frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/fimmu.2022.1080596

OPEN ACCESS

EDITED BY
Daniel López,
Carlos III Health Institute (ISCIII), Spain

REVIEWED BY
Brian M Baker,
University of Notre Dame,
United States
Anastas Dimitrov Pashov,
Bulgarian Academy of Sciences (BAS),
Bulgaria

*CORRESPONDENCE
Charlotte M. Deane
deane@stats.ox.ac.uk

SPECIALTY SECTION
This article was submitted to Vaccines and Molecular Therapeutics, a section of the journal Frontiers in Immunology

RECEIVED 26 October 2022
ACCEPTED 07 December 2022
PUBLISHED 09 January 2023

CITATION
Raybould MIJ, Nissley DA, Kumar S and Deane CM (2023) Computationally profiling peptide: MHC recognition by T-cell receptors and T-cell receptor-mimetic antibodies.
Front. Immunol. 13:1080596.
doi: 10.3389/fimmu.2022.1080596

Computationally profiling peptide:MHC recognition by T-cell receptors and T-cell receptor-mimetic antibodies

Matthew I. J. Raybould[1], Daniel A. Nissley[1], Sandeep Kumar[2] and Charlotte M. Deane[1]*

[1]Oxford Protein Informatics Group, Department of Statistics, University of Oxford, Oxford, United Kingdom, [2]Biotherapeutics Discovery, Boehringer Ingelheim, Ridgefield, CT, United States

T-cell receptor-mimetic antibodies (TCRms) targeting disease-associated peptides presented by Major Histocompatibility Complexes (pMHCs) are set to become a major new drug modality. However, we lack a general understanding of how TCRms engage pMHC targets, which is crucial for predicting their specificity and safety. Several new structures of TCRm:pMHC complexes have become available in the past year, providing sufficient initial data for a holistic analysis of TCRms as a class of pMHC binding agents. Here, we profile the complete set of TCRm:pMHC complexes against representative TCR:pMHC complexes to quantify the TCR-likeness of their pMHC engagement. We find that intrinsic molecular differences between antibodies and TCRs lead to fundamentally different roles for their heavy/light chains and Complementarity-Determining Region loops during antigen recognition. The idiotypic properties of antibodies may increase the likelihood of TCRms engaging pMHCs with less peptide selectivity than TCRs. However, the pMHC recognition features of some TCRms, including the two TCRms currently in clinical trials, can be remarkably TCR-like. The insights gained from this study will aid in the rational design and optimisation of next-generation TCRms.

KEYWORDS

TCR, mimetic, antibody, peptide, MHC, HLA, TCR-likeness, structural biology

1 Introduction

The human adaptive immune system relies upon B-cells and T-cells that use characteristic membrane-bound immunoglobulins, B-cell receptors (BCRs) and T-cell receptors (TCRs), to recognise a broad range of pathogenic antigens, many of which are proteinaceous. BCRs, and their secreted soluble analogues, antibodies, recognise

complete soluble or membrane-bound extracellular proteins. T-cell receptors (TCRs), meanwhile, are focused through thymic development to recognise fragments of intracellularly- or extracellularly-derived peptides presented on cell surfaces by either a class I or class II polymorphic major histocompatibility complex (pMHCs) (1).

Despite their different natural roles, the binding domains of antibodies and TCRs bear several commonalities. They are both comprised of two analogously gene-recombined chains (termed 'heavy/light' (H/L) and 'beta/alpha' (B/A) for antibodies and TCRs respectively) and six complementarity-determining region (CDR) loops that together constitute most of their binding sites. These similarities have long motivated efforts to understand whether antibodies can engage pMHCs with TCR-like specificity (2). 'TCR-mimetic antibodies' (TCRms) that specifically recognise fragments of the intracellular proteome could offer pinpoint recognition of aberrant cells, transforming immunohistochemistry and immunotherapy.

TCRms also offer a number of practical advantages over TCRs in terms of soluble drug development. TCR:pMHC binding affinities lie in the 1-100M range (3, 4), meaning they must be affinity-engineered for use as a monovalent binding arm that recognises low copy number pMHCs (5). By contrast, antibody:antigen monovalent binding frequently occurs at the required range of affinities for therapeutic effect (low nM-pM) (6, 7). Therapeutic antibody development pipelines are also more established than their TCR equivalents, facilitating TCRm clinical translation and adaptation to multispecific formats exploiting proven cancer immune-modulation and T-cell redirection strategies (8–11).

Early studies sought to elicit natural TCRms *via* allogenic mouse immunisation and established that BCRs can be raised against non-self peptide:non-self MHC complexes. Though most antibodies were able to engage the MHC regardless of presented peptide, a smaller fraction were peptide-dependent (*i.e.* at least somewhat TCR-mimetic) (12–15).

TCRm isolation strategies shifted towards the use of *in vitro* phage-display libraries (14, 16), both sidestepping the need to account for self-tolerance and enabling rounds of positive and negative selection to enrich for stronger, more peptide-dependent binding. By 2020, these libraries had produced a variety of TCRms against a wide range of both class I and class II pMHC targets (15, 17–21), although the extent of their peptide specificity, and thus the breadth of their applicability, was still highly variable. Most have been used as chemical probes, for which the required peptide specificity is lower than that required of a therapeutic administered across heterogeneous (MHC-compatible) populations.

Experimental peptidome binding assays performed on three early-generation TCRms (17, 19, 20) suggested that they were unlikely to achieve the levels of specificity required for therapeutic applications (22, 23). However, several TCRms with apparently high specificity have been reported in the past year, fostering renewed interest in this therapeutic modality and resulting in a more than doubling of the number of crystal structures of TCRm:pMHC complexes (24–29). These include an anti-Wilms' Tumor Antigen 1 (WT1) TCRm (24) and an anti-alpha ferroprotein (AFP) TCRm (29) that have both progressed to clinical trials. Two neoantigen peptide:MHC-specific TCRms were also identified that achieve complete selectivity over their wildtype peptide equivalents each differing by just a single residue mutation (25, 26).

Here, we harness this recent increase in structural data on TCRm:pMHC complexes to computationally dissect their molecular recognition properties. We outline where and why TCRm:pMHC binding features tend to align with and differ from representative antibody:antigen and TCR:pMHC complexes. High-throughput interaction profiling of static complexes reveals that molecular differences between antibodies and TCRs result in differential CDR involvement in the pMHC binding event. This tends to lead to more variable peptide sensitivity, but does not preclude some TCRms from recognising pMHCs with similar features to those seen across TCRs. We also perform all-atom simulations which reveal that energetic hotspots in the MHC can play a key role in TCRm binding. TCRs seem to avoid this behavior, instead reliably exploiting energy hotspots on the peptide surface. Finally, we highlight TCR-like pMHC recognition features in the first TCRms to achieve sufficient specificity to reach the clinic [11D06 (24) and AFP-TCRm (29)] *versus* a TCRm with several known off-targets (ESK1). Overall, our analysis begins to quantify TCR-likeness across TCRms, enabling rational TCRm selection, optimisation, and design based on the natural cognate partners of pMHCs.

2 Results

We began our analysis by identifying sets of representative antibody:antigen, TCR:pMHC and TCRm:pMHC complexes from SAbDab (6, 7) and STCRDab (4) (see Methods).

