Molecular Simulation
Vol. 37, No. 10, September 2011, 837–845

Taylor & Francis
Taylor & Francis Group

Molecular dynamics study of the human insulin B peptide SHLVEALYLVCGERGG complexed with HLA-DQ8 reveals important hydrogen bond interactions

Athanassios Stavrakoudis*

Department of Economics, University of Ioannina, Ioannina, Greece

(*Received 5 August 2010; final version received 23 February 2011*)

The basis of proper recognition of pathogens and tumours is provided by adaptive immunity. This immunological reaction of the recognition function of T-cell receptors on T lymphocytes detects antigenic peptides bound to major histocompatibility complex (MHC) molecules. Structural insight into this process has few grown considerably in the last years. In some of the cases, antigens are self-protein fragments causing autoimmunity diseases. Type 1 diabetes is such a disease connected with the human leukocyte antigen-DQ8 molecule, a class II MHC glycoprotein. Its crystal structure, complexed with LVEALYLVCGERGG peptide (insulin B peptide), has been solved, and important information about the significance of P1, P4 and P9 binding pockets has been discovered. The complex structure also revealed an unusual large number of intermolecular hydrogen bonds between insulin B peptide and MHC molecule. To further investigate the dynamics of peptide/MHC interactions, we perform molecular dynamic simulations in explicit water. Analysis of the results provided useful information of the binding of the peptide antigen to MHC molecule, which is supported by numerous hydrogen bonds besides the electrostatic (P1 and P9 pockets) or hydrophobic interactions (P4). Results also allowed some implications to be drawn for the role of residues located outside of the binding groove.

Keywords: HLA-DQ8; MHC/peptide interactions; molecular dynamics; principal component analysis; type 1 diabetes

1. Introduction

Major histocompatibility complex (MHC) class II molecules are proteogenic receptors that lie on the cell surface and display a range of peptides, recognised by T-cell receptors of CD4(+) T-helper cells [1]. These receptors have essential functionality in defence processes to invade micro-organisms [2]. However, in many cases, they recognise self-peptides which can lead to auto-immune response causing autoimmune diseases. One such disease is type 1 diabetes [3].

The structure of human leukocyte antigen (HLA)-DQ8 complexed with insulin B peptide has been identified by crystallographic methods [4]. HLA-DQ8 belongs to class II MHC glycoproteins. Peptide SHLVEALYLVCGERGG found in this complex is associated with diabetes autoimmune disease [5,6]. Along with HLA-DQ2 in humans and I-Ag7 in non-obese diabetic mice, HLA-DQ8 is connected with risk factors for increased susceptibility to type 1 diabetes. X-ray structure of peptide/HLA-DQ8 complex revealed important structural characteristics of the peptide loading to MHC molecule. This structure has also been used as a template for modelling HLA-DQ7 [7] and HLA-DQ2 [8] molecules in complexed form with peptide antigens. Until now, most of the focus on specific peptide/MHC interactions has been around three binding pockets of HLA-DQ8 molecule, namely P1, P4 and P9 that locate the Glu3, Tyr6 and Glu11 residues, respectively. Other, also important, interactions between the peptide and MHC molecule have been relatively underestimated. For example, backbone atoms of Cys9, which occupied the P7 binding pocket, were found to form two hydrogen bonds ($Asn\alpha69{:}N^{\delta 2}$–Cys9:O and Cys9:N–$Tyr\beta30{:}O^{\eta}$) with both α and β MHC chains. Moreover, side chain of Cys9 was also found to form a (somewhat weak) hydrogen bond with side chain of Tyrβ47. The importance of these three hydrogen bonds formed by one single residue is clearly evident.

Molecular dynamics simulations [9] is a well-established tool for studying the structure and dynamics of immunological complexes [10–12], and for facilitating drug design [13]. It has been shown that such simulations enhance our understanding of biological interactions of the immunological synapse processes at atomic level [14], can be applied in studying protein/peptide complexes of immunological interest [15–17] and provide a useful framework for studying peptide/MHC interactions [18–20] or model peptide-free MHC structure [21,22]. In this work, molecular dynamics simulations were utilised to investigate the dynamics properties of a peptide/HLA-DQ8 complex. Important interactions between the peptide and MHC molecule are highlighted and implications on the T cell receptor (TCR) binding of the complex are drawn. As in other cases, this computational approach is complementary

*Email: astavrak@cc.uoi.gr

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927022.2011.566607
http://www.informaworld.com

to the experimental results [23] and can help in better understanding of the biological interactions at atomic level.

