



Structure-based, rational design of T cell receptors

V. Zoete¹, M. Irving^{1,2}, M. Ferber¹, M. A. Cuendet^{1,3} and O. Michielin^{1,2,4,5*}

¹ Molecular Modeling Group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

² Lausanne Cancer Center, Lausanne, Switzerland

³ Department of Chemistry, New York University, New York, USA

⁴ Department of Research, University Hospital Center and University of Lausanne, Lausanne, Switzerland

⁵ Ludwig Center for Cancer Research of the University of Lausanne, Lausanne, Switzerland

Edited by:

Bruno Laugel, Cardiff University
School of Medicine, UK

Reviewed by:

Salvatore Valitutti, Institut National de
la Santé et de la Recherche Médicale,
France

Roland Hugo Stote, Centre National
de la Recherche Scientifique, France
Pierre J. Rizkallah, Institute of
Infection and Immunity, School of
Medicine, Cardiff University, UK

*Correspondence:

O. Michielin, Molecular Modeling
Group, Swiss Institute of
Bioinformatics, Quartier
UNIL-Sorge – Batiment Genopode,
CH-1015 Lausanne, Switzerland
e-mail: olivier.michielin@unil.ch

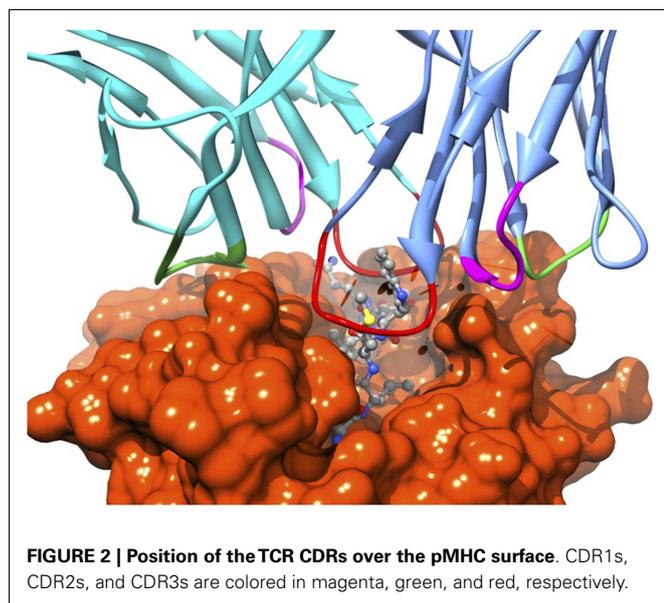
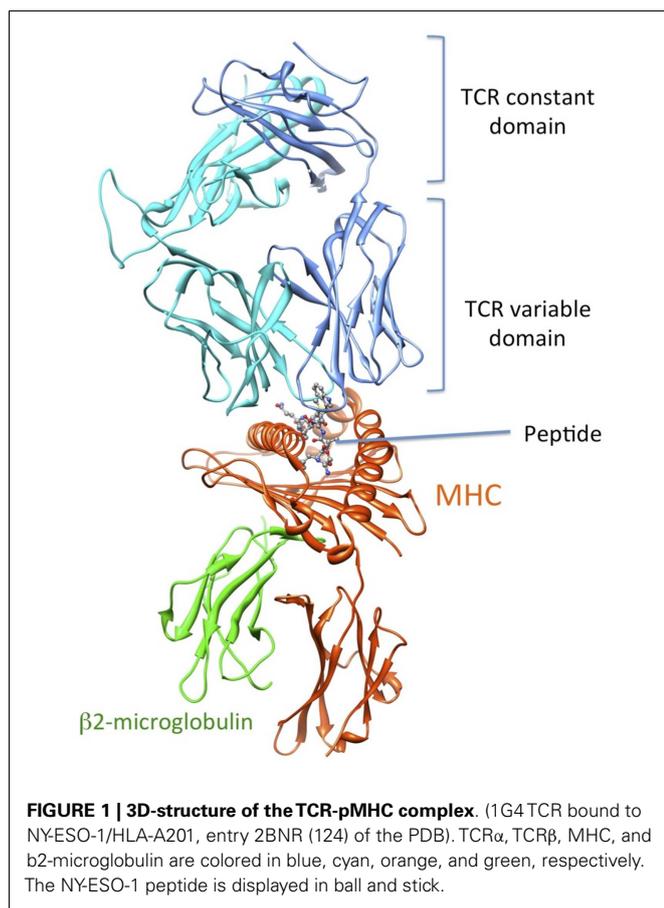
Adoptive cell transfer using engineered T cells is emerging as a promising treatment for metastatic melanoma. Such an approach allows one to introduce T cell receptor (TCR) modifications that, while maintaining the specificity for the targeted antigen, can enhance the binding and kinetic parameters for the interaction with peptides (p) bound to major histocompatibility complexes (MHC). Using the well-characterized 2C TCR/SIYR/H-2K(b) structure as a model system, we demonstrated that a binding free energy decomposition based on the MM-GBSA approach provides a detailed and reliable description of the TCR/pMHC interactions at the structural and thermodynamic levels. Starting from this result, we developed a new structure-based approach, to rationally design new TCR sequences, and applied it to the BC1 TCR targeting the HLA-A2 restricted NY-ESO-1_{157–165} cancer-testis epitope. Fifty-four percent of the designed sequence replacements exhibited improved pMHC binding as compared to the native TCR, with up to 150-fold increase in affinity, while preserving specificity. Genetically engineered CD8⁺ T cells expressing these modified TCRs showed an improved functional activity compared to those expressing BC1 TCR. We measured maximum levels of activities for TCRs within the upper limit of natural affinity, $K_D = \sim 1 - 5 \mu\text{M}$. Beyond the affinity threshold at $K_D < 1 \mu\text{M}$ we observed an attenuation in cellular function, in line with the “half-life” model of T cell activation. Our computer-aided protein-engineering approach requires the 3D-structure of the TCR-pMHC complex of interest, which can be obtained from X-ray crystallography. We have also developed a homology modeling-based approach, TCRRep 3D, to obtain accurate structural models of any TCR-pMHC complexes when experimental data is not available. Since the accuracy of the models depends on the prediction of the TCR orientation over pMHC, we have complemented the approach with a simplified rigid method to predict this orientation and successfully assessed it using all non-redundant TCR-pMHC crystal structures available. These methods potentially extend the use of our TCR engineering method to entire TCR repertoires for which no X-ray structure is available. We have also performed a steered molecular dynamics study of the unbinding of the TCR-pMHC complex to get a better understanding of how TCRs interact with pMHCs. This entire rational TCR design pipeline is now being used to produce rationally optimized TCRs for adoptive cell therapies of stage IV melanoma.

Keywords: molecular modeling, protein-engineering, TCR, TCR-pMHC, immunotherapy, adoptive transfer, cancer

INTRODUCTION

Recognition by the CD8⁺ T cell receptor (TCR) of immunogenic peptide (p) presented by class I major histocompatibility complexes (MHC) is a key event in the specific immune response against virus-infected cells or tumor cells. Binding of the TCR to the pMHC complex leads to T cell activation and killing of the target cell (1). The TCR is composed of two chains, α and β , that pair on the surface of the T cell to form a heterodimeric receptor on the surface of the T cell. Each chain is composed of a constant domain that anchors the protein in the cell membrane and of a variable domain that confers antigen recognition (Figure 1). The TCR contacts pMHC molecules via the 6 complementarity-determining regions (CDR), three each from the α and β chains

(Figure 2). These CDRs constitute the hypervariable regions of the two V domains, called $V\alpha$ and $V\beta$ (2–4). They are generated by somatic gene rearrangement and negatively selected in the thymus against reactivity with endogenous pMHCs. CDR3 α and CDR3 β are the most diverse regions of the TCR and thus play a major role in antigen specificity. The CDR1 and CDR2 loops of the α and β chains predominantly make contact with the MHC molecule. The strength of the interaction between TCR and pMHC has been shown to play an important role in the T cell activation (5–9). However, the kinetics of the TCR/pMHC interaction is also determinant in T cell activation (10, 11). Consequently, understanding the biophysical properties of the TCR/pMHC interaction is of great interest for the prediction of the T cell activation, and for



the rational TCR optimization toward improved adoptive transfer cancer therapy (12, 13).

This review will focus on the different computer-aided techniques we developed and used to study the TCR-pMHC complex

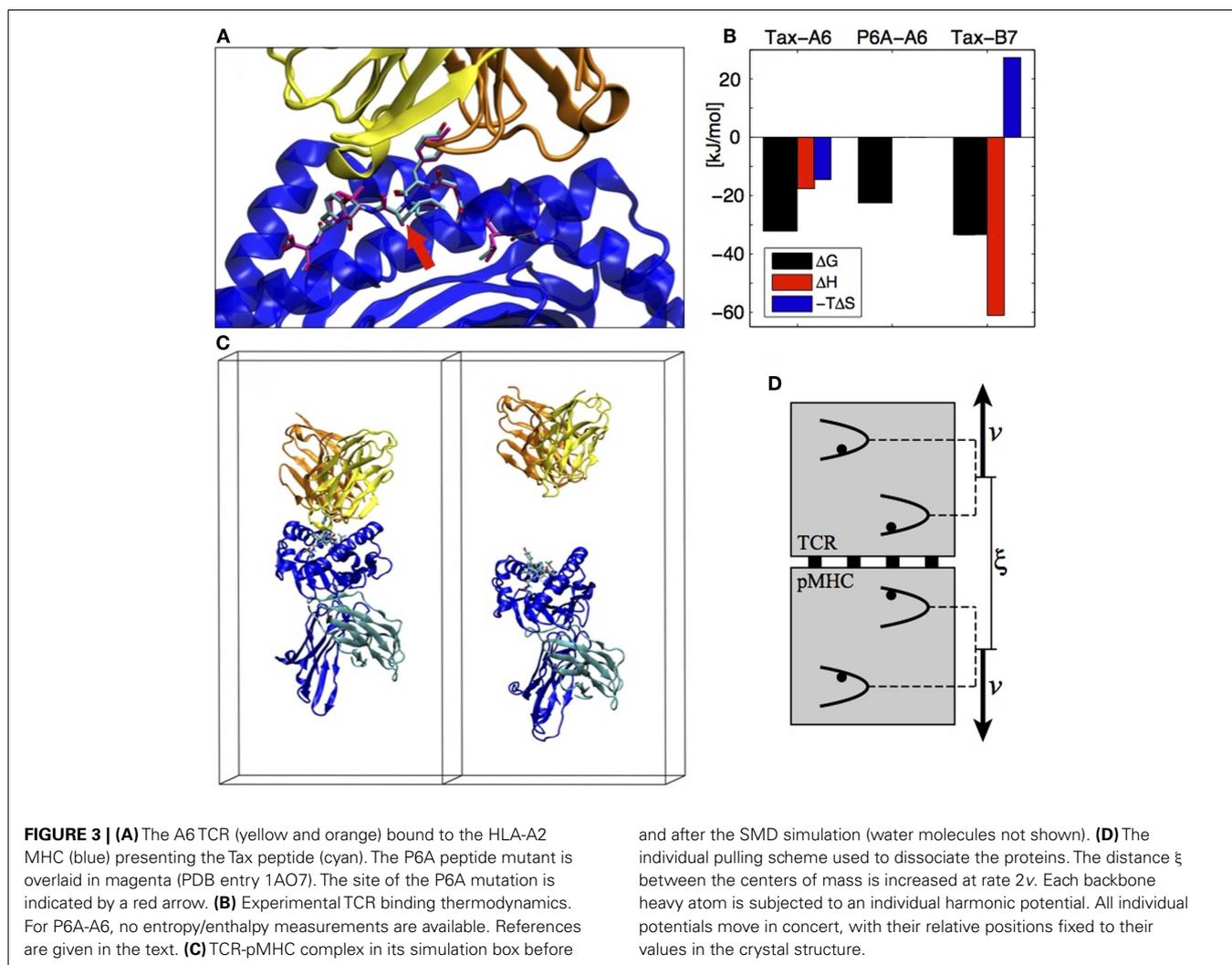
from a structural and thermodynamic point of view. First we present the results obtained by steered molecular dynamics (SMD) simulations of molecular recognition events occurring during the TCR-pMHC complex formation. Second, we describe the different approaches we developed to derive structural models of TCR-pMHC complexes for a large TCR repertoire. Third, we summarize our approaches to estimate the binding free energy for the TCR association to pMHC. Finally, we present our *in silico* structure-based protein-engineering approach that enables the fine-tuning of TCR-pMHC binding parameters.

INVESTIGATING TCR-pMHC INTERACTIONS USING STEERED MD SIMULATIONS

The structure of ~66 TCR-pMHC complexes are known to date (14) among which ~25 are unique complexes. The compared features of these structures were the object of extensive reviews (1, 15, 16). In addition, binding kinetics and thermodynamics of many complexes have been measured (17). The general picture that emerges is that T cell activation requires TCR-pMHC binding to fall within a certain range of affinity and kinetics. However, clear structural determinants of TCR specificity have remained elusive. A single mutation can change a peptide from agonist to antagonist, but the same TCR can recognize various peptides with different binding modes. Thus TCRs can be exquisitely specific while displaying a high degree of cross-reactivity. In addition, TCRs use extremely varied thermodynamic strategies to bind to pMHCs, ranging from entropy-favored to entropy-opposed.

A system that exemplifies both the sensitivity of TCR recognition and its potential for cross-reactivity with different binding strategies is the Tax nonapeptide (LLFGYPVYV) from the HTLV-1 virus presented by the HLA-A0201 MHC. This pMHC is a strong agonist for the A6 TCR but the P6A peptide mutant (Pro replaced by Ala at position 6, see **Figure 3A**) dramatically reduces the binding affinity and abrogates T cell activation (18). On the other hand, the B7 TCR, which has the same α chain as A6 but a different β chain, is also activated by the Tax peptide presented by the same MHC. B7 binds with an affinity similar to A6, but the binding is entropically opposed, whereas A6 binding is entropically favored, outlining a completely different binding mechanism (19) (see **Figure 3B**).

