JBC RESEARCH ARTICLE

Determining T-cell receptor binding orientation and Peptide-HLA interactions using cross-linking mass spectrometry

Received for publication, October 9, 2024, and in revised form, March 8, 2025 Published, Papers in Press, March 26, 2025,
https://doi.org/10.1016/j.jbc.2025.108445

Thomas Powell*, Vijaykumar Karuppiah, Saher Afshan Shaikh, Robert Pengelly, Nicole Mai, Keir Barnbrook, Amit Sharma, Stephen Harper, Martin Ebner, and Andrew J. Creese
From the Immunocore Limited, Abingdon, United Kingdom

Reviewed by members of the JBC Editorial Board. Edited by Robert Haltiwanger

T cell receptors (TCRs) recognize specific peptides presented by human leukocyte antigens (HLAs) on the surface of antigen-presenting cells and are involved in fighting pathogens and cancer surveillance. Canonical docking orientation of TCRs to their target peptide-HLAs (pHLAs) is essential for T cell activation, with reverse binding TCRs lacking functionality. TCR binding geometry and molecular interaction footprint with pHLAs are typically obtained by determining the crystal structure. Here, we describe the use of a cross-linking tandem mass spectrometry (XL-MS/MS) method to decipher the binding orientation of several TCRs to their target pHLAs. Cross-linking sites were localized to specific residues and their molecular interactions showed differentiation between TCRs binding in canonical or reverse orientations. Structural prediction and crystal structure determination of two TCR-pHLA complexes validated these findings. The XL-MS/MS method described herein offers a faster and simpler approach for elucidating TCR-pHLA binding orientation and interactions.

T cells play a crucial role in the immune response to multiple diseases, including fighting against cancer and infectious diseases (1, 2). T cell receptors (TCRs) are heterodimeric membrane-spanning proteins on the surface of T cells that recognize specific linear peptides in the context of major histocompatibility complex (MHC) (human leukocyte antigen (HLA) in humans) on antigen-presenting cells. The highly polymorphic class I HLA molecules bind to peptides (pHLA), derived from intracellular proteins, that are typically 8 to 13 amino acids long (3). The ability to target intercellular proteins *via* pHLA, and the identification of disease-specific peptide antigens through mass spectrometry, has opened new avenues for immunotherapeutic treatments.

In 2022, tebentafusp became the first TCR-based therapeutic to be approved by the FDA (4). Several other TCR-based investigational drugs, both as soluble bispecific molecules and as engineered T cells, have recently been developed and are currently in clinical trials for oncology indications targeting various antigens (5, 6).

TCRs bind to the pHLAs diagonally to enable optimal positioning of all six complementarity-determining regions (CDRs) on to the HLA surface, as well as ensuring full coverage of the target peptide. This binding geometry, which places the TCR alpha chain above the first half of the target peptide and the TCR beta chain above the second half of the peptide is defined as canonical orientation. In some circumstances, deviations of the canonical binding orientation have been reported and attributed to non-functional T-cell signaling (7–9). This suggests a link between permissible TCR binding geometries and T cell potency, with the molecular rules governing these interactions, dependencies on co-receptors, and signaling components starting to emerge (10).

Binding orientations of TCRs to pHLAs are typically discerned using crystal structures, which provide high-level molecular interaction details. However, this technique can be time-consuming and the bottleneck of the formation of suitable crystals hampers obtaining valuable information. Computational prediction of TCR-pHLA complex structures has been developed making use of the AlphaFold2 deep learning program with varying levels of accuracy (11). Alternatively, experimental determination of TCR binding orientation quickly and reliably is advantageous.

Covalent cross-linking hyphenated to tandem mass-spectrometry (XL-MS/MS) is effective at monitoring the higher order structure of proteins and protein complexes (12–15). Hetero-bifunctional cross-linkers contain two functional groups and form covalent bonds between proximal amino acids, and these fragments can be analyzed by LC-MS/MS (13).

Diazirene-based cross-linkers have recently been developed (16–18). In these studies, the cross-linking reaction is activated by UV light (19, 20). In contrast to other methods, photo cross-linkers are less restricted in the residues they can bind to. In the first stage of cross-linking, the amino group of the cross-linker readily reacts with a lysine side chain or protein N-terminus (or to a lesser extent, serine, threonine or tyrosine side chains). Subsequent UV activation of the diazirene moiety yields a highly reactive carbene intermediate that can react with either the side chain or backbone carbon of any other amino acid (21, 22). Sulfosuccinimidyl 4,4′-azipentanoate (sulfo-SDA) is a diazirene-based photo cross-linker and has

* For correspondence: Thomas Powell, thomas.powell@immunocore.com.

Crown Copyright © 2025 Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

recently been shown to be effective at generating high-density cross-linking data (15).

In this study, we present an XL-MS/MS method to study the orientation of TCRs while bound to their pHLA targets. The analyzed cross-linked fragments provide information on TCR-pHLA interactions, and broad TCR orientation on the pHLA surface when at least one fragment linkage was detected from each of the TCR chains. We show that this information was sufficient to differentiate between TCRs binding in canonical or reverse orientations as well as the shift along the pHLA if the footprint was not centrally focused. These results were corroborated using a combination of computationally predicted complex models and experimentally determined crystal structures. This higher throughput XL-MS/MS-based method provides a broad molecular fingerprint of TCR-pHLA binding orientation and has the potential to influence TCR selection and engineering for immunotherapy.

Results

Characteristics of selected TCRs

To evaluate the XL-MS/MS method for determining TCR and pHLA fragment linkages and to infer the binding orientation, five TCRs were selected, each with specificity towards a unique pHLA, covering a broad spectrum of receptors and antigens. All TCRs were affinity enhanced either by screening mutations in CDRs individually or in combination with 3 CDRs in a chain, using phage display. The selected TCRs, on average, had eight mutations. The affinity of the TCRs to their respective pHLAs, measured using surface plasmon reference at 25 °C, ranged from 0.106 to 2.7 nM, accounting for an average of 240,000-fold affinity enhancement compared to the wild-type (Table S1).

Prediction of TCR-pHLA complex structures

To map the cross-links identified from XL-MS/MS peptide analysis, we initiated structure prediction of the TCR-pHLA complexes using TCRmodel2 (11), which utilizes the Alphafold2 deep learning program (23), with customizations for multiple sequence alignments and template selections.

The top-ranked predicted binding orientation of each TCR-pHLA complex is shown in Figure 1. The TCRmodel2 confidence scores of the predictions ranged from 93 to 55 with the highest and lowest scores for TCR-B and TCR-E, respectively. TCRs A-D were predicted to bind their respective pHLAs canonically, with a broadly central pHLA footprint covering the full length of the peptide. Surprisingly, TCR-E was predicted to bind in the reverse orientation, *i.e.*, alpha chains and beta chains interacting with opposing regions of the HLA when compared to canonical binders. Although reverse binders have been noted before (24), the reliability of predicted models to capture these non-typical binding orientations was uncertain considering the limited experimentally determined training data. The crossing angle parameters of TCR binding orientations are listed in Table S2. TCRs A-D bound to their respective pHLAs within the typical canonical binding range of 30 to 80°, with TCR-E significantly outside of this.