2. Computational methods

2.1 System set-up and simulation

Initial coordinates of HLA-DQ8/insulin B complex were downloaded from the Protein Data Bank [24], PDB access code: 1jk8 [4]. The whole MHC molecule was included in the simulation set-up, as previous studies have shown that omission of α3 or β2m domains can cause serious problems [25]. Hydrogen atoms were added according to CHARMM22–CMAP topology parameters [26,27] using the visual molecular dynamics (VMD) program [28]. Inclusion of CMAP correction term has shown to improve the structure quality and to help avoid undesired artificial helical transitions in the MD trajectory [29,30]. The MHC/peptide complex was solvated in a box with dimensions $98.5 \times 78.4 \times 115.2$ Å, containing TIP3P [31] water molecules. Crystallographic water molecules were retained and treated with the TIP3P model. The size of the box allowed a minimum distance of 18 Å between any atom of the protein complex and the edge of the box. The big sized solvent box has been shown to be essential to avoid the periodicity-induced artefacts, when the particle mesh Ewald (PME) method is applied for electrostatics calculations [32]. Thirty-one Na^+ and 17 Cl^- ions were added to the system to neutralise it and to approximate an ionic strength of 100 mM, using the autoionize plugin of the VMD program. From this point, all subsequent MD runs were performed using the NAMD v2.6 [33] package and periodic boundary conditions were employed. Force field parameters were assigned from the CHARMM22 parameter set with the CMAP correction term for all atoms of the system. Non-bonded van der Waals interactions were gradually turned off at a distance between 10 and 12 Å. The non-bonded pair list was updated every 10 steps. Long-range electrostatics were computed every two steps using the PME method [34,35], with a grid spacing of less than 1 Å. The whole system, consisted of 98,691 atoms, was energy minimised with 2500 steps of conjugate gradients, by restraining the coordinates of protein heavy atoms to the crystallographic positions with a force constant of 50 kcal mol^{-1} Å^2. After minimisation, the temperature of the system was gradually increased with Langevin dynamics [36], using the canonical ensemble (constant number of particles, volume and temperature) (NVT) ensemble, to 310 K, during a period of 6000 steps, by stepwise reassignment of velocities every 1000 steps. Equilibration at 310 K was continued until 100 ps. Equilibration at 310 K, using the NVT ensemble, was continued for another 100 ps, with the force restraining the heavy atom positions to be set at the value of 5 kcal mol^{-1} Å^2. Finally, the positional restraints were totally eliminated and the system was equilibrated for another 300 ps. The simulation was continued under constant pressure, with Langevin piston method [37], thus NPT ensemble, for 20 ns. Pressure was maintained at 1 atm and the temperature was kept at 310 K. Time step for integration of equation of motion during MD runs was set to 2 fs. The results presented here are from the last, isothermalisobaric ensemble, MD run. Snapshots were saved to disk at 1 ps interval for further analysis. Conformational analysis and hydrogen bond interactions were calculated using Carma [38] and Eucb [39] software packages. Secondary structure assignment was performed with STRIDE [40]. Structural figures were prepared with PyMOL (www.pymol.org).

Residues of MHC chains are denoted by lower case Greek letters α and β. Peptide residues do not have chain identifier.

2.2 Principal component analysis

Principal component analysis (PCA) is a standard method for analysing MD trajectories, in which the reduction in the dimensionality of a high-dimensional data–set is desired [41]. The basic idea was to compute the covariance matrix

$$\sigma_{ij} = \big\langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \big\rangle, \tag{1}$$

where $q_1, \ldots, q_{3M}$ are the mass-weighted Cartesian coordinates of an M-atom molecule, and to obtain the eigenvectors $\mathbf{b}^i$ and eigenvalues λ^i of the diagonalised covariance matrix.

The data from $\mathbf{q} = (q_1, \ldots, q_{3M})^{\mathrm{T}}$ can be used to represent the free energy surface:

$$\Delta G(r) = -k_{\mathrm{B}} T [\ln P(r) - \ln P_{\max}], \tag{2}$$

where P denotes the probability distribution of the system along r coordinates and $P_{\max}$ is the maximum probability value, which is subtracted to ensure $\Delta G = 0$ for the minimum energy.

Recently a dihedral-based PCA (dPCA) has been proposed [42–44], which better discriminates the energy landscape of a biomolecule than the traditional Cartesian PCA (cPCA). In the dPCA technique, the dihedral angles are transformed to

$$q_{2n-1} = \cos \phi_n, \quad q_{2n} = \sin \phi_n,$$

where $n = 1, \ldots, N$ is the number of the dihedral angles used in the analysis. In this study, both cPCA and dPCA have been performed using Carma [38].

2.3 Buried surface area

The using calculation of buried surface area (BSA) was performed with the NACCESS program, based on the

formula

$$\mathrm{BSA} = S_\mathrm{p} + S_\mathrm{m} - S_\mathrm{c},$$

where S_c is the surface accessible area of the complex, S_p is the surface accessible area of the peptide and S_m is the surface accessible area of the MHC molecule respectively. The calculations were performed to all frames of MD trajectories in order to get and characterise the times series of BSA.