Detailed aspects of protein-protein interactions can be characterized by SMD simulations, in which the dissociation is actuated by an external force acting on the protein. A typical reaction coordinate for protein-protein dissociation is the distance between the centers of mass of each protein. In the following, we call this reaction coordinate ξ . In the case of the TCR-pMHC, we assumed that the dissociation happens in the direction perpendicular to the cellular membranes (see **Figure 3C**). We also assumed that there is no substantial conformational rearrangement upon dissociation (except possibly in the CDRs and in the peptide), which is supported by the similarity of X-ray structures of bound and unbound TCRs or pMHCs. To enforce these assumptions during the SMD, we devised the individual pulling scheme (20) in which each non-H atom (except CDRs and peptide) is subjected to an individual harmonic potential. As shown on **Figure 3D**, the center of mass distance ξ is increased by collectively shifting the reference positions of the individual potentials. For each TCR-pMHC complex,



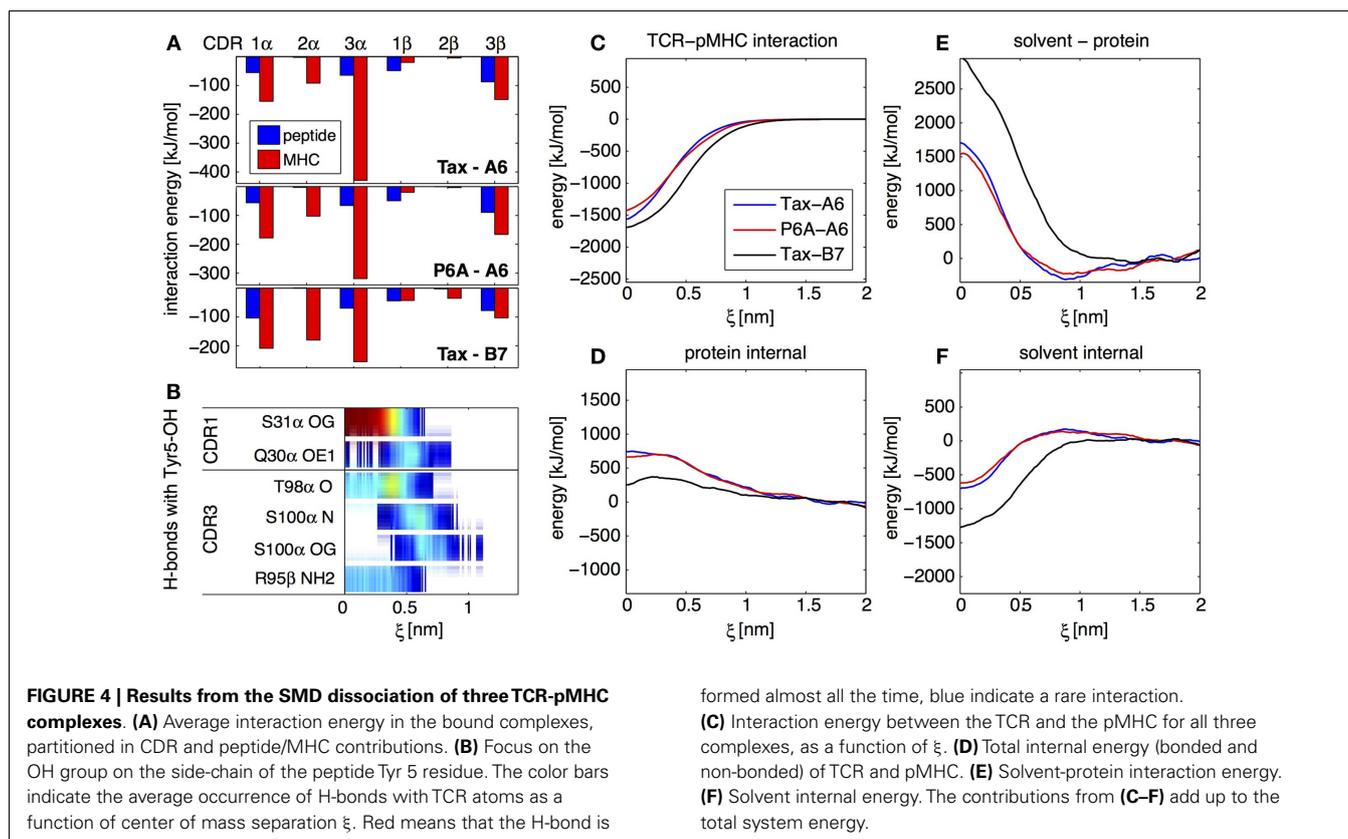
we performed about 150 unbinding trajectories of 4 ns each during which ξ is gradually increased by 2 nm from the bound state distance to reach a final conformation as shown on the right panel of **Figure 3C**. The simulations were performed with the Gromacs software (21) using the Gromos 45a3 force field (22) with explicit water molecules.

With about 150 trajectories we were able to obtain converged ensemble averages of many observables at any given protein-protein separation, including for highly fluctuating quantities involving solvent molecules. In particular, we established maps of H-bonds or of non-polar contacts for all residues of the TCR and pMHC as a function of ξ . As an example, **Figure 4B** shows an H-bond occurrence map for the OH group on the Tyr5 side-chain of the Tax peptide in the Tax-A6 system. We see that α -S31 is the main TCR H-bonding partner in the bound state, but that new H-bonds are formed with α -S100 in the transition state. Overall, our simulations have shown that the number and diversity of H-bonds occurring in a protein complex largely exceeds what is apparent from the crystal structures. Using this methodology, detailed maps such as **Figure 4B** can be established for any interaction to

any atom in the system, depending on the biological question of interest.

For all three TCR-pMHC complexes, we monitored energy variations in different parts of the system upon dissociation. **Figure 4C** shows that there are differences in TCR-pMHC interaction energies between the complexes. But these are largely compensated by effects in the internal protein reorganization energy (**Figure 4D**), solvent-protein interaction energy (**Figure 4E**), and solvent internal energy (**Figure 4F**). As a general lesson for protein-protein interactions, we retain that the solvent plays a key role in two different ways. First, variations of solvation energies exceed contributions from the proteins themselves upon binding. Second, specific water molecules trapped at the interface can influence the binding mechanism and thermodynamics (data not shown). In the present case, these two aspects happen to be also the two major factors differentiating A6 and B7 binding.

Focusing on the bound state of the TCR-pMHC complexes, the decomposition of the average interaction energy among CDRs brings valuable insights. **Figure 4A** illustrates the differences between the A6 and B7 binding modes, with a much less prominent



contribution of CDR3 α for B7. Generally for all complexes, **Figure 4A** shows that, while CDR2 interacts mostly with MHC, both CDR1 and CDR3 interact equally with peptide and MHC, even at large distance, as confirmed by our H-bond occurrence maps (not shown). Therefore, our simulations do not support the two-step model (23) for TCR engagement, in which the CDR1 and CDR2 preferentially contact the MHC at large distance, while the CDR3 establishes final contacts to “read” the peptide mainly at short distances.

Overall, although the P6A-A6 complex has a very different affinity compared to the wild-type Tax-A6, both complexes share very similar features in terms of specific H-bonds or energy contributions. On the other hand, the Tax-B7 complex has a binding affinity similar to that of Tax-A6, but uses a completely different binding mechanism. The B7 TCR creates a very different set of H-bonds and hydrophobic contacts to the pMHC and makes a very different usage of the solvent, which reflects in a different partition of the binding energy. In retrospect, as noted previously by Baker and coworkers (24), it is not so surprising to observe active TCRs with very different binding mechanisms. Indeed, if TCRs are issued from random sequence variation and selection upon pMHC binding affinity and kinetics only, each TCR is likely to adopt its own unique pMHC binding strategy as long as it matches these criteria.

TCR-pMHC HOMOLOGY MODELING

PIONEERING STUDY OF TCR-pMHC HOMOLOGY MODELING

The recent development and use of experimental techniques to determine sequences of TCRs that bind to a pMHC complex (25),

led to the collection of large repertoires of TCR sequences with given pMHC specificities (26, 27). Understanding the selection mechanism that causes this gene usage can be facilitated by the introduction of structural information regarding the underlying TCR-pMHC complexes. This information can be used to identify conserved 3D binding motifs that are not obvious from repertoire sequences alone (28), to suggest explanations regarding the impact of TCR mutation on its affinity for given pMHCs (29) and ultimately to support the rational engineering of TCRs with particular binding properties (7, 30). Experimental structural techniques such as X-ray crystallography or NMR provide direct and valuable information regarding the 3D-structures of macromolecules. Unfortunately, they require the production of the protein, can be time consuming, and are thus hardly applicable to the analysis of large repertoires of tens to hundreds of TCRs. According to the 3D-structure database of the international ImMunoGenetics information system [IMGT/3D-structure-DB (31, 32)] the 3D-structure of 66 TCR-pMHC complexes have been determined experimentally so far. This number is negligible compared to the vast TCR diversity created by genetic rearrangements of the TCR V, D, and J genes. Indeed, the number of unique TCR β chains in blood has been estimated to be of the order of 10^6 (33, 34). There is thus a need for tools able to predict the 3D-structure of TCR-pMHC complexes from the amino acid sequences of their components.

The pioneering work of Michielin et al. (29) provided a remarkable demonstration of the feasibility and the predictive ability of TCR-pMHC 3D-structure modeling. The authors used a murine

T1 TCR, specific for a photoreactive derivative of the *Plasmodium berghei* circumsporozoite (PbCS) 253–260 nonapeptide presented by the K_D class I MHC (35). Fifty mutants involving the TCR's CDR, the MHC's $\alpha 1$ and $\alpha 2$ helices and the peptide were prepared and the association constants between the TCR and the pMHC were measured (35). A first homology model was built for the wild-type TCR-pMHC complex with the MODELLER program (36–38), using the four TCR structures available at that time (39–42) and the structure of H-2Kb MHC (43) as templates. These structures provided good templates for the β -sheet framework of the TCR, and of the α helices/ β -sheet of the MHC grooves. The high secondary structure content of these regions imposed strong restrictions on their backbone conformation according to the MODELLER algorithm, which facilitates the modeling of those parts. In contrast, the CDR loops of TCRs have a very low level of sequence identity and no specific secondary structure, which obviously limits the efficiency of modeling by homology. The CDR loops conformation obtained in the first homology model were thus refined using a simulated annealing technique (44), followed by clustering the generated conformations based on their relative Cartesian coordinate root mean square deviation (RMSD). A final conformation was chosen from a well-populated, low-energy cluster, whose structure was compatible with the experimental mutational data. All but three of the 50 mutations found qualitative explanation in the model in terms of breaking of a significant TCR/pMHC interaction. In addition, the model suggested that a TCR pocket could form upon binding to accommodate the peptide hapten, explaining the high level of affinity of the T1 TCR for this pMHC ($K_D \sim 10$ nM), and demonstrating predictive capabilities for the modeling approach that go beyond reproducing only the structural features present in the templates. Since more X-ray structures of TCR-ligand complexes are continuously determined,

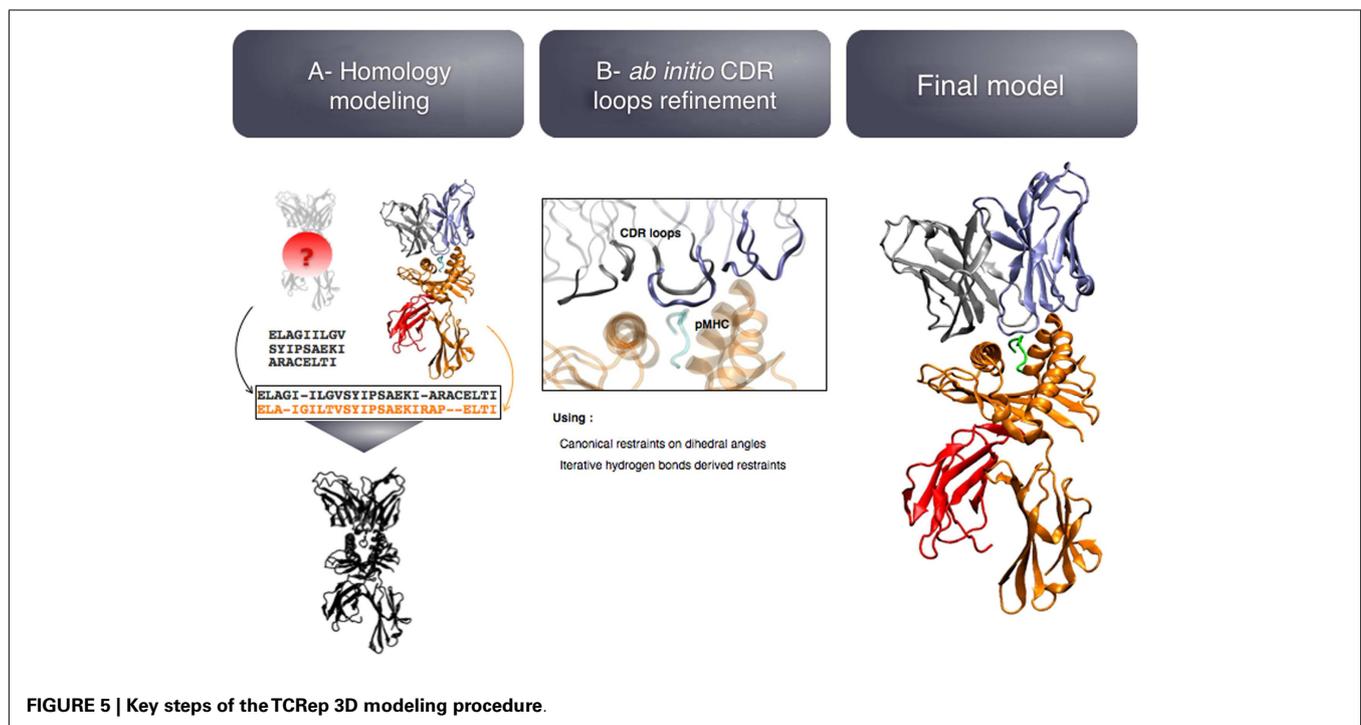
it could be expected that the range of applicability and the accuracy of such a modeling approach would improve, since there is no limitation to the number of simultaneous templates that can be used.

TCRep 3D

The first study described above (29) led to the development of TCRep 3D, as a generalization of the TCR-pMHC modeling approach (28). TCRep 3D is an approach dedicated to the prediction of high-quality 3D-structures that can provide a functional insight on the interaction between a TCR and a pMHC. It includes by design minimal input and optimal automation, to analyze wide sets of sequences of TCRs belonging to a common TCR repertoire.

The modeling pipeline is composed of two modules (see **Figure 5**): (i) homology modeling of the TCR-pMHC complex and (ii) *ab initio* CDR loops structure optimization. First, the user provides the sequence of the target complex and a list of preferred templates. By default, all the TCR-pMHC templates of the Protein Data Bank are used. The global structure of the complex is modeled by homology. It could be possible to couple this step to computer-aided approaches for the docking of peptide antigens into MHC molecules (45), in case the peptide binding mode could not be correctly predicted by homology modeling. Each CDR loop is then subsequently refined while the rest of the complex remains rigid. The MODELLER (38) software is used for the two modules.

The originality of TCRep 3D resides in the specific bias that we impose to the system during the structural sampling. Canonical restraints available from the literature (46) are added to the backbone dihedral angles of CDR1 and 2 to limit the conformational space accessible to the loops. We developed an iterative sampling



method that identifies potential hydrogen bonds between TCR and pMHC, and converts them into modeling restraints. The scoring function was adapted accordingly, to favor structures that satisfy most of the canonical restraints, and display potential hydrogen bonds. We demonstrated that TCRRep 3D is significantly more efficient than common loop modeling approaches in predicting CDR loops conformations.

At the time of the study, TCRRep 3D produced one TCR-pMHC structure in 7 days on a single CPU. However, the modeling can be parallelized on a computing grid, and the computation time scales efficiently with the number of CPUs used, allowing the user to quickly model a large number of sequences.

TCRRep 3D has been successfully applied to experimentally determined sets of sequences of TCRs that recognize given cancer epitopes.

