frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 02 December 2024
DOI 10.3389/fimmu.2024.1491656

OPEN ACCESS

EDITED BY
Jonathan S. Duke-Cohan,
Dana–Farber Cancer Institute, United States

REVIEWED BY
Brian G. Pierce,
University of Maryland, College Park, United States
Cory Ayres,
University of Notre Dame, United States

*CORRESPONDENCE
Charlotte M. Deane
deane@stats.ox.ac.uk
Hashem Koohy
hashem.koohy@rdm.ox.ac.uk

RECEIVED 05 September 2024
ACCEPTED 24 October 2024
PUBLISHED 02 December 2024

CITATION
McMaster B, Thorpe CJ, Rossjohn J, Deane CM and Koohy H (2024) Quantifying conformational changes in the TCR:pMHC-I binding interface. *Front. Immunol.* 15:1491656. doi: 10.3389/fimmu.2024.1491656

Quantifying conformational changes in the TCR:pMHC-I binding interface

Benjamin McMaster[1,2], Christopher J. Thorpe[3,4], Jamie Rossjohn[5,6], Charlotte M. Deane[2*] and Hashem Koohy[1*]

[1]Koohy Lab, Medical Research Council Translational Immune Discovery Unit (MRC TIDU), Weatherall Institute of Molecular Medicine (WIMM), Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom, [2]Oxford Protein Informatics Group, Department of Statistics, University of Oxford, Oxford, United Kingdom, [3]Open Targets, Wellcome Genome Campus, Hinxton, United Kingdom, [4]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, United Kingdom, [5]Rossjohn Lab, Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Melbourne, VIC, Australia, [6]Rossjohn Lab, Institute of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, United Kingdom

Background: T cells form one of the key pillars of adaptive immunity. Using their surface bound T cell antigen receptors (TCRs), these cells screen millions of antigens presented by major histocompatibility complex (MHC) or MHC-like molecules. In other protein families, the dynamics of protein-protein interactions have important implications for protein function. Case studies of TCR:class I peptide-MHCs (pMHC-Is) structures have reported mixed results on whether the binding interfaces undergo conformational change during engagement and no robust statistical quantification has been done to generalise these results. Thus, it remains an open question of whether movement occurs in the binding interface that enables the recognition and activation of T cells.

Methods: In this work, we quantify the conformational changes in the TCR:pMHC-I binding interface by creating a dataset of 391 structures, comprising 22 TCRs, 19 MHC alleles, and 79 peptide structures in both unbound (apo) and bound (holo) conformations.

Results: In support of some case studies, we demonstrate that all complementarity determining region (CDR) loops move to a certain extent but only CDR3α and CDR3β loops modify their shape when binding pMHC-Is. We also map the contacts between TCRs and pMHC-Is, generating a novel fingerprint of TCRs on MHC molecules and show that the CDR3α tends to bind the N-terminus of the peptide and the CDR3β tends to bind the C-terminus of the peptide. Finally, we show that the presented peptides can undergo conformational changes when engaged by TCRs, as has been reported in past literature, but novelly show these changes depend on how the peptides are anchored in the MHC binding groove.

Conclusions: Our work has implications in understanding the behaviour of TCR:pMHC-I interactions and providing insights that can be used for modelling Tcell antigen specificity, an ongoing grand challenge in immunology.

KEYWORDS

TCR, MHC, peptide, HLA, conformational changes, T cell antigen specificity, structural biology

1 Introduction

T cells are essential cells of the adaptive immune system, responsible for identifying and eliminating foreign pathogens or malfunctioning cells to maintain homeostasis. To discriminate between foreign antigens and self-peptides, T cells use their surface bound T cell antigen receptors (TCRs) to scan linearised peptides presented by major histocompatibility complex (MHC) molecules.

TCRs are hetero-dimeric molecules, consisting of an α- and β-chain (or a γ- and δ-chain for a smaller subset) with a constant region that anchors them to the cell membrane and a variable domain responsible for antigen binding (1). The antigen binding region is formed of six protein loops (three on the α-chain and three on the β-chain) known as the complementarity determining regions (CDRs). These loops are the product of a stochastic gene rearrangement process known as V(D)J recombination that occurs during T cell development in the thymus and and results in the huge breadth of T cell diversity (2). Unlike their B cell counterparts, known as B cell receptors or antibodies, TCRs do not undergo any further modifications to influence their antigen specificity. On the antigen presentation side, class I MHC molecules are found on all nucleated cells and are responsible for displaying peptide fragments of degraded proteins on the cell surface. Based on recognition from T cells, these molecules mark cells' internal states as diseased or healthy, helping to ensure homeostasis through the removal of the diseased cells. The class I peptide-MHC (pMHC-I) complex is formed of a membrane-anchored domain and an antigen-binding domain with a groove holding the peptide created by two α-helices and a floor formed of seven anti-parallel β-strands. The complex is stabilised by a β_2-microglobulin protein underneath the antigen binding domain. Recognition of a pMHC-I by a T cell can start a cascade of signalling molecules leading to an immune response.

Proteins by nature can be dynamic entities; they can exist in multiple conformations and use movements to carry out specific biological functions. For example, kinesins undergo large conformational changes when phosphorylated to "walk" down cytoskeletal structures and transport other molecules around the cell (3). When proteins interact with other proteins, there can often be conformational changes to improve the selectivity and strength of binding to one another (4). These changes in structure come with entropic and enthalpic considerations and are thought of in three modalities: 1) the "lock-and-key" model states that neither protein moves and the shapes fit together incurring little to no free energy penalty during binding, 2) the "induced-fit" model assumes that one protein moves while the other remains fixed, or 3) the "pre-existing equilibrium" (also called "conformational selection") states that the proteins exist as conformational ensembles and that when the conformations are both right the proteins can bind (4).

Emerging reports have sought to quantify these principles in antibodies and show that they undergo some conformational change in their CDR regions when binding to antigens (5). In past literature of TCR and pMHC-I interactions, there has been evidence of some conformational changes and plasticity (6–12). Kjer-Nielsen et al. show that the CDR loops of a TCR move to form the bound complex with a pMHC-I (8, 9). Tynan et al. also show that the binding of a TCR onto a pMHC-I molecule flattens a bulging peptide (12). Contrarily, Chen et al. show in a different TCR:pMHC-I both the TCR and pMHC-I maintain their shape a fit together in a "lock-and-key" mode (11). Other past studies have argued that structural rearrangements in the TCR:pMHC-I binding interface affect T cell activation and function (13). Armstrong et al. analysed these early structures from the 9 unique TCRs available in bound and unbound forms (14). Their work concluded that all CDR loops undergo conformational changes but CDR3α and CDR3β have the largest movements. They also argue that TCR:pMHC interactions fall somewhere into the paradigms of "induced-fit" and "pre-existing equilibrium" protein-protein interactions. Since Armstrong et al.'s work in 2008, over 500 new TCR structures have been deposited in the RCSB protein data bank (PDB) (15) and little has been done to conduct a similar analysis on the larger amount of available structures. Questions remain on how past findings generalise to broader TCR and pMHC-I interactions and to quantify the degree and type of movements these molecules may undergo between the unbound (*apo*) TCR and pMHC-I complexes, and the bound (*holo*) TCR:pMHC-I complex.

In this work, we present an analysis of the conformational changes observed during the engagement of TCRs and pMHC-Is. We leverage databases containing TCR and pMHC-I structures to curate a dataset of 391 structures with 22 TCRs binding 19 MHC alleles and 79 different peptides, with both unbound (*apo*) and bound (*holo*) forms of TCRs and pMHC-Is. Using this dataset, we conduct a robust statistical analysis of the amount and type of movement these entities undergo when coming into contact with one another at a scale not previously done in the literature. Our analysis reveals that all CDR loops undergo conformational change between *apo* and *holo* states but that only the CDR3 loops are flexible. By mapping the contacts made between TCR CDR loops and the surface of the pMHC-I complex, we show that the interactions occur in a constrained space for each CDR loop and that both CDR3 loops are equally involved in peptide binding, with the CDR3α focused on the N-terminus of the peptide and the CDR3β focused on the C-terminus of the peptide. We also show that peptides can undergo conformational change when engaged by TCRs and this movement is dictated by how the peptide is anchored in the MHC binding groove. Our work provides a quantitative picture of TCR engagement with pMHC-I molecules, generating insights into the behaviour of T cell antigen recognition.