3. Results and discussion

3.1 Backbone dynamics of the bound peptide

Figure 1 shows the root mean square fluctuation (RMSF) values of the protein C^α atoms. As it was expected, the bound peptide showed limited flexibility with the exception of the N- and C-terminal parts. Residues Leu1 and Val2 were located outside of the binding groove and showed RMSF values 0.288 and 0.153 nm, respectively. C-terminal residues Arg12–Gly13–Gly14 showed increasing RMSF values from 0.096 to 0.173 nm. The central part of the peptide, residues Glu3 to Glu11, that constitute the 9mer binding motif to the MHC groove showed limited flexibility as revealed from RMSF analysis. Values less than 0.1 nm were recorded for these residues although value 0.105 nm was found for Glu3. The higher flexibility of this residue seems that it was influenced by the increased mobility of N-terminal neighbours where the highest RMSF values were recorded. Relatively small RMSF values of C^α atoms were also found for the helical part of the MHC molecule (binding groove). Thus, almost all of chain A residues 50–80 or chain B residues 55–90 showed RMSF values of approximately 0.1 nm or less. This fact indicates the conformational stability of the MHC binding groove and its relative rigidity. Some parts of the MHC molecule, however, were found considerably flexible, for example residues 105–115 of chain B. These residues are far away from the binding groove and do not interplay with peptide/MHC interactions. Apparently, there is evidence that the inclusion of the whole MHC molecule, without neglecting the α3 or β2m chains, is necessary to accurately reproduce peptide/MHC interactions [25].

Time series of root mean square displacement (RMSD) values of backbone atoms relatively to the initial atom positions (X-ray) shown in Figure 1(C). Chain A of are HLA-DQ8 molecule showed limited RMSD fluctuation between 0.110 and 0.207 nm, whereas trajectory values averaged at 0.153 nm (0.016 nm). Chain B of MHC was found to experience relatively higher deviation from the initial structure. RMSD values fluctuated between 0.153 and 0.373 nm and averaged at 0.256 nm (0.054 nm). Despite being more flexible than chain A, chain B also

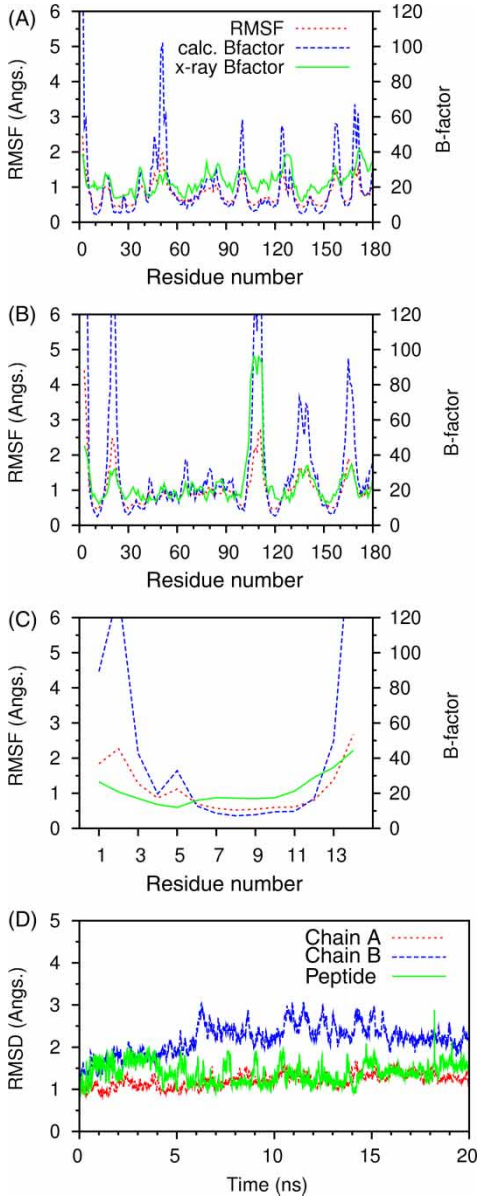

Figure 1. (A) RMSFs of the C^α atoms of MHC chain α, (B) RMSFs of C^α atoms of MHC chain β, (C) RMSFs of the C^β atoms of the peptide, (D) RMSD of backbone atoms of the three chains of the complex. The sub-graphs A–C also display the X-ray with found B-factor of C^α atoms and the calculated B-factor: $(8\pi^2/3)\mathrm{RMSF}^2$.

remained stable during molecular dynamics trajectory. Averages values of backbone RMSD atoms of 0.3 nm are very commonly observed. Peptide's backbone atoms showed small RMSD values with trajectory averaged at 0.136 nm (0.022 nm) which fluctuate between 0.077 and 0.270 nm. It is worthy to note that even smaller values were recorded for 3–12 fragments of the peptide sequence. Thus, the average value of the whole trajectory for these fragments was found only to be 0.077 nm (0.011 nm). The reduced RMSD values of a shorter fragment relative to the

whole peptide sequence are not surprising at all. However, the finding that the exclusion of only four residues (out of 14) resulted in drastic reduction of average RMSD values has to be underlined. This is also in line with the RMSF analysis of peptide's C^{α} atoms previously analysed, and advocates the observation that peptide's residues outside the binding groove experienced increased flexibility. Overall, from RMSD and RMSF analysis, it is evident that the complex remained stable under MD conditions with the peptide well 'engulfed' in the MHC's binding groove. A bundle of structures obtained from the MD trajectory demonstrating this is shown in Figure 2.

Time series of the peptide's backbone ϕ and ψ dihedral angles are shown in Figure 3. Backbone dihedral angles of all residues showed very limited fluctuation around the initial values. This is consistent with RMSD analysis and demonstrates the *immobilisation* of the peptide upon binding to the MHC groove. It is interesting to note that this observation applies well to the residues Val2 and Arg12, which are located outside of the main binding site. Only Gly13 s backbone dihedral angles showed considerable deviations from initial values, especially its ψ backbone dihedral angle fluctuated in the range ($-60°$ to 90°).