We found twelve pMHC-binding antibodies, eleven of which have binding modes that transect the peptide binding groove (17–20, 24–27) and one alloantibody (30) that engages only the MHC and so was not classified as a TCRm. After filtering for interface redundancy (see Methods), we identified 10 representative TCRm:pMHC complexes (9 MHC class I, 1 MHC class II), 60 representative TCR:pMHC complexes (52 MHC class I, 8 MHC class II), and 824 representative antibody:antigen complexes. Unless otherwise stated, properties of TCRs or TCRms engaging MHC class I and class II are pooled as they are both engaged by the same genetic class of TCR (αβ).

2.1 pMHC complexes offer an unusually broad binding surface that is engaged differently by TCR and TCRm CDRs

To quantitatively compare the properties of immunoglobulin:antigen interfaces, we computed the buried surface area (BSA) and

patterns of formal interactions across our representative complexes (see Methods, Supplementary Methods).

2.1.1 Global interface properties

Calculating BSA over the whole immunoglobulin:antigen interface (Figure 1A) reveals that pMHC binding events result in atypically broad interfaces relative to general antigen complexes (TCRs : 1852.6Å^2, sd: 243.8Å^2 and TCRms : 2015.7Å^2, sd: 183.9Å^2; *versus* general antibodies : 1496.7Å^2, sd: 468.7Å^2).

The flat pMHC topology appears to place constraints on the CDRH3/CDRB3 length of their cognate immunoglobulins. Though assembled *via* a common VDJ recombination mechanism, TCR CDRB3s only have a length range of 10-17 (31), while antibody CDRH3s span lengths of 5-30+ (32–34). The TCRm CDRH3s range from length 10 to length 16, biased to the lower end of the range sampled in natural antibodies, with a mean value closer to that seen across TCRs, and entirely within the relatively narrow band of TCR CDRB3 or CDRA3 loop lengths (31, 35) (Supplementary Table 2). This suggests that shorter CDR3 lengths render TCRs/TCRms unable to achieve sufficient interactions to bind a pMHC, while longer lengths may result in destabilising clashes with the pMHC surface.

A residue-level interaction analysis of the immunoglobulin: pMHC complexes shows that the broad interface comprises different total numbers of interactions in TCRs and TCRms (μ: 23.2, sd: 6.2; μ: 27.0, sd: 3.4 respectively), however both classes of immunoglobulin use a similar balance of hydrophobic, aromatic, and polar interactions (Supplementary Table 1).

2.1.2 Binding properties by immunoglobulin region

Only considering the immunoglobulin contribution to the BSA, the TCRm profile (μ: 1015.8Å^2, sd: 111.0Å^2) is again closer to typical TCRs (μ: 934.4Å^2, sd: 127.2Å^2) than typical antibodies (μ: 707.5Å^2, sd: 260.0Å^2) (Figure 1B).

However, dissecting this BSA by CDR contributions demonstrates that antibody and TCR CDR loops play different roles in antigen binding. For example, antibody CDRL2 loops lie unburied in over 50% of general antigen complexes while TCR CDRA2 loops are buried in the pMHC interface to a much greater extent, a greater proportion of the time (Figure 1C). Equally, when considering the relative contributions of CDR3s to pMHC recognition, we find that TCRm binding tends to be biased towards burial of CDRH3 and away from CDRL3, more typical of general antibody:antigen complexes (Figure 1D), while TCRs exploit their CDR3 loops more evenly, if anything with a slight bias towards the CDRA3 loop (the genetic equivalent to CDRL3).

The differences in CDR usage in pMHC binding can be related to the fact that some CDR loops have markedly different length preferences in antibodies than TCRs. For example, antibody CDRH2s and CDRL2s have median IMGT lengths of 8 and 3, respectively (32) In contrast, TCR CDRB2 and CDRA2 loops have median lengths of 6 and 5 (35). Similarly, while antibody CDRH3 and CDRL3 loops have a median length of 15 and 9, TCR CDRB3 and CDRA3 loops have median lengths of 12 and 13 (31, 32, 35). The more even balance in CDR lengths between equivalent CDR loops on the VA and VB chains is consistent with the lower observed bias towards VDJ-chain dominated binding (Figure 1D).

In summary, the differences in the molecular properties of antibodies and TCRs have a direct impact on the roles of their CDR loops during pMHC recognition, with no apparent functional link between antibody/TCR chains made by analogous gene recombination mechanisms.

2.1.3 Binding properties by antigen region

BSAs computed only across the antigen (Figure 1E) reveal that the expected result that TCRs (μ: 918.2Å^2, sd: 122.0Å^2) and TCRms (μ: 999.8Å^2, sd: 79.5Å^2) tend to bury a larger total area of the antigen interface relative to general antibodies (μ: 789.0Å^2, sd: 222.7Å^2).

Although TCRs and TCRms bury a similar area of the pMHC, splitting this area into contributions by the peptide and MHC reveals that the TCRms are strongly biased towards a larger MHC BSA, with the effect that TCRms tend to recognise a smaller proportion of peptide surface during their pMHC recognition events (Figure 1F). This can also be expressed as the total number of peptide residues buried at least to some extent. Regardless of peptide length, no cognate TCR or TCRm has yet been found that can bury every peptide residue. However, for example in ninemer peptides presented by MHC class I, burial of eight peptide residues is not uncommon across TCRs (17.2%) but has not yet been observed in any TCRms (Figure 1G). Only 1/7 (14.3%) ninemer-pMHC class I-binding TCRms buries seven peptide residues (11D06 [7BBG]), while at least this many peptide residues are buried in 13/29 (44.8%) of corresponding TCRs.

Though we observe that TCRms have a lower proportion of pMHC BSA from the peptide, they do engage a similar number of peptide residues using a similar number of immunoglobulin residues as TCRs (Figures 1H, I). This also results in a similar number of formal interactions (Figure 1J). However, the BSA signal that TCRm:pMHC recognition is disproportionately biased towards MHC is recapitulated in the fact that TCRms tend to make more formal interactions to MHC residues than do TCRs (Figure 1K). Dissecting these contributions by immunoglobulin region shows this predominantly originates in TCRm CDRH3s being less peptide-focused than TCR CDRB3 loops (Figure 1L). These profiles suggest that, on aggregate, the current set of isolated TCRms are likely to be less peptide-selective than typical *in vitro*/*in vivo*-selected TCRs.