2 Results

We began our analysis by creating a dataset of TCRs and pMHC-Is structures in both their *apo* and *holo* forms. These structures were collected from the STCRDab (15) and histo.fyi (16) and were subject to the quality screening and alignment procedures described in Section 4.1. The TCR gene usage, MHC alleles, and peptide redundancy of the dataset are visualised in

Figures 1A–C. The dataset contains 22 unique TCRs as defined by their IMGT CDRs, 19 unique MHC alleles, and 79 unique peptides.

We compared the TCRs in the dataset to a background of TCRs sampled from the Observed TCR Space (OTS) database (17) to ascertain how representative the TCRs used in this analysis are to an expected distribution of TCRs. We compared common sequence properties of the molecules including V gene usage, CDR length, and amino acid composition of the CDR loops. Figure 1D shows that for the most part, V genes are well represented within our dataset, as indicated by the near zero percent enrichment or depletion. However, there are several notably enriched genes: TRAV12, TRAV14, TRAV21, TRBV6, and TRBV19, as well as some depleted genes: TRAV8, TRAV13, TRBV5, and TRBV20. Although these results indicate some bias in the dataset, some of these enrichments/depletions are expected as this work focuses solely on TCRs interacting with pMHC-I derived from CD8+ T cells unlike OTS that contains TCRs derived from both CD4+ and CD8+ T cells. For example, TRAV12 and TRAV21 have been shown to be enriched in populations of CD8+ T cells whereas TRAV8 is comparatively enriched in CD4+ T cells, explaining its depletion here (18). When comparing CDR lengths between the selected structures and OTS background (Figure 1E), it seems the distributions are well matched with the top loop length matching in 5 of the 6 loop types (the selected structures for this analysis are lacking CDR3α loops that are 12 amino acids long). A comparison of amino acid composition in the CDR loops (shown in Supplementary Figure S1) also shows no major differences between the selected structures and the OTS background. The TCRs selected for this analysis are primarily human (84.70%), with a small proportion of mouse TCRs binding mouse MHC molecules, and are comparable to the species of the TCR sequences sampled from OTS (see Supplementary Figure S2).

The representation of canonical loop classes was also assessed in the selected structures. The process for assigning loops to canonical classes is described in Section 4.3. Figure 1F highlights the coverage of canonical classes in our dataset. 86.96% of canonical loop classes are represented in our dataset meaning most standard configurations of loop conformations are included.

Although the data contains some biases, through these results we show the dataset of structures in *apo* and *holo* conformations is representative of the broader TCR:pMHC-I interactions.

2.1 TCRs and MHCs undergo significant conformational changes between *apo* and *holo* states

Our analysis shows that all six CDR loops undergo conformational change when a TCR engages with a pMHC-I. Figures 2A, B depict and quantify these changes respectively, with a dashed red line in Figure 2B as a visual aid of the expected noise based on other reports of general noise in crystallography data (5). Figure 2C categorises the movements of all loops between *apo* and *holo* states. The quantification of each loop is reported as the backbone root mean squared deviation (RMSD) between *apo* and *holo* conformations after the TCRs are aligned on their framework regions (see Section 4.2). The CDR1α, CDR2α, and CDR3α loops have a mean change of 1.69 Å, 1.33 Å, and 2.38 Å respectively, with

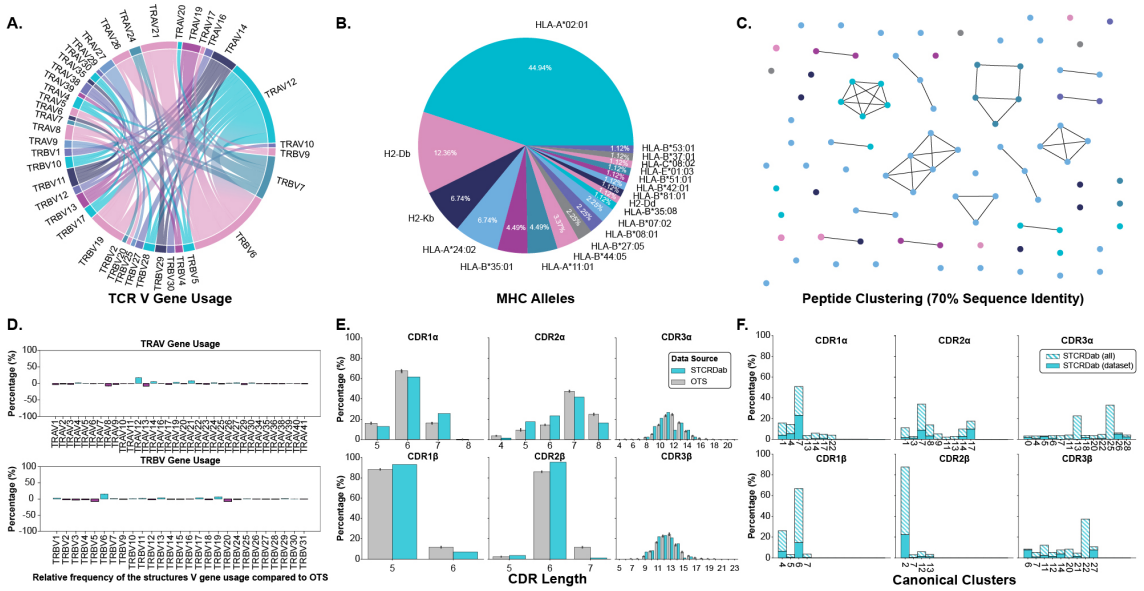

FIGURE 1
Description of TCRs and pMHC-Is in the *apo-holo* analysis dataset. **(A)** Pairings of V genes used by the TCRs in the analysis. **(B)** Proportion of different MHC alleles in the dataset. **(C)** Clustering of peptides in the dataset based on sequence identity. Clusters are formed from peptides with 70% sequence identity and each peptide is coloured by the MHC allele presenting it according to the colouring in panel **(B)**. **(D)** Comparison of V gene usage between the structure dataset and a background of TCRs taken from OTS. **(E)** Comparison of CDR lengths between the structure dataset and a background of TCRs taken from OTS. **(F)** Dataset coverage of canonical loop clusters from the whole STCRDab.