- (a) In a study on HLA-A*0201/Melan-A-specific CD8 T cells (47), the modeling of the TCR-pMHC 3D-structure revealed the structural feature that explained how two distinct sets of TCR performed differently in recognizing a naturally occurring decamer variant of the Melan-A peptide. One of the TCR subsets could not make proper interactions with the glutamic acid at position 1 of the peptide because of the location and structural properties of the CDR1 α (see **Figures 6A,B**).
- (b) The analysis of HLA-A*0201/NY-ESO-1_{157–165} specific CD8⁺ T cells from five melanoma patients showed a preferential usage of three V β genes. Additionally, experimental evidence on the importance of the Met4-Trp5 pair of the NY-ESO-1_{157–165} antigen were found, suggesting that those two contact residues make critical interactions with the TCR, regardless of the gene segment usage (26). The modeling of the corresponding TCR-pMHC structures revealed a striking mechanism of selection through the presence of a single conserved glycine residue situated in the center of all CDR3 β . An *in vitro* experimental functional study of mutations of this amino acid combined with *in silico* modeling of several mutants was performed. All mutations resulted in dramatic structural changes associated with complete experimental loss of affinity of the TCR to NY-ESO-1/HLA-A*0201 (28).

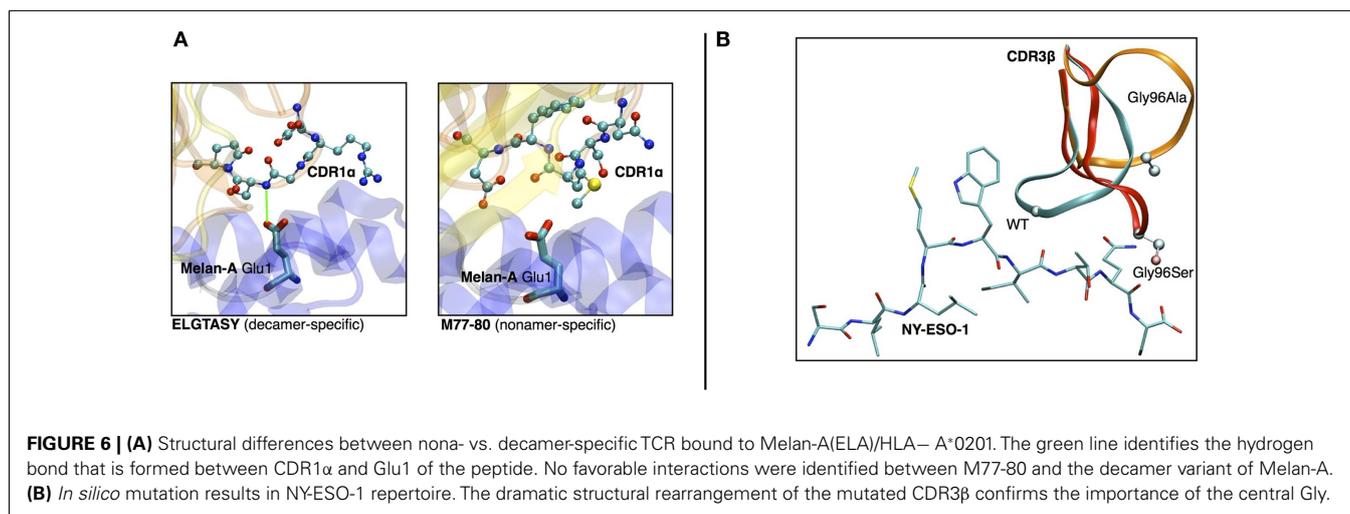
LONG-RANGE DRIVING FORCE FOR TCR ORIENTATION

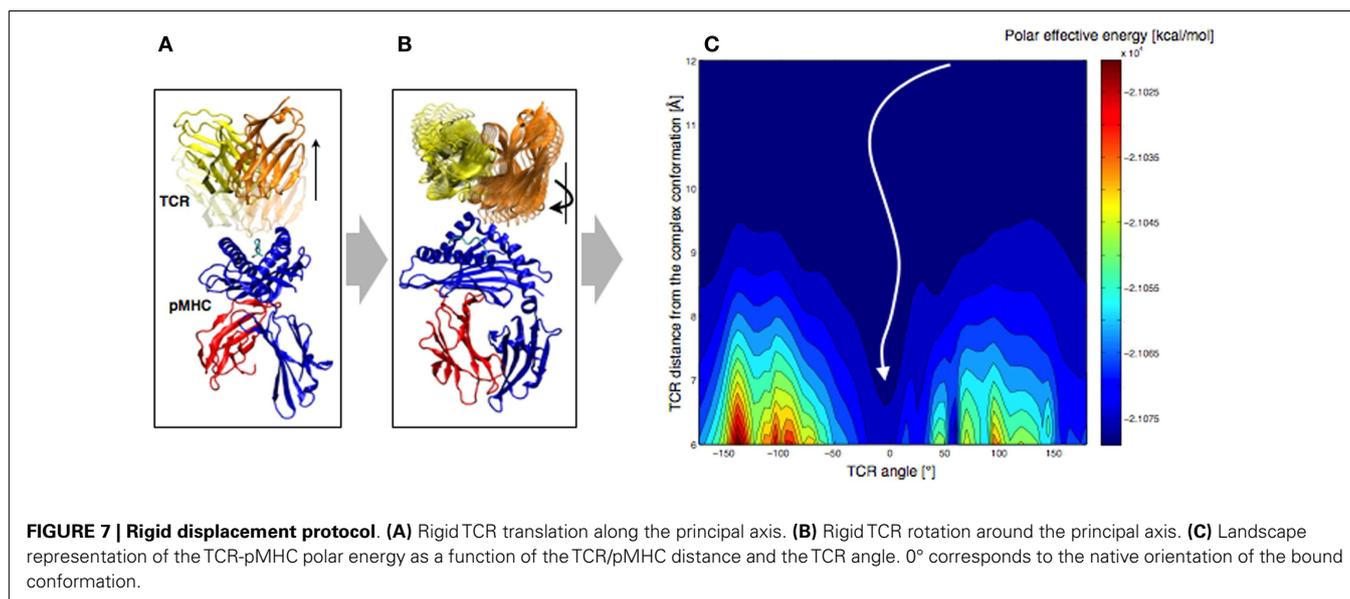
Over the years, successive releases of TCR-pMHC crystal structures have revealed the variety of native binding orientations that the TCR can adopt. Recent studies reported a range of more than 45° in the TCR binding angles relative to the MHC (48), depending on the peptide, the MHC, and the α/β pairing of the TCR. Although the challenge of TCR binding mode prediction has been recurrently discussed, only a few studies have focused on predicting the actual binding mode of given TCR-pMHC (49, 50). Therefore, all methods and applications relied on the existence of at least one TCR-pMHC crystal structure.

In order to understand the molecular basis that governs TCR orientation upon binding, we tested a simplified rigid approach on all published TCR-pMHC crystal structures (48), which allowed scanning quickly multiple orientations of the TCR relative to the pMHC. In this approach, the TCR was moved 6–12 Å away from the pMHC molecule along the TCR principal axis (see **Figure 7A**). Subsequently, the TCR was rotated around that same axis until a complete revolution was obtained (see **Figure 7B**). The effective energy of the system was computed every 5°, as the sum of the intermolecular energy and the solvation free energy, using the CHARMM22 force field (51, 52) in combination with the FACTS implicit solvation model (53).

We demonstrated that the sum of the Coulomb interaction and the electrostatic solvation energies is sufficient to identify the native TCR orientation as the energetic minimum upon rotation (see **Figure 7C**). Importantly, despite the rigid-body simplification, the results were robust upon small structural variations of the TCR such as changes induced by MD simulations. We also tested our approach on crystal structures of unbound TCRs, which were confronted to pMHCs. Accurate energy minima were also identified, suggesting that perfect shape complementarity is not required to obtain a reliable signal. The long-distance interactions during the TCR approach appear to be independent of the binding process itself, since the binding orientation is reliably identified without considering either short-range energy terms or CDR induced fit upon binding.

Furthermore, we decomposed the effective energy into per-residue contributions, in an approach that is similar to the





Molecular Mechanics – Generalized Born Surface Area (MM-GBSA) energy decompositions (54). The contributions of structural sub-groups to the profile of the TCR/pMHC interaction energy during rotation were calculated, to estimate their role in the overall orientation. Results showed that most of the driving force (>90%) leading to the orientation of the TCR is defined by CDR1,2/MHC interactions. This is in agreement with previous observations, revealing a ring of charged residues at the pMHC interface, which interacts with CDR1,2 with complementary charges (48). We reported that the role of the CDR3/peptide interaction is of lesser importance at long-distance.

In turn, such knowledge of the structure may be used as a preliminary approach in the process of modeling protein/protein interactions. More specifically, the rigid search for an energetic minimum upon TCR rotation may become a complementary module of TCRep 3D, to search for the correct binding mode, after modeling the TCR and the pMHC independently. We attempted to predict the binding mode of the A6 TCR with tax/HLA-A*0201, after modeling the TCR by homology. The effective energy minimum upon rotation was computed for 500 homology models, and we obtained an average shift of 12.2° from the orientation of the crystal structure. This demonstrated the potential of the approach as a component of a TCR-pMHC structural prediction pipeline (55). The approach is also easily applicable to other types of protein complexes, provided that the association is also driven by long-range electrostatic interactions.

FREE ENERGY CALCULATIONS

APPLICATION OF THE THERMODYNAMIC INTEGRATION METHOD FOR TCR-pMHC BINDING FREE ENERGY DIFFERENCES

T cell receptor recognition can exhibit exquisite specificity upon single peptide mutation. In the A6/Tax/HLA-A0201 complex described above (Tax-A6), mutating the Ala at position 6 to a Pro (P6A) turns the Tax peptide from a strong agonist into a weak antagonist. These systems were extensively studied experimentally (18, 56) and the binding free energy difference

between the Tax-A6 and the P6A-A6 complexes was found to be $\Delta\Delta G = 2.90 \pm 0.20$ kcal/mol. (see **Figure 3B**). These results are difficult to rationalize from the structure alone, as there is almost no difference between the conformations of the Tax-A6 and P6A-A6 complexes (see **Figure 3A**, red arrow). To gain a better understanding of the effect of the mutation on TCR recognition, we used free energy simulation to analyze in detail the origin of the binding free energy difference (57).

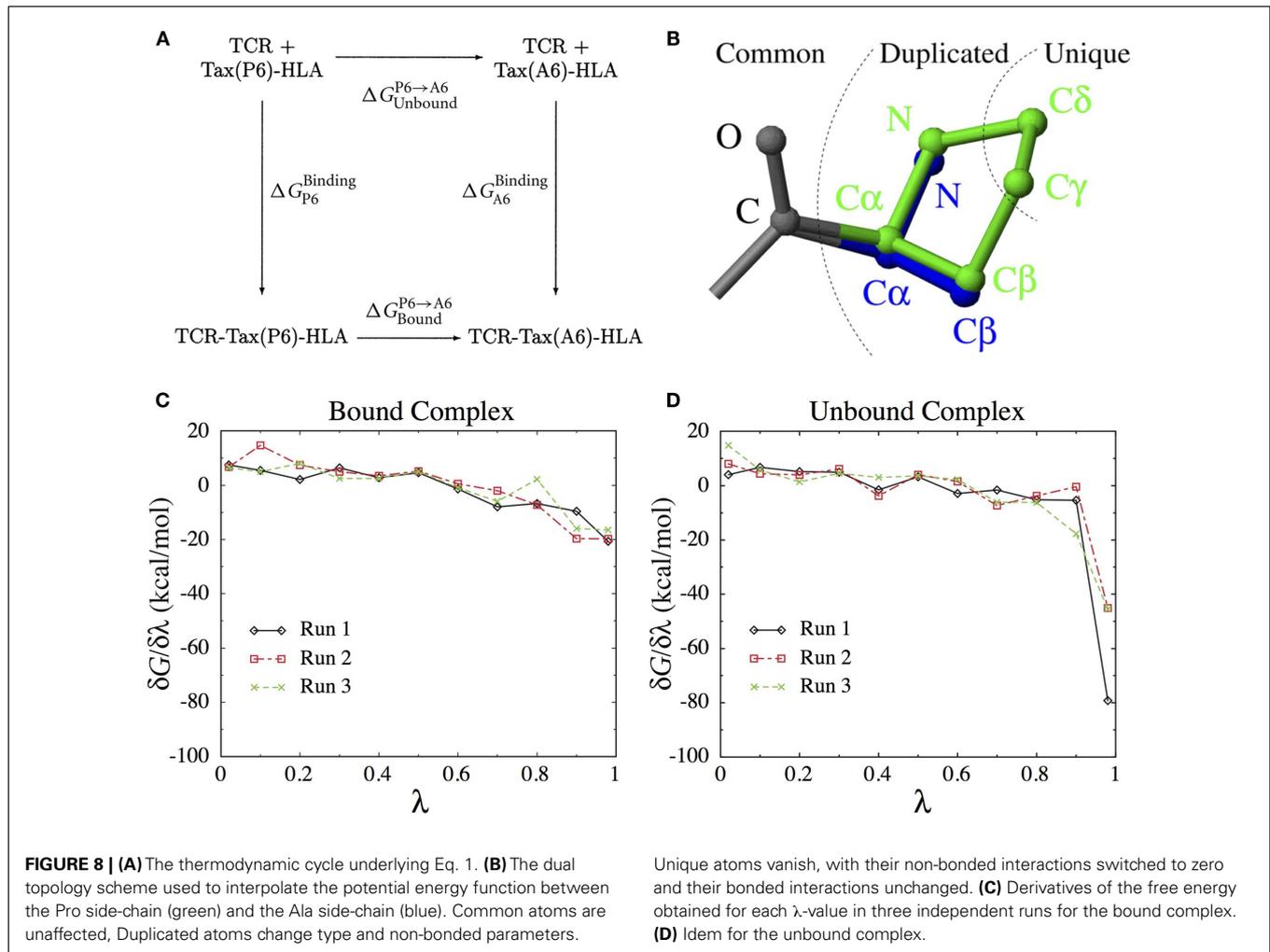
As we will see below, calculating the binding free energy by simulating the entire TCR-pMHC unbinding process itself is difficult. Instead, the present method uses the thermodynamic cycle shown in **Figure 8A** to reformulate the problem,

$$\Delta\Delta G_{P6A}^{\text{Binding}} = \Delta G_{P6}^{\text{Binding}} - \Delta G_{A6}^{\text{Binding}} = \Delta G_{\text{Bound}}^{P6 \rightarrow A6} - \Delta G_{\text{Unbound}}^{P6 \rightarrow A6} \quad (1)$$

This means that we can obtain $\Delta\Delta G_{P6A}^{\text{Binding}}$ by computing the P6 → A6 mutation free energy in both the unbound and the bound states. Among the different methods available to calculate mutation free energy differences (58, 59), we chose thermodynamic integration (60, 61). We define an interpolated potential energy function $U(r, \lambda)$ that is equivalent to the potential energy function of the wild-type for $\lambda = 0$ (Pro) and to that of the P6A mutant for $\lambda = 1$ (Ala). The free energy difference can be obtained through

$$\Delta G_{\text{Bound}}^{P6 \rightarrow A6} = \int_0^1 \left\langle \frac{\partial U(r, \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda.$$

Here, $\langle \cdot \rangle_{\lambda}$ represents an ensemble average at fixed λ . In practice, we perform a set of simulations at discrete λ -values and evaluate the integral above numerically. The number and position of λ -values required for accurate integration depends on the smoothness of the $\langle \frac{\partial U}{\partial \lambda} \rangle_{\lambda}$ function. To construct an appropriate interpolating potential energy function, a dual topology scheme was used, as shown on **Figure 8B**. For vanishing atoms, only the non-bonded interactions are scaled, while the bonded interactions are left unchanged (62). As we did not use soft-core potentials



(63) for vanishing atoms, special care was taken in the limit of $\lambda \rightarrow 0$ and $\lambda \rightarrow 1$ to deal with the singularities of the Coulomb and Lennard-Jones potentials.