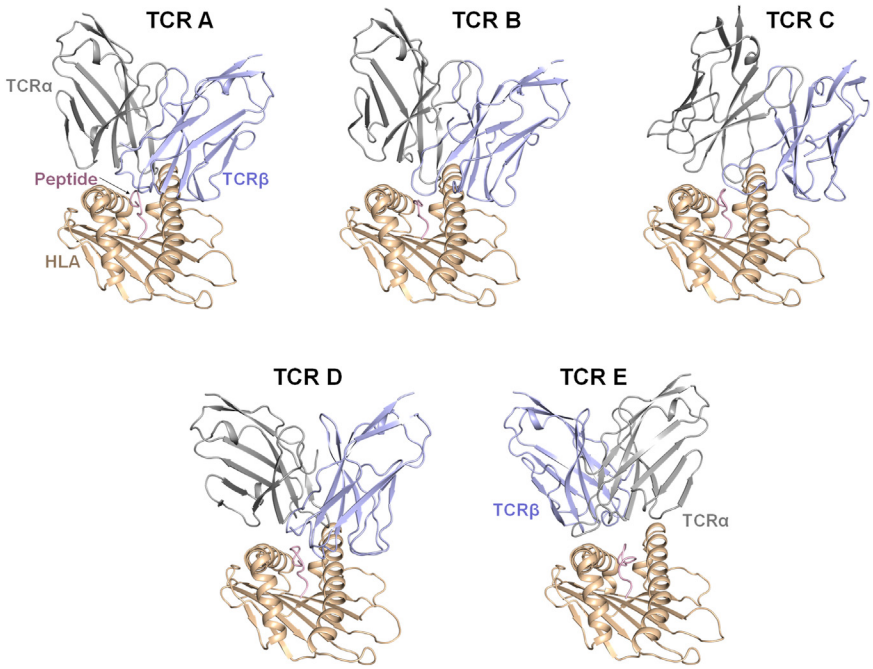

Figure 1. Predicted structures of TCR-pHLA complexes using the TCRmodel2 deep learning tool. Cartoon depictions of the HLA (wheat), peptide (*light pink*), TCRα (*gray*), and TCRβ (*blue*) are shown. TCRs A-D were predicted to bind their respective pHLA in the canonical orientation and TCR-E was predicted to bind in the reverse orientation. For TCR A-D complexes, the model with the highest confidence score (ranked 0) was shown, and for the TCR-E complex, the second highest confidence score model (ranked 1) was shown. The models were aligned using the HLA chain. The HLA α3 domain, β2m, and TCR constant domains were not shown in the models predicted by TCRmodel2.

Intermolecular cross-links reveal information about TCR-pHLA binding orientation

TCR-pHLA complexes were formed by mixing the components in an equimolar ratio and then cross-linked by adding sulfo-SDA and exposing them to UV light. The mixture was digested with trypsin and analyzed by LC-MS/MS. The data were studied for intermolecular cross-links between the TCR alpha or beta chain with either the HLA or target peptide. The positions of each identified intermolecular cross-link for each complex are listed in Table 1. For the TCR-A-pHLA complex, two sets of intermolecular cross-linked amino acids were identified. The first of the cross-links was predicted to be between residues TCR-Aα K68 and HLA E166. The ion containing these cross-linking sites was fragmented (Fig. 2*A* and Table S3) which revealed a partially maintained covalent cross-link. This is due to an additional fragmentation between the sulfo-SDA linker and the HLA E166. As a result, two possible b- and y-ion series are predicted for both peptides. A prominent y ion series was observed for TCR region LNASLDKSSGR. Ions corresponding to the theoretical masses of y_6, y_8, y_9, and y_{10} were noted, but with an additional mass of +82.0419. These ions were hypothesized to be a result of the additional fragmentation cleavage between the sulfo-SDA linker and HLA E166. The observation of the b_6 ion at *m/z* 614.3133 confirms the cross-linking site at TCR-Aα K68. Similarly, ions hypothesized to be a result of the fragmentation of HLA peptide AYLEGTCVEWLR were interrogated. The ion series running from y_4-y_{10}, as well as the absence of a y_3 ion at ∼ *m/z* 851.9, confirms cross-linking at HLA E166 (peptide position 9).

The second of the two reported cross-links in TCR-A was analyzed to be between TCR-Aβ D57 and HLA K68 (Fig. 2*B* and Table S4). In the predicted TCR-A-pHLA structure, the distances between TCR-Aα K68 (Nζ atom) and HLA E166 (Cα atom), and TCR-Aβ D57 (Cα atom) and HLA K68 (Nζ atom) were 6.3 Å and 8.5 Å respectively (Fig. 3), well within the expected cross-linking distances for sulfo-SDA (<∼15 Å) (25).

For the TCR-B-pHLA complex, like TCR-A, two sets of cross-linked amino acids were observed. The first cross-link was between TCR-Bα Y52 and HLA E154 (Fig. S1*A* and Table S5). The base peak in the spectrum results from the co-elution of a singly-charged ion of high abundance. Nevertheless, the comprehensive sequence coverage provided confidence in the peptide assignment and localization of modification sites. The second cross-link was between TCR-Bβ K58 and HLA K68 (Fig. S1*B* and Table S6). The distance measurements from the predicted structure for the two cross-links were 9.5 Å (TCR-Bα Y52 and HLA E154) and 9.9 Å (TCR-Bβ K58 and HLA K68) (Fig. 3).

Two sets of cross-linked amino acids were also identified for the TCR-C-pHLA complex (Fig. S1, *C* and *D*, and Tables S7, S8). Analysis of the fragment ions localized the first cross-link to be between TCR-Cα K53 and HLA E154. The distance measurements from the predicted structure for this cross-link was 9.9 Å. The second link was determined to be between TCR-Cα M1 and HLA E58. As the N-terminal methionine was

Table 1
Summary of cross-linked peptides obtained following tryptic digestion of TCRs incubated with pHLA

TCR peptide			pHLA/Target peptide			LC-MS reports		
Sequence	TCR	Xlink AA	Sequence	Xlink AA	Xlink distance (Å) (N^{ζ}-Cα)	*m/z*	RT	ΔPPM
LNASLD**K**SSGR	A	αK68	AYLEGTCV**E**WLR	E166	6.3 (predicted model)	909.1279	19.6	2.8
QILGQGPELLVQFQDESVVD**D**SQLPK	A	βD57	V**K**AHSQTHR	K68	8.5 (predicted model)	1007.5309	20.5	3.3
KGPELLMYT**Y**SSGNK	B	αY52	WEAAHVA**E**QLR	E154	9.5 (Cα-Cα, predicted model)	770.3891	18.5	−0.3
QTLGQGPEFLTYFQNEAQLE**K**SR	B	βK58	V**K**AHSQTHR	K68	9.9 (predicted model)	766.5958	18.0	2.1
FFIQGY**K**TNVSNEVASLFIPADR	C	αK53	WEAAHVA**E**QLR	E154	6.8 (crystal structure)	1336.3595	24.6	4.7
MLAK	C	αM1	APWIEQEGPE**Y**WDGETR	E58	N/A	869.4114	19.1	−1.4
MNAGVTQTPK	D	βM1	WEAA**H**VA**E**QLR	H151/A153	N/A	813.0814	15.7 and 16.5	0.1
FTAQLN**K**ASQYVSLLIR	D	αK68	**Y**LENGK	Y171	17.6 (predicted model)	919.5060	22.9	1.9
LIYYSASG**G**TTDK	D	βG54	Target Peptide	7	-	-	18.0	0.2
SPELIMSI**Y**ETSIK	E	αE53	SLAGGLDDM**K**A	K10	8.4 (crystal structure)	923.8029	24.8	0.2

Cross-linked amino acids (AA) are highlighted in bold.