10.3389/fimmu.2024.1491656

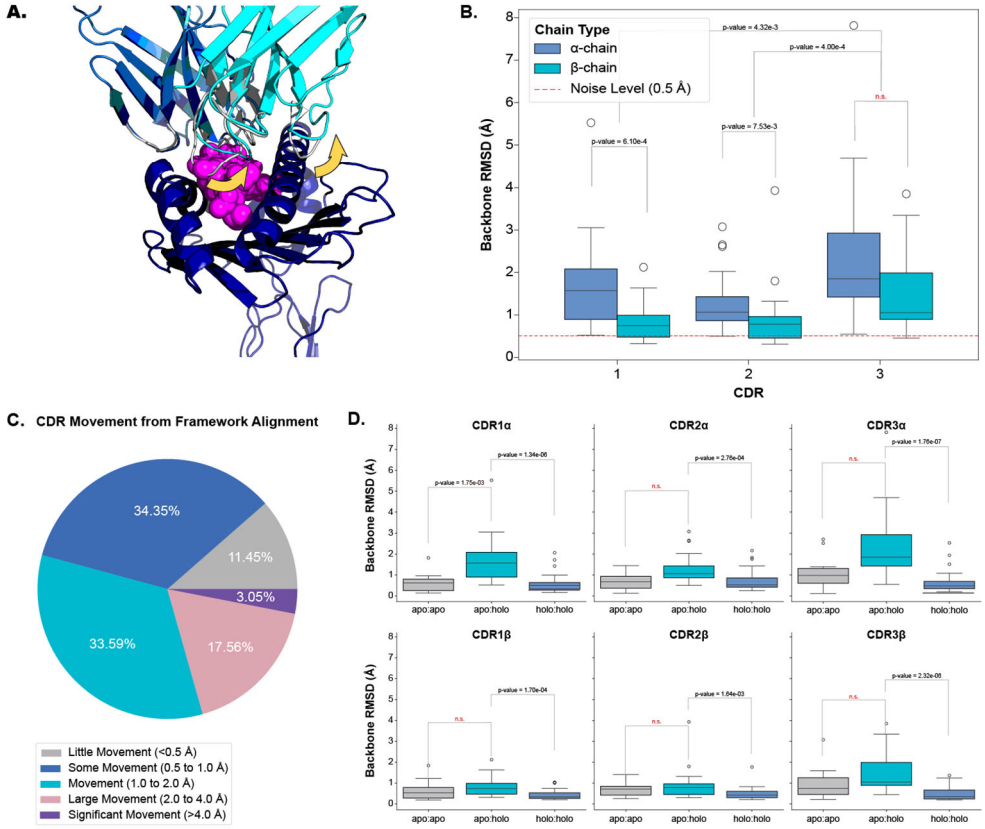

FIGURE 2
Quantifying the movement of each CDR loop after alignment on the TCR framework regions. **(A)** An example of CDR movement between the *apo* PDB ID 4jfh TCR (grey) and the *holo* PDB ID 4jfd TCR:pMHC-I structure (blue and cyan) as denoted by the yellow arrows. **(B)** Comparison of each loop between *apo* and *holo* states. There is an overall significant difference in the amount of movement each loop undergoes based on the Kruskal-Wallis test (p-value of 1.20×10^{-6}; significance level < 0.05). **(C)** Percentage of different movement categories from all CDRs. **(D)** Comparison of different movement types for each CDR loop. *apo:apo* refers to changes between different *apo* structures of the same TCR (10 TCRs), *apo:holo* refers to changes between *apo* and *holo* structures (22 TCRs), and *holo:holo* refers to changes between different *holo* structures of the same TCR (30 TCRs). There are significant differences between the movement types based on a p-value of 1.28×10^{-20} from a Kruskal-Wallis test (significance level < 0.05).

a standard deviation (SD) of 1.12 Å, 0.72 Å, and 1.67 Å. Similarly, the CDR1β, CDR2β, and CDR3β, loops have a mean change of 0.82 Å, 0.91 Å, and 1.50 Å respectively, with a of SD 0.45 Å, 0.77 Å, and 0.98 Å. Performing a Kruskal-Wallis test reveals a significant difference between the amount of conformation each loop undergoes (p-value of 1.20×10^{-6} at a significance level of 0.05). *Post hoc* Wilcoxon rank-sum tests with adjusted significance levels using a Bonferroni correction show that the α-chain moves more than the β-chain for the CDR1 and CDR2 loops but the same significance could not be determined for the CDR3 loops although the mean movement is higher in the CDR3α compared to CDR3β. When considering both chains together, our *post hoc* tests indicate that the CDR3 loops move more than both the CDR1 and CDR2 loops.

To ascertain whether these movements are the result of an engagement between the molecules or noise between different crystal structures, we compared the changes between *apo* and *holo* structures (22 TCRs) to those between different *apo* structures (10 TCRs) and between different *holo* structures (30 TCRs) of the same TCR. Figure 2D illustrates that there is a difference between these comparisons and performing a Kruskal-Wallis test on the comparisons yields significant results at a significance level of 0.05 (the p-value of the test is 1.28×10^{-20}). Performing *post hoc* Wilcoxon rank-sum tests with adjusted significance levels using the Bonferroni correction shows that the changes between *apo* and *holo* structures are significantly higher than between the changes of different *holo* structures for all six loop types (our baseline used in this analysis). Although an increased trend was observed, there was not enough statistical power to distinguish the changes between *apo* structures of the same TCR and the changes between the *apo* and *holo* structures for all loops except CDR1α loops.

We further investigated whether a TCR could exhibit both flexible and rigid behaviours depending on the pMHC-I context. We found that there were no cases where TCRs had larger than 1.0 Å backbone RMSD and smaller than 0.5 Å backbone RMSD between *apo* and *holo* conformations. This indicates that TCRs may operate in only one mode of protein-protein interaction.

10.3389/fimmu.2024.1491656

Our analysis shows that all six CDR loops undergo conformational change between *apo* and *holo* states and that there is some variation in the amount of change between loop types.

2.2 Only CDR3 loops deform when binding pMHC-Is

The results of the previous section encompass two different types of movements in the TCR: (1) bulk movements driven by changes in the anchors of CDR loops relative to the framework region and (2) loop flexibility driven by deformation in the CDR loops themselves. Again, anecdotal reports in the literature suggest differing views on the amount of flexibility CDR loops have. In some cases it has been shown that the CDR loops maintain a rigid body-like state, keeping their general shape while engaging with pMHC-I molecules (11). Other studies suggest that all CDR loops have some type of plasticity, disrupting their canonical forms, when engaging with pMHC-I molecules (9). In antibodies, non-CDR-H3 loops have practically no backbone deformation, and only a small amount of deformation is seen in CDR-H3 loops (5).

Here, we investigate the flexibility of TCR CDR loops to classify each loop as a rigid body, where all movement is the result of bulk movements, or as plastic entities, where flexibility also adds to the overall movement of the loop. We investigate this in several different modalities to establish robust descriptions of the deformation (Figures 3A–D).

In the first instance, we perform the same comparison of backbone RMSD of all CDR loops (see Section 2.1) but this time, we superimpose the *apo* and *holo* forms of each loop before measuring the differences to consider only the movement in the loops themselves and not bulk movements resulting from the framework region. Figure 3B depicts the quantification of these measurements. What becomes apparent is that only the CDR3α and CDR3β loops have their medians above the noise level line annotated on the figures, showing that these are the only significantly flexible loops. These findings are supported by *post hoc* Wilcoxon rank-sum tests that show only the differences between apo-holo comparisons and the holo-holo background for the CDR3α and CDR3β are significant (see Supplementary Figure S4).

In the second case, we look at the heavy atom RMSD between *apo* and *holo* states of each residue in the loops. Again, the *apo* and *holo* forms of the loops are superimposed before measuring the differences between them. The CDR loops are numbered using the IMGT numbering system so that each residue number is structurally equivalent between different loops and loop lengths (19). Figure 3C shows that the CDR1 and CDR2 loops (from both the α- and β-chain) have uniform-like profiles where every residue moves about the same amount, whereas the CDR3α and CDR3β loops have normal-like distributions meaning the middle residues move more than those towards the anchors of the loops.

Further, we analyse canonical clustering for each loop type and ascertain whether each kind of loop is moving from its canonical

