On Conformations of Peptides Bound to Class I Major Histocompatibility Complexes

Jinbu Wang and Brian Y. Chen*
chen@cse.lehigh.edu
Department of Computer Science and Engineering, Lehigh University
Bethlehem, PA

ABSTRACT

The binding of peptides to major histocompatibility complex class I (MHCI) molecules is a key step for regulating the immune response. Hundreds of crystal structures of the MHCI in complex with the peptide, which sometimes include the T-cell receptor(TCR), have been produced to help understand this key interaction. To further understand the conformations of peptides bound to MHCI, we examined all peptide structures in peptide-MHCI dimers or TCR-peptide-MHCI trimers. We observed that peptides with the same sequence in different peptide-MHCI/TCR-peptide-MHCI complexes can either take similar conformations in both backbone and side chain, similar conformations in backbone and distinct conformations in side chains, and also distinct conformations in both backbone and side chains. These conformational varieties can be used to support the modeling of peptide-MHC complexes and the structure-based prediction of the binding affinity of the MHC-peptide complex.

CCS CONCEPTS

• **Applied computing → Molecular structural biology**.

KEYWORDS

Protein-Ligand Interactions, Protein Structure Comparison

ACM Reference Format:
Jinbu Wang and Brian Y. Chen. 2019. On Conformations of Peptides Bound to Class I Major Histocompatibility Complexes. In *10th ACM Int'l Conference on Bioinformatics, Computational Biology and Health Informatics (ACM-BCB '19), September 7–10, 2019, Niagara Falls, NY, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3307339.3343868

1 INTRODUCTION

Deleterious activations of the immune response spur important questions in human health relating to autoimmune reactions [8] and reactions to tissue transplantation [3]. Recognition by the T-cell receptor (TCR) of peptides presented by the Major Histocompatibility Complex (MHC) receptors [2, 6, 10] is a central step in these activations. Therefore, the exact nature of peptide-MHC interactions is an important point of study relating to these immune responses[7, 9, 10]. The conformation of the bound peptide is especially significant because it has selective interactions with both the MHC and the TCR, and the detailed orientations of its backbone and sidechains depends on a number of structural factors relating to the MHC and TCR.

As an overview, the known MHC class I (MHCI) structures in the PDB databank [1, 11] are composed of two polypeptide chains, a heavy chain and the non-covalently attached light chain known as β_2-microglobulin. The heavy chain is divided into thee domains $\alpha1$, $\alpha2$ and $\alpha3$ (Figure 1). The MHC-I binding groove is resides between $\alpha1$ and $\alpha2$ (Figure 2). Each of these domains contributes four strands to form an eight strand β-sheet on the bottom of the binding groove. Two long interrupted helices from each domain forms the side of the groove. The bound peptide is located between two helices in an extended conformation. When there are nine, the residues of the peptide are enumerated P1-P9. The TCR in the TCR-peptide-MHCI complex is bound to the peptide and parts of the MHCI on the opposite side of the MHCI.

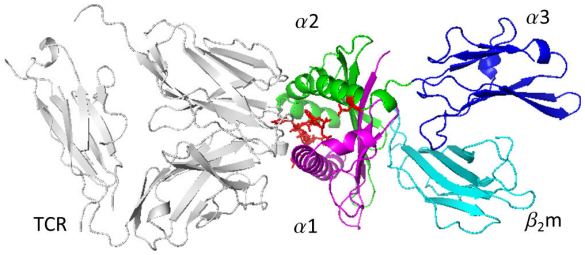

Figure 1: MHC class I complex structure of T-cell receptor, peptide and HLA-A2 (pdb: 1AO7). (A) Overall complex structure with T-cell receptor in light grey, $\alpha1$ in magenta, $\alpha2$ in green, $\alpha3$ in cyan and β_2m in blue. The bound peptide is shown as red stick model between T-cell receptor(TCR) and $\alpha1$-$\alpha2$ domains. All figures were produced using PyMOL [4].