Starting from the crystal structures for both the unbound (64) and the bound complex (18), MD simulations were performed in the CHARMM program (51) with the CHARMM22 force field (52). The proteins were locally solvated in a sphere of 16 Å surrounding the peptide using the stochastic boundary method (65). The set of λ -values used were $\lambda = 0.02, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.98$. After 100 ps of initial equilibration, data was collected for 30 ps at each λ -value, separated by 10 ps equilibration time after each λ update. Three simulations with independent initial velocities were produced.

The simulations were structurally stable with average RMSD of non-H atoms no greater than 0.8 Å with respect to the crystal structures. The average free energy derivatives obtained in the three different runs are shown on **Figures 8C,D** for the bound and unbound states, respectively. In the unbound state (**Figure 8D**), the derivative takes very large values close to $\lambda = 1$, due to the interaction of the vanishing Pro atoms with the solvent molecules. This does not happen in the bound state, because the vanishing

atoms are concealed from the solvent in the void created by the protein pocket.

Three different schemes were tested for the extrapolation to $\lambda = 0$ and $\lambda = 1$, using linear, quadratic or $\lambda^{-3/4}$ functions. The integration over all λ -values was performed using the trapezoidal rule. The final result is $\Delta \Delta G_{P6A}^{\text{Binding}} = 2.9 \pm 1.1$ kcal/mol, which compares very favorably with the experimental value of 2.9 ± 0.2 kcal/mol.

One of the major strengths of the method lies in the linearity of Eq. 1, which allows decomposing as a sum of contributions from different types of interactions and/or of different parts of the system. Given that the total free energy difference is in good agreement with experiment, there is a good chance that the decomposition provides meaningful insights on the mechanisms leading to TCR specificity.

A notable contribution to $\Delta \Delta G$ (+0.64 kcal/mol) arises from the difference in solvation free energy of the mutated residue: in the unbound structure, the Tax P6 residue is solvent-exposed with around 35% of its surface accessible to water molecules. The A6 mutant has fewer exposed hydrophobic groups, which entails a more modest solvation penalty in the unbound state.

This stabilizes the unbound state of the P6A mutant relative to the wild-type, which in turn makes P6A-A6 binding less favorable. The rest of the peptide contributes a modest 0.38 kcal/mol to $\Delta\Delta G$.

The total contribution of the TCR in the bound state is around +0.8 kcal/mol destabilizing the P6A mutant, the most significant part of which is due to the CDR3 α loop. Most of this energy arises from the van der Waals term, in accord with the fact that the TCR provides good surface complementarity for the hydrophobic side-chain of the Pro residue. Since the pocket is already present in the Tax-A6 complex, there is no large free energy cost needed to induce it, in contrast to what is found in the solvent. The TCR residues that contribute the most to the TCR specificity for the wild-type peptide are N30 from the CDR1 α , D99, and S100 from the CDR3 α , and G97, L98, A99, G100, G101 from the CDR3 β .

The most important contribution to $\Delta\Delta G$ (1.26 kcal/mol) arises from the difference in interactions with the MHC. This is due to a conformational change that takes place in the Tax P6 region upon TCR engagement. In the Tax system, the cost of this conformational change is balanced by a very favorable interaction of the Pro ring with hydrophobic residues of the MHC groove, which does not take place with the shorter Ala side-chain. This is an example of how a conformational change taking place along the physical binding pathways translates into a free energy contribution along the alchemical pathways.

Overall, it emerges that the total binding free energy difference between the wild-type and the mutant peptide consists of four contributions that are similar in magnitude. The self-interaction of the peptide and the change in the interaction between the peptide and the three portions of its environment (TCR, HLA-A2, and solvent) all contribute between 0.5 and 1.2 kcal/mol to stabilizing the wild-type complex. This important result was not evident from the X-ray structures or the experimental data. Interestingly, these calculations show that accurate free energy differences could be obtained although most of the complexity of this system was ignored in our relatively short simulations including only the mutated side-chain and its local environment.

ASSESSING THE APPLICABILITY OF THE JARZYNSKI IDENTITY TO CALCULATE TCR-pMHC BINDING FREE ENERGY PROFILES

In classical thermodynamics, the dissipative work $W_{A \rightarrow B}$ needed to bring a system from state A to state B is greater than the free energy difference ΔG_{AB} between the two states, with equality only in adiabatic conditions. Conversely, a recent result in non-equilibrium statistical mechanics, the Jarzynski identity (JI) states that (66, 67)

$$e^{-\beta\Delta G_{AB}} = \left\langle e^{-\beta W_{A \rightarrow B}} \right\rangle_0.$$

Here, $\beta = (k_B T)^{-1}$, with T the temperature and k_B the Boltzmann constant. The average $\langle \cdot \rangle_0$ is taken over canonically distributed initial conditions in state A. The JI was proven to hold in the case of thermostated molecular dynamics (66–69). The JI was applied with some success to simulations of small molecular systems (68–71). Given the biological importance of the TCR-pMHC system, we were compelled to determine if the JI could be employed to calculate protein–protein binding free energy profiles from SMD trajectories of large protein–protein complexes

(72). The free energy profile, or potential of mean force (PMF), is very relevant because in addition to the total binding free energy it provides estimates of k_{on} and k_{off} from the free energy barrier height.

In SMD, an external potential of the form $u(\xi(r), t)$ is used to steer the system from position A at time 0 to B at time t along the reaction coordinate ξ . This potential can be a simple harmonic potential or take a more complex form such as in the individual pulling scheme (72) presented in Section “Investigating TCR-pMHC Interactions Using Steered MD Simulations.” If we are interested in the PMF along ξ , the JI can be written as (73),

$$e^{-\beta G(\xi)} \propto \left\langle \delta[\xi(r) - \xi] e^{-\beta W(t)} \right\rangle_0.$$

Importantly, $W(t)$ is the work accumulated by the perturbed system (including u), defined as (74),

$$W(t) = \int_0^t dt \frac{\partial u}{\partial t}(\xi(r), t).$$

For each of the Tax-A6 and P6A-A6 complexes, we performed 150 trajectories starting from independent conformation of the bound complex. The TCR and pMHC centers of mass separation was increased by 2 nm over 4 ns. The resulting work profiles are shown on **Figure 9A** and the distribution of final work values is shown on **Figure 9B**. As expected if the system is not too far from equilibrium (71, 75), the distribution (in its central part) close to a Gaussian. To obtain the PMF, the work profiles $W(t)$ collected from multiple SMD simulations have to be postprocessed with three distinct operations:

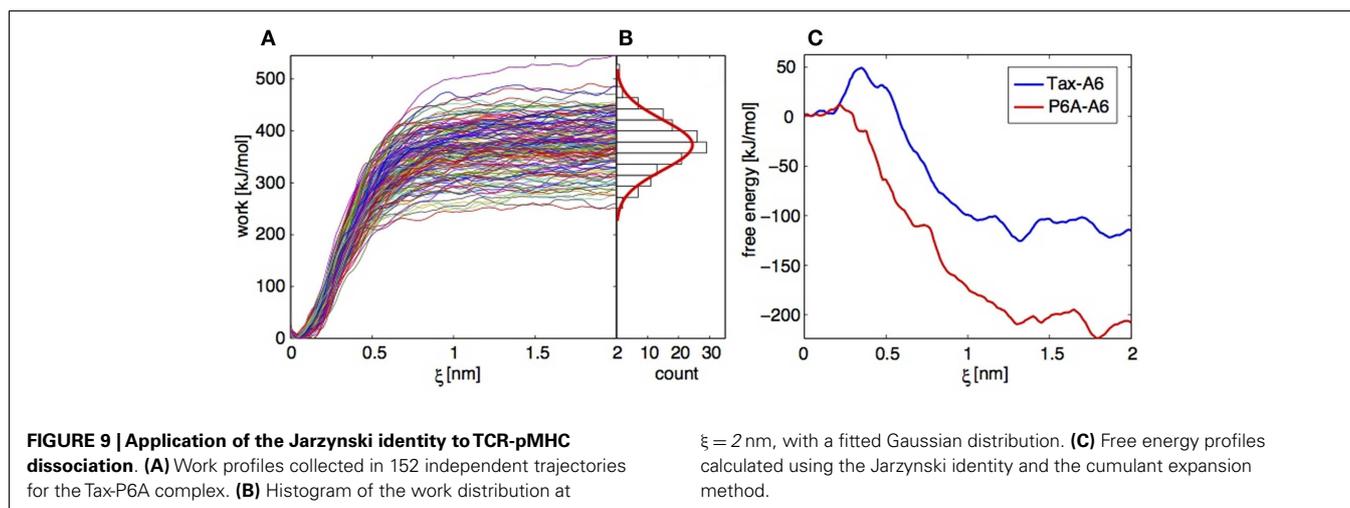
1. Reduce from the biased system to the physical system (unbias);
2. Average over all $\xi(r)$ visited during the evolution to recover $G(\xi)$;
3. Estimate the exponential average $\langle e^{-\beta W} \rangle_0$.

Operations 1 and 2 are performed within a modified weighted histogram scheme (73) adapted to the case of the individual potentials (72). Operation 3 can be more problematic. If we apply direct exponential averaging, the estimated PMF is very close to the lowest measured work profile at a given ξ . In the case where the true free energy value lies in the unsampled lower tail of the work distribution, direct averaging will result in a large overestimation. Indeed, for the Tax-A6 complex, the work profiles of **Figure 9A** would result in a dissociation free energy around 250 kJ/mol, compared to the experimental value of 32.2 kJ/mol (18).

Instead of direct averaging, if we assume that the work distribution is Gaussian, we can estimate $G(\xi)$ with a second-order cumulant expansion (71, 75),

$$G(\xi) = \bar{W}(\xi) - \frac{\beta}{2} \sigma^2(\xi).$$

Here, $\bar{W}(\xi)$ is the mean and σ the standard deviation of the work values at ξ obtained by applying operations 1 and 2 above to each trajectory independently. The resulting PMFs are shown on **Figure 9C**. The final dissociation free energies are –110 and



–210 kJ/mol for Tax-A6 and P6A-A6, respectively, which is a severe underestimation of the experimental values of 32.5 and 22.5 kJ/mol. We note that the calculated values should be corrected for translational and rotational entropic contributions to be compared to standard state free energy measurements, but this does not improve the results.

Our results show that direct averaging produces a strong overestimation of $G(\xi)$, which could be fixed only by sampling an extremely large number of trajectories to get enough low work values. Conversely, the cumulant expansion produces a strong underestimation of $G(\xi)$, which shows that the real work distribution has a shorter lower tail than the Gaussian distribution. Repeating the calculations with datasets of similar sizes and slightly different conditions showed the reproducibility of these findings.

Overall, this example illustrates the severe sampling difficulties that hamper the application of the JI to systems with sizeable dissociation. These difficulties have been evidenced by other researchers in systematic convergence studies of the JI method (76–79).

ESTIMATION OF RESIDUE CONTRIBUTION TO THE BINDING FREE ENERGY OF THE TCR/pMHC ASSOCIATION USING MM-PB(GB)SA

In the Molecular Mechanics – Poisson Boltzman Surface Area (MM-PBSA), or its variant the MM-GBSA, the binding free energy, ΔG_{bind} , is written as the sum of the gas phase contribution, $\Delta H_{\text{bind}}^{\text{gas}}$, the desolvation free energy of the system upon binding, ΔG_{desolv} , and an entropic contribution, $-T\Delta S$ (80):

$$\Delta G_{\text{bind}} = \Delta H_{\text{bind}}^{\text{gas}} + \Delta G_{\text{desolv}} - T\Delta S.$$

The term $\Delta H_{\text{bind}}^{\text{gas}}$ is constituted by the van der Waals (ΔH_{vdw}) and electrostatic (ΔH_{elec}) interaction energies between the two partners in the complex, and their conformational energy change upon binding, ΔH_{intra} :

$$\Delta H_{\text{bind}}^{\text{gas}} = \Delta H_{\text{elec}} + \Delta G_{\text{vdw}} + \Delta H_{\text{intra}}.$$

ΔG_{desolv} is the difference between the solvation free energy, ΔG_{solv} , of the complex and that of the isolated parts. ΔG_{solv}

is divided into the electrostatic, $\Delta G_{\text{elec, solv}}$, and the non-polar, $\Delta G_{\text{np, solv}}$, contributions,

$$\Delta G_{\text{solv}} = \Delta G_{\text{elec, solv}} + \Delta G_{\text{np, solv}}.$$

$\Delta G_{\text{elec, solv}}$ is calculated by solving the exact Poisson or the Poisson–Boltzmann (PB) equation (81, 82) in MM-PBSA, and the much faster but approximate Generalized Born (GB) model (83) in MM-GBSA (84). The term $\Delta G_{\text{np, solv}}$, which can be considered as the sum of a cavity term and a solute–solvent van der Waals term, is assumed to be proportional to the solvent accessible surface area (SASA) (85),

$$\Delta G_{\text{np, solv}} = \gamma \text{SASA} + b.$$

The entropy term can be decomposed into translational, S_{trans} , rotational, S_{rot} , and vibrational S_{vib} , contributions. These terms are calculated using standard equations of statistical mechanics (54, 86).

In the standard MM-PB(GB)SA protocol, all these energy terms are typically averaged over several hundreds of frames extracted from multi-nanosecond MD simulation trajectories, generally performed in explicit solvent. Those explicit water molecules are removed prior to energy calculations. In principle, three trajectories should be performed, one for the complex and one for each of the isolated partners, and the energy terms calculated using the corresponding simulation. However, a frequent, less computationally demanding, approximation consists in performing only one MD simulation for the complex (54). The terms relative to one isolated partner are then calculated after removing the atoms of the other partner in the frames extracted from the MD simulation of the complex. In this variant, the conformational energy change upon association is therefore neglected ($\Delta H_{\text{intra}} = 0$), and the influence of conformational changes on the other energy terms are not captured.