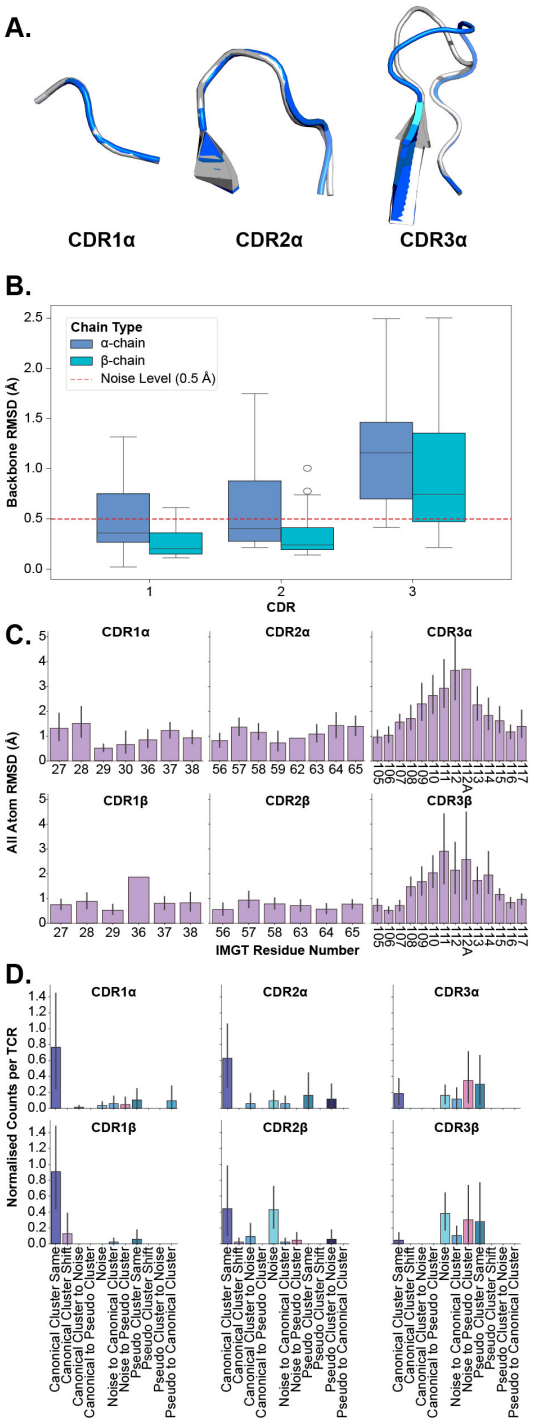

FIGURE 3
Examining the deformation of each CDR loop after superposition of *apo* and *holo* forms. **(A)** An example of loop deformation in the α-chain CDR loops between the *apo* (grey; PDB ID 2bnu) and *holo* (blue; PDB ID 2bnr) states. **(B)** Comparison of each loop between *apo* and *holo* states after aligning loops. **(C)** Per-residue RMSD changes of each CDR loop after aligning loops. The CDR loops are numbered following the IMGT numbering system to account for loops of different lengths. There are cases where only one TCR has a certain IMGT residue and thus no error bars are on those bars. **(D)** Normalised counts of shifting cluster types between *apo* and *holo* states for each type of loop.

cluster group or remaining in the same structural cluster (see Section 4.3 for a description of the annotations). Figure 3D shows that for CDR1 and CDR2 loops on both the α- and β-chain, the mode is to stay in the same canonical cluster between *apo* and *holo* states. For CDR3α and CDR3β, there is a larger variety of modes, including remaining un-clustered as noise, only forming canonical clusters in the *holo* state, and only clustering into pseudo clusters that remain between *apo* and *holo* states. The counts of these plots are listed in Table 1.

We also look at the changes in backbone dihedral angles using the D-score metric defined in Section 4.2. The results of this analysis, shown in Supplementary Figure S5, show that only CDR3 loops exhibit backbone flexibility between the *apo* and *holo* conformational states and the background of different *holo* states.

Figure 3A illustrates these conclusions, where CDR1α and CDR2α look nearly identical between *apo* and *holo* states, but CDR3α has large changes in the middle of the loop. These results show that in general the CDR1 and CDR2 loops act as rigid bodies, deforming very little when engaging with pMHC-Is, but the CDR3 loops undergo plastic deformation to enable the interactions with the pMHC-I molecule.

2.3 Identifying TCR contact points on pMHC-Is

The CDR loops have been well-established as the pMHC-I binding portion of the TCR (1), but the equivalent relevant binding portion of the pMHC-I has been less studied collectively across many different structures. Thus, we set out to map out the interacting regions from the pMHC-I perspective to conduct further analyses. As described in Section 4.5, we mapped the contacts of TCR CDR loops onto the pMHC-I surface. Figure 4A depicts these contacts on the surface of the MHC molecule and the peptide.

When looking at the loops contacting the peptide (Figure 4B), there is a trend that the CDR3α forms the majority of the contacts with the first half of peptides (p1-p5) and that the CDR3β forms majority of the contacts in the second half of the peptide (p6-p9). This description of loop engagement is in support of past literature that, despite its lack of D-gene segments, the CDR3α loop has a large structural diversity (20) and is important in determining the specificity of a TCR (21–24). These results also show that other loops, such as CDR1α and CDR1β, can form part of the peptide binding interface as has been previously documented (25). These trends extend to other non-nonamer peptides shown in Supplementary Figure S7.

Through this analysis, we were also able to identify the fingerprint of TCR interactions on MHC molecules and determine specific residue positions that correspond to the TCR contact points of the MHC. The conserved binding mode of TCRs on pMHC-Is has been well described in past literature (1, 25). Figure 4A displays the dominant loop contacting each residue and Figure 4C shows the percentage of contacts each loop makes with the MHC residues by IMGT number.

Our generated contact maps support that both TCR chains are equally involved in contacting the peptide and each chain has a preferred side of the peptide. The contact maps also provide an alternative means to quantify the diagonal binding mode TCRs take across the MHC.

2.4 Peptides undergo varying amounts of conformational change based on MHC anchor locations

Establishing that TCR CDR loops undergo conformational changes when binding pMHC-Is, we set out to investigate the changes pMHC-Is undergo between *apo* and *holo* states on the other side of the complex. Figure 5A shows that the majority of the conformational changes happen in the peptide, and that the MHC molecule undergoes little structural movements between *apo* and *holo* forms. When comparing the peptide, MHC TCR contact positions (as

TABLE 1 Numerical counts of the different types of cluster shifts undertaken by each loop type between *apo* and *holo* states.

Movement Type	CDR1α	CDR2α	CDR3α	CDR1β	CDR2β	CDR3β
Canonical to Pseudo Cluster	–	–	–	–	–	–
Canonical Cluster to Noise	1	5	–	–	8	–
Canonical Cluster Same	66	54	16	78	38	4
Canonical Cluster Shift	–	–	–	11	2	–
Noise	3	8	14	–	37	33
Noise to Canonical Cluster	5	5	10	2	2	9
Noise to Pseudo Cluster	4	–	30	–	4	26
Pseudo to Canonical	8	–	–	–	–	–
Pseudo Cluster to Noise	–	10	–	–	5	–
Pseudo Cluster Same	9	14	26	5	–	24
Pseudo Cluster Shift	–	–	–	–	–	–

10.3389/fimmu.2024.1491656

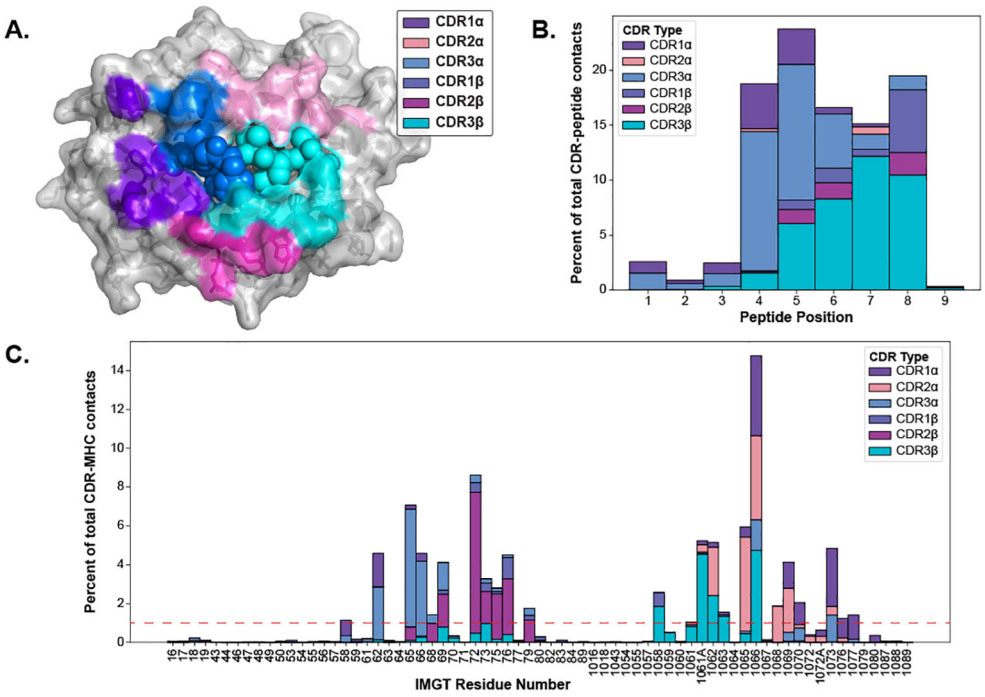

FIGURE 4
Visualizing the contacts made between TCR and pMHC-Is. **(A)** The contacts made between TCR CDR loops and pMHC-Is. Here, all of the contacts (< 5 Å between heavy atoms) that make up over 1% of an MHC residue contact (denoted by the red line in panel *C*) for all of the TCRs bound to pMHC-Is in the STCRDab (15) are mapped to a reference MHC molecule (PDB ID 1hhi). Notably, there are no residues dominantly contacted by CDR1β at this threshold. **(B)** Distribution of contacts made between CDR loops and nonamer-peptides from the STCRDab structures. **(C)** Distribution of contacts made between CDR loops and MHC molecules from the STCRDab structures.

described in Section 2.3), and non-TCR contact positions on the antigen binding domains of the MHC I molecules, there is significantly increased movement of the peptide. Performing a Kruskal-Wallis test shows significant differences (at a 0.05 significance level) between the regions, and *post hoc* Wilcoxon rank-sum tests with Bonferroni corrections show the peptide has an increased movement over all MHC domains. There is no significant difference between the TCR contact positions and other positions in the antigen binding domains not involved directly in TCR interaction.