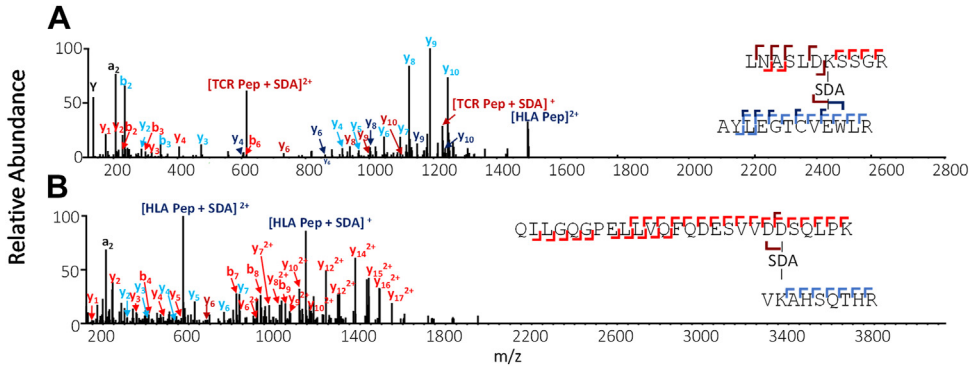

Figure 2. MS/MS spectra for TCR-A-pHLA cross-linked peptides identified at m/z (*A*) 1336.3595 and (*B*) 869.4114. b and y ions are shown in *red* or *blue* depending on their corresponding peptide. Ions resulting from the additional cleavage adjacent to the SDA cross-linker are shown in *dark red* or *dark blue*. Some less abundant ions are not labeled in the spectra. A full list of assignments is included in the relevant supplemental tables.

not included in the predicted model, distance measurement was not carried out for this linkage.

It is interesting to note that in both TCR-A and TCR-B complexes, HLA K68 was involved in forming the cross-links with the beta chain, indicating this lysine residue can play a crucial role in studying TCR-pHLA interactions using linkers. Similarly, HLA E154 was involved in cross-links with the alpha chain of TCR-B and TCR-C, showing preference by sulfo-SDA for this HLA residue. Distance measurements between the reactive groups of connected residues in the predicted TCR-pHLA structures were below 10 Å for the three TCR complexes, indicating this distance range was highly suited for mediating linkages by sulfo-SDA. Importantly, these cross-links can be formed only when TCRs A-C

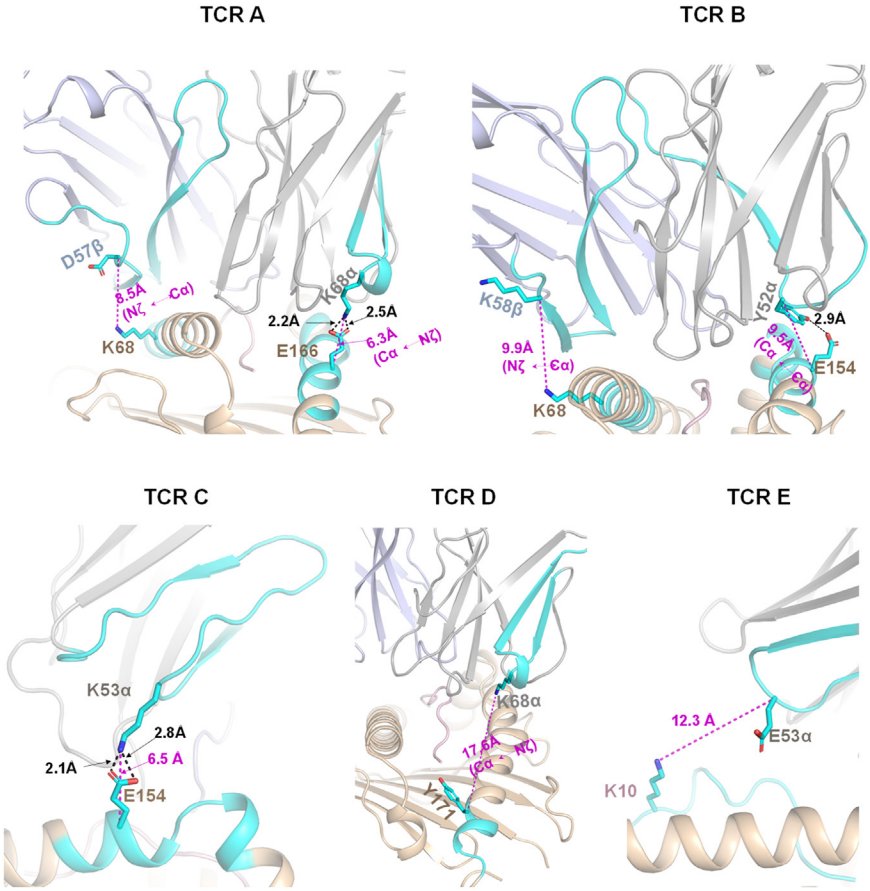

Figure 3. Analysis of MS linkage residues between TCR and pHLA in the TCRmodel2 predicted models. Linkage residues that were identified from MS peptides for TCR *A*, *B*, and *D* in complex with respective pHLA were shown in sticks. The HLA, peptide, TCRα, and TCRβ are shown as cartoons and are colored in *wheat*, *light pink*, *gray*, and *blue*, respectively. The *black* dotted lines indicate hydrogen bonds. The *magenta dotted line* shows the distance between the lysine side chain Nζ atom or backbone Cα atom of the reactive residue of one chain and the backbone Cα atom of the reactive residue of the nearby chain. The region identified in MS peptides is colored in *cyan*.