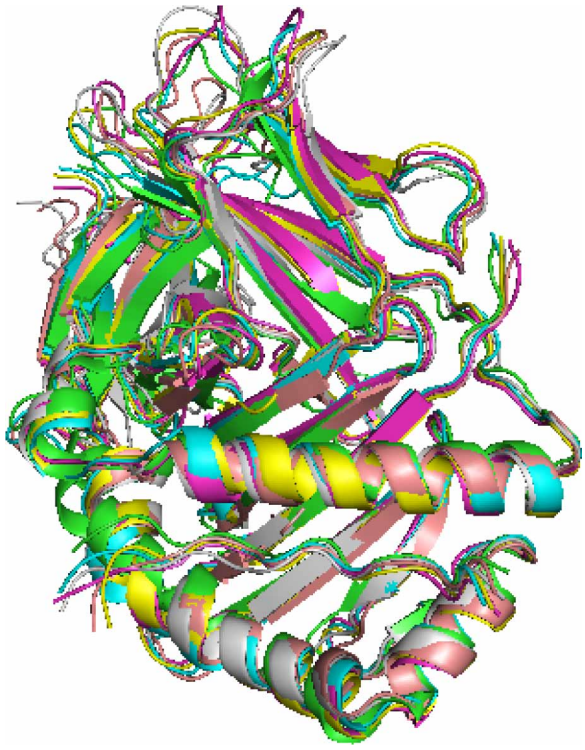

Figure 2. Ribbon representation of five representative structures obtained from MD trajectory (one every 4 ns) superimposed to the crystal structure of peptide/HLA-DQ8 (green colour online).

Conformational space coverage and principal components analysis

A critical question in all MD simulations of large biomolecule is about the conformational space sampling. PCA techniques offer a standard framework to analyse MD trajectories and to address such questions. As it has been recently proposed, dihedral PCA can be advantageous sometimes in comparison with classical cPCA [42]. Here, both techniques were applied to the peptide's backbone atoms, in order to have a more spherical view of the backbone dynamics of the peptide.

Figure 4 shows the comparison of dihedral and cPCA of the peptide's backbone atoms over the first three components, considering the whole peptide sequence. The same analysis was also applied to the Glu3–Glu11 sequence, thus considering only the nine residues within MHC's binding groove. The corresponding results are shown in Figure 5. As it can be seen from these graphs, the peptide visited limited conformational space, in line with the fact that there is no much space for manoeuvring in MHC's binding site. The peptide was found in one main conformational cluster, although it can be argued that two others are also present, although with definitely decreased occurrence. There is no significant difference between cPCA and dPCA results. Both cPCA and dPCA suggested the existence of one main conformational cluster, at approximately the same ratio. But the refinement of conformers distribution in dPCA is seen much better than that in cPCA. From this perspective, the methodology proposed in other published works [42–44] is considered as superior also in this case. Moreover, the dPCA method performed better in cluster discrimination than the cPCA method in the case of Glu3–Glu11 peptide fragments, thereby the residues extending from MHC's binding site were excluded from the PCA (Figure 5).

3.2 Hydrogen bond interactions

A plethora of hydrogen bond interactions between the peptide and the HLA-DQ8 molecule have been identified in the MD trajectory. Both backbone and side chain atoms of the peptide formed a considerable number of hydrogen bonds, found in the X-ray structure, which was used as a starting conformation of the simulation. Table 1 lists the hydrogen bonds between the peptide and HLA-DQ8 as well as the percentage of their conservation during MD run.

The P1 position of the HLA-DQ8 molecule accommodated the Glu3 residue of the peptide. The P1 site is very polar [45]. Glu3 forms the peptide and Argα52 from MHC molecule remained hydrogen bonded through their side chains for most of the simulation time.

The Ala4 residue, located at P2 position, was found to form a pair of hydrogen bond interactions with Asnβ52 side chain. It is quite unusual for a neutral residue of a