Investigating the movement in peptides further, we found that the anchoring strategy employed by the MHC molecule dictates how a peptide will move when engaged by a TCR. We describe our procedure for identifying MHC anchors in Section 4.6. Figure 5B illustrates two different anchoring modes, and the results on the peptide conformational changes between *apo* and *holo* states. These plots show, that when a peptide is anchored in positions p2 and p9, as is common with the allele-specific motifs that comprise the majority of the structural data, the distribution of movement is uni-modal-like where the farthest points from anchors in the middle of the peptide have the largest RMSD changes. However, when peptides have an additional anchor in the p5 position, the distribution becomes bi-modal-like, meaning the movement is

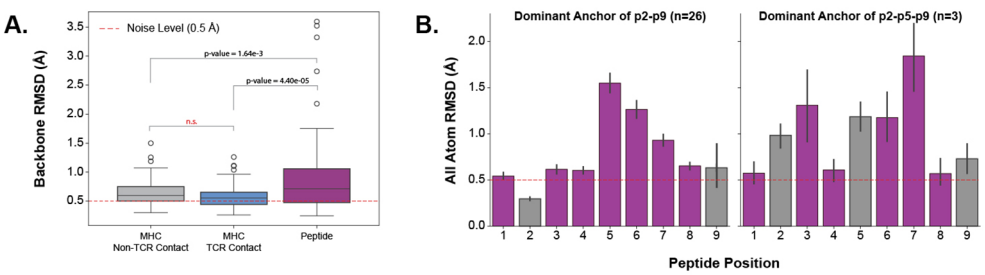

FIGURE 5
Quantifying the movement of pMHC-Is between *apo* and *holo* states. **(A)** Comparison of different parts of the pMHC-I complex between *apo* and *holo* states. There is a significant difference between these components as reported by a p-value of 3.32×10^{-6} from a one-way ANOVA test. **(B)** Effects of different peptide anchoring strategies on the deformation pattern of each peptide residue between *apo* and *holo* states.

restricted in the middle of the peptide and compensated for by freedom between anchoring residues.

These results show that the peptide is the component of the pMHC-I interface that undergoes the most change between *apo* and *holo* states and that these changes depend on how the peptide is anchored in the MHC binding groove.

3 Discussion

In this analysis, we have demonstrated the conformational changes occurring between interacting αbTCR and pMHC-I molecules. In particular, we showed that all six CDR loops undergo statistically significant movement between *apo* and *holo* states, but only the CDR3α and CDR3β loops have plasticity when engaging with pMHC-Is. We also map the interactions between TCRs and pMHC-Is, highlighting the importance of CDR3α and CDR3β in making contacts with the peptides. Finally, we show that the peptides also undergo conformational changes but these changes are restricted by the way the peptides are anchored in the MHC binding groove. Our results generalise the phenomena anecdotally reported in past literature to all TCR:pMHC-I interactions, yield new insights into the TCR:pMHC-I binding, and provide considerations for modelling of TCR:pMHC-I structures and T cell antigen specificity.

This work solely focussed on understanding the interactions between TCRs and pMHC-Is, as opposed to other types of TCR interactions with presentation molecules, due to the higher availability of data and tools. We speculate that we would observe the same trends in CDR movements between *apo* and *holo* states of TCR and class II peptide-MHCs (pMHC-IIs) as the coverage of TCR genes, CDR lengths, and canonical classes was comparable to OTS (see Section 2). Contrarily, we suspect there will be less conformational change in the core of MHC class II bound peptides as the peptide is chelated by a set of hydrogen bonds between the backbone and conserved residues in the MHC class II antigen binding domains, leading to a more rigid and linear structure for the peptide that is highly similar across structures of different alleles and peptides (1). In contrast, MHC class I peptides are mainly held by hydrogen bonds to their termini, with the centre of the peptide bulging to accommodate different-length peptides. As the resources available for studying TCR:pMHC-II and other non-conventional TCR interactions (for example interactions with CD1 and MR1 molecules) become more abundant, the analysis framework of this work can be applied to study these systems.

With recent advances in machine learning applied to protein structure prediction (26), many new approaches have been developed to predict the structures of adaptive immune receptors from sequence to allow repertoire scale analysis of immune receptor structures (27). Our results show important considerations for these structural prediction methods that currently do little to consider differences in the conformation of immune receptor CDRs. Here, we show that not only is there a significant difference between the *apo* and *holo* states of TCRs but also that there is movement between different *apo* states of the same TCR (see Figure 2D). These differences may contribute to the "poor" performance of structure prediction methods at predicting CDR loops (28). Further, in Figure 2B, we show that the α-chain has a significantly larger movement than the β-chain CDR1 and CDR2 loops and although there is not enough statistical power to support this, the CDR3α loop has a larger change than the CDR3β loops. This fits with the narrative of recent reports of larger structural diversity in the CDR3α than the CDR3β when predicting repertoires of TCR structures (20) and results of benchmarking studies that show the CDR3α loops are harder to predict than the CDR3β loops (28).

Other works have taken similar approaches to ours in mapping out either TCR contacts with peptides (29) or TCR contacts with class I MHCs (30). These contact maps provide useful CDR loop-specific pseudosequences of the MHC surface that contact TCRs. Training predictive models on these structurally constrained sequences may improve the ongoing challenge of predicting T cell antigen specificity (31). Other works have illustrated the value of this type of information to model the MHC restriction of TCR repertoires (32).

We compared our conclusions with recent literature conducting a similar study in antibodies (5). Antibodies have similar variable region structures to TCRs with two chains (heavy, H, and light, L) and six CDR loops forming the antigen binding domain of the molecule. However, these immune cell receptors target protein antigens directly, without the need for peptide linearisation and presentation. Liu et al. report that the conformational changes are confined to the CDR-H3 loops, both in large-scale movements and plasticity, the other five loops undergo very little conformational change. Liu et al.'s results corroborate past literature on how antibody loops behave, where the CDR-H3 loop drives specificity and the other CDR loops act as stabilisation for the H3 loop (33). In contrast, our work shows that both CDR3α and CDR3β have peptide-directed movement and plasticity, and the other CDRs also have movement that may or may not be involved in the stabilisation of CDR3 loops. This difference between receptor types may result from the fundamental differences between TCR and antibody binding of antigens (23). However, we draw similar conclusions to Liu et al. that the CDR1 and CDR2 loops are mostly found in the same canonical cluster group between *apo* and *holo* states.

