless of the presence of coreceptor. This was an unexpected result, given its affinity (3.9 μ M) should be squarely within the bounds of what should be expected for an agonist peptide.

Closer examination revealed some differences in biophysical characteristics that were not readily apparent from the SPR affinity measurements. When performing a '2D' affinity measurement based upon cell adhesion (129), p3A1 had a twofold lower affinity than QL9. The coreceptor independent p4B10 showed a corresponding increase in affinity relative to the canonical peptide. This result correlated with a difference seen when assaying the interaction via tetramer staining of cells. When TCR tetramers were used to stain APCs expressing the library peptides, staining for the various peptides was comparable. However, when the experiment was flipped, with pMHC tetramers staining T cells, staining with p3A1 tetramer was the weakest of all of the ligands. These results collectively suggest part of the nonagonism for p3A1 may be derived from limits imposed in the context of a cell-cell interaction (2D) versus free in solution (3D). When the TCR is readily accessible in solution (as it is in SPR or when being used as a staining reagent), there is a robust binding interaction. When the TCR is in the context of other proteins (on a cell for either 2D affinity measurements or as a target for pMHC tetramer staining), p3A1 shows a markedly reduced binding signal. Such discordance appears to suggest that the TCR, when on the T cell, does not have complete orientational freedom to engage the pMHC, as it does in solution.

This conundrum was partially illuminated when the crystal structures of a representative peptide from each library was solved in complex with 42F3 and L^d (Fig. 3). The TCR and MHC contact peptides, p4B10 and p5E8 respectively, each shared a nearly identical binding topology with the natural QL9 ligand (Fig. 3A). The CDR1 and CDR2 contacts are entirely conserved. To accommodate the different peptides, the TCR CDR3 loops are involved in a unique set of contacts for each peptide (Fig. 3B,C). p5E8, in spite of the library's conservative mutagenesis strategy, showed a drastic difference: in order to accommodate the evolved P6 Trp, the peptide's register flipped. The p5 position became an MHC contact while P6 contacted the TCR (Fig. 3D). That these two peptides maintained an essentially identical binding orientation with comparable affinities while including major differences in peptide sequence is a testament to the cross-reactivity afforded by the TCR CDR3 loops and, in our opinion, to the ability of the germline TCR-MHC contacts to steer TCR binding into its typical canonical footprint.

For the peptide from the random library, p3A1, the story deviates. Far from maintaining the WT orientation, p3A1 causes a drastic shift in both the pMHC-TCR binding angle and center of interface (Figs 3 and 4A). The TCR rotates ~38° and moves ~7Å relative to QL9 and the other library peptides. The canonical germline contacts are also drastically altered: Tyr50α still maintains a contact with the MHC Tyr155, but the orientation and contact network for the other conserved germline contacts seen for 2C and 42F3 TCR ligands are altered.

Nevertheless, the focus of TCR CDR3 loops on the peptide epitope is maintained, as is the general α/β polarity of the complex (Fig. 3). This is a noteworthy feature of the selected peptide: even with a randomly generated peptide that causes TCR docking to diverge, and even though selections were conducted in the absence of coreceptor or any other potential biasing agent, the binding footprint only diverges so far. *A priori*, there is no reason a peptide epitope could not be found that flips the TCR polarity or even binds to another prominent surface-exposed epitope on the TCR such as the TCR Cβ FG loop. We believe that the fact that even at its most deviant the pMHC-TCR interaction remains within a series of recognizable norms is a strong argument for the influence of evolutionary conserved germline contacts.

Given that the search for novel TCR ligands has been underway for decades, it is fair to ask why high affinity/no activity ligands have not been previously characterized. We can offer a few theories. First, many of the studies to find and characterize APLs or other peptide

ligands have relied upon functional activity as either the primary readout of a hit or as an essential validation step. Needless to say, these approaches by their very design will be blind to peptides not able to induce a signal. Second, our selections were conducted in a system free from the confines of co-receptors selections done in the presence of CD3/4/8 could potentially diminish the signal of a ligand with an aberrant binding pattern. Third, yeast display allowed us a library design and size that allowed for the selection from a hugely diverse array of peptides. Even with a library of 10^8 unique clones heavily randomized at 7 out of 9 positions, our selections returned a single peptide that produced this binding orientation. Given even these large libraries hugely undersample the theoretical diversity, it is entirely possible that other libraries simply did not contain the diversity necessary to find strongly binding nonagonist peptides, or that the library was 'unlucky' in its sequence coverage. Finally, we believe that the careful construction and optimization of the MHC used as the basis of selections to be crucial: the L^d platform went through multiple steps of library mutagenesis and rational design before used for these studies, with careful consideration applied to ensure the peptide and TCR contact surfaces of the MHC remained as close to WT as possible. The careful creation of robust yet biologically relevant MHC scaffolds will be a crucial determinant for future library-based experiments.

Biological implications of high affinity nonagonists

The discovery of a new class of peptide, nonstimulatory high-affinity pMHC ligands, leads to a series of questions with potentially profound implications (*Fig. 4A,B*). (i) At which point in the signaling cascade does the nonstimulating peptide diverge from agonist peptides with comparable biophysical properties? Is the defect manifested at the cell surface? Or is this effect due to a decoupling of multiple TCR signaling pathways, perhaps through a nonstandard phosphorylation of CD3 ITAMs, from a delay in CD3 phosphorylation, allowing resident phosphatases to squelch signaling before it can even start, or from a complete lack of TCR activation, indicating that the binding of these nonstimulating peptides produces a signal that is qualitatively completely divergent from an agonist?

