Rapid Assessment of T-Cell Receptor Specificity of the Immune Repertoire

Xingcheng Lin[1,2,*], Jason T. George[1,3,*], Nicholas P. Schafer[1,4], Cecilia Clementi[1,4], José N. Onuchic[1,5,†], and Herbert Levine[1,6,†]

[1]Center for Theoretical Biological Physics, Rice University, Houston, TX

[2]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA

[3]Medical Scientist Training Program, Baylor College of Medicine, Houston, TX

[4]Departments of Chemistry, Rice University, Houston, TX

[5]Department of Physics and Astronomy, Rice University, Houston, TX

[6]Department of Physics, Northeastern University, Boston, MA

[*]Equal contribution

[†]To whom correspondence should be addressed: jonuchic@rice.edu, h.levine@northeastern.edu

Abstract

Accurate assessment of TCR-antigen specificity at the whole immune repertoire level lies at the heart of improved cancer immunotherapy, but predictive models capable of high-throughput assessment of TCR-peptide pairs are lacking. Recent advances in deep sequencing and crystallography have enriched the data available for studying TCR-p-MHC systems. Here, we introduce a pairwise energy model, RACER, for rapid assessment of TCR-peptide affinity at the immune repertoire level. RACER applies supervised machine learning to efficiently and accurately resolve strong TCR-peptide binding pairs from weak ones. The trained parameters further enable a physical interpretation of interacting patterns encoded in each specific TCR-p-MHC system. When applied to simulate thymic selection of a T-cell repertoire, RACER accurately estimates recognition rates for tumor-associated neoantigens, foreign peptides, and allogeneic T-cells, thus demonstrating its capability in handling the large computational challenge of reliably identifying tumor antigen-specific T-cells at the level of an individual patient's immune repertoire.

1 Introduction

The advent of new strategies that unleash the host immune system to battle malignant cells represents one of the largest paradigm shifts in treating cancer and has ushered in a new frontier of cancer immunotherapy [1]. Various treatments have emerged, including checkpoint blockade therapy [2, 3, 4], tumor antigen vaccine development [5, 6], and the infusion of a donor-derived admixtures of immune cells [7]. A majority of successful treatments to-date rely on the anti-tumor potential of the CD8+ T-cell repertoire, a collection of immune cells capable of differentiating between malignant cells and normal tissue by recognizing tumor-associated neoantigens (TANs) detectable on the cell surface [8]. Therefore, accurately assessing a T-cell repertoire's ability to preferentially recognize cancer cells by recognizing their tumor antigens lies at the heart of optimizing cancer immunotherapy.

T-cell clones, each with a uniquely generated T-cell receptor (TCR), undergo a thymic negative selection process whereby immature T-cells are trained to interface with numerous peptides presented on a major histocompatibility complex (p-MHC) molecule and to avoid self-peptide recognition [9, 10, 11, 12]. During negative selection, each TCR interfaces with many ($\sim 10^4$) self-peptides [13] presented by thymic medullary epithelial cells, surviving only if it does not bind too strongly. This process imparts surviving T-cells with central tolerance to major self-peptides and influences many of the recognition properties of the resultant repertoire. Although T-cells also undergo more subtle and reversible systems-level peripheral tolerance tests in response to dynamic pathogen encounters [14, 15], negative selection is responsible for generating durable tolerance.

To understand the effects of tolerance on T-cell recognition at the cell population level, prior theoretical models have focused on studying general statistical properties of the resulting T-cell repertoire, which has been estimated to consist of $\sim 10^7$ unique TCR clones [16]. Properties pertinent

to immunotherapy and self vs. non-self discrimination include the effectiveness of the repertoire at discerning tumor from self-antigens [17], the repertoire's ability to impart immunity against current and future threats [18, 19], and the extent of selection pressure that the repertoire exerts on an evolving cancer population [20, 21]. All such approaches rely on a detailed study of the tumor-immune interaction, which is a formidable task, owing to the large theoretical upper limits of TCR diversity ($\sim 10^{20}$) [22] and antigen space ($\sim 20^{9}$) [23]. Computational tools are therefore needed for rapid and accurate assessment of TCR specificity that are capable of differentiating between self and foreign antigens at the repertoire level.

Previously published methods have elucidated the molecular mechanism of TCR-p-MHC interactions. To-date, state-of-art experimental investigations have integrated solved crystal structures [24, 25] with in-depth peptide sequencing [26, 27] to probe the physiochemical hallmarks of various epitope-specific TCRs. Publicly available crystal structures have allowed researchers to identify detailed structural features determining the binding specificity of TCR-p-MHC pairs at atomistic resolution. These efforts have enabled subsequent analyses that attribute differences in TCR-p-MHC binding affinity to subtle geometrical changes in the binding interface, as well as the side-chain packings of residue pairs. However, despite existing structures for a few individual examples, solving crystal or cryo-EM structures for large-scale investigation of TCR-p-MHC interactions still remains a formidable task.

Alternate methods utilizing either template-based structural modeling [28] or rigid-body docking [29] have increased calculation speeds to, at most, one structure per minute. Thus it is unlikely in the foreseeable future that such approaches will be used to investigate the number of TCR-peptide interactions necessary to study the problem at the immune repertoire level, which can easily require the assessment of more than 10^{9} pairs simultaneously [13]. Additionally, identifying the correct

structure is only the first step toward evaluating the TCR-peptide binding affinities, as detailed quantification of binding affinity relies on the interplay of energetic interactions among different residue pairs on the TCR-p-MHC interface. As a useful approximation, one can implement statistical scores calculated from docking algorithms as a surrogate measure of binding affinity [29], but such scoring functions impart inherent bias favoring existing crystal structures in the database and consequently have a diminished ability to differentiate between strong- and weak-binding peptides in cases where crystal structures are unavailable.

One way to address the aforementioned problems is to first obtain an energy model governing TCR-p-MHC interactions, which, when applied to a TCR-peptide pair, evaluates binding energy. Inspired by similar approaches utilized for investigations of protein folding based on peptide primary sequence [30, 31, 32, 33], we implement a coarse-grained, while chemically-accurate energy model to accelerate energy evaluation time. The model is optimized to recognize strong TCR-peptide interactions based on experimentally obtained structures and sequences, with the goal of defining interactions pertinent to the TCR-peptide environment so that affinity could subsequently be determined solely by primary sequence data. The growing number of TCR-p-MHC solved crystal structures and their sequences provide a plethora of pertinent information on TCR-peptide interactions. Combining this experimental data with a supervised machine learning approach to obtain an optimized energy model has enabled us to efficiently harness as much of this information as possible, thereby providing a tractable means to extract pertinent TCR-peptide residue interactions. Here, we parametrize a pairwise energy model by maximizing the specificity of TCRs for differentiating between strong and weak binding peptides, in an analogous optimization scheme previously developed to understand protein folding and binding more generally [30, 33, 34, 35, 36, 37, 38]. The optimized energy model, which we refer to as the Rapid Coarse-grained Epitope TCR (RACER), can accurately iden-

tify strong-binding peptides from weak ones across various TCRs and validation tests. Guided by a physical mechanism, the optimized parameters of RACER proved instrumental in understanding the specific interactions leading to TCR recognition. Furthermore, RACER can be applied in a straightforward manner to study all TCR-peptide interactions at the whole-repertoire level using modest computational resources. We demonstrate RACER's ability by generating population-level behavior consistent with prior theoretical estimates of TCR recognition [17], thereby demonstrating proof of its utility in evaluating the post-selection adaptive immune system's ability to recognize TANs and foreign antigens.

2 Results

2.1 Overview of the supervised machine learning approach used to obtain the TCR-peptide energy model

RACER's optimization protocol (Fig. 1) utilizes high-throughput deep sequencing data on TCR-peptide interactions across a large peptide library [26], as well as known physical contacts between TCRs and peptides obtained from deposited crystal structures [39]. Adapting an approach previously implemented for studying folded proteins [30, 33], the RACER optimization strategy trains a pairwise energy model which maximizes TCR-peptide binding specificity. The energy model was optimized by maximizing the z-score defined to separate the affinities of experimentally determined strong-binding peptides, called 'strong binders' hereafter, from computationally generated, randomized decoys [1]. The optimized residue type-dependent force field can be used, in combination with

[1]The z-score is defined as the difference between the averaged binding energies of strong binders and decoys, divided by the standard deviation of the decoy energies

the contact map for each TCR-peptide system, to calculate the binding energies of an ensemble of TCR-peptide systems. These RACER-generated binding energies can subsequently be utilized to mimic TCR thymic negative selection, thereby enabling estimation of post-selection TCR recognition probabilities of TANs and foreign peptides. Lastly, the generated binding affinities were used on a large-scale simulation of hematopoietic stem cell transplant (HSCT).

Figure 2 summarizes RACER's predictive performance. The optimized energy model correctly resolves experimentally predetermined strong and weak binders. These known strong binding peptides for a given TCR, as well as the randomly generated decoy peptides that represent weak binders, were used to train the energy model for correct binding specificity prediction (See Method section for details). The resulting energy model was then applied to calculate binding energies for the training peptides, in addition to test peptides independently verified experimentally as weak binders. The model was tested on three independent TCRs: 2B4, 5CC7 and 226. Although the experimentally identified weak binders were omitted from the training set, RACER effectively resolves binding energies between experimentally determined strong and weak binders, with z-scores larger than 3.5 in all cases (Fig. 2a), highlighting the predictive power of this approach.

Strong binders for a particular TCR should only account for a small subset of the entire antigen space. Moreover, their prediction is of direct clinical importance in applications to immunotherapy, thus warranting further assessment of RACER's predictive power. To test RACER's ability to identify strong binders, we performed a leave-one-out-cross-validation (LOOCV) test, using data from TCR 2B4 as an example. For each test iteration, one known strong binder was withheld from the training set of 44 strong binders. The same optimization protocol was applied to train the energy model by using the remaining 43 peptides and then predicting the binding energy of the withheld peptide. This prediction was then compared to binding energies of known weak binders, and the

procedure was repeated 44 times for each peptide. Among the 44 test cases, our model is able to accurately distinguish the withheld strong binder in 43 cases (Fig. 2b). This is in stark contrast to a cluster-based attempt at strong binder identification based on peptide sequences alone, which at best correctly identifies 19 out of 44 strong binders (See SI for details).

