

In-Silico Design of Peptides for Inhibition of HLA-A*03-KLIETYFSK Complex as a New Drug Design for Treatment of Multiples Sclerosis Disease

Zahra Ghobadi

Shahrekord University

karim mahnam (✉ mahnam.karim@sku.ac.ir)

Shahrekord University <https://orcid.org/0000-0002-9906-4957>

Mostafa Shakhshi-Niaei

Shahrekord University

Research

Keywords: Multiple Sclerosis, HLA-A*03, Peptide Design, T Cell Receptor (TCR), Molecular Dynamics Simulation, MM/GBSA Binding Free Energy

Posted Date: June 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-629231/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Journal of Molecular Graphics and Modelling on November 1st, 2021. See the published version at <https://doi.org/10.1016/j.jmgm.2021.108079>.

Abstract

Multiple sclerosis is recognized as a chronic inflammatory disease. Human leukocyte antigen (HLA) plays an important role in initiating adaptive immune responses. HLA class I is present in almost all nucleated cells and presents the cleaved endogenous peptide antigens to cytotoxic T cells. HLA-A*03 is one of the HLA class I alleles, which is reported as substantially related HLA to MS disease. In 2011, structure of the HLA-A*03 in complex was identified with an immunodominant proteolipid protein (PLP) epitope (KLIETYFSK). This complex has been reported as an important autoantigen-presenting complex in MS pathogenesis. In this study, new peptides were designed to bind to this complex that may prevent specific pathogenic cytotoxic T cell binding to this autoantigen-presenting complex and CNS demyelination. Herein, 14 new helical peptides containing 19 amino acids were designed and their structures were predicted using the PEP-FOLD server. Binding of each designed peptide to the mentioned complex was then performed. A mutation approach was used by the BeAtMuSiC server to improve binding affinity of the designed peptide. In each position, amino acid substitutions leading to an increase in binding affinity of the peptide to the mentioned complex were determined. Finally, the resulting complexes were simulated for 40 ns using AMBER18 software. The results revealed that out of 14 designed peptides, "WRYWWKDWAKQFRQFYRWF" peptide exhibited the highest affinity for binding to the mentioned complex. This peptide can be considered as a potential drug to control multiple sclerosis disease in patients carrying the HLA-A*03 allele.

Introduction

Approximately 2.5 million individuals suffer from multiple sclerosis (MS) worldwide, as a persistent inflammatory neurological disorder of the brain and spinal cord, which is a prevalent cause of significant physical impairment, fatigue, and pain in young adults, especially women [1].

Differences in clinical symptoms are correlated with spatiotemporal distribution within pathological sites of lesions in the central nervous system (CNS) [2]. These lesions are characteristic of MS and are caused by the blood-brain-barrier (BBB) crossing of immunogenic cells, which causes inflammation, demyelination, gliosis, and neuroaxonal degeneration, finally leading to neuronal death [3]. Macrophages dominate the infiltrate, followed by CD8⁺ T cells, while fewer CD4⁺ T, B, and plasma cells can also be found [4]. Many overlapping genetic and environmental risk factors contribute to an immune response to CNS myelin antigens [5]. Several alleles of the human leukocyte antigen (HLA) or major histocompatibility complex (MHC) on the short arm of chromosome 6 [6p21.3] were identified as basis of genetic susceptibility to MS [6]. MHC class I (HLA A, B, and C) molecules present endogenous peptides to CD8⁺ cytotoxic cells. The CD4⁺ T cells recognize the processed exogenous antigens located on MHC class II (HLA DRB1, DPB1, and DQB1) alleles existing on antigen-presenting cells, such as B cells to stimulate their antibody production [7]. The influence of both HLA classes of I and II genes on MS susceptibility may reflect important role regarding antigen recognition of HLA class I-dependent CD8⁺ T cells and HLA class II-dependent CD4⁺ T cells in pathophysiology of MS [8].

CD8⁺ T cells have a major role in pathogenesis of autoimmune diseases like MS. CD8⁺ T cells recognize antigen peptides on MHC class I via the clonotypically-expressed $\alpha\beta$ TCR and the lineage-specific CD8 co-receptor [9]. These cells can then kill target cells presenting the peptide antigens. TCR-HLA crystal structures of the peptide complex, combined with functional experiments have demonstrated that the complementarity-determining region (CDR) loops of TCR engage both the presented peptide and the HLA presentation platform in a generally conserved diagonal orientation, with the TCR- α chain position over the MHC- α_2 helix and the TCR- β chain over the MHC- α_1 helix (Fig. S1) [10, 11].

HLA-A3 (HLA-A*03) is one of the class I HLA alleles, which is firmly related to MS. The presence of this allele has been reported as a predisposing risk factor for patients with MS in Khuzestan Province, Iran [12]. Supplementary effect of class I HLA loci (e.g., HLA-A*03) has been described to double the risk of MS [13, 14].

In 2011, structure of the HLA-A*03 in complex with an immunodominant proteolipid protein (PLP) epitope (KLIETYFSK) was determined [15]. PLP makes up 50% of whole myelin sheath protein [16]. Protein-protein interactions (PPIs) play pivotal roles in majority of biological processes. Therefore, the improved approaches to target and disrupt PPIs would provide tools for chemical biology and lead to therapeutic development [17]. Peptide-based inhibitors have more advantages than small-molecule drugs. Firstly, targeting protein interactions using small molecules is considered to be difficult due to size of proteins and the lack of well-defined binding pockets [18, 19]. Secondly, the peptide has high specificity towards a target and may not accumulate in organs. This would minimize the side effects and toxicity [20]. Peptides modulating PPIs can be directly derived from crystallographic interface, or through screening of peptide sequences that do not originate from natural proteins using phage display library screening protocol [21].

Besides, computational methods combining simulations of molecular dynamics and calculations of binding free energy could present both structural and energetic outlook of peptide inhibitors to evaluate binding affinity of peptides [21]. In the present study, it is attempted to prevent formation of trimolecular complexes (HLA-A*03-PLP₍₄₅₋₅₃₎-TCR) and then, inhibit proliferation of the activated T cells. In other words, herein, rational design of binding peptides towards HLA-A*03-PLP₍₄₅₋₅₃₎ complex is studied.

Computational Methods

Structure Preparation

X-ray crystallographic structure of the complex of HLA-A*03 with PLP (protein data bank (PDB) code: 2XPG) (Resolution: 2.60 Å) was obtained from the research collaboratory for structural bioinformatics (RCSB) site (<http://www.rcsb.org/pdb>). The particular PDB entry was selected because it contains the main immunodominant epitope PLP₄₅₋₅₃ involved in MS, as well as an HLA-A*03 Allele from a patient with MS [15]. This structure contains three chains (A, B, and C). Chain A (α chain of HLA-A*03 Allele) and chain B (β_2 -microglobulin), and chain C (PLP₄₅₋₅₃) have 274, 98, and 9 residues, respectively. The TCR recognizes both immunodominant epitope PLP and the HLA-A*03 molecules. Normally, chain B or β_2 M has a non-covalently link with chain A in HLA class I structure. For reducing computational cost, chain B was omitted in this study. The p-HLA complex containing chains A and C is described in this study.

Identification of the HLA Hot-Spot Residues for Binding to TCR

The lack of available triple molecular crystal structures is a serious obstacle to determine the most important residues of the mentioned p-HLA complex for binding to TCR. Initially, sequence similarity of HLA-A*03 alleles with other class I, HLA proteins was investigated. Sequence similarity was determined using pBLAST (<https://blast.ncbi.nlm.nih.gov/pBlast>) and Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) tools. Sequence identity of HLA-A*03 alleles with HLA-A*11 and HLA-A*02 was equal to 97 and 93%, respectively. The used TCR databanks included TCR structural repertoire database (<https://tcldb.ibbr.umd.edu/>) [22], VDJdb web browser (<https://vdjdb.cdr3.net>) [23], STCRDab database (<http://opig.stats.ox.ac.uk/webapps/stcrdab/>) [24], and immune epitope database and analysis resource (<http://www.iedb.org/>) [25]. The selected complexes for this study consisted of an MHC molecule (HLA-A*02 or HLA-A*11), a 9- or 10-mer peptide, and a TCR. Both TCR subunits were different in the TCR beta variable gene (TR β V) and TCR alpha variable gene (TR α V). Only 12 structures that were different in CDR1, CDR2, and CDR3 sequences were retained (Table S1). Inter-subunit interactions of p-HLA-TCR selected complexes were analyzed by three different computational approaches. Firstly, Robetta, a fast computational alanine scanning method was used to predict energetically important amino acids in the selected p-HLA-TCR complex (<http://robetta.bakerlab.org/alaninescan>) [26]. Secondly, knowledge-based FADE and Contacts2 (KFC2) server was used to predict hot-spots by analyzing chemical and physical properties in surrounding p-HLA-TCR PPIs interface (https://mitchell-web.ornl.gov/KFC_Server/index.php) [27]. Finally, the HotRegion database was used to predict the hot-spots by accessible surface area and knowledge-based pair energies of each amino acid residue in the target complex (<http://prism.cccb.ku.edu.tr/hotregion/>) [28]. Web-based servers used their characteristic calculations and automatically predicted the hot-spots in inter-subunit interactions of the p-HLA-TCR selected complexes. These computational approaches allowed us to predict possible points of interaction with TCR located on the α_1 and α_2 -helix of α chain of the HLA-A*03 molecule. Interface residues indicated in this step with 9 residues of PLP epitope were selected as binding site for all peptides docking (active residue) in ClusPro and HADDOCK servers.

