University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Food for Health Papers & Publications | Food for Health

11-28-2016

A generalized framework for computational design and mutational scanning of T-cell receptor binding interfaces

Timothy P. Riley

Cory M. Ayres

Lance M. Hellman

Nishant K. Singh

Michael Cosiano

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unl.edu/ffhdocs

 Part of the Biochemical Phenomena, Metabolism, and Nutrition Commons, Dietetics and Clinical Nutrition Commons, Gastroenterology Commons, Medical Microbiology Commons, and the Medical Nutrition Commons

This Article is brought to you for free and open access by the Food for Health at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Food for Health Papers & Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Timothy P. Riley, Cory M. Ayres, Lance M. Hellman, Nishant K. Singh, Michael Cosiano, Jennifer M. Cimons, Michael J. Anderson, Kurt H. Piepenbrink, Brian G. Pierce, Zhiping Weng, and Brian M. Baker

Copyright Oxford University Press. Used by permission.

Protein Eng Des Sel. 2016 Dec; 29(12): 595–606.
Published online 2016 Nov 28. doi: 10.1093/protein/gzw050
PMCID: PMC5181382
PMID: 27624308

A generalized framework for computational design and mutational scanning of T-cell receptor binding interfaces

Timothy P. Riley,[1] Cory M. Ayres,[1] Lance M. Hellman,[1] Nishant K. Singh,[1] Michael Cosiano,[1] Jennifer M. Cimons,[1] Michael J. Anderson,[1] Kurt H. Piepenbrink,[1,4] Brian G. Pierce,[2] Zhiping Weng,[3] and Brian M. Baker[1,*]

[1]Department of Chemistry & Biochemistry and the Harper Cancer Research Institute, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, IN, 46556, USA
[2]Institute for Bioscience and Biotechnology Research, University of Maryland, 9600 Gudelsky Drive, Rockville, MD, 20850, USA
[3]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA, 01605, USA
Contributed by [4]Present address: Institute of Human Virology, University of Maryland School of Medicine, 725 West Lombard Street, Baltimore, MD 21201, USA.
[*]To whom correspondence should be addressed. E-mail: ude.dn@rekab-nairb
Contributed by

Edited by: Andrew Miranker

Received 2016 May 19; Revised 2016 Aug 19; Accepted 2016 Aug 23.
This article has been cited by other articles in PMC.

Associated Data

Supplementary Materials [attached]
Supplementary Data
supp_29_12_595__index.html (825 bytes)
GUID: 26050EB5-574F-4D7C-A839-8D148591475F
supp_gzw050_updatedsupplementaldatafile.pdf (2.3M)
GUID: E8EBA806-3FF9-411A-916D-826D736B4BD1

Abstract

T-cell receptors (TCRs) have emerged as a new class of therapeutics, most prominently for cancer where they are the key components of new cellular therapies as well as soluble biologics. Many studies have generated high affinity TCRs in order to enhance sensitivity. Recent outcomes, however, have suggested that fine manipulation of TCR binding, with an emphasis on specificity may be more valuable than large affinity increments. Structure-guided design is ideally suited for this role, and here we studied the generality of structure-guided design as applied to TCRs. We found that a previous approach, which successfully optimized the binding of a therapeutic TCR, had poor accuracy when applied to a broader set of TCR interfaces. We thus sought to develop a more general purpose TCR design framework. After assembling a large dataset of experimental data spanning multiple interfaces, we trained a new scoring function that accounted for unique features of each interface. Together with other improvements, such as explicit inclusion of molecular flexibility, this permitted the design new affinity-enhancing mutations in multiple TCRs, including those not used in training. Our approach also captured the impacts of mutations and substitutions in the peptide/MHC ligand, and recapitulated recent findings regarding TCR specificity, indicating utility in more general mutational scanning of TCR–pMHC interfaces.

Keywords: affinity, mutational scanning, specificity, structure-guided design, T-cell receptor

Introduction

αβ T cells utilize clonotypic T-cell receptors (TCRs) to recognize antigens and initiate cellular immune responses. TCRs have emerged as a new class of therapeutics, most prominently for the treatment of cancer. Although in many ways similar to antibodies, TCRs differ in the complexity of the receptor-ligand interface: whereas antibodies can be elicited to almost any antigen, TCRs are restricted to linear peptide antigens presented by class I or class II MHC proteins (pMHC), with the TCR invariably contacting both (Rossjohn *et al*., 2015). Additionally, TCRs do not undergo affinity maturation, and, similar to naive antibodies, bind with weak-to-moderate affinities and reduced specificity (Baker *et al*., 2012).

Recent advances have highlighted the potential therapeutic uses for TCRs with altered binding properties. As T-cell potency can be improved with antigen affinity (Varela-Rohena *et al*., 2008; Zhao *et al*., 2007), clinical trials with gene-modified T cells have explored the use of engineered, high affinity TCRs for improved antigen targeting (Linette *et al*., 2013). High affinity TCRs are also used as the antigen recognition component of soluble reagents designed to redirect naive, unmodified T cells (Oates and Jakobsen, 2013).

Multiple methods have been used to generate high affinity TCRs, with the majority created using yeast or phage display (Bowerman *et al*., 2009; Holler *et al*., 2000; Li *et al*., 2005; Varela-Rohena *et al*., 2008; Zhao *et al*., 2007). However, recent findings have shown that careful control is necessary when modifying TCRs. Due to their cross-reactive nature, enhancing affinity may

introduce new reactivities: improving affinity against one antigen can improve affinity towards others, leading to reactivity towards antigens that might otherwise be ignored by T cells expressing the wild-type receptor. This could include self-antigens, leading to possible off-target recognition (Zhao *et al*., 2007). Such an outcome is believed to have led to fatal autoimmunity in a recent clinical trial that used a high affinity TCR to target a melanoma antigen (Linette *et al*., 2013). The likelihood of such an outcome may be increased if added 'glue' is directed more towards the MHC protein than the peptide. Additionally, the relationship between TCR affinity and potency is not well understood. Although some very high affinity TCRs show considerable sensitivity (Varela-Rohena *et al*., 2008), in other cases improving affinity outside an optimal window or above a threshold has led to decreased potency (Stone and Kranz, 2013).

Although *in vitro* evolution has been used to generate the majority of high affinity TCRs, structure-guided computational design offers the potential for finer control over affinity and specificity. Not only can interactions be manipulated in a way that more appropriately addresses peptide specificity, affinity increments can in principle be more tightly controlled. Towards these goals, structure-guided design has been used to modify a small number of TCRs (Haidar *et al*., 2009; Malecek *et al*., 2014; Pierce *et al*., 2014; Zoete *et al*., 2013). Recently, we used structure-guided design to engineer variants of the DMF5 TCR, which has been used clinically in immunotherapy for melanoma and continues to serve as a model TCR for improving cancer immunotherapy (Johnson *et al*., 2009). Building on an approach originally developed for the well-studied A6 TCR (Haidar *et al*., 2009), we successfully engineered nanomolar affinity variants of DMF5 with altered specificity, and found excellent agreement between prediction and experiment for both structure and affinity (Pierce *et al*., 2014).

Here we addressed the generality of our TCR design efforts. We found that our approach successfully used with DMF5 performed poorly with additional mutations and other, unrelated TCRs. This may be attributable to the complexity of TCRs and their interfaces with pMHC, such as varying binding geometries, sub-optimal packing and differing amounts of receptor and ligand flexibility (Baker *et al*., 2012, Rossjohn *et al*., 2015). We therefore sought to develop a more generalizable framework for TCR design. After assembling and modeling a large training set of experimental binding data spanning multiple TCR–pMHC systems, we trained a candidate score function which outperformed those used previously. Performance was further enhanced by optimizing the scoring methodology and including information on receptor/ligand flexibility as well as buried water. This allowed for successful design of new, affinity-enhancing mutations in multiple TCRs, including an unrelated receptor not used in training. The new design framework was also successful in recapitulating positive and negative effects of mutations to the MHC protein as well as substitutions in the peptide, and captured emerging themes in TCR specificity (Adams *et al*., 2016). Although there are avenues for improvement, these new developments greatly extend the applicability of structure-guided design for the manipulation and screening of TCR binding properties, and suggest ways for computational screening for peptide antigenicity.

Materials and Methods

Crystal structure processing and design parameters

For structural modeling, Rosetta with the Talaris2013 score function was used (Das and Baker, 2008; Kaufmann *et al*., 2010; Leaver-Fay *et al*., 2013; Moretti *et al*., 2013), using the PyRosetta interface (Chaudhury *et al*., 2010). Native crystal structures were brought to local energy minima through multiple cycles of backbone minimization and rotamer optimization with heavy atom restraints (Bradley *et al*., 2005). Following structure minimization, the desired TCR, MHC, or peptide mutation was computationally introduced followed by three independent Monte Carlo based simulated annealing trajectories of the TCR CDR loops. This was performed using Rosetta's LoopMover_Refine_CCD mover with 3 outer cycles and 10 inner cycles, using an initial metropolis acceptance criteria of 2.2 that decreased linearly to 0.6 (Canutescu and Dunbrack, 2003). The large number of resulting packing operations introduced some minor variability when scoring the models. Therefore, the unweighted score terms for the three trajectories were averaged and stored for point mutation energy calculations (Kellogg *et al*., 2011). When screening TCR point mutations, TCR residue positions with a center of mass within 10 Å (DMF5 and B7) or 15 Å (DMF4) of a peptide heavy atom were selected for design. For peptide screens, all positions other than the primary anchors of the $MART1_{26(27L)–35}$ peptide underwent the design procedure. The design process sampled every amino acid (19 mutations and the wild-type residue) at each specified position in triplicate. Wild-type complexes were modeled and included in scoring to account for impacts of minimization and conformational sampling. For double mutants, both mutations were introduced simultaneously followed by a minimum of six independent minimization trajectories to account for additional structural impacts.

Score function training

To develop a new score function for predicting changes in binding $\Delta\Delta G°$, we considered Rosetta full atom terms in addition to dynamically derived terms (bound and free order parameters and RMS fluctuations). Multiple linear regression was performed in MATLAB 2015b using measured $\Delta\Delta G°$ values. A stepwise elimination protocol was used to remove contextually insignificant terms. A *k*-fold ($k = 10$) cross validation was performed with the data points and significant predictor terms (Arlot and Celisse, 2010).

Modeling explicit water molecules and sarcosine

To model and score buried water molecules and the non-standard sarcosine, explicit TIP3P waters and sarcosine parameters were enabled in Rosetta. Water molecules were placed at their initial crystallographic coordinates followed by 100 high resolution docking trials to coordinate the water molecule in the pocket of the interfaces. The water coordinates were then fixed in position relative to the pMHC for TCR point mutation modeling.

Molecular dynamics simulations of bound and free structures

Molecular dynamics simulations were calculated utilizing the AMBER molecular dynamics suite (Salomon-Ferrer *et al*., 2013) as previously described (Ayres *et al*., 2016). Results for the free and bound A6 and DMF5 were taken from these simulations, with other simulations following the same protocol. Briefly, coordinates for the complexes with the LC13, B7 and DMF4 TCRs were obtained from PDB accession codes 1MI5, 1BD2 and 3QDM. Coordinates for the free

$Tax_{11–19}$/HLA-A2 complex were from 1DUZ. For the LC13, B7 and DMF4 TCRs, coordinates for the free TCRs were obtained by stripping away the pMHC. Prior to simulation, starting systems were charge neutralized with explicit Na^+ counterions and solvated with explicit SPC/E water. Following this, systems were energy minimized and heated to 300 K with solute restraints. Afterwards, solute restraints were gradually relaxed and followed with 2 ns of simulation with no solute restraints for equilibration, after which 100 ns production trajectories for all systems were calculated. Trajectories were calculated using GPU-accelerated code (Götz *et al*., 2012; Salomon-Ferrer *et al*., 2013). Trajectory analysis including calculation of RMSF values used the ccptraj from the AMBER suite (Roe and Cheatham, 2013). Order parameters were calculated using isotropic reorientational eigenmode dynamic analysis using vectors defined from the Cα to Cβ (or Cα to H for glycine) atoms (Prompers and Brüschweiler, 2002). For double mutants, descriptors were averaged between the two positions for scoring purposes (i.e. for mutant XY, the RMSF of position X is averaged with the position Y RMSF to give an RMSF descriptor for XY).