MM-PB(GB)SA has been used successfully to identify the hot-spots of protein–protein association and to determine the effect of mutations on association processes (84, 87–90). Two approaches can be considered. The computational alanine scanning (CAS)

approach (91) is directly comparable to its experimental counterpart. In this approach, ΔG_{bind} values are calculated for the wild-type system, as well as for several mutants in which one residue has been replaced by an alanine. The difference in ΔG_{bind} between the wild-type system and the mutants can be used to estimate the role played by each residue in the association process. Alternatively, it is possible to perform a binding free energy decomposition (BFED) for the wild-type system (84). In this approach, the contributions to ΔG_{bind} arising from groups of atoms, typically single residues, or even backbone or side-chain, are estimated from the wild-type system by performing a pairwise decomposition of the MM-PB(GB)SA terms (88, 90). BFED, which requires only one binding free energy calculation, is faster than CAS. In addition, BFED provides the possibility to study contributions from non-mutable groups of atoms, such as the backbone. However, contrarily to CAS, BFED results cannot be compared directly to an experimental alanine scanning.

To assess the ability of the MM-GBSA approach to identify quantitatively the hot-spots residues for the TCR/pMHC association, we performed a study of the 2C TCR/SIYR/H-2Kb system using both the CAS and BFED methods (90). This system was chosen because both the experimental 3D-structure and the results of an experimental alanine scanning were available at that time (92). A very good correlation was found between the residue contributions to ΔG_{bind} from both methods, with a correlation coefficient of 0.94, highlighting the interest in the faster BFED approach. A correlation coefficient of $R = 0.67$ was found between experimentally determined activity differences for alanine mutants and the calculated binding free energy changes upon mutation.

Our results also showed that BFED provided a more detailed and reliable description of the interactions between the TCR and pMHC molecules when including entropic terms. When the entropy was taken into account, the correlation coefficient was increased to 0.72. It was noticeable that the correlation obtained when neglecting the entropy term, which is very computationally expensive to calculate, was sufficient to quantify and rank the importance of the residues for TCR/pMHC association. Altogether, these pioneering results suggested that the BFED for the TCR-pMHC system provides a detailed and reliable enough description of the interactions between the molecules to be used as an *in silico* investigation tool in TCR protein-engineering.

COMPUTER-AIDED PROTEIN ENGINEERING

BACKGROUND

Patients with diverse types of cancer develop tumor-specific CD4⁺ and CD8⁺ T cell responses. Although these responses are typically unable to contain solid tumor growth or hematological malignancies, clinical studies have revealed the adoptive transfer of *ex vivo* expanded autologous tumor-specific T cells to be a promising immunotherapeutic approach to cancer treatment (13). A limitation, however, is that TCRs which bind tumor associated/self antigen are often of relatively low affinity. TCRs generally bind pMHC in the range of $K_D = 1\text{--}100\ \mu\text{M}$ (93). However, as a result of the thymic negative selection whereby T cells with high-affinity TCRs for “self” antigens are eliminated to prevent autoimmunity, TCRs specific for “self” tumor associated antigens tend to be weaker-binders compared to TCRs specific for “non-self” peptides (94).

Thus, the development of tumor-targeting TCRs to endow them with optimal binding properties, both in terms of fine-specificity against the targeted pMHC, and kinetic/affinity parameters that confer maximum cellular responsiveness, is a field of intense research toward cancer immunotherapy development (95).

The relative importance of the roles played by the TCR/pMHC binding affinity (K_D), and individual kinetic parameters (k_{on} and k_{off}), on T cell activation, has been intensively studied recently (11). The emerging consensus hypothesizes the existence of a TCR/pMHC “dwell-time” (96) enabling the sequential interaction of TCRs with a rare antigenic pMHC complex – a process known as “serial triggering” – and conferring an optimal T cell activation (97). It has also been demonstrated that both k_{on} and k_{off} define the “effective half-life” of a TCR/pMHC interaction (98). Thus, K_D , k_{on} , and k_{off} do all contribute to T cell activation and their optimizations should be addressed concomitantly by TCR engineering techniques for cancer immunotherapy.

As introduced earlier, TCRs contact pMHC antigens via the six CDRs (Figure 2), with CDR3 α and CDR3 β mainly bound over the peptide, and CDR1 and CDR2, α and β , making more contacts with MHC. It could thus be expected that mutations of CDR3s would be more likely to maintain peptide specificity than mutants of CDR1s and CDR2s. Indeed, several mutagenesis studies produced high-affinity TCRs bearing mutations on CDR3 α and CDR3 β that were found to be peptide-specific (99–102). However, other studies also discovered high-affinity mutants in CDR1 and CDR2 retaining peptide specificity despite the close proximity of the mutated region to MHC residues (99, 103). This indicates that all six CDRs can serve as a focus for mutagenesis to generate higher-affinity TCRs, while still potentially retaining substantial peptide specificity (2, 93).

Several efforts have been performed to optimize TCRs (93), which mainly consist in experimental yeast (100, 101, 104–106), phage (18, 102, 103, 107), and mammalian cell (108, 109) display techniques. These approaches were able to increase the affinity of the TCR by a factor of $100\text{--}10^6$, leading to K_D as low as 26 pM (102). Although very efficient to increase the K_D , these techniques lead to TCRs bearing multiple mutations, without providing a straightforward control of the effect of each one. Such TCRs are prone to alloreactivity due to peptide-independent binding of MHCs (95, 110).

Detailed control of the effect of each mutation at the atomic level can be provided by *in silico* rational protein-engineering techniques (111–115). Recently, Haidar et al. (116) engineered the human A6 TCR for enhanced affinity toward the Tax peptide/HLA-A2 MHC complex. Rapidly, the authors created a set of 219 fitted scoring functions, aiming to reproduce the binding affinity change upon 648 mutations of the ovomucoid turkey inhibitor molecule, using energy and statistically derived potential terms. Each function was then tested against the affinity changes of a first set of 11 A6 TCR mutants, and evaluated by correlation. The function reproducing the best the affinity changes on A6 TCR (named ZAFFI score) was retained to suggest new mutations. Due to the significant number of non-binding mutations generated using only the ZAFFI score, the authors further developed the ZAFFI filter function. The latter, trained by a Monte Carlo method on the 36 first A6 TCR point mutations, was employed to filter out mutations

with potentially bad electrostatic contacts. In total 59 mutants were tested. Twelve were found to be better binders than the wild-type TCR, as measured experimentally by Surface Plasmon resonance (SPR). It must be noted that some non-binders were generated on purpose to help training the ZAFFI score and filter. All mutations found in this study to increase the binding were hydrophobic substitutions that enhanced the interface complementarity. No mutation introducing new significant electrostatic contacts, and thus potentially increasing the selectivity of the TCR/pMHC binding, was found positive. Despite the use of fitted scoring function and filter that hampers its straightforward translation to non-TCR systems, this interesting study illustrated the feasibility of a rational *in silico* approach to design TCR with higher affinities. It opened the road to new approaches, with physically sound and non-fitted universal free energy estimates, straightforward transferability, and high success rates.

METHODOLOGY AND RESULTS ON BC1 TCR

Encouraged by the results of our BFED method in reproducing the outcome of an alanine scanning experiment on the TCR/SIYR/H-2Kb system (90), we decided to develop a new structure-based approach, based on MM-GBSA free energy calculations, to rationally design new TCR sequences. Our approach can be divided into several steps. First, the importance of each wild-type TCR residue for the TCR/pMHC association is estimated using a MM-GBSA BFED. Then, based on the TCR-pMHC structure and the residue contributions to the binding, mutations are designed for the residues showing the most promising opportunities of enhancement for the interaction with the pMHC. These putative sequence modifications are finally selected for experimental testing based on the estimated binding free energy gain, $\Delta\Delta G_{\text{bind}}$. The latter was obtained by calculating the contribution of each residue to the binding free energy change upon a given mutation, $\Delta\Delta G_{\text{bind}}^{\text{res}}$, as the difference between the residue contribution for the mutated complex, $\Delta G_{\text{bind}}^{\text{res,mut}}$, and that for the wild-type complex, $\Delta G_{\text{bind}}^{\text{res}}$:

$$\Delta\Delta G_{\text{bind}}^{\text{res}} = \Delta G_{\text{bind}}^{\text{res,mut}} - \Delta G_{\text{bind}}^{\text{res}}$$

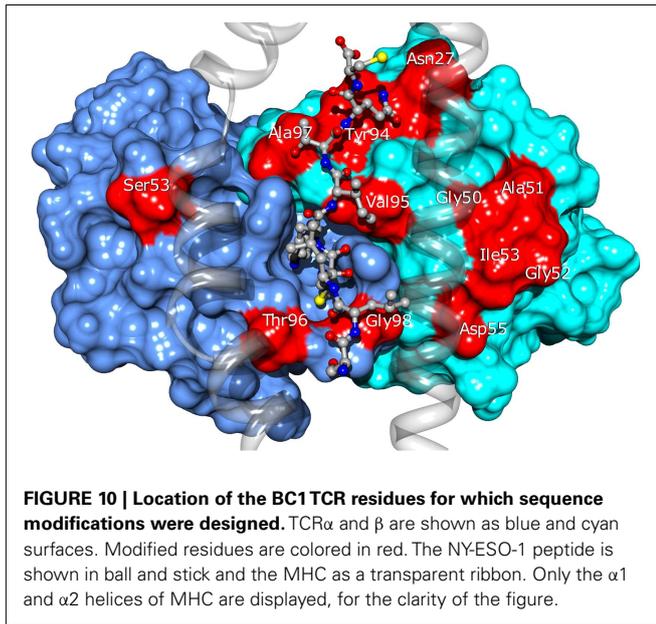
$\Delta G_{\text{bind}}^{\text{res}}$ and $\Delta G_{\text{bind}}^{\text{res,mut}}$ values were calculated from MD simulations of the wild-type and the corresponding mutated TCR-pMHC, respectively. The binding free energy difference upon a given mutation, $\Delta\Delta G_{\text{bind}}$, was finally obtained by summing the $\Delta\Delta G_{\text{bind}}^{\text{res}}$ values over the mutated residue and all the residues in contact with it. This local summation was preferred to a summation over all residues of the TCR-pMHC complex since it suppresses the errors arising from residues far from the site of the sequence modification, and making no contribution to the mutation effect.

We have applied our computer-aided protein-engineering approach to the BC1 TCR (26, 117). This TCR was discovered from a long-surviving melanoma patient (LAU #155) with a naturally occurring CD8⁺ T cell response against the immunodominant cancer-testis epitope NY-ESO-1_{157–165} (SLLMWITQC), presented by the commonly expressed MHC class I allele HLA-A*0201. The expression profile of NY-ESO-1 makes it an attractive target for cancer immunotherapy. It is indeed expressed by several

solid tumors, including melanoma, as well as hematological malignancies (myelomas, lymphomas, and leukemias), while in normal tissue its expression is limited to the testis cells (118–122). Interestingly, recent studies have shown that an immune response against NY-ESO-1 can convey an important clinical benefit for the patient. Seventy-seven percent of patients treated with the CTLA-4 blocking antibody ipilimumab showed favorable clinical outcome if they had a detectable CD8⁺ T cell responses against NY-ESO-1, compared to only 14% otherwise (123). Our *in silico* approach was facilitated by the existence of the crystal structure of the 1G4 TCR bound to NY-ESO-1/HLA-A*201 (124), PDB ID 2BNR, available in the protein databank (125). The sequence of the latter differs only from that of the BC1 TCR by four residues: Thr95 α , Ser96 α , Asn97 β , and Thr98 β of the 1G4 TCR, are replaced by Gln95 α , Thr96 α , Ala97 β , and Ala98 β in BC1 TCR, respectively. These sequence modifications were introduced in the 1G4 X-ray structure before applying our approach. Noticeably, three crystal structures of free and bound 1G4 TCR were solved after we performed our simulations (126). These affinity-enhanced TCRs contain mutations in the CDR3 loops or in both the CDR2 and CDR3 loops, which were obtained by *in vitro* directed evolution (102). They revealed that the binding mode for the high-affinity TCRs was identical to that of the wild-type TCR, with only limited changes in the mutated CDRs. A previous assessment by Zhao et al., of six phage library-derived 1G4 TCR variants demonstrated three categories of TCR specificity related to affinity; (i) super-high-affinity TCR (26 pM) which completely lacked specificity, (ii), mid-range affinity TCR (5 and 85 nM) that were specific only in the absence of CD8 co-engagement, and, (iii), intermediate-range affinity TCR (0.4 and 4 μ M) that maintained specificity (107). By taking a rational computer-aided approach to TCR engineering, we were able to design a new original set of TCR variants. We enhanced the TCR/pMHC binding interactions in a “step-by-step” manner, targeting change to specific kinetic parameters, and limiting overall gain in affinity as well as potential for cross-reactivity.

Twenty four of the most promising mutations identified using this approach, spanning 11 different residues of the CDRs (see **Figure 10**), were introduced into the BC1 TCR sequence. The engineered proteins were produced using a mammalian cell expression system, purified and tested by titration ELISA. We found a qualitative agreement between the calculated $\Delta\Delta G_{\text{bind}}$ values and the experimental results. Thirteen (54%) of the mutations proposed by our approach showed improved affinity for the pMHC, compared to the wild-type TCR (Zoete et al., in preparation). We obtained an excellent correlation of $R = 0.85$ between the calculated $\Delta\Delta G_{\text{bind}}$ and the measured co-logarithm of the optical density measured by ELISA titration. Only three outliers were found: V β A51V, Y94N, and A51D. Their presence might be explained by conformational rearrangements upon mutation, minimally accessible during the MD simulations. This correlation illustrates the efficiency with which the $\Delta\Delta G_{\text{bind}}$ calculated with our method allows for the rational selection of TCR sequence modifications potentially increasing its affinity for pMHC.

For quantitative TCR/pMHC binding measurements some of the TCRs were refolded from bacterial inclusion bodies and analyzed by SPR. We began by producing TCR



having one amino acid change. The best single replacement TCR had 52-fold improvement in K_D (Zoete et al., in preparation). Then, we gradually increased affinity by step-wise combination of rationally selected replacements in both the β -chain and the α -chain. As predicted by modeling and assessed by SPR binding assays, we found the following progression in K_D (**Table 1**): WT (21.4 μ M) > β -G50A (4.6 μ M) > β -G50A + A51E (1.9 μ M) > β -G50A + A51E + A97L (0.91 μ M) > β -G50A + β A51E- β A97L + α -S53W (0.14 μ M) (Zoete et al., in preparation) (7). Our method thus allows for the design of TCRs with fine tuned K_D values, potentially leading to T cells with optimal activity (7). This is in contrast with experimental display approaches which tend to select the tightest-binding TCRs that can potentially surpass an “affinity threshold of specificity,” as has been reported for the 1G4 TCR (107). We found a very satisfying correlation between the calculated energies and the pK_D and k_{off} values measured by SPR (Zoete et al., in preparation).