In order to further characterize RACER's predictive power, an independent set of K_{d} values measured by surface plasmon resonance (SPR) [26] were compared with predicted affinity. The SPR experiments were performed over 9 independent peptides for each of the three TCRs. The free energies, $k_{\mathrm{B}}T\log(K_{\mathrm{d}})$, were compared with calculated binding energies from RACER for quantitatively testing the accuracy of binding affinity predictions. Lower binding energies indicate stronger binding affinity so that a positive correlation between the $k_{\mathrm{B}}T\log(K_{\mathrm{d}})$ values and calculated binding energies implies a successful prediction. As shown in Fig. 2c, RACER was able to correctly predict the binding affinities of these 9 peptides for all TCRs, with an average Pearson correlation coefficient of 0.74, and an average Spearman's rank correlation coefficient of 0.65.

2.2 RACER's residue type-dependent interactions were optimized specifically for TCR-peptide recognition

The specific crystal structures, as well as the sequences of the TCR-p-MHC system, determine a unique contact pattern that we used to encode a system-specific force field distinct from the one typically used for studying protein folding. To test RACER, we considered TCR 2B4 as an initial example (Fig. 3). The crystal structure of TCR 2B4 (Fig. 3a) reveals that there are many contacts formed between Threonine (T), Asparagine (N) and Alanine (A) residues. The observed probabilities of contact formation between these residues are significantly enriched in the set of strong

binders, while much less so in the decoy set (Fig. 3b), so that the optimized energy model shows the strongest attractions between the A-T, A-N residue pairs. This relative enrichment contrasts with Tryptophan (W) which frequently interacts with Alanine (A) in both the strong and decoy binders. As a result, the optimized energy model does not favor the interaction between those two types of residues.

Previous hydrophobicity-based protein force fields are well-established in the protein folding literature. One goal of this approach is to extract relevant information on interactions unique to TCR-peptide binding. We demonstrate the necessity of identifying TCR-peptide-specific content by showing that the standardized interaction strengths from a force field that was optimized for protein folding cannot fully distinguish between strong and weak binders. To test the performance of a typical protein-folding force field, we used the default direct-contact parameters from the original AWSEM force field [32], which was optimized based on deposited folded proteins, as well as the MJ force field [40], which is based on the probability distribution of contacting residues from the deposited protein database, to calculate the binding energies between the TCR 2B4 and peptides in the same way as RACER. Neither force field fully resolves these groups, with z-scores of 0.69 and 1.28, respectively (Fig. S1). A similar performance was observed utilizing TCRs 5CC7 and 226, effectively demonstrating that RACER encodes pertinent structural information necessary for studying the TCR-peptide system.

2.3 RACER shows transferability across different TCRs

Given RACER's ability to identify strong binders for a fixed choice of TCR, it is natural to next assess the transferability of RACER for binding specificity prediction across different TCRs. Toward this end, we used the energy model optimized using structural data for one of the three TCRs to

predict the TCR-peptide binding energies of the remaining two. Although the z-scores measured for TCRs that were not included in the training set are lower than the one that was used in energy model optimization, RACER is still able to distinguish a majority of strong binders from weak binders, with an average z-score of 1.8 (Fig. 4a). This demonstrates that, despite genetic diversity, distinct TCR-p-MHC systems still share similar structural-sequence patterns.

We next assessed whether RACER's predictive accuracy could be improved by the incorporation of additional TCRs in the optimization step. Toward this end, we subsequently included crystal structures for the remaining TCRs (5CC7 and 226) as two additional strong binders into our training set (of 2B4) for the transferability assessment (See Method for details). This procedure was repeated for each of the three TCRs. The newly optimized energy model demonstrates significant improvement in z-scores. This result suggests that future incorporation of additional crystal structures of target TCRs may lead to improved resolution of strong and weak binders via refinement of the optimized energy model.

2.4 Energy model optimized T-cell repertoire generation

One of the greatest strengths of RACER lies in its ability to simulate and quantify pairwise interactions between an entire T-cell repertoire and antigen landscape. Toward this end, pairwise interactions were calculated between 10^5 simulated TCRs and 10^4 self-peptides, requiring 5,000 CPU hours on an Intel(R) Xeon(R) CPU E5-2650 v2. The maximum binding energy for each TCR over all self-peptides defines a selection curve (Fig. 5a) for which there corresponds a percentage of surviving T-cells at each negative selection threshold in thymic selection. Although empirical estimates of the percentage of surviving TCRs during thymic negative selection vary between 20% and 50% [41, 42, 43], we calculate relevant recognition behavior for all selection rates, restricting our analysis

to 50% when applicable.

Most self-peptides present in thymic selection are expected to participate in the deletion of self-reactive T-cells. Thus, a reasonable model of thymic selection would feature a majority of self-peptides contributing to the selection of immature T-cells. A rank order of these self-peptides based on their ability to recognize unique T-cells characterizes the extent to which each self-peptide is utilized in thymic selection. The RACER-derived rank order generates reasonable behavior with a majority of self-peptides participating in selection (Fig. S4b). One key issue related to adaptive immune recognition of TANs is the level of recognition efficiency of peptides closely related to self (e.g. point mutants) relative to foreign peptide recognition. Comparison of the post-selection individual TCR recognition of foreign and point-mutant self-peptides shows little difference in recognition potential (Fig. 5b,c), with variances of these estimates overlapping with one another and in line with previous theoretical estimates (Fig. S4c). An analogous comparison of the post-selection TCR repertoire recognition probability of foreign and mutant peptides demonstrates this minimal difference is maintained at the aggregate immune system level (Fig. 5d). Lastly, our prior theoretical model posited thymic selection as an optimization problem with a survival cutoff of $1/e$ resulting in the production of maximally efficient thymic selection [17]. Calculating the product of survival and recognition probabilities yields a broad curve with large values located at intermediate survival cutoffs, including the previously predicted optimal survival cutoff (Fig. S2d). Taken together, these results exhibit consistency with previous studies and reinforce the utility of RACER for performing repertoire-level analyses.

2.5 Simulating allogeneic cancer immunotherapy

Allogeneic hematopoietic stem cell transplant (HSCT) is a particularly powerful application of cancer immunotherapy that utilizes the TCR repertoire of a healthy donor to target and eliminate host cancer cells [44]. Due to the complexity of balancing desirable recognition of tumor antigens with the risk of potentially fatal graft-versus-host-disease, a majority of the recommendations which guide this therapy are empirically driven. We harnessed the power of our model for rapid evaluation of T-cell recognition to simulate HSCT by comparing the host- and donor-derived post-selection TCR pools. Both thymic selection processes demonstrated self-consistency in the binding profile of randomly generated TCRs and self-peptides in thymic selection (Fig. 6a). The proportion of donor-derived TCRs reactive to host self-peptides defines the alloreactivity rate and contributes to undesirable graft-versus-host disease [45], a clinically important consideration in HSCT therapy. Empirical estimates of the alloreactivity rate vary substantially, up to 20% in unmatched pairs [46, 47]. For MHC-matched host and donor, the estimates are lower, ranging from 0.1% to 10% [48, 49]. Assuming only host-donor single nucleotide polymorphism differences in roughly 6% of the thymic self-peptide pool, our calculations yield an alloreactivity rate of order 0.1% (Fig. 6b) in agreement with the lower limit of empirical estimates. HSCT, though effective in treating hematological malignancies [45], is an invasive medical procedure with decision-making largely guided by prior empirical outcomes. An *a priori* assessment of putative tumor antigens would substantially improve treatment risk-benefit assessment. We demonstrate this by studying the ability of a post-selection T-cell repertoire to recognize foreign (resp. mutant) peptides, representing mis-localized self-peptides (resp. TANs) present in cancer cells, repeated for various levels of pre-selection T-cell diversity ($\sim 5 \cdot 10^3$ to 10^5). Our results highlight the ability of our approach to quantify the predicted benefit

of transplant (Fig. 6c,d).

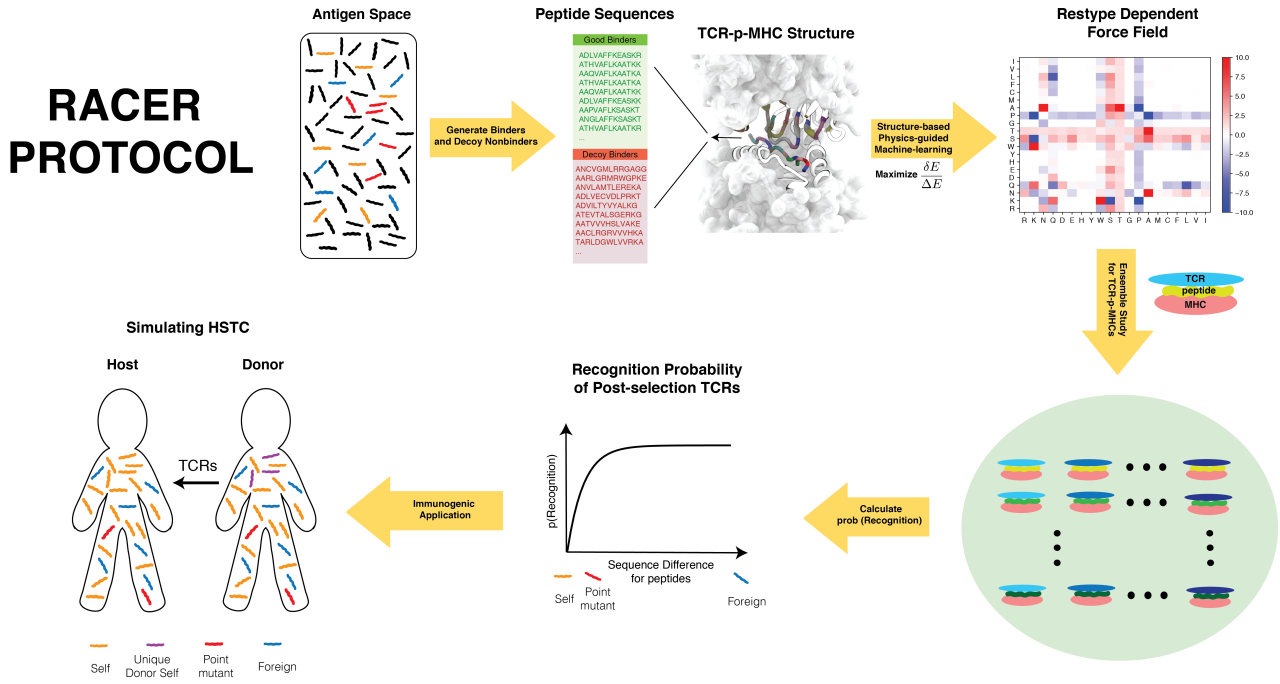

Figure 1: Summary of the modeling approach employed in this study. The optimization of RACER starts from a series of TCR binders obtained from the deep-sequencing experiments [26], as well as the corresponding TCR-p-MHC crystal structures deposited in the database [39]. The sequences of the strong binders, as well as the generated decoy binders from randomizing the non-anchoring sequences of the strong binders, are collected for parameterizing a pairwise energy model which maximizes the energetic gap between the strong binders and a randomized set of decoys. The resulting energy model can be used to quickly evaluate the binding affinities of an ensemble of TCR-peptide interactions at the population level. The calculated binding affinities can be used for simulating the negative selection process in the thymus, as well as measuring the recognition probability of the post-selection TCRs. Finally, this kind of ensemble study can be used to for immunogenic applications that require information of the entire T-cell repertoires, such as simulation of hematopoietic stem cell transplant (HSCT).