Through this work, we have focused on understanding the dynamics of TCR:pMHC-I interactions, but counterintuitively, we have used static structures that offer only a "snapshot" of TCR state to infer these dynamics. The reason for this is the availability of data from static X-ray crystallography is much higher than any alternatives for TCR:pMHC-I interactions. There is only one solution nuclear magnetic resonance (NMR) structure and only eight structures resolved from cryogenic electron microscopy (cryo-EM) in the STCRDab (15) with little comparison between *apo* and *holo* states. These alternative structural biology techniques may provide the field with more ways to study the dynamics of TCR:pMHC-I interactions as they are conducted in more native-like solution environments. The effects of crystal packing on the *apo* conformations in X-ray crystallography data have been discussed in other works and possibly contribute to differences in conformational states between *apo* structures (34). Using solution NMR, Hawse et al. (35) showed the dynamics of the CDR3β loop of

a TCR after binding to pMHC-I and highlight the overall reduction of motility of the other CDR binding loops. Other experimental techniques such as hydrogen-deuterium exchange have been used in past works to measure pMHC-I flexibility (36–38) as well as the rigidification of TCRs when binding pMHC-I (39). We thus hope that our investigation of TCR:pMHC-I conformational changes further prompts the study and data collection of the dynamics of these interactions.

Outside of experimental approaches, molecular dynamics (MD) simulations offer a computational alternative for exploring the dynamics of protein interactions. Several groups have sought to understand similar questions regarding the dynamics of TCR:pMHC-I interactions using these simulations. Tripathi and Salunke explored the conformational changes of the IG4 TCR in complex with the tumour epitope NY-ESO peptide (SLLMWITQC) (40). Their work closely supports the results here, showing that most of the conformational change occurs in the CDR3 loops and that the CDR loops of the β-chain move less than those of the α-chain. They state that these findings support the paradigm of induced fit occurring between TCRs and pMHC-Is. Other groups have focused on the conformational changes of pMHC-Is when engaged by TCRs. MD simulations show that the peptide and MHC molecule greatly affect each other's flexibility. Hawse et al. show that the peptide amino acid composition modulates the MHC molecules flexibility (36) and Pöhlmann et al. (41) show that MHC polymorphisms affect the flexibility of peptides, both having implications in TCR specificity (42). Many works show that both the TCR and pMHC-I undergo a rigidification after binding to one another (43–45) and correspondingly that the CDR loops, peptide, and MHC α2-helix are more motile in the *apo* state providing evidence for pre-existing equilibrium binding between these molecules (46). These results would be difficult to validate with X-ray crystallography data alone as it is challenging to capture the full range of protein motion from these snapshots. Finally, Alba et al. illustrate the uniqueness of TCR:pMHC-I interactions through MD simulations, showing the hydrogen bonding and conformational effects of a peptide are unique to the interacting TCR (47). These help validate our lack of statistical significance between comparisons of *apo* and *holo* pMHC-Is and comparisons of *holo* pMHC-I forms with different TCRs (see Supplementary Figure S8) as, based on Alba et al., there is an expected heterogeneity in *holo* pMHC-I conformations. Overall, past work using MD simulations closely supports our results and provides more insight into the dynamics of TCR:pMHC-I interactions.

The introduction outlined three paradigms for protein-protein interactions: "lock-and-key", "induced-fit", and "pre-existing equilibrium". Here, we have provided evidence that conformational change is a key characteristic of TCR:pMHC-I molecules, meaning both "induced-fit" and "pre-existing equilibrium" seem plausible to describe the interactions. To discriminate between these modes, further work would need to be done. "Pre-existing equilibrium" could be determined by studying the unbound forms of these molecules and seeing if the *holo* states of these molecules are found within the range of *apo* forms. With limited data on this, further data collection using techniques such as NMR and cryo-EM of *apo* TCRs or pMHC-Is would be essential. In both "induced-fit" and "pre-existing equilibrium" paradigms, the interactions incur an energetic penalty as some structural rearrangements would be required for an interaction to occur. However, the flexibility of proteins can be seen as an evolutionary advantage as it allows for broader specificity of interactions from fewer stochastic sequence rearrangements events (4, 48). Thus, the conformational changes observed here support the growing evidence that not only does our immune system rely on sequence diversity for protection from harm-causing pathogens, but structural diversity also plays a role (20).

A limitation throughout this study is the availability of non-redundant crystal structures and representation to the αβTCR and pMHC-I repertoire as a whole. Although there were several hundred structures in the dataset, careful normalisation was required to not bias the results towards overly representative TCRs or pMHC-Is. Performing more crystallography work to increase the size and diversity of the available datasets would be an obvious solution, but a highly resource-intensive endeavour. In earlier work, we have discussed the ability of machine learning models for structure predictions in overcoming some of the challenges in producing large amounts of structural data for a broader analysis of TCR and pMHC-I repertoires (27). We speculate that this may be a promising way to overcome the biases and limitations of currently available crystallography data in the TCR:pMHC-I field, however, others have reported that these methods have limitations in predicting novel CDR shapes (49).

4 Methods

4.1 Curating *apo* and *holo* TCR and pMHC-I structures

The *apo* TCR structures and the *holo* TCR:pMHC-Is were collected from the STCRDab (15) and the *apo* pMHC-Is structures were collected from histo.fyi (16) on April 17th, 2024. The structures were all unbound αβTCRs, unbound pMHC-Is, or αβTCRs bound to pMHC-Is. Structures with a resolution greater than 3.50 Å or missing residues in the TCR CDR loops or the peptide were removed from the dataset. The unbound TCRs and pMHC-Is were matched to the TCR:pMHC-I complexes based on the TCR's CDR sequences or the pMHC-I's peptide sequence and allele name. Only data points with both an *apo* and *holo* form were kept in the dataset. The resulting dataset contains 391 structures coming from 301 PDB entries and encompassing 22 unique TCRs, 19 MHC alleles, and 79 peptides. The exact structures are listed in Supplementary Table S1 and are grouped by TCR and pMHC-I in Supplementary Tables S2, S3 respectively.

All selected structures were renumbered using the same version of ANARCI (50) to provide consistent IMGT numbering. For every *holo* TCR:pMHC-I complex, the *apo* TCRs were aligned to the *holo* TCRs on their framework regions and the *apo* pMHC-Is were aligned to the *holo* pMHC-Is on the strands forming the floor of the MHC binding groove. For some of the calculations, *holo* structures

were aligned to each other in the same manner where one of the *holo* forms is aligned on either the TCR framework region or the floor of the MHC binding groove. This created the final datasets used for calculations in the rest of the analysis and these aligned structures are provided as part of the provided code (see *Data Availability Statement*).

4.2 Measuring conformational changes between states

The difference in states was measured using RMSD throughout the analysis. RMSD is defined as:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} \quad (1)$$

Where d_i is the distance between atoms for all atoms being compared. When larger entities were being compared, for example, the entire CDR loops, the peptide, or parts of the MHC antigen binding domain, the measure was taken using the backbone atoms (N, C_α, C, O) of these entities. When comparing residues to one another, the RMSD was taken for all heavy atoms (non-hydrogen) to include information about the amino-acid side chains. Alternative measures including measuring the difference of the centre of mass of heavy atoms in each entity, the distance between C_α positions, the difference in χ-angles, and D-scores were used in other plots seen in the Supplementary Information (see Supplementary Figures S5, S6). The D-scores, used in other similar analysis by North et al. and Gupta et al. (51, 52), are defined as:

$$d_\theta(\theta_1, \theta_2) = 2(1 - \cos(\theta_1 - \theta_2)) \quad (2)$$

$$\text{D-score}(A, B) = \sum_{i}^{n} (d(\phi_i^A, \phi_i^B) + d(\psi_i^A, \psi_i^B)) \quad (3)$$

The score combines changes in both backbone dihedral angles. For all analyses, the results were normalised by the type of TCR or pMHC-I being compared. This was done by taking the mean of all results for that entity type before plotting or computing statistics so that over-represented TCR or pMHC-I structures did not bias the results.

4.3 Clustering CDR loops based on structure

Clustering of loop structures and annotation of canonical forms was done similar to previous works by Wong et al. and Greenshields-Watson et al. (49, 53). All CDR loops of the αβTCRs with a resolution of 3.50 Å and below were taken from the STCRDab (15). For every loop of the same type (CDR1α, CDR2α, CDR3α, CDR1β, CDR2β, CDR3β), a pairwise distance matrix was created following the same procedure: the two loops being compared were aligned on the backbone of the five anchor residues flanking each side (N and C termini) of the loops. The distance between their backbones was computed using the length-independent distance measure of dynamic time warping (DTW) (54). The HDBScan clustering algorithm (55) with a minimum cluster size cutoff of five was used on the distance matrices to create clusters of similar loop structures. Clusters were assigned as canonical clusters if they contained more than two unique sequences, otherwise, they were labelled as pseudo clusters following the approach of Wong et al. (53).