In-Silico Design of Peptide

The first peptide for inhibition of HLA-A*03-PLP₍₄₅₋₅₃₎ was designed innovatively by the above-described crystal structures interface contacts. TCR-HLA interfacial contacts were analyzed by the PISA server [29], using a cut-off value of 4 Å for contacts. Considering length of α_1 and α_2 -helix of HLA-A*03 molecule on which the hot/warm spots were distributed, some helical 19-mer peptides were designed in this research. Before docking, structure of the first peptide was modeled using the PEP-FOLD server (<https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/>) [30]. The best 3D model was selected according to the PEP-FOLD server, considering the lowest energy model that indicates peptide stability [31]. The binding affinity obtained from the docking results suggested that these peptides were not suitable and they required to be improved. To this end, several steps of affinity maturation were performed by the BeAtMuSiC server (<http://babylone.ulb.ac.be/beatmusic/>) [32]. This server employed a mutation strategy, and any residue was substituted with all other 19 amino acids separately. This server uses the known structure of a protein-protein complex to evaluate change in binding affinity between two proteins caused by single mutations in their sequence [33]. Each substitution was accompanied by an evaluation of peptides binding affinity and several physicochemical properties, using ClusPro docking server [34] and ProtParam (<https://web.expasy.org/protparam/>) web tool [35]. ClusPro assigns a weighted ranking to binding energy of the protein complex. In parallel, molecular docking was separately conducted using the HADDOCK server to validate interaction mode of each peptide-HLA-A*03-PLP₍₄₅₋₅₃₎ complex. This server is a

powerful docking platform that uses a data-driven approach supporting a wide variety of experimental data [36]. Molecular docking was used for obtaining the most possible binding pose for each designed peptide and determining probable binding site of each peptide into the mentioned p-HLA complex. Finally, at each position, amino acid substitutions leading to an increase in binding affinity and stability of the peptides to the p-HLA complex were distinguished and mutations were subsequently used in combinations to design the second-generation peptide libraries. Similarly, subsequent peptide libraries were generated until a peptide that had more binding affinity relative to reference peptide was identified.

Several physicochemical properties of peptides including molecular weight, net charge at pH=7, pI, aliphatic index, grand average of hydropathicity index (GRAVY) were measured using ProtParam and PepCalc methods [37]. Generally, a positive GRAVY value implies hydrophobicity while a negative GRAVY value suggests a protein's hydrophilicity degree [38]. Besides, antigenicity of the peptide sequence was predicted using the VaxiJen 2.0 server [39] and the predicted antigenic peptides server (<http://imed.med.ucm.es/Tools/antigenic.pl>). Finally, 14 proper peptides were selected for investigation via MD simulation (Table 1).

MD Simulations

The AMBER18 MD simulation package [40] was used to perform MD simulations of p-HLA complex bound to each of 14 different peptides (Table 1). Calculations were conducted using the SANDER GPU-accelerated version [41] with Amber ff14SB force field [42], in periodic boundary conditions. In all cases, the peptide was neutralized by adding sodium ions and was solved in a truncated octahedron box filled with TIP3P water model molecules [43], which led to formation of a solvent shell with a distance of at least 10 Å between each box face and the solute using the tleap module. About 7,000 water molecules were added in solvation stage. The total number of atoms was about 35,000 for each system. Structural minimization was performed in two steps to eliminate any steric conflicts. At first, water molecules and sodium were relaxed with harmonic constraint potential of 2.0 kcal/molÅ² on all atoms of both p-HLA protein and designed peptides. Then, the entire system was subsequently minimized by 5,000 steps of the steepest descent minimization followed by 5,000 steps of conjugate gradient minimization. Langevin thermostat was used [44] with coupling coefficient of 2 ps and force constant of 2 kcal/molÅ² on the complex. After minimization, MD simulations were initiated by gradually heating each system in canonical ensemble (NVT) from 0 to 300 K in three steps for 200 ps. Restraint constant of heating process was equal to 100, 10, and 1 kcal/molÅ², respectively. Then, density equilibration of 100 ps was performed on the complex by removing all the restraints in isothermal–isobaric (NPT) ensemble. Berendsen barostat with a pressure-coupling constant of 2 ps was used to maintain pressure at 1 bar [45]. Ultimately, production runs were carried out for 40 ns of MD simulations in the NPT ensemble at constant temperature of 300 K for each complex. During MD simulation, particle mesh Ewald (PME) approach was used to treat long-range electrostatic interactions [46] and the SHAKE algorithm restricted all covalent bond lengths involving hydrogen atoms [47]. The cut-off distance for long-range van der Waals (VDW) interaction term was equal to 10 Å. Time step was set to 2 fs, and trajectory coordinates were recorded every 2 ps. Post-MD analyses were performed, using the CPPTRAJ module in Amber18 software [48]. Visualization of trajectories was performed using VMD software [49].

Free Energy Calculations and Per-Residue Contributions

For investigating interaction of each peptide from energetic perspective, binding free energy was calculated based on trajectories of MD simulations by molecular mechanics/generalized born surface area (MM/GBSA) and molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) methods using mmpbsa.py module [50]. These calculations were carried out following the “single trajectory method”, where snapshots of binding partners were extracted from MD trajectories of pHLA-peptide complexes. The single trajectory method neglects energetic contributions due to conformational changes leading to a drastic reduction in statistical uncertainty of free energy components [51]. Binding free energy is a significant thermodynamic parameter that provides the detailed information about the peptide-protein interaction. It also offers a clear understanding of binding mechanism including their different energy contributions to molecular identification, such as enthalpy and entropy. For further exploring the detailed interaction information of each peptide, free energy decomposition was performed using MM/GBSA method to identify key residues responsible for binding energy. Free energy decomposition calculations were done during the last 10 ns of MD trajectories using a single trajectory approach averaged over 1,003 snapshots with 10 ps intervals. The residues contributing less than -2 kcal/mol were considered to be very significant for binding of p-HLA complex with peptide and these residues were designated as hot-spot amino acids.

During simulation process using the CPPTRAJ module, the number of hydrogen bonds formed versus simulation time was calculated to detect complex stability and confirm important active site residues interacting with each peptide. Maximum hydrogen-acceptor distance of 3.5 Å and a minimum hydrogen donor-acceptor angle of 135° were allocated to hydrogen bonds [52-54]. The frames adopted for calculation were taken from the entire 40 ns of trajectories at intervals of 2 ps. Furthermore, for exactly determining how hydrogen bonds play dominant roles in maintaining peptide stability in MD simulations, hydrogen bond occupations formed during this period were

calculated [53]. The occupations vary from 1 to 100%, and a higher percentage represents a more stable hydrogen bond. Hydrogen bonds were only considered if their occupancies were more than 20%.

Results And Discussion

Docking and Structural Properties of the Peptides

In this study, residues of 1-283 refer to the HLA-A*03 and residues of 55-85 refer to the α_1 -helix and residues of 139-175 represent the α_2 -helix of the HLA-A*03 allele, and residues of 275-283 show the PLP immunodominant epitope in HLA-A*03- PLP_[45-53] complex (p-HLA complex) (Fig. 1). Also, residues of P1-P19 refer to positions 1-19 of each peptide. P-HLA complex in the presence of each peptide (triple complexes) is named as complexes of 1 - 14.

Docking of the peptide into the p-HLA binding site is one of commonly used molecular modeling methods for determining the desired orientation of each peptide. Totally, 14 designed peptides were docked into the p-HLA complex at the TCR-binding interface. The obtained active residues for HADDOCK server explained in Methodology Section were residues of 58, 62, 63, 64, 65, 66, 67, 69, 71, 72, 73, 74, 75, and 76 from α_1 helix and residues of 146, 147, 152, 153, 154, 155, 157, 158, 159, 161, 163, and 167 from α_2 helix and residues of 275-283 from PLP neuroantigen. Also, all 19 residues of the designed peptides were considered as active residues for the HADDOCK server. The predicted binding score is presented in Table 1 for each peptide calculated by the HADDOCK web server.

The results of docking showed that the peptide 14 (WRYWWKDWAKQFRQFYRWF) had the highest binding affinity than other peptides. This relatively high binding affinity exhibited by peptide 14 can be significantly attributed to replacement of Glu residues with Tyr at location 3 and Phe with Ala at position 9 of peptide 1 (WREWWKDWFKQFRQFYRWF). Of course, residue 3 is more important than residue 9, because it has more contribution in hydrogen binding formation between peptide and p-HLA complex. Selected peptides were assessed for grand average hydropathicity (GRAVY) and antigenic region. Positive and negative GRAVY values represent hydrophilicity and hydrophobicity, respectively. All peptides had a strong degree of hydrophilicity (negative GRAVY value) without any antigenic region in the peptides' sequence (Table 1). Increased peptide hydrophilicity may lead to passage across the blood-brain barrier (BBB) membrane.

Results of MD Simulations

Sometimes, docking results are unreliable because some of the best-docked conformations become separate from the protein within a short time of MD simulation. Thus, herein, MD simulations were performed to make sure that the docked peptide stays at the p-HLA binding site.

Changes in root mean square deviation (RMSD) relative to their initial structure during simulation showed stability of backbone atoms (Fig. 2(a)).

The average values of backbone RMSD of p-HLA are summarized in Table 2. P-HLA complex in the presence of each peptide was characterized with an initially high RMSD fluctuation followed by a decrease in values of backbone RMSD during the last 30 ns (Fig. 2(a)). The highest backbone RMSD of p-HLA belonged to system 14 (p-HLA in complex with peptide 14). However, the average backbone RMSD value of the p-HLA complex indicated strong fluctuations in certain regions. Then, RMSD of the residues forming the α_1 -helix (Fig. S2(a)), α_2 -helix (Fig. S2(b)), and PLP-bound neuroantigen (Fig. S2(c)) of the p-HLA complex was calculated separately (Table 2). Comparing Fig. 2(a) with Figs. S2(a-c) showed that RMSD values of these three regions were much lower than those of the p-HLA in the complex. Since, the alpha chain of class I HLA molecules consists of three domains of α_1 , α_2 , and α_3 , it was concluded that high fluctuation observed in Fig. 2(a) will be attributable to α_3 helix. In this study, the β_2 M chain was omitted to decrease the computational cost. The absence of the β_2 M domain in simulation led to high RMSD of the α_3 domain. In class I of HLA molecules, it was found that β_2 M interacts with all three domains of α_1 , α_2 , and α_3 whereas the α_3 domain is mostly responsible for binding to β_2 M [55]. As shown in Figs. S2(a-c), the three mentioned regions were distinguished by RMSD fluctuations of less than 2.5 Å, followed by a relative decrease in RMSD. After approximately 30 ns, RMSD values converged to about 2 Å.