10.3389/fimmu.2022.1080596

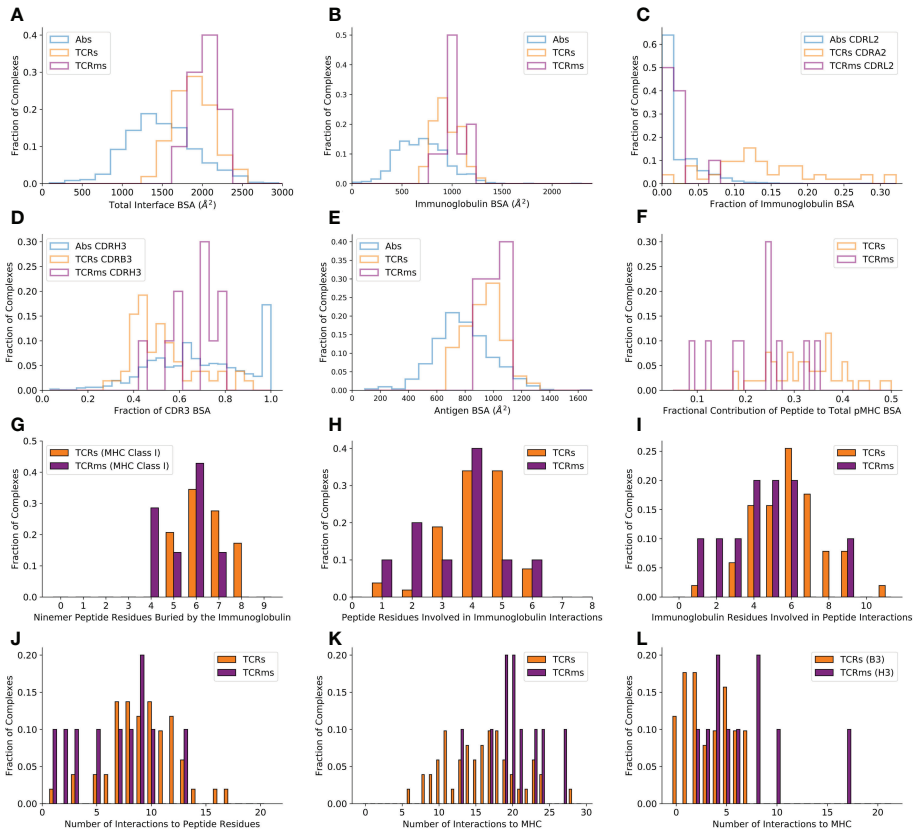

FIGURE 1
Buried surface area (BSA) and interaction profiles across the representative complexes from the Structural Antibody Database (6, 7) and Structural T-cell Receptor Database (4). **(A)** The total BSA across the immunoglobulin:antigen interface. **(B)** The immunoglobulin portion of the total interface BSA. **(C)** The fractional contribution of CDRL2 (antibodies, TCRms) or CDRA2 (TCRs) to immunoglobulin BSA. **(D)** The fractional contribution of CDRH3 (antibodies, TCRms) or CDRB3 (TCRs) to CDR3 BSA. **(E)** The antigen/pMHC portion of the total interface BSA. **(F)** The fractional contribution of peptide BSA to pMHC BSA. **(G)** The number of peptide residues buried in each immunoglobulin to ninemer peptide: MHC Class I complexes. **(H)** The number of peptide residues involved in binding interactions to the TCRms/TCRs. **(I)** The number of immunoglobulin residues involved in binding interactions to the peptide across the TCRms/TCRs. **(J)** The number of interactions between the immunoglobulin and the peptide across the TCRms/TCRs. **(K)** The number of interactions to MHC residues across the TCRms/TCRs. **(L)** The number of interactions to MHC residues made by the CDRH3 (TCRms) or CDRB3 (TCRs) loop.

2.2 TCRms can approach pMHCs with a diagonal orientation, but this does not guarantee TCR-like pMHC recognition

It has been previously shown that TCRs converge around diagonal engagement of pMHCs with the centre of mass of the beta chain sitting over the C-terminus of the MHC α_1 helix and the centre of mass of the alpha chain sitting over the C-terminus of the MHC α_2 helix (15, 36). This is quantified by the 'docking/crossing angle', calculated as the intercept of the line connecting conserved centres of mass within the variable region and the axis of the peptide binding groove. Our 60 representative TCR:pMHC complexes almost entirely comply with this canonical binding definition (μ: 45.8°, sd: 17.0°, Supplementary Figure 2A), with one outlier [Protein Data Bank (37) identifier (PDB ID) 4Y19 (38)] that engages pMHC with a diagonal but reverse polarity mode (136.7°; *i.e.* the VB chain sits above the MHC α_2 helix while the VA chain occludes the MHC α_1 helix, Supplementary Figure 2B).

To visualise their binding orientations, we aligned our set of TCRm complexes along the canonical peptide binding groove axis (Figure 2). While two of the TCRms bound in a non-diagonal fashion, the rest (80% of non-redundant TCRms) adopted a diagonal pMHC binding mode that fell within the range of absolute docking angles set by TCRs (Figure 2). However, in contrast to the TCRs, TCRm diagonal pMHC binding is more frequently achieved using reverse polarity (where VH is in the position of VA and VL is in the position of VB), reinforcing the notion that antibody and TCR chains with similar gene recombination mechanisms are not necessarily analogous in terms of pMHC recognition.

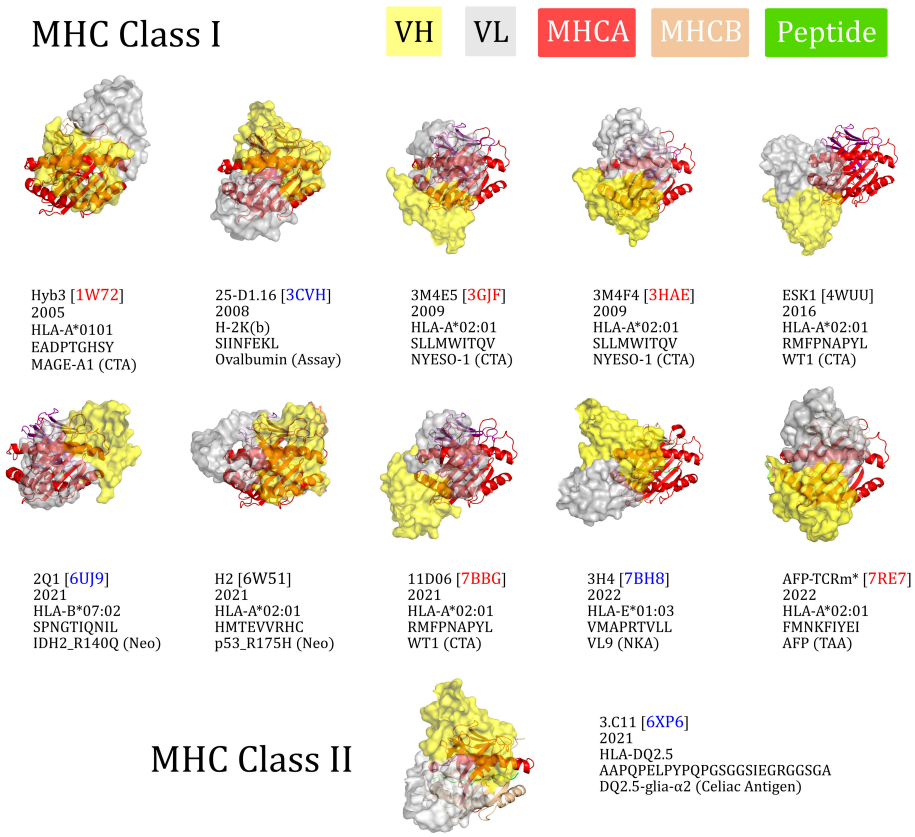

FIGURE 2
The eleven TCR-mimetic antibodies (TCRms) with solved structures in complex with their cognate pMHC, aligned by the MHC residues. The perspective is set so that the x-axis is the vector through the C_α atoms of the peptide anchor residues. 3M4E5 and 3M4F4 are closely related and derive from the same screening campaign, so are jointly represented by 3M4E5 in the analysis. Metadata is supplied below each structure: the TCRm name and Protein Data Bank (PDB) identifier, year of release in the PDB, MHC allele, antigen amino acid sequence, and antigen common name. ESK1 and H2 fall outside the range of diagonality seen in TCR:pMHC binding (PDB codes written in black). Diagonally-binding TCRms with VH in the position of VB and VL in the position of VA (canonical polarity binders) have blue PDB codes, while those with VH in the position of VA and VL in the position of VB (reverse polarity binders) have red PDB codes. CTA, cancer testis antigen; MHCA, MHC alpha chain; MHCB, MHC beta chain; Neo, neoantigen; NKA, natural killer cell antigen; TAA, tumor-associated antigen. *We have labelled the TCRm from 7RE7 as AFP-TCRm as it has no name in the seminal paper (29).