3 Discussion

In this paper, we have presented an optimized, pairwise molecular energy model that can be utilized to quickly assess TCR-peptide interactions and distinguish strong binding pairs using their primary sequences. The approach, which we call RACER, enables rapid and accurate TCR-peptide binding assessments with computational efficiency capable of evaluating interactions between putative

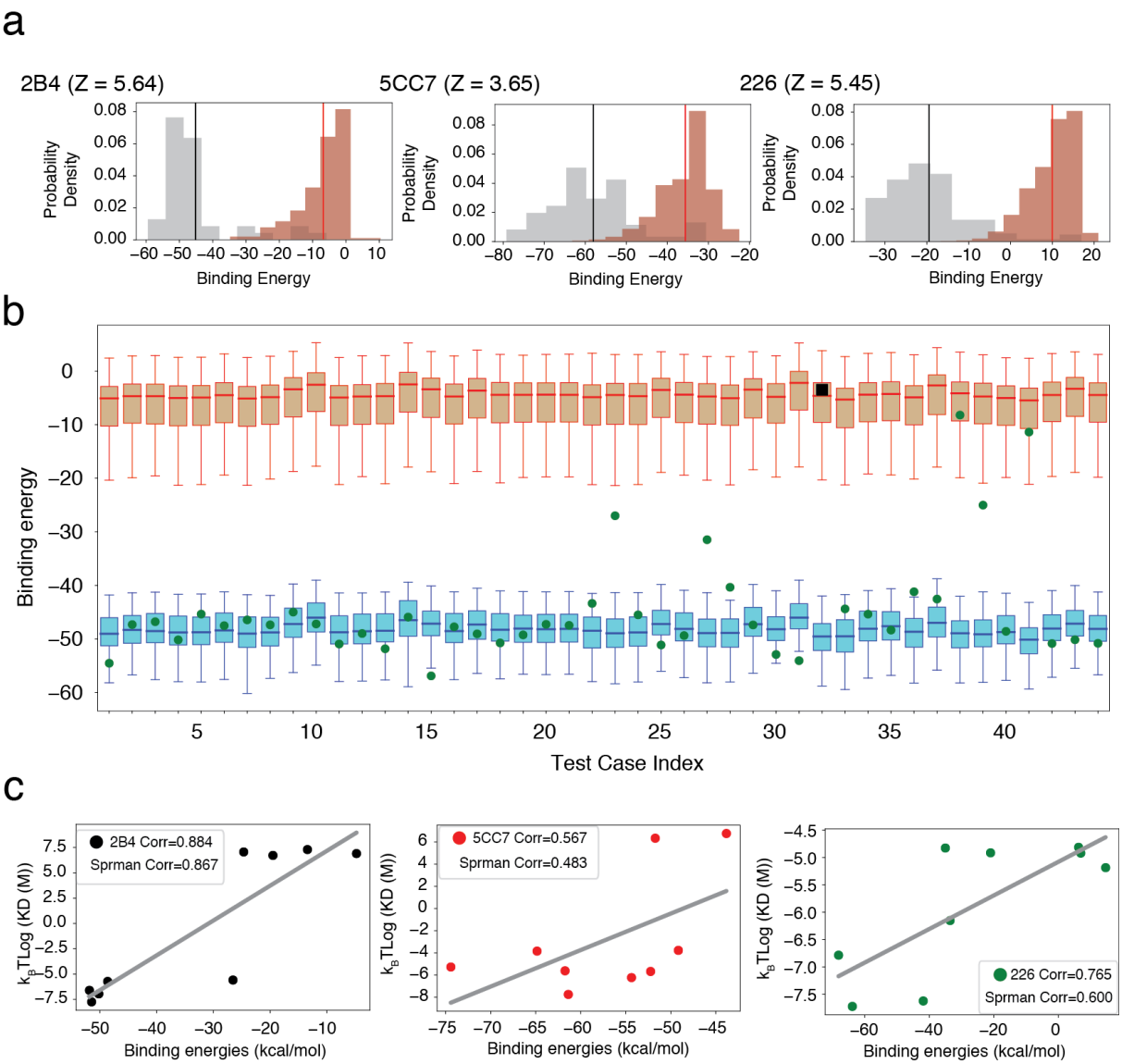

Figure 2: RACER can fully separate the strong binders of a specific TCR from its weak binders. **a,** For three TCRs (2B4, 5CC7 and 226) whose strong and weak binders have been experimentally determined [26], the calculated binding energies based on RACER can well separate the strong binders from the weak ones of each individual TCR. **b.** In the leave-one-out-cross-validation exemplified using the TCR 2B4, RACER can successfully recognize the withheld strong binders in 43 out of 44 tests, where the predicted binding energies of the withheld test binder (green) is lower than the median (red bar) of the experimentally determined weak binders. The only exception is marked as a black square. **c,** In a completely independent testing data measured by surface plasmon resonance (SPR) [26], the calculated binding energies of testing peptides correlate well with their experimentally determined dissociation constant K_d. Corr: Pearson correlation coefficient. Sprman Corr: Spearman's rank correlation coefficient. More details of selected peptides for each TCR are reported in figure S3.

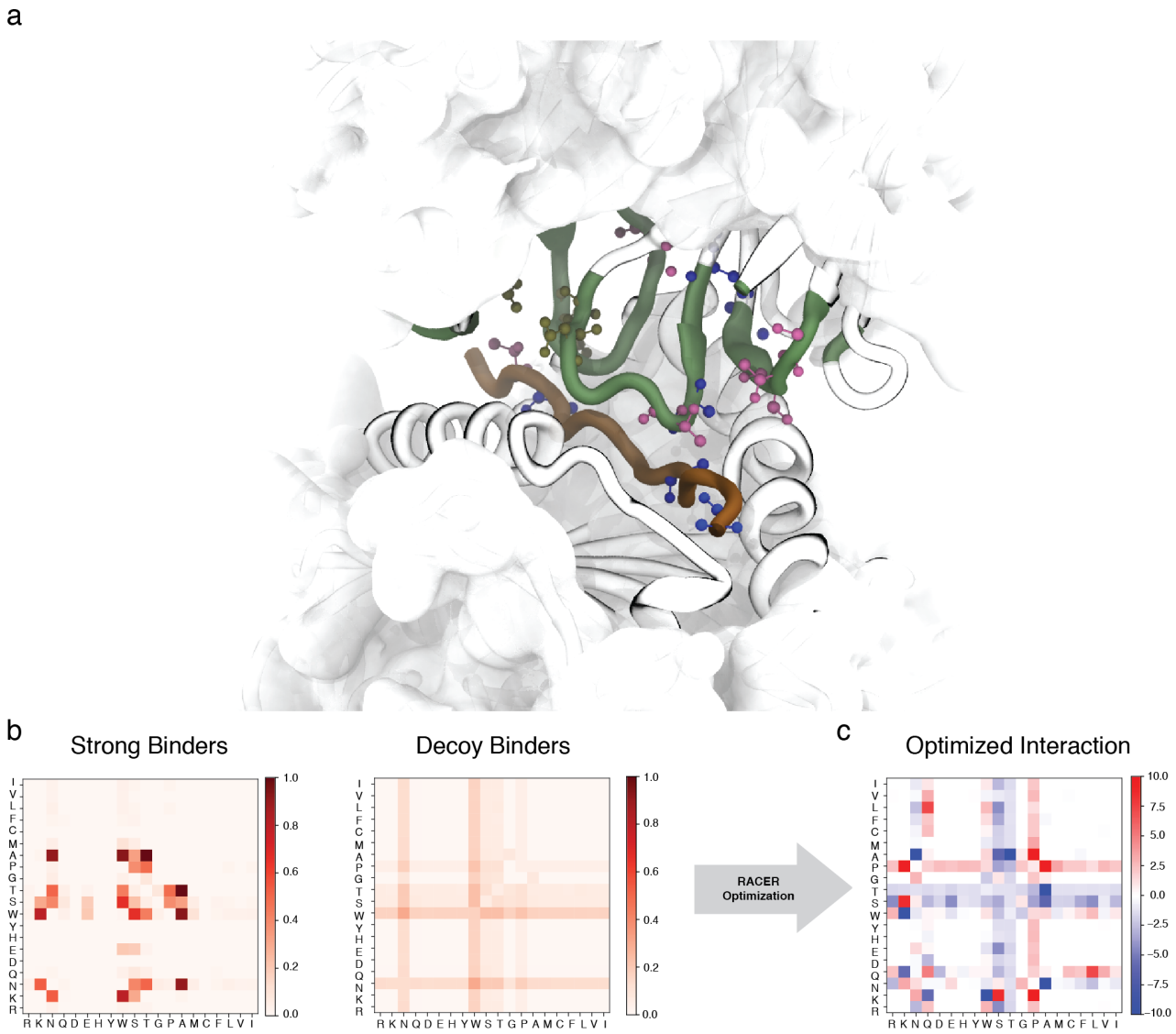

Figure 3: The specific contacts from TCR-peptide structure dictates a optimized energy model different from those of a typical protein-folding force field. **a,** The 3D crystal structure of the 2B4 TCR (PDBID: 3QIB). The residues pairs that are in contact between the TCR and peptide are color-highlighted as green (TCR) and orange (peptide). Also shown are residues Alanine (blue), Threonine (magenta) and Asparagine (tan) which are prevalent in this structure (CPK representation [50]). **b,** The probability of contacts formation between each two of the 20 amino acids in the set of strong binders (up) and the set of randomized decoy binders (bottom) of the TCR 2B4. **c,** The residue-based interaction strength determined by RACER for the TCR 2B4. A more negative value indicates a stronger attractive interaction between the corresponding two residues.