Furthermore, RMSD of all peptides experienced various degrees of fluctuations at first but gradually tended to converge (Fig. 2(b)). Also, probability distribution of RMSD values over the last 10 ns of simulation was calculated for all RMSD values (data not shown) and the Gaussian shape of probability showed no statistical errors in simulation time and sampling.

Besides, in 14 simulated complexes, the RMSD values of the backbone atoms of PLP peptide varied between 0.2 - 1.3 Å (Fig. S2(c) and Table 2). Since, the PLP was surrounded by HLA residues from all directions, α_1 and α_2 helices and beta floor of HLA, and peptides'

residues, there was hardly a space for conformational movements and fluctuation in MD simulation (Table 2).

Root Mean Square Fluctuation (RMSF)

For understanding fluctuations in residues and investigating the effect of peptide binding on p-HLA complex dynamics, RMSF of residues during the last 10 ns of MD trajectories was calculated and the results are presented in Fig. 3(a).

Also, using the formula mentioned in the literature [56], overall fluctuations (F_{overall}) or total flexibility for the three described regions, peptides, and p-HLA complexes were calculated. Binding peptides led to different F_{overall} values in the p-HLA complex and change in the overall structure of p-HLA (Fig. S3, Table 2). In the presence of peptide 14, the p-HLA complex had the highest F_{overall} value relative to other peptides. In other words, peptide 14 induced the most structural fluctuation in p-HLA complex. P-HLA complex in interaction with peptide 11 had the lowest F_{overall} value and there were slight changes in the p-HLA complex vibration.

Also, interaction of different peptides with PLP peptides in the HLA binding site led to different fluctuations in PLP peptides. In fact, each peptide represents particular new positioning of PLP (antigenic peptide presentation mode) in HLA-A*03 binding grooves. Fig. S3 provides a comparison of the PLP conformation in the HLA-A*03 binding groove in the presence of each peptide. The least and the most F_{overall} values of PLP belonged to binding peptides of 7 and 12, respectively. In other words, binding peptide of 12 led to more fluctuation or more structural change in PLP in the binding groove of HLA relative to other peptides. Madura in a study showed TCR-induced alteration in interaction of HLA with anchoring antigenic peptide (PLP). T cell antigen recognition required an anchor residue shift at N-terminus of the antigenic peptide [57]. Results of another study showed that latter portion of the antigenic peptide is "lifted" from the binding groove by 1–1.5 Å upon TCR binding, beginning at residue of 4 and continuing through the C-terminal residues [11]. Our results also indicated that binding of the designed peptides changes F_{overall} of PLP peptide (Residues of 275-283 in Fig. 3(a)). It may reflect weaker interactions of peptides with class I HLA proteins at their C-terminal ends [58]. Similarly, in the presence of peptides 3 and 11, the C-terminal of PLP detaches from their binding pockets and stretches outside the HLA-A3 binding groove because, these peptides make strong hydrogen bonds with C-terminal of PLP (Table S3).

Motions or dynamics of α_1 - α_2 helices and β -strand floor are primarily influenced by each peptide binding [58]. Our results showed that the RMSF values of α_1 and α_2 helices and binding groove ($\alpha_1+\alpha_2+\beta$ -strand floor) changed in binding of different peptides (Table 2). In the presence of peptides 11 and 13, HLA-A*03 binding groove showed the lowest and most F_{overall} values, respectively. Then, peptide 13 induced the most structural changes in binding groove of HLA relative to other peptides. Knapp et al., showed that only flexibility of α_{2-1} and α_{2-2} residues changed in interaction with TCR [59].

Also, RMSF of the designed peptides was calculated during the last 10 ns of MD simulation and the results are presented in Fig. 3(b) and Table 2. The least F_{overall} value of peptides belonged to peptides 7 and 11. Then, these peptides had the least fluctuation in binding to p-HLA complex. Also, the most fluctuation belonged to peptide 12.

Radius of Gyration (RG)

In the presence of each peptide, level of compaction of α_1 and α_2 helix of the HLA-A*03 molecule and PLP was determined by RG (Rg) so that, lower Rg indicates more compactness of the p-HLA complex. Fig. 4 and Table 3 show RG of p-HLA complex and each peptide measured over the last 10 ns of simulation. The results indicated partial structural changes in p-HLA complex in binding of different peptides. According to Fig. 4, it was concluded that the p-HLA complex was more compact while binding to peptide 11 relative to other peptides. Compaction of p-HLA complex in binding to peptide 11 was compatible with the lowest F_{overall} value of p-HLA complex for complex 11 (Table 2). In other words, compaction can lead to low fluctuation.

According to Table 3, the PLP was partially uncompact while binding of peptide 14 to p-HLA complex relative to other peptides. The most Rg value of PLP peptide was seen in triple complex 14 that was compatible with the highest F_{overall} value in PLP peptide in complex with peptides 12 and 14 (Fig. S4 and Tables 2 and 3).

Average values of Rg were less for α_1 -helix than α_2 -helix as the second α_2 -helix was not completely straight and contained kinks at residues of 150 and 165 such that, one can distinguish three α -helical sub-segments (Fig. 1). The most Rg and F_{overall} values of α_2 helix belonged to complex containing peptide 13. Then, these parameters were compatible in α_2 helix. (Fig. S4 and Tables 2 and 3).

Fig. S4 and Tables 2 and 3 show that binding of the designed peptides can lead to induction of different tertiary structures in p-HLA complex and PLP and α_1 and α_2 structures. Radius of the designed peptides was similar, but the least and most average Rg values of the

designed peptides belonged to peptides 8 and 7 (Table 3).

Solvent Accessible Surface Area (SASA)

SASA shows the amount of exposure of a given region to solvent medium. Figs. S5(a-c) depict SASA values of α_1 and α_2 helices, and PLP versus simulation time, also the average values of SASA are summarized in Table 4. Here, decreased values of SASA denote to coverage of the mentioned regions by peptides. The results of analyzing SASA suggested that peptides 5, 11, and 1 decreased SASA of PLP more than other peptides, in other words, these peptides covered surface of PLP peptide better than other peptides and could inhibit TCR binding to p-HLA complex. In the presence of peptide 3, PLP peptide had the most SASA and the least coverage with PLP peptide.

The most coverage (the least SASA) and least coverage of α_1 -helix belonged to peptides 9 and 13, respectively.

Peptides 8 and 14 had the least and most SASA values during the last 10 ns of MD simulation, respectively. Similar to Rg results, the estimated SASA values for the α_1 -helix were less than α_2 -helix values. Optimal α_2 helix coverage was not feasible due to kinks at residues of 150 and 165.

Center of Mass Distance

The distance between center of mass (COM) of the p-HLA complex and peptides was calculated. Fig. 5 demonstrates the distance based on all Ca atoms plotted versus simulation time. The average COM values of the last 10 ns of the simulation are summarized in Table 5. The least and most COM distance values belonged to peptides 9 and 2, respectively. Of course, the average COM is not a precise parameter, because it depends on peptide size and initial position (obtained from docking). Also, probability distribution of center of distances was obtained (the plot was not shown). The Gaussian shape of this plot showed no statistical error in MD simulation and sampling.

Binding Free Energy

The mentioned parameters, such as RMSD, RMSF, SASA, and COM are qualitative parameters for investigation of peptide binding to p-HLA complex. For investigating binding mode of peptides to p-HLA complex quantitatively, binding free energy was calculated during the last 10 ns of MD simulation after MD simulation. In a previous study, Hou et al., [60] showed that MM/GBSA calculation was more proper for indication of relative binding free energies. Considering computational efficiency of the MM/GBSA approach, it was decided to calculate binding free energy by this approach (Table 5). The best negative binding free energy belonged to peptide 14 with a value of about -75 kcal/mol indicating favorable protein-peptide interaction. The calculated values were overestimated because of the lack of entropic and energetic contributions. Entropic contributions were unfavorable in the case of flexible peptides. Energetic contributions or conformational changes occur due to using the single trajectory approach [61]. Binding free energy results were compatible with docking results.

Also, for investigating contributions of binding free energy of each residue, energies are decomposed into individual residue contributions using the MM/GBSA approach. This type of calculations not only rationalize molecular recognition processes but also can guide identification of residues that mimic determinants of binding in protein-protein complexes (hot-spots)[62]. Contribution of residues in the total binding free energy of each peptide is presented in Figs. S6-18. Only those residues whose contributions to effective energy (ΔG_{eff}) were equal or less than -2 kcal/mol were considered. For understanding the main binding driving forces of peptides to the p-HLA complex, ΔG values were divided into two terms: electrostatic interaction plus polar solvation energies ($\Delta E_{\text{ele}} + \Delta G_{\text{GB}}$) and Van der Waals interaction energy plus non-polar solvation energies ($\Delta E_{\text{vdw}} + \Delta G_{\text{SA}}$)[63].