Convergence upon diagonal pMHC binding across TCRs is thought to be driven by improved typical TCR specificity, achieved by positioning the most hypervariable loops within interaction distance of the peptide, the key locus of pMHC variability (35). We therefore surveyed the properties of diagonally *versus* non-diagonally engaging TCRms to investigate to what extent this property correlates with more TCR-like pMHC recognition (Supplementary Figure 1).

The property distributions indicate that TCRm diagonal engagement is not systematically linked with total pMHC BSA (Supplementary Figure 1E) nor a higher fraction of peptide buried surface area (Supplementary Figure 1F). Some diagonal modes result in few formal interactions between the TCRm CDR3s and the peptide (e.g. Hyb3, with just two interactions to the peptide from CDRH3), while others result in numbers large even by TCR standards (Figure 1J; 3M4E5 has three formal interactions to the peptide from CDRH3 and seven from CDRL3). However, one property that appears to be systematically linked to TCRm diagonal engagement is the ability to bury a greater number of peptide residues; where N is the total number of peptide residues, non-diagonal pMHC class-I modes bury a maximum of N-4 residues, while diagonal modes frequently bury more, up to a currently-observed maximum of N-2 (Supplementary Figure 1G).

Overall these mixed results show that not all diagonal TCRm pMHC binding modes yield TCR-like pMHC engagement profiles, and that this property alone is likely insufficient for capturing the specificity of a TCRm.

10.3389/fimmu.2022.1080596

2.3 TCRms and TCRs have different trends in binding energetics even in identical pMHC contexts

We next performed molecular dynamics studies on three pMHC contexts for which we have a crystal structure in the PDB of at least one TCRm partner and at least one natural/affinity-enhanced TCR partner, allowing us to investigate their binding energetics. We selected three cases studies in which HLA-A*0201 presents a different peptide antigen:

(1) Wilms' Tumor 1 (WT1) antigen (RMFPNAPYL), bound to two TCRms and one affinity-enhanced TCR; ESK1 (PDB ID: 4WUU), 11D06 (PDB ID: 7BBG), and a7b2 (PDB ID: 6RSY), respectively.

(2) New York esophageal squamous cell carcinoma 1 (NY-ESO-1) antigen (SLLMWITQV), bound to one TCRm, one natural TCR, and three affinity-enhanced TCRs; 3M4E5 (PDB ID: 3GJF), sp3.4 (PDB ID: 6Q3S), NYE_S1 (PDB ID: 6RPB), NYE_S2 (PDB ID: 6RPA), and NYE_S3 (PDB ID: 6RP9), respectively.

(3) p53_R175H neoantigen (HMTEVVRHC), bound to one TCRm and three natural TCRs; H2 (PDB ID: 6W51), 1a2 (PDB ID: 6VQO), 12-6 (PDB ID: 6VRM), and 38-10 (PDB ID: 6VRN), respectively.

To characterize the energetic differences between the bound and unbound states of each complex, we ran sets of thirty 5-ns all-atom explicit solvent molecular dynamics simulations, using Molecular Mechanics Generalized-Born Surface Area (MMGBSA) (39–41) to compute overall free-energy changes upon binding as well as the per-residue contributions from both the pMHC and TCR/TCRm (see Methods, Supplementary Methods, Supplementary Figures 3–5). We used an ensemble sampling approach, *i.e.* a set of short statistically independent simulations initiated with different random velocities from the same starting structure, to ensure thorough sampling of the bound configurations.

Quantifying the relative contributions of the peptide, MHC α_1 helix, and MHC α_2 helix to the overall free energy change of the pMHC (see Supplementary Methods for equations), we find that both natural TCRs and affinity-enhanced TCRs tend to gain a larger proportion of their binding energy through the peptide than do TCRms (Figure 3, affinity-enhanced TCRs labelled TCR*). Within a given pMHC case study, no single TCRm achieved a higher proportion of binding energy from the peptide than any TCR. It is striking that some peptide antigens appear more tractable than others in terms of achieving a significant contribution to the overall binding energy. For example, even an affinity-enhanced TCR engaging the WT1 pMHC only reaches around 40% peptide contribution to binding free energy, while natural tumor-infiltrating lymphocytes complementary to p53_R175H achieve up to 65%.

We also considered the energies contributed by individual residues to identify interaction hotspots (Table 1; Supplementary Figures 2–4, Supplementary Table 4). Arbitrarily, we define interaction hotspots as individual residues that have a predicted free energy of -7 kcal/mol or stronger. All the natural TCRs assessed have at least one hotspot residue in the peptide, suggesting this may be a frequent feature of natural peptide recognition and consistent with current hypotheses on TCR recognition (43, 44). That the three p53_R175H specific

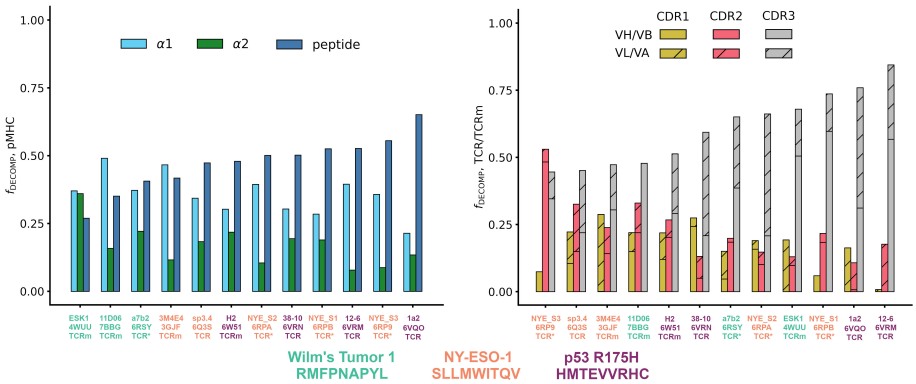

FIGURE 3
The energetics of pMHC binding of TCR-mimetic antibodies (TCRms), natural T-cell receptors (TCRs) and affinity-enhanced T-cell Receptors (TCR*s). (LHS) Relative contributions of the peptide, the MHC α_1 helix, and the MHC α_2 helix to the MMGBSA free energy change (42). Individual values were computed with Supplementary Equations 1-3, and then rank ordered by increasing peptide contribution. (RHS) Relative contributions to binding free energy of the CDR1, CDR2, and CDR3 loops for the heavy/beta chain (no hatch) and light/alpha chain (diagonal hatch), calculated as described in Methods. Complexes were rank ordered by increasing CDR3 contribution. Note that the CDR1, CDR2, and CDR3 bars do not sum to exactly one for all complexes as small negative values of the fraction of the DECOMP energy (f_{DECOMP}), representing CDRs that experience repulsive interactions with the pMHC, were passed to zero for visual clarity. NY-ESO-1: New York esophageal squamous cell carcinoma 1.

10.3389/fimmu.2022.1080596

TABLE 1 Decomposed per-residue energetic profiles for the case study TCR and TCRm complexes, grouped by pMHC region (peptide, MHC α_1 helix, or MHC α_2 helix).