peptide targets and a complete immune repertoire. RACER requires only ∼0.02s for evaluating one TCR-peptide pair, thousands of times faster than the current state-of-the-art, while preserving sufficient prediction accuracy (Figs. 2, 4). Consequently, our approach can be used to study TCR-p-

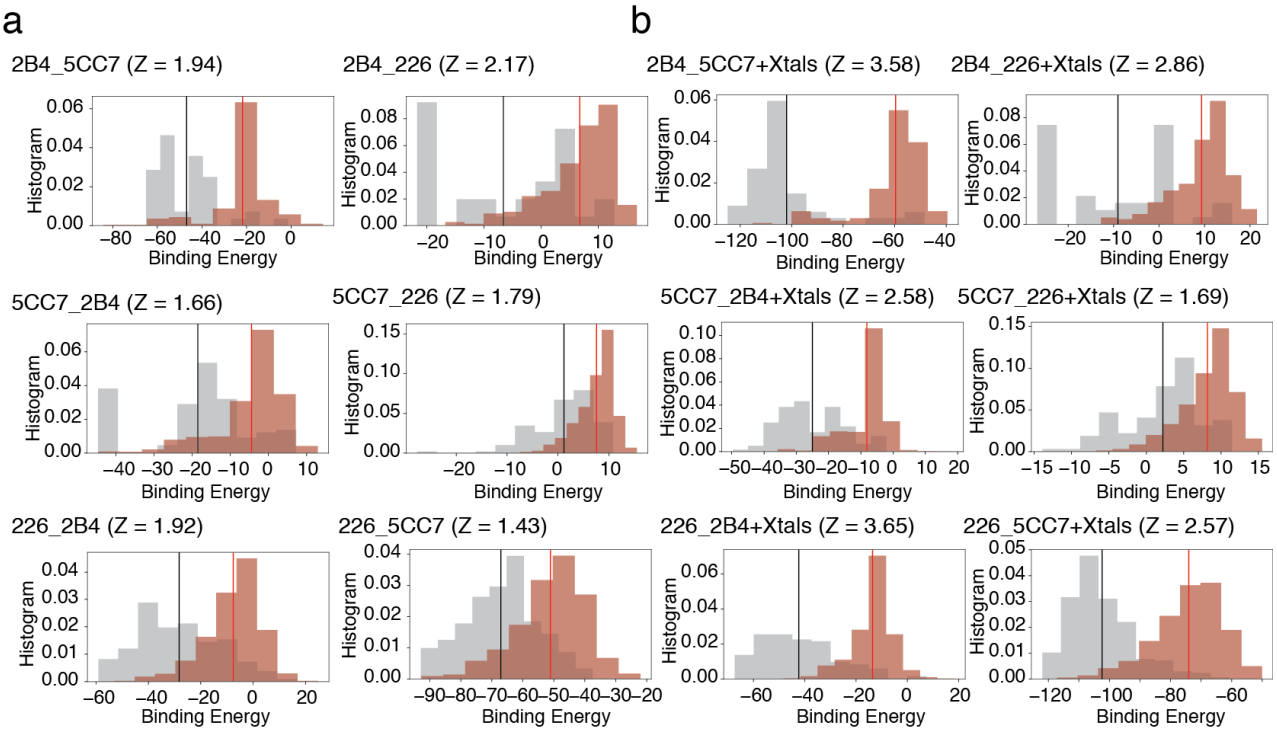

Figure 4: RACER shows transferability in terms of predicting TCR-p-MHC interactions across different TCRs. **a,** The energy model trained based on one TCR (e.g. 2B4) is capable of resolving the experimentally determined strong binders from weak binders of the other two TCRs (e.g., 5CC7 and 226). **b,** By adding strong binders from crystal structures of the other two TCRs into training sets, RACER can be further improved for identifying the experimentally determined strong binders. The title of each figures follows the format of "target_training TCRs", e.g., "2B4_5CC7" means using the energy model trained from the TCR 5CC7 for predicting the peptide binding affinities of the TCR 2B4. "Xtals" means the strong binders from the crystal structures of the other two TCRs were added into the training set.

MHC interactions at a population level sufficient for modeling *in vivo* thymic selection and immune repertoire antigen recognition.

3.1 Specificity v.s. Generality of the optimized energy model

The unique topology of the TCR-p-MHC structure endows a system-specific contact map between interacting TCRs and peptides. Moreover, the arrangement of contacts gives rise to a post-optimization energy model distinct from the traditional hydrophobic-hydrophilic interaction patterns [51] used for protein folding, which have previously been utilized in force fields such as the MJ potential [40] to

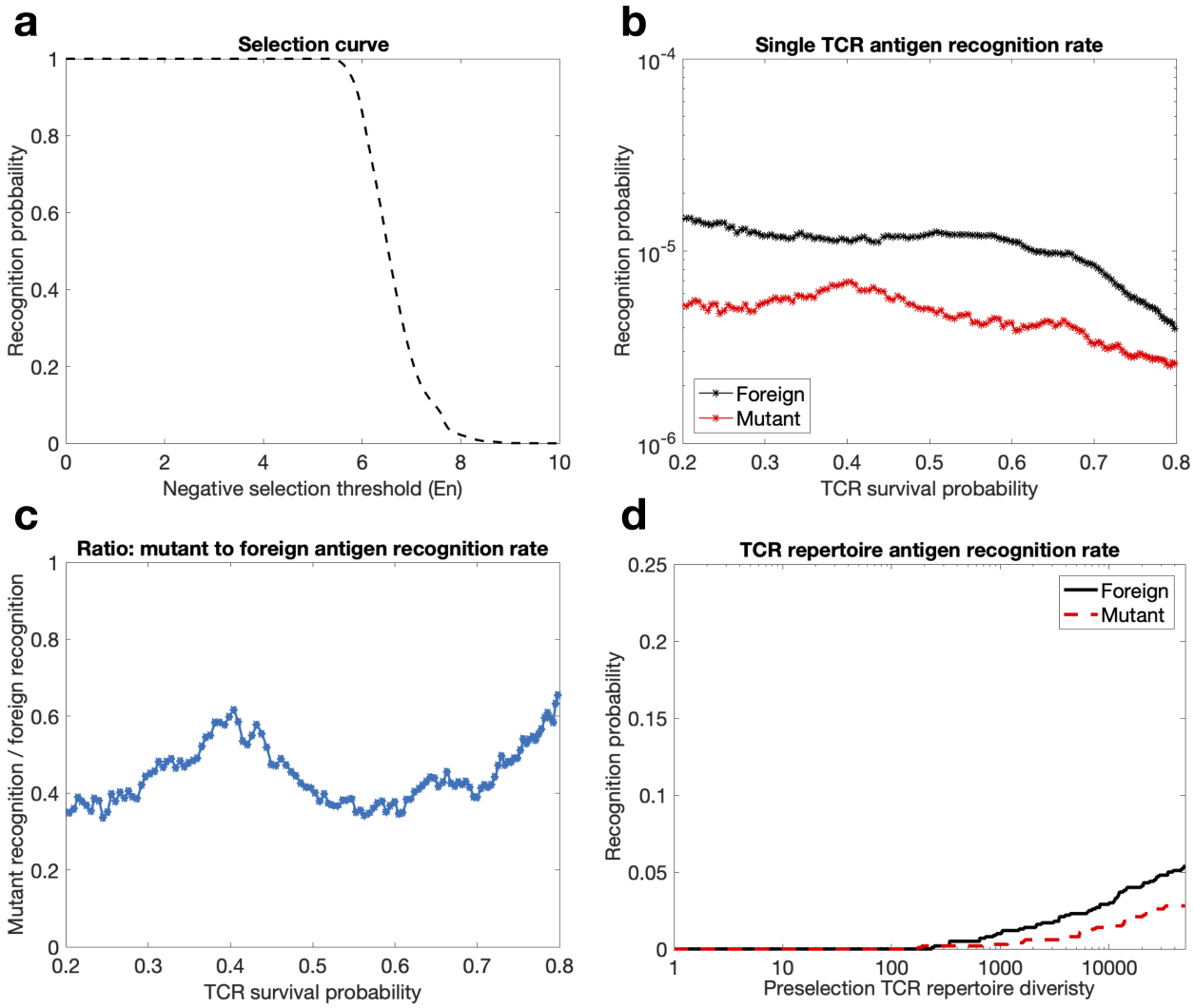

Figure 5: T-cell repertoire simulations of thymic selection and antigen recognition in the RACER model. **a,** T-cell recognition probability as a function of binding energy cutoff based on simulated thymic selection with $N_n = 10^4$ self-peptides ($N_t = 10^5$ simulated peptides). **b,** The probability of recognizing foreign and point-mutant antigens for individual post-selection T-cells is plotted as a function of percentage of TCRs surviving negative selection (ordinate of the graph in panel a, simulations averaged over all post-selection TCRs with pairwise interactions amongst 10^3 random peptides and 10^3 point-mutant peptides). **c,** The ratio of individual TCR recognition of mutant to foreign peptide recognition as a function of TCR survival percentage. **d,** The probability of T-cell repertoire recognition of foreign (black) and mutant (red) peptides as a function of preselection TCR repertoire diversity, with negative selection thresholds giving 50% survival.

study TCR-peptide binding. Significantly, the sequences and structures of TCR-peptide systems were found experimentally to be relatively conserved among various peptides [26, 27, 25]. The preserved sequence and structural features could dramatically limit the physiochemical space explorable by TCR-peptide residue pairs. Since RACER is optimized on the TCR-peptide system, it features