(ii) Biophysically, what causes the observed signaling differences? We have previously discussed some possibilities for this that generally fall into two categories. First, it is possible that even though the SPR affinities and kinetics between agonist and nonagonist peptides are comparable, the complexes behave differently in the context of the physiological pMHC-TCR ultrastructure that exists between a T cell and APC (*Fig. 4C*). These differences may arise because of a conformation that disfavors receptor oligomerization, prevents engagement of co-receptors, or partially occludes the binding site. Experimental evidence exists, as the 2D kinetic measurements for the agonist and nonagonist peptides show a divergence in affinities and on rates that was not apparent via SPR (127). It is also possible that even if the pMHC-TCR interaction is not biophysically distinct in the context of cells, this interaction cannot produce a productive signaling output. There is a significant body of literature demonstrating that the adage of 'any dimer will do' does not necessarily apply to all receptor-ligand systems (130). Studies using anti-TCR or – CD3 antibodies show that antibody crosslinking produces a range of signaling potencies (with some notable examples not signaling at all), presumably based upon what epitope is recognized (131).

(iii) Can the nonagonist peptides perform secondary signaling functions? Peptides that do not directly serve as agonists have been shown to be functional as co-agonists in pMHC 'pseudodimers' (combining a pMHC agonist with a pMHC null peptide), providing a signaling response even if the coagonist has no detectable affinity for the cognate TCR (3, 4, 132). Additionally, low affinity pMHC interactions serve as a source of tonic TCR signaling in the periphery (5). Given both of these roles are presumably able to be accommodated by a

large variety of null peptides, it will be informative to see, for example, if agonist/nonagonist pMHC pseudodimers can activate T-cell signaling.

(iv) What are the immunological implications of high-affinity nonagonist pMHC/TCR interactions? The peptides found through the yeast display library were discovered through assaying for binding instead of activity in a system devoid of potential geometric restraints imposed by cellular superstructures or co-receptors. It is likely this method allows for the discovery of nonagonist peptides that would not otherwise be found in functional based schemes. The p3A1 peptide itself does not exist in nature; however, the primary amino acid sequence itself is unremarkable and relatively close analogues do exist in various microbial organisms. This begs the question – what becomes of the immune response if an immunodominant epitope of a microbial challenge can produce this nonstimulating phenotype?

Conclusion and future prospective

The immense combinatorial diversity of the adaptive immune system has been both a great source of scientific interest and a hindrance in defining general rules for its ability to recognize and respond to antigen. We believe that an embrace of this diversity has begun to allow for better understanding than a mosaic of model systems could ever allow. Quantitating and assaying the natural immune repertoire as a whole has made great advances for establishing the sequence, scope, and specificity of the immune response. Generating diversity through peptide and TCR libraries has allowed for rigorous testing of hypotheses of structure and function that would not otherwise be possible.

The information provided by diversity-oriented approaches will provide the ability to assess and then manipulate the interplay between the immune response and disease where the causative antigens are not known. Changes in the immune response can be measured as a whole, with peptide ligands then able to be rapidly identified and perhaps even used as a diagnostic or treatment. Creation and discovery of pMHC and TCR variants directly in the context of a native T cell-APC system may eventually be able to expedite this process (133-135).

We also believe that the techniques and reagents being developed in the studies detailed here have great synergy with advances in biophysical techniques that have allowed for a more detailed mechanistic understanding of antigen receptor function. As one recent example, affinity maturation of human CD4 for class II MHC enabled the first structure determination of a complete TCR-pMHC-CD4 ternary complex (136, 137). We speculate that harnessing the full complement of diversity-based approaches to interrogate T-cell biology is only just beginning.

Acknowledgments

MEB is supported by National Science Foundation and Stanford Graduate pre-doctoral fellowships. KCG is supported by NIH R01-AI48540 and is an Investigator of the Howard Hughes Medical Institute. The authors have no conflicts of interest to declare.

References

1. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature*. 1988; 334:395–402. [PubMed: 3043226]
2. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol*. 2006; 24:419–466. [PubMed: 16551255]

3. Irvine DJ, Purbhoo MA, Krogsgaard M, Davis MM. Direct observation of ligand recognition by T cells. *Nature*. 2002; 419:845–849. [PubMed: 12397360]
4. Juang J, Ebert PJ, Feng D, Garcia KC, Krogsgaard M, Davis MM. Peptide-MHC heterodimers show that thymic positive selection requires a more restricted set of self-peptides than negative selection. *J Exp Med*. 2010; 207:1223–1234. [PubMed: 20457759]
5. Hochweller K, Wabnitz GH, Samstag Y, Suffner J, Hammerling GJ, Garbi N. Dendritic cells control T cell tonic signaling required for responsiveness to foreign antigen. *Proc Natl Acad Sci USA*. 2010; 107:5931–5936. [PubMed: 20231464]
6. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*. 2008; 26:1135–1145. [PubMed: 18846087]
7. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet*. 2008; 9:387–402. [PubMed: 18576944]
8. Weinstein JA, Jiang N, White RA 3rd, Fisher DS, Quake SR. High-throughput sequencing of the zebrafish antibody repertoire. *Science*. 2009; 324:807–810. [PubMed: 19423829]
9. Warren RL, et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res*. 2011; 21:790–797. [PubMed: 21349924]
10. Robins HS, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood*. 2009; 114:4099–4107. [PubMed: 19706884]
11. Arstila TP, Casrouge A, Baron V, Even J, Kanellopoulos J, Kourilsky P. A direct estimate of the human alphabeta T cell receptor diversity. *Science*. 1999; 286:958–961. [PubMed: 10542151]
12. Freeman JD, Warren RL, Webb JR, Nelson BH, Holt RA. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res*. 2009; 19:1817–1824. [PubMed: 19541912]
13. Klarenbeek PL, et al. Human T-cell memory consists mainly of unexpanded clones. *Immunol Lett*. 2010; 133:42–48. [PubMed: 20621124]
14. Robins HS, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire. *Sci Transl Med*. 2010; 2:47ra64.
15. Venturi V, et al. A mechanism for TCR sharing between T cell subsets and individuals revealed by pyrosequencing. *J Immunol*. 2011; 186:4285–4294. [PubMed: 21383244]
16. Wang C, et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc Natl Acad Sci USA*. 2010; 107:1518–1523. [PubMed: 20080641]
17. Huseby ES, et al. How the T cell repertoire becomes peptide and MHC specific. *Cell*. 2005; 122:247–260. [PubMed: 16051149]
18. Logunova NN, et al. Restricted MHC-peptide repertoire predisposes to autoimmunity. *J Exp Med*. 2005; 202:73–84. [PubMed: 15998789]
19. Yin L, et al. A single T cell receptor bound to major histocompatibility complex class I and class II glycoproteins reveals switchable TCR conformers. *Immunity*. 2011; 35:23–33. [PubMed: 21683626]
20. Davis MM, Altman JD, Newell EW. Interrogating the repertoire: broadening the scope of peptide-MHC multimer analysis. *Nat Rev Immunol*. 2011; 11:551–558. [PubMed: 21760610]
21. Jenkins MK, Moon JJ. The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. *J Immunol*. 2012; 188:4135–4140. [PubMed: 22517866]
22. Moon JJ, et al. Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity*. 2007; 27:203–213. [PubMed: 17707129]
23. Obar JJ, Khanna KM, Lefrancois L. Endogenous naive CD8+ T cell precursor frequency regulates primary and memory responses to infection. *Immunity*. 2008; 28:859–869. [PubMed: 18499487]
24. Newell EW, Klein LO, Yu W, Davis MM. Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. *Nat Methods*. 2009; 6:497–499. [PubMed: 19543286]
25. Bendall SC, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science*. 2011; 332:687–696. [PubMed: 21551058]