Protein expression and purification

Expression and refolding of soluble constructs of the DMF5, B7 and DMF4 TCRs and HLA-A2 were performed as previously described (Davis-Harrison *et al*., 2005). Briefly, the TCR α and β chains, the HLA-A2 heavy chain and β_2-microglobulin (β_2m) were generated in *Escherichia coli* as inclusion bodies, which were isolated and denatured in 8 M urea. TCR α and β chains were diluted in TCR refolding buffer (50 mM Tris (pH 8), 2 mM EDTA, 2.5 M urea, 9.6 mM cysteamine, 5.5 mM cystamine, 0.2 mM PMSF) at a 1:1 ratio. HLA-A2 and β_2m were diluted in MHC refolding buffer (100 mM Tris (pH 8), 2 mM EDTA, 400 mM l-arginine, 6.3 mM cysteamine, 3.7 mM cystamine, 0.2 mM PMSF) at a 1:1 ratio in the presence of excess peptide. TCR and pMHC complexes were incubated for 24 h at 4°C. Afterward, complexes were desalted by dialysis at 4°C and room temperature respectively, then purified by anion exchange followed by size-exclusion chromatography. Refolded protein absorptions at 280 nm were measured spectroscopically and concentrations determined with appropriate extinction coefficients. Mutations in TCR α and β chains were generated by whole-plasmid mutagenesis and confirmed by sequencing. Peptides were synthesized and purified commercially.

Surface plasmon resonance

Surface plasmon resonance experiments were performed with a Biacore 3000 instrument using CM5 sensor chips as previously described (Davis-Harrison *et al*., 2005). In all experiments, TCR was immobilized to the sensor chip via standard amine coupling and pMHC complex was injected as analyte. Experiments were performed at 25°C in 20 mM HEPES (pH 7.4), 150 mM NaCl, 0.005% Nonidet P-20. All experiments were steady-state experiments measuring RU vs. concentration of injected analyte, and were performed with TCRs coupled onto the sensor chip at 400–2000 response units. Injected pMHC spanned a concentration range of 0.1–150 μM at flow rates of 5 μl/min. Data were processed with BiaEvaluation 4.1 and fit using a 1:1 binding model utilizing MATLAB 2015b.

Results

Application of earlier structure-guided design methods to the B7 TCR

Based on previous work with the A6 TCR (Haidar *et al*., 2009), we recently described a modeling and scoring scheme to predict the structural and energetic effects of point mutations within interfaces with the αβ TCR DMF5 (Pierce *et al*., 2014). Using this approach we identified several affinity-enhancing mutations in DMF5 which when combined led to affinity enhancements towards pMHC of up to 400-fold. To explore the generality of this approach, we applied the same methodology to the B7 TCR (Ding *et al*., 1998), which binds the human T-cell lymphoma virus $Tax_{11–19}$ peptide presented by HLA-A2 with a similar affinity and orientation as the A6 TCR (Fig. (Fig.1a).1a). The A6 and B7 TCRs also share the same germline-derived Vβ chain, although crystallographic structures and biophysical studies of A6 and B7 with $Tax_{11–19}$/HLA-A2 showed structural and thermodynamic differences in binding (Davis-Harrison *et al*., 2005). We modeled 740 point mutations in the B7-$Tax_{11–19}$/HLA-A2 interface using Rosetta (Das and Baker, 2008; Kaufmann *et al*., 2010) and the scheme described in Pierce *et al*. (2014). As performed previously, effects were determined by scoring the complex, then separating the components and separately scoring the TCR and pMHC in order to calculate a 'binding score' (Kortemme and Baker, 2002). Based on these scores, nine mutations were selected for predicted enhancements to binding affinity and chosen for experimental testing.

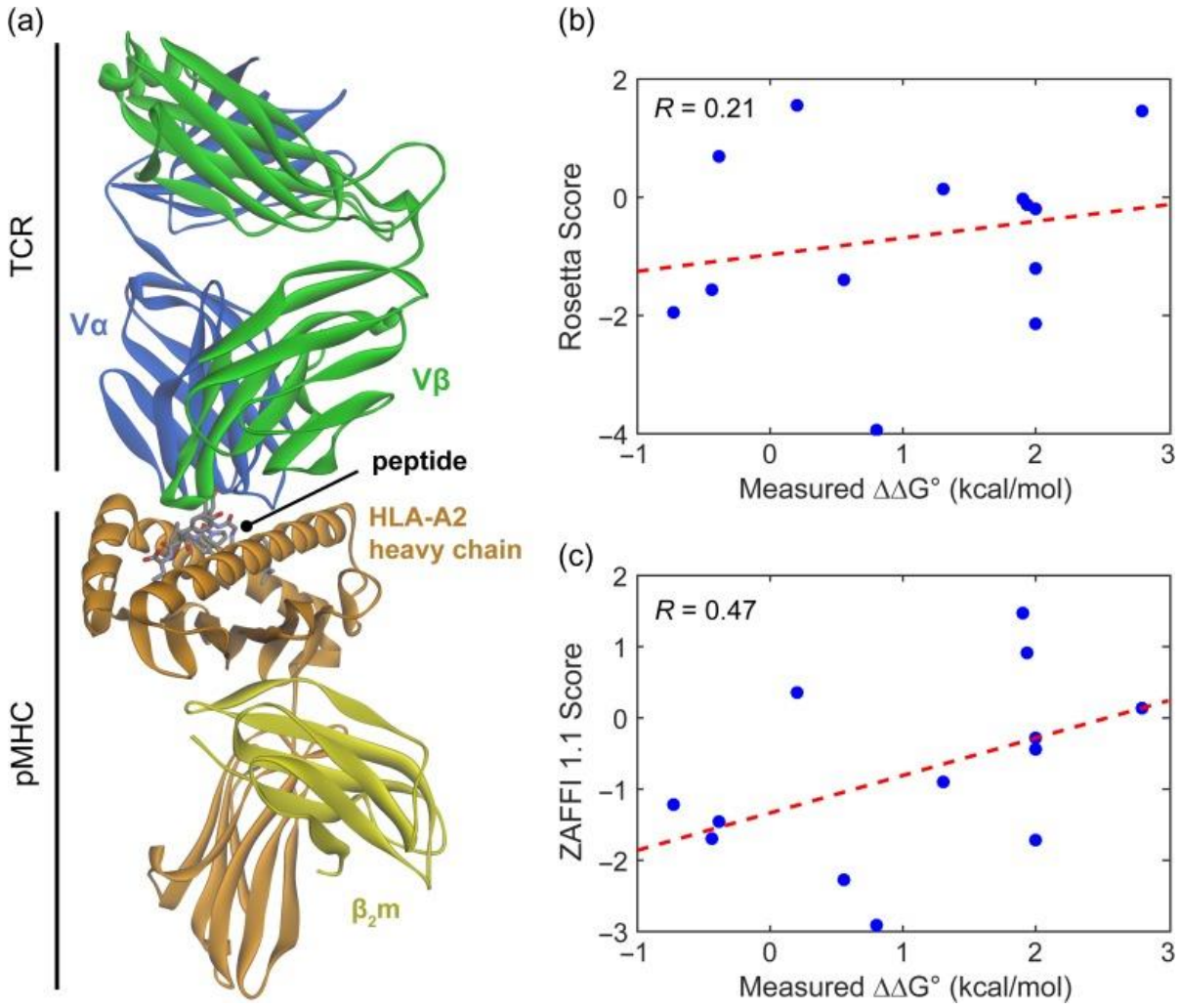

Fig. 1 Mutations in the interface between the B7 TCR and Tax_{11-19}/HLA-A2 are scored poorly with the Rosetta interface and ZAFFI 1.1 functions. (**a**) Structural overview of the B7 TCR–pMHC complex. (**b**) Score vs. experimental ΔΔG° for point mutations modeled with Rosetta and scored with the Rosetta interface function. The best fit line and correlation coefficient is indicated. (**c**) As with panel b, scored with the ZAFFI 1.1 function (Haidar *et al*., 2009; Pierce *et al*., 2014).

We performed mutagenesis using soluble B7 gene constructs, expressed and purified the mutant and wild-type proteins, and measured their binding affinities toward $Tax_{11–19}$/HLA-A2 using surface plasmon resonance (Table S1 and Fig. S1). Three of the mutations (S27αM, S50αY, G99βY) led to moderately enhanced affinity towards $Tax_{11–19}$/HLA-A2, although the remaining six mutations weakened affinity or led to no detectable binding. Including four additional B7 mutations studied previously (Piepenbrink *et al*., 2013), the correlation between the predicted and experimental change in binding energy was low with the Rosetta interface score function (R = 0.21; Fig. Fig.1b).1b). Utilizing the ZAFFI score function first developed for the A6 TCR and refined with the DMF5 TCR (Haidar *et al*., 2009; Pierce *et al*., 2014) led to an improved but

still weak correlation (R = 0.47; Fig. Fig.1c).1c). Thus, the TCR design approach developed for the A6 TCR and later applied to DMF5 performs poorly with the B7 TCR.

Collection of new data to train a score function for HLA-A2-restricted TCRs

In light of the low correlations between prediction and experiment with the B7 TCR, we aimed to develop a more generalizable framework for modeling and predicting point mutations across multiple TCR–pMHC interfaces. We collected 96 independent ΔΔG° values resulting from single amino acid mutations from four TCR–pMHC interfaces (A6-Tax$_{11–19}$/HLA-A2; B7-Tax$_{11–19}$/HLA-A2; DMF5-MART1$_{27–35}$/HLA-A2; and DMF5-MART1$_{26(27L)–35}$/HLA-A2). This data originated from our previously published structure-guided design efforts with the A6 and DMF5 TCRs (Haidar *et al*., 2009; Pierce *et al*., 2014) as well as our recent double mutant cycle deconstruction of the A6 interface (Piepenbrink *et al*., 2013). We also included the additional data with B7 described above, and performed new binding measurements in the DMF5-MART1$_{26(27L)–35}$/HLA-A2 interface (Table S1 and Fig. S1). We restricted the dataset to high quality measurements with low experimental error (< 0.5 kcal/mol).

The point mutations in our dataset covered a broad range of mutation types as described in Table S2. The ΔΔG° values ranged from −1.8 to 2.8 kcal/mol and were approximately normal in distribution (Fig. (Fig.2a).2a). The median ΔΔG° value of the selected dataset was 0.5 kcal/mol with a standard deviation of 1.1 kcal/mol. When comparing the 29 mutations that improved binding, it became evident the majority of affinity-enhancing mutations resulted from replacement of small or polar residues with large hydrophobic or amphipathic residues (Fig. (Fig.22b).

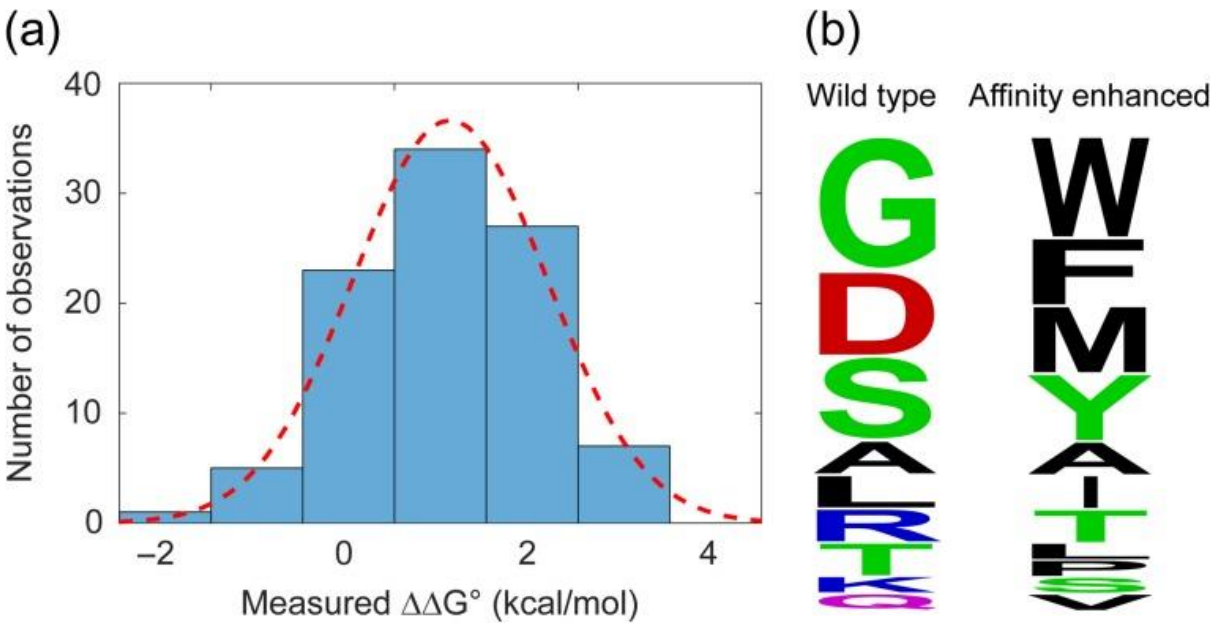

Fig. 2 Experimental ΔΔG° values of TCR point mutations are normal in distribution and affinity-enhancing mutations are predominantly hydrophobic or amphipathic. (**a**) The 96 point mutations collected in different TCR–pMHC interfaces were approximately normal in distribution with a

median ΔΔG° value of 0.5 kcal/mol and a standard deviation of 1.1 kcal/mol. (**b**) Sequence logos of the 29 mutations that improved binding (ΔΔG° < 0).

