

RESEARCH

Open Access



# Conformational flexibility of a free and TCR-bound pMHC-I protein investigated by long-term molecular dynamics simulations

Lisa Tomasiak, Rudolf Karch\*  and Wolfgang Schreiner

From 4th International Workshop on Computational Methods for the Immune System Function (CMISF 2020) Virtual. 16-19 December 2020

## Abstract

**Background:** Major histocompatibility complexes (MHCs) play a crucial role in the cell-mediated adaptive immune response as they present antigenic peptides (p) which are recognized by host T cells through a complex formation of the T cell receptor (TCR) with pMHC. In the present study, we report on changes in conformational flexibility within a pMHC molecule upon TCR binding by looking at molecular dynamics (MD) simulations of the free and the TCR-bound pMHC-I protein of the LC13-HLA-B\*44:05-pEEYLQAFY complex.

**Results:** We performed long-term MD simulations with a total simulation time of 8  $\mu$ s, employing 10 independent 400 ns replicas for the free and the TCR-bound pMHC system. Upon TCR ligation, we observed a reduced dynamic flexibility in the central residues of the peptide and the MHC  $\alpha$ 1-helix, altered occurrences of hydrogen bonds between the peptide and the MHC, a reduced conformational entropy of the peptide-binding groove, as well as a decreased solvent accessible surface area.

**Conclusions:** In summary, our results from 8  $\mu$ s MD simulations indicate a restricted conformational space of the MHC peptide-binding groove upon TCR ligation and suggest a minimum simulation time of approximately 100 ns for biomolecules of comparable complexity to draw meaningful conclusions. Given the relatively long total simulation time, our results contribute to a more detailed view on conformational flexibility properties of the investigated free and TCR-bound pMHC-I system.

**Keywords:** Molecular dynamics, Major histocompatibility complex, T cell receptor

## Background

Major histocompatibility complex type I (MHC-I) or II (MHC-II) molecules present antigenic peptides (p) which are recognized by the host T cells and hence play a crucial role in the cell-mediated adaptive immune response.

The T cell activation and the subsequent immune reaction are initiated by the T cell receptor (TCR) forming a complex with the peptide-presenting MHC (pMHC) [1]. MHC-I molecules are membrane bound proteins consisting of three  $\alpha$ -domains ( $\alpha$ 1,  $\alpha$ 2,  $\alpha$ 3) and a  $\beta$ 2-microglobulin, where  $\alpha$ 1 and  $\alpha$ 2 form the peptide-binding groove with two  $\alpha$ -helices on the sides and eight  $\beta$ -strands on the floor [2, 3].

While crystallographic data have played a considerable role in increasing our understanding of the structural

\*Correspondence: rudolf.karch@meduniwien.ac.at

Institute of Biosimulation and Bioinformatics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria



basis of TCR-pMHC engagement, they only provide a static picture and do not give comprehensive insight into the inter- and intra-molecular dynamics. In a recent study by Fodor et al. [4], the shortcomings of exclusively looking at static pictures of biomolecules were emphasized through uncovering significant conformational plasticity of human pMHC-I systems by crystallographic ensemble refinement techniques and molecular dynamics simulations. Their findings indicate that structural differences between pMHC and TCR-pMHC conformations might not necessarily be due to the binding process itself, which was proposed by former studies looking at static structures [5], but rather due to intrinsic flexibility. As the interest in the dynamics of immune response related molecules grew [6], the structure and dynamics of the *peptide* have generally been the focus of both experimental and computational studies. In a recent review, Ayres et al. [7] concluded that the mobility of peptides within MHC binding grooves impacts TCR binding, influencing association rates and entropic penalties. Peptides with higher mobility are therefore recognized more weakly by TCRs, with correspondingly reduced antigenicity. But these authors also discussed the fact that peptides can influence the motions of the MHC peptide-binding groove, especially of the  $\alpha 1$ - and  $\alpha 2$ -helix [8–10]. The focus on conformational dynamics of the MHC has grown in the past years and effects on TCR recognition and associated disease susceptibility has received great attention [11–13].

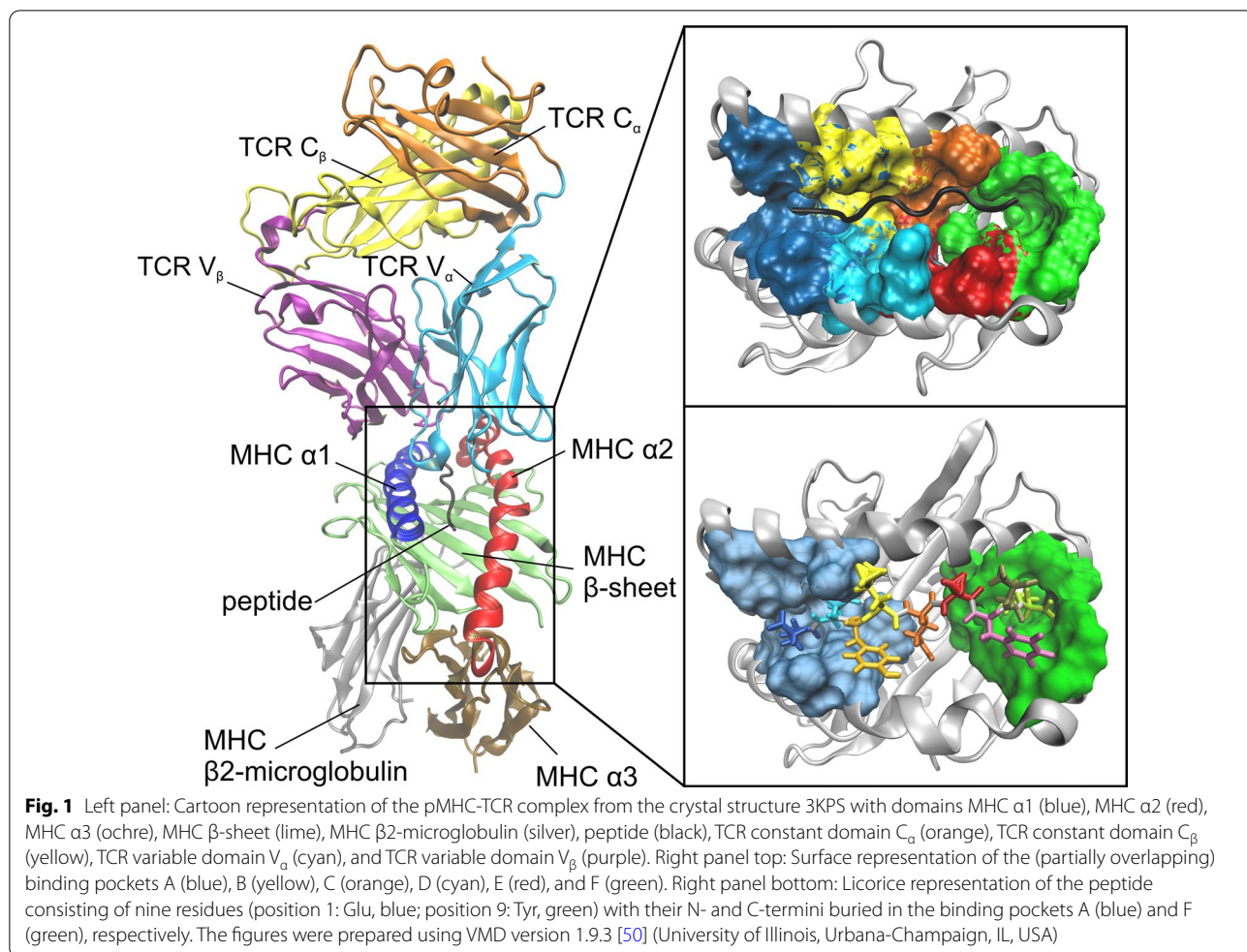
Molecular dynamics (MD) simulations have been used to provide information that goes beyond static structures and to give insight into the dynamic behavior of the TCR-pMHC complex. Narzi et al. [14] used 400 ns MD simulations to study two human MHC-I subtypes, HLA-B\*27:05 and HLA-B\*27:09, presenting a viral and three self-peptides. They could show that differences in the dynamic behavior, such as increased flexibility of the  $\alpha 1$ -helix and/or an opening of the binding groove of HLA-B\*27:05 compared to HLA-B\*27:09, are caused by very small structural differences in MHC subtypes (only one amino acid exchange in position 116). Gur et al. [15] conducted a similar study looking at HLA-B51 and HLA-B52 which only differ by two amino acids located in the B pocket [16] of the antigen-binding groove, while only HLA-B51 is associated with Behçet's disease. Using all-atom conventional MD simulations of 4.8  $\mu$ s in total length, these authors investigated the dynamics of the two HLA alleles bound to three different peptides and found that in HLA-B51 all peptides fluctuated to a larger extent due to looser binding compared to HLA-B52. Knapp and Deane [17] investigated the free and LC13 TCR-bound HLA-B\*08:01 in complex with the

Epstein-Barr viral peptide and 172 single-point mutations of the peptide using 100 ns MD simulations. In the bound state, these authors reported more hydrogen bonds (H-bonds) between the peptide and the MHC as well as altered flexibility patterns in the MHC helices and the peptide.