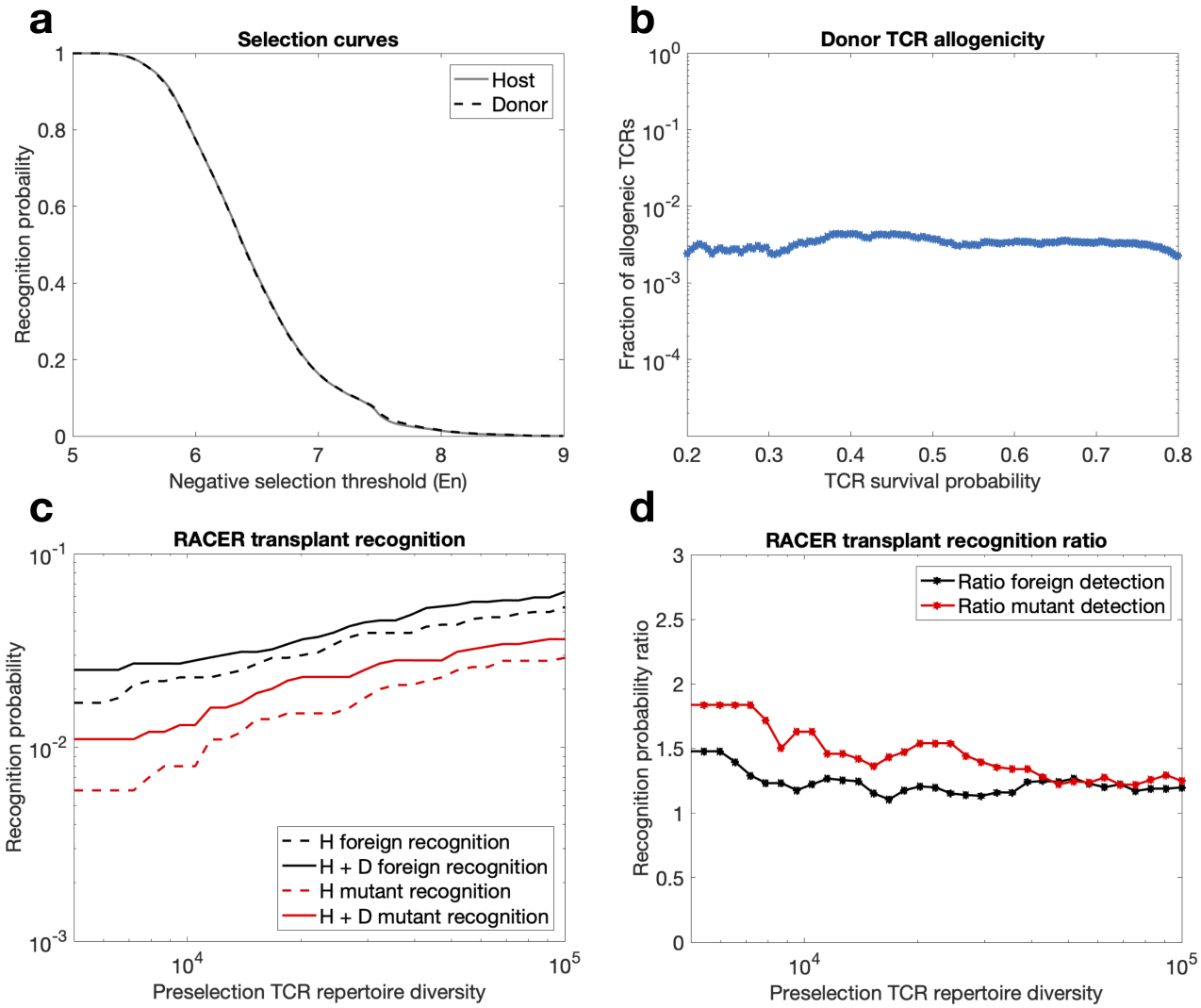

Figure 6: T-cell repertoire simulations of hematopoietic stem cell transplant in the RACER model. **a,** T-cell recognition probability as a function of binding energy cutoff based on simulated thymic selection with $N_n = 10^4$ self-peptides ($N_t = 10^5$ simulated peptides) for distinct, randomly-generated TCRs and self-peptides in host and donor selection. **b,** The fraction of donor-derived post-selection TCRs which recognize host self-peptide, assuming host and donor proteomes differ by single nucleotide polymorphism in 600 out of 10^4 thymic self-peptides, plotted as a function of percentage of TCRs surviving negative selection (ordinate of the graph in panel a). **c,** The recognition probability of foreign (black) and point-mutant (red) peptides by host-derived post-selection TCR repertoires (dashed), in addition to the recognition probability by both host and donor TCRs repertoires (solid) is plotted as a function of preselection TCR repertoire diversity. **d,** The corresponding ratio of host and donor recognition to recognition solely based on host repertoire is plotted as a function of preselection TCR repertoire diversity (negative selection thresholds chosen to generate 50% survival).

system-dependent parameters distinct from those used for describing generic protein-protein interactions (Fig. 3). This hypothesis is strongly supported by the observation that RACER is capable of

identifying strong binders of corresponding TCRs (Fig. 2) while previous methods fall short.

The departure of RACER from a typical protein-folding force field also reflects a myriad of other environmental factors that become implicitly encoded into the optimized energy model. Because we are interested in resolving strong binders from weak ones with a finite dataset, our optimization amplifies the distinction between these two sets of binders by enlarging their energetic gap in the training process. In the thymus, however, it is likely that peptides still abide by the rule of typical protein interactions, and the thermo-stability of strong binders only tip the balance over the weak binders by several $k_{\mathrm{B}}T$. Other factors in the thymus, however, may further enhance the distinction between strong and weak binders by introducing cooperativity in the selection processes, such as quorum sensing [52] or kinetic proofreading [53, 54, 55]. By maximizing the z-score between strong and weak binders, RACER approximates these mechanisms by learning an effective binding energy which amplifies small difference in thermo-stability among candidate binders. This approximation, however, affects neither the identification of the strong binders of a specific TCR nor the subsequent ensemble study of peptide recognition, since only the order of binding affinities among individual TCR-p-MHC pairs matters for our results.

3.2 Structural information from available crystal structures adds higher-order interactions into the pairwise model

Our pairwise model offers a novel avenue for developing models that incorporate information contained in available protein structures. Recent studies support the importance of primary sequence higher-order correlations in describing biomolecular interactions [56]. Prior investigations have applied artificial neural networks for predicting strong binders of TCR [57] and p-MHC [58] molecules

based only on the primary sequences. Although deep learning can implicitly account for higher-order interactions, such approaches may still be limited by the available sequences that can be identified from experiments. By using a pairwise potential, RACER incorporates higher-order interactions by including existing crystal structures. The highly accurate prediction of RACER, in turn, demonstrates that such a structurally educated pairwise model is able to resolve the specificity of TCR-p-MHC interactions in a biological environment, justifying the linear constitutive assumption which sums up the binding energies of individual TCR-peptide residue pairs for quantifying the interactions between TCRs and peptides, utilized here and in prior theoretical analyses [17, 11]. Moreover, the predictive accuracy of RACER can be further improved by including additional strong binders from crystal structures that are deposited in the database (Fig. 4), thus providing a mechanism for additional refinement and improvements in predictive accuracy as more sequence and structural data become available.

3.3 Recognition of foreign and point-mutated self-peptides

Our approach, which utilizes structural information to assess binding, promises to simulate differences in the resulting T-cell repertoire, and hence recognition of TANs, across patients and cancer subtypes. Applying our model to $\text{CDR}_3\alpha$, β chains, obtained from T-cell sequencing, together with possible TAN lists generated by deep sequencing of cancer populations could provide a rapid method of generating clinically actionable information for cancer specific TCRs in the form of putative TCR-TAN pairs. While we focused our analysis on a single MHC restriction, an identical approach may be constructed using the crystal structure of another TCR-p-MHC pair, together with several known strong and weak binder candidates.

The relative efficacy of targeting TANs remains an important question with significant clinical

implications. We have shown that RACER can readily simulate full-scale thymic selection to produce a T-cell repertoire. RACER-derived donor alloreactivity rates fell on the lower end of experimentally obtained estimates. It is likely that additional differences between the pattern of expression in host and donor thymic environments, including differences in peptide pattern expression, contributes to variant allogenicity. The overall agreement in post-selection behavior between this study and our previous theoretical investigation is reassuring for both approaches. Taken together, these results imply that thymic selection affords little to no recognition protection of peptides closely related to self, thus supporting the notion that T-cells undergoing central tolerance to thymic self-peptides are virtually memorizing a list of antigens to avoid. Given that a large class of TANs are generated via point mutations in self-peptide, this result also provides a quantitative argument for the efficacy of immunotherapies which target point-mutated neoantigens.

Rapid assessment of pairwise TCR-p-MHC interactions enables a number of applications with immediate clinical merit, including optimized HSCT. Here, the balance of beneficial graft-versus-tumor and undesirable graft-versus-host disease is guided by previous empirical successes (and failures). Assessing the alloreactivity rates of donor immune cells prior to HSCT would generate substantial value for potential transplant recipients. Here we have provided proof-of-principle that this approach can handle sufficient interactions to simultaneously quantify potential alloreactivity alongside predicted transplant efficacy from identified host and donor TCRs and self-antigens. One limitation of transplant therapy is the reduced T-cell diversity of the post-transplant repertoire [59]. The RACER model not only promises to identify *a priori* putative TCR-peptide targets, but also the level of diversity required for adequate recognition of the target by the donor immune system.

While important, studying TCR-p-MHC pairwise interactions on the scale of an entire T-cell repertoire is only one factor influencing adaptive immune system recognition. Signaling between

other adaptive immune system elements (including helper T-cells and natural killer cells) and intracellular factors which influence antigen generation, abundance, and availability on the cell surface also affect recognition rates. Encouraged by the RACER model's reasonable selection and recognition behavior, we propose this optimized framework as the first of its kind tool for tackling general questions regarding the interactions between the T-cell repertoire and relevant antigen landscape. Although we calculate static antigen recognition probabilities, the temporal tumor-immune interaction leads to dynamic co-evolution [21] reliant on the quality, abundance, and systems-level signaling of antigens. In the setting of HSCT, the availability of time series assessments of immune cell repertoires, self-peptides, and tumor antigens promises to inform optimal treatment strategies based on the donor immune system and host cancer population.