Development of a generalized TCR–pMHC scoring function

We next developed computational structural models of all 96 point mutations for training generalized TCR prediction models. We extended our strategy by adapting techniques for modeling the effects of interface mutations shown to be successful in recent community-wide assessments. Mutations were modeled with the standard Talaris2013 score function allowing for off-rotamer sampling and limited backbone flexibility in the CDR loops (Leaver-Fay *et al*., 2013; Moretti *et al*., 2013). Additionally, side chains of residues within a 10 Å sphere of any CDR loop residue were repacked in response to each mutation and resulting CDR loop movements. Each point mutation was modeled in triplicate and scores averaged for further analysis. Analysis of the mutation models identified one with an anomalously high repulsive clash score and another where a residue was forced into an unusual high energy rotamer. Both of these mutations were excluded from further training and comparisons, leaving a dataset of 94 point mutations and their structural models.

To develop a generalizable TCR scoring function, we considered 16 full-atom Rosetta terms commonly used for protein design and structure prediction (Leaver-Fay *et al*., 2013; Moretti *et al*., 2013). Using the Rosetta terms as predictor variables and experimental binding energies of the dataset described above as the response variable, we used multi-linear regression to parameterize a starting score function for estimating the effect of the various point mutations on ΔΔG°. The most significant contributors to the model ($P < 0.05$) described van der Waals attractive forces and solvation effects. However, the correlation between binding score and ΔΔG° remained low ($R = 0.43$; Fig. Fig.3a).3a). Thus we did not explore removing insignificant features at this stage in favor of obtaining a more robust prediction model.

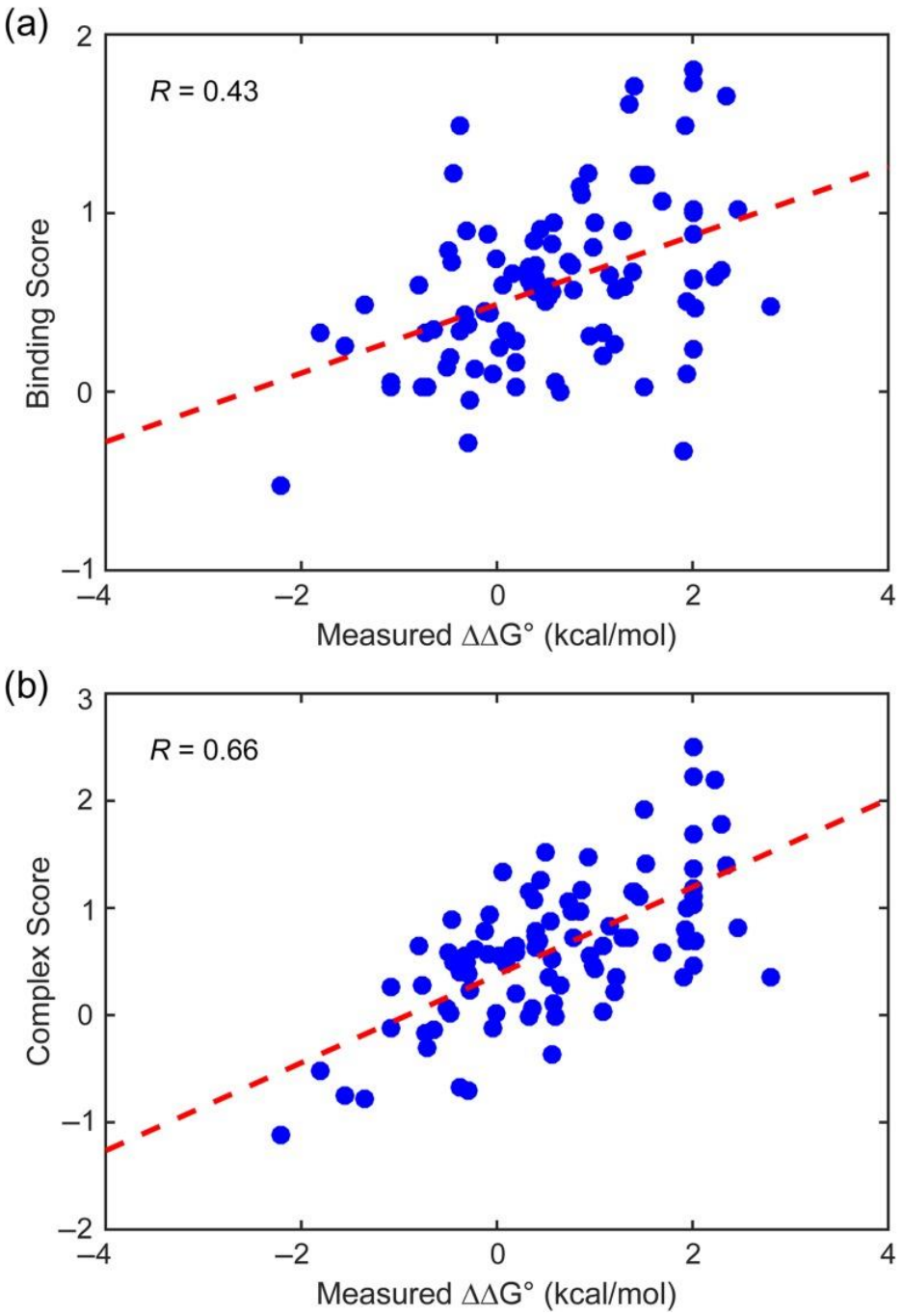

Fig. 3 Relative TCR–pMHC complex scores correlate better with affinity than binding scores. (**a**) Scores vs. experimental ΔΔG° for modeled point mutations. Scores were determined by scoring each complex and two free proteins (i.e. binding score = $score_{complex} - (score_{TCR} + score_{pMHC})$). The wild-type 'binding score' was then subtracted from each mutant binding score. After parameterization of Rosetta structural terms, relative binding scores were plotted vs. experimental ΔΔG°. (**b**) As with panel a, but scores determined and parameterized by scoring only the wild-type and each mutant complex, yielding 'complex scores' as described in the text.

Ideally, binding energy calculations would utilize structural information for both the free and bound molecules (Kortemme and Baker, 2002; Vreven *et al*., 2012). However, structures of free TCRs and pMHCs can vary between free and bound states (Armstrong *et al*., 2008), and the large surface areas of receptor and ligand binding sites possess significant conformational degrees of freedom. We thus focused only on relative effects by scoring only TCR–pMHC complexes, rather than scoring the complex and the two free proteins as described above. We refer to the difference in scores between wild type and mutant complexes as 'complex scores'. This approach comes with a limitation in that complex scores do not account for energies in the free TCR associated with making the mutation (i.e. the ΔG° for TCR WT ↔ TCR mutant). Ideally these would be subtracted when examining the impact of a mutation on binding. There are two potentially significant consequences to this. First, an improved complex score could arise solely due to improved contacts within the TCR (i.e better TCR stability). We minimized the impact of this by focusing on sites that are in proximity to the ligand and thus more likely to influence binding. Second, any effects on binding stemming from conformational changes in the free TCR will be ignored.

Using the same 16 full-atom Rosetta terms, a multi-linear regression of complex scores vs. ΔΔG° yielded an improved function (R = 0.66 for complex scores, vs. R = 0.43 for binding scores; Fig. Fig.3b).3b). Despite the theoretical limitations noted above, complex scores are therefore more applicable for our framework and were used for all further calculations. The improvement using complex scores may reveal underlying limitations in the energy function terms and/or limitations in recapitulating conformational differences between free and bound TCRs as noted above, leading to inaccuracies when 'binding scores' are computed. The inherently weak affinities and correspondingly poor quality of TCR–pMHC interfaces (compared, e.g. to high affinity antibody-antigen interfaces) could also contribute to why complex scores outperform binding scores.

Refinement of the regression model to include flexibility and validation of terms

Although utilization of complex scores improved the correlation between prediction and experiment, we sought to identify additional predictors of TCR binding affinity that might further improve performance. One of the differences between TCRs is their degree of binding loop flexibility, particularly for the hypervariable CDR3α and CDR3β loops (Scott *et al*., 2011). Although various methods for conformational sampling such as stochastic loop perturbations or generation of structural ensembles exist (Feixas *et al*., 2014; Sinko *et al*., 2013; Tuffery and Derreumaux, 2012), these are computationally expensive. To more simply address the impacts of TCR loop flexibility, we considered descriptors from molecular dynamics (MD) simulations of the free and bound TCRs. We recently described a comprehensive MD study of the free and bound A6 and DMF5 TCRs (Ayres *et al*., 2016) using an experimentally benchmarked simulation methodology (Scott *et al*., 2011, Scott *et al*., 2012). We performed similar simulations on the free and bound B7 TCR. From these simulations root mean square (RMS) fluctuations for each α carbon were determined along with Cα-Cβ (Cα-H for glycine) and Cα-C order parameters to quantify nanosecond timescale backbone flexibility (Fig. S2). Due to the time that would be required to simulate dozens or hundreds of mutations, only the wild-type TCRs and their complexes were simulated. Fluctuation values and order parameters were then treated as 'positional modifiers' for each amino acid position, biasing positions for design based on their

relative flexibility in the wild-type free and bound structures. Although necessary for throughput, this approach makes the limiting assumption that any given mutation does not impact backbone flexibility on the nanosecond timescale.

To determine if inclusion of RMS fluctuations and/or order parameters could lead to an improved scoring function, we included these six terms along with the 16 full-atom Rosetta terms in a multi-linear regression of complex scores vs. ΔΔG°, coupled with a stepwise elimination protocol (Hocking, 1976). This fit identified six significant ($P < 0.05$) features: four structural terms (van der Waals attractive and repulsive forces, solvation energies and sidechain hydrogen bonding) and two flexibility terms (RMS fluctuations for α carbons of the free and bound structures). A structural term weighting Ramachandran angle propensities was borderline significant ($P = 0.11$), but was retained to help identify and exclude structural models with residues forced into unrealistic conformations.

The regression models estimated the weights of the RMS fluctuation features to be negative, suggesting flexible positions are more favorable to target for design (although mobility in the complex was weighted more heavily as discussed below). To critically examine the significance of this determination, models with and without the fluctuation terms in addition to the five Rosetta terms were generated and compared. Akaike information criterion (AIC) (Akaike, 1998) found the incorporation of features describing flexibility resulted in a 99.8% likelihood of a superior prediction model. Bayesian information criterion (BIC) (Kass and Raftery, 1995) more strongly penalized additional terms, yet also indicated that inclusion of the fluctuation terms improved the regression model beyond random chance (Table (TableII).

Table I.

Inclusion of RMS fluctuations improves the score function regression model

Criteria	**RMSF excluded**	**RMSF included**
R	0.63	0.71
P-value	7.9×10^{-9}	9.0×10^{-11}
AIC	239.2	226.8
BIC	254.4	246.2

Finally, a *k*-fold cross validation ($k = 10$) (Arlot and Celisse, 2010) was used to validate and estimate overall predictive performance. From this analysis, the RMS error (reflecting the difference between experimental and predicted ΔΔG° values) was estimated as 0.81 kcal/mol, with an impressive correlation of 0.71 (Fig. (Fig.4a;4a; note this correlation includes accounting for structural water as described below). For comparison, our previous approach with the Rosetta interface score function yielded a correlation of only 0.16 (Fig. (Fig.4b),4b), and a recent analysis of protein design approaches estimated an average error of 1.2 kcal/mol for protein–protein interactions (Potapov *et al*., 2009). The terms and weights for the final regression model, termed the TR3 score function, are shown in Table TableIIII.