The earlier mentioned studies of various TCR-pMHC systems provide valuable contributions to our current understanding of the adaptive immune response. In a previous work [18], we studied the TCR-pMHC-I system LC13-HLA-B\*44:05-pEEYLQAFTY (see Fig. 1) and evaluated flexibility properties of the pMHC protein upon TCR ligation using MD simulations with a total length of 8  $\mu$ s. This system is of particular interest as the LC13 TCR is selected for recognition of self-HLA-B\*08:01 bound to the Epstein-Barr viral peptide FLRGRAYGL, but alloreacts with HLA-B\*44:05 bound to EEYLQAFTY [5].

In the present study, we extended our previous findings regarding the pMHC-TCR system LC13-HLA-B\*44:05-pEEYLQAFTY by adding additional analyses and considering further properties. In particular, we examined differences in the intramolecular flexibility of the free and the TCR-bound pMHC-I system with emphasis on the MHC binding groove as well as the antigenic peptide, which are involved in the TCR-pMHC complex assembly. We used the root mean square fluctuations (RMSFs) of individual residues of the MHC  $\alpha 1$ - and  $\alpha 2$ -helix as well as of the peptide to quantify the flexibility and to characterize relative movements of intramolecular domains in the free and the bound state. Furthermore, the radius of gyration ( $R_g$ ) of the MHC  $\alpha 1$ -helix, the MHC  $\alpha 2$ -helix, both MHC  $\alpha$ -helices and the peptide was computed to characterize structural differences of MHC domains between the free and the bound state. To measure the compactness of the binding groove, the time-evolution of the inter-domain distance between the MHC  $\alpha 1$ - and  $\alpha 2$ -helix was evaluated for the free and the bound state. Additionally, the total number of H-bonds as well as the residues involved in the formation of the H-bonds were analyzed for the free and the bound state in order to get an understanding of the intermolecular pMHC interaction through hydrogen bonding. As molecular flexibility is directly related to conformational entropy, we estimated changes in conformational entropy of the binding groove upon TCR ligation. Finally, we computed the solvent accessible surface area (SASA), which is another useful descriptor in the energetic analysis of protein–ligand associations.

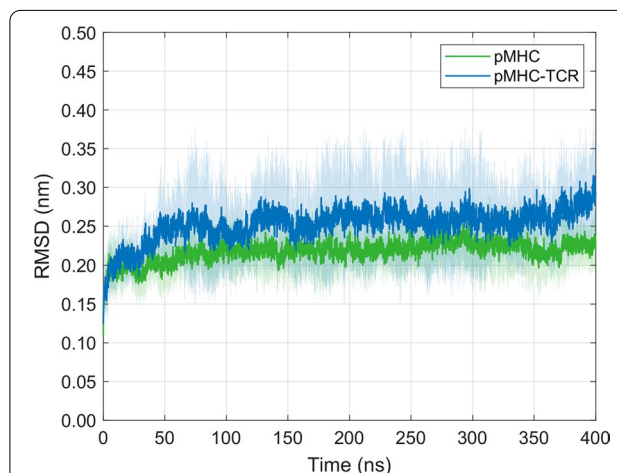
Our results contribute to a more detailed view on flexibility properties of the investigated free and TCR-bound pMHC-I system due to the long simulation time of 8  $\mu$ s.



## Results

### RMSD of the free pMHC molecule and the pMHC-TCR complex

The 3KPS pMHC-TCR protein represents a fairly complex system for MD studies, and one is well advised to examine the degree of equilibration prior to further analysis steps. Figure 2 shows the mean  $\pm$  SD of  $C_{\alpha}$  RMSDs calculated from 10 independent 400 ns simulations each for the free pMHC molecule and the whole 3KPS system. In the free pMHC system, after approximately 50 ns, the mean RMSD levels off and eventually reaches a plateau, indicating an approximately stationary probability distribution for the states of the system sampled by the MD simulations. In the pMHC-TCR complex, however, the mean RMSD increases after around 350 ns. This is caused by the RMSD courses of individual runs (see Additional file 1: Figs. S1, S2), mainly due to the movements of the MHC  $\alpha$ 3 and  $\beta$ 2-microglobulin domains in the respective runs.



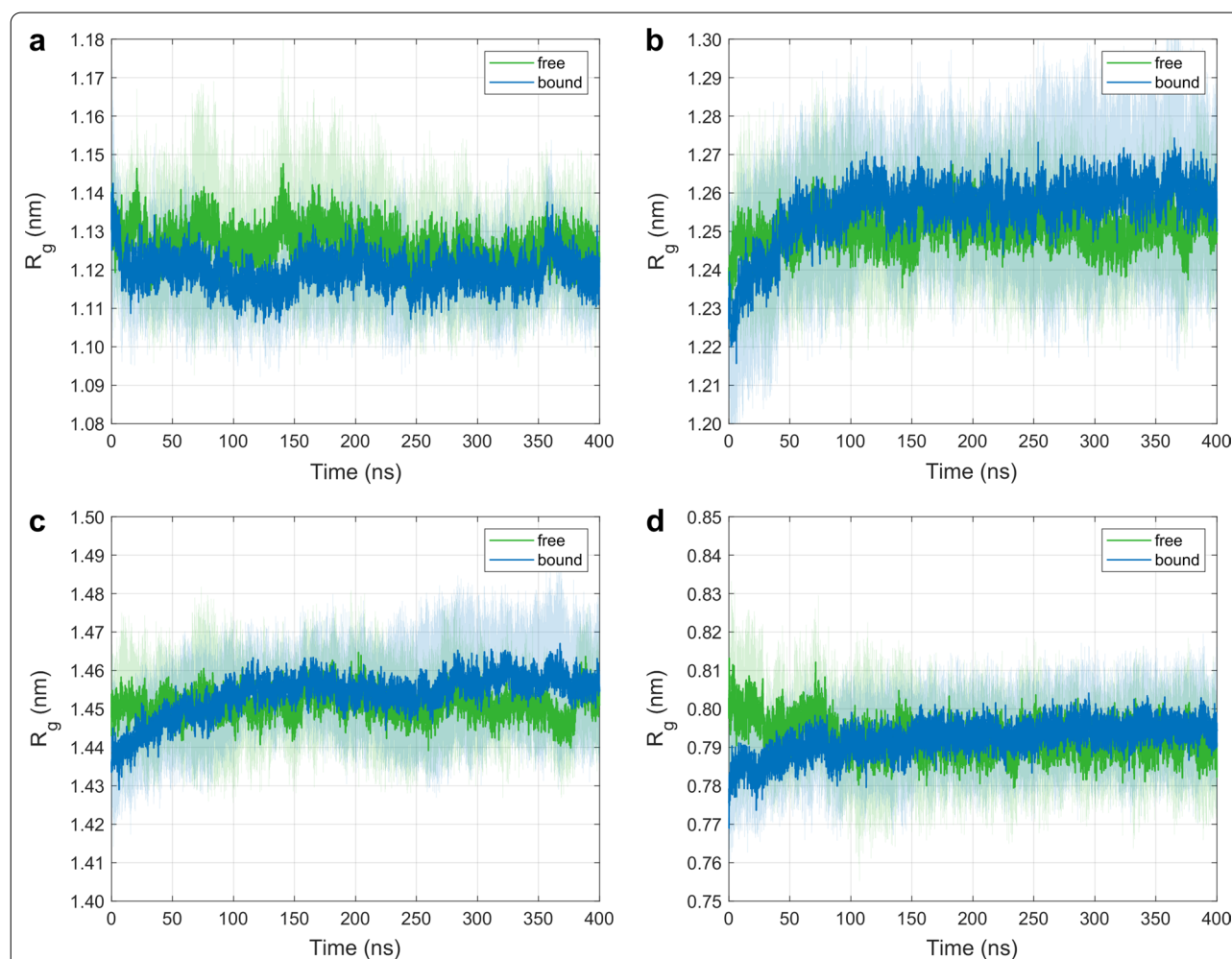
**Fig. 2** Time course of  $C_{\alpha}$  RMSD-values (mean  $\pm$  SD) for ten independent MD simulations of the free pMHC molecule and of the whole pMHC-TCR complex