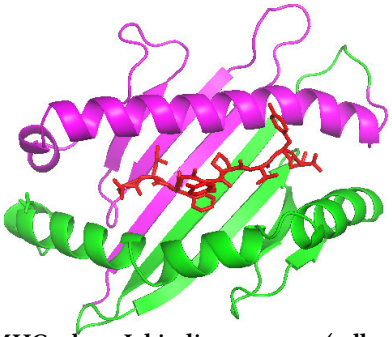

Figure 2: MHC class I binding groove (pdb: 1AO7) with bound peptide (red). The MHCI binding groove is formed from two domains: $\alpha1$ (magenta) and $\alpha2$ (green).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ACM-BCB '19, September 7–10, 2019, Niagara Falls, NY, USA
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6666-3/19/09…$15.00
https://doi.org/10.1145/3307339.3343868

This study examines all available peptide-MHC dimer and trimer structures to build a quantitative picture of how peptides vary in their bound conformations within the space of all human MHCI structures. We focused our examination of all structures in complex with a 9 residue peptide because there exist the largest number of structural examples. Next, our dataset divides into structures with peptides that occur only in peptide-MHC dimers, peptides that occur only in TCR-peptide-MHC trimers, and peptides that occur in both dimers and trimers. In each case, we measured the structural variation between the different crystal structures of the same peptides, computing both backbone and sidechain variation. These surveying efforts do not represent the frequency by which structural variations may occur, but they do reveal important examples that may run counter to standard intuitions. In this way, exhaustive surveillance efforts, first examined here, can uncover important details particular to the peptide-MHC system that could be overlooked by default modelling efforts.

In existing work, Schueler-Furman et al. [9] predicted peptide sidechain conformations based on bound peptides from 23 MHCI-peptide complex structures by constructing a specialized rotamer library. Their results showed that the predicted peptide structure selects the correct backbone from a set of backbone structures of other known structures with the same allele and peptide length.

Tong et al.[12] presented a protocol for modeling peptides in MHCI-peptide complexes. They analyzed 41 MHCI-peptide complexes and measured regional structural conservation in bound peptides. Lower Cα RMSD values at the N and C termini revealed that terminal residues were more constrained. They utilized this observation to initialize their modeling protocol.

Fagerberg et al. [5] developed an ab initio structure prediction approach to predict the bound peptide structure in MHCI-peptide complexes. They proposed a molecular dynamics approach based on a set of 41 MHCI-peptide complexes, where peptides were shorter than eleven amino acids. This prediction method can predict complexes of natural or modified antigenic peptides in their MHC environment with the aim to perform rational structure-based optimizations of tumor vaccines.

Relative to existing work, this paper offers an exhaustive analysis of all available human MHCI structures to offer a foundation on which modelling can be performed and evaluated with maximum information. Modelling and docking are better informed because the dependence of ligand conformations is assessed relative to the presence of MHC and TCR partners. The exhaustive nature of this study permits us to assess whether a ligand structure is an appropriate template, or if divergent information exists in potential templates. Likewise, by modeling the ligand of a structure that exists, a modeling technique can be assessed against an exhaustive catalog of all crystallized ligands with documented binding partners. Finally, this survey reveals consistencies and inconsistencies between all structure data and the common structural assumptions that might be made otherwise.

2 METHODS

2.1 Dataset Construction

We begin with more than 3000 MHCI related protein 3D structures from the PDB data bank [11]. Through sequence analysis, we reduced this initial set to 1008 structures containing MHCI domains. Within this set there are 468 structures of MHCI-bound peptides with 9 residues, 100 structures with 8 residues, and 136 with 10 residues. The binding groove of the MHCI fold has conserved residues that close both ends, restricting ligand peptides to be shorter than 10 residues. In practice, nine residue peptides are studied most. For this reason, we restricted this initial survey to peptides with exactly nine residues. From the 468 structures with 9 residue peptides, 278 were found to be bound to human MHCI structures, which we examine exclusively.

Since we are examining conformational variation in peptide binding conformations, we examine only peptides that occur in multiple structures. There are 107 structures in the 278 human MHCI complexes that contain a peptide that can be found in the bound conformation with more than one MHCI structure. These 107 structures represent the dataset used in this study. These 107 structures contain 32 distinct peptides bound to 8 distinct MHCI alleles. 55 of the structures are peptide-MHCI dimers and 52 are TCR-peptide-MHCI trimers. We use these structures because they can explain how the same ligand might or might not vary in different structural contexts.

2.2 Sequentially Identical MHCI structures

MHCI proteins from the same allele have the same sequence, but crystal structures often vary in the five N-terminal amino acids. To ensure precision in the exact classification of MHCI structures, we generated a multiple sequence alignment of the heavy chain of all 107 MHCI structures. Excluding the the first five positions of the alignment, all MHCI heavy chains with otherwise-identical sequences had identical alleles. We refer to them as being "sequentially identical" for brevity, with the express understanding that we are excluding variations that might exist in the first five positions of the sequence alignment.