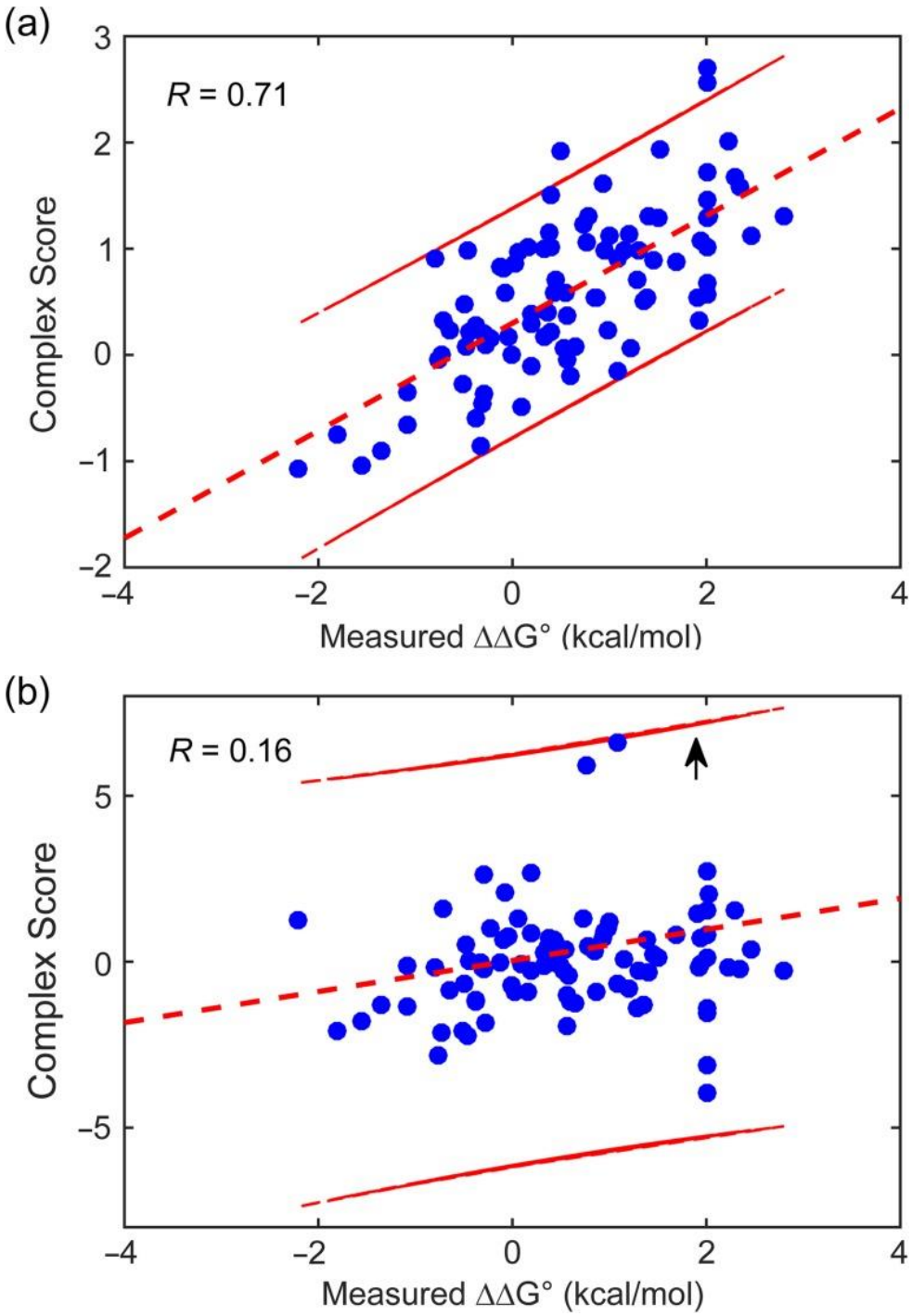

Fig. 4 The TR3 score function outperforms our previous TCR design methodology. (**a**) Complex score vs. experimental ΔΔG° for 94 point mutations modeled with Rosetta and scored with the TR3 function. The best fit line, 95% confidence interval, and correlation coefficient is indicated. (**b**) Performance of our previous methodology applied to the same data. An off-scale prediction score of 26 (DMF5 G28αL) is denoted by a black arrow and the best fit line and correlation are indicated.

Table II.

Terms and their statistics in the TR3 score function

Term	Weight	Error[a]	*P*-value[b]
Intercept	2.29	0.35	<0.001
Fa_atr	0.21	0.03	<0.001
Fa_rep	0.05	0.01	0.005
Fa_sol	0.18	0.08	<0.001
Hbond_sc	0.34	0.09	0.008
Rama	0.12	0.05	0.119
RMSF_bound	−0.82	0.30	0.049
RMSF_free	−0.36	0.10	0.003
Estimated error: 0.81 kcal/mol[c]			

[a]Determined as 1.96 standard deviations of *k*-fold cross-validation weights.
[b]*P*-value for the *F* statistic of the hypotheses test that the corresponding coefficient is equal to zero.
[c]Average test RMS error from *k*-fold cross validation.

Accounting for energetically significant structural water improves predictions

Rosetta utilizes an implicit solvation model to estimate solvation energies associated with bulk water (Lazaridis and Karplus, 1999). However, TCR–pMHC interfaces are large and buried water molecules are often observed crystallographically. In some instances these structural waters play key roles in the interface that would not be captured with an implicit solvation model (Jiang *et al*., 2005). Indeed, many predicted mutations which filled the void of an interfacial water molecule in the interface with the DMF5 TCR resulted in a falsely favorable score. For example, Ser99 in the DMF5 β chain contacts the peptide, but is also involved in a complex water-mediated hydrogen bond network linking the peptide to the TCR (Fig. (Fig.5a).5a). The predicted impacts of mutations at this position did not correlate well with experiment (Fig. (Fig.5b),5b), consistent with a determination that this water molecule is structurally and energetically significant. To directly account for it, the buried water in the DMF5 interface was docked into its corresponding pocket and treated explicitly in modeling and scoring. This improved the agreement between prediction and experiment for Ser99β point mutations without altering the predictions for distant residues (Fig. (Fig.5c).5c). Further design efforts incorporated this technique when buried water molecules were observed crystallographically in the interface between peptide and TCR (i.e. the DMF4 and LC13 TCRs as described later).

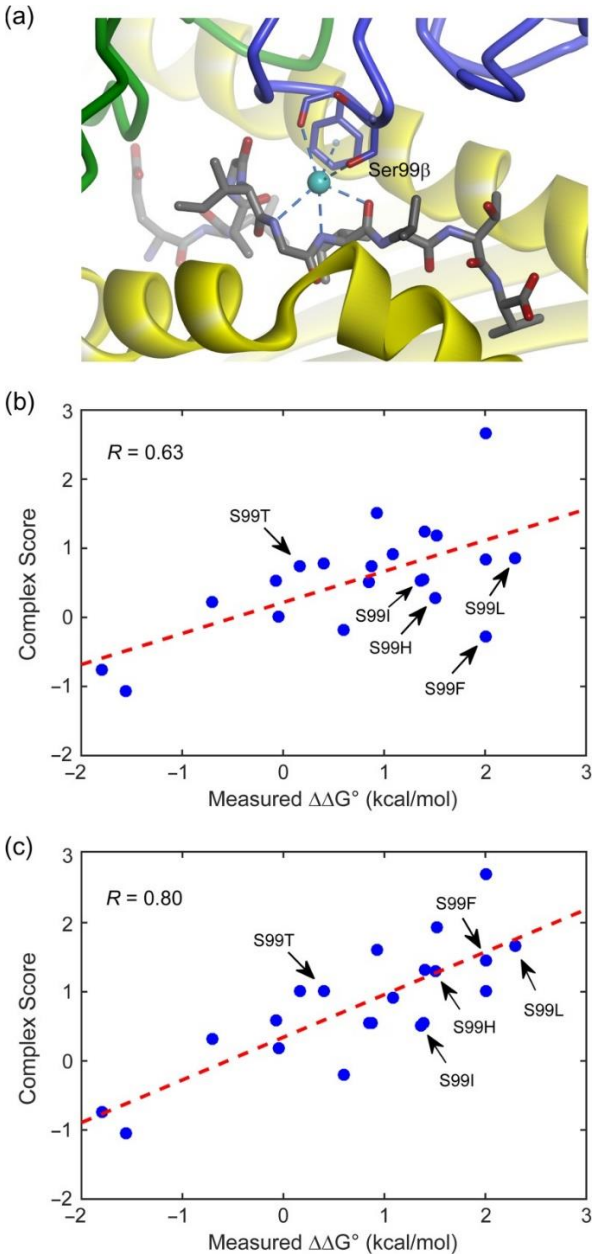

Fig. 5 Accounting for buried structural water improves predictions. (**a**) A buried water molecule observed crystallographically in the DMF5-MART1$_{26(27L)-35}$/HLA-A2 interface forms multiple electrostatic interactions between the TCR and peptide. The sidechain of Ser99 of the DMF5 β chain is indicated. (**b**) The correlation between prediction and experiment for models of DMF5 point mutants scored with TR3 is 0.63 when the buried water molecule is ignored. Five mutations at position 99β are indicated and are responsible for the low correlations. (**c**) The correlation between prediction and experiment for DMF5 point mutants improves to 0.80 when the buried water molecule is treated explicitly. The predicted effects of the five mutations at position 99β agree better with experiment as shown.

Validation with new TCR mutations and combinations to further modulate affinity

We next collected additional data to assess the performance of our new framework on mutations not used in training. We screened for new mutations in the interfaces with the DMF5 and B7 TCRs (DMF5-MART1$_{26(27L)–35}$/HLA-A2 and B7-Tax$_{11–19}$/HLA-A2). To emphasize peptide specificity, only positions with a center of mass within 10 Å of a peptide heavy atom were selected for design. A total of 18 sites in both DMF5 and B7 were modeled and scored with all 20 amino acids (684 point mutations in total and 36 wild-type controls). As expected, most mutations were predicted to have deleterious effects on binding. However, several mutations were predicted to enhance affinity, most at sites where mutations have previously been shown to favorably impact binding (Table S1). The two predicted to be most favorable (G99βW for B7; D26αF for DMF5) were both generated, and the impact on binding assessed experimentally. Both mutations improved binding as predicted. The ΔΔG° for G99βW in B7 was −0.5 kcal/mol; for D26αF in DMF5 it was −0.4 kcal/mol. The value for D26αF was less than observed previously with tyrosine or tryptophan at this position (−1.8 and −1.6 kcal/mol, respectively), suggesting that the amphipathic character of tyrosine and tryptophan may be advantageous for enhancing TCR affinity as discussed below.

Our previous designs for the A6 and DMF5 TCRs combined multiple mutations to generate molecules which bound in the nanomolar range (Haidar *et al*., 2009; Pierce *et al*., 2014). The approximate additive effects of mutations in both interfaces were captured by our new framework with the TR3 score function after averaging the RMSF positional values of each of the mutations. To ask if our new framework also allowed for this in another TCR, we combined the S27αM and G99βY mutations in the B7 receptor, which together improved the B7 affinity for Tax$_{11–19}$/HLA-A2 7-fold, from 1.5 μM to 220 nM (Fig. (Fig.6).6). These mutations are ~27 Å apart, and were correctly predicted to be additive when combined (ΔΔG° = −1.2 kcal/mol, complex score = −0.77).