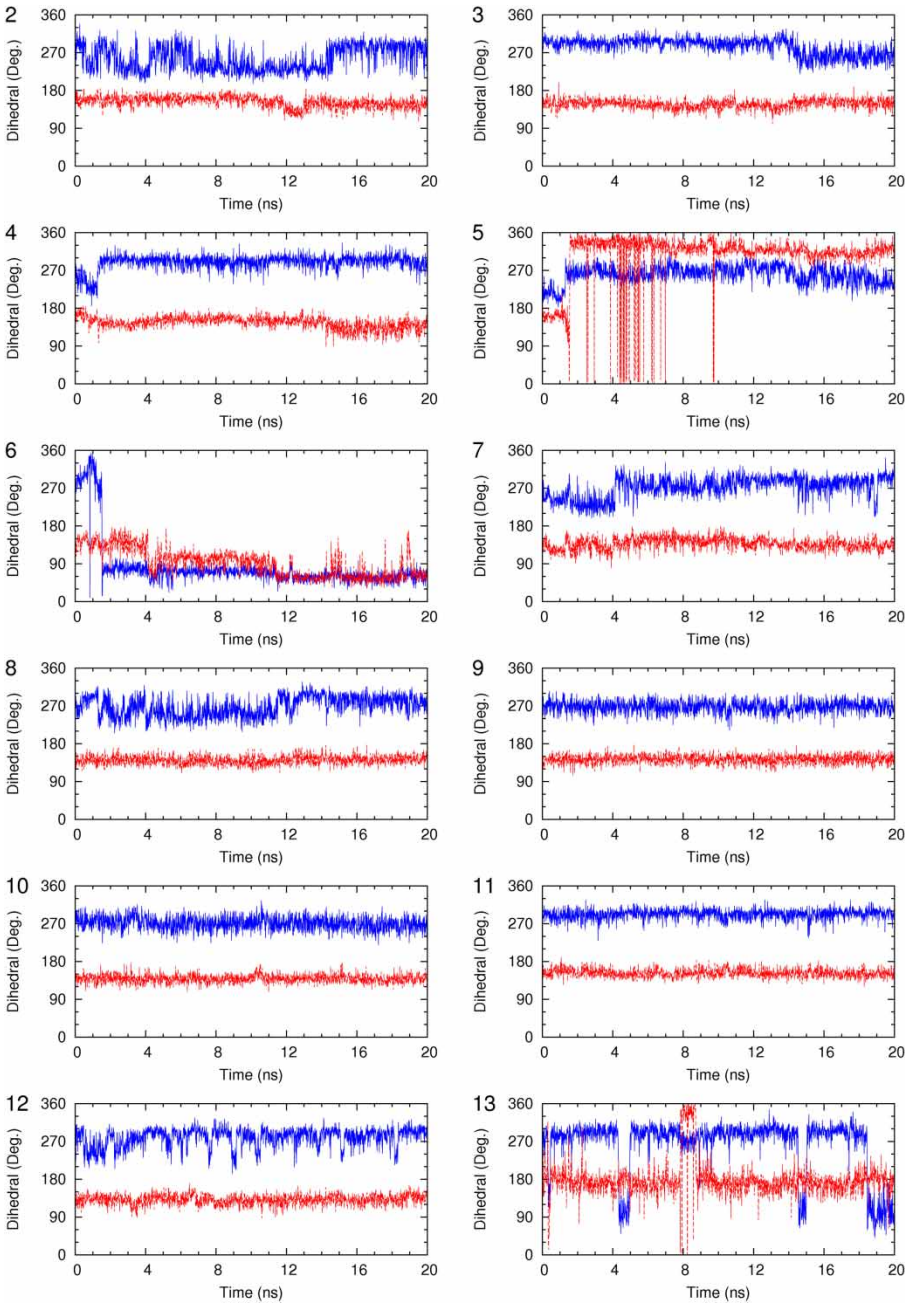

Figure 3. Time series of backbone dihedral angles of the peptide. Blue line corresponds to ϕ dihedral angle and red line corresponds to ψ dihedral angle (colour online).

peptide to form two hydrogen bonds with MHC molecule. However, in this case, backbone oxygen was found to form hydrogen bond with Asnβ82:$N^{\delta 2}$ and backbone amide group was found to form hydrogen bond interaction with Asnβ82:$O^{\delta 1}$. These two hydrogen bonds were conserved for 74 and 38% of the simulation time, respectively. The existence of at least one hydrogen bond for the majority of the trajectory frames, of two hydrogen bonds for a considerable amount of simulation time, indicates the importance of the Ala4 (P2 position) contribution to peptide's binding. It is also an alarm that alanine residue is not so neutral in peptide/protein interactions.

Backbone oxygen atom of Tyr6 made an important hydrogen bond interaction with Asnα62:$N^{\delta 2}$ which was found in 84% of the trajectory frames. This residue was found to be deeply bound into P4 binding pocket through hydrophobic interactions. However, the backbone/side chain hydrogen bond enhanced the binding ability of the peptide.

Backbone Leu7 amide group and side chain carboxyl group of Gluβ74 were found in hydrogen bond state for