2.3 Computing RMSD between ligand backbones and sidechains

Protein structure alignments were computed with Ska [13]. Molecule figures were drawn with PYMOL [4] and the root-mean-square deviations (RMSD) of backbones and sidechains were also computed with PYMOL. Sidechain RMSD is computed on all non-hydrogen sidechain atoms, which is possible because we are comparing the conformations of identical ligands.

3 RESULTS

In our dataset, peptides with nine residues that are found in complex with more than one MHCI can be divided into three categories. The first category includes structures of peptides that are only found in a dimer with the same MHCI allele. The second category includes structures of peptides that are only found in a trimer with the same MHCI allele and TCR. Finally, the third category includes structures of peptides that are found in both a dimer or a trimer of the above types. These categories leave out possibilities that were not observed in our dataset, such as the case where the same peptide is found in complex with different MHCI alleles. In each case, we report the backbone and side chain conformations of the bound peptides.

3.1 Similarity and Variation in multiple structures of the same MHCI-peptide dimer

Within our dataset, multiple structures of the same MHCI-peptide dimer are found in both very similar conformations and in different conformations. There are 5 examples of the same peptide found in two complexes with the same MHCI allele, where the peptide conformation changes very little. All comparisons of these peptide structures revealed very low RMSD values, as can be seen in Table 1. The highest backbone RMSD is 0.645 Å and the highest sidechain RMSD is 0.987 Å. These low backbone and sidechain RMSD values indicate degrees of structural similarity bordering on identity. The degree of similarity is also apparent in the ligand conformations shown in Figure 3.

There are also nine examples of peptides bound to the same MHCI allele where substantial conformational variations can be observed in the peptide (Table 2). In four cases, peptide ILKEPVHGV is bound to HLA A*02:01 in four different structures, none of which involve the TCR. While the different structures of the peptide exhibit relatively minimal backbone variations, higher sidechain RMSDs indicate differences in sidechain conformation. In the second example, peptide RRKWRRWHL, which is found in complex with HLA B*27:05 in 5 distinct PDB structures. Two distinct conformations were found in these 5 structures: 1OGT, 5IB3 and 5IB4 have a similar conformation with relatively small backbone (0.140 Å) and sidechain rmsd (0.508 Å). 5IB3 and 5IB4 also have a similar backbone conformation, with rmsd=0.115 Å, which is very distinct from that of 1OGT, 5IB3 and 5IB4 and apparent in their overall average RMSD. These two distinct conformations are illustrated in Figure 4. The average rmsds of backbone and sides between peptides in these two conformations are 1.488 and 4.865. The most changed residue conformation is P5 Arginie(R) and second is P6 Arginie(R).

These observations demonstrate that even basic assumptions about peptide-MHC complexes should be carefully considered in the context of existing data. It would be easy to assume that, when the peptide and the MHC are identical, that they would exhibit similar complex structures. It would be easy to arbitrarily pick one structure as a modeling template. However, actual crystal structures tell a different story, indicating that the same ligand can be found in different stable conformations with the same MHC, and that alternative modeling templates might need to be considered.

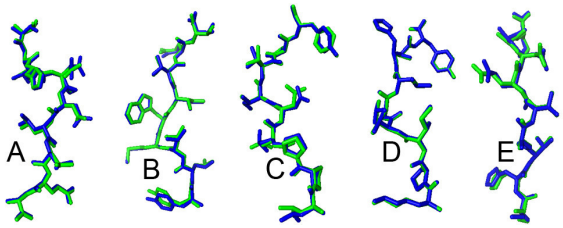

Figure 3: Peptides bound to the same MHCI complex in similar conformations. Five peptides were each found twice in complex with the same MHCI. Their conformations were very similar each time. These pairs are shown in A) 3D25 and 3FT3, B) 4L29 and 4L3C, C) 1A9B and 1A9E, D) 2CIK and 2H6P, and E) 4O2C and 4O2E.