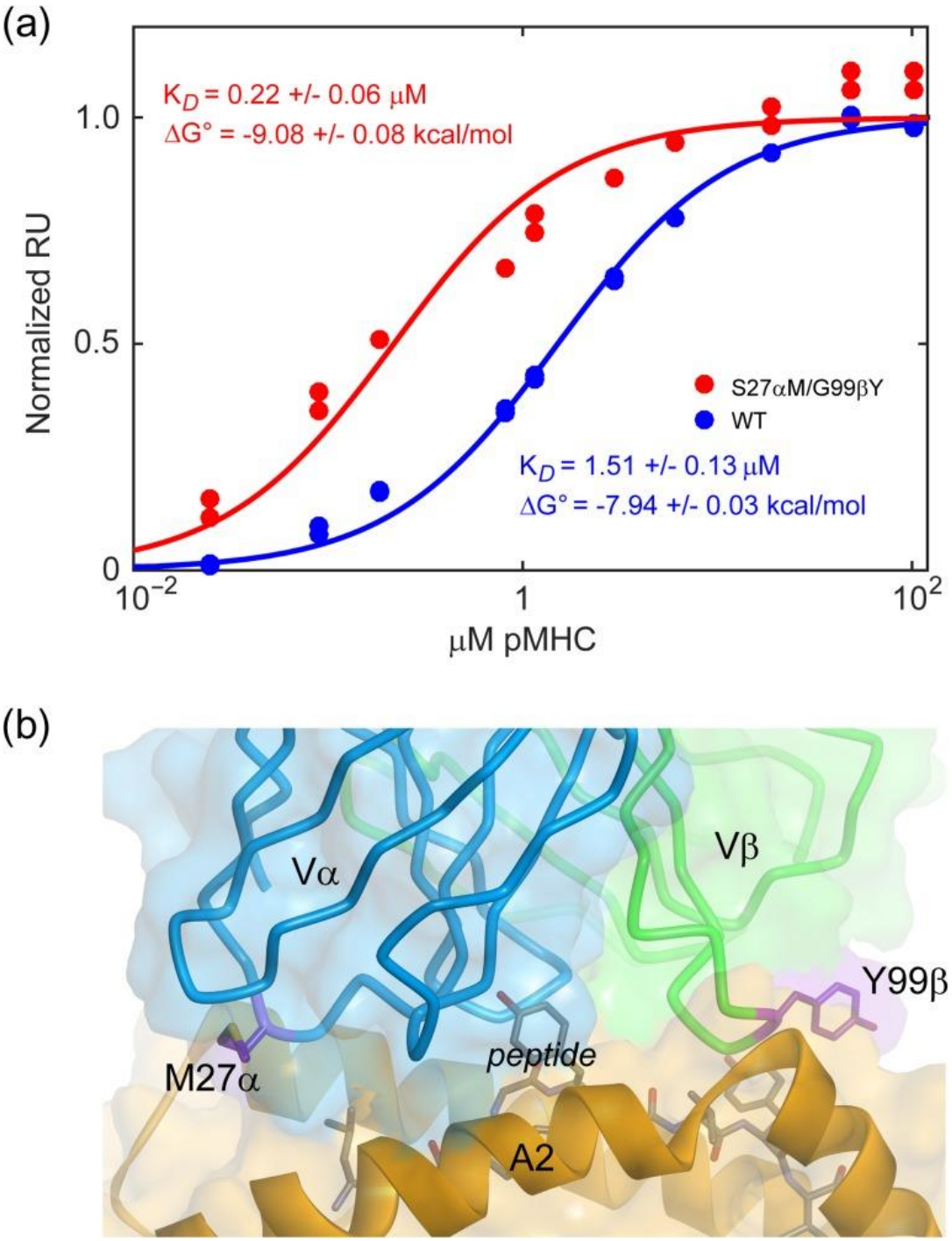

Fig. 6 Combining two computationally designed B7 mutations yields nanomolar binding affinity. (**a**) Combining the S27αM and G99βY mutations in the B7 TCR improves binding to Tax_{11-19}/HLA-A2 7-fold, from 1.5 μM to 220 nM. (**b**) The sites of the S27αM and G99βY mutations in the B7 TCR are separated by ~27 Å and are predicted to improve affinity independently through improved van der Waals interactions with the pMHC.

To investigate broader applicability, mutations in another TCR not used in training were modeled and scored. The DMF4 TCR also recognizes MART1 antigens presented by HLA-A2, but utilizes different α and β chains than DMF5, A6 and B7 (Borbulevych *et al*., 2011; Johnson *et al*., 2009). As performed with the A6, B7 and DMF5 TCRs, MD simulations of the free wild-type DMF4 TCR and its complex with $MART1_{26(27L)–35}$/HLA-A2 were performed and used along with Rosetta to simulate 960 structures (19 mutations at 48 sites, and 48 wild-type controls) in the DMF4-$MART1_{26(27L)–35}$/HLA-A2 interface. Several mutations in the α chain were favorably ranked based on their ability to fill an interfacial void near the N-terminus of the peptide. Three of these mutations were selected for experimental investigation (S26αW, N29αW and T92αW). Although the N29αW mutation was of particular interest as it provided another opportunity to investigate a structural water, this mutant could not be folded from inclusion bodies. This left two mutations for experimental testing. As predicted, both of these enhanced DMF4 binding affinity, with ΔΔG° values of −0.4 and −0.9 kcal/mol (Table S1). These mutations were also predicted and found to be additive when combined: together the S26αW and T92αW mutations improved the affinity of the DMF4 TCR 10-fold, from 60 to 6 μM (ΔΔG° of −1.4 kcal/mol).

Overall, when applied to data outside of our training set, our new modeling and scoring procedure recapitulated the effects of multiple mutations in the B7, DMF5 and DMF4 TCRs and permitted the identification of new affinity-enhancing mutations in all three receptors. The RMS error between predicted and experimentally determined impacts on binding was 1.5 kcal/mol, higher than observed with training and cross-validation but still lower than observed with our previous methodology (Fig. (Fig.7a,7a, black points).

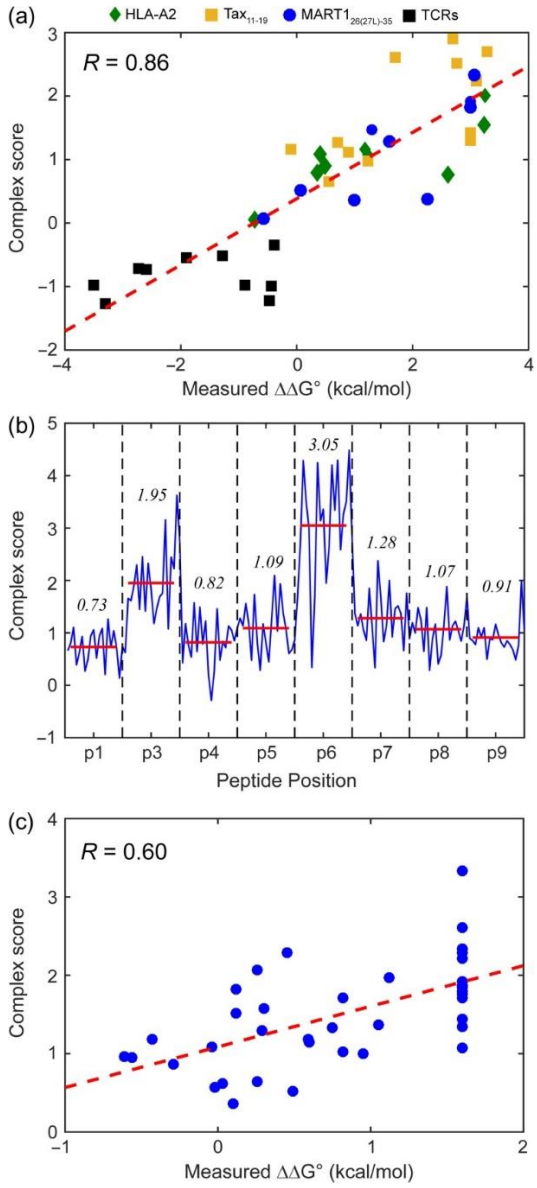

Fig. 7 Performance of our improved framework on new TCR mutations, HLA-A2 mutations and peptide variations. (**a**) All point mutation data examined in evaluating our new approach, including TCR, peptide and HLA-A2 data, plotted together, excluding data used in training. The overall correlation between prediction and experiment is 0.86. (**b**) The predicted effects of $MART1_{26(27L)-35}$ peptide substitutions on the binding of DMF5 to $MART1_{26(27L)–35}$/HLA-A2 indicate amino acids that are more tolerating of or more sensitive to substitutions. Position 6 near the center of the peptide is particularly sensitive. Each segment of the plot shows the complex scores for all 20 amino acids substituted at the indicated position. Solid lines and numbers in each segment show the average scores for all 20 amino acids at that position. (**c**) Performance is more limited on a system involving a more diverse, non-HLA-A2 restricted TCR. The impact of mutations in the LC13 TCR with FLR/HLA-B8 are predicted with a correlation coefficient of 0.60 (ΔΔG° values of mutations with no detectable binding were reported previously as 1.6 kcal/mol) (Borg *et al*., 2005).

Including HLA-A2 mutations in validation

αβ TCRs show MHC restriction, i.e. they recognize peptides only when presented by MHC proteins (Zinkernagel and Doherty, 1974). Many studies have examined the effects of mutations in the α helices of MHC binding groove as a means to determine energetically significant positions that might guide restriction, including a recent comprehensive analysis of the binding of A6 TCR to the $Tax_{11–19}$/HLA-A2 complex. Of nine published mutations, eight weakened affinity and one enhanced affinity (Piepenbrink *et al*., 2013). To recapitulate this data *in silico*, we modeled the impact of mutations in HLA-A2 on the binding of A6 to $Tax_{11–19}$/HLA-A2, incorporating free and bound flexibility through MD simulations as described above. The effects of these mutations were well captured, with RMS error between prediction and experiment of 1.0 kcal/mol (Fig. (Fig.7a,7a, green points). Thus our new framework is applicable not only to TCRs, but can predict the energetics associated with mutations in the HLA-A2 side of the interface as well.

Computational scanning of peptide variants

TCRs are broadly cross-reactive and recognize a multitude of antigenic peptides, a requirement of the fixed size of the T-cell repertoire (Mason, 1998). Additionally, altering TCR binding by changing peptide sequence is another approach for modulating TCR binding and immune responses (McMahan and Slansky, 2007; Piepenbrink *et al*., 2009). Quantitative data for how eight substitutions in the $Tax_{11–19}$ peptide impact the binding of the A6 TCR is available (Davis-Harrison *et al*., 2007; Piepenbrink *et al*., 2013), and we collected new alanine scanning data for recognition of four more $Tax_{11–19}$ variants by B7 (Table S1). As with the HLA-A2 mutations above, we used our new modeling and scoring approach to assess how these peptide variants impact recognition by A6 and B7. The impacts on binding ΔΔG° were recapitulated well, with an RMS error of 0.9 kcal/mol (Fig. (Fig.7a,7a, yellow points).

To further demonstrate the utility of our approach for assessing peptide variations, residues in the $MART1_{26(27L)–35}$ peptide were computationally varied to cover all 20 amino acids, and, after completing a MD simulation of the $MART1_{26(27L)–35}$/HLA-A2 complex, scored for impact on DMF5 binding. All peptide substitutions were predicted to be unfavorable, although mutations at the P3 and P6 positions were predicted to have the most dramatic impacts (Fig. (Fig.7b).7b). This outcome is consistent with recent findings on TCR specificity, which suggest the existence of peptide 'hotspots' of reduced structural and chemical diversity, outside of which greater variation is permitted (Adams *et al*., 2016).

Next, eight $MART1_{26(27L)–35}$ peptide variants with a broad range of complex scores were selected for experimental testing with DMF5 (Table S1). We also examined a peptide with a non-standard sarcosine (*N*-methyl glycine) substituted for Gly6 of the peptide to help test the implications of treating structured water explicitly in the DMF5 interface as discussed above and shown in Fig. Fig.5.5. Overall, there was a good correlation between ΔΔG° and binding score for the nine $MART1_{26(27L)–35}$ peptide variants explored experimentally, with experiment and prediction differing with an RMS error of 0.9 kcal/mol (Fig. (Fig.7a,7a, blue points). The experiments with

the sarcosine-modified peptide led to improved binding as predicted, leading to a 3-fold affinity enhancement in affinity (ΔΔG° of −0.6 kcal/mol). The affinity enhancement is attributable to the increased van der Waals interactions to Thr102 of the TCR while maintaining the solvated state of polar atoms in the surrounding pocket.

Overall performance and exploration of an even more diverse, non-HLA-A2 interface

To explore the overall performance of our new approach, we examined the new TCR mutations, HLA-A2 mutations and peptide variants described above together as one large test set. These amounted to 40 independent ΔΔG° measurements distinct from the training set from five different TCR–pMHC interfaces. We also included the double mutants in the DMF5, B7 and A6 TCRs. Altogether, performance was excellent, with predicted and experimental impacts on binding agreeing with an impressive correlation coefficient of 0.86 and a RMS error of 1.1 kcal/mol, spanning a large range of ~7 kcal/mol in binding free energy (Fig. (Fig.7a,7a, all points). Complex scores again showed improved performance over binding scores, as scoring the 40 test set mutations using binding scores yielded a weaker correlation coefficient (R = 0.66) and larger RMS error (2.8 kcal/mol) (Fig. S3).