4 Materials and Methods

4.1 Details of the Hamiltonian used in our optimization

To evaluate the binding energies on the basis of a structurally motivated molecular energy model, the framework of a coarse-grained protein energy model, AWSEM force field [32], was utilized for calculating the binding energies between the T-cell receptors (TCRs) and the peptide on top of a MHC molecule. AWSEM is a coarse-grained model with each residue described by the positions of its 3 atoms – Cα, Cβ and O atoms (except for glycine, which does not have Cβ atoms) [32]. We used the Cβ atom (except for glycine, where the Cα atom was used) of each residue to calculate inter-residue interactions. The original AWSEM Hamiltonian includes both bonded and non-bonded interactions.

$$V_{\text{total}} = V_{\text{bonded}} + V_{\text{nonbonded}} \tag{1}$$

Since those residue pairs that contribute to the TCR-peptide binding energy, specifically those from the CDR loops and peptides, are in separate protein chains, only non-bonded interactions are considered. $V_{\text{nonbonded}}$ is composed of three terms:

$$V_{\text{nobonded}} = V_{\text{pairwise}} + V_{\text{burial}} + V_{\text{database}} \tag{2}$$

Among them, V_{burial} is a one-body term describing the propensity of residues to be buried in or exposed on the surface of proteins. V_{database} is a protein sequence-specific term that uses information from existing protein database, such as secondary and tertiary interactions, to ensure locally accurate chemistry of protein structure. Since the TCR-p-MHC system features pairwise interactions between a TCR and its corresponding peptide, only the term V_{pairwise} is used for this study.

The pairwise Hamiltonian of AWSEM potential describes the interactions between any two non-bonded residues and can be further separated into two terms:

$$V_{\text{pairwise}} = V_{\text{direct}} + V_{\text{mediated}} \tag{3}$$

V_{direct} captures the direct protein-protein interaction of residues that are in between 4.5 and 6.5 Å. The functional form of V_{direct} is

$$V_{\text{direct}} = \sum^{N} \gamma_{ij}(a_i, a_j)\Theta^I_{ij} \tag{4}$$

in which $\Theta^I_{ij} = \frac{1}{4}(1 + \tanh[5.0 \cdot (r_{\text{ij}} - r^I_{\text{min}})])(1 + \tanh[5.0 \cdot (r^I_{\text{max}} - r_{\text{ij}})])$ is a switching function capturing the effective range of interactions between two residues (here taken between $r^I_{\text{min}} = 4.5\mathring{A}$ and $r^I_{\text{max}} = 6.5\mathring{A}$). Thus, two residues are defined to be "in contact" if their distance falls between

4.5 $\mathring{A}$ and 6.5 $\mathring{A}$. $\gamma_{ij}(a_i, a_j)$ describes the residue-type dependent interaction strength, and is the most important parameter that enters the optimization of RACER.

Similarly, V_{mediated} describes the longer range interactions of two residues separated between 6.5 and 9.5 Å. Depending on the local density of residue environment, V_{mediated} can be further divided into a protein-mediated term and a water-mediated term.

$$V_{\text{mediated}} = -\sum \Theta_{ij}^{II}(\sigma_{ij}^{wat}\gamma_{ij}^{wat}(a_i, a_j) + \sigma_{ij}^{prot}\gamma_{ij}^{prot}(a_i, a_j)) \tag{5}$$

where $\sigma_{ij}^{wat} = \frac{1}{4}(1 - \tanh[7.0 \cdot (\rho_i - 2.6)])(1 - \tanh[7.0 \cdot (\rho_j - 2.6)])$ and $\sigma_{ij}^{prot} = 1 - \sigma_{ij}^{wat}$ are switching functions indicating the local environment based on the density of each residue. $\Theta_{ij}^{II} = \frac{1}{4}(1+\tanh[5.0\cdot(r_{\text{ij}} - r_{\text{min}}^{II})])(1+\tanh[5.0\cdot(r_{\text{max}}^{II} - r_{\text{ij}})])$ with $r_{\text{min}}^{II} = 6.5\mathring{A}$ and $r_{\text{max}}^{II} = 9.5\mathring{A}$. By optimizing $\gamma_{ij}^{wat}(a_i, a_j)$, $\gamma_{ij}^{prot}(a_i, a_j)$ together with $\gamma_{ij}(a_i, a_j)$, V_{mediated} ensures a more accurate description of long-range interaction between two non-bonded residues, but increases computational expense when evaluating the binding energy in between TCR and peptides by more than 5-fold compared with V_{direct}. Since the use of V_{direct} sufficiently separates strong binders from weak ones, only V_{direct} is employed for calculating the binding energies throughout our manuscript, for computational efficiency in studying the TCR repertoire.

4.2 Optimization of RACER to maximize specificity of TCR-peptide recognition

For each interaction type, the $\gamma_{ij}(a_i, a_j)$ parameters constitute a 20-by-20 matrix of parameters that describes the pairwise interaction between any two residues i, j, each with one of the 20 residue types, a_i, a_j. Guided by the principle of minimum frustration [34], $\gamma_{ij}(a_i, a_j)$ was previously op-

timized self-consistently to best separate the folded states from the misfolded states of proteins. Distilled into the mathematical details, the energy model was optimized to maximize the functional $\delta E/\Delta E$, where δE is the energy gap between the folded and misfolded proteins, and ΔE measures the standard deviation of the energies of the misfolded states. An energy model was optimized based on a pool of selected protein structures [60], where a series of decoy structures were generated by either threading the sequences along the existing crystal structures, or by biasing the proteins into molten globule structures using MD simulations [33]. The resultant γ parameter thus entails an energy model that facilitates the folding of proteins with given sequences.

Motivated by this idea, RACER was parametrized to maximize the z-scores for fully separating TCR strong binders from weak ones. Strong binders were chosen to be those top peptides that survive and enriched by more than 50 copies after four rounds of experimental deep sequencing processes (details in Section Data used in our analyses) [26], together with the peptides present in the deposited crystal structures [39]. The decoy sequences were generated by randomizing the non-anchoring residues of each strong binder for 1000 copies, and excludes the strong-binder sequences. The γ parameters were then optimized to maximize the stability gap between strong and randomized set of decoy binders, $\delta E = A\gamma$, and the standard deviation of decoy energies, $\Delta E = \gamma B \gamma$, where the matrix A and B are defined as:

$$\begin{aligned} A_i &= \langle \phi_i \rangle_{\mathrm{wb}} - \phi_{\mathrm{sb}} \\ B_{i,j} &= \langle \phi_i \phi_j \rangle_{\mathrm{wb}} - \langle \phi_i \rangle_{\mathrm{wb}} \langle \phi_j \rangle_{\mathrm{wb}} \end{aligned} \tag{6}$$

In the above Eq. 6, ϕ_i is the functional form for each interaction type, either V_{direct} or V_{mediated}. ϕ_i also summaries the probability of contacts formation (contact map) between pairs of amino acids in a specific TCR-peptide system. The subscripts “wb” stands for “weak binders” and “sb” stands

for "strong binders". The optimization of $\delta E/\Delta E = A\gamma/\sqrt{\gamma B\gamma}$ can be performed effectively by maximizing the functional objective $R = A\gamma - \lambda_1\sqrt{\gamma B\gamma}$, where the Lagarange multiplier λ_1 sets the energy scale. The solution of this optimization gives $\gamma \propto B^{-1}A$. In the practice of protein folding, this optimization was performed in an iterative way where the optimized parameters were used for generating a new set of decoy protein structures. In this study, since different peptides are structurally degenerate on top of MHC as observed from experiments [26], only one round of optimization was performed.

4.3 Data input used in our analyses

A deep-sequencing technique was developed to assess the binding affinity of a diverse repertoire of MHC-presented peptides towards a certain type of TCR [26]. Specifically, 3 types of TCRs: 2B4, 5CC7 and 226, were used for selecting peptides upon four rounds of purification. The peptides that survive and enriched with multiple copies bind strongly with the corresponding TCR. In contrast, the peptides present initially but become extinct during purification represent experimentally determined weak binders. For each of the 3 TCRs, the peptides that end up with more than 50 copies after the purification process, together with the peptides presented in the crystal structures, were selected as strong binders. 1000 decoy sequences were generated for each of the strong binders by randomizing the non-anchoring residues. Both strong binders and decoys were included in the training set. In addition, to test the performance of RACER, peptides having at least 7 copies initially but disappearing during purification were selected as experimentally determined weak binders and were assigned to the test set for each TCR. To test the transferability of the model, we used weak-binding peptides of two different TCRs (e.g., 5CC7 and 226) as additional test sets distinct from the TCR used in training (e.g., 2B4).

When structural data for a specific TCR-peptide pair of interest is unavailable, we built the structure by homology modeling [61], based on a known TCR-peptide crystal structure incorporating the same TCR. Since potential steric clashes after switching peptide sequences may disfavor the strong binders used in our training set, we used Modeller [61] to relax structures of strong binders before including them in the training process. Likewise, the binding energies of the experimentally determined weak binders were also evaluated after structural relaxation. The structural relaxation adds several seconds of computational time for each TCR-peptide pair, and thus poses a challenge for large scale repertoire analysis. However, the coarse-grained nature of RACER framework may significantly reduce the probability of side-chain clashes after switching peptide sequences. To test the accuracy of our model prediction without structural relaxation, we calculated the binding energies of strong and weak binders of TCR 2B4 by only switching the peptide sequences, omitting any structural adjustment. Our result (Fig. S2) shows comparable accuracy in separating strong from weak binders, similar to that reported in Fig. 2. Encouraged by the accuracy of our coarse-grained model without relaxation, we modeled large pairwise collections of TCR-peptide interactions by only altering their corresponding sequences.