Antigen	Name [PDB]	TCRm/TCR	Peptide Hotspots[1]	Num. Peptide/α_1/α_2 Hotspots
WT1	ESK1 [4WUU]	TCRm	**R**MFPNAPYL	0/0/0
	11D06 [7BBG]	TCRm	**R**MFP**N**APYL	0/1/0
	a7b2 [6RSY]	TCR[2]	RMFP**N**AP**Y**L	1/0/0
NYESO-1	3M4E5 [3GJF]	TCRm	SLL**M**W**I**TQV	1/1/0
	sp3.4 [6Q3S]	TCR	SLL**M**W**I**T**Q**V	2/0/0
	NYE_S3 [6RP9]	TCR[2]	SLLMW**I**TQV	0/0/0
	NYE_S2 [6RPA]	TCR[2]	SLL**M**W**I**TQV	1/0/0
	NYE_S1 [6RPB]	TCR[2]	SLL**M**W**I**T**Q**V	1/0/0
p53_R175H	H2 [6W51]	TCRm	HMTEVV**R**H**C**	1/0/0
	12-6 [6VRM]	TCR	HMTEVV**R**HC	1/0/0
	38-10 [6VRN]	TCR	HMTEVV**R**H**C**	1/0/0
	1a2 [6VQO]	TCR	HMTEVV**R**H**C**	1/0/0

[1]'Hotspots' are defined as residues predicted to have an attractive per-residue contribution of ≥ 7 kcal/mol to free energy based on DECOMP analysis (bold red text). 'Semi-hotspots' are defined as residues that contribute between -4 and -7 kcal/mol (bold black text). [2]Affinity-enhanced TCRs.

TCRs found in tumor-infiltrating lymphocytes achieve more neoantigen recognition at R7 rather than H8 (the site of somatic mutation) shows that this binding hotspot does not necessarily lie at the residue that distinguishes self from non-self.

By contrast, only TCRms were found to have hotspots that lie within the MHC (3M4E5 and 11D06). Despite their origins in different light chain loci (IGLV2-11 and IGKV1-5, respectively; Supplementary Table 5), the hotspot occurs at the same residue in the MHC α_1 helix (R65), which in both cases forms a salt bridge to D56 of CDRL2. That such non-peptide hotspots were not readily observed in TCR formats might indicate that current TCRms are prone to higher inherent affinity for the MHC and thus poorer typical peptide specificity. We observed from the crystal structures that in TCRm binding events that led to an energetic hotspot (3GJF, 7BBG), R65 tends to adopt a more extended conformation that protrudes towards the immunoglobulin, whereas in the TCR binding events without hotspots (6RSY, 6Q3S, 6RP9, 6RPA, 6RPB, 6VRM, 6VRN, 6VQO), R65 tends to adopt conformations that orient the plane of its guanidino group in line with that of the MHC α_1 helix (Supplementary Figure 6). This observation holds throughout our molecular dynamics simulations, as measured by the tilt of the plane of the guanidino group relative to the plane of the MHC α_1 helix over each of the 30 independent trajectories (Supplementary Figure 7, Supplementary Methods).

Finally, to characterize the differences in the energetics of CDR binding to the pMHCs, we computed the fractional contributions of each CDR to the total CDR free energy change (Figure 3). All TCRms and most TCR contexts were CDR3-dominated in their binding energetics. The balance of CDR[H/B]3 to CDR[L/A]3 energy reflected the picture seen in the general analysis of interface properties: TCRms gain more binding energy through CDRH3 than CDRL3, while in TCRs either CDR3 can be energetically dominant.

3 Discussion

TCRs exhibit a range of pMHC specificities, with a degree of polyspecificity considered an advantageous property to maximise the TCR repertoire's breadth of antigen recognition (45). However, whether due to evolutionary constraints or thymic selection mechanisms, TCRs have been observed to converge around several pMHC recognition properties including a canonical docking polarity, orientation, and the positioning of their most variable CDR loops above the peptide binding groove of the MHC (46). Regardless of underpinning mechanisms, this suggests there are conserved properties of pMHC recognition that grant TCRs the basal level of peptide specificity necessary to avoid widespread autoimmunity *in vivo*. The properties of their interfaces should therefore shed light on pMHC recognition features that ought to be generally advantageous to other molecules seeking to mimic them.

In this paper we have shown that, despite their broader genetic similarities, the idiotypic molecular configurations of antibody and TCR CDR loops contribute to the observed differences in their pMHC recognition tendencies. So far we lack evidence that genetically-equivalent antibody CDR loops can precisely imitate the burial profiles of TCR CDR loops in

their canonical binding footprint. However, relaxing the requirement for loop and chain equivalency, TCRms do exhibit a spectrum of TCR-likeness in the way they engage pMHC.

Within this spectrum, it is currently difficult to set thresholds for how TCR-like a TCRm needs to be to be clinically viable. Three of the TCRms, ESK1, 11D06, and AFP-TCRm, may begin to shed light on the answer. ESK1 is potentially the least TCR-like TCRm analysed in this study, engaging the pMHC with an orthogonal binding mode (Figure 2), the lowest fraction of binding energy to the peptide of any simulated complex (Figure 3), and several burial/interaction profiles on or outside of the bounds seen previously in TCRs (Supplementary Figure 8). By contrast, 11D06 and AFP-TCRm are considerably more TCR-like, engaging the pMHC diagonally (Figure 2) and with far fewer MHC interactions than ESK1 (Supplementary Figure 8). 11D06 buries the highest number of peptide residues, and AFP-TCRm the highest fraction of peptide/pMHC surface area, of any class-I TCRm to date.

Though ESK1 was reported many years before 11D06 and AFP-TCRm, it has not yet progressed through preclinical development and has several known human proteome off-targets (23), while 11D06 and AFP-TCRm have already advanced to in-human clinical trials. This suggests that increased TCR-likeness may be beneficial for clinical progression. However, it is worth reiterating that even these clinical TCRms are not perfectly TCR-like. For example, 11D06 harbors an interaction hotspot in the MHC, has a smaller percentage of pMHC binding energy from the peptide than any TCR assessed in the simulation study, and its VJ-recombined CDR3 loop plays no role in binding. This implies that some TCR-like pMHC recognition properties could be more important than others for therapeutic development.

One repeating feature that distinguishes the three TCRms able to recognise a large proportion of peptide residues without extensive MHC interaction (3M4E5 and both clinical-stage TCRms) is a salt bridge between IMGT residue D56 at the start of CDRL2 and residue R65 on the HLA-A2 MHC α_1 helix. While we were preparing this manuscript for publication, a new TCRm ('MA2') was released to the PDB that engages a decameric MART-1 antigen:HLA-A2 complex (https://www.rcsb.org/structure/7TR4); it too exploits a D56-R65 salt bridge, scanning across peptide with a reverse polarity diagonal binding mode. Though TCRm D56-R65 salt bridges seem to result in an energetic hotspot on the MHC (Supplementary Figures 3, 4), this may be compensated for by helping to orient the TCRm in a highly peptide-sensitive 'reverse polarity' pMHC recognition mode. We note that polar and/or charged interactions to R65 are also frequently observed in TCR interactions to HLA-A, commonly originating from a tyrosine (Y) or aspartic acid (D) in the CDRB2 loop (47). Therefore, it may be worth investigating whether the recurrent D56-R65 interaction can be exploited in future anti-HLA-A2 TCRm drug development pipelines; several kappa and lambda antibody germline genes encode D56 or E56, and these could form the basis of a targeted screening library (Supplementary Table 5).