Importantly, we found that our approach was able to predict successfully mutations toward both non-polar and polar residues, contrary to previous studies where only designed mutations toward non-polar residues were successful in increasing the experimental affinity (116). To take place, polar interactions, such as hydrogen bond or ionic interaction, require an appropriate match of chemical functionalities and precise geometrical constraints between interacting partners. Therefore, they provide an essential contribution to the directionality and the specificity of molecular recognition (127). This point is critical to the development of TCRs for immunotherapies. It is indeed essential that TCRs maintain their specificity, and do not acquire novel antigen specificities that might cause off-target toxicity upon the adoptive transfer of genetically engineered T cells to patients. Two rationally designed sequence modifications provide interesting examples regarding the detailed control provided by our approach on the binding process. Mutation β -A51E introduces an ionic interaction between the new glutamate side-chain and MHC Arg75

Table 1 | Binding affinity of rationally designed BC1 TCR variants [from Ref. (7)] measured by SPR.

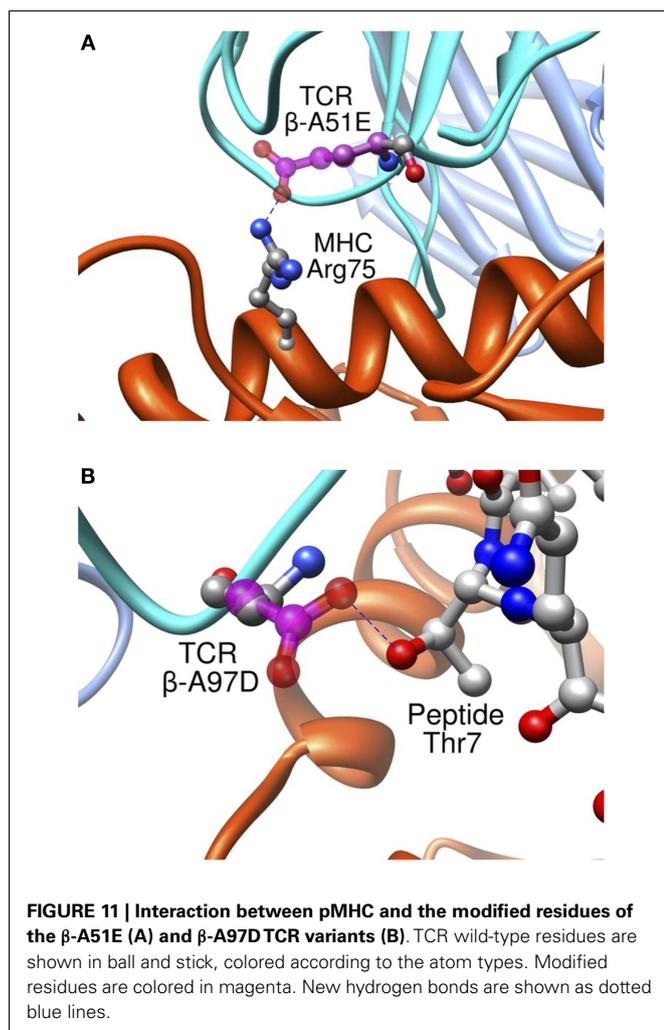
TCR mutations	K_D (μ M)
β -V49I	Near non-binding
Wild-type	21.4
β -G50A	4.62
β -A97L	2.69
β -G50A + A51E	1.91
β -G50A + A51E + A97L	0.91
α -S53W + β -G50A + A51E	0.4
α -S53W + β -G50A + A51E + A97L	0.14
β -G50A + A51I + G52Q + I53T	0.015

(**Figure 11A**), translating into a calculated favorable $\Delta\Delta G_{bind}$. This replacement produced a fourfold experimental improvement in K_D as measured by SPR (8). This mutation thus increases the affinity for the pMHC through better interactions with the MHC molecule, and therefore cannot be expected to increase the selectivity for the peptide antigen. Interestingly, we also designed the β -A97D variant, which introduces a new hydrogen bond between TCR and the peptide Thr7 side-chain (**Figure 11B**). The overall effect of this modification was somewhat unfavorable to the binding, as shown by the experimental titration ELISA and calculated $\Delta\Delta G_{bind}$ values (Zoete et al., in preparation). However, despite the decreased activity, this mutation is interesting for the putative gain in selectivity for the NY-ESO-1 system thanks to the new polar interaction taking place between the TCR and the peptide.

An experimental alanine scanning of the peptide supported this idea. The affinity of several TCR mutants (β -G50A/A51E, A97L, A97D, I53W, V95L, and Y94N) for the different mutated pMHC were measured by titration ELISA and compared to the results for the wild-type TCR. All mutated TCR showed affinities similar to the wild-type TCR for all the peptide mutants. The only exception was the β -A97D TCR mutant, which did not bind to the peptide T7A mutant contrary to the wild-type TCR. This experimental result is in good agreement with the *in silico* data showing that all those TCR mutations modified the interactions with the MHC rather than with the peptide with the exception of β -A97D which exchanges a hydrogen bond with the peptide Thr7 side-chain. The β -A97L TCR mutant, which makes non-polar contact with the side-chain of peptide Thr7 but no hydrogen bond, is less affected by the Thr7 mutation to Ala than the β -A97D TCR variant. Thus, although the β -A97D modification is somewhat unfavorable to the binding, it could be useful to introduce it in engineered TCRs targeting NY-ESO-1, in addition to other variations increasing the affinity, in order to obtain more selective TCRs.

The selectivity of TCRs designed by our approach was tested experimentally. We observed no interaction of our TCRs with any non-cognate pMHC complexes. In addition, tetramer binding studies with an extensive panel of non-cognate pMHC revealed that the TCR variants expressed at the cell-surface were also HLA-A*0201/NY-ESO-1_{157–165}-specific (7).

The binding free energy calculations used in our approach are physics-based and not reliant on *ad hoc* model fitting. We



can thus expect that our design strategy is highly transferable to any protein/protein interaction of known structure and biological interest.

INTERMEDIATE TCR/pMHC BINDING PARAMETERS CONFER MAXIMUM BIOLOGICAL RESPONSES

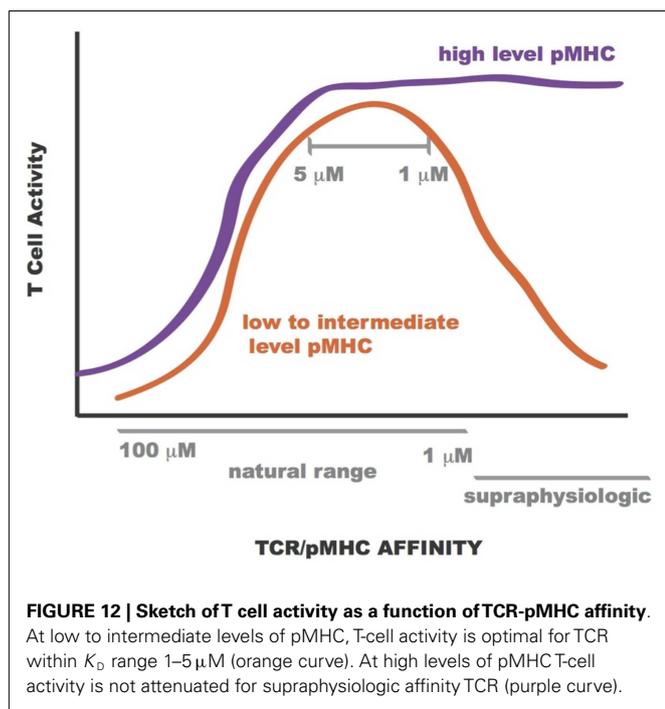
In order to efficiently screen our extensive panel of modeled BC1 TCR variants we began by establishing a mammalian cell expression system for TCR production. HEK-293 cells were PEI co-transfected with pHYK8 plasmids encoding the α - and β -chains, each under control of the CMV promoter. Following the strategy of Chang et al. an acidic-basic zipper was incorporated to facilitate heterodimeric chain pairing (128). Following 5–7 days culture in serum-free medium, the TCR variants were HIS-tag purified from the supernatants (yields were up to ~ 3 mg/L) and compared to the wild-type TCR by titration ELISA for binding plate-captured pMHC. Single and multiple amino acid replacements were assessed in the α - and β -chains. Over 60% showed enhanced pMHC binding as compared to the wild-type TCR. Further, binding of the TCR variants against an Ala replacement scan of the NY-ESO-1_{157–165} peptide (SLLMWITQC) revealed near identical patterns of recognition, suggesting a conserved

mechanism of binding. For all TCR variants, M160, W161, I162, and Q164 were critical contact residues as their replacement with Ala abolished TCR/pMHC binding (7, 8).

The TCR chains were subcloned into pGMT7 for their expression as inclusion bodies in BL21(DE3)pLys bacterial cells and subsequent TCR refolding by dialysis (129). The binding affinity and kinetics for a panel of TCRs having various combinations of amino acid replacements in CDR2- β , CDR3- β , and CDR2- α , were measured by SPR using the BIAcore 3000. Most natural TCRs bind pMHC with weak affinity, in the range of 1–100 μ M, as a result of slow association (10^3 – 10^4 $M^{-1}s^{-1}$), and fast rates of dissociation (typically a half-life of seconds at 37°C) (1, 10, 130), reflecting the need for T lymphocytes to detect a virtually limitless repertoire of foreign epitopes while avoiding autoreactivity, and the fact that they do not undergo somatic hypermutation as do antibodies. The eight TCRs chosen amongst our panel incrementally increased in affinity from 21.4 μ M for the wild-type one, to extreme physiologic affinity at 15 nM (summarized in **Table 1**).

Lentiviral constructs were built to assess the activity of the rationally designed TCRs in transduced primary human CD8⁺ T cells and identify those able to confer maximum activity levels. A range of functions were assessed for the transduced CD8⁺ T cells including Ca²⁺ flux, intracellular signaling, cell-surface TCR clustering, target-cell killing, and cytokine/chemokine release. All activity levels for affinity-enhanced TCRs uniformly increased from that of the wild-type TCR, reaching a peak for TCRs within the upper range of natural affinity, 1–5 μ M. Beyond this affinity, in the supra-physiological range, however, activity levels began to drop, both in the presence and absence of CD8 co-receptor engagement, usually reaching a minimum for the extreme supra-physiological affinity TCR (1, 7, 8, 10, 130). However, under experimental conditions in which the T cells were exposed to target-cells pulsed with high concentration of NY-ESO-1_{157–165} peptide this attenuation in activity for the supra-physiological affinity TCR was no longer observed. This phenomenon is illustrated in **Figure 12**. Importantly, no non-specific reactivity was observed for any of the rationally designed TCRs (7, 8).

Over the years two main models of T cell activation have emerged [reviewed in Stone et al. (11)]. For the “affinity model” (101, 131) the total number of TCRs bound to pMHCs at equilibrium is thought to regulate T cell activity levels. The “half-life model” proposes that the TCR must stay bound with sufficiently strength/duration for productive signaling and also enable the serial engagement of the “rare” antigenic pMHC complex by adjacent TCR for the amplification of signal (i.e., the TCRs must have an optimal dwell-time) (96, 97, 132–134). Although the half-life of a TCR/pMHC interaction has traditionally been calculated from k_{off} values ($t_{1/2} = \ln 2/k_{off}$), recent work by Aleksic et al. has elegantly demonstrated that for TCRs having faster association rates there can be rapid TCR/pMHC re-engagement rather than a lateral diffusion of TCRs in the cell membrane to prolong the effective half-life of the TCR/pMHC interaction. Thus, both k_{off} and k_{on} can contribute to TCR/pMHC dwell or confinement time (135). Overall our findings correspond to the “half-life” model of T cell activation. Presumably the supra-physiological affinity TCRs are limiting the serial engagement but at high peptide concentrations this is not an issue as the TCR can find more pMHC molecules

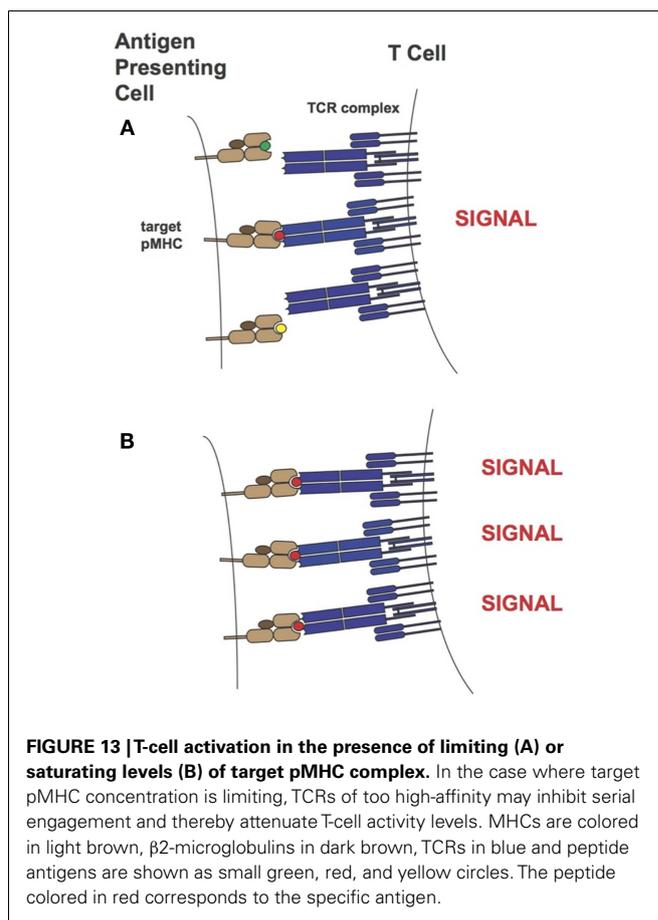


on the target-cell-surface. This is demonstrated in **Figure 13**. Our work also demonstrated the value of a modeling approach because TCRs with particular binding parameters, i.e., falling within an optimal affinity range, either through faster on-rates or slower off-rates, can be developed to enable maximum T cell activity levels.

CONCLUSION

We found that the synergy between *in silico* design, and *in vitro* testing using both soluble molecules and transfected cells is key to the design of improved TCRs to be used in adoptive cell transfer therapies. Using an innovative structural approach based on recent *in silico* techniques, we have developed a method to rationally optimize the surface of the TCR to increase its affinity for pMHC. As opposed to library screening, we propose a step-wise, incremental optimization of the TCR. This allows the preservation of the favorable initial properties of the TCR through the various cycles. Our optimization method can selectively target contacts involving the MHC or the peptide, thus maintaining a good balance between overall affinity and specificity. In addition, modifying only a few amino acids reduces the risk of obtaining undesired cross-recognition or raising anti-TCR antibodies *in vivo*. This rational optimization approach is therefore very promising to the design of TCRs for adoptive T cell immunotherapy clinical trials.

The application of our TCR engineering approach to the tumor-targeting BC1 TCR targeted to the A2/NY-ESO-1_{157–165} antigen identified several original single mutations of the CDR loops conferring increased experimental affinity for pMHC compared to the wild-type TCR. T cells expressing some of the affinity-enhanced TCR showed better overall functionality, including improved killing of both peptide-loaded T2 cells and melanoma



tumor cell lines, higher proliferative capacity, and increased levels of cytokine/chemokine secretion, as compared to wild-type TCR expressing T cells. For all functional assays, we observed a gain in CD8⁺ T cell activity level with increase in affinity, with a peak at an affinity of \sim 1–5 μM . Beyond this affinity we observed a progressive decrease in activity levels. We are currently testing the relative activity of the different TCRs *in vivo* in a mouse model. The successful candidates are planned to enter a phase I clinical trial program for stage IV melanoma. As the methods presented here are general and transferable to any TCR-pMHC complex, other cancer types will follow shortly.