4.4 Sampling TCRs from OTS

We sampled OTS (17) to generate a background of TCR sequences to which we could compare the structures used in this analysis. All of OTS was downloaded on July 23rd, 2024. We selected all six CDR sequences and the TRAV, TRBV, TRAJ, and TRBJ gene identities based on the assigned call and removed redundant entries. The resulting dataset was then uniformly sampled 10 times, each containing 1000 unique TCRs. This created a dataset of representative TCRs that we could compare against our selected structures.

4.5 Mapping contacts between TCRs and pMHC-Is

The contacts made between TCRs and pMHC-Is were determined by finding all heavy atoms (nonhydrogen) that were less than 5 Å apart between the two structures. All αβTCRs contacting pMHC-Is presenting peptide antigens with a resolution under (and including) 3.50 Å in the STCRDab (as of April 2024) (15) were considered.

4.6 Incorporating MHC Motif Atlas data to identify MHC peptide anchors

To determine the anchoring strategy of each pMHC-I, we annotated our datasets with information from the MHCMotifAtlas (56). For each MHC allele in the MHCMotifAtlas, we created a simplified peptide motif based on the proportion of amino acid usage at each peptide position. These amino acids were classified as "dominant" residues if they were at over 60% of the observed amino acids in that position, "high" if they were between 30% and 60% of the observed amino acids, "medium" if they were between 20% and 30% of the observed amino acids, "low" if they were between 10% and 20% of the observed amino acids, and "very-low" if they were below 10% of the observed amino acids. Where there were "dominant" or "high" amino acids, it was assumed that these residues corresponded to residues necessary to anchor the peptide in the binding groove. Using these anchor residues, we annotated our dataset with anchors at the positions where peptide residues matched the motif anchor residues, the peptide lengths matched the observed simplified motif lengths, and the MHC alleles were the same. These annotations created two distinct groups corresponding to MHC alleles that anchor the peptide in positions 2 and 9 (p2-p9) or anchor at positions 2, 5, and 9 (p2-p5-p9). Where only anchors at

position 2 or position 9 were found (a minor subset of the data), we added these to the p2-p9 group as it was assumed that the peptide would be anchored by another amino acid type at the missing end.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors. The link to our code and repository is https://github.com/benjiemc/tcr-pmhc-interface-analysis.

Author contributions

BM: Conceptualization, Data curation, Formal analysis, Visualization, Software, Writing – original draft, Writing – review & editing. CJT: Conceptualization, Data curation, Formal analysis, Supervision, Writing – review & editing. JR: Supervision, Writing – review & editing. CMD: Supervision, Methodology, Writing – review & editing. HK: Conceptualization, Supervision, Writing – review & editing, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by funding from the UK Medical Research Council grant number MC UU 12010/3 to HK, the UK Medical Research Council grant number MC UU 00008 to BM, an ARISE Fellowship from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement number 945405 to CT. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

Acknowledgments

We would like to thank A. Greenshields-Watson, M. Raybould, F. Spoendlin, and N. Quast for their insightful conversations and expertise in some technical implementations throughout the project. We would also like to thank the developers of PyMol, Pandas, NumPy, Matplotlib and Seaborn for their contributions to the open-source community. These tools and packages were used to conduct the analysis and generate the figures for this work.

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1491656/full#supplementary-material

References

1. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol.* (2006) 24:419–66. doi: 10.1146/annurev.immunol.23.021704.115658

2. Turner SJ, Doherty PC, McCluskey J, Rossjohn J. Structural determinants of T-cell receptor bias in immunity. *Nat Rev Immunol.* (2006) 6:883–94. doi: 10.1038/nri1977

3. Ali I, Yang WC. The functions of kinesin and kinesin-related proteins in eukaryotes. *Cell Adhes Migr.* (2020) 14:139–52. doi: 10.1080/19336918.2020.1810939

4. Goh CS, Milburn D, Gerstein M. Conformational changes associated with protein–protein interactions. *Curr Opin Struct Biol.* (2004) 14:104–9. doi: 10.1016/j.sbi.2004.01.005

5. Liu C, Denzler LM, Hood OE, Martin AC. Do antibody CDR loops change conformation upon binding? *mAbs.* (2024) 16:2322533. doi: 10.1080/19420862.2024.2322533

6. Garcia KC, Degano M, Pease LR, Huang M, Peterson PA, Teyton L, et al. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science.* (1998) 279:1166–72. doi: 10.1126/science.279.5354.1166

7. Reiser JB, Grégoire C, Darnault C, Mosser T, Guimezanes A, Schmitt-Verhulst AM, et al. A T cell receptor CDR3beta loop undergoes conformational changes of unprecedented magnitude upon binding to a peptide/MHC class I complex. *Immunity.* (2002) 16:345–54. doi: 10.1016/s1074-7613(02)00288-1

8. Kjer-Nielsen L, Clements CS, Brooks AG, Purcell AW, McCluskey J, Rossjohn J. The 1.5 A crystal structure of a highly selected antiviral T cell receptor provides evidence for a structural basis of immunodominance. *Struct (London England: 1993).* (2002) 10:1521–32. doi: 10.1016/s0969-2126(02)00878-x

9. Kjer-Nielsen L, Clements CS, Purcell AW, Brooks AG, Whisstock JC, Burrows SR, et al. A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity. *Immunity.* (2003) 18:53–64. doi: 10.1016/s1074-7613(02)00513-7

10. Reiser JB, Darnault C, Gregoire´ C, Mosser T, Mazza G, Kearney A, et al. CDR3 loop flexibility contributes to the degeneracy of TCR recognition. *Nat Immunol.* (2003) 4:241–7. doi: 10.1038/ni891

11. Chen JL, Stewart-Jones G, Bossi G, Lissin NM, Wooldridge L, Choi EML, et al. Structural and kinetic basis for heightened immunogenicity of T cell vaccines. *J Exp Med.* (2005) 201:1243–55. doi: 10.1084/jem.20042323

12. Tynan FE, Reid HH, Kjer-Nielsen L, Miles JJ, Wilce MCJ, Kostenko L, et al. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat Immunol.* (2007) 8:268–76. doi: 10.1038/ni1432

13. Krogsgaard M, Prado N, Adams EJ, He Xl, Chow DC, Wilson DB, et al. Evidence that Structural Rearrangements and/or Flexibility during TCR Binding Can Contribute to T Cell Activation. *Mol Cell.* (2003) 12:1367–78. doi: 10.1016/S1097-2765(03)00474-X

14. Armstrong KM, Piepenbrink KH, Baker BM. Conformational changes and flexibility in T-cell receptor recognition of peptide–MHC complexes. *Biochem J.* (2008) 415:183–96. doi: 10.1042/BJ20080850

15. Leem J, de Oliveira SH, Krawczyk K, Deane CM. STCRDab: the structural t-cell receptor database. *Mol Cell.* (2018) 46:D406–12. doi: 10.1093/nar/gkx971

16. Thorpe C. Histo.fyi — An interactive exploration of the structure and function of MHC molecules(2024). Available online at: https://www.histo.fyi/ (Accessed April 17, 2024).

17. Raybould MI, Greenshields-Watson A, Agarwal P, Aguilar-Sanjuan B, Olsen TH, Turnbull OM, et al. The observed t cell receptor space database enables paired-chain repertoire mining, coherence analysis and language modelling. *bioRxiv.* (2024) 43:114704. doi: 10.1101/2024.05.20.594960

18. Xue Z, Wu L, Tian R, Liu Z, Bai Y, Sun D, et al. Disease associated human TCR characterization by deep-learning framework TCR-DeepInsight. (2023). doi: 10.1101/2023.05.22.541406. Preprint.