For instance, in system 14, the main energy contribution belonged to the residues of Arg 2, Tyr 3, Arg 17, Trp 8, Trp 4, Trp 18, Phe 12, Phe 15, and Trp 5 of peptide and residues of Arg 65, Phe 281, Glu 58, Gln 62, and Glu 152 of the complex (Fig. 6). Electrostatic interaction and polar solvation energy (ΔG_{GB}) played the main roles in binding of the residues of Arg 2 and Arg 17 of peptide, and Glu 152 of the complex. Binding energy contribution of the other residues was mainly due to van der Waals interaction (ΔE_{vdw}) and non-polar solvation energy (ΔG_{SA}) (Fig. 6).

Analysis of Hydrogen Bond

Hydrogen bonds play an important role in stabilizing peptides inside the HLA binding groove [59]. Loss of residue-residue hydrogen bond connections is a potential explanation of the increased residue flexibility (F_{overall}) as noted in Table 2. McMahon et al., showed a hydrogen bond network between PLP peptide and HLA-A3 peptide-binding groove in X-ray structure of the p-HLA complex [15]. Hydrogen bonding occupancy in the presence of each peptide between the PLP and HLA residues is shown in Table S2 of supplementary material,

throughout 40 ns of MD simulation. Also, the average of number of hydrogen bonds was calculated during the last 30 ns of MD simulation in this study. During MD trajectory, these hydrogen bonds fully changed in the presence of peptides in most of them. For instance, in the presence of peptide 14, it was found that Glu 4 of PLP peptide made a hydrogen bond with Tyr 171 of HLA chain. The newly developed H-bond was defined with high occupancy of 85% and an average H-bond distance of 2.75 Å. McMahon et al., demonstrated a conventional hydrogen bond with Tyr 171 formed with N-terminal of PLP residues [13]. In our study, a pocket was occupied by the Glu 4 of PLP. As can be seen in Table S2, instead of Tyr 7, stated in the study by McMahon et al., N-terminal residue (Lys 275) of PLP preferred to form a hydrogen bond with Arg 6 and N atoms of the binding groove (HB occupancy= 53%).

For studying origin of stronger binding affinity of peptides to p-HLA complex in detail, hydrogen bond was analyzed for all peptides. The average number of hydrogen bonds was calculated between each peptide and the p-HLA complex during simulation (Fig. 7). Peptide 14 exhibited the highest total H-bond occupancy between the peptides and the p-HLA complex (547%) compared to other peptides (Table 5). Calculation of the average of number of hydrogen bonds between peptide 14 and p-HLA complex during simulation time indicated that tyrosine substitution at location 3 of peptide 14 provided appropriate distance and angle for growth of more hydrogen bonds over simulation time (Figs. 7 and 8). Within peptide 14, it was found that the formed hydrogen bonds between Arg 2 of peptide and Arg 65 from HLA's α_1 -helix exhibited high H-bond total occupancy of 130 % with an average H-bond distance of 2.8 Å. As can be seen in Table 6, the stated hydrogen bonds were created by the backbone carbonyl group of Arg 2 whereas, Arg 2 and sidechain atom formed hydrogen bonds with Glu 58 atoms from HLA's α_1 -helix (H-bond occupancy= 69 %). H-bond distance was responsible for high bond strength in peptide 14 complex. There were five hydrogen bonds in peptide 14 that reached more than 50% of their occupation while 3, 3, 2, 2, and 3 hydrogen bonds with the mentioned occupancy percentages were found in peptides of 11, 10, 12, 5, and 3, respectively.

H-bond distance and occupancy of 14 respective peptides throughout simulation time are presented in Table S3 for other peptides. In peptide 3, two of three strong hydrogen bonds resulted from interaction of C-terminal residues of PLP with Arg 2 of peptide 3.

Similarly, in peptide 11, hydrogen bonds between C-terminal of PLP (Lys 283) and Arg 2 of peptide 11 played an important role in hydrogen bond strength and increase in total occupancy (Table S3).

Strength of the mentioned hydrogen bonds was well confirmed by free energy decomposition analysis (Fig. S5-S17). In peptides of 3 and 11, the total binding energy contribution of Arg 2 was equal to -14.3 and -16.2 kcal/mol, respectively.

As shown in Table S2, in the presence of peptide 7, not only C-terminal of PLP was properly positioned inside the F-pocket (ASP 74 and ASP 77) but also position 3 of PLP was able to maintain the hydrogen bond described in the study by McMahon et al., and hydrogen bond was formed between Ile3 of PLP and Tyr 99 of C-pocket confirming low F_{overall} value of PLP in binding of peptide 7 to p-HLA complex (Table 2).

Without taking into account entropic contribution, the calculated MM/GBSA energies were equal to the available experimental enthalpies. MM/GBSA binding free energy of peptide 14 with p-HLA was about -75.1 kcal/mol, which is in fair agreement with the TCR-pHLA complex enthalpies, usually ranging from -10 to -30 kcal/mol, measured by van't Hoff analysis or calorimetric experiments [64] and theoretical procedures. Also, enthalpies values reported for TCR-HLA-A2-(MP₅₈₋₆₆) and TCR-HLA-A2-SL9 ternary complexes reactions were equal to -23 and -10.4 kcal/mol, respectively [65]. Non-covalent interactions at PPI sites consist of van der Waals interactions, hydrogen bonds, and salt bridges. Interaction surfaces of both TCR and HLA-peptide complex are large so that, approximately a thousand square angstroms are buried when an interaction event occurs and most of the contacts are mediated by van der Waals interactions. Both hydrogen bonds and salt bridges, which are comparatively rare in TCR and p-HLA PPI interfaces, add specificity because of their dependence on stereochemistry and geometry of the interactions. Although, TCR-p-HLA interactions bury large amounts of surface area, they seem to have relatively less optimal shape complementarity (dissociation constants in the micromolar range) relative to cytokine-receptor interactions (dissociation constants in the nanomolar range); this low shape complementarity can lead to low average affinity of TCR-p-HLA interactions [66]. On the other hand, PPIs are extremely difficult to target because of their high surface area, and they are normally flat and shapeless, characterizing the interaction [57].

Murray described a conserved HLA motif (a.a., R65-X-X-K-A-X-S-Q72) and found that the TCR's CDR3 α loop hooks up to R65-joint and this position predisposed the HLA-A lineage to TCR alloreactivity. Furthermore, CDR3 α loop hooks up to Arg 65-joint and maintains its CDR2 α loop at a distance of about 4 Å from polymorphic amino acid positions of the α_2 helix [67]. As a result, blocking Arg 65 of the conserved motif of p-HLA with Arg 2 of peptide 14 (Table 6) led to inhibition of binding of TCR to the p-HLA complex. Besides, an overall H-bond occupancy of 194.7% was observed between Arg 17 of peptide 14 and Glu 152 of α_2 -helix (Table 6). HLA class I supertypes have been classified according to similar peptide-binding motifs. HLA-A*0301, -A*1101, -A*3101, -A*3301, and -A*6801 alleles belong to the A3-like

superfamily [68]. Here, the HLA-A*03 allele sequence was compared with other A3 superfamily alleles and it was found that at position 152, only the HLA-A*03 allele has glutamine residue. Specificity of the peptide 14 is explained based on association with sites that are distinctive to HLA-A*03. Furthermore, T cell maturation optimizes the α/β TCR repertoire for recognition of major histocompatibility complex (HLA) class-I polymorphism [67]. The interaction with α_2 -helix was accomplished via a remarkably compatible conjunction involving Gln155 as stated in the study conducted by Murray. Carbonyl group of Gln155 formed partially hydrogen bonds with Gln 11 and hydrogen atoms of peptide 14 (H-bond occupancy: 20%). Since, proximity of CDR2 α loop over HLA (residues of 151–158) is essential for interaction with α_2 helix, it appears that due to a strong hydrogen bond with Glu 152, the α_2 helix is not recognizable by TCR. Furthermore, Murray stated that HLA-A24, -23 does not have Arg 65, but interestingly, has conjunction involving CDR3 α contact with Glu 62 of the α_1 helix. As shown in Table 6, a strong hydrogen bond formed between Trp 4 of peptide 14 and Gln 62 of p-HLA complex (H-bond occupancy: 61%). Also, Gln 62 was partially occupied with Trp 5 atoms. Therefore, peptide 14 can form strong hydrogen bonds with polymorphic and non-polymorphic contact sites of TCR. Interaction of HLA's residue with peptide 14 is compatible with results of a former study on a TCR-pHLA triple complex model [11, 69, 70]. Tripathi identified hydrogen bond patterns and electrostatics in IG4TCR-HLA-A2-tumor epitope (NY-ESO) ternary complexes interface with six hydrogen bonds. In the study by Tripathi, three hydrogen bonds with H-bond occupancy of 90%, one case with H-bond occupancy of 77%, and two cases with H-bond occupancy of 40-50% were observed [71]. The hydrogen bond developed between peptide 14 and the p-HLA complex in the present research showed higher total HB occupancies compared to those reported in the study by Tripathi (total HB occupancies: 547%). This is consistent with free energy decompositions, which revealed that, in general, peptide residues of P2, P17, P8, P4, P18, P12, P15, and P5 had the largest contribution to effective free energy, respectively (Fig. 6). Studying free energy decomposition also revealed that van der Waals interactions and non-polar part of solvation free energy represent binding strength of the peptides, whereas binding specificity is determined by electrostatic interactions and polar part of solvation free energy, as confirmed in the literature [71]. Also, peptide 14 typically had contact with the exposed side chain of the PLP neuroantigen (Phe 281) and partly with C-terminal residues (Lys 283) (Fig. 6).

Conclusion

Pathophysiology of MS is extremely complicated including several immune system branches, and it has yet to be fully understood. Current MS therapies are often non-specific, resulting in inhibition of the general immune response to combat against pathogenic infections. As a result, more antigen-specific therapies are needed to avoid this general suppression. In this study, a therapeutic peptide design strategy was proposed to interrupt the TCR-pHLA interaction. Results of the post-MD analyses revealed that the designed peptide 14 was highly capable of binding to the TCR-binding sites of p-HLA complex in-silico, in a selective manner. This peptide could induce a local conformational change in p-HLA complex structure that led to an increase in total hydrogen bond occupancy and favorable protein-peptide interaction. It is hoped that findings of this research introduce a new strategy for treatment of MS disease.