### Radius of gyration

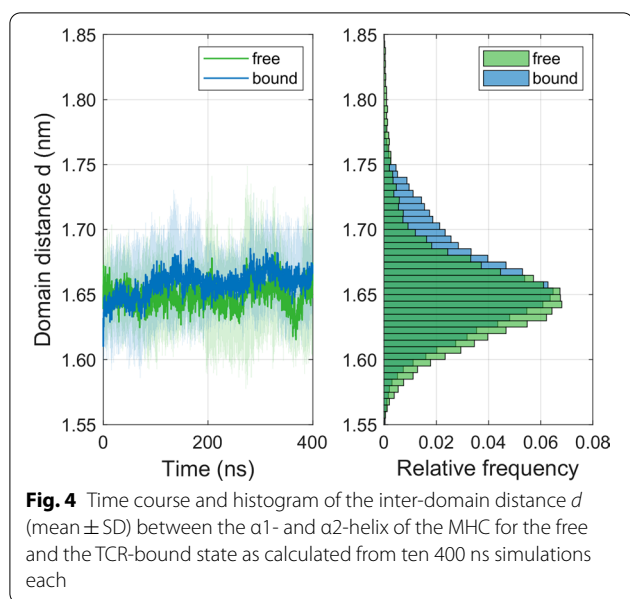
We employed the  $R_g$ -descriptor to measure differences in the compactness of MHC-domains between the free and the bound state. Figure 3 displays time courses of  $R_g$ -values (mean  $\pm$  SD) as obtained from 10 independent 400 ns MD simulations for the free and the TCR-bound pMHC molecule. While the compactness of the  $\alpha$ 1-helix is increased in the bound state (Fig. 3a), the  $\alpha$ 2-helix shows a reversed behavior (Fig. 3b). For the MHC-peptide (Fig. 3d),  $R_g$ -values are systematically elevated upon complex formation. The decreased compactness of the  $\alpha$ 2-helix upon TCR binding outweighs the opposite trend of the  $\alpha$ 1-helix, so that the two joint helices (Fig. 3c) and the peptide show a similar behavior. Even though these differences are small, they remain manifest after an equilibration phase during the first 100 ns of the simulations.

### Inter-domain distance between MHC $\alpha$ -helices

The proper presentation of the peptide to a TCR by the MHC binding groove is central to a successful initiation of an immune response. We therefore evaluated the time evolution of the inter-domain distance  $d$  between the  $\alpha$ 1- and  $\alpha$ 2-helix of the MHC binding groove in the free and the TCR-bound state, see Fig. 4. After a cross-over at approximately 50 ns, the time averaged distance between the helices only shows slight differences between the free and the bound state:  $\langle d_{\text{free}} \rangle = 1.65$  nm,  $\langle d_{\text{bound}} \rangle = 1.66$  nm, where  $\langle \dots \rangle$  denotes the time average over the pooled 10 trajectories. After an initial equilibration of approximately 50 ns duration, mean distances in the bound state are systematically larger than in the free state.



**Fig. 3** Time course of the radius of gyration  $R_g$  of  $C_\alpha$  atoms (mean  $\pm$  SD) for ten independent MD simulations of 400 ns each for the free and the TCR-bound state: **a** MHC  $\alpha$ 1-helix. **b** MHC  $\alpha$ 2-helix. **c** Both MHC  $\alpha$ -helices. **d** MHC peptide



#### Root mean square fluctuations in the pMHC molecule

We evaluated  $C_{\alpha}$  RMSF-values of individual residues of the MHC  $\alpha 1$ - and  $\alpha 2$ -helix as well as of the peptide, where the first 50 ns of individual simulations were excluded from the analysis. RMSF-values are generally decreased upon binding with the TCR or at most remain unaffected, with the exception of the MHC  $\alpha 2$ -helix, see Fig. 5. Most striking differences between free and bound RMSF-values were observed for residues at positions 4–6 (Leu, Gln, Ala) of the peptide as well as residues at positions 151–153 (Arg, Val, Ala) within the  $\alpha 2$ -helix and less pronounced changes were observed within the  $\alpha 1$ -helix at positions 70–73 (Asn, Thr, Gln, Thr). N- and C-terminal positions (residues at positions 1 and 9, respectively) of the peptide remain comparably stable upon TCR-binding with respect to the RMSE.

#### H-Bonds of the peptide with the MHC binding groove

Figure 6a displays the time course of the total number of H-bonds (mean  $\pm$  SD) between the peptide and the MHC binding groove, i.e., the lateral  $\alpha$ -helices and the  $\beta$ -sheet underneath, as obtained from 10 independent MD simulations each for the free and the bound state. After an equilibration phase of 50 ns, the dynamics of the mean total number of H-bonds virtually does not change with time, nor do the differences between the free and the bound states with respect to their mean values:  $9.7 \pm 0.2$  (free) versus  $9.8 \pm 0.3$  (bound). There are systematic differences between the free and the bound state in the individual simulations (data not shown), albeit in mutually opposite directions, i.e., approximately half of the

simulations show an increased, while the other half show a decreased total number of H-bonds in the bound state. Figure 6b shows a cartoon representation of the MHC binding groove, where residues with H-bonds of a relative occurrence  $\geq 25\%$  (see the following paragraph) in either the bound, the free or both states are labelled with their respective 3-letter codes.