26. Newell EW, Sigal N, Bendall SC, Nolan GP, Davis MM. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity*. 2012; 36:142–152. [PubMed: 22265676]
27. Chen DS, et al. Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. *PLoS Med*. 2005; 2:e265. [PubMed: 16162034]
28. Ge X, et al. Peptide-MHC cellular microarray with innovative data analysis system for simultaneously detecting multiple CD4 T-cell responses. *PLoS One*. 2010; 5:e11355. [PubMed: 20634998]
29. Kwong GA, et al. Modular nucleic acid assembled p/MHC microarrays for multiplexed sorting of antigen-specific T cells. *J Am Chem Soc*. 2009; 131:9695–9703. [PubMed: 19552409]
30. Sabatino JJ Jr. Huang J, Zhu C, Evavold BD. High prevalence of low affinity peptide-MHC II tetramer-negative effectors during polyclonal CD4+ T cell responses. *J Exp Med*. 2011; 208:81–90. [PubMed: 21220453]
31. Landais E, et al. New design of MHC class II tetramers to accommodate fundamental principles of antigen presentation. *J Immunol*. 2009; 183:7949–7957. [PubMed: 19923463]
32. Davis MM, et al. Ligand recognition by alpha beta T cell receptors. *Annu Rev Immunol*. 1998; 16:523–544. [PubMed: 9597140]
33. Garcia KC, et al. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science*. 1998; 279:1166–1172. [PubMed: 9469799]
34. Valitutti S, Müller S, Cella M, Padovan E, Lanzavecchia A. Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature*. 1995; 375:148–151. [PubMed: 7753171]
35. Valitutti S, Lanzavecchia A. Serial triggering of TCRs: a basis for the sensitivity and specificity of antigen recognition. *Immunol Today*. 1997; 18:299–304. [PubMed: 9190117]
36. McKeithan TW. Kinetic proofreading in T-cell receptor signal transduction. *Proc Natl Acad Sci USA*. 1995; 92:5042–5046. [PubMed: 7761445]
37. Kiecke MC, Shusta EV, Boder ET, Teyton L, Wittrup KD, Kranz DM. Selection of functional T cell receptor mutants from a yeast surface-display library. *Proc Natl Acad Sci USA*. 1999; 96:5651–5656. [PubMed: 10318939]
38. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD. Isolating and engineering human antibodies using yeast surface display. *Nat Protoc*. 2006; 1:755–768. [PubMed: 17406305]
39. Boder ET, Wittrup KD. Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol*. 1997; 15:553–557. [PubMed: 9181578]
40. Shusta EV, Kiecke MC, Parke E, Kranz DM, Wittrup KD. Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J Mol Biol*. 1999; 292:949–956. [PubMed: 10512694]
41. Shusta EV, Holler PD, Kiecke MC, Kranz DM, Wittrup KD. Directed evolution of a stable scaffold for T-cell receptor engineering. *Nat Biotechnol*. 2000; 18:754–759. [PubMed: 10888844]
42. Holler PD, Holman PO, Shusta EV, O'Herrin S, Wittrup KD, Kranz DM. In vitro evolution of a T cell receptor with high affinity for peptide/MHC. *Proc Natl Acad Sci USA*. 2000; 97:5387–5392. [PubMed: 10779548]
43. Holler PD, Chlewicki LK, Kranz DM. TCRs with high affinity for foreign pMHC show self-reactivity. *Nat Immunol*. 2003; 4:55–62. [PubMed: 12469116]
44. Chlewicki LK, Holler PD, Monti BC, Clutter MR, Kranz DM. High-affinity, peptide-specific T cell receptors can be generated by mutations in CDR1, CDR2 or CDR3. *J Mol Biol*. 2005; 346:223–239. [PubMed: 15663940]
45. Holler PD, Lim AR, Cho BK, Rund LA, Kranz DM. CD8(-) T cell transfectants that express a high affinity T cell receptor exhibit enhanced peptide-dependent activation. *J Exp Med*. 2001; 194:1043–1052. [PubMed: 11602635]
46. Holler PD, Kranz DM. Quantitative analysis of the contribution of TCR/pepMHC affinity and CD8 to T cell activation. *Immunity*. 2003; 18:255–264. [PubMed: 12594952]
47. Jones LL, et al. Different thermodynamic binding mechanisms and peptide fine specificities associated with a panel of structurally similar high-affinity T cell receptors. *Biochemistry*. 2008; 47:12398–12408. [PubMed: 18973345]