Abbreviations

MS; multiple sclerosis, p-HLA; HLA-A*03-PLP_(45–53) complex, PPI; protein-protein interaction; MD; molecular dynamics, MHC; major histocompatibility complex, SASA; solvent accessible surface area, RMSD; root mean square deviation, Rg; radius of gyration

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

There is no funding for this project.

Authors' contributions

All authors contribute and read and approved the final manuscript.

Acknowledgment

We would like to appreciate the supports provided by the University of Shahrekord, Iran for the implementation of this project.

References

1. Ghozeud E, et al. Potential biomarker and therapeutic LncRNAs in multiple sclerosis through targeting memory B cells. *Neuromol Med.* 2020;22(1):111–20.
2. Kearney H, et al. Cervical cord lesion load is associated with disability independently from atrophy in MS. *Neurology.* 2015;84(4):367–73.
3. Frischer JM, et al. The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain.* 2009;132(5):1175–89.
4. Dendrou CA, Fugger L, Friese MA. Immunopathology of multiple sclerosis. *Nat Rev Immunol.* 2015;15(9):545–58.
5. Olsson T, Barcellos LF, Alfredsson L. Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nature Reviews Neurology.* 2017;13(1):25.
6. Goodin DS, et al. Highly conserved extended haplotypes of the major histocompatibility complex and their relationship to multiple sclerosis susceptibility. *PloS one.* 2018;13(2):e0190043.
7. Werneck LC, et al. Multiple sclerosis: disease modifying therapy and the human leukocyte antigen. *Arq Neuropsiquiatr.* 2018;76(10):697–704.
8. Muñoz-Culla M, Irizar H, Otaegui D. The genetics of multiple sclerosis: review of current and emerging candidates. The application of clinical genetics. 2013;6:63.
9. Clement M, et al. Targeted suppression of autoreactive CD8 + T-cell activation using blocking anti-CD8 antibodies. *Scientific reports.* 2016;6(1):1–14.
10. Rossjohn J, et al. T cell antigen receptor recognition of antigen-presenting molecules. *Annu Rev Immunol.* 2015;33:169–200.
11. Wang Y, et al., How an alloreactive T-cell receptor achieves peptide and MHC specificity. *Proceedings of the National Academy of Sciences*, 2017. **114**(24): p. E4792-E4801.
12. Galehdari H, et al., Analysis of HLA-A* 03 in Multiple Sclerosis Patients in Khuzestan Province, Iran. *Gene, Cell and Tissue*, 2018. 5(3).
13. Burfoot RK, et al. SNP mapping and candidate gene sequencing in the class I region of the HLA complex: searching for multiple sclerosis susceptibility genes in Tasmanians. *Tissue antigens.* 2008;71(1):42–50.
14. Leon VJ, et al. Study of HLA alleles in multiple sclerosis in Castilla y Leon (Spain). *Hum Immunol.* 2005;8(66):40.
15. McMahon RM, et al. Structure of HLA-A* 0301 in complex with a peptide of proteolipid protein: insights into the role of HLA-A alleles in susceptibility to multiple sclerosis. *Acta Crystallogr Sect D: Biol Crystallogr.* 2011;67(5):447–54.
16. Badawi AH, Siahaan TJ. Immune modulating peptides for the treatment and suppression of multiple sclerosis. *Clinical immunology.* 2012;144(2):127–38.
17. Fletcher JM, et al. De novo coiled-coil peptides as scaffolds for disrupting protein–protein interactions. *Chemical science.* 2018;9(39):7656–65.
18. Chidipi B, et al. Bioengineered Peptibodies as Novel Therapeutic Potassium Channel Blockers. *Circulation.* 2018;138(Suppl_1):A16680–0.
19. Arivajiagane A, et al. In silico structure-based design of enhanced peptide inhibitors targeting RNA polymerase PAN-PB1C interaction. *Comput Biol Chem.* 2019;78:273–81.
20. Fletcher S, Hamilton AD. Targeting protein–protein interactions by rational design: mimicry of protein surfaces. *Journal of the Royal Society Interface.* 2006;3(7):215–33.
21. Lu M-C, et al. A systematic molecular dynamics approach to the study of peptide Keap1–Nrf2 protein–protein interaction inhibitors and its application to p62 peptides. *Mol Biosyst.* 2016;12(4):1378–87.
22. Gowthaman R, Pierce BG. TCR3d: The T cell receptor structural repertoire database. *Bioinformatics.* 2019;35(24):5323–5.

23. Bagaev DV, et al. VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res.* 2020;48(D1):D1057–62.
24. Leem J, et al. STCRDab: the structural T-cell receptor database. *Nucleic acids research.* 2018;46(D1):D406–12.
25. Vita R, et al., The immune epitope database (IEDB) 3.0. *Nucleic acids research,* 2015. 43(D1): p. D405-D412.
26. Kortemme T, Kim DE, Baker D. Computational alanine scanning of protein-protein interfaces. *Science's STKE.* 2004;2004(219):pl2–2.
27. Zhu X, Mitchell JC. KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. *Proteins: Struct Funct Bioinf.* 2011;79(9):2671–83.
28. Cukuroglu E, Gursoy A, Keskin O. HotRegion: a database of predicted hot spot clusters. *Nucleic acids research.* 2012;40(D1):D829–33.
29. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology.* 2007;372(3):774–97.
30. Thévenet P, et al. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic acids research.* 2012;40(W1):W288–93.
31. López-Martínez R, et al. Inhibition of influenza A virus infection in vitro by peptides designed in silico. *PloS one.* 2013;8(10):e76876.
32. Dehouck Y, et al. BeAtMuSiC: prediction of changes in protein–protein binding affinity on mutations. *Nucleic acids research.* 2013;41(W1):W333–9.
33. Gao F, et al. Cooperation of B cell lineages in induction of HIV-1-broadly neutralizing antibodies. *Cell.* 2014;158(3):481–91.
34. Comeau SR, et al. ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic acids research.* 2004;32(suppl_2):W96–9.
35. Garg VK, et al. MFPPi–multi FASTA ProtParam interface. *Bioinformatics.* 2016;12(2):74.
36. De Vries SJ, Van Dijk M, Bonvin AM. The HADDOCK web server for data-driven biomolecular docking. *Nature protocols.* 2010;5(5):883.
37. Lear S, Cobb SL. Pep-Calculator.com: a set of web utilities for the calculation of peptide and peptoid properties and automatic mass spectral peak assignment. *J Comput Aided Mol Des.* 2016;30(3):271–7.
38. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology.* 1982;157(1):105–32.
39. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform.* 2007;8(1):1–7.
40. Case DA, et al. The Amber biomolecular simulation programs. *J Comput Chem.* 2005;26(16):1668–88.
41. Salomon-Ferrer R, et al. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J Chem Theory Comput.* 2013;9(9):3878–88.
42. Maier JA, et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* 2015;11(8):3696–713.
43. Jorgensen WL, et al. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79(2):926–35.
44. Liu J, Li D, Liu X. A simple and accurate algorithm for path integral molecular dynamics with the Langevin thermostat. *J Chem Phys.* 2016;145(2):024103.
45. Berendsen HJ, et al. Molecular dynamics with coupling to an external bath. *J Chem Phys.* 1984;81(8):3684–90.
46. 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.
47. Ryckaert J-P, Ciccotti G, Berendsen HJ. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys.* 1977;23(3):327–41.
48. Roe DR, Cheatham TE III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput.* 2013;9(7):3084–95.
49. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996;14(1):33–8.
50. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery.* 2015;10(5):449–61.
51. Gohlke H, Case DA. Converging free energy estimates: MM-PB (GB) SA studies on the protein–protein complex Ras–Raf. *J Comput Chem.* 2004;25(2):238–50.
52. Fu T, et al. Understanding the molecular mechanism of binding modes of Aurora A inhibitors by long time scale GPU dynamics. *Journal of Theoretical Computational Chemistry.* 2013;12(08):1341003.

53. Liu M, et al. Investigating the impact of Asp181 point mutations on interactions between PTP1B and phosphotyrosine substrate. *Scientific reports*. 2014;4(1):1–8.
54. Wang Q, et al. Molecular mechanism of the inhibition and remodeling of human islet amyloid polypeptide (hIAPP1–37) oligomer by resveratrol from molecular dynamics simulation. *J Phys Chem B*. 2015;119(1):15–24.
55. Hassan MI, Ahmad F. Structural diversity of class I MHC-like molecules and its implications in binding specificities. *Advances in protein chemistry structural biology*. 2011;83:223–70.
56. Basu S, Sen S. Do homologous thermophilic–mesophilic proteins exhibit similar structures and dynamics at optimal growth temperatures? a molecular dynamics simulation study. *J Chem Inf Model*. 2013;53(2):423–34.
57. Madura F, et al. TCR-induced alteration of primary MHC peptide anchor residue. *Eur J Immunol*. 2019;49(7):1052–66.
58. Zacharias M, Springer S. Conformational flexibility of the MHC class I $\alpha 1$ - $\alpha 2$ domain in peptide bound and free states: a molecular dynamics simulation study. *Biophysical journal*. 2004;87(4):2203–14.
59. 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.
60. Hou T, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model*. 2011;51(1):69–82.
61. Garcia KC, et al. The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nature immunology*. 2009;10(2):143–7.
62. Metz A, et al. Hot spots and transient pockets: predicting the determinants of small-molecule binding to a protein–protein interface. *J Chem Inf Model*. 2012;52(1):120–33.
63. Miller III. B.R., et al., MMPBSA. py: an efficient program for end-state free energy calculations. *J Chem Theory Comput*. 2012;8(9):3314–21.
64. Krogsgaard M, Davis MM. How T cells' see'antigen. *Nature immunology*. 2005;6(3):239–45.
65. Anikeeva N, et al. Distinct molecular mechanisms account for the specificity of two different T-cell receptors. *Biochemistry*. 2003;42(16):4709–16.
66. Roomp K, Domingues FS. Predicting interactions between T cell receptors and MHC-peptide complexes. *Molecular immunology*. 2011;48(4):553–62.
67. Murray JS. An old Twist in HLA-A: CDR3 α Hook up at an R65-joint. *Frontiers in immunology*. 2015;6:268.
68. Lichtenfeld M, et al. T cell receptor cross-recognition of an HIV-1 CD8 + T cell epitope presented by closely related alleles from the HLA-A3 superfamily. *Int Immunol*. 2006;18(7):1179–88.
69. Baker BM, et al. Identification of a crucial energetic footprint on the $\alpha 1$ helix of human histocompatibility leukocyte antigen (Hla)-A2 that provides functional interactions for recognition by tax peptide/Hla-A2–specific T cell receptors. *The Journal of experimental medicine*. 2001;193(5):551–62.
70. Culshaw A, et al. Germline bias dictates cross-serotype reactivity in a common dengue-virus-specific CD8 + T cell response. *Nature immunology*. 2017;18(11):1228.
71. Tripathi SK, Salunke DM. Exploring the different states of wild-type T-cell receptor and mutant conformational changes towards understanding the antigen recognition. *Journal of Biomolecular Structure Dynamics*. 2021;39(1):188–201.