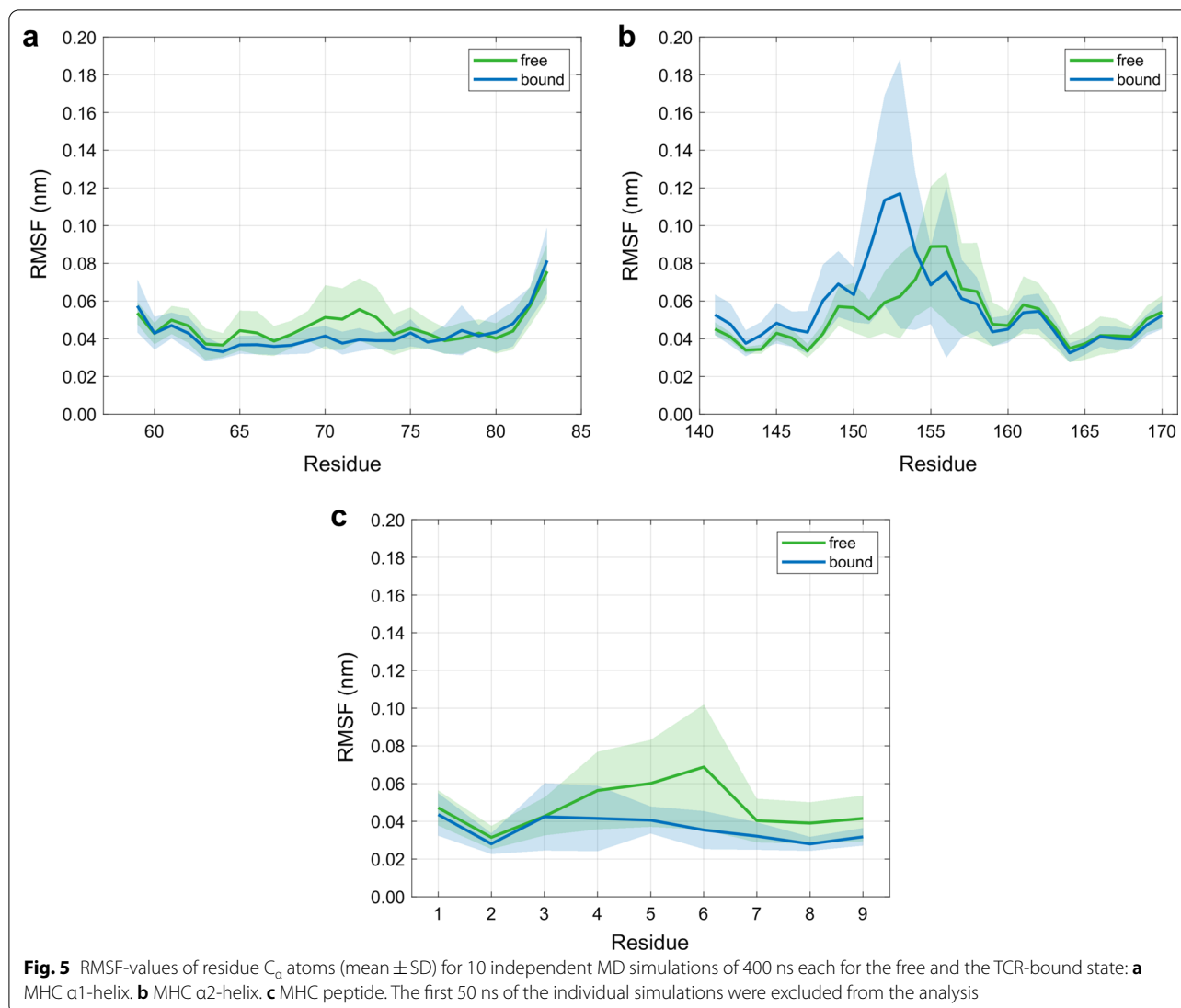
In addition to the progression of the total number of H-bonds between the peptide and the MHC binding groove, we analyzed residues engaged in forming the respective interactions, excluding the first 50 ns of the simulations. Figure 7 shows residues of the peptide and the binding groove, which are involved in H-bond interactions that were present more than 80% of the total simulation time in at least one of the 20 MD runs. Displayed are relative occurrences (i.e., average percentages of H-bonds weighted by the number of their appearances within the 10 individual runs each for the free and the bound state, irrespective of donors/acceptors located in the main or the side chain). Figure 7 demonstrates larger relative occurrences of H-bonds at the N-terminus (Glu1) of the peptide in the free state and larger relative occurrences of H-bonds at the C-terminus (Tyr9) of the peptide in the bound state. While many H-bonds were highly conserved during TCR ligation (e.g., Tyr3-Asp156, Tyr9-Tyr84), some others appeared only in the free state (e.g., Glu2-Tyr99, Tyr9-Thr80) or only in the bound state (e.g., Tyr3-Tyr99, Thr8-Glu76). Also, the comparably short presence of H-bonds involving peptide residue Ala6 is consistent with its large positional fluctuations in the free state, see Fig. 5c for the respective RMSF-values.

#### Configurational entropy of the binding groove

Figure 8 displays estimates of configurational entropies of the pMHC binding groove for each of the 10 runs for the free and the TCR-bound pMHC molecule, where the first 50 ns of the simulations were excluded from the analysis. Although we observed an overlapping range of entropic energy values  $T \cdot S$  of the free and the bound states at  $T = 310$  K, the difference  $T \Delta S = -98.43$  kJ/mol between the respective mean values of entropic energies ( $5060.91 \pm 93.16$  kJ/mol bound versus  $5159.33 \pm 80.84$  kJ/mol free, mean  $\pm$  SD) is statistically significant (two-sided  $t$ -test for independent samples,  $p = 0.021$ ) with a standard error of the difference  $T \Delta S$  of the mean values of 39.0 kJ/mol.

#### Discussion

We performed a series of long-term MD simulations to analyze conformational and flexibility properties of a pMHC class I molecule upon TCR ligation. In particular, we conducted ten independent 400 ns runs each for an unbound and TCR-bound pMHC system, summing



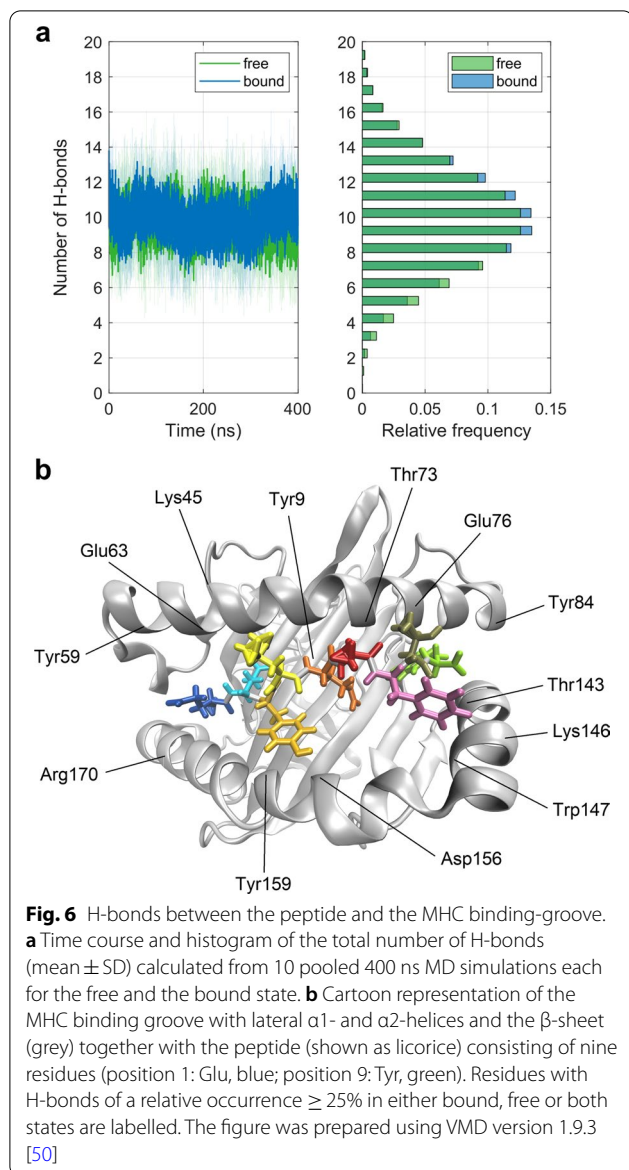
up to a total simulation time of 8  $\mu$ s. We chose this design of the simulations, since conformational space has been reported of being more efficiently sampled by a series of relatively short and independent MD simulations than by a single long one [19].

### Backbone RMSD

Mean RMSD-values were adequately relaxed to equilibrium after an initial relaxation time of approximately 50 ns, suggesting that the simulations had generated stable trajectories, thus providing a sound basis for further analysis. Therefore, we removed the first 50 ns in the calculation of RMSF-values, in the analysis of H-bonds with VMD, and in the calculation of conformational entropies of the pMHC complex.

### Radius of gyration

To characterize the compactness of a structure, we evaluated the radius of gyration  $R_g$  of the MHC  $\alpha$ 1-helix, the MHC  $\alpha$ 2-helix, both MHC  $\alpha$ -helices, and the peptide. We observed that relaxation times for equilibration were quite different both with respect to the domains considered (whole complex and, e.g., the peptide) and with respect to the descriptor (RMSD and  $R_g$ ), see Figs. 2 and 3. The different trend in  $R_g$ -values of the  $\alpha$ 1-helix and the  $\alpha$ 2-helix upon binding might be due to inherent structural variations between the two helices: While the  $\alpha$ 1-helix constitutes a pure  $\alpha$ -helix along all of its residues, the  $\alpha$ 2-helix is composed of three parts [20], joined by coils at residues 151 (Arg) and 162 (Gly), see Fig. 1.