48. Willcox BE, et al. TCR binding to peptide-MHC stabilizes a flexible recognition interface. *Immunity*. 1999; 10:357–365. [PubMed: 10204491]
49. Weber KS, Donermeyer DL, Allen PM, Kranz DM. Class II-restricted T cell receptor engineered in vitro for higher affinity retains peptide specificity and function. *Proc Natl Acad Sci USA*. 2005; 102:19033–19038. [PubMed: 16365315]
50. Donermeyer DL, Weber KS, Kranz DM, Allen PM. The study of high-affinity TCRs reveals duality in T cell recognition of antigen: specificity and degeneracy. *J Immunol*. 2006; 177:6911–6919. [PubMed: 17082606]
51. Persaud SP, Donermeyer DL, Weber KS, Kranz DM, Allen PM. High-affinity T cell receptor differentiates cognate peptide-MHC and altered peptide ligands with distinct kinetics and thermodynamics. *Mol Immunol*. 2010; 47:1793–1801. [PubMed: 20334923]
52. Richman SA, et al. Structural features of T cell receptor variable regions that enhance domain stability and enable expression as single-chain ValphaVbeta fragments. *Mol Immunol*. 2009; 46:902–916. [PubMed: 18962897]
53. Aggen DH, Chervin AS, Insaiddo FK, Piepenbrink KH, Baker BM, Kranz DM. Identification and engineering of human variable regions that allow expression of stable single-chain T cell receptors. *Protein Eng Des Sel*. 2011; 24:361–372. [PubMed: 21159619]
54. Li Y, et al. Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat Biotechnol*. 2005; 23:349–354. [PubMed: 15723046]
55. Weidanz JA, Card KF, Edwards A, Perlstein E, Wong HC. Display of functional alphabeta single-chain T-cell receptor molecules on the surface of bacteriophage. *J Immunol Methods*. 1998; 221:59–76. [PubMed: 9894898]
56. Loset GA, Lunde E, Bogen B, Brekke OH, Sandlie I. Functional phage display of two murine alpha/beta T-cell receptors is strongly dependent on fusion format, mode and periplasmic folding assistance. *Protein Eng Des Sel*. 2007; 20:461–472. [PubMed: 17925331]
57. Boulter JM, et al. Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Eng*. 2003; 16:707–711. [PubMed: 14560057]
58. Dunn SM, et al. Directed evolution of human T cell receptor CDR2 residues by phage display dramatically enhances affinity for cognate peptide-MHC without increasing apparent cross-reactivity. *Protein Sci*. 2006; 15:710–721. [PubMed: 16600963]
59. Sami M, et al. Crystal structures of high affinity human T-cell receptors bound to peptide major histocompatibility complex reveal native diagonal binding geometry. *Protein Eng Des Sel*. 2007; 20:397–403. [PubMed: 17644531]
60. Colf LA, et al. How a single T cell receptor recognizes both self and foreign MHC. *Cell*. 2007; 129:135–146. [PubMed: 17418792]
61. Jones LL, Colf LA, Stone JD, Garcia KC, Kranz DM. Distinct CDR3 conformations in TCRs determine the level of cross-reactivity for diverse antigens, but not the docking orientation. *J Immunol*. 2008; 181:6255–6264. [PubMed: 18941216]
62. Garcia KC, Adams JJ, Feng D, Ely LK. The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nat Immunol*. 2009; 10:143–147. [PubMed: 19148199]
63. Zhao Y, et al. High-affinity TCRs generated by phage display provide CD4+ T cells with the ability to recognize and kill tumor cell lines. *J Immunol*. 2007; 179:5845–5854. [PubMed: 17947658]
64. Liddy N, et al. Monoclonal TCR-redirected tumor cell killing. *Nat Med*. 2012; 18:980–987. [PubMed: 22561687]
65. Varela-Rohena A, et al. Control of HIV-1 immune escape by CD8 T cells expressing enhanced T-cell receptor. *Nat Med*. 2008; 14:1390–1395. [PubMed: 18997777]
66. Purbhoo MA, et al. Quantifying and imaging NY-ESO-1/LAGE-1-derived epitopes on tumor cells using high affinity T cell receptors. *J Immunol*. 2006; 176:7308–7316. [PubMed: 16751374]
67. Zhang B, et al. Induced sensitization of tumor stroma leads to eradication of established cancer by T cells. *J Exp Med*. 2007; 204:49–55. [PubMed: 17210731]
68. Buonpane RA, et al. Neutralization of staphylococcal enterotoxin B by soluble, highaffinity receptor antagonists. *Nat Med*. 2007; 13:725–729. [PubMed: 17515896]