In parallel to direct applications of the existing approach, we will take advantage of the methodological work presented above to improve our *in silico* TCR optimization method. The development of homology-based approaches to model the 3D-structure of TCR-pMHC complexes potentially extends the use of our method to TCR repertoires for which no X-ray structure is available. The detailed picture of the TCR/pMHC interaction emerging from the MD simulations showed the presence of single water molecules trapped at the interface (20). Including these interfacial waters in the BFED scheme would improve accuracy in some cases. Close residues at protein-protein interfaces can display collaborative effects that result in the non-additivity of their contributions to the binding free energy, which will have to be taken into account in the next generation of TCR optimization

methods. Noticeably, because the binding free energy decomposition method used is physics-based, without any *ad hoc* parameters, our *in silico* techniques are straightforwardly transferable to other types of macromolecular protein complexes.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Center of Competence in Research and the Swiss National Science Foundation

REFERENCES

- Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* (2006) **24**:419–66. doi:10.1146/annurev.immunol.23.021704.115658
- Burrows SR, Chen Z, Archbold JK, Tyman FE, Beddoe T, Kjer-Nielsen L, et al. Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. *Proc Natl Acad Sci U S A* (2010) **107**:10608–13. doi:10.1073/pnas.1004926107
- Garcia KC, Adams EJ. How the T cell receptor sees antigen—a structural view. *Cell* (2005) **122**:333–6. doi:10.1016/j.cell.2005.07.015
- Godfrey DI, Rossjohn J, McCluskey J. The fidelity, occasional promiscuity, and versatility of T cell receptor recognition. *Immunity* (2008) **28**:304–14. doi:10.1016/j.immuni.2008.02.004
- Davis MM, Boniface JJ, Reich Z, Lyons D, Hampl J, Arden B, et al. Ligand recognition by alpha beta T cell receptors. *Annu Rev Immunol* (1998) **16**:523–44. doi:10.1146/annurev.immunol.16.1.523
- Dietrich P-Y, Le Gal F-A, Dutoit V, Pittet MJ, Trautman L, Zipperli A, et al. Prevalent role of TCR alpha-chain in the selection of the preimmune repertoire specific for a human tumor-associated self-antigen. *J Immunol* (2003) **170**:5103–9.
- Irving M, Zoete V, Hebeisen M, Schmid D, Baumgartner P, Guillaume P, et al. Interplay between T cell receptor binding kinetics and the level of cognate peptide presented by major histocompatibility complexes governs CD8⁺ T cell responsiveness. *J Biol Chem* (2012) **287**:23068–78. doi:10.1074/jbc.M112.357673
- Schmid DA, Irving MB, Posevitz V, Hebeisen M, Posevitz-Fejfar A, Sarria J-CF, et al. Evidence for a TCR affinity threshold delimiting maximal CD8 T cell function. *J Immunol* (2010) **184**:4936–46. doi:10.4049/jimmunol.1000173
- Valmori D, Dutoit V, Schnuriger V, Quiquerez A-L, Pittet MJ, Guillaume P, et al. Vaccination with a Melan-A peptide selects an oligoclonal T cell population with increased functional avidity and tumor reactivity. *J Immunol* (2002) **168**:4231–40.
- Krogsgaard M, Davis MM. How T cells “see” antigen. *Nat Immunol* (2005) **6**:239–45. doi:10.1038/ni1173
- Stone JD, Chervin AS, Kranz DM. T-cell receptor binding affinities and kinetics: impact on T-cell activity and specificity. *Immunology* (2009) **126**:165–76. doi:10.1111/j.1365-2567.2008.03015.x
- Restifo NP, Dudley ME, Rosenberg SA. Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat Rev Immunol* (2012) **12**:269–81. doi:10.1038/nri3191
- Rosenberg SA, Restifo NP, Yang JC, Morgan RA, Dudley ME. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. *Nat Rev Cancer* (2008) **8**:299–308. doi:10.1038/nrc2355
- Khan JM, Cheruku HR, Tong JC, Ranganathan S. MPID-T2: a database for sequence-structure-function analyses of pMHC and TR/pMHC structures. *Bioinformatics* (2011) **27**:1192–3. doi:10.1093/bioinformatics/btr104
- Gras S, Burrows SR, Turner SJ, Sewell AK, McCluskey J, Rossjohn J. A structural voyage toward an understanding of the MHC-I-restricted immune response: lessons learned and much to be learned. *Immunol Rev* (2012) **250**:61–81. doi:10.1111/j.1600-065X.2012.01159.x
- Wang J-H, Reinherz EL. The structural basis of $\alpha\beta$ T-lineage immune recognition: TCR docking topologies, mechanotransduction, and co-receptor function. *Immunol Rev* (2012) **250**:102–19. doi:10.1111/j.1600-065X.2012.01161.x
- Armstrong KM, Insaiddo FK, Baker BM. Thermodynamics of T-cell receptor-peptide/MHC interactions: progress and opportunities. *J Mol Recognit* (2008) **21**:275–87. doi:10.1002/jmr.896
- Ding YH, Baker BM, Garboczi DN, Biddison WE. Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. *Immunity* (1999) **11**:45. doi:10.1016/S1074-7613(00)80080-1
- Davis-Harrison RL, Armstrong KM, Baker BM. Two different T cell receptors use different thermodynamic strategies to recognize the same peptide/MHC ligand. *J Mol Biol* (2005) **346**:533–50. doi:10.1016/j.jmb.2004.11.063
- Cuendet MA, Zoete V, Michielin O. How T cell receptors interact with peptide-MHCs: a multiple steered molecular dynamics study. *Proteins* (2011) **79**:3007–24. doi:10.1002/prot.23104
- Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Mod* (2001) **7**:306.
- Daura X, Mark AE, van Gunsteren WF. Parametrization of aliphatic CHn united atoms of GROMOS96 force field. *J Comput Chem* (1998) **19**:535–47. doi:10.1002/(SICI)1096-987X(19980415)19:5<535::AID-JCC6>3.0.CO;2-N
- Wu LC, Tuot DS, Lyons DS, Garcia KC, Davis MM. Two-step binding mechanism for T-cell receptor recognition of peptide MHC. *Nature* (2002) **418**:552–6. doi:10.1038/nature00920
- Gagnon SJ, Borbulevych OY, Davis-Harrison RL, Baxter TK, Clemens JR, Armstrong KM, et al. Unraveling a hotspot for TCR recognition on HLA-A2: evidence against the existence of peptide-independent TCR binding determinants. *J Mol Biol* (2005) **353**:556–73. doi:10.1016/j.jmb.2005.08.024
- Boudinot P, Marriotti-Ferrandiz ME, Pasinquier LD, Benmansour A, Cazeneuve P-A, Six A. New perspectives for large-scale repertoire analysis of immune receptors. *Mol Immunol* (2008) **45**:2437–45. doi:10.1016/j.molimm.2007.12.018
- Derré L, Bruyninx M, Baumgaertner P, Ferber M, Schmid D, Leimgruber A, et al. Distinct sets of alphabeta TCRs confer similar recognition of tumor antigen NY-ESO-1_{157–165} by interacting with its central Met/Trp residues. *Proc Natl Acad Sci U S A* (2008) **105**:15010–5. doi:10.1073/pnas.0807954105
- Romero P, Dunbar PR, Valmori D, Pittet M, Ogg GS, Rimoldi D, et al. *Ex vivo* staining of metastatic lymph nodes by class I major histocompatibility complex tetramers reveals high numbers of antigen-experienced tumor-specific cytolytic T lymphocytes. *J Exp Med* (1998) **188**:1641–50. doi:10.1084/jem.188.9.1641
- Leimgruber A, Ferber M, Irving M, Hussain-Kahn H, Wiekowski S, Derré L, et al. TCRRep 3D: an automated *in silico* approach to study the structural properties of TCR repertoires. *PLoS ONE* (2011) **6**:e26301. doi:10.1371/journal.pone.0026301
- Michielin O, Luescher I, Karplus M. Modeling of the TCR-MHC-peptide complex. *J Mol Biol* (2000) **300**:1205–35. doi:10.1006/jmbi.2000.3788
- Zoete V, Irving MB, Michielin O. MM-GBSA binding free energy decomposition and T cell receptor engineering. *J Mol Recognit* (2010) **23**:142–52. doi:10.1002/jmr.1005
- Ehrenmann F, Lefranc M-P. IMGT/3D structure-DB: querying the IMGT database for 3D structures in immunology and immunoinformatics (IG or antibodies, TR, MH, RPI, and FPIA). *Cold Spring Harb Protoc* (2011) **2011**:750–61. doi:10.1101/pdb.prot5637
- Ehrenmann F, Kaas Q, Lefranc M-P. IMGT/3D structure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhSF. *Nucleic Acids Res* (2010) **38**:D301–7. doi:10.1093/nar/gkp946
- 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**:958–61. doi:10.1126/science.286.5441.958

34. Robins HS, Campregher PV, Srivastava SK, Wachter A, Turtle CJ, Khsai O, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* (2009) **114**:4099–107. doi:10.1182/blood-2009-04-217604
35. Kessler B, Michielin O, Blanchard CL, Apostolou I, Delarbre C, Gachelin G, et al. T cell recognition of hapten. Anatomy of T cell receptor binding of a H-2kd-associated photoreactive peptide derivative. *J Biol Chem* (1999) **274**:3622–31. doi:10.1074/jbc.274.6.3622
36. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, et al. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* (2006) **5**:Unit5.6.
37. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* (2000) **29**:291–325. doi:10.1146/annurev.biophys.29.1.291
38. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* (1993) **234**:779–815. doi:10.1006/jmbi.1993.1626
39. Bentley GA, Boulout G, Karjalainen K, Mariuzza RA. Crystal structure of the beta chain of a T cell antigen receptor. *Science* (1995) **267**:1984–7. doi:10.1126/science.7701320
40. Fields BA, Ober B, Malchioldi EL, Lebedeva MI, Braden BC, Ysern X, et al. Crystal structure of the V alpha domain of a T cell antigen receptor. *Science* (1995) **270**:1821–4. doi:10.1126/science.270.5243.1821
41. Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* (1996) **384**:134–41. doi:10.1038/384134a0
42. Garcia KC, Degano M, Pease LR, Huang M, Peterson PA, Teyton L, et al. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* (1998) **279**:1166–72. doi:10.1126/science.279.5354.1166
43. Fremont DH, Stura EA, Matsumura M, Peterson PA, Wilson IA. Crystal structure of an H-2Kb-ovalbumin peptide complex reveals the interplay of primary and secondary anchor positions in the major histocompatibility complex binding groove. *Proc Natl Acad Sci U S A* (1995) **92**:2479–83. doi:10.1073/pnas.92.7.2479
44. Fiser A, Do RK, Sali A. Modeling of loops in protein structures. *Protein Sci* (2000) **9**:1753–73. doi:10.1110/ps.9.9.1753
45. Fagerberg T, Cerottini J-C, Michielin O. Structural prediction of peptides bound to MHC class I. *J Mol Biol* (2006) **356**:521–46. doi:10.1016/j.jmb.2005.11.059
46. Al-Lazikani B, Lesk AM, Chothia C. Canonical structures for the hypervariable regions of T cell alphabeta receptors. *J Mol Biol* (2000) **295**:979–95. doi:10.1006/jmbi.1999.3358
47. Derré L, Ferber M, Touvrey C, Devevre E, Zoete V, Leimgruber A, et al. A novel population of human melanoma-specific CD8 T cells recognizes Melan-AMART-1 immunodominant nonapeptide but not the corresponding decapeptide. *J Immunol* (2007) **179**:7635–45.
48. Khan JM, Ranganathan S. Understanding TR binding to pMHC complexes: how does a TR scan many pMHC complexes yet preferentially bind to one. *PLoS ONE* (2011) **6**:e17194. doi:10.1371/journal.pone.0017194
49. Roomp K, Domingues FS. Predicting interactions between T cell receptors and MHC-peptide complexes. *Mol Immunol* (2011) **48**:553–62. doi:10.1016/j.molimm.2010.10.014
50. Varani L, Bankovich AJ, Liu CW, Colf LA, Jones LL, Kranz DM, et al. Solution mapping of T cell receptor docking footprints on peptide-MHC. *Proc Natl Acad Sci U S A* (2007) **104**:13080–5. doi:10.1073/pnas.0703702104
51. Broocks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. *J Comput Chem* (2009) **30**:1545–614. doi:10.1002/jcc.21287
52. Mackerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* (1998) **102**:3586–616. doi:10.1021/jp973084f
53. Haberthür U, Cafilisch A. FACTS: fast analytical continuum treatment of solvation. *J Comput Chem* (2008) **29**:701–15. doi:10.1002/jcc.20832
54. Zoete V, Meuwly M, Karplus M. Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. *Proteins* (2005) **61**:79–93. doi:10.1002/prot.20528
55. Ferber M, Zoete V, Michielin O. T-cell receptors binding orientation over peptide/MHC class I is driven by long-range interactions. *PLoS ONE* (2012) **7**:e51943. doi:10.1371/journal.pone.0051943
56. Baker BM, Gagnon SJ, Biddison WE, Wiley DC. Conversion of a T cell antagonist into an agonist by repairing a defect in the TCR/peptide/MHC interface: implications for TCR signaling. *Immunity* (2000) **13**:475–84. doi:10.1016/S1074-7613(00)00047-9
57. Michielin O, Karplus M. Binding free energy differences in a TCR-peptide-MHC complex induced by a peptide mutation: a simulation analysis. *J Mol Biol* (2002) **324**:547–69. doi:10.1016/S0022-2836(02)00880-X
58. Christ CD, Mark AE, van Gunsteren WF. Basic ingredients of free energy calculations: a review. *J Comput Chem* (2010) **31**:1569–82. doi:10.1002/jcc.21450
59. Cuendet MA, Zoete V, Michielin O. Molecular Dynamics-based Free Energy Simulations. In: Schwede T, Peitsch M editors. *Computational Structural Biology*. Singapore: World Scientific Publishing (2008). p. 513–48.
60. Gao J, Kuczera K, Tidor B, Karplus M. Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science* (1989) **244**:1069–72. doi:10.1126/science.2727695
61. Kirkwood JG. Statistical mechanics of fluid mixture. *J Chem Phys* (1935) **3**:300–13. doi:10.1063/1.1749657
62. Boresch S, Karplus M. The role of bonded terms in free energy simulations: I. Theoretical analysis. *J Phys Chem A* (2013) **103**:103–18. doi:10.1021/jp981628n
63. Beutler TC, Mark AE, van Schaik RC, Gerber PB, van Gunsteren WF. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem Phys Lett* (1994) **222**:529–39. doi:10.1016/0009-2614(94)00397-1
64. Madden DR, Garboczi DN, Wiley DC. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* (1993) **75**:693–708. doi:10.1016/0092-8674(93)90490-H
65. Brooks CL, Brünger A, Karplus M. Active site dynamics in protein molecules: a stochastic boundary molecular-dynamics approach. *Biopolymers* (1985) **24**:843–65. doi:10.1002/bip.360240509
66. Jarzynski C. A non-equilibrium equality for free energy differences. *Phys Rev Lett* (1997) **78**:2690–3. doi:10.1103/PhysRevLett.78.2690
67. Jarzynski C. Equilibrium free energy differences from non-equilibrium measurements: a master equation approach. *Phys Rev E* (1997) **56**:5018–35. doi:10.1103/PhysRevE.56.5018
68. Cuendet MA. Statistical mechanical derivation of Jarzynski's identity for thermostated non-Hamiltonian dynamics. *Phys Rev Lett* (2006) **96**:120602. doi:10.1103/PhysRevLett.96.120602
69. Cuendet MA. The Jarzynski identity derived from general Hamiltonian or non-Hamiltonian dynamics reproducing NVT or NPT ensembles. *J Chem Phys* (2006) **125**:144109–144109–12. doi:10.1063/1.2338535
70. Isralewitz B, Davis-Harrison RL, Baudry J, Armstrong KM, Gullingsrud J, Baker BM, et al. Steered molecular dynamics investigations of protein function. *J Mol Graphics Modell* (2001) **19**:13. doi:10.1016/S1093-3263(00)00133-9
71. Park S, Schulten K. Calculating potentials of mean force from steered molecular dynamics information. *J Chem Phys* (2004) **120**:5946–61. doi:10.1063/1.1651473
72. Cuendet MA, Michielin O. Protein-protein interaction investigated by steered molecular dynamics: the TCR-pMHC complex. *Biophys J* (2008) **95**:3575–90. doi:10.1529/biophysj.108.131383
73. Hummer G, Szabo A. Free energy reconstruction from non-equilibrium single-molecule pulling experiments. *Proc Natl Acad Sci U S A* (2001) **98**:3658–61. doi:10.1073/pnas.071034098
74. Schurr JM, Fujimoto BS. Equalities for the non-equilibrium work transferred from an external potential to a molecular system. analysis of single-molecule extension experiments. *J Phys Chem B* (2003) **107**:14007–19. doi:10.1021/jp0306803
75. Hendrix DA, Jarzynski C. A “fast growth” method of computing free energy differences. *J Chem Phys* (2001) **114**:5974. doi:10.1063/1.1353552