Tables

Table 1. Docking binding score and hydrophilicity properties of fourteen selected peptides.

| Peptide number | Peptide sequence | HADDOCK score | ClusPro score | GRAVY |
|----------------|---------------------|------------------|------------------|-------|
| 1 | WREWWKDWFKQFRQFYRWF | -112.6±2.2 | -602.4 | -1.57 |
| 2 | WREWWKDWAKQFRQFYRWF | -119.7±4.5 | -662.2 | -1.62 |
| 3 | WREWWQDWFKQFRQFYRAF | -126.1±3.4 | -739.7 | -1.41 |
| 4 | WREWWKDWFKQFRQFYRKF | -131.9±6.8 | -783.3 | -1.73 |
| 5 | WREWWKDWFKQFRQFYRQF | -133.6±5.1 | -961.4 | -1.7 |
| 6 | WREWWKDWFKKFRQFYRAF | -137.7±2.9 | -979.5 | -1.45 |
| 7 | WREWWKEWFKQFRQFYRAF | -141.8±2.2 | -1073.5 | -1.43 |
| 8 | WDEWWKEWFKQFRQFYRAF | -140.8±9.7 | -1092.4 | -1.37 |
| 9 | WWEWWKEWFKQFRQFYRAF | -144.1±3.7 | -1163.9 | -1.24 |
| 10 | WREWWQDWFKQFRQFYRTF | -145.8±8.3 | -1188.7 | -1.54 |
| 11 | WREWWKDWFKQFRQFYRAF | -146.7±9.4 | -1223.1 | -1.43 |
| 12 | WDEWWKEWFKQFRQFYRDF | -150.7±5.8 | -1247.4 | -1.63 |
| 13 | WREWWQDWFKQFRQFYRAF | -148.9±7.6 | -1250.6 | -1.41 |
| 14 | WRYWWKDWAKQFRQFYRWF | -158.7±0.9 | -1265.5 | -1.51 |

Table 2. The average values of RMSD and overall fluctuations (F_{overall}) for the three described regions, peptides, and p-HLA complexes

| | RMSD p-HLA complex | RMSD plp fragment | RMSD α_1 - helix | RMSD α_2 - helix | RMSD designed peptide | F _{overall} p-HLA complex | F _{overall} plp fragment | F _{overall} α_1 - helix | F _{overall} α_2 - helix | F _{overall} binding grooves | F _{overall} designed peptide |
|----------|--------------------------|-------------------------|-------------------------------|-------------------------------|-----------------------------|--|---|---|---|--|---|
| System1 | 2± 0.3 | 0.4± 0.1 | 0.7± 0.1 | 0.6± 0.1 | 0.8± 0.2 | 8.8 | 6.3 | 8.3 | 7.9 | 8.4 | 9.4 |
| System2 | 2.2± 0.5 | 0.8± 0.1 | 1.2± 0.2 | 0.9± 0.2 | 0.7± 0.2 | 10.5 | 7.6 | 8.6 | 9 | 8.8 | 10.5 |
| System3 | 1.4± 0.3 | 0.6± 0.2 | 0.6± 0.1 | 0.7± 0.1 | 1.3± 0.4 | 9.6 | 7.1 | 11.2 | 7.8 | 9.4 | 11.5 |
| System4 | 2.5± 0.7 | 0.6± 0.1 | 0.6± 0.1 | 1.3± 0.4 | 1.2± 0.5 | 8.5 | 6.4 | 8.9 | 8.9 | 8.7 | 8.9 |
| System5 | 1.6± 0.3 | 0.4± 0.1 | 0.7± 0.1 | 0.8± 0.1 | 0.7± 0.2 | 8.5 | 7.4 | 9.6 | 10.4 | 9.3 | 11.3 |
| System6 | 1.7± 0.3 | 0.4± 0.1 | 0.9± 0.1 | 0.9± 0.2 | 0.7± 0.1 | 10.7 | 7.6 | 7.6 | 10.2 | 9.6 | 8.5 |
| System7 | 1.4± 0.2 | 0.6± 0.1 | 0.5± 0.1 | 0.8± 0.2 | 1.9± 0.9 | 6.9 | 4.7 | 6.1 | 7.6 | 6.9 | 7.3 |
| System8 | 1.5± 0.3 | 0.5± 0.1 | 0.7± 0.2 | 0.8± 0.2 | 1± 0.2 | 9.5 | 8 | 10.9 | 9.1 | 10.3 | 11.4 |
| System9 | 2.5± 0.5 | 0.4± 0.1 | 0.4± 0.1 | 0.6± 0.1 | 0.6± 0.1 | 9.5 | 5 | 6.3 | 9.7 | 8.8 | 10.7 |
| System10 | 2± 0.5 | 0.4± 0.1 | 0.6± 0.1 | 1± 0.2 | 0.6± 0.1 | 8.3 | 7.3 | 9.1 | 8.4 | 8.3 | 9.7 |
| System11 | 1.7± 0.4 | 0.5± 0.1 | 0.8± 0.2 | 0.7± 0.1 | 0.7± 0.2 | 6.4 | 5.1 | 6.2 | 6.4 | 6.4 | 7.3 |
| System12 | 1.9± 0.3 | 0.5± 0.2 | 0.9± 0.2 | 0.7± 0.1 | 0.9± 0.1 | 10 | 8.8 | 9.8 | 10.3 | 9.8 | 11.6 |
| System13 | 2.6± 0.7 | 0.5± 0.1 | 0.9± 0.1 | 1.2± 0.1 | 1± 0.3 | 10.4 | 6.3 | 10 | 11.6 | 11.3 | 9.2 |
| System14 | 3.7± 1.2 | 0.4± 0.1 | 0.9± 0.2 | 0.8± 0.1 | 1.1± 0.3 | 10.9 | 8.1 | 11.5 | 10.9 | 10.8 | 10.6 |

Table 3. The average values of radius of gyration (RG) for the three described regions, peptides, and p-HLA complexes

| | Rg p-HLA complex | Rg PLP neuroantigen fragment | Rg α_1 -helix | Rg α_2 -helix | Rg designed peptide |
|----------|------------------|------------------------------|----------------------|----------------------|---------------------|
| System1 | 22.6±0.2 | 8.9±0.1 | 9.7±0.1 | 10±0.1 | 9.6±0.1 |
| System2 | 21.3±0.2 | 9.5±0.1 | 9.4±0.1 | 9.7±0.1 | 9±0.1 |
| System3 | 21.6±0.1 | 8.6±0.1 | 9.7±0.1 | 10.3±0.1 | 9.7±0.1 |
| System4 | 22.4±0.2 | 8.8±0.1 | 10±0.1 | 10.3±0.1 | 9.5±0.1 |
| System5 | 21.8±0.2 | 8.7±0.1 | 9.3±0.1 | 10.2±0.1 | 10±0.1 |
| System6 | 20.7±0.1 | 8.9±0.1 | 9.6±0.1 | 10.2±0.2 | 9.5±0.1 |
| System7 | 21.5±0.2 | 8.6±0.1 | 9.4±0.1 | 9.9±0.1 | 10.2±0.5 |
| System8 | 22.4±0.1 | 9.7±0.1 | 9.5±0.1 | 10.2±0.1 | 8.6±0.1 |
| System9 | 23±0.1 | 9.1±0.1 | 9.4±0.1 | 10.1±0.1 | 9.3±0.1 |
| System10 | 21.6±0.2 | 8.7±0.1 | 9.5±0.1 | 10.1±0.1 | 9.4±0.1 |
| System11 | 20.6±0.2 | 8.9±0.1 | 9.6±0.1 | 10.3±0.1 | 9.9±0.1 |
| System12 | 22.1±0.2 | 9.4±0.1 | 9.5±0.1 | 10.2±0.1 | 9.1±0.2 |
| System13 | 22.7±0.2 | 9±0.1 | 10.3±0.1 | 10.5±0.1 | 9.6±0.2 |
| System14 | 22.3±0.3 | 10±0.1 | 9.6±0.1 | 10.1±0.1 | 9.6±0.2 |