69. Moza B, et al. Structural basis of T-cell specificity and activation by the bacterial superantigen TSST-1. *The EMBO J.* 2007; 26:1187–1197.
70. Buonpane RA, Moza B, Sundberg EJ, Kranz DM. Characterization of T cell receptors engineered for high affinity against toxic shock syndrome toxin-1. *J Mol Biol.* 2005; 353:308–321. [PubMed: 16171815]
71. Kieke MC, Sundberg E, Shusta EV, Mariuzza RA, Wittrup KD, Kranz DM. High affinity T cell receptors from yeast display libraries block T cell activation by superantigens. *J Mol Biol.* 2001; 307:1305–1315. [PubMed: 11292343]
72. Peng KW, Holler PD, Orr BA, Kranz DM, Russell SJ. Targeting virus entry and membrane fusion through specific peptide/MHC complexes using a high-affinity T-cell receptor. *Gene Ther.* 2004; 11:1234–1239. [PubMed: 15215883]
73. Hsieh CS, Lee HM, Lio CW. Selection of regulatory T cells in the thymus. *Nat Rev Immunol.* 2012; 12:157–167. [PubMed: 22322317]
74. Sloan-Lancaster J, Allen PM. Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. *Annu Rev Immunol.* 1996; 14:1–27. [PubMed: 8717505]
75. Evavold BD, Sloan-Lancaster J, Allen PM. Tickling the TCR: selective T-cell functions stimulated by altered peptide ligands. *Immunol Today.* 1993; 14:602–609. [PubMed: 8305133]
76. Unanue ER. Altered peptide ligands make their entry. *J Immunol.* 2011; 186:7–8. [PubMed: 21172872]
77. Evavold BD, Allen PM. Separation of IL-4 production from Th cell proliferation by an altered T cell receptor ligand. *Science.* 1991; 252:1308–1310. [PubMed: 1833816]
78. Sloan-Lancaster J, Shaw AS, Rothbard JB, Allen PM. Partial T cell signaling: altered phospho-zeta and lack of zap70 recruitment in APL-induced T cell anergy. *Cell.* 1994; 79:913–922. [PubMed: 8001128]
79. Sloan-Lancaster J, Allen PM. Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. *Annu Rev Immunol.* 1996; 14:1–27. [PubMed: 8717505]
80. Hogquist KA, Jameson SC, Heath WR, Howard JL, Bevan MJ, Carbone FR. T cell receptor antagonist peptides induce positive selection. *Cell.* 1994; 76:17–27. [PubMed: 8287475]
81. Hedrick SM. Positive selection in the thymus: an enigma wrapped in a mystery. *J Immunol.* 2012; 188:2043–2045. [PubMed: 22345700]
82. Ding YH, Baker BM, Garboczi DN, Biddison WE, Wiley DC. Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. *Immunity.* 1999; 11:45–56. [PubMed: 10435578]
83. Reay PA, Kantor RM, Davis MM. Use of global amino acid replacements to define the requirements for MHC binding and T cell recognition of moth cytochrome c (93-103). *J Immunol.* 1994; 152:3946–3957. [PubMed: 7511662]
84. Pinilla C, et al. Exploring immunological specificity using synthetic peptide combinatorial libraries. *Curr Opin Immunol.* 1999; 11:193–202. [PubMed: 10322159]
85. Sospedra M, Pinilla C, Martin R. Use of combinatorial peptide libraries for T-cell epitope mapping. *Methods.* 2003; 29:236–247. [PubMed: 12725789]
86. Davenport MP, et al. HLA class I binding motifs derived from random peptide libraries differ at the COOH terminus from those of eluted peptides. *J Exp Med.* 1997; 185:367–371. [PubMed: 9016886]
87. Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature.* 1991; 351:290–296. [PubMed: 1709722]
88. Stevens J, Wiesmuller KH, Barker PJ, Walden P, Butcher GW, Joly E. Efficient generation of major histocompatibility complex class I-peptide complexes using synthetic peptide libraries. *J Biol Chem.* 1998; 273:2874–2884. [PubMed: 9446598]
89. Fleckenstein B, et al. New ligands binding to the human leukocyte antigen class II molecule DRB1*0101 based on the activity pattern of an undecapeptide library. *Eur J Biochem.* 1996; 240:71–77. [PubMed: 8797837]
90. Tallquist MD, Yun TJ, Pease LR. A single T cell receptor recognizes structurally distinct MHC/peptide complexes with high specificity. *J Exp Med.* 1996; 184:1017–1026. [PubMed: 9064319]

91. Degano M, Garcia KC, Apostolopoulos V, Rudolph MG, Teyton L, Wilson IA. A functional hot spot for antigen recognition in a superagonist TCR/MHC complex. *Immunity*. 2000; 12:251–261. [PubMed: 10755612]
92. Hiemstra HS, et al. The identification of CD4+ T cell epitopes with dedicated synthetic peptide libraries. *Proc Natl Acad Sci USA*. 1997; 94:10313–10318. [PubMed: 9294207]
93. Hiemstra HS, et al. Definition of natural T cell antigens with mimicry epitopes obtained from dedicated synthetic peptide libraries. *Journal of immunology*. 1998; 161:4078–4082.
94. Pinilla C, Appel JR, Houghten RA. Investigation of antigen-antibody interactions using a soluble, non-support-bound synthetic decapeptide library composed of four trillion (4×10^{12}) sequences. *Biochem J*. 1994; 301:847–853. [PubMed: 7519851]
95. Uda K, Wiesmuller KH, Kienle S, Jung G, Walden P. Self-MHC-restricted peptides recognized by an alloreactive T lymphocyte clone. *J Immunol*. 1996; 157:670–678. [PubMed: 8752916]
96. Judkowski V, Pinilla C, Schroder K, Tucker L, Sarvetnick N, Wilson DB. Identification of MHC class II-restricted peptide ligands, including a glutamic acid decarboxylase 65 sequence, that stimulate diabetogenic T cells from transgenic BDC2.5 nonobese diabetic mice. *J Immunol*. 2001; 166:908–917. [PubMed: 11145667]
97. Gundlach BR, Wiesmuller KH, Junt T, Kienle S, Jung G, Walden P. Specificity and degeneracy of minor histocompatibility antigen-specific MHC-restricted CTL. *J Immunol*. 1996; 156:3645–3651. [PubMed: 8621898]
98. Gundlach BR, Wiesmuller KH, Junt T, Kienle S, Jung G, Walden P. Determination of T cell epitopes with random peptide libraries. *J Immunol Methods*. 1996; 192:149–155. [PubMed: 8699011]
99. Kan-Mitchell J, et al. Degeneracy and repertoire of the human HIV-1 Gag p17(77-85) CTL response. *J Immunol*. 2006; 176:6690–6701. [PubMed: 16709828]
100. Blondelle SE, Moya-Castro R, Osawa K, Schroder K, Wilson DB. Immunogenically optimized peptides derived from natural mutants of HIV CTL epitopes and peptide combinatorial libraries. *Biopolymers*. 2008; 90:683–694. [PubMed: 18481808]
101. Wooldridge L, et al. A single autoimmune T cell receptor recognizes more than a million different peptides. *J Biol Chem*. 2012; 287:1168–1177. [PubMed: 22102287]
102. Wilson DB, et al. Immunogenicity. I. Use of peptide libraries to identify epitopes that activate clonotypic CD4+ T cells and induce T cell responses to native peptide ligands. *J Immunol*. 1999; 163:6424–6434. [PubMed: 10586032]
103. Maynard J, et al. Structure of an autoimmune T cell receptor complexed with class II peptide-MHC: insights into MHC bias and antigen specificity. *Immunity*. 2005; 22:81–92. [PubMed: 15664161]
104. Blake J, Johnston JV, Hellstrom KE, Marquardt H, Chen L. Use of combinatorial peptide libraries to construct functional mimics of tumor epitopes recognized by MHC class I-restricted cytolytic T lymphocytes. *J Exp Med*. 1996; 184:121–130. [PubMed: 8691125]
105. Hemmer B, et al. Identification of high potency microbial and self ligands for a human autoreactive class II-restricted T cell clone. *J Exp Med*. 1997; 185:1651–1659. [PubMed: 9151902]
106. Dogan I, et al. Phage-displayed libraries of peptide/major histocompatibility complexes. *Eur J Immunol*. 2004; 34:598–607. [PubMed: 14768065]
107. Zhao H, Hermesen T, Stet RJ, Skjodt K, Savelkoul HF. Peptide-binding motif prediction by using phage display library for SasaUBA*0301, a resistance haplotype of MHC class I molecule from Atlantic Salmon (*Salmo salar*). *Mol Immunol*. 2008; 45:1658–1664.
108. Le Doussal J, Piqueras B, Dogan I, Debre P, Gorochoff G. Phage display of peptide/major histocompatibility complex. *J Immunol Methods*. 2000; 241:147–158. [PubMed: 10915857]
109. Vest Hansen N, Ostergaard Pedersen L, Stryhn A, Buus S. Phage display of peptide / major histocompatibility class I complexes. *Eur J Immunol*. 2001; 31:32–38. [PubMed: 11169435]
110. Kurokawa MS, et al. Expression of MHC class I molecules together with antigenic peptides on filamentous phages. *Immunol Lett*. 2002; 80:163–168. [PubMed: 11803048]