There are several limitations to this study, primary amongst which is the relative paucity of TCRm structures, and particularly pMHC contexts, upon which to make conclusions about general TCRm pMHC recognition. Selection bias that cannot be accounted for through non-redundancy filtering alone is also likely to influence our conclusions. For example, it is likely that TCRms that have more promising specificity profiles would be more likely to be subjected to crystallographic analysis, which may artificially increase their apparent average TCR-likeness. The TCR-likeness of TCRms may also be expected to increase over time as *in vitro* negative selection techniques become ever more rigorous (e.g. negative baiting against panels of molecular mimicking off-targets, single point mutants of the target, or general broad screening panels of representative self or artificial peptides loaded onto MHC). Additionally, a notable trend in recent years has been to solve TCR:pMHC complexes that 'break the rules' rather than adhere to them (48), which may interfere with interpretability of the properties of our 'representative TCRs' as reflective of typical TCR behavior. Finally, neither our method for calculating solvent-accessible surface area (SASA), nor for calculating changes in free energy, account for induced fit that may occur during the binding event.

Nevertheless even with these early data, some general properties of TCRms and TCRs, such as their balance of peptide *versus* MHC recognition, show marked differences. As more TCRms and immune mobilising monoclonal T-cell receptors against cancer (immTACs) are developed and structurally characterized, it should be insightful to evaluate their binding properties and assess which pMHC recognition features are most reliably linked with successful clinical progression. For now, our TCR:pMHC complex profiles offer an initial set of benchmarks for the rational computational derisking of novel TCR-mimetic modalities.

4 Methods

4.1 Numbering and region definitions

To enable a direct comparison between antibodies and TCRs, and between heavy and light chains, the IMGT numbering scheme and CDR definitions were used throughout this work [CDR1: IMGT residues 27-38, CDR2: IMGT residues 56-65, CDR3: IMGT residues 105-117 (49)]. ANARCI was used to number all sequence inputs (50). Where necessary, MHC chains were renumbered to enable a direct comparison between TCRm and TCR complexes.

4.2 Structure datasets

SAbDab (6, 7) and STCRDab (4) were downloaded on 30th September, 2022. All complexes were stripped of explicit hydrogens, heteroatoms, and water molecules, and only immunoglobulins with protein antigens were considered. These databases were then mined for particular subcategories as follows:

1. TCR-mimetic antibody (TCRm) complexes were identified by filtering SAbDab with the search terms 'HLA' and 'MHC', followed by manual validation.

2. Representative sets of non-redundant high quality antibody:antigen and TCR:pMHC (αβ only) complexes were derived by first filtering SAbDab and STCRDab for structures of complexes solved by X-ray crystallography to 2.5 Å resolution. For each class of immunoglobulin separately, the IMGT-defined CDR sequences (49) were concatenated in a consistent order and used as inputs to greedy clustering by cd-hit (51) (80% sequence identity threshold), to create 'paratope clusters'. To identify cases where chemically-similar paratopes bind to significantly different antigens, we performed a second round of clustering over the paratope clusters using the concatenated sequence of the antigen (s) associated with each antibody (concatenated in descending length order), or the sequence of the presented peptide for the TCRs. cd-hit was run at an 80% sequence identity threshold, allowing a minimum alignment length of as little as 20%. This ensured truncated antigens were not considered as different targets (e.g. an anti-coronavirus antibody solved in complex with the receptor-binding domain would be considered the same context as a complex of the same antibody binding to the full-length spike protein). This resulted in sets of 824, 52, and 8 representative antibody:antigen, TCR:pMHC (class 1) and TCR:pMHC (class 2) complexes respectively.

3. Complexes of TCRs and TCRms engaging the same peptide were obtained by searching SAbDab and STCRDab by antigen sequence for peptide fragments with a 100% sequence identity match.

4.3 Buried surface area calculations

Complete immunoglobulins were separated from their antigens (*i.e* the co-ordinates of the partner were deleted in a copy of the original file), and the difference in solvent-accessible surface area between each residue in the original complex and artificially-generated 'apo' state was recorded. Per-residue solvent-accessible surface area (SASA) was calculated using an in-house implementation of the Shrake and Rupley algorithm (52), applying a probe radius of 1.4Å. Residues were defined as buried upon binding if their SASA decreased when the co-ordinates of the binder partner were re-introduced into the system.

4.4 Interaction mapping

The Arpeggio (53) software package was used to assign interactions across each set of complexes. Arpeggio featurises input protein complexes with predicted interaction types based on the relative distances and orientations of PDB atom types; see (53) for definitions. In our analysis, we only considered the four most attractive interaction types: hydrophobic, aromatic, hydrogen bond and salt bridge interactions. We disregarded weak hydrogen bond assignments, as they ignore the relative angular dependence of hydrogen bond acceptor and donor. To interpret arpeggio interaction outputs on a per-residue level requires a consolidation of atom-level interaction profiles, which we performed according to the following post-processing steps:

1. Where two residues had multiple hydrophobic atom-atom contacts this was recorded as one residue-residue "hydrophobic" interaction.

2. Where two residues had multiple aromatic atom-atom contacts this was recorded as one residue-residue "aromatic" interaction.

3. A residue-residue pair can be labelled as contributing both a hydrophobic and an aromatic interaction due to their different electrostatic origins.

4. Polar interactions between oppositely charged residues were initially labelled as 'salt bridge/hydrogen bond'. Once all interactions were recorded the salt bridge was assigned to the pair of residues with the closest atom-atom pair and all other interactions were considered hydrogen bonds, ensuring only one salt bridge was recorded per positive/negative charge pair.

4.5 Molecular dynamics simulations and analysis

pMHC-TCR and pMHC-TCRm complexes were prepared in AmberTools21 (54) (see Supplementary Table 6 for rebuilt sections) and all simulations run in OpenMM v7.5 (55). Cis peptide bonds and chiral centers with incorrect stereochemistry within rebuilt sections of protein structures were identified and corrected using the Cispeptide and Chirality (56) plugins of Visual Molecular Dynamics (57) v1.8.3. Corrected peptide bonds were maintained in the trans configuration with periodic torsion restraints with force constant 500 kJ/mol. These restraints were removed during the final unrestrained equilibration along with positional restraints on C_α atoms. Complexes were then solvated in orthorhombic boxes with a minimum of 1.4 nm between protein atoms and the box edge. Sodium and chloride ions were added to neutralize each system and bring the salt concentration to 0.15 M.

All simulations were carried out using FF14SB (58) and TIP3P (59) forcefield parameters with a Langevin thermostat (friction coefficient 1 ps^{-1}) and, for constant-pressure simulations, a Monte Carlo barostat (pressure 1 bar). Complexes were minimized and then heated to 298 K at

constant volume with protein atoms restrained. These restraints were relaxed over a set of five 100-ps simulations at constant pressure, culminating with 100-ps of unrestrained equilibration. Finally, 5-ns production simulations at constant pressure were run. Thirty independent replicas of this protocol were executed for each complex (60). MMGBSA calculations were performed with MMPBSA.py (42) using 100 frames collected every 40 ps from the final 4 ns of each 5-ns trajectory (i.e., 3000 frames per complex). Per-residue energy contributions to the binding energy were computed using the MMPBSA.py DECOMP functionality (42). Further detail on all simulation methods can be found in the Supplementary Information.

4.6 Visualizations

All manuscript visualizations were created using open-source PyMOL or matplotlib version 3.5.2.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://zenodo.org/record/7220531.

Author contributions

All authors designed the research; MR curated the datasets and analyzed the crystal complexes, DN performed/analyzed the molecular dynamics simulations; all authors analyzed the aggregate data. All authors contributed to the article and approved the submitted version.