Table 4. The average values of Solvent Accessible Surface Area (SASA) (\AA^2) for the three described regions and designed peptides

| | SASA of plp neuroantigen fragment | SASA α_1 -helix | SASA α_2 -helix | SASA designed peptide |
|----------|-----------------------------------|------------------------|------------------------|-----------------------|
| System1 | 999.6±10.1 | 1659.2±28 | 1719±17.8 | 1904.5±47 |
| System2 | 1055.2±10.4 | 1613.3±36.8 | 1678.8±22.2 | 1799.3±44.4 |
| System3 | 1064.2±9.4 | 1650.6±35.8 | 1724.7±25.5 | 1877.4±48.1 |
| System4 | 1029.3±11.1 | 1653.2±31.6 | 1760.4±28.5 | 1932.2±58.8 |
| System5 | 989.9±11 | 1590.5±25.2 | 1789.5±24.8 | 1962.7±43.1 |
| System6 | 1011±9.8 | 1622.4±23.3 | 1756.8±43.9 | 1883.7±44.6 |
| System7 | 1047.5±10 | 1620.2±38.9 | 1722.4±19.1 | 2034.4±57.7 |
| System8 | 1033.6±15.5 | 1614.1±25.6 | 1746.2±28.4 | 1790.6±57 |
| System9 | 1041±9.2 | 1552.4±24.2 | 1705.6±31.7 | 1873.2±33.7 |
| System10 | 1027±10 | 1618.6±25.8 | 1719.9±21.8 | 1876.6±38.4 |
| System11 | 998.3±11.4 | 1596.2±30.2 | 1736.5±19.5 | 2011.8±40.8 |
| System12 | 1054.2±7.8 | 1641±42.5 | 1777.2±22.1 | 1882.4±52.6 |
| System13 | 1047.1±10.2 | 1758.2±30.9 | 1794.4±31.8 | 1931±49.2 |
| System14 | 1059.9±9.4 | 1602±22.1 | 1742.6±27 | 2060.2±50.9 |

Table 5. The average values of H-bonds number and distance between the center of mass (COM) of the p-HLA complex and peptides, Total H-bonds occupancy, MMGBSA Binding free energy (kcal/mole) during the last 10 ns of MD simulation

| | Binding free energy (MMGBSA) | H-bonds number | Total H-bonds occupancy (%) | Distance (Å) |
|-----------|------------------------------|----------------|-----------------------------|--------------|
| Peptide1 | -12±6.8 | 3.6±1.8 | 131 | 24.2±0.6 |
| Peptide2 | -17.4±4.2 | 2.1±1.1 | 85 | 30.1±0.4 |
| Peptide3 | -39.3±9.7 | 7.1±2.3 | 459 | 24.7±0.7 |
| Peptide4 | -38.9±5.5 | 1.4±1.1 | 0 | 24.9±0.5 |
| Peptide5 | -39.1±6.5 | 3.6±1.3 | 198 | 25.6±0.4 |
| Peptide6 | -45.1±6.5 | 2.5±1.5 | 72 | 25±0.4 |
| Peptide7 | -52.6±6.9 | 6.3±2.1 | 251 | 23±0.5 |
| Peptide8 | -51.6±8.1 | 6.1±2.2 | 225 | 26.5±0.3 |
| Peptide9 | -51.6±5.5 | 3.6±1.5 | 119 | 22.19±0.6 |
| Peptide10 | -55±6.2 | 5.2±1.7 | 346 | 23.3±0.3 |
| Peptide11 | -54.3±5.5 | 6.4±1.7 | 447 | 24.5±0.5 |
| Peptide12 | -65.4±6.8 | 6.6±2.1 | 292 | 23.7±0.3 |
| Peptide13 | -59.2±6.9 | 5.1±1.9 | 152 | 25±0.4 |
| Peptide14 | -75.1±6.1 | 7.4±1.8 | 547 | 24.5±0.3 |

Table 6. Percentage of occupancy and average distance (Å) between peptide 14 and p-HLA complex active site residues during the last 10 ns of MD simulation time

| Acceptor | Donor | Occupancy % | Avg. Distance (Å) | Avg. Angle |
|-------------|-------------|-------------|-------------------|------------|
| ARG_2p@O | ARG_65@NH1 | 73.40 | 2.8 | 151.8 |
| ARG_2p@O | ARG_65@NH2 | 57.14 | 2.8 | 149.2 |
| GLU_152@OE2 | ARG_17p@NH1 | 63.62 | 2.7 | 157.6 |
| GLU_152@OE1 | ARG_17p@NH2 | 61.59 | 2.7 | 160 |
| GLU_152@OE1 | ARG_17p@NH1 | 38.96 | 2.8 | 155.5 |
| GLU_152@OE2 | ARG_17p@NH2 | 30.61 | 2.8 | 159.1 |
| GLU_58@OE1 | ARG_2p@NH2 | 35.72 | 2.8 | 157.6 |
| GLU_58@OE2 | ARG_2p@NH2 | 34.18 | 2.8 | 157.5 |
| GLU_58@OE1 | ARG_2p@NE | 27.16 | 2.8 | 156.9 |
| GLU_58@OE2 | ARG_2p@NE | 26.30 | 2.8 | 157.1 |
| GLN_62@OE1 | TRP_4p@N | 61.1 | 2.8 | 153.9 |
| GLN_62@OE1 | TRP_5p@N | 22.81 | 2.8 | 161.7 |
| GLN_155@O | GLN_11p@NE2 | 20.18 | 2.8 | 157.4 |

Figures

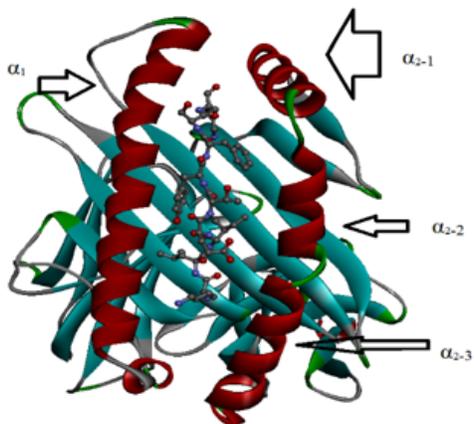
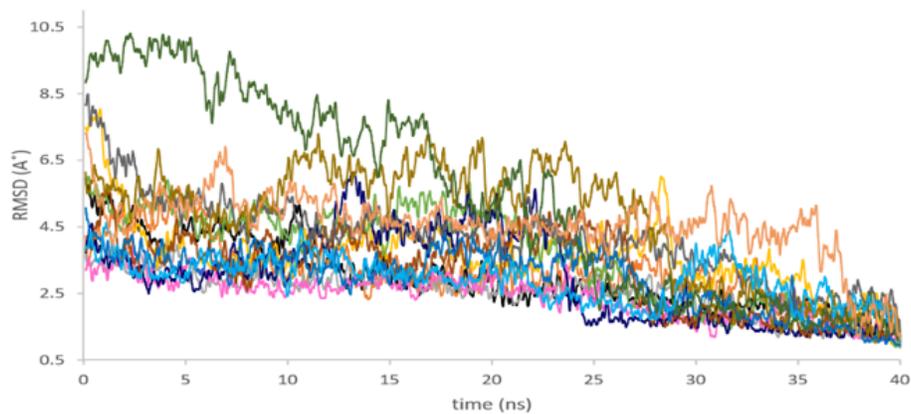


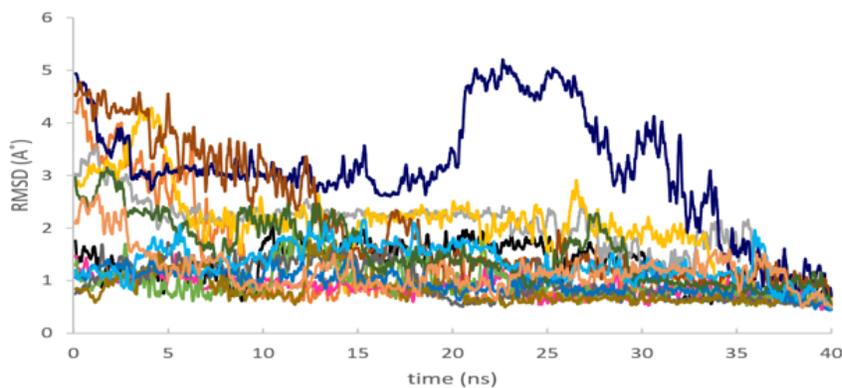
Figure 1

Cartoon representation of the MHC class I molecule HLA-A*0301 in complex with an antigenic peptide (PLP) (McMahon et al., 2011; peptide sequence: KLIETYFSK; 2XPG). The α -helices of the MHC protein are in red (β -strands in light blue). The view is into the α 1- α 2 domain peptide binding cleft (peptide in atomic resolution). The β 2m protein and the α 3 domain are located below the α 1- α 2 domain

2A



2B

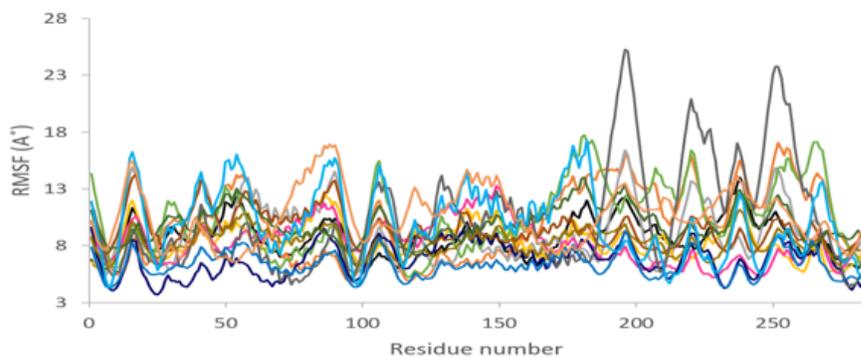


— system 1 — system 2 — system 3 — system 4 — system 5 — system 6 — system 7
 — system 8 — system 9 — system 10 — system 11 — system 12 — system 13 — system 14

Figure 2

RMSD of the P-HLA complex (2a) and designed peptide (2b).