111. Crawford F, Huseby E, White J, Marrack P, Kappler JW. Mimotopes for alloreactive and conventional T cells in a peptide-MHC display library. *PLoS Biol.* 2004; 2:E90. [PubMed: 15094798]
112. Grabherr R, Ernst W, Oker-Blom C, Jones I. Developments in the use of baculoviruses for the surface display of complex eukaryotic proteins. *Trends Biotechnol.* 2001; 19:231–236. [PubMed: 11356285]
113. Grabherr R, Ernst W, Doblhoff-Dier O, Sara M, Katinger H. Expression of foreign proteins on the surface of *Autographa californica* nuclear polyhedrosis virus. *Biotechniques.* 1997; 22:730–735. [PubMed: 9105625]
114. Wang Y, et al. Using a baculovirus display library to identify MHC class I mimotopes. *Proc Natl Acad Sci USA.* 2005; 102:2476–2481. [PubMed: 15699351]
115. Wen F, Sethi DK, Wucherpfennig KW, Zhao H. Cell surface display of functional human MHC class II proteins: yeast display versus insect cell display. *Protein Eng Des Sel.* 2011; 24:701–709. [PubMed: 21752831]
116. Crawford F, et al. Use of baculovirus MHC/peptide display libraries to characterize T-cell receptor ligands. *Immunol Rev.* 2006; 210:156–170. [PubMed: 16623770]
117. Macdonald WA, et al. T cell allorecognition via molecular mimicry. *Immunity.* 2009; 31:897–908. [PubMed: 20064448]
118. Jordan KR, et al. Baculovirus-infected insect cells expressing peptide-MHC complexes elicit protective antitumor immunity. *J Immunol.* 2008; 180:188–197. [PubMed: 18097019]
119. Starwalt SE, Masteller EL, Bluestone JA, Kranz DM. Directed evolution of a single-chain class II MHC product by yeast display. *Protein Eng.* 2003; 16:147–156. [PubMed: 12676983]
120. Wen F, Esteban O, Zhao H. Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library. *J Immunol Methods.* 2008; 336:37–44. [PubMed: 18448115]
121. Esteban O, Zhao H. Directed evolution of soluble single-chain human class II MHC molecules. *J Mol Biol.* 2004; 340:81–95. [PubMed: 15184024]
122. Boder ET, Bill JR, Nields AW, Marrack PC, Kappler JW. Yeast surface display of a noncovalent MHC class II heterodimer complexed with antigenic peptide. *Biotechnol Bioeng.* 2005; 92:485–491. [PubMed: 16155952]
123. Jiang W, Boder ET. High-throughput engineering and analysis of peptide binding to class II MHC. *Proc Natl Acad Sci USA.* 2010; 107:13258–13263. [PubMed: 20622157]
124. Brophy SE, Holler PD, Kranz DM. A yeast display system for engineering functional peptide-MHC complexes. *J Immunol Methods.* 2003; 272:235–246. [PubMed: 12505727]
125. Jones LL, et al. Engineering and characterization of a stabilized $\alpha 1/\alpha 2$ module of the class I major histocompatibility complex product Ld. *J Biol Chem.* 2006; 281:25734–25744. [PubMed: 16815841]
126. Varani L, et al. Solution mapping of T cell receptor docking footprints on peptide-MHC. *Proc Natl Acad Sci USA.* 2007; 104:13080–13085. [PubMed: 17670943]
127. Adams JJ, et al. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity.* 2011; 35:681–693. [PubMed: 22101157]
128. Hornell TM, et al. Alloreactive and syngeneic CTL are comparably dependent on interaction with MHC class I α -helical residues. *J Immunol.* 1999; 163:3217–3225. [PubMed: 10477590]
129. Huang J, et al. The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness. *Nature.* 2010; 464:932–936. [PubMed: 20357766]
130. Livnah O, et al. An antagonist peptide-EPO receptor complex suggests that receptor dimerization is not sufficient for activation. *Nat Struct Biol.* 1998; 5:993–1004. [PubMed: 9808045]
131. Yoon ST, Dianzani U, Bottomly K, Janeway CA Jr. Both high and low avidity antibodies to the T cell receptor can have agonist or antagonist activity. *Immunity.* 1994; 1:563–569. [PubMed: 7600285]
132. Li QJ, et al. CD4 enhances T cell sensitivity to antigen by coordinating Lck accumulation at the immunological synapse. *Nat Immunol.* 2004; 5:791–799. [PubMed: 15247914]