For blind assessment of TCR specificity, we ask whether we can improve prediction accuracy if there are available strong binders determined in crystal structures of the target TCRs. To test this, we added contact maps calculated from the crystal structures of the other two TCRs as two additional strong binders in the training set. For example, in the case of TCR 2B4, the contact maps from the crystal structures of TCR 5CC7 and 226 were added into the training set of TCR 2B4, constituting a total of 46 strong binders. The test shows a significant improvement in predicting the binding specificity of TCR 5CC7 and 226 (Fig. 4b).

4.4 The leave-one-out cross validation

The Leave-one-out cross validation (LOOCV) was used to test the predictive power of RACER on its ability to identify strong binders. Specifically, one of the 44 strong binders of the TCR 2B4 was removed from the training set, and its predicted binding energy E_{pred} was compared with the experimentally determined weak binders. If the median of the weak binders is larger than E_{pred} (a larger binding energy is associated with smaller affinity), the testing strong binder is successfully identified. The performance of RACER is compared with that from the clustering of peptide sequences using the algorithm from CD-Hit [62] (See SI for details).

4.5 Comparing the correlation of binding energies with the K_{d} from SPR experiments

As another independent experiment, surface plasmon resonance (SPR) was performed to assess the binding affinities of the three TCRs towards 9 selected peptides [26]. The correlation between the predicted binding energies from RACER and the dissociation constant K_{d} evaluated from the SPR experiments thus constitutes a separate set of tests for the accuracy of RACER. The Pearson correlation coefficient and the Spearman's rank correlation coefficient between $k_{\mathrm{B}}T\log(K_{\mathrm{d}})$ and predicted binding energies were used to quantify this correlation.

4.6 Evaluation of different TCR-pMHC interactions used for statistical study

In order to assess the statistical behavior of the inferential model, we calculated the pairwise binding interactions between a simulated T-cell population of size $N_t = 10^5$ and $N_n = 10^4$ thymic self-peptides. The binding energies from optimized energy model calculations leaves a molecular

multiplier as a free parameter. Since only the relative rank order of binding affinities matters for the following statistical study, for ease of comparison to prior models, the inverse of binding energies were taken to be representative of TCR-p-MHC binding affinity, and these values were shifted and scaled to the interval $[0, 10]$. Each T-cell, by definition, survives if the maximal interaction over all self-peptides does not exceed some upper threshold. Selection thresholds were chosen to achieve 50% [23], and post-selection T-cells were quantified for their ability to recognize random non-self-antigens and tumor neoantigens that differ from one of the N_n thymic self peptides by one residue. 50% selection of TCRs result in approximately $5 \cdot 10^5$ surviving, for which pairwise interactions are generated against 10^3 random and 10^4 point-mutated self-peptides, representing foreign and tumor-associated neoantigens, respectively (randomly generated peptides were checked to ensure non-membership in the set of thymic self-peptides). Estimates of individual TCR recognition probability were calculated by averaging the $5 \cdot 10^4$-by-10^3 indicator matrix of binders vs. non-binders. Estimates of TCR repertoire recognition probability was calculated by assessing the 1-by-10^3 vector indicating the presence or absence of at least 1 recognizing TCR. The post-selection individual and repertoire T-cell recognition probabilities of random and point-mutant antigens were then compared with previously derived analytic results for a random energy model [17]. Potency refers to a measure of the number of unique TCRs recognized by a given self-peptide. For this, we find the most potent self-peptide, which is the one that eliminates the largest number of the approximately $5 \cdot 10^4$ TCRs in selection. The next most potent self-peptide recognizes the largest number of remaining TCRs not recognized by the first. This procedure iteratively defines a relative ordering of self-peptides along with a potency index given by the number of recognized TCRs.

5 Acknowledgments

The authors would like to thank Dr. Michael E. Birnbaum for fruitful discussion on systems-level TCR-antigen specificity. This work is supported by the Center for Theoretical Biological Physics sponsored by the National Science Foundation (NSF) Grant NSF PHY-1427654. JNO was also supported by NSF CHE-1614101. HL was also supported by the NSF grant PHY-1935762. JTG was supported by National Cancer Institute of NIH (F30CA213878). JNO is a CPRIT Scholar in Cancer Research sponsored by the Cancer Prevention and Research Institute of Texas.

References

[1] J. Couzin-Frankel, "Cancer Immunotherapy," Science, vol. 342, pp. 1432–1433, Dec. 2013.

[2] D. R. Leach, M. F. Krummel, and J. P. Allison, "Enhancement of antitumor immunity by CTLA-4 blockade," Science (New York, N.Y.), vol. 271, pp. 1734–1736, Mar. 1996.

[3] H. O. Alsaab, S. Sau, R. Alzhrani, K. Tatiparti, K. Bhise, S. K. Kashaw, and A. K. Iyer, "PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome," Frontiers in Pharmacology, vol. 8, p. 561, 2017.

[4] M. Sadelain, I. Rivire, and S. Riddell, "Therapeutic T cell engineering," Nature, vol. 545, no. 7655, pp. 423–431, 2017.

[5] P. A. Ott, Z. Hu, D. B. Keskin, S. A. Shukla, J. Sun, D. J. Bozym, W. Zhang, A. Luoma, A. Giobbie-Hurder, L. Peter, C. Chen, O. Olive, T. A. Carter, S. Li, D. J. Lieb, T. Eisenhaure, E. Gjini, J. Stevens, W. J. Lane, I. Javeri, K. Nellaiappan, A. M. Salazar, H. Daley, M. Seaman, E. I. Buchbinder, C. H. Yoon, M. Harden, N. Lennon, S. Gabriel, S. J. Rodig, D. H.

Barouch, J. C. Aster, G. Getz, K. Wucherpfennig, D. Neuberg, J. Ritz, E. S. Lander, E. F. Fritsch, N. Hacohen, and C. J. Wu, "An immunogenic personal neoantigen vaccine for patients with melanoma," Nature, vol. 547, no. 7662, pp. 217–221, 2017.

[6] P. Johansen, T. Storni, L. Rettig, Z. Qiu, A. Der-Sarkissian, K. A. Smith, V. Manolova, K. S. Lang, G. Senti, B. Mllhaupt, T. Gerlach, R. F. Speck, A. Bot, and T. M. Kndig, "Antigen kinetics determines immune reactivity," Proceedings of the National Academy of Sciences, vol. 105, pp. 5189–5194, Apr. 2008.

[7] J. J. Molldrem, K. Komanduri, and E. Wieder, "Overexpressed differentiation antigens as targets of graft-versus-leukemia reactions," Current Opinion in Hematology, vol. 9, pp. 503–508, Nov. 2002.

[8] A. K. Abbas, A. K. Abbas, A. H. Lichtman, and S. Pillai, Cellular and molecular immunology. 2018.

[9] H. von Boehmer, "Thymic selection: a matter of life and death," Immunology Today, vol. 13, pp. 454–458, Nov. 1992.

[10] G. J. Nossal, "Negative selection of lymphocytes," Cell, vol. 76, pp. 229–239, Jan. 1994.

[11] A. Kosmrlj, A. K. Jha, E. S. Huseby, M. Kardar, and A. K. Chakraborty, "How the thymus designs antigen-specific and self-tolerant T cell receptor sequences," Proceedings of the National Academy of Sciences, vol. 105, pp. 16671–16676, Oct. 2008.

[12] Y. Elhanati, A. Murugan, C. G. Callan, T. Mora, and A. M. Walczak, "Quantifying selection in immune receptor repertoires," Proceedings of the National Academy of Sciences of the United States of America, vol. 111, pp. 9875–9880, July 2014.

[13] J. Ishizuka, K. Grebe, E. Shenderov, B. Peters, Q. Chen, Y. Peng, L. Wang, T. Dong, V. Pasquetto, C. Oseroff, and others, "Quantitating T cell cross-reactivity for unrelated peptide antigens," The Journal of Immunology, vol. 183, no. 7, pp. 4337–4345, 2009. Publisher: Am Assoc Immnol.

[14] M. M. Davis, "Not-So-Negative Selection," Immunity, vol. 43, pp. 833–835, Nov. 2015.

[15] C. F. Arias, M. A. Herrero, J. A. Cuesta, F. J. Acosta, and C. Fernndez-Arias, "The growth threshold conjecture: a theoretical framework for understanding T-cell tolerance," Royal Society Open Science, vol. 2, p. 150016, July 2015.

[16] R. J. De Boer and A. S. Perelson, "How diverse should the immune system be?," Proceedings. Biological Sciences, vol. 252, pp. 171–175, June 1993.

[17] J. T. George, D. A. Kessler, and H. Levine, "Effects of thymic selection on T cell recognition of foreign and tumor antigenic peptides," Proceedings of the National Academy of Sciences of the United States of America, vol. 114, no. 38, pp. E7875–E7881, 2017.

[18] A. Mayer, V. Balasubramanian, A. M. Walczak, and T. Mora, "How a well-adapting immune system remembers," Proceedings of the National Academy of Sciences, vol. 116, pp. 8815–8823, Apr. 2019.

[19] G. Altan-Bonnet, T. Mora, and A. M. Walczak, "Quantitative immunology for physicists," Physics Reports, 2020. Publisher: Elsevier.

[20] J. T. George and H. Levine, "Stochastic modeling of tumor progression and immune evasion," Journal of Theoretical Biology, vol. 458, pp. 148–155, 2018.

[21] J. T. George and H. Levine, "Sustained coevolution in a stochastic model of cancer–immune interaction," Cancer Research, vol. 80, no. 4, pp. 811–819, 2020.

[22] V. I. Zarnitsyna, B. D. Evavold, L. N. Schoettle, J. N. Blattman, and R. Antia, "Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire," Frontiers in Immunology, vol. 4, p. 485, Dec. 2013.

[23] A. J. Yates, "Theories and Quantification of Thymic Selection," Frontiers in Immunology, vol. 5, 2014.

[24] T. P. Riley, L. M. Hellman, M. H. Gee, J. L. Mendoza, J. A. Alonso, K. C. Foley, M. I. Nishimura, C. W. Vander Kooi, K. C. Garcia, and B. M. Baker, "T cell receptor cross-reactivity expanded by dramatic peptideMHC adaptability," Nature Chemical Biology, vol. 14, pp. 934–942, Oct. 2018.