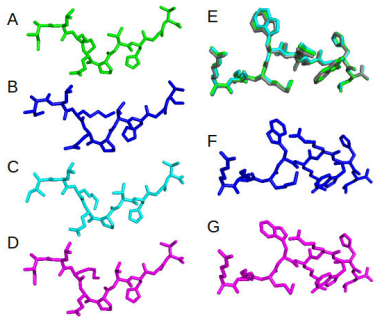

Figure 4: Peptides bound to the same MHCI in different conformations

3.2 Similarity in peptides found with only in the TCR-peptide-MHCI trimer

A second category of peptides is found only in complex with both the MHCI and the TCR. Our exhaustive survey found seven pairs of structures containing sequentially identical MHCI-peptide-TCR trimers. Across pairs, and one triplet, that have identical peptides, peptide conformation was found to change very little. The backbone and sidechain RMSDs between peptides are quite small. The highest backbone RMSD was 0.563 Å and the highest sidechain RMSD was 0.983 Å. The similarly of the conformations of these peptides can be seen in Figure 5.

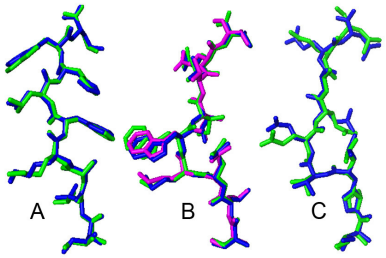

Figure 5: Distinct conformations of bound peptides in MHC-peptide complexes. Peptide structures 4MNQ and 5MEN (A), 2BNQ, 3GJF and 3HAE (B), 6BJ2 and 6BJ3 (C)

3.3 Similarity and Variations in peptides found with both dimers and trimers

The remaining peptides in our dataset were found in both peptide-MHC dimers as well as TCR-peptide-MHC trimers. We observed examples of peptides that maintain the same bound conformation in both dimers and trimers, examples of peptides that have separate conformations for the dimer and the trimer, and also several examples of varying conformations in one or both states.

Eleven peptides were found to form similar conformations in peptide-MHC dimers and TCR-peptide-MHCI trimers (see Table 4. Since the peptide is found in both configurations and it exhibits little conformational change in these cases, it appears that these peptides are not substantially influenced in their conformation by TCR binding. Among these examples, the highest backbone RMSD is 0.723 Å with the peptide SLYNTIATL in 1T20, 5NMH and 5NMF. The conformations of these bound peptides are shown in Figure 6

Table 1: RMSDs of peptides bound to the same MHCI complex in similar conformations

Allele	Peptide	PDBID	RMSD	
			backbone	Sidechain
HLA_A02:01	VLHDDLLEA	3D25, 3FT3	0.337	0.987
	YLLMWITQV	4L29, 4L3C	0.279	0.402
HLA_B*35:01	LPPLDITPY	1A9B,1A9E	0.212	0.804
	KPIVVLHGY	2CIK, 2H6P	0.069	0.806
HLA_B*39:01	SHVAVENAL	4O2C, 4O2E	0.645	0.809

Table 2: RMSDs of Peptides bound to the same MHCI complex in different conformations

Allele	Peptide	Conformation	PDBID	Minmax RMSD[a]	
				backbone	Sidechain
HLA_A*02:01	ILKEPVHGV	Fig.4 A	1HHJ	0.458	1.701
		Fig.4 B	1P7Q		
		Fig.4 C	2X4U		
		Fig.4 D	6EWA		
HLA_B*27:05	RRKWRRWHL	Fig.4 E	1OGT, 5IB3, 5IB4	2.931	12.132
		Fig.4 F	5IB1		
		Fig.4 G	5IB2		

[a] Minmax RMSD stands for the minimum of the maximum RMSDs of the nine amino acids' backbone/side chains in the peptide between MHCI-peptide structures in all different conformations.

Table 3: Similar peptide conformation in TCR-peptide-MHCI complex

Allele	Peptide	PDBID	RMSD	
			Backbone	Sidechain
HLA_A*02:01	ILAKFLHWL	4MNQ, 5MEN	0.242	0.481
	SLLMWITQV	2BNQ, 3GJF, 3HAE	0.289	0.676
HLA_B*35:01	IPLTEEAEL	6BJ2, 6BJ3	0.565	0.983

Table 4: RMSDs of Peptides that maintain similar conformations in MHC-peptide dimers and TCR-peptide-MHCI trimers