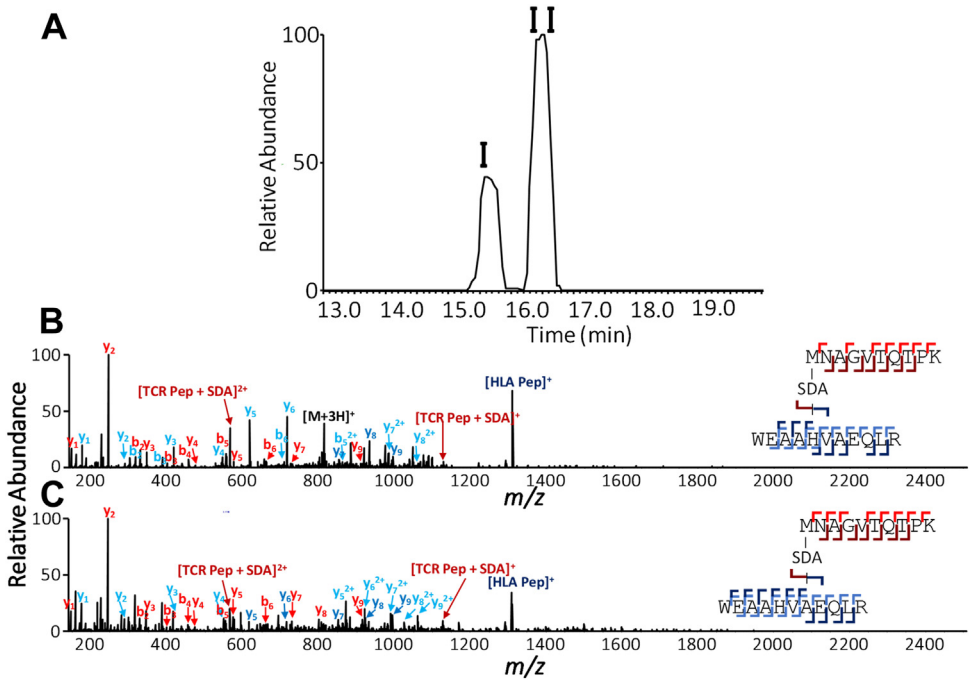

Figure 4. Differentiation of different cross-linked species using liquid chromatography. *A*, extracted ion chromatogram for TCR-D-pHLA cross-linked peptides at m/z 813.07– 813.09, obtained following LC-MS and the corresponding HCD MS/MS spectra for (*B*) elution point I and (*C*) elution point II. b and y ions are shown in *red* or *blue* depending on their corresponding peptide. Ions resulting from the additional cleavage adjacent to the cross-linker are shown in *dark red* or *dark blue*. Some less abundant ions are not labeled in the spectra. A full list of assignments is included in the relevant supplemental tables.

bound canonically, in agreement with the predicted structures.

Analysis of the TCR-D-pHLA complex revealed multiple cross-links between TCR-Dβ M1 and two HLA amino acids. Figure 4*A* shows the extracted ion chromatogram of ∼*m/z* 813. Two elution peaks were observed with retention times of ∼15.7 (I) and 16.5 (II) minutes. Analysis of the MS/MS data for peak I localized the cross-link between TCR-Dβ M1 (peptide position 1) and HLA H151 (peptide position 5) (Fig. 4*B* and Table S9). The HCD MS/MS data acquired over peak II again indicated that TCR-Dβ M1 is the cross-linking site, but the retention time shift suggests a structural change (Fig. 4*C* and Table S10). The ions at *m/z* 872.4561 and 921.9916 in the fragment spectrum over peak II were lacking in the corresponding fragment spectrum in peak I. These data indicate HLA A153 was the cross-linking site. An additional cross-link was identified between TCR-Dα K68 and HLA Y171 (Fig. 3*D* and Table S11). The distance between these two residues in the predicted structure was 17.6 Å, noticeably larger than the distances predicted for the cross-links identified for TCRs A-C, suggesting a potential structure violation.

Interestingly, a third cross-link was identified between TCR-Dβ G54 and a residue in the target peptide antigen (Fig. S2*B* and Table S12). This data suggests the target peptide antigen was partially or fully exposed when the TCR was bound to the pHLA. This is contrary to the modeled structure which predicted the TCR positioned in a more central orientation.

Additionally, the larger measured distance between cross-linked residues TCR-Dα K68 and HLA Y171 suggests lower accuracy for this predicted structure. Nevertheless, the cross-linking sites identified clearly revealed TCR-D binds in canonical orientation, in agreement with the predicted structure. However, the XL-MS/MS results TCR-D binds towards the N-terminus of the target peptide, in comparison to TCRs A-C which are more central.

Intermolecular cross-links reveal TCR-E binds in reverse orientation

Structural prediction showed TCR-E binds in a reverse orientation, that is, the position of alpha and beta chains on the pHLA surface was interchanged. MS/MS analysis identified one cross-link in the TCR-E-pHLA complex between residues TCR-Eα E53 and peptide antigen K10 (Fig. S3 and Table S13). This cross-link was feasible only if the alpha chain interacted with the second half of the peptide antigen and the beta chain engaged with the first half of the peptide antigen, the exact opposite of the typical canonical binding orientation. This confirms that TCR-E was a reverse binder in agreement with the predicted structure. The distance between the cross-linked residues in the predicted structure was 12.3 Å.

Crystal structures validate XL-MS/MS observations

To experimentally verify the correlation observed between the XL-MS/MS peptide linkages and predicted structures, we determined representative TCR-pHLA complexes using X-ray crystallography. TCR-C and TCR-E in complex with their respective pHLA readily crystallized and their structures were determined to be 2.77 and 1.86 Å resolutions, respectively.