### Inter-domain distance between MHC $\alpha$ -helices

Inter-domain distances  $d$  between the  $\alpha$ 1- and  $\alpha$ 2-helices are slightly elevated in the bound state after an initial equilibration phase of approximately 50 ns, where we observed a cross-over of the respective time courses, see Fig. 4. This behavior is consistent with the time course of  $R_g$ -values for both  $\alpha$ -helices, see Fig. 3c.

### Flexibility within the pMHC molecule

Flexibility properties of the peptide and the MHC are important for TCR ligation [6, 21, 22]. Changes in flexibility of pMHC domains (peptide and  $\alpha$ -helices) upon TCR binding were characterized by  $C_\alpha$  RMSF-values

of the respective residues, see Fig. 5. In addition, Fig. 9 illustrates the extent of movements of the peptide in two representative 400 ns runs of the free and of the bound state. In line with previous studies [23, 24], we observed a reduced dynamic flexibility upon TCR ligation within the central residues of the peptide and less pronounced also at the MHC  $\alpha$ 1-helix. As the authors in [7] have pointed out, this reduced conformational heterogeneity upon binding leads to increased entropic costs for binding and to a weakened affinity. Changes of RMSF-values upon TCR ligation, in particular within the peptide and the  $\alpha$ 2-helix, reflect a certain degree of adaptation of the pMHC complex to the TCR, thus pointing to an induced fit binding model rather than supporting the conformational selection paradigm [25].

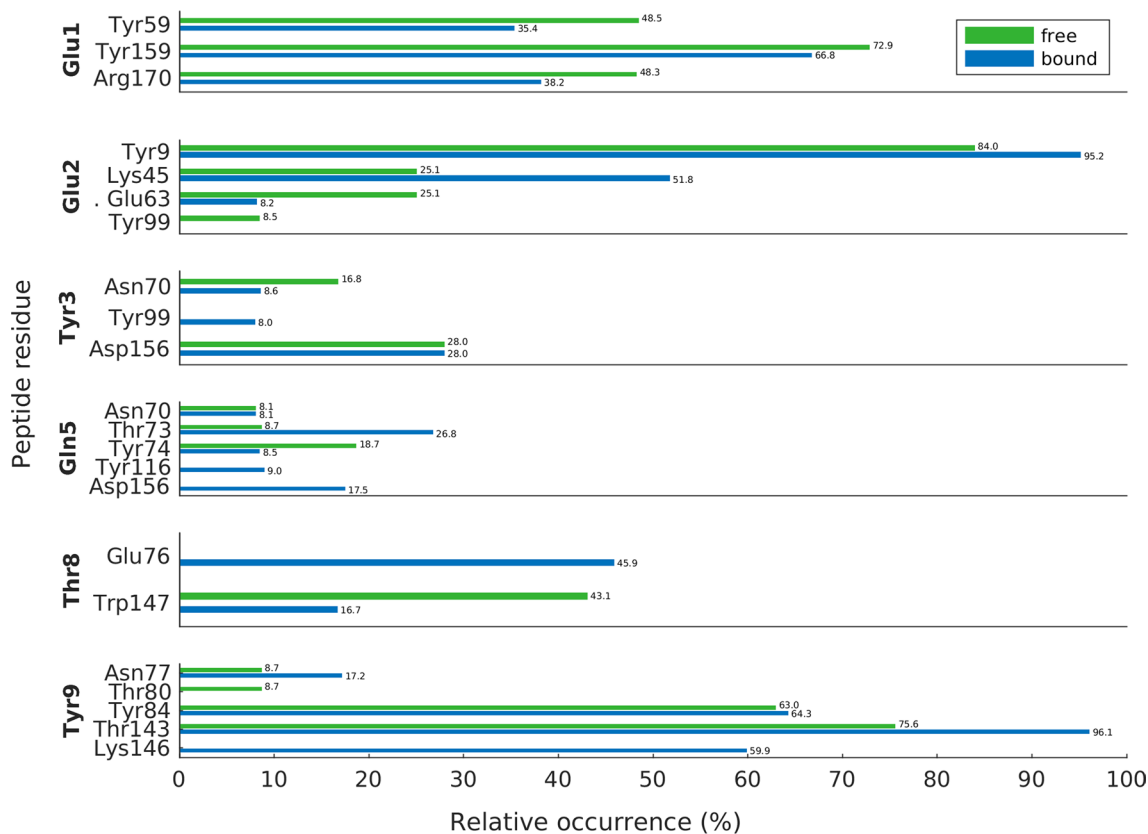
### H-Bonds between the peptide and the MHC binding groove

H-bonds between the peptide and the MHC binding groove (i.e., the lateral  $\alpha$ -helices and the subjacent  $\beta$ -sheet) anchor the peptide in the MHC and thus guarantee appropriate presentation to the TCR interface. The RMSF-values demonstrate that the N- and C-termini of the peptide (residue positions 1 and 9) are much less affected by TCR-binding than its central residues (see Fig. 5) as these termini act as anchor sites through a network of hydrogen bonds which stabilize the peptide-binding domain [20]. In agreement with other studies [26], most H-bonds of the terminal residues of the peptide were retained both in the free and in the bound state, see Fig. 7. H-bonds between peptide Glu1 and  $\alpha$ 2-helix Tyr159, peptide Glu2 and  $\alpha$ 2-helix Tyr9 as well as between peptide Tyr9 and  $\alpha$ 2-helix Thr143 were present with the highest relative occurrences (Glu1-Tyr159: 72.9% free, 66.8% bound; Glu2-Tyr9: 84.0% free, 95.2% bound; Tyr9-Thr143: 75.6% free, 96.1% bound), suggesting an adequate anchoring of the peptide ends in their respective binding pockets A and F [1, 16], see also Fig. 1.

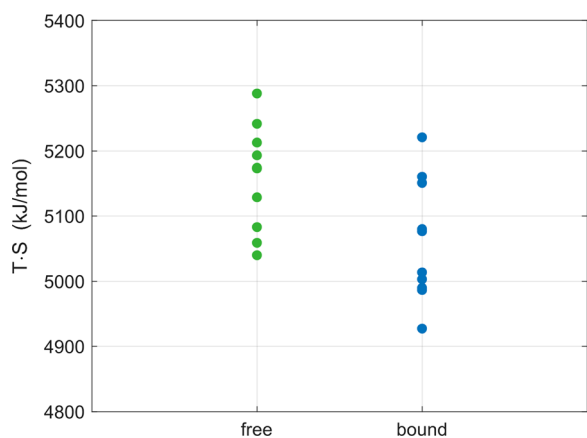
### Conformational entropy of the MHC binding groove

TCR recognition of a pMHC complex is guided by the binding affinity of the TCR, which is directly related to the change in Gibbs free energy  $\Delta G_{\text{bind}}$  upon binding.  $\Delta G_{\text{bind}}$  represents the difference of an enthalpic ( $\Delta H_{\text{bind}}$ ) and an entropic ( $T\Delta S_{\text{bind}}$ ) term,  $\Delta G_{\text{bind}} = \Delta H_{\text{bind}} - T\Delta S_{\text{bind}}$ . Therefore, configurational entropy  $S$ , which is an essential contribution to  $\Delta S_{\text{bind}}$ , can play an important role in modulating free energy changes during the association of pMHC molecules with a TCR, i.e., during the transition from a high entropy (disordered) unbound state to a low entropy (ordered) bound state [27, 28].