Allele	Peptide	PDBID		RMSD	
		MHCI-peptide	TCR-peptide-MHCI	backbone	Sidechain
HLA_A*01:01	EADPTGHSY	3BO8	1W72	0.362	0.368
HLA_A*02:01	ALWGFFPVL	1B0G	1LP9	0.246	0.388
	SLYNTIATL	1T20, 5NMH	5NMF	0.723	1.559
	SLYNTVATL	1T21, 1T22, 2V2W	5NME	0.620	1.287
	NLVPMVATV	2X4R, 3GSO	3GSN, 5D2L, 5D2N	0.577	1.080
	RMFPNAPYL	3HPJ	4WUU	0.280	1.284
	GLCTLVAML	3MRE	3O4L	0.273	1.310
HLA_B*08:01	FLRGRAYGL	1M05	1MI5, 3FFC, 3SJV	0.380	0.824
	HSKKKCDEL	4QRQ	4QRP	0.387	1.239
	LSSPVTKSF	2RFX	3VH8, 5B38, 5B39	0.133	0.415
HLA_C*04:01	QYDDAVYKL	1QQD	1IM9	0.557	0.835

The dataset contained 8 peptides with different conformations for the MHC-peptide dimer and for the TCR-peptide-MHCI trimer (See Table 5. Both backbone RMSD and sidechain RMSD are relative high these cases. The highest backbone RMSD is 1.023 Å between 3PWL and 3PWP. The highest sidechain RMSD is 2.682, between 3H7B and 3H9S. It is clear that these peptides are assuming different conformations between the dimeric and trimeric states. Note that in both dimer and trimer structures that the sequence of the MHCI is identical; the difference lies in the presence or absence of the TCR.

To further examine the effects of conformational change in dimers and trimers, we examined the positions of individual amino acids in the peptide. The amino acids with the most conformational changes are P3 and P5 for the peptide ALGIGILTV, P5 and P6 for the peptide SLLMWITQC, P4 and P5 for the peptide SLFNTIAV, P6 and P7 for the peptide ALGIGILTV, P3 and P5 for the peptide

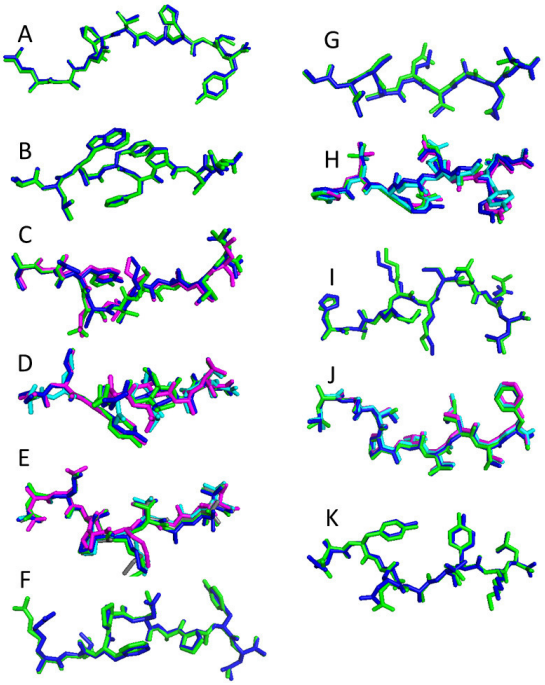

Figure 6: Peptides that maintain similar conformations in MHC-peptide dimers and TCR-peptide-MHCI trimers. A) 3BO8 and 1W72. B) 1B0G and 1LP9. C) 1T20, 5NMH and 5NMF. D) 1T21, 1T22, 2V2W and 5NME. E) 2X4R, 3GSO, 3GSN, 5D2L and 5D2N. F) 3HPJ and 4WUU. G) 3MRE and 3O4L. H) 1M05, 6BJ2 and 6BJ3. I) 4QRQ and 4QRP. J) 2RFX, 3VH8, 5B38 and 5B39. K) 1QQD and 1IM9.

MLWGYLQYV, P5 and P6 for the peptide LGYGFVNYI, P7 for the peptide GILGLVFTL, P5 and P6 for the peptide GILEFVFTL. The conformations of these peptides are shown in Figure 7. These variations never affect the P2 anchor residue, and only once affect the secondary anchor P7, whereas the amino acids in between are more often involved in conformational variation.