Funding

The authors declare that this study received funding from Boehringer Ingelheim. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

The authors thank Drs. Nicolas Sabarth, Niksa Kastrapeli, Andrew Nixon, Justin Scheer, Noah Pefaur, Srinath Kasturirangan, Paul Adams, Renate Konopitzkey and several other researchers at Boehringer Ingelheim for numerous insightful discussions and their critical reading of this manuscript.

Conflict of interest

SK was employed by Boehringer Ingelheim.

The remaining authors declare that the research was conducted in the absence of any commercial of financial relationship that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.1080596/full#supplementary-material

References

1. Cano RLE, Lopera HDE. Introduction to T and B lymphocytes *Autoimmunity from bench to bedside*. (El Rosario: University Press) (2013).

2. Wylie DE, Klinman NR. The murine B cell repertoire responsive to an influenza-infected syngeneic cell line. *J Immunol* (1981) 127:194–8. doi: 10.4049/jimmunol.127.1.194

3. 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

4. Leem J, de Oliveira SHP, Krawczyk K, Deane CM. STCRDab: the structural T-cell receptor database. *Nucleic Acids Res* (2018) 46:D406–12. doi: 10.1093/nar/gkx971

5. Nathan P, Hassel JC, Rutkowski P, Baurain J-F, Butler MO, Schlaak M, et al. Overall survival benefit with tebentafusp in metastatic uveal melanoma. *New Engl J Med* (2021) 385:1196–206. doi: 10.1056/NEJMoa2103485

6. Dunbar J, Konrad K, Leem J, Baker T, Fuchs A, Georges G, et al. SAbDab: the structural antibody database. *Nucleic Acids Res* (2014) 42:D1140–6. doi: 10.1093/nar/gkt1043

7. Schneider C, Raybould MIJ, Deane CM. SAbDab in the age of biotherapeutics: Updates including SAbDab-nano, the nanobody structure tracker. *Nucleic Acids Res* (2022) 50:D1368–72. doi: 10.1093/nar/gkab1050

8. Slaney CY, Wang P, Darcy PK, Kershaw MH. CARs versus BiTEs: A comparison between T cell–redirection strategies for cancer treatment. *Cancer Discovery* (2018) 8:924–34. doi: 10.1158/2159-8290.CD-18-0297

9. He Q, Liu Z, Liu Z, Lai Y, Zhou X, Weng J. TCR-like antibodies in cancer immunotherapy. *J Hematol Oncol* (2019) 12:99. doi: 10.1186/s13045-019-0788-4

10. Yu L, Wang J. T Cell-redirecting bispecific antibodies in cancer immunotherapy: recent advances. *J Cancer Res Clin Oncol* (2019) 145:941–56. doi: 10.1007/s00432-019-02867-6

11. Blanco B, Compte M, Lykkemark S, Sanz L, Alvarez-Vallina L. T Cell-redirecting strategies to 'STAb' tumors: Beyond CARs and bispecific antibodies. *Trends Immunol* (2019) 40:243–57. doi: 10.1016/j.it.2019.01.008

12. Dadaglio G, Nelson CA, Deck MB, Petzold SJ, Unanue ER. Characterization and quantitation of peptide–MHC complexes produced from hen egg lysozyme using a monoclonal antibody. *Immunity* (1997) 6:727–38. doi: 10.1016/S1074-7613(00)80448-3

13. Porgador A, Yewdell JW, Deng Y, Bennink RN, Germain JR. Localization, quantitation, and in situ detection of specific peptide–MHC class I complexes using a monoclonal antibody. *Immunity* (1997) 6:715–26. doi: 10.1016/S1074-7613(00)80447-1

14. Denkberg G, Reiter Y. Recombinant antibodies with T-cell receptor-like specificity: Novel tools to study MHC class I presentation. *Autoimmun Rev* (2006) 5:252–7. doi: 10.1016/j.autrev.2005.07.004

15. Høydahl LS, Frick R, Sandlie I, Løset GÅ. Targeting the MHC ligandome by use of TCR-like antibodies. *Antibodies* (2019) 8:32. doi: 10.3390/antib8020032

16. Chames P, Hufton SE, Coulie PG, Uchanska-Ziegler B, Hoogenboom HR. Direct selection of a human antibody fragment directed against the tumor T-cell epitope HLA-A1–MAGE-A1 from a nonimmunized phage-fab library. *Proc Natl Acad Sci* (2000) 97:7969–74. doi: 10.1073/pnas.97.14.7969

17. Hulsmeyer M, Chames P, Hillig RC, Stanfield RL, Held G, Coulie PG, et al. A major histocompatibility Complex·Peptide-restricted antibody and T cell receptor molecules recognize their target by distinct binding modes: CRYSTAL STRUCTURE OF HUMAN LEUKOCYTE ANTIGEN (HLA)-A1·MAGE-A1 IN COMPLEX WITH FAB-HYB3. *J Biol Chem* (2005) 280:2972–80. doi: 10.1074/jbc.M411323200

18. Mareeva T, Martinez-Hackert E, Sykulev Y. How a T cell receptor-like antibody recognizes major histocompatibility complex-bound peptide. *J Biol Chem* (2008) 283:29053–9. doi: 10.1074/jbc.M804996200

19. Stewart-Jones G, Wadle A, Hombach A, Shenderov E, Held G, Fischer E, et al. Rational development of high-affinity T-cell receptor-like antibodies. *Proc Natl Acad Sci* (2009) 106:5784–8. doi: 10.1073/pnas.0901425106

20. Ataie N, Xiang J, Cheng N, Brea EJ, Lu W, Scheinberg DA, et al. Structure of a TCR-mimic antibody with target predicts pharmacogenetics. *J Mol Biol* (2016) 428:194–205. doi: 10.1016/j.jmb.2015.12.002

21. Li Y, Jiang W, Mellins ED. TCR-like antibodies targeting autoantigen-mhc complexes: a mini-review. *Front Immunol* (2022) 13:968432. doi: 10.3389/fimmu.2022.968432

22. Holland CJ, Crean RM, Pentier JM, de Wet B, Lloyd A, Srikannathasan V, et al. Specificity of bispecific T cell receptors and antibodies targeting peptide-HLA. *J Clin Invest* (2020) 130:2673–2688. doi: 10.1172/JCI130562

23. Gejman RS, Jones HF, Klatt MG, Chang AY, Oh CY, Chandran SS, et al. Identification of the targets of T-cell receptor therapeutic agents and cells by use of a high-throughput genetic platform. *Cancer Immunol Res* (2020) 8:672–84. doi: 10.1158/2326-6066.CIR-19-0745

24. Augsberger C, Hänel G, Xu W, Pulko V, Hanisch LJ, Augustin A, et al. Targeting intracellular WT1 in AML with a novel RMF-peptide-MHC-specific T-cell bispecific antibody. *Blood* (2021) 138:2655–69. doi: 10.1182/blood.2020010477

25. Hwang MS, Miller MS, Thirawatananond P, Douglass J, Wright KM, Hsiue EH-C, et al. Structural engineering of chimeric antigen receptors targeting HLA-restricted neoantigens. *Nat Commun* (2021) 12:5271. doi: 10.1038/s41467-021-25605-4

26. Hsiue EH-C, Wright KM, Douglass J, Hwang MS, Mog BJ, Pearlman AH, et al. Targeting a neoantigen derived from a common TP53 mutation. *Science* (2021) 371:eabc8697. doi: 10.1126/science.abc8697