Thermal motions on a picosecond to nanosecond time-scale are universal in protein dynamics and are believed



**Fig. 7** Relative occurrences of H-bonds between the residues of the peptide and the binding groove (i.e., average percentage of H-bonds weighted by the number of their appearance in the 10 individual runs each for the free and the bound state, irrespective of donors/acceptors located in the main or the side chain)

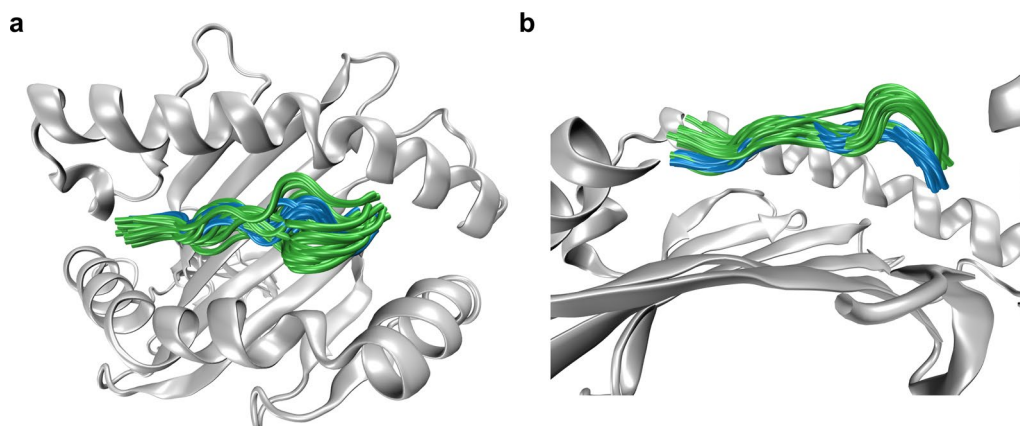


**Fig. 8** Dot plot of entropy estimates for the pMHC binding groove for each of the 10 runs of the free and the TCR-bound pMHC molecule, where the first 50 ns of the simulations were excluded from the analysis. Mean  $\pm$  SD of entropic energy values  $T \cdot S$  at  $T = 310$  K:  $5060.91 \pm 93.16$  kJ/mol (bound) versus  $5159.33 \pm 80.84$  kJ/mol (free). Difference of mean values:  $T \Delta S = -98.43$  kJ/mol. Standard error of  $T \Delta S$ : 39.0 kJ/mol. The difference  $T \Delta S$  of respective mean values was statistically significant (two-sided  $t$ -test for independent samples,  $p = 0.021$ )

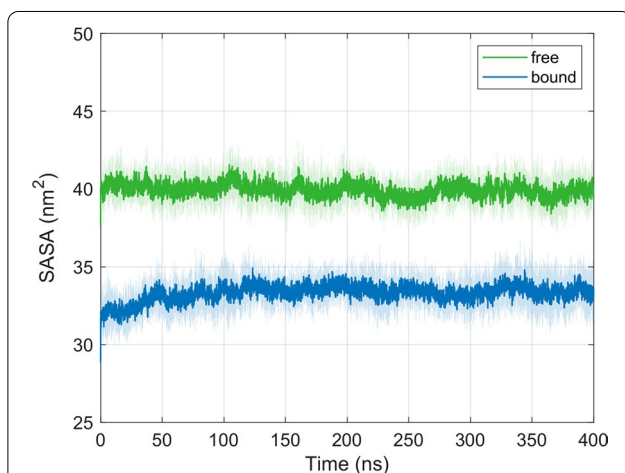
to be directly related to protein conformational entropy. Motions on these timescales can be directly studied with nuclear magnetic resonance (NMR) experiments and can also be captured in MD simulations with realistic force fields and adequate sampling techniques. Although absolute entropies depend on the models of molecular motions, relative entropies were reported to be model independent [28]. To estimate conformational entropy changes of the MHC binding groove upon TCR ligation, we used the quasi-harmonic approximation developed by Schlitter [29], which is based on a previous work of Karplus and Kushick [30] on entropy estimation from the covariance matrix  $\sigma$  of atomic coordinate fluctuations.

In the 3KPS system, we observed a value of  $T \Delta S = -98.43$  kJ/mol for the mean entropic energy change of the peptide-binding groove upon TCR ligation. This value is consistent with results reported by Narzi et al. [14], who estimated conformational entropies of the binding groove of MHC class I proteins in complex with viral and self-peptides from MD simulations. On the other hand, our observed decrease in configurational entropy is also consistent with the reduced flexibility of





**Fig. 9** Snapshots illustrating peptide movements. **a** Cartoon representation of overlaid snapshots of MHC peptides at every 40 ns from representative 400 ns runs of the free (green) and the bound (blue) states to illustrate the extent of peptide movements. **b** Side view with parts of the  $\alpha$ 2-helix removed for better visibility of the peptide snapshots. The figures were prepared using VMD version 1.9.3 [50]



**Fig. 10** Time course of the solvent accessible surface area SASA (mean  $\pm$  SD) of the MHC binding groove for the free and the TCR-bound state as calculated from ten 400 ns simulations each

the  $\alpha$ 1-helix and parts of the  $\alpha$ 2-helix upon binding, see Fig. 5.

#### Solvent accessible surface area of the MHC binding groove

The solvent accessible surface area (SASA) constitutes another useful descriptor in the energetic analysis of protein–ligand associations, e.g., in estimating solvation free energy changes. SASA values were calculated using the GROMACS tool *sasa* [31]. Figure 10 shows the time course of SASA values (mean  $\pm$  SD) of the MHC binding groove in the free and the TCR-bound state as calculated from ten independent simulations of 400 ns each. After an initial transient phase of about 50 ns in the bound state, the SASA values remained fairly stable both in the

free and in the bound state. The decrease upon TCR ligation by about 7 nm<sup>2</sup> with respect to the free state is consistent with the view of a detachment of water molecules from the TCR–pMHC interface upon binding [32].

#### Limitations

The membrane environment of pMHC and TCR proteins has been identified as an essential part of the pMHC–TCR interaction and the T cell response [33]. Since the seminal study of Wan et al. [34] of a membrane-embedded TCR–pMHC–II complex, many research groups have investigated the influence of the membrane on structural and thermodynamic properties of pMHC–TCR systems, e.g., [26, 35, 36]. In particular, Bello et al. [36] performed 100 and 150 ns MD simulations of peptide-bound and peptide-free MHC-II molecules in aqueous and membrane-bound configurations and reported that the presence of the membrane might restrict the conformational flexibility of the peptide-binding groove. Furthermore, co-receptors and various signaling proteins are involved in the formation of the immunological synapse and should also be included in future MD studies of pMHC–TCR systems [21].

Despite known issues associated with the Berendsen-barostat [37], we chose this algorithm to render our results better comparable with previous studies, e.g., Wan et al. [34], Narzi et al. [14], Bello and Correa-Basurto [35]. Our total simulation time of 8  $\mu$ s for both the bound and the free system is still too short to capture processes with time scales far beyond this limit, such as slow complex motions or the initiation of TCR-signaling. To overcome the limits of simulation time, various approaches have been reported in the literature, such as coarse-grained

methods, elastic network models, or steered MD simulations [38–40].

## Conclusions

In the present study, we used a series of long-term MD simulations to examine dynamic features of various domains of the HLA-B\*44:05-pEEYLQAFTY complex upon binding with an  $\alpha\beta$  LC13 TCR. As a consequence of TCR-binding, we observed a reduced dynamic flexibility in the central residues of the peptide and the MHC  $\alpha$ 1-helix, altered occurrences of H-bonds between the peptide and the MHC, a reduced conformational entropy of the peptide-binding groove, as well as a decreased SASA. Moreover, our results suggest a minimum simulation time of approximately 100 ns for biomolecules of comparable complexity to obtain meaningful results. Furthermore, due to the large variability of individual RMSD-values as depicted in Fig. 2, we suggest a reasonable number of replicas for each system (depending on the specific descriptors studied) to draw reliable conclusions from conventional all-atom MD simulations in explicit solvent.

## Methods

### Simulation systems

We employed two different sets of start configurations for the MD simulations: The first set was made up of the crystal structure of a LC13  $\alpha\beta$  TCR, a nine-mer self-peptide (EEYLQAFTY) derived from the ATB binding cassette protein ABCD3, and a MHC molecule of the HLA-B\*44:05 type (PDB-ID: 3KPS) [5], see Fig. 1. This system holds the advantage of being available at a high spatial resolution of 2.7 Å. In the second set, MD simulations were started from unbound pMHC-molecules assembled from the crystal structure of the original 3KPS complex.

### Molecular dynamics simulations

We conducted all-atom MD simulations in explicit water with GROMACS 5.1.1 [41], using the Amber99sb-ildn force field [42] and a rhombic dodecahedral box with 2 nm minimum distance between each protein and the box boundaries. The 3KPS complex comprised 396,833 atoms, representing 826 protein residues and 127,634 solvent molecules. To reach a physiological salt concentration of 0.15 mol/L, the proteins were solvated in SPC water [43] and neutralized by replacing 748 solvent molecules with 386  $\text{Na}^+$  and 362  $\text{Cl}^-$  ions.