We also observed peptides with varying conformations in one or both of the dimer or trimer states. The peptide AAGIGILTV was found in 3 conformations across five structures, one conformation in the MHCI-peptide dimer and two conformations in the TCR-peptide-MHCI trimer (Table 6. The backbone RMSD between these conformations was 0.889 Å and the sidechain RMSD between these conformations was 2.447 Å. As can be inferred from these values, large conformational variations occurred in the side chains of the peptide. Further investigation revealed that the RMSDs between tyrosine P5 residues was very high, such as 4.570 Å between 1DUZ and 4E5X, 2.234 Å between 1DUZ and 1AO7, and 4.930 Å between 4E5X and 1AO7. The conformations of this peptides are shown in Figure 8B.

In contrast with the earlier peptide, a second peptide, LLFGYPVYV, formed three conformations in dimers and one conformation in trimer structures (See Table 7). The average backbone RMSD and sidechain RMSD were again quite high with backbone RMSD equal to 0.889 Åand sidechain RMSD Å. These high RMSDs indicate that the 4 conformations are quite different. The sidechain RMSDs

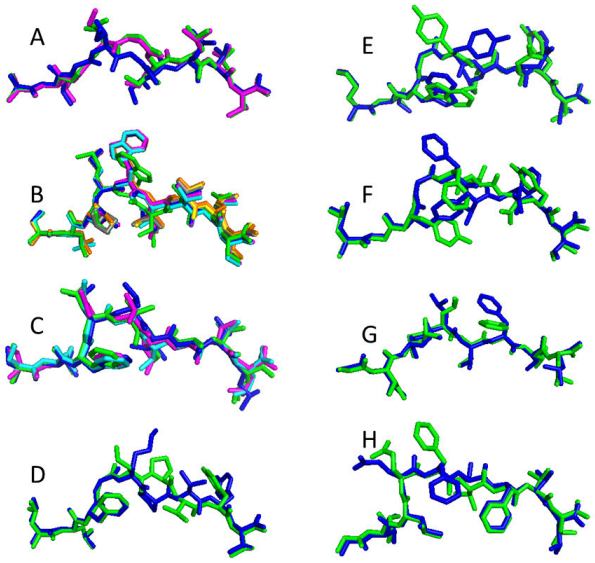

Figure 7: Peptide ligands with different conformations in MHC-peptide dimers and TCR-peptide-MHCI trimers. A) 1JHT,2GTZ and 4EUP. B) 1S9W, 2BNR, 2F53, 5F54, 2P5E, 2P5W and 2PYE. C) 1T1W, 2C7U, 5NMK, 5NMG. D) 2GIT and 2GJ6. E) 3H7B and 3H9S. F) 3PWL and 3PWP. G) 5HHN and 5HHM. H) 5HHP and 5HHO.

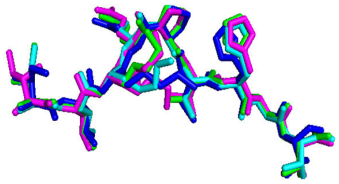

Figure 8: Various conformations of bound peptides in MHC-peptide complexes (pdbs: 1HHJ, 1P7A, 2X4U, 6EWA).

of the tyrosine P5 of the peptide are also relatively high, such as 4.570 Å between 1DUZ and 4E5X, 2.234 Å between 1DUZ and 1AO7, and 4.930 Å between 4E5X and 1AO7. The conformations of this peptides are shown in Figure 8A.

Finally, the peptide GILGFVFTL was had one conformation shared between dimers and some trimer complexes, but it also exhibited a second conformation TCR-peptide-MHCI trimers. The average backbone and sidechain RMSDs between all pairs of peptide conformations are 0.519 Åand 1.856 Å, respectively (Table 8. The sidechain RMSDs of the P5 and P7 amino acids (both phenylalanine) between these two conformations are 2.995 Åand 1.881 Å. Both conformations of the GILGFVFTL peptide are shown in Figure 10.