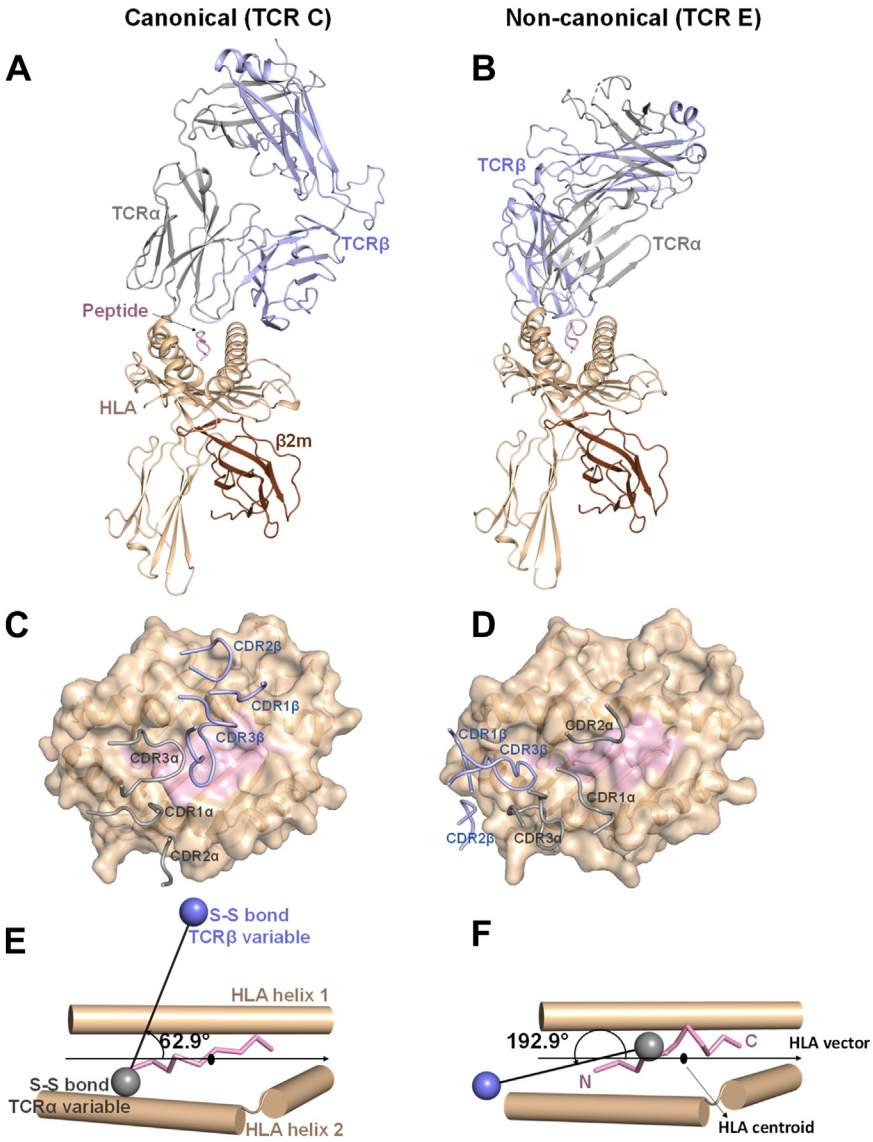

Figure 5. Crystal structures of TCR-C-pHLA and TCR-E-pHLA complexes. *A* and *B*, cartoon representations of the complexes. The HLA, β2m, and peptides are colored *light brown*, *dark brown*, and *pink*, respectively. The TCRα and TCRβ chains are colored *grey* and *blue* respectively. *C* and *D*. *top* view of the complexes showing HLA and peptides as surfaces and TCR CDRs as cartoon tubes. The rest of the TCR domains were not shown for clarity. The chains are colored as in a. *E* and *F*. crossing angle calculations for the TCR-pHLA complexes. The HLA helices are shown as cylinders and the peptides as ribbons. The position of the disulfide bonds in the TCRα and TCRβ variable domains are denoted as spheres and the line connecting the domains represents the TCR vector. The HLA vector (with direction from peptide N-terminus to C-terminus) passes through its centroid. The angle between the vectors for the complexes is indicated.

TCR-C was confirmed to bind canonically to its cognate pHLA with the alpha chain positioned toward the N-terminal of the binding peptide and the beta chain positioned towards the C-terminal of the binding peptide. The crossing angle of the TCR-C-pHLA binding orientation was measured at 62.9°, which is within the typical canonical binding range of 30 to 80°. βCDR2, βCDR3, and all three alpha CDRs made contacts with the HLA, and αCDR1, αCDR3, and βCDR3 made contacts with the peptide, providing a total buried surface area of 662 Å2 and 333 Å2 on the HLA and peptide respectively (Fig. 5, *A*, *C* and *E*). Concurring with the predicted model from TCRmodel2, the crystal structure revealed TCR-E bound to its cognate pHLA in the reverse orientation. This non-canonical TCR had a crossing angle of 192.9°. All six CDRs made contact with the pHLA, providing a total buried surface area of 1016 Å2 and 226 Å2 on the HLA and peptide respectively (Fig. 5, *B*, *D* and *F*).

Though TCRmodel2 was able to predict canonical and reverse binding orientation for TCR-C and TCR-E respectively, differences were observed when compared with the crystal structures. Overlay of the predicted and crystal structure of the TCR-C-pHLA complex, when aligned overall, revealed an all-atom RMSD of 1.19 Å with some differences in the CDR3 conformations and pHLA binding region (Fig. S4, *A* and *B*). As a comparison, distance measurements between the TCR-Dα K53 and HLA E154 crosslinked residues in the

predicted and crystal structures were very similar (6.5 Å in TCRmodel2 versus 6.8 Å in the crystal structure) (Figs. 3*C* and S5*A*).

For the TCR-E-pHLA complex, the overlay of the predicted and crystal structures showed some differences in the pHLA binding region, with an all-atom RMSD of 2.81 Å (Fig. S4, *C* and *D*). Distance measurements between the identified MS peptide pair linkage residues also showed some variability (12.3 Å in TCRmodel2 and 8.4 Å in crystal structure) (Figs. 3*C* and S5*B*), though both are well within the expected binding distances.

Irrespective of the differences in binding orientations observed between predicted and crystal structures for TCR-C and TCR-E, both models agreed with identified XL-MS/MS-based amino acid linkages. This clearly demonstrates the potential of this experimental technique to decipher TCR-pHLA binding orientations through molecular linkages.

Integrating cross-links into predicted models

To investigate if including XL-MS distances of the cross-linked residue pairs as part of structure prediction can aid in the accuracy of the models generated for TCR-pHLA complexes, we utilized the recently released Alphalink2 tool (26). Stahl K *et al.* showed that upon integrating crosslinking mass spectrometry data with AlphaFold-Multimer, Alphalink2 can substantially improve modeling performance by influencing the identification of interfaces and model selection, including on challenging antibody-antigen complexes.

When examined for the prediction of the five TCR-pHLA complexes by including XL-MS distances, Alphalink2 performance was similar or slightly better for TCR prediction when compared to TCRmodel2, though Alphalink2 fails very often to predict target peptide placement. We compared the two structure prediction tools, both utilizing unique Alphafold2 customizations, especially for multiple sequence alignments, on the predicted Cα-Cα distances with XL-MS distances and where possible RMSD of the TCR variable domains (when aligned based on HLA) with crystal structure.

In general, models generated by Alphalink2 were poor for target peptide modeling, and in the worst cases, target peptides were placed completely outside of the HLA groove. The reasons for this performance, localized to the target peptide chain, are not clear and unseen with TCRmodel2. However, even with incorrect target peptide conformation, when all models from Alphalink2 output for TCR-E-pHLA complex were aligned with the crystal structure, the overall TCR orientation of the best Alphalink2 model (rank 1, RMSD of 6 Å) was slightly better than the best model from TCRmodel2 (rank 1, RMSD 7.3 Å). When segregated to individual variable domains, the Alphalink2 model aligned relatively similarly for the alpha chain (RMSD 7.7 Å vs 6.4 Å) but better for the beta chain (RMSD 3.6 Å vs 8.1 Å) when compared with the TCRmodel2 model (Fig. S6, *A* and *B*). A similar comparison between the top-ranked models (rank 0) from Alphalink2 and TCRmodel2 for TCR-C-pHLA complex, when assessed using the crystal structure, showed that the Alphalink2 performance (barring peptide conformation) was slightly better for TCR orientation (RMSD 5.7 Å vs 6.9 Å) and in Cα-Cα distance of the crosslinked residue pair (11.1 Å vs 12.1 Å, crystal structure – 11.2 Å) (Fig. S6, *C*–*E*).

For TCR-A, TCR-B, and TCR-D, where we do not have crystal structures, we compared model performance based on the Cα-Cα distance of the crosslinked residue pairs, with selected models (due to incorrect target peptide placement for top-scoring models). For TCR-A-pHLA predictions, distances of the two crosslinked residue pairs were highly comparable between the two tools (<1 Å difference) with a relatively closer interface (below 13.5 Å) and very similar binding modes, implying both tools predicted high confidence models. For TCR-D-pHLA predictions, the Alphalink2 model had lower Cα-Cα distance for the two crosslinks (15.8 Å & 14 Å vs 9.8 Å & 18.7 Å), suggesting TCR orientation was better captured in the Alphalink2 model.

Overall, these results showed that integrating XL-MS distances with structure prediction could be advantageous, at least in cases when the confidence of the predicted models from TCRModel2 was suspected to be low. Alphalink2 predictions for TCR-pHLA complexes suffered from poor target peptide modeling and had longer run times (include comparison here). Future improvements in peptide modeling would likely improve the overall Alphalink2 performance to allow regular use for TCR-pHLA structure predictions.