The systems used in development and testing all involved the class I MHC protein HLA-A2. To explore how our new framework performed when additional diversity was included, we used it to assess the impact of mutations between the interface of the LC13 TCR and the class I MHC protein HLA-B:08:01 (HLA-B8) presenting the FLR peptide (sequence FLRGRAYGL). The structure of the LC13-FLR/HLA-B8 complex has been determined, as have ΔΔG° values for 39 alanine or glycine mutations in the various LC13 CDR loops (Borg *et al*., 2005). After completing MD simulations of LC13 and its complex, we applied our approach to this dataset, recapitulating the effects of these mutations with an overall correlation of 0.60 and an RMS error of 1.0 kcal/mol (Fig. (Fig.7c).7c). While errors are still within the range obtained with our previous methodology (Pierce *et al*., 2014), the correlation is weaker than what we achieved with HLA-A2-restricted systems.

There are at least two possible reasons for the weaker performance with the LC13 TCR. First, many of the 39 mutations in the LC13 interface result in very weak or no detectable binding, with ΔΔG° values reported simply as above an upper limit of 1.6 kcal/mol (corresponding to a 15-fold weakening of affinity). The limited accuracy of these measurements will affect the correlation between prediction and experiment. As evidence of this, binary metrics demonstrated good predictive performance when separating affinity increasing mutations from affinity decreasing mutations (ROC AUC = 0.84; Fig. S4). Second, our reliance on HLA-A2-restricted systems in parameterization of the new TR3 score function could result in an inherent bias. HLA-A2 and HLA-B8 differ by 42 amino acids, 32 of which are in the peptide binding domain (Robinson *et al*., 2011). In addition to different energetic contributions, these differences could alter the structural and dynamic responses to mutations in ways that are poorly captured by our framework.

Discussion

TCRs have emerged as a new class of immunological therapeutics, most prominently for cancer, where they are the key components of new cellular immunotherapies as well as soluble biologics (Oates and Jakobsen, 2013; Restifo *et al*., 2012). There is significant interest in enhancing TCR affinity to improve antigen sensitivity, and accordingly, numerous high affinity TCRs have been generated. Although T-cell potency has been shown to improve with affinity, questions remain about the existence of optimal affinity windows or thresholds and the merits of large vs. incremental improvements in binding affinity (Stone and Kranz, 2013). Additionally, following adverse events in clinical trials (Morgan *et al*., 2013; Parkhurst *et al*., 2011), there is a growing recognition of the importance of evaluating and controlling specificity in affinity-enhanced or otherwise modified TCRs.

In principle, structure-guided computational design offers the potential for fine manipulation of TCR binding properties. Structure-guided design has been used to generate a small number of high affinity TCRs, as well as manipulate binding specificity (Haidar *et al*., 2009; Malecek *et al*., 2014; Pierce *et al*., 2014; Zoete *et al*., 2013). However, although the TCR–pMHC structural database has grown significantly in recent years, wide-scale application of structure-guided TCR design is hindered by several complexities. These include the complex architecture of the TCR–pMHC interface (Baker *et al*., 2012; Rossjohn *et al*., 2015), as well as the varying degrees of diversity and molecular flexibility in both receptor and ligand (Borbulevych *et al*., 2009; Insaidoo *et al*., 2011; Scott *et al*., 2011). We demonstrated the limited applicability of current TCR design approaches here by showing that our prior approach used to successfully engineer the clinically relevant DMF5 TCR performed poorly when applied to the unrelated B7 TCR.

To generate an improved framework for structure-guided TCR design, we assembled a large database of mutations from four TCR–pMHC interfaces and used this in developing a new, 'general purpose' approach to TCR design, including a novel score function. Similar to other score functions trained to predict binding affinity (Kortemme and Baker, 2002), the results heavily weighted van der Waals attractive forces and solvation and dampened repulsive terms. We also accounted for molecular flexibility via a novel cost-effective approach. For data in the training set we were able to achieve a correlation between predicted effect on binding and experimental $\Delta\Delta G°$ of 0.71. When applied to new data from multiple TCRs, HLA-A2 and peptides, we obtained an impressive correlation of 0.86 and low RMS error of 1.1 kcal/mol, which is near the expected upper limit due to error in experimental data (Potapov *et al*., 2009). As with other studies, the slope of predicted vs. experimental $\Delta\Delta G°$ was <1, indicating that impacts on binding affinity are typically under-estimated, giving some indications of ways to improve. The potential for further improvements is also found with the better performance observed with complex vs. binding scores, despite the theoretical limitations associated with complex scores.

Accounting for flexibility is an important aspect of our improved framework, as varying degrees of CDR loop, MHC and peptide flexibility is a characteristic feature of TCRs and pMHC complexes (Borbulevych *et al*., 2009; Insaidoo *et al*., 2011; Scott *et al*., 2011). As with other efforts in protein design, we relied on MD simulations to incorporate flexibility. However, as opposed to simulating structures to identify alternate configurations or generate structural

ensembles (Feixas *et al*., 2014; Sinko *et al*., 2013; Tuffery and Derreumaux, 2012), we added 'positional modifiers' that report on amino-acid level motional properties as terms in the score function. We chose this approach as it greatly simplifies the treatment of flexibility, requiring only single MD trajectories for the free wild-type TCR and the TCR–pMHC complex. Of the properties considered, Cα RMS fluctuations were most significant and were incorporated into the final function. The weights for these terms were negative, indicating that more flexible positions are more favorable for design. There is some anecdotal evidence to support this: in the A6 TCR, the hypervariable CDR3β loop is by far the most mobile, and multiple mutations within this loop improve A6 binding (Haidar *et al*., 2014; Li *et al*., 2005; Scott *et al*., 2011). This could reflect a form of the 'fly-casting' effect, in which mobile sites in a receptor are more adept at finding compatible partner sites in a ligand (Shoemaker *et al*., 2000). Interestingly, the flexibility weights were larger for residues in the bound state. While residual mobility in TCR–pMHC interfaces has been observed and this term could be accounting for this (Hawse *et al*., 2014; Reboul *et al*., 2012), it is also possible that in the complex this term helps overcome limited conformational sampling in modeling. Further work is needed to explore this, along with whether using longer or additional MD simulations can yield further improvements, at the expense of throughput.

Solvent considerations can be important in structure-guided design, as buried water molecules can play critical roles in protein binding (Janin, 1999; Rodier *et al*., 2005). As seen in other systems (de Graaf *et al*., 2005), we demonstrate that explicitly modeling water in TCR interfaces can improve predictions. This was most useful for the DMF5 TCR, although buried waters were incorporated when modeling the DMF4 and LC13 interfaces as well (chosen in these cases because they were buried in the interface and participated in multiple hydrogen bonds with the TCR, peptide or MHC). Because we relied on crystallographically observed water, there is a corresponding demand on the resolution and quality of the original crystal structure. As TCR–pMHC interfaces may be poorly packed and crystallographic resolutions low, incorporation of approaches to predict the location of water molecules not observed crystallographically could lead to further improvements (Bui *et al*., 2007; Jiang *et al*., 2005).

Our improved framework for TCR design permitted the identification of new affinity-enhancing mutations in multiple interfaces, including the DMF4 TCR which was not used in training and uses different Vα and Vβ genes than those in the training set (Borbulevych *et al*., 2011; Johnson *et al*., 2009). The enhancements to affinity are relatively modest, but as noted above fine control may be most desirable when manipulating TCR affinity. Additionally, when combined these mutations can yield larger improvements, as shown for the B7, A6, DMF4 and DMF5 TCRs. Our approach also accounted for the relative effects of alanine and glycine mutations in another TCR, LC13, although the poorer performance with this receptor suggests that application to class I MHC proteins other than HLA-A2 (such as the HLA-B8 allele recognized by LC13) could require additional training outside of HLA-A2 systems to improve generality. By extension, we can expect that application towards other MHC proteins, particularly class II or nonclassical MHC proteins, will also require further effort.

Mutations which enhanced affinity tended to be (although were not exclusively) those that replaced small polar or charged residues with large hydrophobic or amphipathic amino acids. The significance of this is unclear; this was seen in our training data as well as new mutations.

Increasing buried hydrophobic surface area is a well-known strategy for enhancing binding, and the rigid, bulky and amphipathic nature of tyrosine and tryptophan provide structural and chemical utility (Koide and Sidhu, 2009). Highly antigenic peptides have a tendency to be enriched in hydrophobic amino acids in TCR contact sites (Chowell *et al*., 2015), potentially indicating that such amino acids are indeed optimal for enhancing TCR affinity.

What do these results indicate about introducing charged or polar amino acids? While electrostatic interactions can contribute to specificity, their contributions to affinity can vary due to high desolvation penalties (Bosshard *et al*., 2004; Hendsch and Tidor, 1994). Additionally, electrostatic interactions have strict geometrical dependences. Due to these complexities, protein design algorithms are recognized to have limited success modeling electrostatics (Fleishman *et al*., 2011; Procko *et al*., 2013; Stranges and Kuhlman, 2013). Thus introduction of charged or polar interactions may be unintentionally disfavored during design. Overcoming this potential limitation is important, as an over-reliance on select amino acids limits applicability. Further, accurately accounting for electrostatic effects could provide another means to selectively engineer TCR specificity, irrespective of their impacts on binding affinity (Blevins *et al*., 2016; Stadinski *et al*., 2016).

The availability of polar/charged mutations that improve TCR binding will be helpful in further assessing and improving our design framework. Molecular evolution experiments can provide such data, although because multiple mutations are often found in affinity-matured molecules, identifying the impact of individual mutations can be difficult. Perhaps a more promising source for such data could be deep mutational scanning experiments, which sample the effects of every amino acid at multiple positions in one experiment (Fowler and Fields, 2014; Whitehead *et al*., 2012). Our recent deep mutational scanning experiments with the A6 and variant TCRs provide several promising examples of polar/charged mutations that appear to favorably impact binding (Harris *et al*., 2016), and careful analysis of these will be helpful.

Lastly, our approach was also able to account for the effects of mutations in the HLA-A2 protein as well as peptide substitutions. This raises the possibility of using computational design not only for engineering TCRs to modulate their binding properties, but also ligands with enhanced affinity for select TCRs. Such an approach has been proposed as a novel means for peptide-based vaccine design (McMahan and Slansky, 2007; Piepenbrink *et al*., 2009), and could be useful in the development of new T-cell detection or imaging reagents. Additionally, the capacity to accurately score peptide variants could allow for computationally assessing the cross-reactivity of TCRs or the reactivities of peptide sets. This could prove useful for predicting and controlling off target toxicity for TCRs used clinically or identifying reactive self-antigens in autoimmunity, transplantation or vaccination.

Supplementary data

Supplementary data are available at *PEDS* online

Supplementary Data:

Acknowledgements

We thank Moushumi Hossain for collecting TCR binding data for select peptide variants.

Funding

This work was supported by the National Institutes of Health [GM103773; and GM118166] and a grant from the Carole and Ray Neag Comprehensive Cancer Center.