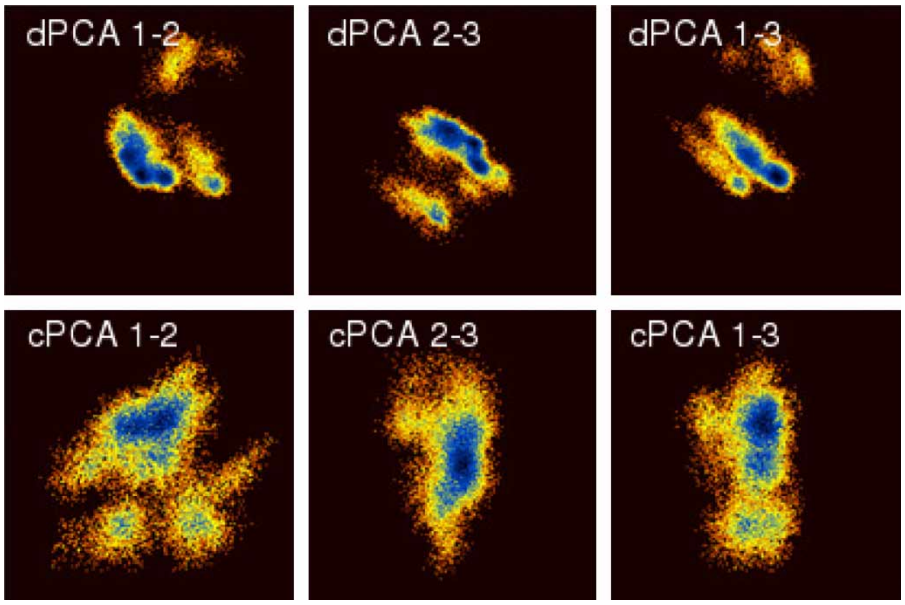

Figure 4. dPCA and cPCA of the backbone atoms of the peptide. The diagrams are pseudo-colour representations of the density functions ($\Delta G = -k_B \ln(p/p_{max})$) corresponding to the fluctuations of the N, C′, C^{α} atoms on the top three eigenvectors. Cartesian PCA is denoted as cPCA, whereas dihedral PCA is denoted as dPCA.

approximately 50% of the time. In the X-ray structure, the distance between Leu7:N and Gluβ74:$O^{\varepsilon 2}$ atoms was found to be 0.316 nm suggesting a hydrogen bond. However, side chain $O^{\varepsilon 1}$ and $O^{\varepsilon 2}$ atoms of Gluβ74 competed as acceptors of the hydrogen bond during the MD trajectory. Moreover, approximately half of the trajectory the hydrogen bond was not found at all. Time series analysis revealed that this specific bond broke 726 times during the 20 ns trajectory. Formation and breaking of this hydrogen bond was almost uniformly distributed during the trajectory.

Another relatively weak hydrogen bond was observed between Val8: and Asnα62:δ1 atoms. The corresponding distance in the X-ray structure was found to be 0.361 nm. This hydrogen bond was found in approximately 21% of the trajectory frames and, as in the previous case, uniformly distributed over simulation time.

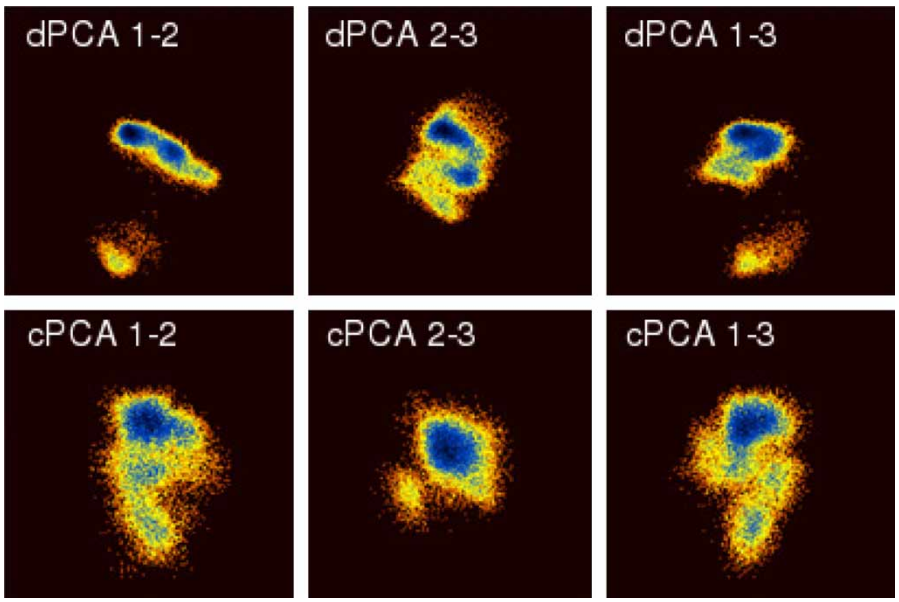

Figure 5. dPCA and cPCA of the backbone atoms of the peptide, taking into consideration only the Glu3–Glu11 residues. The diagrams are pseudo-colour representations of the density functions ($\Delta G = -k_B \ln(p/p_{max})$) corresponding to the fluctuations of the N, C′, C^{α} atoms on the top three eigenvectors. Cartesian PCA is denoted as cPCA, whereas dihedral PCA is denoted as dPCA.

Table 1. Hydrogen bonds found between the peptide and the HLA molecule.