27. Frick R, Høydahl LS, Petersen J, du Pré F, Kumari S, Bernstsen G, et al. A high-affinity human TCR-like antibody detects celiac disease gluten peptide–MHC complexes and inhibits T cell activation. *Sci Immunol* (2021) 6:eabg4925. doi: 10.1126/sciimmunol.abg4925

28. Li D, Brackenridge S, Walters LC, Swanson O, Harlos K, Rozbesky D, et al. Mouse and human antibodies bind HLA-e-leader peptide complexes and enhance NK cell cytotoxicity. *Commun Biol* (2022) 5:271. doi: 10.1038/s42003-022-03183-5

29. Liu C, Liu H, Dasgupta M, Hellman LM, Zhang X, Qu K, et al. Validation and promise of a TCR mimic antibody for cancer immunotherapy of hepatocellular carcinoma. *Sci Rep* (2022) 12:12068. doi: 10.1038/s41598-022-15946-5

30. Gu Y, Wong YH, Liew CW, Chan CEZ, Murali TM, Yap J, et al. Defining the structural basis for human alloantibody binding to human leukocyte antigen allele HLA-A*11:01. *Nat Commun* (2019) 10:893. doi: 10.1038/s41467-019-08790-1

31. Soto C, Bombardi RG, Kozhenikov M, Sinkovits RS, Chen EC, Branchizio A, et al. High frequency of shared clonotypes in human T cell receptor repertoires. *Cell Rep* (2020) 32:107882. doi: 10.1016/j.celrep.2020.107882

32. Raybould MIJ, Marks C, Krawczyk K, Taddese B, Nowak J, Lewis AP, et al. Five computational developability guidelines for therapeutic antibody profiling. *Proc Natl Acad Sci* (2019) 116:4025–40301. doi: 10.1073/pnas.1810576116

33. Kovaltsuk A, Leem J, Kelm S, Snowden J, Deane CM, Krawczyk K. Observed antibody space: A resource for data mining next-generation sequencing of antibody repertoires. *J Immunol* (2018) 201:2502–9. doi: 10.4049/jimmunol.1800708

34. Olsen TH, Boyles F, Deane CM. Observed antibody space: A diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Sci* (2022) 31:141–6. doi: 10.1002/pro.4205

35. Wong WK, Leem J, Deane CM. Comparative analysis of the CDR loops of antigen receptors. *Front Immunol* (2019) 10:2454. doi: 10.3389/fimmu.2019.02454

36. Dunbar J, Knapp B, Fuchs A, Shi J, Deane CM. Examining variable domain orientations in antigen receptors gives insight into TCR-like antibody design. *PloS Comput Biol* (2014) 10:e1003852. doi: 10.1371/journal.pcbi.1003852

37. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Res* (2000) 28:235–42. doi: 10.1093/nar/28.1.235

38. Bering DX, Kleijwegt FS, Wiede F, van der Slik AR, Loh KL, Petersen J, et al. T Cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. *Nat Immunol* (2015) 16:1153–61. doi: 10.1038/ni.3271

39. Chen F, Liu H, Sun H, Pan P, Li Y, Li D, et al. Assessing the performance of the MM/PBSA andMM/GBSA methods. 6. capability to predictprotein–protein binding free energies and re-rankbinding poses generated by protein–protein docking. *Phys Chem Chem Phys* (2016) 18:22129. doi: 10.1039/c6cp03670h

40. Wang E, Sun H, Wang J, Wang Z, Liu H, Zhang JZH, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem Rev* (2019) 119:9478–508. doi: 10.1021/acs.chemrev.9b00055

41. Onufriev AV, Case DA. Generalized born implicit solvent models for biomolecules. *Annu Rev Biophys* (2019) 48:275–96. doi: 10.1146/annurev-biophys-052118-115325

42. Miller BRIII, McGee D, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA.py: An efficient program for end-state free energy calculations. *J Chem Theory Comput* (2012) 8:3314–21. doi: 10.1021/ct300418h

43. Adams JJ, Narayanan S, Liu B, Birnbaum ME, Kruse AC, Bowerman NA, et al. T Cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* (2011) 35:681–93. doi: 10.1016/j.immuni.2011.09.013

44. Singh NK, Riley TP, Baker SCB, Borrman T, Weng Z, Baker BM. Emerging concepts in TCR specificity: Rationalizing and (maybe) predicting outcomes. *J Immunol* (2017) 199:2203–13. doi: 10.4049/JIMMUNOL.1700744

45. Sewell AK. Why must T cells be cross-reactive? *Nat Rev Immunol* (2012) 12:669–77. doi: 10.1038/nri3279

46. Zareie P, Szeto C, Farenc C, Gunasinghe SD, Kolawole EM, Nguyen A, et al. Canonical T cell receptor docking on peptide:MHC is essential for T cell signaling. *Science* (2021) 372:eabe9124. doi: 10.1126/science.abe9124

47. Blevins SJ, Pierce BG, Singh NK, Riley TP, Wang Y, Spear TT, et al. How structural adaptability exists alongside HLA-A2 bias in the human $a\beta$ TCR repertoire. *Proc Natl Acad Sci* (2016) 113:E1276–85. doi: 10.1073/pnas.1522069113

48. Szeto C, Lobos CA, Nguyen AT, Gras S. TCR recognition of peptide–MHC-I: Rule makers and breakers. *Int J Mol Sci* (2021) 22:68. doi: 10.3390/ijms22010068

49. Lefranc M-P, Pommié C, Ruiz M, Giudicelli V, Foulquier E, Truong L, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and ig superfamily V-like domains. *Dev Comp Immunol* (2003) 27:55–77. doi: 10.1016/S0145-305X(02)00039-3

50. Dunbar J, Deane CM. ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* (2015) 32:298–300. doi: 10.1093/bioinformatics/btv552

51. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* (2006) 22:1658–9. doi: 10.1093/bioinformatics/btl158

52. Shrake A, Rupley JA. Environment and exposure to solvent of protein atoms. lysozyme and insulin. *J Mol Biol* (1973) 79:351–71. doi: 10.1016/0022-2836(73)90011-9

53. Jubb HC, Higueruelo AP, Ochoa-Montaño B, Piit WR, Ascher DB, Blundell TL. Arpeggio: A web server for calculating and visualising interatomic interactions in protein structures. *J Mol Biol* (2017) 429:365–71. doi: 10.1016/j.jmb.2016.12.004

54. Salomon-Ferrer R, Case DA, Walker RC. An overview of the amber biomolecular simulation package. *Wiley Interdiscip Rev Compu Mol Sci* (2013) 3:198–210. doi: 10.1002/wcms.1121

55. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol* 13:e1005659. doi: 10.1371/journal.pcbi.1005659

56. Schreiner E, Trabuco LG, Freddolino PL, Schulten K. Stereochemical errors and their implications for molecular dynamics simulations. *BMC Bioinf* (2011) 12:190. doi: 10.1186/1471-2105-12-190

57. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph* (1996) 14:22–38. doi: 10.1016/0263-7855(96)00018-5

58. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* (2015) 11:3696–713. doi: 10.1021/acs.jctc.5b00255

59. Jorgensen WL, Chandrasekhar J, Madura JD. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* (1983) 79:926. doi: 10.1063/1.445869

60. Knapp B, Ospina L, Deane CM. Avoiding false positive conclusions in molecular simulation: The importance of replicas. *J Chem Theory Comput* (2018) 14:6127–38. doi: 10.1021/acs.jctc.8b00391