19. Lefranc MP, 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

20. Quast NP, Abanades B, Guloglu B, Karuppiah V, Harper S, Raybould MIJ, et al. T-cell receptor structures and predictive models reveal comparable alpha and beta chain structural diversity despite differing genetic complexity. (2024). doi: 10.1101/2024.05.20.594940. Preprint.

21. Borg NA, Ely LK, Beddoe T, Macdonald WA, Reid HH, Clements CS, et al. The CDR3 regions of an immunodominant T cell receptor dictate the 'energetic landscape' of peptide-MHC recognition. *Nat Immunol.* (2005) 6:171–80. doi: 10.1038/ni1155

22. Coles CH, Mulvaney RM, Malla S, Walker A, Smith KJ, Lloyd A, et al. TCRs with distinct specificity profiles use different binding modes to engage an identical peptide–HLA complex. *J Immunol Author Choice.* (2020) 204:1943–53. doi: 10.4049/jimmunol.1900915

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

24. Henderson J, Nagano Y, Milighetti M, Tiffeau-Mayer A. Limits on inferring T-cell specificity from partial information. (2024). doi: 10.48550/arXiv.2404.12565. Preprint.

25. 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

26. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* (2021) 596:583–9. doi: 10.1038/s41586-021-03819-2

27. McMaster B, Thorpe C, Ogg G, Deane CM, Koohy H. Can AlphaFold's breakthrough in protein structure help decode the fundamental principles of adaptive cellular immunity? *Nat Methods.* (2024) 21:766–76. doi: 10.1038/s41592-024-02240-7

28. Abanades B, Wong WK, Boyles F, Georges G, Bujotzek A, Deane CM. ImmuneBuilder: Deep-learning models for predicting the structures of immune proteins. *Commun Biol.* (2023) 6:1–8. doi: 10.1038/s42003-023-04927-7

29. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature.* (2017) 547:94–8. doi: 10.1038/nature22976

30. Gupta S, Sgourakis NG. A structure-guided approach to predict MHC-I restriction of T cell receptors for public antigens. (2024). doi: 10.1101/2024.06.04.597418. Preprint.

31. Hudson D, Fernandes RA, Basham M, Ogg G, Koohy H. Can we predict T cell specificity with digital biology and machine learning? *Nat Rev Immunol.* (2023) 23:511–21. doi: 10.1038/s41577-023-00835-3

32. Ortega MR, Pogorelyy MV, Minervina AA, Thomas PG, Walczak AM, Mora T. Learning predictive signatures of HLA type from t-cell repertoires. (2024). doi: 10.1101/2024.01.25.577228

33. Guloglu B, Deane CM. Specific attributes of the VL domain influence both the structure and structural variability of CDR-H3 through steric effects. *Front Immunol.* (2023) 14:1223802. doi: 10.3389/fimmu.2023.1223802

34. Fernández-Quintero ML, Pomarici ND, Loeffler JR, Seidler CA, Liedl KR. T-cell receptor CDR3 loop conformations in solution shift the relative Vα-Vβ Domain distributions. *Front Immunol.* (2020) 11. doi: 10.3389/fimmu.2020.01440

35. Hawse WF, De S, Greenwood AI, Nicholson LK, Zajicek J, Kovrigin EL, et al. TCR scanning of peptide/MHC through complementary matching of receptor and ligand molecular flexibility. *J Immunol (Baltimore Md.: 1950).* (2014) 192:2885–91. doi: 10.4049/jimmunol.1302953

36. Hawse WF, Gloor BE, Ayres CM, Kho K, Nuter E, Baker BM. Peptide modulation of class I major histocompatibility complex protein molecular flexibility and the implications for immune recognition *. *Front Immunol.* (2013) 288:24372–81. doi: 10.1074/jbc.M113.490664

37. van Hateren A, Anderson M, Bailey A, Werner JM, Skipp P, Elliott T. Direct evidence for conformational dynamics in major histocompatibility complex class I molecules. *J Biol Chem.* (2017) 292:20255–69. doi: 10.1074/jbc.M117.809624

38. Fabian H, Huser H, Narzi D, Misselwitz R, Loll B, Ziegler A, et al. HLA-B27 subtypes differentially associated with disease exhibit conformational differences in solution. *J Mol Biol.* (2008) 376:798–810. doi: 10.1016/j.jmb.2007.12.009

39. Hawse WF, Champion MM, Joyce MV, Hellman LM, Hossain M, Ryan V, et al. Cutting edge: evidence for a dynamically driven T cell signaling mechanism. *J Immunol.* (2012) 188:5819–23. doi: 10.4049/jimmunol.1200952

40. Tripathi SK, Salunke DM. Exploring the different states of wild-type T-cell receptor and mutant conformational changes towards understanding the antigen recognition. *J Biomol Struct Dyn.* (2021) 39:188–201. doi: 10.1080/07391102.2019.1708795

41. Pöhlmann T, Böckmann RA, Grubmüller H, Uchanska-Ziegler B, Ziegler A, Alexiev U. Differential peptide dynamics is linked to major histocompatibility complex polymorphism *. *J Biol Chem.* (2004) 279:28197–201. doi: 10.1074/jbc.C400128200

42. Reboul CF, Meyer GR, Porebski BT, Borg NA, Buckle AM. Epitope flexibility and dynamic footprint revealed by molecular dynamics of a pMHC-TCR complex. *PloS Comput Biol.* (2012) 8:e1002404. doi: 10.1371/journal.pcbi.1002404

43. Knapp B, Deane CM. T-cell receptor binding affects the dynamics of the peptide/MHC-I complex. *J Chem Inf Model.* (2016) 56:46–53. doi: 10.1021/acs.jcim.5b00511

44. Tomasiak L, Karch R, Schreiner W. Conformational flexibility of a free and TCR-bound pMHC-I protein investigated by long-term molecular dynamics simulations. *BMC Immunol.* (2022) 23:36. doi: 10.1186/s12865-022-00510-7

45. Alba J, D'Abramo M. The full model of the pMHC-TCR-CD3 complex: A structural and dynamical characterization of bound and unbound states. *Cells.* (2022) 11:668. doi: 10.3390/cells11040668

46. Karch R, Stocsits C, Ilieva N, Schreiner W. Intramolecular domain movements of free and bound pMHC and TCR proteins: A molecular dynamics simulation study. *Cells.* (2019) 8:720. doi: 10.3390/cells8070720

47. Alba J, Di Rienzo L, Milanetti E, Acuto O, D'Abramo M. Molecular dynamics simulations reveal canonical conformations in different pMHC/TCR interactions. *Cells.* (2020) 9:942. doi: 10.3390/cells9040942

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

49. Greenshields-Watson A, Abanades B, Deane CM. Investigating the ability of deep learning-based structure prediction to extrapolate and/or enrich the set of antibody CDR canonical forms. *Front Immunol.* (2024) 15:1352703. doi: 10.3389/fimmu.2024.1352703

50. Dunbar J, Deane CM. ANARCI: Antigen receptor numbering and receptor classification. *Bioinformatics.* (2016) 32:298–300. doi: 10.1093/bioinformatics/btv552

51. North B, Lehmann A, Dunbrack RL. A new clustering of antibody CDR loop conformations. *J Mol Biol.* (2011) 406:228–56. doi: 10.1016/j.jmb.2010.10.030

52. Gupta S, Nerli S, Kutti Kandy S, Mersky GL, Sgourakis NG. HLA3DB: Comprehensive annotation of peptide/HLA complexes enables blind structure prediction of T cell epitopes. *Nat Comm.* (2023) 14:6349. doi: 10.1038/s41467-023-42163-z

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

54. Meert W, Hendrickx K, Van Craenendonck T, Robberechts P, Blockeel H, Davis J. Dtaidistance. (2022). doi: 10.5281/zenodo.7158824

55. McInnes L, Healy J, Astels S. hdbscan: Hierarchical density based clustering. *J Open Source Softw.* (2017) 2:205. doi: 10.21105/joss.00205

56. Tadros DM, Eggenschwiler S, Racle J, Gfeller D. The MHC Motif Atlas: A database of MHC binding specificities and ligands. *Nucleic Acids Res.* (2023) 51:D428–37. doi: 10.1093/nar/gkac965