4 DISCUSSION

We have presented a comprehensive analysis of the range of conformations that can be observed in all existing structures of a 9 residue peptide bound in peptide-MHCI dimer or an TCR-peptide-MHCI trimer. While the measurements made cannot be expected to represent all peptide conformations in dimers and trimers, the

Table 5: RMSD of Peptides with different conformations in MHC-peptide dimers and TCR-peptide-MHCI trimers

Allele	Peptide	PDBID		RMSD	
		MHCI-peptide	TCR-peptide-MHCI	backbone	Sidechain
HLA_A*02:01	ALGIGILTV	1JHT, 2GTZ	4EUP	0.725	1.509
	SLLMWITQC	1S9W	2BNR, 2F53, 2F54 2P5E, 2P5W, 2PYE	0.688	2.357
	SLFNTIAVL	1T1W, 2C7U, 5NMK	5NMG	0.554	1.519
	LLFGKPVYV	2GIT	2GJ6	0.952	1.899
	MLWGYLQYV	3H7B	3H9S	0.711	2.682
	LGYGFVNYI	3PWL	3PWP	1.023	2.531
	GILGLVFTL	5HHN	5HHM	0.776	2.214
	GILEFVFTL	5HHP	5HHO	0.705	2.442

Table 6: various peptide conformations in MHCI-peptide complex

Allele	Peptide	Complex	PDBID	RMSD	
				Backbone	Sidechain
HLA_A*02:01	ILKEPVHGV	MHCI-peptide	1HHJ	0.459	1.208
			1P7Q		
			2X4U		
			6EWA		

Table 7: various peptide conformations in MHCI-peptide complex and TCR-peptide-MHCI complex

Allele	Peptide	Complex	PDBID	RMSD	
				Backbone	Sidechain
HLA_A*02:01	LLFGYPVYV	MHCI-peptide	1DUZ, 1HHK, 1IM3	0.889	2.502
			4E5X		
			5IRO		
		TCR-peptide-MHCI	1AO7, 1BD2, 4FTV		
	AAGIGILTV	MHCI-peptide	2GUO	1.665	2.447
		TCR-peptide-MHCI	3QDJ, 3QEQ		
			6EQA, 6EQB		

Table 8: Various peptide conformations in TCR-peptide-MHCI

Allele	Peptide	Conformation	Complex	PDBID	RMSD	
					Backbone	Sidechain
HLA_A*02:01	GILGFVFTL	1	MHCI-peptide	1HHI, 2VLL	0.519	1.856
			TCR-peptide-MHCI	1OGA, 2VLJ, 2VLK, 2VLR 5E6I, 5ISZ, 5TEZ		
		2	TCR-peptide-MHCI	5JHD		

survey revealed several interesting observations that both support and, interestingly, contradict standard structural assumptions.

First, we observed that there are examples of structural conservation in dimer conformations. It is natural to assume that a crystal structure without the TCR would always impart looser conformational constraints on the ligand, and this was not always true. We observed several examples where peptides in dimer structures maintained tightly consistent conformations. Indeed, we observed other peptides whose bound conformation did not substantially change when a TCR is added. We also observed an example of a peptide, GILGFVFTL, that exhibits two different conformations in the trimer complex, which one might expect to be more tightly constrained. These observations quantify ways in which there may be greater generality in the conformational space of bound peptides than might be expected, and also that

Second, we found that the conformations of same bound peptide in different MHCI-peptide complexes had similar conformations such as peptide VLHDDLLEA in 3D25 and 3FT3 (Table 1 and different conformations such as the peptide RRKWRRWHL in 1OGT and 5IB1 (Table 2. Conformational changes usually happened in P3, P4,P5,P6 and P7 amino acids. The conformations of same bound peptide between MHCI-peptide complexes and TCR-peptide complexes

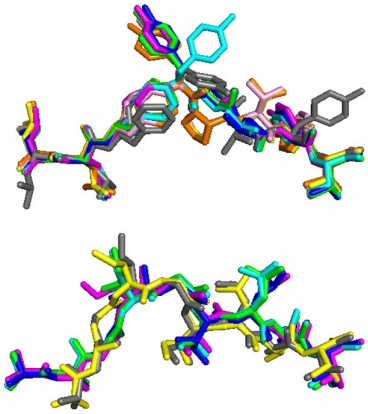

Figure 9: Various conformations in MHCI-peptide complexes and TCR-peptide-MHCI Complexes. A) Peptide structures of 1DUZ, 1HHK, 1IM, 4E5X, 5IRO, 1AO7, 1BD2 and 4FTV. B) Peptide structures of 2GUO, 3QDJ, 3QEQ, 6EQA and 6EQB.