3a



3b

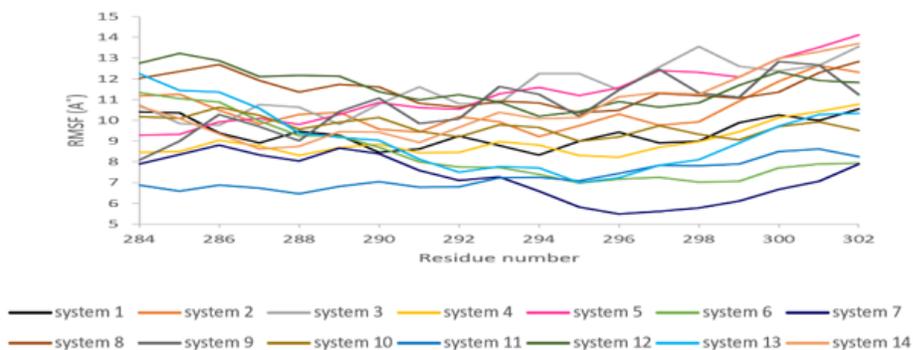


Figure 3

Per-residue backbone RMSF atoms for p-HLA complex (3a) and designed peptide (3b) during the last 10 ns of MD simulation

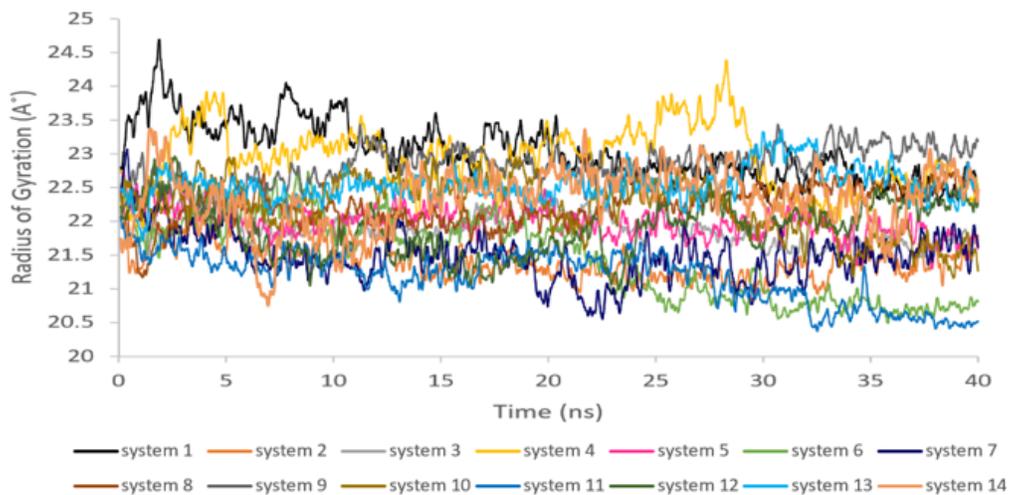


Figure 4

Fluctuation of Rg during simulation time for p-HLA complex

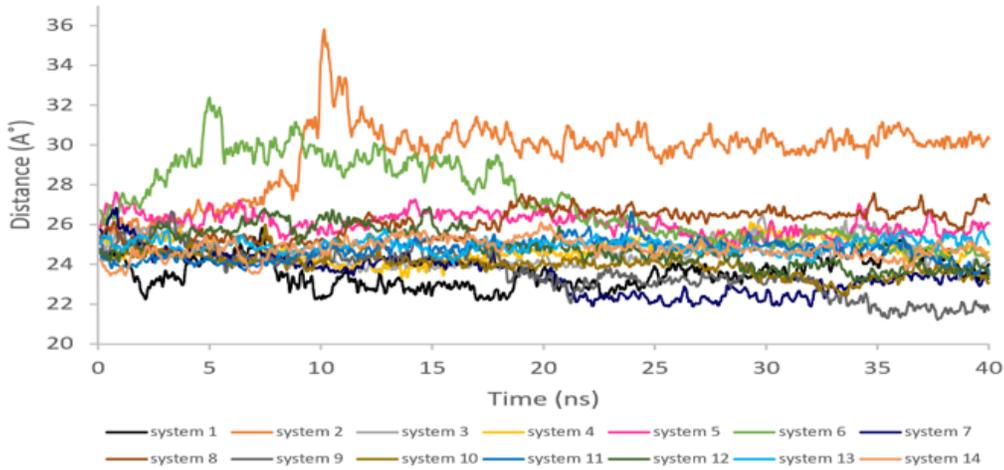


Figure 5

Distance between the center of mass (COM) of the p-HLA complex and peptides during the simulation time

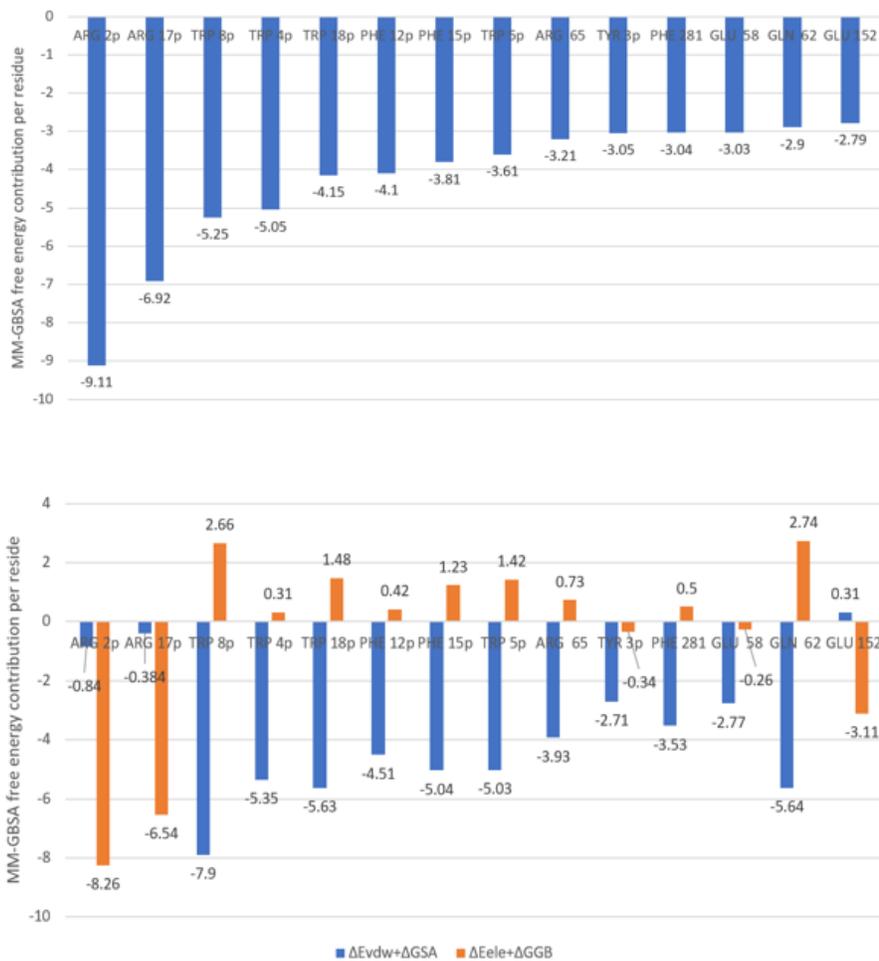


Figure 6

Per-residue energy (kcal/mol) decomposition of binding free energy analysis for peptide 14 to p-HLA complex (character p after residue number refer to peptide number)

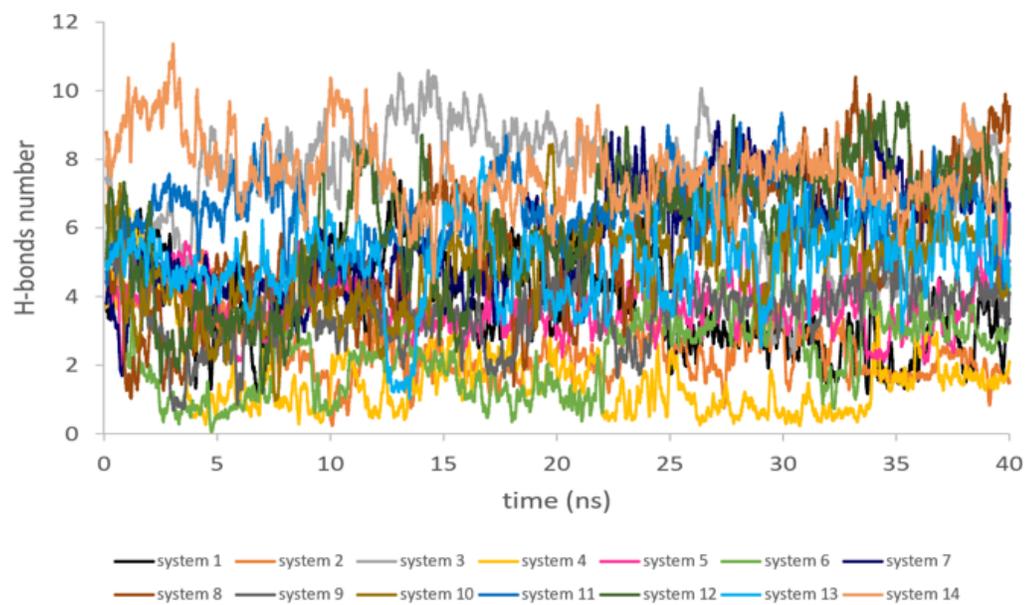


Figure 7

The number of hydrogen bonds during MD simulation between p-HLA complex and peptides.