Discussion

In this study, we have shown that by using XL-MS/MS, we were able to distinguish distinct protein footprints of TCRs and their targets. Several other mass spectrometry-based methods are available to monitor contact sites, such as hydrogen-deuterium exchange (HDX), and fast photochemical oxidation of proteins (FPOP) (27, 28), Each of these methods involves labeling of solvent-accessible residues in the native state, followed by subsequent digestion and analysis by LC-MS/MS. Although both techniques are effective in epitope mapping, the resolution for these techniques is typically limited to the peptide level, which often results in an overestimation of epitopes (29). Further, both HDX and FPOP measure "protection", rather than "binding". Therefore, any allosteric conformational change undergone during binding that results in reduced solvent exposure will also be identified as an epitope. In these respects, XL-MS is advantageous as it purely measures contact points and can do so to amino acid resolution. However, in comparison, it may underestimate the epitope, due to the limited reactivity of the linker used. This is highlighted in TCR-E where only one cross-link was noted. In future studies, a combination of several mass spectrometry methods would provide a more thorough picture of TCR binding.

In our study, cross-linking does offer a distinct advantage. Localizing contact sites to individual residues enables us to pinpoint specific regions in the TCRs and pHLAs that were colocalized, for example, the alpha chain over HLA helix two in canonical binders (Fig. 5*C*). Thereby, using this

methodology, we can distinguish between canonical and reverse binders, which other mass spectrometry methods cannot discern.

A potential limitation of XL-MS is the possibility of structural changes during the cross-linking step. Previous studies have suggested that whilst overall protein fold is retained, even at higher cross-linker concentrations, some small local structural changes may occur (30). Significant structural changes in our data are unlikely, given the consistency in measured distances between cross-linked residues in the crystal structures.

Reversed docking modes that confer antigen specificity and functionality, have been noted in the literature, but are much less common than canonical binders and are understood to be limited in their ability to contribute to the immune response (24). The XL-MS method presented in our study provides a system to more rapidly discover TCRs that bind in reverse orientation, in comparison to crystallographic techniques. A more thorough and complete understanding of TCR binding modes may prove crucial for the optimization and development of TCR-based immunotherapies. Additionally, during candidate selection, it is important to select TCRs that have optimal engagement with the target peptide antigen to ensure high specificity. Understanding the nuances of TCR binding, including reverse docking, can help in selecting T cells with the most effective antigen recognition capabilities.

Furthermore, the reverse docking mode is a testament to the ability of the TCR to adapt to different MHCs and peptide shapes. Highlighting how TCRs may engage with their ligands in multiple conformations can further our knowledge of how T cells recognize a vast array of antigens with a limited repertoire of TCRs.

Recent advances in deep learning-based artificial intelligence applied to protein science have improved the accuracy of predicting antibody epitopes and their interactions. However, unconventional binding, such as reverse docking, challenges some of the existing models of TCR-pMHC interactions. It necessitates that predictive models should account for alternative binding modes, which could be important for designing vaccines and predicting immune responses.

Our data emphasized that using predicted models without experimental data has inherent risks and accurately predicting conformational epitopes remains challenging. This is highlighted in the TCR-D modeled structure, which was predicted to bind with a broadly central footprint, yet the identification of a cross-link to the target peptide suggests the TCR is instead positioned towards the N-terminal of the target peptide. Recent work has illustrated the need to integrate computational data, which may unreliably predict a structure with mass spectrometry data collected in the laboratory, which may suffer in resolution and density (31, 32).

Ultimately, a more high-throughput and automated version of this assay may be designed. An interesting observation was the repeated identification of certain TCR residues readily forming cross-links in canonically bound HLAs, for example, K68 was predicted as a cross-linking site in three of the four canonical binders. A version of this method may be developed which involves searching for a small number of specific cross-linked tryptic peptides determined *in silico* prior to data analysis.

The mass spectrometry data uncovered several interesting insights into the behavior of sulfo-SDA-initiated cross-links under HCD. While the identification and localization of these types of cross-links has been well established, the respective MS/MS data were not always discussed in detail and many fragmentation patterns remain elusive (15, 33). It is of particular interest that we observe an additional fragmentation under HCD on either side of the cross-linker as this type of fragmentation typically targets C-N bonds. The bond fragmented is predicted to be between the sulfo-SDA cross-linker and the peptide, formed during UV activation. We hypothesize this bond was comparatively the weakest in the peptide. This was supported in HCD MS/MS spectra where many of the prominent peaks correspond to the expected b and y ions but with the additional mass of the cross-linker (+82.0149). Identification of this additional fragmentation product aided in the localization of each cross-linked peptide (example in Fig. 5*B* and Table S11). Understanding that two adjacent sets of b and y ions were produced in these instances was essential in providing accurate identification of the cross-linking sites.

Further work is needed to fully elucidate the nature of the cross-links under different fragmentation conditions which is outside of the scope of the work described here. Fully understanding the behavior of covalent cross-linkers under various mass spectrometry fragmentation conditions will increase confidence when localizing cross-links to specific residues. This will ultimately lead to more reliable data when investigating the structural dynamics of proteins.

Conclusion

In conclusion, we demonstrate the ability of XL-MS to investigate the nature of binding between TCRs and their target pHLAs using cross-linking to directly probe proximal residues in intermolecular complexes. We have shown this method is able to differentiate between canonical and reverse binders. Additionally, we can provide some measure of the TCR location with respect to the pHLA. With advantages in speed and ease over adjacent structural methods, XL-MS offers an alternative for screening TCR interactions when broad footprint information is sufficient. While the work described here was on affinity-enhanced TCRs (nM range), future work will aim to demonstrate the applicability of this technique to wild-type TCRs (μM range). Furthermore, the principles discussed here can be expanded to other protein-protein complexes with nM binding affinities, for example, antigen-antibody complexes. We have also highlighted the need to determine binding epitopes using a combination of experimental techniques and computational methods, with ongoing efforts to enhance accuracy and practical application.

Materials and methods

Protein expression and purification

Five affinity-matured TCRs (TCR A, B, C, D, and E) and their respective pHLA complexes were expressed, refolded,

and purified as described previously (34). Briefly, TCR alpha and beta chains were expressed separately as inclusion bodies in *E. coli* and refolded together. HLA heavy chain and beta-2-microglobulin were expressed separately as inclusion bodies in *E. coli* and refolded together with chemically synthesized peptide (Peptide Protein Research Ltd). Both TCRs and pHLAs were purified using both anion exchange and size exclusion chromatography.