133. Kessels HW, van Den Boom MD, Spits H, Hooijberg E, Schumacher TN. Changing T cell specificity by retroviral T cell receptor display. *Proc Natl Acad Sci USA*. 2000; 97:14578–14583. [PubMed: 11121060]
134. Chervin AS, Aggen DH, Raseman JM, Kranz DM. Engineering higher affinity T cell receptors using a T cell display system. *J Immunol Methods*. 2008; 339:175–184. [PubMed: 18854190]
135. Boen E, Crownover AR, McIlhaney M, Korman AJ, Bill J. Identification of T cell ligands in a library of peptides covalently attached to HLA-DR4. *J Immunol*. 2000; 165:2040–2047. [PubMed: 10925287]
136. Yin Y, Wang XX, Mariuzza RA. Crystal structure of a complete ternary complex of T-cell receptor, peptide-MHC, and CD4. *Proc Natl Acad Sci USA*. 2012; 109:5405–5410. [PubMed: 22431638]
137. Wang XX, et al. Affinity maturation of human CD4 by yeast surface display and crystal structure of a CD4-HLA-DR1 complex. *Proc Natl Acad Sci USA*. 2011; 108:15960–15965. [PubMed: 21900604]

Fig. 1. Sources of diversity in pMHC-TCR interactions
Left: Overall binding topology of representative pMHC-TCR complex. Right: close-up of pMHC-TCR binding interface. TCR CDR1 and CDR2 (boxed in cyan) are encoded in the germline and primarily bind MHC (green). The TCR CDR3 interface (boxed in magenta) is uniquely derived for each receptor and primarily binds the peptide (yellow). The table (below) lists areas of diversity. Overall theoretical TCR diversity is 10^{15} unique receptors. The potential peptidome displayed on MHC is a comparable number.

Fig. 2. Synthetic versus cell-based methods for discovery of new TCR peptide ligands

(A) Split-pool libraries use mixtures of peptides to activate a T-cell clone of interest. The response of a set amino acid at a given position can lead to information about residues that are favored or disfavored. Schematic data is provided for three peptide positions (P1-P3). In this data, Asp and Glu are favored for P1 (left), Pro is disfavored for P2 (middle), and Lys is strongly favored for P3 (right). This information can be combined to find an optimized peptide ligand and even to deorphanize a TCR. (B) Cell-based display systems can be used to generate libraries of diverse peptides. Library members are then selected for peptide binding to MHC or for pMHC binding to TCR. This approach can optimize peptide affinity for MHC (top), higher TCR affinity (middle), or be used to identify new peptides of interest (bottom).

Fig. 3. Diversity of structural features exhibited by yeast display-derived peptides recognized by 42F3 TCR
 (A) Overall binding topology for p5E8, p4B10, P3A1, and QL9 shows a marked difference for p3A1. (B) 42F3 CDR3 loops (red and blue) adopt different topologies to bind to peptides (yellow) with markedly different chemistries. (C) Peptide-CDR3 interaction map showing van der Waals contacts (black) and hydrogen bonds (red). (D) Electron density maps of peptides from L^d-42F3 complex crystal structures.

Fig. 4. Potential mechanisms for high affinity, non-stimulatory pMHC ligands

(A) 42F3 TCR binding footprint for agonist (QL9, black) and nonagonist (p3A1, red) peptides. While there is a large deviation in binding footprint, the overall TCR binding polarity is maintained. (B) Schematic of binding/function differences between agonist and high affinity nonagonist. (C) Potential models for agonist (yellow) vs. nonagonist (red) peptides. Efficient pMHC binding may not be possible in the context of a T cell (left), binding may not lead to productive signaling (center), or TCR oligomerization may be disrupted (right).