After energetically minimizing the systems using the method of steepest-descent with 10,000 steps, equilibration runs of 100 ps in length were performed at NVT conditions with the temperature set to 310 K by means of a Berendsen-thermostat with a time constant

of 0.1 ps and position restraint MD. The following equilibration runs in NPT ensembles were conducted under the control of a Berendsen-barostat set to 1 bar with a time constant of 1.0 ps.

Production runs of 400 ns each were performed using the LINCS algorithm [44] for all bonds and virtual sites for hydrogen atoms, thus allowing for a time-step of 4 fs [41]. A cut-off of 1.47 nm was used for van der Waals interactions and the particle-mesh Ewald (PME) algorithm [45] was applied to calculate long-range electrostatic Coulomb interactions with a cut-off of 1.4 nm. Temperature coupling was conducted with the velocity-rescaling algorithm [46] at a temperature of 310 K. Isotropic pressure coupling at 1 bar was performed using the Berendsen algorithm [47]. Atom coordinates were saved every 40 ps to a trajectory file to yield 10,000 frames for each run. We carried out 20 independent MD simulations employing different initial velocities. Ten runs were performed for the whole complex and another 10 runs for the unbound pMHC molecule, resulting in a total simulation time of 8  $\mu$ s.

### Structural analysis

Before analyzing the trajectories, translational and rotational motions of the proteins relative to their respective first frames (i.e., the equilibrated structures) were removed by means of the *trjconf* tool supplied with GROMACS.

To assess stability and adequate equilibration of the simulated systems, we computed the RMSD( $t$ ) (root mean square deviation) for each time-frame  $t$  with respect to the first frame of each trajectory with the GROMACS *rms* tool, considering only  $C_\alpha$  atoms of the protein backbone:

$$\text{RMSD}(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N |\mathbf{r}_i(t) - \mathbf{r}_i^{\text{ref}}|^2}$$

where  $\mathbf{r}_i(t)$  is the position of atom  $i$  at time  $t$ ,  $N$  is the total number of  $C_\alpha$  atoms within the protein backbone, and  $\mathbf{r}_i^{\text{ref}}$  is the position of atom  $i$  in the selected reference structure (the first frame of each trajectory in our case). The descriptor RMSD( $t$ ) thus represents the square root of averaged squared distances between the atom positions at time  $t$  and the positions in the reference structure, thereby rating molecular deformations at time  $t$  with respect to the reference positions at time zero.

To characterize the flexibility of individual atoms  $i$  or atom groups (e.g., residues), we evaluated the root mean square fluctuation RMSF( $i$ ) with the *rmsf* tool of GROMACS:

$$\text{RMSF}(i) = \sqrt{\frac{1}{T} \sum_{t_j=1}^T |\mathbf{r}_i(t_j) - \langle \mathbf{r}_i \rangle|^2}$$

where  $T$  is the number of time steps over which the average is computed,  $t_j$  is the time step of frame  $j$  within the respective trajectory, and  $\langle \mathbf{r}_i \rangle$  is the time-averaged position of atom  $i$ . Contrary to the  $\text{RMSD}(t)$ , which represents an average over all atoms for a specific time  $t$ , the  $\text{RMSF}(i)$  is an average over time for a specific atom  $i$ . As such, the  $\text{RMSF}(i)$  constitutes a statistical measure of the deviation between the positions of an atom  $i$  during a simulation and its time-averaged position  $\langle \mathbf{r}_i \rangle$ .

Molecular flexibility is directly related to conformational entropy, since increased molecular flexibility populates more states and thus entails higher entropy values [7]. To estimate changes in conformational entropy  $S$  of the binding groove upon TCR ligation, we employed the quasi-harmonic approximation developed by Schlitter [29, 48] as supplied with the GROMACS tools *covar* and *anaeig*:

$$S = \frac{k}{2} \ln \det \left( \frac{kTe^2}{\hbar^2} \mathbf{M}\boldsymbol{\sigma} + \mathbf{I} \right)$$

where  $\mathbf{M}$  and  $\mathbf{I}$  are the mass matrix and the identity matrix, respectively, and  $e$  is Euler's number.  $\boldsymbol{\sigma}$  is the covariance matrix of the coordinate fluctuations computed from the MD trajectories,

$$\sigma_{ij} = (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle)$$

where  $x_1, \dots, x_{3N}$  are the Cartesian coordinates of the  $N$  atom system and  $\langle \dots \rangle$  stands for the time average.  $k$ ,  $T$ , and  $\hbar$  are Boltzmann's constant, absolute temperature, and  $\hbar = h/(2\pi)$ , where  $h$  denotes Planck's constant.

To describe relative movements of the  $\alpha 1$ - and  $\alpha 2$ -helices in the MHC molecule, we calculated the distance  $d$  between the geometric centers of these domains based on the  $C_\alpha$  atoms of the protein backbones.

To probe differences in compactness between free and bound states, we also analyzed the radius of gyration  $R_g$  of  $C_\alpha$  atoms in selected domains with the *gyrate* tool of GROMACS:

$$R_g = \sqrt{\frac{\sum_{i=1}^N |\mathbf{r}_i(t)|^2 m_i}{\sum_{i=1}^N m_i}}$$

Here,  $m_i$  means the mass of atom  $i$  and  $\mathbf{r}_i$  denotes the position of atom  $i$  with respect to the center of mass of the molecule or a domain within the molecule.  $R_g$  is defined as the mass-weighted root mean square distance of a group of atoms from their common center of mass

[49] and thus serves as a measure of the compactness of a structure.

Finally, we evaluated the time course of H-bonds between the peptide and the MHC binding groove using the *hbond* tool available within GROMACS. To examine the residues participating in the formation of H-bonds, we employed the *Hydrogen bonds* analysis tool of the VMD software [50]. H-bonds between donor- and acceptor-atoms were identified by means of the following geometrical criteria: (a) Donor-acceptor distance  $\leq 3.5$  Å, (b) hydrogen-donor-acceptor angle  $\leq 30^\circ$ . The donor-acceptor distance of 3.5 Å stems from the first minimum of the radial distribution function of SPC water.

#### Abbreviations

MD: Molecular dynamics; MHC: Major histocompatibility complex; pMHC: Peptide-MHC complex; RMSD: Root mean square deviation; RMSF: Root mean square fluctuation;  $S$ : Configurational entropy in kJ/(mol K); SASA: Solvent accessible surface area; SD: Standard deviation;  $T$ : Absolute temperature in Kelvin (K); TCR: T cell receptor.

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12865-022-00510-7>.

**Additional file 1.** Supplementary Figures.

#### Acknowledgements

We thank Michael Cibena (Institute of Biosimulation and Bioinformatics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna) for his help running the simulations and preparing the manuscript. The authors gratefully acknowledge the helpful comments of the reviewers.

#### About this supplement

This article has been published as part of *BMC Immunology* Volume 23 Supplement 1 2022: Selected papers from the 4th International Workshop on Computational Methods for the Immune System Function (CMISF 2020): *immunology*. The full contents of the supplement are available at <https://bmcimmunol.biomedcentral.com/articles/supplements/volume-23-supplement-1>.

#### Author contributions

RK conceived and planned the study. LT and RK performed MD simulations, evaluated the results and wrote the paper. WS provided computational resources and reviewed the paper. RK served as corresponding author. All authors have read and approved the final version of the manuscript.

#### Funding

For the research costs no external funding was used. Publication costs were funded by the Medical University of Vienna.