References

1. Adams J.J., Narayanan S., Birnbaum M.E. et al. (2016) Nat. Immunol., 17, 87–94. [PMC free article] [PubMed] [Google Scholar]
2. Akaike H. (1998) Parzen E., Tanabe K. and Kitagawa G (eds), Selected Papers of Hirotugu Akaike. Springer New York, New York, NY: pp.215–222. [Google Scholar]
3. Arlot S. and Celisse A. (2010) A survey of cross-validation procedures for model selection. Statistical Surveys, 4, 40–79. [Google Scholar]
4. Armstrong K.M., Piepenbrink K.H. and Baker B.M. (2008) Biochem. J., 415, 183–196. [PMC free article] [PubMed] [Google Scholar]
5. Ayres C.M., Scott D.R., Corcelli S.A., and Baker B.M. (2016) Sci. Rep., 6, 1–4. [Google Scholar]
6. Baker B.M., Scott D.R., Blevins S.J. and Hawse W.F. (2012) Immunol. Rev., 250, 10–31. [PubMed] [Google Scholar]
7. Blevins S.J., Pierce B.G., Singh N.K., Riley T.P., Wang Y., Spear T.T., Nishimura M.I., Weng Z. and Baker B.M. (2016) Proc. Natl. Acad. Sci., 113, E1276–E1285. [PMC free article] [PubMed] [Google Scholar]
8. Borbulevych O.Y., Piepenbrink K.H., Gloor B.E., Scott D.R., Sommese R.F., Cole D.K., Sewell A.K. and Baker B.M. (2009) Immunity, 31, 885–896. [PMC free article] [PubMed] [Google Scholar]
9. Borbulevych O.Y., Santhanagopolan S.M., Hossain M. and Baker B.M. (2011) J. Immunol., 187, 2453–2463. [PMC free article] [PubMed] [Google Scholar]
10. Borg N.A., Ely L.K., Beddoe T. et al. (2005) Nat. Immunol., 6, 171–180. [PubMed] [Google Scholar]
11. Bosshard H.R., Marti D.N. and Jelesarov I. (2004) J. Mol. Recognit., 17, 1–16. [PubMed] [Google Scholar]
12. Bowerman N.A., Crofts T.S., Chlewicki L., Do P., Baker B.M., Christopher Garcia K. and Kranz D.M. (2009) Mol. Immunol., 46, 3000–3008. [PMC free article] [PubMed] [Google Scholar]
13. Bradley P., Misura K.M.S. and Baker D. (2005) Science, 309, 1868–1871. [PubMed] [Google Scholar]

14. Bui H.-H., Schiewe A.J. and Haworth I.S. (2007) J. Comput. Chem., 28, 2241–2251. [PubMed] [Google Scholar]
15. Canutescu A.A. and Dunbrack R.L. (2003) Protein Sci., 12, 963–972. [PMC free article] [PubMed] [Google Scholar]
16. Chaudhury S., Lyskov S. and Gray J.J. (2010) Bioinformatics, 26, 689–691. [PMC free article] [PubMed] [Google Scholar]
17. Chowell D., Krishna S., Becker P.D., Cocita C., Shu J., Tan X., Greenberg P.D., Klavinskis L.S., Blattman J.N. and Anderson K.S. (2015) Proc. Natl. Acad. Sci., 112, E1754–E1762. [PMC free article] [PubMed] [Google Scholar]
18. Das R. and Baker D. (2008) Annu. Rev. Biochem., 77, 363–382. [PubMed] [Google Scholar]
19. Davis-Harrison R.L., Armstrong K.M. and Baker B.M. (2005) J. Mol. Biol., 346, 533–550. [PubMed] [Google Scholar]
20. Davis-Harrison R.L., Insaidoo F.K. and Baker B.M. (2007) Biochemistry, 46, 1840–1850. [PubMed] [Google Scholar]
21. de Graaf C., Pospisil P., Pos W., Folkers G. and Vermeulen N.P.E. (2005) J. Med. Chem., 48, 2308–2318. [PubMed] [Google Scholar]
22. Ding Y.H., Smith K.J., Garboczi D.N., Utz U., Biddison W.E. and Wiley D.C. (1998) Immunity, 8, 403–411. [PubMed] [Google Scholar]
23. Feixas F., Lindert S., Sinko W. and McCammon J.A. (2014) Biophys. Chem., 186, 31–45. [PMC free article] [PubMed] [Google Scholar]
24. Fleishman S.J., Whitehead T.A., Strauch E.-M. et al. (2011) J. Mol. Biol., 414, 289–302. [PMC free article] [PubMed] [Google Scholar]
25. Fowler D.M. and Fields S. (2014) Nat. Methods, 11, 801–807. [PMC free article] [PubMed] [Google Scholar]
26. Götz A.W., Williamson M.J., Xu D., Poole D., Le Grand S. and Walker R.C. (2012) J. Chem. Theory Comput., 8, 1542–1555. [PMC free article] [PubMed] [Google Scholar]
27. Haidar J.N., Pierce B., Yu Y., Tong W., Li M. and Weng Z. (2009) Proteins, 74, 948–960. [PMC free article] [PubMed] [Google Scholar]
28. Haidar J.N., Zhu W., Lypowy J., Pierce B.G., Bari A., Persaud K., Luna X., Snavely M., Ludwig D. and Weng Z. (2014) J. Mol. Biol., 426, 1583–1599. [PubMed] [Google Scholar]
29. Harris Daniel T., Singh Nishant K., Cai Q., Smith Sheena N., Vander Kooi Craig W., Procko E., Kranz David M. and Baker Brian M. (2016) Structure, 24, 1142–1154. [PMC free article] [PubMed] [Google Scholar]
30. Hawse W.F., De S., Greenwood A.I., Nicholson L.K., Zajicek J., Kovrigin E.L., Kranz D.M., Garcia K.C. and Baker B.M. (2014) J. Immunol., 192, 2885–2891. [PMC free article] [PubMed] [Google Scholar]
31. Hendsch Z.S. and Tidor B. (1994) Protein Sci., 3, 211–226. [PMC free article] [PubMed] [Google Scholar]
32. Hocking R.R. (1976) Biometrics, 32, 1–49. [Google Scholar]
33. Holler P.D., Holman P.O., Shusta E.V., O'Herrin S., Wittrup K.D. and Kranz D.M. (2000) Proc. Natl. Acad. Sci. USA, 97, 5387–5392. [PMC free article] [PubMed] [Google Scholar]

34. Insaidoo F.K., Borbulevych O.Y., Hossain M., Santhanagopolan S.M., Baxter T.K. and Baker B.M. (2011) J. Biol. Chem., 286, 40163–40173. [PMC free article] [PubMed] [Google Scholar]
35. Janin J. (1999) Structure, 7, R277–R279. [PubMed] [Google Scholar]
36. Jiang L., Kuhlman B., Kortemme T. and Baker D. (2005) Proteins, 58, 893–904. [PubMed] [Google Scholar]
37. Johnson L.A., Morgan R.A., Dudley M.E. et al. (2009) Blood, 114, 535–546. [PMC free article] [PubMed] [Google Scholar]
38. Kass R.E. and Raftery A.E. (1995) J. Am. Stat. Assoc., 90, 773–795. [Google Scholar]
39. Kaufmann K.W., Lemmon G.H., DeLuca S.L., Sheehan J.H. and Meiler J. (2010) Biochemistry, 49, 2987–2998. [PMC free article] [PubMed] [Google Scholar]
40. Kellogg E.H., Leaver-Fay A. and Baker D. (2011) Proteins, 79, 830–838. [PMC free article] [PubMed] [Google Scholar]
41. Koide S. and Sidhu S.S. (2009) ACS Chem. Biol., 4, 325–334. [PMC free article] [PubMed] [Google Scholar]
42. Kortemme T. and Baker D. (2002) Proc. Natl. Acad. Sci. USA, 99, 14116–14121. [PMC free article] [PubMed] [Google Scholar]
43. Lazaridis T. and Karplus M. (1999) Proteins, 35, 133–152. [PubMed] [Google Scholar]
44. Leaver-Fay A., O'Meara M.J., Tyka M. et al. (2013) Amy E.K. (ed), Methods in Enzymology. Academic Press; pp.109–143. [PMC free article] [PubMed] [Google Scholar]
45. Li Y., Moysey R., Molloy P.E. et al. (2005) Nat. Biotechnol., 23, 349–354. [PubMed] [Google Scholar]
46. Linette G.P., Stadtmauer E.A., Maus M.V. et al. (2013) Blood, 122, 863–871. [PMC free article] [PubMed] [Google Scholar]
47. Malecek K., Grigoryan A., Zhong S., Gu W.J., Johnson L.A., Rosenberg S.A., Cardozo T. and Krogsgaard M. (2014) J. Immunol., 193, 2587–2599. [PMC free article] [PubMed] [Google Scholar]
48. Mason D. (1998) Immunol. Today, 19, 395–404. [PubMed] [Google Scholar]
49. McMahan R.H. and Slansky J.E. (2007) Semin. Cancer Biol., 17, 317–329. [PMC free article] [PubMed] [Google Scholar]
50. Moretti R., Fleishman S.J., Agius R. et al. (2013) Proteins, 81, 1980–1987. [PMC free article] [PubMed] [Google Scholar]
51. Morgan R.A., Chinnasamy N., Abate-Daga D. et al. (2013) J. Immunother., 36, 133–151. [PMC free article] [PubMed] [Google Scholar]
52. Oates J. and Jakobsen B.K. (2013) OncoImmunology, 2, e22891. [PMC free article] [PubMed] [Google Scholar]
53. Parkhurst M.R., Yang J.C., Langan R.C. et al. (2011) Mol. Ther., 19, 620–626. [PMC free article] [PubMed] [Google Scholar]
54. Piepenbrink K.H., Blevins S.J., Scott D.R. and Baker B.M. (2013) Nat. Commun., 4, 1–9. [Google Scholar]
55. Piepenbrink K.H., Borbulevych O.Y., Sommese R.F., Clemens J., Armstrong K.M., Desmond C., Do P. and Baker B.M. (2009) Biochem. J., 423, 353–361. [PMC free article] [PubMed] [Google Scholar]

56. Pierce B.G., Hellman L.M., Hossain M., Singh N.K., Vander Kooi C.W., Weng Z. and Baker B.M. (2014) PLoS Comput. Biol., 10, e1003478. [PMC free article] [PubMed] [Google Scholar]
57. Potapov V., Cohen M. and Schreiber G. (2009) Protein Eng. Des. Sel., 22, 553–560. [PubMed] [Google Scholar]
58. Procko E., Hedman R., Hamilton K. et al. (2013) J. Mol. Biol., 425, 3563–3575. [PMC free article] [PubMed] [Google Scholar]
59. Prompers J.J. and Brüschweiler R. (2002) J. Am. Chem. Soc., 124, 4522–4534. [PubMed] [Google Scholar]
60. Reboul C.F., Meyer G.R., Porebski B.T., Borg N.A. and Buckle A.M. (2012) PLoS Comput. Biol., 8, e1002404. [PMC free article] [PubMed] [Google Scholar]
61. Restifo N.P., Dudley M.E. and Rosenberg S.A. (2012) Nat. Rev. Immunol., 12, 269–281. [PMC free article] [PubMed] [Google Scholar]
62. Robinson J., Mistry K., McWilliam H., Lopez R., Parham P. and Marsh S.G. (2011) Nucleic Acids Res., 39, D1171–1176. [PMC free article] [PubMed] [Google Scholar]
63. Rodier F., Bahadur R.P., Chakrabarti P. and Janin J. (2005) Proteins, 60, 36–45. [PubMed] [Google Scholar]
64. Roe D.R. and Cheatham T.E. (2013) J. Chem. Theory Comput., 9, 3084–3095. [PubMed] [Google Scholar]
65. Rossjohn J., Gras S., Miles J.J., Turner S.J., Godfrey D.I. and McCluskey J. (2015) Annu. Rev. Immunol., 33, 169–200. [PubMed] [Google Scholar]
66. Salomon-Ferrer R., Case D.A. and Walker R.C. (2013) Wiley Interdiscip. Rev. Comput. Mol. Sci., 3, 198–210. [Google Scholar]
67. Salomon-Ferrer R., Götz A.W., Poole D., Le Grand S. and Walker R.C. (2013) J. Chem. Theory Comput., 9, 3878–3888. [PubMed] [Google Scholar]
68. Scott D.R., Borbulevych O.Y., Piepenbrink K.H., Corcelli S.A. and Baker B.M. (2011) J. Mol. Biol., 414, 385–400. [PMC free article] [PubMed] [Google Scholar]
69. Scott Daniel R., Vardeman Ii Charles F., Corcelli Steven A. and Baker Brian M. (2012) Biophys. J., 103, 2532–2540. [PMC free article] [PubMed] [Google Scholar]
70. Shoemaker B.A., Portman J.J. and Wolynes P.G. (2000) Proc. Natl. Acad. Sci., 97, 8868–8873. [PMC free article] [PubMed] [Google Scholar]
71. Sinko W., Lindert S. and McCammon J.A. (2013) Chem. Biol. Drug Des., 81, 41–49. [PMC free article] [PubMed] [Google Scholar]
72. Stadinski B.D., Shekhar K., Gomez-Tourino I., Jung J., Sasaki K., Sewell A.K., Peakman M., Chakraborty A.K. and Huseby E.S. (2016) Nat. Immunol., 17, 946–955. [PMC free article] [PubMed] [Google Scholar]
73. Stone J.D. and Kranz D. (2013) Front. Immunol., 4, 1–16. [Google Scholar]
74. Stranges P.B. and Kuhlman B. (2013) Protein Sci., 22, 74–82. [PMC free article] [PubMed] [Google Scholar]
75. Tuffery P. and Derreumaux P. (2012) J. R. Soc. Interface, 9, 20–33. [PMC free article] [PubMed] [Google Scholar]
76. Varela-Rohena A., Molloy P.E., Dunn S.M. et al. (2008) Nat. Med., 14, 1390–1395. [PMC free article] [PubMed] [Google Scholar]
77. Vreven T., Hwang H., Pierce B.G. and Weng Z. (2012) Protein Sci, 21, 396–404. [PMC free article] [PubMed] [Google Scholar]