Figure 10: Different conformations in MHCI-peptide complexes and TCR-peptide-MHCI Complexes. Peptide structures of 1HHI, 2VLL, 1OGA, 2VLJ, 2VLK, 2VLR, 5E6I, 5ISZ, 5TEZ and 5JHD.

also have a similar pattern, distinct conformational changes usually usually happened in amino acids P3, P4, P5, P6 and P7. The amino acids of P1, P2 P8 and P9 have less conformational changes in different complex structures. By analyzing of all these conformations in 107 MHCI-peptide/TCR-peptide-MHCI complexes from pdb databank, one peptide can form different conformations or maintain a similar conformation in different MHCI-peptide complexes/TCR-peptide-MHCI complexes. Due to the limited number of known MHCI-peptide/TCR-peptide-MHCI complexes, it is not clear if conformational changes are sequence dependent.

Together, these observations suggest ways in which geometric survey information could be collected for applications in knowledge based structure prediction and for building structure based maps of the extent of conformational variations that could occur at the peptide binding site.

REFERENCES

[1] Pamela J Bjorkman, MA Saper, B_ Samraoui, William S Bennett, JL t Strominger, and DC Wiley. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329, 6139 (1987), 506.

[2] Caitlin D Castro, Adrienne M Luoma, and Erin J Adams. 2015. Coevolution of T-cell receptors with MHC and non-MHC ligands. *Immunological reviews* 267 (Sept. 2015), 30–55. Issue 1. https://doi.org/10.1111/imr.12327

[3] Nicolas Degauque, Sophie Brouard, and Jean-Paul Soulillou. 2016. Cross-Reactivity of TCR Repertoire: Current Concepts, Challenges, and Implication for Allotransplantation. *Frontiers in immunology* 7 (2016), 89. https://doi.org/10.3389/fimmu.2016.00089

[4] Warren L DeLano et al. 2002. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* 40, 1 (2002), 82–92.

[5] Theres Fagerberg, Jean-Charles Cerottini, and Olivier Michielin. 2006. Structural prediction of peptides bound to MHC class I. *Journal of molecular biology* 356, 2 (2006), 521–546.

[6] Morgan Grau, Paul R Walker, and Madiha Derouazi. 2018. Mechanistic insights into the efficacy of cell penetrating peptide-based cancer vaccines. *Cellular and molecular life sciences : CMLS* 75 (Aug. 2018), 2887–2896. Issue 16. https://doi.org/10.1007/s00018-018-2785-0

[7] Lu He, Anne S De Groot, Andres H Gutierrez, William D Martin, Lenny Moise, and Chris Bailey-Kellogg. 2014. Integrated assessment of predicted MHC binding and cross-conservation with self reveals patterns of viral camouflage. *BMC bioinformatics* 15 Suppl 4 (2014), S1. https://doi.org/10.1186/1471-2105-15-S4-S1

[8] Lina Petersone, Natalie M Edner, Vitalijs Ovcinnikovs, Frank Heuts, Ellen M Ross, Elisavet Ntavli, Chun J Wang, and Lucy S K Walker. 2018. T Cell/B Cell Collaboration and Autoimmunity: An Intimate Relationship. *Frontiers in immunology* 9 (2018), 1941. https://doi.org/10.3389/fimmu.2018.01941

[9] O Schueler-Furman, R Elber, and H Margalit. 1998. Knowledge-based structure prediction of MHC class I bound peptides: a study of 23 complexes. *Folding & design* 3 (1998), 549–564. Issue 6. https://doi.org/10.1016/S1359-0278(98)00070-4

[10] Stefan StevanoviÄĞ. 2002. Structural basis of immunogenicity. *Transplant immunology* 10 (Aug. 2002), 133–136. Issue 2-3.

[11] Joel L Sussman, Dawei Lin, Jiansheng Jiang, Nancy O Manning, Jaime Prilusky, Otto Ritter, and Enrique E Abola. 1998. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. *Acta Crystallographica Section D: Biological Crystallography* 54, 6 (1998), 1078–1084.

[12] Joo Chuan Tong, Tin Wee Tan, and Shoba Ranganathan. 2004. Modeling the structure of bound peptide ligands to major histocompatibility complex. *Protein science : a publication of the Protein Society* 13 (Sept. 2004), 2523–2532. Issue 9. https://doi.org/10.1110/ps.04631204

[13] A S Yang and B Honig. 2000. An integrated approach to the analysis and modeling of protein sequences and structures. I. Protein structural alignment and a quantitative measure for protein structural distance. *Journal of molecular biology* 301 (Aug. 2000), 665–678. Issue 3. https://doi.org/10.1006/jmbi.2000.3973