Donor	Acceptor	Distance (nm)	% Occurrence
Hisβ81:$N^{\varepsilon 2}$	Leu1:O	0.612	13
Hisβ81:$N^{\varepsilon 2}$	Val2:O	0.300	46
Argα52:N^{ε}	Glu3:$O^{\varepsilon 1}$	0.373	97
Argα52:$N^{\eta 2}$	Glu3:$O^{\varepsilon 1}$	0.263	88
Glu3:N	Glu3:$O^{\varepsilon 2}$	0.396	89
Asnβ82:$N^{\delta 2}$	Ala4:O	0.307	74
Ala4:N	Asnβ82:$^{\delta 1}$	0.281	38
Asnα62:$N^{\delta 2}$	Tyr6:O	0.307	84
Tyr6:O^{η}	Glyβ13:O	0.512	17
Leu7:N	Gluβ74:$O^{\varepsilon 1,2}$	0.316	52
Val8:N	Asnα62:$^{\delta 1}$	0.361	21
Asnα69:$N^{\delta 2}$	Cys9:O	0.295	100
Cys9:N	Tyrβ30:O^{η}	0.290	97
Cys9:S^{γ}	Tyrβ30:O^{η}	0.365	15
Cys9:S^{γ}	Tyrβ47:O^{η}	0.332	38
Trpβ61:$N^{\varepsilon 1}$	Gly10:O	0.293	99
Hisα68:$N^{\varepsilon 2}$	Glu11:O	0.287	70
Glu11:N	Asnα69:$^{\delta 1}$	0.276	100
Argα76:$N^{\eta 2}$	Glu11:$O^{\varepsilon 1}$	0.266	100
Argα76:$N^{\eta 1}$	Glu11:$O^{\varepsilon 2}$	0.304	100
Arg12:N^{ε}	Gly14:$O^{\tau 1,2}$		92
Arg12:$N^{\eta 2}$	Gly14:$O^{\tau 1,2}$		85

Notes: The table lists the corresponding donor–acceptor atom pair, the distance found in the X-ray structure and the percentage of hydrogen bond occurrence in the MD trajectory. Greek letters α and β denote the α and β chains of the MHC molecule, respectively.

One of the strongest hydrogen bonds found between the peptide and the HLA-DQ8 during the current simulation was between Asnα69:$N^{\delta 2}$ and Cys9:O atoms. This hydrogen was also detected in the X-ray structure of the complex (distance 0.295 nm) and it was conserved for approximately 100% of the simulation time. A similar very well-conserved hydrogen bond was also found between Cys9:N and Tyrβ30:O^{η} atoms. Initial distance was found to be 0.290 nm, although 97% of the trajectory frames satisfied the hydrogen bond criteria for this pair of donor/acceptor. It is the second case, after Ala4, that both backbone polar groups of the same residue are involved in hydrogen bond interactions. It is also interesting to note that significantly higher percentages were recorded for both hydrogen bonds in which backbone atoms of Cys9 (P7 binding pocket) were involved, in comparison with Ala4 (P2 binding pocket). The side chain of the Cys9 residue acted as hydrogen bond donor with Tyrβ47:O^{η}. The distance between Cys9:S^{γ} and Tyrβ47:O^{η} was found to be 0.332 nm in the X-ray structure of the peptide/HLA-DQ8 complex. Approximately 38% of the trajectory frames met the hydrogen bond criteria for this pair of atoms, which classifies this interaction of moderate power. However, the existence of three hydrogen bond interactions in only one residue might be considered quite an interesting case and underlines the importance of the P7 binding pocket that locates the Cys9 residue.

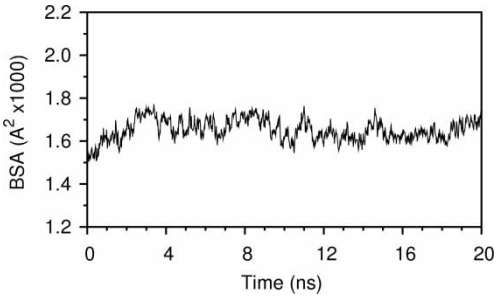

Figure 6. Time evolution of the BSA of the peptide/HLA-DQ8 interface.

The Gly10 residue was found in a relatively exposed position in the X-ray structure. For example, backbone amide group was found to be hydrated for 92% of the simulation time. On the contrary, the carboxyl group of the Gly10 residue accepted a hydrogen bond from Trpβ61:$N^{\varepsilon 1}$ atom for approximately 99% of the simulation time. The initial distance between Trpβ61:$N^{\varepsilon 1}$ and Gly10:O atoms was found to be 0.293 nm.

The P9 binding pocket of MHC molecule accommodated the Glu11 residue. Its acidic side chain was found to form (almost unbreakable) hydrogen bond with side chain of Argα76. Moreover, Glu11:N–Asnα69:$O^{\delta 1}$ was also perfectly conserved during MD trajectory. The importance of this third hydrogen bond is clearly seen from the RMSF analysis. Glu3 residue also made side chain hydrogen bond contacts with the MHC molecule. In contrast with Glu11 residue, backbone atoms of Glu3 residue did not participate in hydrogen bond interactions with the MHC molecule. The resulted RMSF values of backbone C^{α} atoms, 0.105 and 0.077 nm, for Glu3 and Glu11, respectively, reflect perfectly this situation and underline very well the importance of backbone interactions. The Glu3 residue made less contact and it was found to show increased flexibility. From this point of view, P9 binding pocket seems to be suitable for binding other amino acids with backbone interactions, of course with the penalty of reduced affinity.

3.3 BSA calculations

The initial BSA value was found to be 1418.6 Å^2. The BSA value varied between 1443.3 and 1845.2 Å^2 with an average value of 1651.8 (57.5) Å^2. Time series of this quantity are shown in Figure 6. The slight increase in the BSA value and its stability in the MD trajectory perfectly underlies the strength of peptide/MHC binding.