76. Oberhofer H, Dellago C, Geissler PL. Biased sampling of non-equilibrium trajectories: can fast switching simulations outperform conventional free energy calculation methods? *J Phys Chem B* (2005) **109**:6902–15. doi:10.1021/jp044556a
77. Rudolph MG, Zuckerman DM, Stanfield RL, Woolf TB, Wilson IA. Systematic finite-sampling inaccuracy in free energy differences and other non-linear quantities. *J Stat Phys* (2004) **114**:1303. doi:10.1023/B:JOSS.0000013961.84860.5b
78. Shirts MR, Pande VS. Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J Chem Phys* (2005) **122**:144107. doi:10.1063/1.1873592
79. Ytreberg FM, Swendsen RH, Zuckerman DM. Comparison of free energy methods for molecular systems. *J Chem Phys* (2006) **125**:184114. doi:10.1063/1.2378907
80. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* (2000) **33**:889–97. doi:10.1021/ar000033j
81. Gilson MK, Honig B. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins* (1988) **4**:7–18. doi:10.1002/prot.340040104
82. Gilson MK, Honig BH. Energetics of charge-charge interactions in proteins. *Proteins* (1988) **3**:32–52. doi:10.1002/prot.340030104
83. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc* (1990) **112**:6127–9. doi:10.1021/ja00172a038
84. Gohlke H, Kiel C, Case DA. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J Mol Biol* (2003) **330**:891–913. doi:10.1016/S0022-2836(03)00610-7
85. Amidon GL, Yalkowsky SH, Anik ST, Valvani SC. Solubility of non-electrolytes in polar solvents. V. Estimation of the solubility of aliphatic monofunctional compounds in water using a molecular surface area approach. *J Phys Chem* (2013) **79**:2239–46. doi:10.1021/j100588a008
86. Tidor B, Karplus M. The contribution of vibrational entropy to molecular association. The dimerization of insulin. *J Mol Biol* (1994) **238**:405–14. doi:10.1006/jmbi.1994.1300
87. Guo J, Wang X, Sun H, Liu H, Yao X. The molecular basis of IGF-II/IGF2R recognition: a combined molecular dynamics simulation, free-energy calculation and computational alanine scanning study. *J Mol Model* (2012) **18**:1421–30. doi:10.1007/s00894-011-1159-4
88. Lafont V, Schaefer M, Stote RH, Altschuh D, Dejaegere A. Protein-protein recognition and interaction hot spots in an antigen-antibody complex: free energy decomposition identifies “efficient amino acids.” *Proteins* (2007) **67**:418–34. doi:10.1002/prot.21259
89. Oberlin M, Kroemer R, Mikol V, Minoux H, Tastan E, Baurin N. Engineering protein therapeutics: predictive performances of a structure-based virtual affinity maturation protocol. *J Chem Inf Model* (2012) **52**:2204–14. doi:10.1021/ci3001474
90. Zoete V, Michielin O. Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: application to the TCR-p-MHC complex. *Proteins* (2007) **67**:1026–47. doi:10.1002/prot.21395
91. Massova I, Kollman PA. Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. *J Am Chem Soc* (1999) **121**:8133–43. doi:10.1021/ja990935j
92. 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**:251–61. doi:10.1016/S1074-7613(00)80178-8
93. Richman SA, Kranz DM. Display, engineering, and applications of antigen-specific T cell receptors. *Biomol Eng* (2007) **24**:361–73. doi:10.1016/j.bioeng.2007.02.009
94. Cole DK, Pumphrey NJ, Boulter JM, Sami M, Bell JI, Gostick E, et al. Human TCR-binding affinity is governed by MHC class restriction. *J Immunol* (2007) **178**:5727–34.
95. Johnson LA, Morgan RA, Dudley ME, Cassard L, Yang JC, Hughes MS, et al. Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood* (2009) **114**:535–46. doi:10.1182/blood-2009-03-211714
96. Kalergis AM, Boucheron N, Doucey MA, Palmieri E, Goyarts EC, Vegh Z, et al. Efficient T cell activation requires an optimal dwell-time of interaction between the TCR and the pMHC complex. *Nat Immunol* (2001) **2**:229–34. doi:10.1038/85286
97. Valitutti S. The serial engagement model 17-years after: from tcr triggering to immunotherapy. *Front Immunol* (2012) **3**:272. doi:10.3389/fimmu.2012.00272
98. Aleksic M, Dushek O, Zhang H, Shenderov E, Chen J-L, Cerundolo V, et al. Dependence of T cell antigen recognition on T cell receptor-peptide MHC confinement time. *Immunity* (2010) **32**:163–74. doi:10.1016/j.immuni.2009.11.013
99. 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**:223–39. doi:10.1016/j.jmb.2004.11.057
100. 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 U S A* (2000) **97**:5387–92. doi:10.1073/pnas.080078297
101. Holler PD, Kranz DM. Quantitative analysis of the contribution of TCR/pepMHC affinity and CD8 to T cell activation. *Immunity* (2003) **18**:255–64. doi:10.1016/S1074-7613(03)00019-0
102. Li Y, Moysey R, Molloy PE, Vuidepot A-L, Mahon T, Baston E, et al. Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat Biotechnol* (2005) **23**:349–54. doi:10.1038/nbt1070
103. Dunn SM, Rizkallah PJ, Baston E, Mahon T, Cameron B, Moysey R, 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**:710–21. doi:10.1110/ps.051936406
104. Bowerman NA, Crofts TS, Chlewicki L, Do P, Baker BM, Christopher Garcia K, et al. Engineering the binding properties of the T cell receptor:peptide:MHC ternary complex that governs T cell activity. *Mol Immunol* (2009) **46**:3000–8. doi:10.1016/j.molimm.2009.06.012
105. 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**:308–21. doi:10.1016/j.jmb.2005.08.041
106. 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**:1305–15. doi:10.1006/jmbi.2001.4560
107. Zhao Y, Bennett AD, Zheng Z, Wang QJ, Robbins PF, Yu LYL, 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**:5845–54.
108. 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**:175–84. doi:10.1016/j.jim.2008.09.016
109. 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 U S A* (2000) **97**:14578–83. doi:10.1073/pnas.97.26.14578
110. Hawkins RE, Gilham DE, Debets R, Eshhar Z, Taylor N, Abken H, et al. Development of adoptive cell therapy for cancer: a clinical perspective. *Hum Gene Ther* (2010) **21**:665–72. doi:10.1089/hum.2010.086
111. Barrozo A, Borstnar R, Marloie G, Kamerlin SCL. Computational protein engineering: bridging the gap between rational design and laboratory evolution. *Int J Mol Sci* (2012) **13**:12428–60. doi:10.3390/ijms131012428
112. Karanickolas J, Kuhlman B. Computational design of affinity and specificity at protein-protein interfaces. *Curr Opin Struct Biol* (2009) **19**:458–63. doi:10.1016/j.sbi.2009.07.005
113. Pantazes RJ, Maranas CD. OptCDR: a general computational method for the design of antibody complementarity determining regions for

- targeted epitope binding. *Protein Eng Des Sel* (2010) **23**:849–58. doi:10.1093/protein/gzq061
114. Pantazes RJ, Grisewood MJ, Maranas CD. Recent advances in computational protein design. *Curr Opin Struct Biol* (2011) **21**:467–72. doi:10.1016/j.sbi.2011.04.005
 115. Samish I, MacDermaid CM, Perez-Aguilar JM, Saven JG. Theoretical and computational protein design. *Annu Rev Phys Chem* (2011) **62**:129–49. doi:10.1146/annurev-physchem-032210-103509
 116. Haidar JN, Pierce B, Yu Y, Tong W, Li M, Weng Z. Structure-based design of a T-cell receptor leads to nearly 100-fold improvement in binding affinity for pepMHC. *Proteins* (2009) **74**:948–60. doi:10.1002/prot.22203
 117. Le Gal F-A, Ayyoub M, Dutoit V, Widmer V, Jäger E, Cerottini J-C, et al. Distinct structural TCR repertoires in naturally occurring versus vaccine-induced CD8⁺ T-cell responses to the tumor-specific antigen NY-ESO-1. *J Immunother* (2005) **28**:252–7. doi:10.1097/01.cji.0000161398.34701.26
 118. Atanackovic D, Arfsten J, Cao Y, Gnjatic S, Schnieders F, Bartels K, et al. Cancer-testis antigens are commonly expressed in multiple myeloma and induce systemic immunity following allogeneic stem cell transplantation. *Blood* (2007) **109**:1103–12. doi:10.1182/blood-2006-04-014480
 119. Chen YT, Scanlan MJ, Sahin U, Türeci O, Gure AO, Tsang S, et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc Natl Acad Sci U S A* (1997) **94**:1914–8. doi:10.1073/pnas.94.5.1914
 120. Liggins AP, Lim SH, Soilleux EJ, Pulford K, Banham AH. A panel of cancer-testis genes exhibiting broad-spectrum expression in haematological malignancies. *Cancer Immun* (2010) **10**:8.
 121. Lim SH, Austin S, Owen-Jones E, Robinson L. Expression of testicular genes in haematological malignancies. *Br J Cancer* (1999) **81**:1162–4. doi:10.1038/sj.bjc.6690824
 122. Marsh S, Parham P, Barber LD. *The HLA Factsbook*. (1999).
 123. Yuan J, Adamow M, Ginsberg BA, Rasalan TS, Ritter E, Gallardo HF, et al. Integrated NY-ESO-1 antibody and CD8⁺ T-cell responses correlate with clinical benefit in advanced melanoma patients treated with ipilimumab. *Proc Natl Acad Sci U S A* (2011) **108**:16723–8. doi:10.1073/pnas.1110814108
 124. Chen J-L, Stewart-Jones G, Bossi G, Lissin NM, Wooldridge L, Choi EML, et al. Structural and kinetic basis for heightened immunogenicity of T cell vaccines. *J Exp Med* (2005) **201**:1243–55. doi:10.1084/jem.20042323
 125. Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S, et al. The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res* (2013) **41**:D475–82. doi:10.1093/nar/gks1200
 126. Sami M, Rizkallah PJ, Dunn S, Molloy P, Moysey R, Vuidepot A, 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. doi:10.1093/protein/gzm033
 127. Hubbard RE, Haider MK. Hydrogen bonds in proteins: role and strength. *Encyclopedia of Life Science* (2001). 1–6.
 128. Chang HC, Bao Z, Yao Y, Tse AG, Goyarts EC, Madsen M, et al. A general method for facilitating heterodimeric pairing between two proteins: application to expression of alpha and beta T-cell receptor extracellular segments. *Proc Natl Acad Sci U S A* (1994) **91**:11408–12. doi:10.1073/pnas.91.24.11408
 129. Boulter JM, Glick M, Todorov PT, Baston E, Sami M, Rizkallah P, et al. Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Eng* (2003) **16**:707–11. doi:10.1093/protein/gzg087
 130. Janin J. The kinetics of protein-protein recognition. *Proteins* (1997) **28**:153–61. doi:10.1002/(SICI)1097-0134(199706)28:2<153::AID-PROT4>3.0.CO;2-G
 131. Tian S, Maile R, Collins EJ, Frelinger JA. CD8⁺ T cell activation is governed by TCR-peptide/MHC affinity, not dissociation rate. *J Immunol* (2007) **179**:2952–60.
 132. Kersh EN, Shaw AS, Allen PM. Fidelity of T cell activation through multistep T cell receptor zeta phosphorylation. *Science* (1998) **281**:572–5. doi:10.1126/science.281.5376.572
 133. McKeithan TW. Kinetic proofreading in T-cell receptor signal transduction. *Proc Natl Acad Sci U S A* (1995) **92**:5042–6. doi:10.1073/pnas.92.11.5042
 134. 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**:148–51. doi:10.1038/375148a0
 135. Dushek O, Aleksic M, Wheeler RJ, Zhang H, Cordoba S-P, Peng Y-C, et al. Antigen potency and maximal efficacy reveal a mechanism of efficient T cell activation. *Sci Signal* (2011) **4**:ra39. doi:10.1126/scisignal.2001430

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 May 2013; paper pending published: 10 June 2013; accepted: 19 August 2013; published online: 12 September 2013.

Citation: Zoete V, Irving M, Ferber M, Cuendet MA and Michielin O (2013) Structure-based, rational design of T cell receptors. *Front. Immunol.* **4**:268. doi: 10.3389/fimmu.2013.00268

This article was submitted to *T Cell Biology*, a section of the journal *Frontiers in Immunology*.

Copyright © 2013 Zoete, Irving, Ferber, Cuendet and Michielin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.