78. Whitehead T.A., Chevalier A., Song Y. et al. (2012) Nat. Biotechnol., 30, 543–548. [PMC free article] [PubMed] [Google Scholar]
79. Zhao Y., Bennett A.D., Zheng Z. et al. (2007) J. Immunol., 179, 5845–5854. [PMC free article] [PubMed] [Google Scholar]
80. Zinkernagel R.M. and Doherty P.C. (1974) Nature, 251, 547–548. [PubMed] [Google Scholar]
81. Zoete V., Irving M., Ferber M., Cuendet M. and Michielin O. (2013) Front. Immunol., 4, 1–9. [PMC free article] [PubMed] [Google Scholar]

Articles from Protein Engineering, Design and Selection are provided here courtesy of **Oxford University Press**

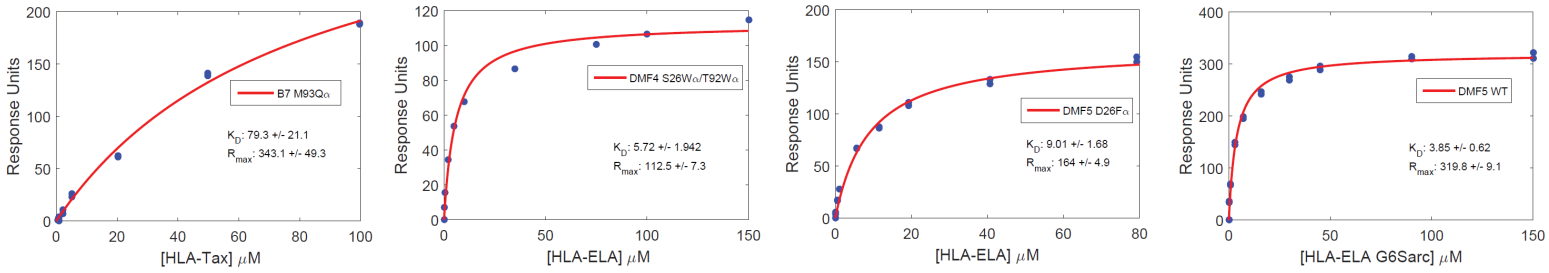

Figure S1. Representative TCR-pMHC SPR binding data for experiments shown in Table S1.

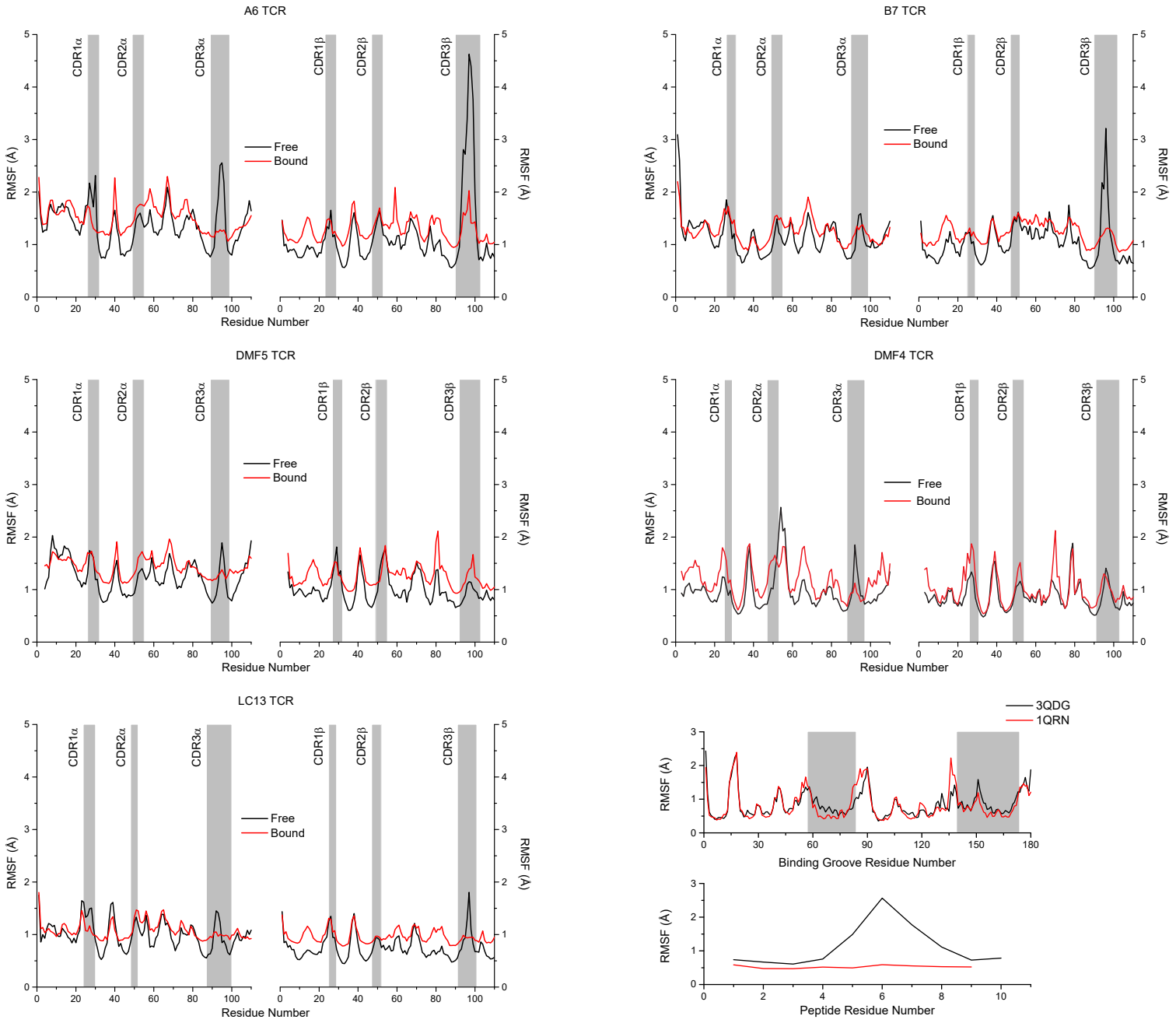

Figure S2. Root mean square fluctuations from MD simulations of free and bound TCRs and pMHC complexes. For TCRs, shaded boxes indicate the locations and values of the six CDR loops. Data for the A6 and B7 TCRs is from Ayres et al., 2016.

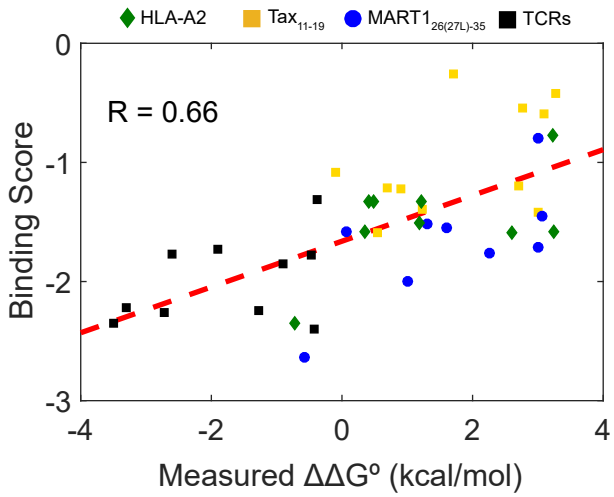

Figure S3. Performance of our improved framework on new TCR mutations, HLA-A2 mutations, and peptide variations when evaluated using binding rather than complex scores. All point mutation data examined in evaluating our new framework, including TCR, peptide, and HLA-A2 data, are plotted together, excluding data used in training. The overall correlation between prediction and experiment with binding scores is 0.66, compared to 0.86 when using complex scores (compare with Fig. 7a).

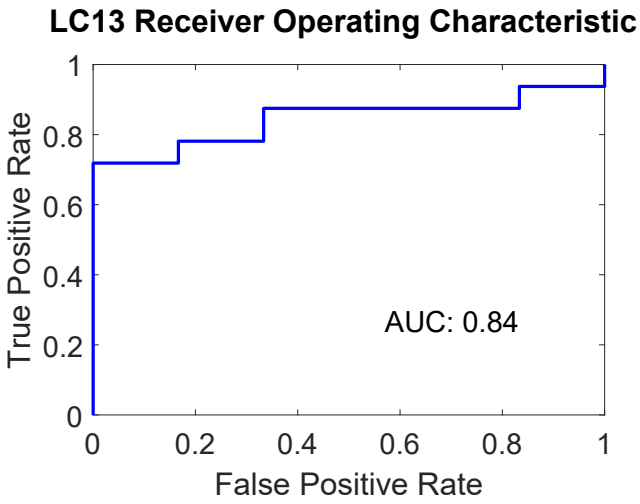

Figure S4. Receiver operating characteristic (ROC) curve for predictions in the LC13 system. The area under the curve is 0.84, indicating good predictive performance when separating affinity increasing mutations from affinity decreasing mutations

Table S1. New binding data for TCR mutations and peptide substitutions in the B7, DMF5, and DMF4 TCR-pMHC interfaces

TCR	Peptide[a]	TCR mutation or peptide substitution	ΔΔG° (kcal/mol)	Error (kcal/mol)
B7	Tax	S27αM	-0.43	0.08
B7	Tax	D30αQ	>2	ND[b]
B7	Tax	S50αY	-0.73	0.09
B7	Tax	M93αE	>2	ND
B7	Tax	M93αQ	1.94	0.1
B7	Tax	Q102αW	0.56	0.14
B7	Tax	P97βW	>2	ND
B7	Tax	G98βF	0.82	0.09
B7	Tax	G99βY	-0.39	0.11
B7	Tax	G99βW	-0.47	0.08
B7	Tax	S27αM/G99βY	-1.15	0.1
B7	Tax	pF3A	2.7	0.02
B7	Tax	pY5A	3.28	0.11
B7	Tax	pY5F	0.55	0.04
B7	Tax	pY8A	2.76	0.07
DMF5	ELA	D26αF	-0.43	0.1
DMF5	ELA	R27αF	-0.3	0.13
DMF5	ELA	K96αW	-0.65	0.12
DMF5	ELA	T54αI	0.33	0.12
DMF5	ELA	S99βF	>2	ND
DMF5	ELA	S99βH	1.48	0.11
DMF5	ELA	S99βI	1.36	0.09
DMF5	ELA	S99βL	2.27	0.03
DMF5	ELA	S99βT	0.4	0.13
DMF5	ELA	pE1A	0.06	0.19
DMF5	ELA	pE1D	1.3	0.26
DMF5	ELA	pE1F	2.26	0.06
DMF5	ELA	pE1Q	1.0	0.03
DMF5	ELA	pI5E	3.07	0.18
DMF5	ELA	pG6-Sarc	-0.58	0.07
DMF5	ELA	pL8A	>3	ND
DMF5	ELA	pT9A	1.6	0.03
DMF5	ELA	pT9W	>3	ND
DMF4	ELA	S26αW	-0.63	0.04
DMF4	ELA	T92αW	-0.38	0.06

[a]Tax = HTLV $Tax_{11\text{-}19}$ (LLFGYPVYV); ELA = $MART\text{-}1_{26(27L)\text{-}35}$ (ELAGIGILTV)

[b]ND = not determined

Table S2. Descriptive breakdown of training set data

Total mutations in training set	94	
Polar/charged WT residues	56	(60%)
Polar/charged mutant residues	24	(26%)
Mutations with polar/charged WT & mutant residues	11	(12%)
Large hydrophobic/aromatic WT residues[1]	14	(16%)
Large hydrophobic/aromatic mutant residue	41	(44%)
Mutations with large hydrophobic/aromatic WT & mutant residues	7	(7%)
Alanine mutations[2]	22	(23%)
Alanine mutations with large/hydrophobic WT residues	5	(5%)

[1]Large hydrophobic/aromatic residues defined as Y/W/L/I/F/M

[2]Excluding mutations with glycine WT residues