Crystallization and structure determination

The TCR-pHLA complex was prepared by mixing purified TCR and pHLA at a molar ratio of 1:1.1 and concentrating the mixture to ~10 mg/ml. The crystallization drops were set up by dispensing 150 nl of protein solution plus 150 nl of reservoir solution in sitting-drop vapor diffusion format in two-well MRC Crystallization plates (SwissSci) using a Gryphon robot (Art Robbins). The plates were incubated at 20 °C in a Rock Imager 1000 storage system (Formulatrix). Crystals of TCR-C-pHLA complex were grown when the reservoir solution contained 0.2 M Sodium citrate tribasic dihydrate, 0.1 M Bis-Tris propane pH 7.5, and 20% w/v PEG 3350. Crystals of TCR-E-pHLA complex were grown when the reservoir solution contained 0.1 M MMT pH 9.0 and 25% PEG 1500. Crystals were cryoprotected in a reservoir solution containing 30% ethylene glycol and flash-cooled in liquid N2. X-ray data were collected from Diamond Light Source (Oxfordshire) beamlines I04 and I04-1 and processed using the xia2-3dii automated pipeline (35). The structure of the TCR-E-pHLA complex was solved by molecular replacement using PDB codes 5E00 (chains A and B for HLA-B2m), 3QEU (chain A for TCRα), and 4G8F (chain B for TCRβ) as search models in Phaser (36). The structure of the TCR-C-pHLA complex was solved by molecular replacement using PDB codes 5E00 (chains A and B for HLA-B2m), 4OZI (chain E for TCRα), and 4QRP (chain E for TCRβ) as search models in Phaser. The TCR-pHLA models were built using Coot (37) and refined using refmac (38) within the CCP4 suite (39). The structures were validated using molprobity (40). The data collection and refinement statistics are given in Table S14. The structure figures were generated using Pymol (http://www.pymol.org/pymol). The crossing angle was obtained by using a custom script in MOE (https://www.chemcomp.com/MOE-Molecular_Operating_Environment.htm).

Photo cross-linking reaction and sample preparation

60 μg TCR:pHLA complex was crosslinked with sulfo-SDA, using a protein-to-crosslinker molar ratio of 1:200. Cross-linking was carried out in two stages: first, sulfo-SDA (Thermo Fisher Scientific), dissolved in crosslinking buffer (PBS 7.4 (Gibco)) was added to the complex and left to react in the dark for 50 min at room temperature. Cross-linking reactions were quenched by the addition 10 μl 750 mM Tris-HCl buffer. Subsequently, the samples were photoactivated using ultraviolet light irradiation at 365 nm for 50 min at 200,000 μJ/cm2 using a UVP CL-1000 UV crosslinker (Akribis.).

Samples were denatured in 8 M guanidine hydrochloride (Sigma-Aldrich), reduced with 500 mM dithiothreitol, and alkylated with 500 mM iodoacetamide (both Thermo Fisher Scientific). Samples were desalted using Zeba spin columns (Thermo Fisher Scientific) according to the manufacturers' instructions and digested using sequencing grade modified trypsin/Lys-C (Promega, Madison, WI, USA) in 7.5 mM Tris-HCl buffer, pH 7.9.

Data acquisition

Peptides were separated using online RP-LC (Dionex UltiMate 3000), using a binary solvent system consisting of mobile phase A (LC-MS grade water (Thermo Fisher Scientific) with 0.1% formic acid (Thermo Fisher Scientific)) and mobile phase B (LC-MS grade acetonitrile (Thermo Fisher Scientific) with 0.1% formic acid (Thermo Fisher Scientific)). Peptides were separated using an ACQUITY UPLC BEH C18 Column (130 Å, 1.7 μm, 2.1 mm × 150 mm) (Waters) over a linear gradient. Samples were directly eluted *via* a heated electrospray ionization (HESI) source into the mass spectrometer.

All mass spectrometry experiments were performed on a Q Exactive Plus (Thermo Fisher Scientific). Data acquisition was controlled by Chromeleon 7.2 (Thermo Fisher Scientific). Full MS scans were performed at m/z 200 to 2000 and acquired at a resolution of 17,500 at m/z 200. The automatic gain control (AGC) target for the survey scans was 3 × 106 charges with a maximum injection time of 100 ms.

Higher energy collisional dissociation (HCD) was performed at a normalized collision energy of 28%. Fragments were detected at a resolution of 17,500 at m/z 200. The width of the precursor isolation window was 2 Th. AGC target was 1 × 105 charges with a maximum injection time of 200 ms. The loop count was set at 20.

Identification of cross-linked peptides

RAW mass spectrometry files acquired on the Q Exactive Plus were processed into peak lists and converted to mgf files using Msconvert (Version: 3.0.22010) (Proteowizard) (41). Cross-links were identified using both Xi and MeroX software (42, 43). Each identified cross-linked peptide was manually validated regardless of score to ensure the densest possible data.

In Xi analysis, mgf files were submitted to XiSearch for the identification of cross-linked peptides. Each sample was analyzed separately. The database consisted of the sequence for the TCR, along with those of the corresponding binding peptide and HLA. Search parameters were as follows: MS tolerance, 20 ppm; MS/MS tolerance, 20 ppm; enzyme, trypsin; missed cleavages, 3; crosslinker, SDA; fixed modification, carbamidomethylation of cysteine; variable modification, SDA and SDA-loop. Data were subsequently exported to XiFDR for validation. Identification with 5% residue-pair false discovery rate (FDR) was accepted for quantitation.

MS and MS/MS tolerances were set higher than the mass accuracy typically observed in the Q-Exactive Plus to best avoid missing cross-link identifications. Each identification was manually verified.

In MeroX analysis, mgf files were submitted to MeroX for further cross-link identification. The search parameters were the same as those described in Xi (44).

Data availability

X-ray data of TCR-C-pHLA and TCR-E-pHLA complexes were deposited at the Protein Data Bank (PDB) under accession codes 9GV6 and 9GV7, respectively.MS Data are available *via* ProteomeXchange with identifier PXD056652.

Supporting information—This article contains supporting Information.

Acknowledgments—We thank Diamond Light Source, UK for access to beamlines (proposals in 17077 and in 28224) and beamline staff for support with X-ray data collection. We thank Tristan Vaughan and Milos Aleksic for critically reading the manuscript.

Author contributions—A. C., V. K., T. P., M. E., and S. H. writing–review & editing; A. C., M. E., and S. H. supervision; V. K., T. P., and S. A. S. methodology; V. K. and T. P. investigation; T. P. writing–original draft; T. P., N. M., R. P., and A. S. conceptualization; K. B. resources.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AGC, automatic gain control; CDRs, complementarity-determining regions; HLA, human leukocyte antigens; MHC, major histocompatibility complex; pHLAs, peptide-HLAs; sulfo-SDA, Sulfosuccinimidyl 4,4′-azipentanoate; TCRs, T cell receptors; XL-MS/MS, cross-linking tandem mass spectrometry.