#### Availability of data and materials

Data for MD simulations of the 3KPS system were downloaded from the PDB protein data bank: <https://www.rcsb.org/structure/3KPS>. Datasets generated by MD simulations are available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

Received: 4 July 2022 Accepted: 13 July 2022

Published online: 29 July 2022

## References

- Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol*. 2006;24:419–66.
- Madden DR. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol*. 1995;13:587–622.
- Kulski JK, Shiina T, Dijkstra JM. Genomic diversity of the major histocompatibility complex in health and disease. *Cells*. 2019;8(10):1270.
- Fodor J, Riley BT, Borg NA, Buckle AM. Previously hidden dynamics at the TCR-peptide-MHC interface revealed. *J Immunol*. 2018;200(12):4134–45.
- Macdonald WA, Chen Z, Gras S, Archbold JK, Tynan FE, Clements CS, et al. T cell allorecognition via molecular mimicry. *Immunity*. 2009;31(6):897–908.
- Natarajan K, Jiang J, May NA, Mage MG, Boyd LF, McShan AC, et al. The role of molecular flexibility in antigen presentation and T cell receptor-mediated signaling. *Front Immunol*. 2018;9:1657.
- Ayres CM, Corcelli SA, Baker BM. Peptide and peptide-dependent motions in MHC proteins: immunological implications and biophysical underpinnings. *Front Immunol*. 2017;8:935.
- Smith KJ, Reid SW, Stuart DI, McMichael AJ, Jones EY, Bell JI. An altered position of the alpha 2 helix of MHC class I is revealed by the crystal structure of HLA-B\*3501. *Immunity*. 1996;4(3):203–13.
- 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. *J Biol Chem*. 2013;288(34):24372–81.
- Borbulevych OY, Piepenbrink KH, Baker BM. Conformational melding permits a conserved binding geometry in TCR recognition of foreign and self molecular mimics. *J Immunol*. 2011;186(5):2950–8.
- Pöhlmann T, Böckmann R, Grubmüller H, Uchanska-Ziegler B, Ziegler A, Alexiev U. Differential peptide dynamics is linked to major histocompatibility complex polymorphism. *J Biol Chem*. 2004;279(27):28197–201.
- Beerbaum M, Ballaschk M, Erdmann N, Schnick C, Diehl A, Uchanska-Ziegler B, et al. NMR spectroscopy reveals unexpected structural variation at the protein-protein interface in MHC class I molecules. *J Biomol NMR*. 2013;57(2):167–78.
- Hee CS, Beerbaum M, Loll B, Ballaschk M, Schmieder P, Uchanska-Ziegler B, et al. Dynamics of free versus complexed  $\beta$ 2-microglobulin and the evolution of interfaces in MHC class I molecules. *Immunogenetics*. 2013;65(3):157–72.
- Narzi D, Becker CM, Fiorillo MT, Uchanska-Ziegler B, Ziegler A, Bockmann RA. Dynamical characterization of two differentially disease associated MHC class I proteins in complex with viral and self-peptides. *J Mol Biol*. 2012;415(2):429–42.
- Gur M, Golcuk M, Gul A, Erman B. Molecular dynamics simulations provide molecular insights into the role of HLA-B51 in Behçet's disease pathogenesis. *Chem Biol Drug Des*. 2020;96(1):644–58.
- Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J Mol Biol*. 1991;219(2):277–319.
- Knapp B, Deane CM. T-cell receptor binding affects the dynamics of the peptide/MHC-I complex. *J Chem Inf Model*. 2016;56(1):46–53.
- Tomasiaek L, Karch R, Schreiner W. Long-term molecular dynamics simulations reveal flexibility properties of a free and TCR-bound pMHC-I system. In: Proceedings of the 2020 IEEE international conference on bioinformatics and biomedicine (BIBM), 2020; 1295–302.
- Adler M, Beroza P. Improved ligand binding energies derived from molecular dynamics: replicate sampling enhances the search of conformational space. *J Chem Inf Model*. 2013;53(8):2065–72.
- Zacharias M, Springer S. Conformational flexibility of the MHC class I alpha1-alpha2 domain in peptide bound and free states: a molecular dynamics simulation study. *Biophys J*. 2004;87(4):2203–14.
- Kass I, Buckle AM, Borg NA. Understanding the structural dynamics of TCR-pMHC complex interactions. *Trends Immunol*. 2014;35(12):604–12.
- Wieczorek M, Abualrous ET, Sticht J, Alvaro-Benito M, Stolzenberg S, Noé F, et al. Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation. *Front Immunol*. 2017;8:292.
- Madden DR, Garboczi DN, Wiley DC. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell*. 1993;75(4):693–708.
- Tynan FE, Burrows SR, Buckle AM, Clements CS, Borg NA, Miles JJ, et al. T cell receptor recognition of a "super-bulged" major histocompatibility complex class I-bound peptide. *Nat Immunol*. 2005;6(11):1114–22.
- Armstrong KM, Piepenbrink KH, Baker BM. Conformational changes and flexibility in T-cell receptor recognition of peptide-MHC complexes. *Biochem J*. 2008;415(2):183–96.
- Alba J, Rienzo LD, Milanetti E, Acuto O, D'Abramo M. Molecular dynamics simulations reveal canonical conformations in different pMHC/TCR interactions. *Cells*. 2020;9(4):942.
- Frederick KK, Marlow MS, Valentine KG, Wand AJ. Conformational entropy in molecular recognition by proteins. *Nature*. 2007;448(7151):325–9.
- Polyansky AA, Zubac R, Zagrovic B. Estimation of conformational entropy in protein-ligand interactions: a computational perspective. *Methods Mol Biol*. 2012;819:327–53.
- Schlitter J. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem Phys Lett*. 1993;215(6):617–21.
- Karplus M, Kushick JN. Method for estimating the configurational entropy of macromolecules. *Macromolecules*. 1981;14(2):325–32.
- Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M. The double cube lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J Comput Chem*. 1995;16(N3):273–84.
- Horton N, Lewis M. Calculation of the free energy of association for protein complexes. *Protein Sci*. 1992;1(1):169–81.
- He HT, Bongrand P. Membrane dynamics shape TCR-generated signaling. *Front Immunol*. 2012;3:90.
- Wan S, Flower DR, Coveney PV. Toward an atomistic understanding of the immune synapse: large-scale molecular dynamics simulation of a membrane-embedded TCR-pMHC-CD4 complex. *Mol Immunol*. 2008;45(5):1221–30.
- Bello M, Correa-Basurto J. Energetic and flexibility properties captured by long molecular dynamics simulations of a membrane-embedded pMHCII-TCR complex. *Mol Biosyst*. 2016;12(4):1350–66.
- Bello M, Correa-Basurto J. Molecular dynamics simulations to provide insights into epitopes coupled to the soluble and membrane-bound MHC-II complexes. *PLoS ONE*. 2013;8(8):e72575.
- Wong-ekkabut J, Karttunen M. The good, the bad and the user in soft matter simulations. *Biochim Biophys Acta*. 2016;1858(10):2529–38.
- Kmieciak S, Gront D, Kolinski M, Wieteska L, Dawid AE, Kolinski A. Coarse-grained protein models and their applications. *Chem Rev*. 2016;116(14):7898–936.
- Kmieciak S, Kouza MA-O, Badaczewska-Dawid AE, Kloczkowski A, Kolinski AA-O. Modeling of protein structural flexibility and large-scale dynamics: coarse-grained simulations and elastic network models. *Int J Mol Sci*. 2018;19(11):3496.
- Cuendet MA, Michielin O. Protein-protein interaction investigated by steered molecular dynamics: the TCR-pMHC complex. *Biophys J*. 2008;95(8):3575–90.
- Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4(3):435–47.
- Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*. 2010;78(8):1950–8.

43. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. Intermolecular forces: proceedings of the fourteenth Jerusalem symposium on quantum chemistry and biochemistry held in Jerusalem, Israel, Apr 13–16, 1981. Dordrecht: Springer; 1981. p. 331–42.
44. Hess B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J Chem Theory Comput.* 2008;4(1):116–22.
45. Darden T, York D, Pedersen L. Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys.* 1993;98(12):10089–92.
46. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys.* 2007;126(1):014101.
47. Berendsen HJ, Postma JPM, Van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys.* 1984;81(8):3684–90.
48. Schäfer H, Mark AE, Gunsteren WF. Absolute entropies from molecular dynamics simulation trajectories. *J Chem Phys.* 2000;113(18):7809–17.
49. Kumar A, Purohit R. Use of long term molecular dynamics simulation in predicting cancer associated SNPs. *PLoS Comput Biol.* 2014;10(4):e1003318.
50. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996;14(1):33–8.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