Table 1

TCRs displayed in formats for diversification

TCR	Species	Display System	Display Format	Functional?	Evolution Required?	Notes	ref
2C	Mouse	Yeast	Aga2p-V β -linker-V α	Yes	Yes	<i>a</i>	37, 41
3.L2	Mouse	Yeast	Aga2p-V β -linker-V α	Yes	Yes	<i>a</i>	49
C18	Mouse	Yeast	Aga2p-V β -linker-V α	Yes	Yes		52
mWT1	Mouse	Yeast	Aga2p-V α -linker-V β	Yes	Yes		52
CE10	Mouse	Yeast	Aga2p-V β -linker-V α	Yes	Yes		52
2C	Mouse	Yeast	Aga2p-V α Ca (Secreted V β CB)	Yes	No	<i>b</i>	52
CE10	Human	Yeast	Aga2p-V α Ca (Secreted V β CB)	No	N.D.		52
CE10	Human	Yeast	Aga2p-V β CB (Secreted V α Ca)	Yes	No		52
1G4	Human	Yeast	Aga2p-V β -linker-V α	No	No	<i>c</i>	53
A6	Human	Yeast	Aga2p-V β -linker-V α	Yes	Yes		53
868	Human	Yeast	Aga2p-V β -linker-V α	Yes	No	<i>d</i>	53
A6	Human	Phage	V β CB-pIII (Secreted V α Ca)	Yes	No	<i>a</i>	54
1G4	Human	Phage	V β CB-pIII (Secreted V α Ca)	Yes	No	<i>a</i>	54
ILAK	Human	Phage	V β CB-pIII (Secreted V α Ca)	Yes	No		54
LCL3	Human	Phage	V β CB-pIII (Secreted V α Ca)	Yes	No		54
JM22	Human	Phage	V β CB-pIII (Secreted V α Ca)	N.D.	No		54
AH1.23	Human	Phage	V β CB-pIII (Secreted V α Ca)	N.D.	No		54
MM15	Human	Phage	V β CB-pIII (Secreted V α Ca)	Yes	No		54
1A77	Human	Phage	V β CB-pIII (Secreted V α Ca)	N.D.	No		54
CD1d	Human	Phage	V β CB-pIII (Secreted V α Ca)	Yes	No		54
GRb	Human	Phage	V β CB-pIII (Secreted V α Ca)	N.D.	No		54
DOI1.10	Mouse	Phage	V α -linker-V β CB-pVIII	Yes	No		55
DOI1.10	Mouse	Phage	V α -linker-V β CB-pIII	Yes	No		55
anti-p53 TCR	Human	Phage	V α -linker-V β CB-pIII	Yes	No	<i>e</i>	55
4B2A1	Mouse	Phage	V α -linker-V β -pIII	Yes	No		56
4B2A1	Mouse	Phage	V β -linker-V α -pIII	Yes	No		56
4B2A1	Mouse	Phage	V α CL-pIII (V β CH Secreted)	Yes	No		56
4B2A1	Mouse	Phage	V β CL-pIII (V α CH Secreted)	Yes	No	<i>f</i>	56
7A10B2	Mouse	Phage	V α -linker-V β -pIII	Yes	No	<i>f</i>	56

TCR	Species	Display System	Display Format	Functional?	Evolution Required?	Notes	ref
7A10B2	Mouse	Phage	Vβ-linker-Vα-pIII	Yes	No	<i>f</i>	56
7A10B2	Mouse	Phage	VαCL-pIII (VβCH Secreted)	Yes	No	<i>f</i>	56
7A10B2	Mouse	Phage	VβCL-pIII (VαCH Secreted)	Yes	No	<i>f</i>	56

Notes

- ^aHigh affinity versions created
- ^bInter-chain disulfide required for correct fold
- ^cEvolution attempted but no improvement noted
- ^dImprovement from WT seen with mutagenesis
- ^e100-1000 fold more phage needed than for DO11.10
- ^fEfficient display requires rFkpA coexpression

Table 2

pMHCs displayed in formats for diversification

MHC Allele	Display System	Display Format	Properly Folded?	Evolution Required?	Binds Soluble TCR	Activates T cells	ref
I-A ^b	Baculovirus	Peptide-MHCβ-zipper-TM (MHCα-zipper Secreted)	Yes	No	Yes	Yes	111
D ^d	Baculovirus	MHCα.1α.3-TM (peptide-β2m Secreted)	Yes	No	Yes	Yes	114
HLA-DR15	Yeast	Aga2p-MHCα-linker-peptide-MHCβ	Yes	Yes	No	No	115
HLA-DR15	Baculovirus	MHCα-linker-peptide-MHCβ-VSVG anchor	Yes	No	No	Yes	115
HLA-DR15	Baculovirus	Peptide-MHCβ-zipper-VSVG (MHCα-zipper Secreted)	Yes	No	Yes	Yes	115
L ^d	Baculovirus	MHC α.1α.3-TM (peptide-β2m Secreted)	Yes	No	Yes	Yes	116, 118
I-E ^k	Baculovirus	Peptide-MHCβ-zipper-TM (MHCα-zipper Secreted)	Yes	No	N.D.	N.D.	116
HLA-DR52c	Baculovirus	Peptide-MHCβ-zipper-TM (MHCα-zipper Secreted)	Yes	No	N.D.	N.D.	116
HLA-B*4405	Baculovirus	MHC α.1α.3-TM (peptide-β2m Secreted)	Yes	No	Yes	N.D.	117
I-A ^{g7}	Yeast	Aga2p-peptide-MHCβ-linker-MHCα	Yes	Yes	N.D.	N.D.	119
HLA-DR1	Yeast	Aga2p-MHCα-linker-peptide-MHCβ	Yes	No	N.D.	Yes	120
HLA-DR1	Yeast	Aga2p-MHCβ-linker-MHCα	Yes	Yes	N.D.	N.D.	121
HLA-DR1	Yeast	Aga2p-MHCα-linker-MHCβ	Yes	Yes	N.D.	N.D.	121
HLA-DR4	Yeast	Aga2p-peptide-MHCβ (MHCα Secreted)	No	N.D.	N.D.	N.D.	122
HLA-DR4	Yeast	Peptide-MHCβ-Aga2p (MHCα Secreted)	Yes	No	N.D.	No	122
HLA-DR1	Yeast	Aga2p-peptide (MHCα and MHCβ Secreted)	Yes	No	N.D.	N.D.	123
K ^b	Yeast	Aga2p-MHCα.1α.3-linker-β2m	Yes	No	Yes ^a	N.D.	124
K ^b	Yeast	Peptide-linker-β2m-linker-MHCα.1α.3-Aga2p	Yes	No	No	Yes	124
L ^d	Yeast	Aga2p-MHCα.1α.3-linker-β2m	Yes	No ^b	Yes ^a	N.D.	125
L ^d	Yeast	Aga2p-MHCα.1α.2	Yes	Yes	Yes ^a	N.D.	125
L ^d	Yeast	Aga2p-MHCα.1α.2-linker-peptide	Yes	Yes	Yes	N.D.	127

^aStaining seen when exogenous peptide is loaded

^bImprovement seen upon mutagenesis and selection