References

1. Waldman, A. D., Fritz, J. M., and Lenardo, M. J. (2020) A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat. Rev. Immunol.* **20**, 651–668
2. Sun, L., Su, Y., Jiao, A., Wang, X., and Zhang, B. (2023) T cells in health and disease. *Signal Transduct. Target. Ther.* **8**, 235
3. Trolle, T., McMurtrey, C. P., Sidney, J., Bardet, W., Osborn, S. C., Kaever, T., *et al.* (2016) The length distribution of class I–restricted T cell epitopes is determined by both peptide supply and MHC allele–specific binding preference. *J. Immunol.* **196**, 1480–1487
4. Nathan, P., Hassel, J. C., Rutkowski, P., Baurain, J.-F., Butler, M. O., Schlaak, M., *et al.* (2021) Overall survival benefit with tebentafusp in metastatic uveal melanoma. *N. Engl. J. Med.* **385**, 1196–1206
5. Gelmi, M. C., Gezgin, G., Velden, P. A. van der, Luyten, G. P. M., Luk, S. J., Heemskerk, M. H. M., *et al.* (2023) PRAME expression: a target for cancer immunotherapy and a prognostic factor in uveal melanoma. *Investig. Ophthalmol. Vis. Sci.* **64**, 36
6. Yang, H., Buisson, S., Bossi, G., Wallace, Z., Hancock, G., So, C., *et al.* (2016) Elimination of latently HIV-infected cells from antiretroviral therapy-suppressed subjects by engineered immune-mobilizing T-cell receptors. *Mol. Ther.* **24**, 1913–1925
7. Ishina, I. A., Zakharova, M. Y., Kurbatskaia, I. N., Mamedov, A. E., Belogurov, A. A., and Gabibov, A. G. (2023) MHC class II presentation in autoimmunity. *Cells* **12**, 314
8. Sundberg, E. J., Deng, L., and Mariuzza, R. A. (2007) TCR recognition of peptide/MHC class II complexes and superantigens. *Semin. Immunol.* **19**, 262–271
9. Milighetti, M., Shawe-Taylor, J., and Chain, B. (2021) Predicting T cell receptor antigen specificity from structural features derived from homology models of receptor-peptide-major histocompatibility complexes. *Front. Physiol.* **12**, 730908
10. Rossjohn, J., Gras, S., Miles, J. J., Turner, S. J., Godfrey, D. I., and McCluskey, J. (2014) T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 1–32
11. Yin, R., Ribeiro-Filho, H. V., Lin, V., Gowthaman, R., Cheung, M., and Pierce, B. G. (2023) TCRmodel2: high-resolution modeling of T cell receptor recognition using deep learning. *Nucleic Acids Res.* **51**, W569–W576
12. Sinz, A. (2014) The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev. Proteomics* **11**, 733–743
13. O'Reilly, F. J., and Rappsilber, J. (2018) Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat. Struct. Mol. Biol.* **25**, 1000–1008
14. Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2015) Crosslinking and mass spectrometry: an integrated technology to understand the structure and function of molecular machines. *TIBS* **41**, 20–32
15. Müller, F., Graziadei, A., and Rappsilber, J. (2019) Quantitative photo-crosslinking mass spectrometry revealing protein structure response to environmental changes. *Anal. Chem.* **91**, 9041–9048
16. Shigdel, U. K., Zhang, J., and He, C. (2007) Diazirine-based DNA photo-cross-linking probes for the study of protein–DNA interactions. *Angew. Chem. Int. Ed. Engl.* **120**, 96–99
17. Gomes, A. F., and Gozzo, F. C. (2010) Chemical cross-linking with a diazirine photoactivatable cross-linker investigated by MALDI-and ESI-MS/MS. *J. Mass Spectrom.* **45**, 892–899
18. Giese, S. H., Belsom, A., and Rappsilber, J. (2016) Optimized fragmentation regime for diazirine photo-cross-linked peptides. *Anal. Chem.* **88**, 8239–8247
19. Piersimoni, L., Kastritis, P. L., Arlt, C., and Sinz, A. (2021) Cross-linking mass spectrometry for investigating protein conformations and protein–protein Interactions— A method for all seasons. *Chem. Rev.* **122**, 7500–7531
20. Brunner, J. (2003) New photolabeling and crosslinking methods. *Annu. Rev. Biochem.* **62**, 483–514
21. Blencowe, A., and Hayes, W. (2005) Development and application of diazirines in biological and synthetic macromolecular systems. *Soft Matter* **1**, 178–205
22. Dorman, G., and Prestwich, G. D. (1994) Benzophenone photophores in biochemistry. *Biochemistry* **33**, 5661–5673
23. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
24. Gras, S., Chadderton, J., Del Campo, C. M., Farenc, C., Wiede, F., Josephs, T. M., *et al.* (2016) Reversed T cell receptor docking on a major histocompatibility class I complex limits involvement in the immune response. *Immunity* **45**, 749–760
25. Torrents de la Peña, A., Sewall, L. M., de Paiva Froes Rocha, R., Jackson, A. M., Pratap, P. P., Bangaru, S., *et al.* (2023) Increasing sensitivity of antibody-antigen interactions using photo-cross-linking. *Cell Rep. Methods* **3**, 100509
26. Stahl, K., Warneke, R., Demann, L., Bremenkamp, R., Hormes, B., Brock, O., *et al.* (2024) Modelling protein complexes with crosslinking mass spectrometry and deep learning. *Nat. Commun.* **15**, 7866
27. Liu, X. R., Zhang, M. M., and Gross, M. L. (2020) Mass spectrometry-based protein footprinting for higher-order structure analysis: fundamentals and applications. *Chem. Rev.* **120**, 4355–4454
28. Wang, L., and Chance, M. R. (2017) Protein footprinting comes of age: mass spectrometry for biophysical structure assessment. *Mol. Cell. Proteom.* **16**, 706–716
29. Dang, X., Guelen, L., Hulsik, D. L., Ermakov, G., Hsieh, E. J., Kreijtz, J., *et al.* (2023) Epitope mapping of monoclonal antibodies: a comprehensive comparison of different technologies. *mAbs* **15**, 2285285
30. Rozbeský, D., Rosůlek, M., Kukačka, Z., Chmelík, J., Man, P., and Novák, P. (2018) Impact of chemical cross-linking on protein structure and function. *Anal. Chem.* **90**, 1104–1113

31. Drake, Z. C., Seffernick, J. T., and Lindert, S. (2022) Protein complex prediction using Rosetta, AlphaFold, and mass spectrometry covalent labeling. *Nat. Commun.* **13**, 7846
32. Stahl, K., Graziadei, A., Dau, T., Brock, O., and Rappsilber, J. (2023) Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning. *Nat. Biotechnol.* **41**, 1810–1819
33. Pompach, P., Viola, C. M., Radosavljević, J., Lin, J., Jiráček, J., Brzozowski, A. M., *et al.* (2019) Cross-linking/mass spectrometry uncovers details of insulin-like growth factor interaction with insect insulin binding protein imp-L2. *Front. Endocrinol.* **10**, 695
34. Boulter, J. M., Glick, M., Todorov, P. T., Baston, E., Sami, M., Rizkallah, P., *et al.* (2003) Stable, soluble T-cell receptor molecules for crystallization and therapeutics. *Protein Eng.* **16**, 707–711
35. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190
36. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
37. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D.* **66**, 486–501
38. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **67**, 355–367
39. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D.* **67**, 235–242
40. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., *et al.* (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315
41. Holman, J. D., Tabb, D. L., and Mallick, P. (2014) Employing ProteoWizard to convert raw mass spectrometry data. *Curr. Protoc. Bioinformatics* **46**, 13–24
42. Götze, M., Pettelkau, J., Fritzsche, R., Ihling, C. H., Schäfer, M., and Sinz, A. (2014) Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass Spectrom.* **26**, 83–97
43. Iacobucci, C., Götze, M., Ihling, C. H., Piotrowski, C., Arlt, C., Schäfer, M., *et al.* (2018) A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein–protein interactions. *Nat. Protoc.* **13**, 2864–2889
44. [preprint] Mendes, M. L., Fischer, L., Chen, Z. A., Barbon, M., O'Reilly, F. J., Giese, S., *et al.* (2019) An integrated workflow for crosslinking mass spectrometry. *bioRxiv*. https://doi.org/10.1101/355396