[25] N. K. Singh, T. P. Riley, S. C. B. Baker, T. Borrman, Z. Weng, and B. M. Baker, "Emerging Concepts in TCR Specificity: Rationalizing and (Maybe) Predicting Outcomes," Journal of Immunology (Baltimore, Md.: 1950), vol. 199, no. 7, pp. 2203–2213, 2017.

[26] M. E. Birnbaum, J. L. Mendoza, D. K. Sethi, S. Dong, J. Glanville, J. Dobbins, E. zkan, M. M. Davis, K. W. Wucherpfennig, and K. C. Garcia, "Deconstructing the Peptide-MHC Specificity of T Cell Recognition," Cell, vol. 157, pp. 1073–1087, May 2014.

[27] P. Dash, A. J. Fiore-Gartland, T. Hertz, G. C. Wang, S. Sharma, A. Souquette, J. C. Crawford, E. B. Clemens, T. H. O. Nguyen, K. Kedzierska, N. L. La Gruta, P. Bradley, and P. G. Thomas, "Quantifiable predictive features define epitope-specific T cell receptor repertoires," Nature, vol. 547, no. 7661, pp. 89–93, 2017.

[28] R. Gowthaman and B. G. Pierce, "TCRmodel: high resolution modeling of T cell receptors from sequence," Nucleic Acids Research, vol. 46, pp. W396–W401, July 2018.

[29] B. G. Pierce and Z. Weng, "A flexible docking approach for prediction of T cell receptor-peptide-MHC complexes," Protein Science: A Publication of the Protein Society, vol. 22, pp. 35–46, Jan. 2013.

[30] R. A. Goldstein, Z. A. Luthey-Schulten, and P. G. Wolynes, "Protein tertiary structure recognition using optimized Hamiltonians with local interactions.," Proceedings of the National Academy of Sciences, vol. 89, pp. 9029–9033, Oct. 1992.

[31] C. Clementi, H. Nymeyer, and J. N. Onuchic, "Topological and energetic factors: what determines the structural details of the transition state ensemble and en-route intermediates for protein folding? an investigation for small globular proteins," Journal of Molecular Biology, vol. 298, pp. 937–953, May 2000.

[32] A. Davtyan, N. P. Schafer, W. Zheng, C. Clementi, P. G. Wolynes, and G. A. Papoian, "AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing," The Journal of Physical Chemistry B, vol. 116, pp. 8494–8503, July 2012.

[33] N. P. Schafer, B. L. Kim, W. Zheng, and P. G. Wolynes, "Learning To Fold Proteins Using Energy Landscape Theory," Israel Journal of Chemistry, vol. 54, pp. 1311–1337, Aug. 2014.

[34] J. D. Bryngelson and P. G. Wolynes, "Spin glasses and the statistical mechanics of protein folding.," Proceedings of the National Academy of Sciences, vol. 84, pp. 7524–7528, Nov. 1987.

[35] P. E. Leopold, M. Montal, and J. N. Onuchic, "Protein folding funnels: a kinetic approach to the sequence-structure relationship.," Proceedings of the National Academy of Sciences, vol. 89, pp. 8721–8725, Sept. 1992.

[36] V. I. Abkevich, A. M. Gutin, and E. I. Shakhnovich, "Improved design of stable and fast-folding model proteins," Fold Des, vol. 1, no. 3, pp. 221–230, 1996.

[37] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, "THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective," Annual Review of Physical Chemistry, vol. 48, pp. 545–600, Oct. 1997.

[38] J. Wang and G. M. Verkhivker, "Energy Landscape Theory, Funnels, Specificity, and Optimal Criterion of Biomolecular Binding," Physical Review Letters, vol. 90, May 2003.

[39] E. W. Newell, L. K. Ely, A. C. Kruse, P. A. Reay, S. N. Rodriguez, A. E. Lin, M. S. Kuhns, K. C. Garcia, and M. M. Davis, "Structural Basis of Specificity and Cross-Reactivity in T Cell Receptors Specific for Cytochrome *c* I-E ^{k}\$," The Journal of Immunology, vol. 186, pp. 5823–5832, May 2011.

[40] S. Miyazawa and R. L. Jernigan, "Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation," Macromolecules, vol. 18, pp. 534–552, May 1985.

[41] C. Sinclair, I. Bains, A. J. Yates, and B. Seddon, "Asymmetric thymocyte death underlies the CD4:CD8 T-cell ratio in the adaptive immune system," Proceedings of the National Academy of Sciences, vol. 110, pp. E2905–E2914, July 2013.

[42] L. Ignatowicz, W. Rees, R. Pacholczyk, H. Ignatowicz, E. Kushnir, J. Kappler, and P. Marrack, "T cells can be activated by peptides that are unrelated in sequence to their selecting peptide," Immunity, vol. 7, pp. 179–186, Aug. 1997.

[43] J. Zerrahn, W. Held, and D. H. Raulet, "The MHC reactivity of the T cell repertoire prior to positive and negative selection," Cell, vol. 88, pp. 627–636, Mar. 1997.

[44] E. Strnen, M. Toebes, S. Kelderman, M. M. van Buuren, W. Yang, N. van Rooij, M. Donia, M.-L. Bschen, F. Lund-Johansen, J. Olweus, and T. N. Schumacher, "Targeting of cancer neoantigens with donor-derived T cell receptor repertoires," Science (New York, N.Y.), vol. 352, pp. 1337–1341, June 2016.

[45] S. R. Riddell, C. Berger, M. Murata, S. Randolph, and E. H. Warren, "The graft versus leukemia response after allogeneic hematopoietic stem cell transplantation," Blood reviews, vol. 17, no. 3, pp. 153–162, 2003. Publisher: Elsevier.

[46] M. Bevan, R. Langman, and M. Cohn, "H-2 antigen-specific cytotoxic T cells induced by concanavalin A: estimation of their relative frequency," European journal of immunology, vol. 6, no. 3, pp. 150–156, 1976. Publisher: Wiley Online Library.

[47] J. Ashwell, C. Chen, and R. Schwartz, "High frequency and nonrandom distribution of alloreactivity in T cell clones selected for recognition of foreign antigen in association with self class II molecules.," The Journal of Immunology, vol. 136, no. 2, pp. 389–395, 1986. Publisher: Am Assoc Immnol.

[48] E. J. Suchin, P. B. Langmuir, E. Palmer, M. H. Sayegh, A. D. Wells, and L. A. Turka, "Quantifying the frequency of alloreactive T cells in vivo: new answers to an old question," The Journal of Immunology, vol. 166, no. 2, pp. 973–981, 2001. Publisher: Am Assoc Immnol.

[49] J. J. Melenhorst, P. Scheinberg, A. Williams, D. R. Ambrozak, K. Keyvanfar, M. Smith, J. P. McCoy Jr, N. F. Hensel, D. C. Douek, and A. J. Barrett, "Alloreactivity across HLA barriers is mediated by both naive and antigen-experienced T cells," Biology of Blood and Marrow Transplantation, vol. 17, no. 6, pp. 800–809, 2011. Publisher: Elsevier.

[50] W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," Journal of Molecular Graphics, vol. 14, pp. 33–38, Feb. 1996.

[51] L. H. Kapcha and P. J. Rossky, "A Simple Atomic-Level Hydrophobicity Scale Reveals Protein Interfacial Structure," Journal of Molecular Biology, vol. 426, pp. 484–498, Jan. 2014.

[52] T. C. Butler, M. Kardar, and A. K. Chakraborty, "Quorum sensing allows T cells to discriminate between self and nonself," Proceedings of the National Academy of Sciences, vol. 110, pp. 11833–11838, July 2013.

[53] J. J. Hopfield, "Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity," Proceedings of the National Academy of Sciences, vol. 71, pp. 4135–4139, Oct. 1974.

[54] T. W. McKeithan, "Kinetic proofreading in T-cell receptor signal transduction.," Proceedings of the National Academy of Sciences, vol. 92, pp. 5042–5046, May 1995.

[55] O. Stepanek, A. S. Prabhakar, C. Osswald, C. G. King, A. Bulek, D. Naeher, M. Beaufils-Hugot, M. L. Abanto, V. Galati, B. Hausmann, R. Lang, D. K. Cole, E. S. Huseby, A. K. Sewell,

A. K. Chakraborty, and E. Palmer, "Coreceptor Scanning by the T Cell Receptor Provides a Mechanism for T Cell Tolerance," Cell, vol. 159, pp. 333–345, Oct. 2014.

[56] A. J. Riesselman, J. B. Ingraham, and D. S. Marks, "Deep generative models of genetic variation capture the effects of mutations," Nature Methods, vol. 15, pp. 816–822, Oct. 2018.

[57] V. I. Jurtz, L. E. Jessen, A. K. Bentzen, M. C. Jespersen, S. Mahajan, R. Vita, K. K. Jensen, P. Marcatili, S. R. Hadrup, B. Peters, and M. Nielsen, "NetTCR: sequence-based prediction of TCR binding to peptide-MHC complexes using convolutional neural networks," bioRxiv, Oct. 2018.

[58] M. Nielsen, C. Lundegaard, P. Worning, S. L. Lauemller, K. Lamberth, S. Buus, S. Brunak, and O. Lund, "Reliable prediction of T-cell epitopes using neural networks with novel sequence representations," Protein Science, vol. 12, pp. 1007–1017, May 2003.

[59] J. W. Van Heijst, I. Ceberio, L. B. Lipuma, D. W. Samilo, G. D. Wasilewski, A. M. R. Gonzales, J. L. Nieves, M. R. Van Den Brink, M. A. Perales, and E. G. Pamer, "Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation," Nature medicine, vol. 19, no. 3, p. 372, 2013. Publisher: Nature Publishing Group.

[60] G. A. Papoian, J. Ulander, M. P. Eastwood, Z. Luthey-Schulten, and P. G. Wolynes, "From The Cover: Water in protein structure prediction," Proceedings of the National Academy of Sciences, vol. 101, pp. 3352–3357, Mar. 2004.

[61] B. Webb and A. Sali, "Comparative Protein Structure Modeling Using MODELLER," Current Protocols in Bioinformatics, vol. 54, June 2016.

[62] L. Fu, B. Niu, Z. Zhu, S. Wu, and W. Li, "CD-HIT: accelerated for clustering the next-generation sequencing data," Bioinformatics, vol. 28, pp. 3150–3152, Dec. 2012.