4. Conclusions

Classical molecular dynamics simulations were utilised to enlighten the interactions of the peptide and the MHC

molecule. Peptide LVEALYLVCGERGG complexed with the HLA-DQ8 was simulated in explicit water for 20 ns taking advantage of the recent progress in computational technology and parallel computation. Such simulations were shown to enlighten our knowledge of immunological complexes. Interactions involving HLA-DQ8 were focused on P1, P2 and P9 binding pockets. The current study corroborates their significance in peptide loading. However, analysis of the crystal structure, supported by simulation results, revealed important peptide/MHC or peptide/peptide interactions that might have serious consequences in T cell recognition of the peptide/MHC complex.

A series of hydrogen bond interactions between the peptide and MHC molecule, not previously focused, were established. With the exclusion of polar binding pockets P1 and P9, which accommodated acidic residues Glu3 and Glu11, respectively, six out of the seven remaining residues of 9mer core of the peptide made important hydrogen bonds with the MHC molecule. For example, Ala4 (pocket P2), Cys9 (pocket P7) and Gly10 (pocket P8) made important hydrogen bond interactions with the MHC molecule, something that has been somewhat underestimated until now. The existence of approximately nine intra-molecular hydrogen bonds involving backbone atoms of the peptide is something rarely seen in other cases, in which four–five hydrogen bonds are commonly found. This situation can also implicate non-specific binding of peptides by HLA-DQ8 molecule of different sequences, under the penalty of possible reduced affinity, because backbone hydrogen bonds can be formed, in principle, by any residue.

Acknowledgements

Parallel execution of NAMD was performed at the Research Centre for Scientific Simulations (RCSS) of the University of Ioannina. The open source community (Linux, NAMD, GNU, etc) is gratefully acknowledged for releasing to public all the necessary computer software needed for this research.

References

[1] E.Y. Jones, L. Fugger, J.L. Strominger, and C. Siebold, *MHC class II proteins and disease: A structural perspective*, Nat. Rev. Immunol. 6 (2006), pp. 271–282.

[2] E.J. Sundberg, L. Deng, and R.A. Mariuzza, *TCR recognition of peptide/MHC class II complexes and superantigens*, *Seminars in Immunology*, Vol. 19, Elsevier, 2007, pp. 262–271. Available at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2949352

[3] K.W. Wucherpfennig and G.S. Eisenbarth, *Type 1 diabetes*, Nat. Immunol. 2 (2001), pp. 767–768.

[4] K.H. Lee, K.W. Wucherpfennig, and D.C. Wiley, *Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes*, Nat. Immunol. 2 (2001), pp. 501–507.

[5] G. Rajagopalan, A.K. Mangalam, M.M. Sen, S. Cheng, Y.C. Kudva, and C.S. David, *Autoimmunity in HLA-DQ8 transgenic mice expressing granulocyte/macrophage-colony stimulating factor in the beta cells of islets of langerhans*, Autoimmunity 40 (2007), pp. 169–179.

[6] D. Daniel and D.R. Wegmann, *Protection of nonobese diabetic mice from diabetes by intranasal or subcutaneous administration of insulin peptide B-(9-23)*, Proc. Natl Acad. Sci. USA 93 (1996), pp. 956–960.

[7] A. Kosmopoulou, M. Vlassi, A. Stavrakoudis, C. Sakarellos, and M. Sakarellos-Daitsiotis, *T-cell epitopes of the La/SSB autoantigen: Prediction based on the homology modeling of HLA-DQ2/DQ7 with the insulin-B peptide/HLA-DQ8 complex*, J. Comput. Chem. 27 (2006), pp. 1033–1044.

[8] S. Costantini, M. Rossi, G. Colonna, and A.M. Facchiano, *Modelling of HLA-DQ2 and its interaction with gluten peptides to explain molecular recognition in celiac disease*, J. Mol. Graph. Model. 23 (2005), pp. 419–431.

[9] W.F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D.P. Geerke, A. Glättli, P.H. Hünenberger, M.A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N.F.A. van der Vegt, and H.B. Yu, *Biomolecular modeling: Goals, problems, perspectives*, Angew. Chem. Int. Ed. Engl. 45 (2006), pp. 4064–4092.

[10] D. Morikis and J.D. Lambris, *Physical methods for structure, dynamics and binding in immunological research*, Trends Immunol. 25 (2004), pp. 700–707.

[11] B. Mallik and D. Morikis, *Applications of molecular dynamics simulations in immunology: A useful computational method in aiding vaccine design*, Curr. Proteom. 3 (2006), pp. 259–270.

[12] A. Stavrakoudis, *Conformational flexibility in designing peptides for immunology: The molecular dynamics approach*, Curr. Comput.-Aid. Drug Des. 6 (2010), pp. 207–222.

[13] R. Galeazzi, *Molecular dynamics as a tool in rational drug design: Current status and some major applications*, Curr. Comput.-Aid. Drug Des. 5 (2009), pp. 225–240.

[14] J. Wang and M.J. Eck, *Assembling atomic resolution views of the immunological synapse*, Curr. Opin. Immunol. 15 (2003), pp. 286–293.

[15] A. Stavrakoudis, *Cis-trans isomerization of the Epstein-Barr virus determinant peptide EENLLDFVRF after the DM1 TCR recognition of the HLA-B*4405/peptide complex*, FEBS Lett. 585 (2011), pp. 485–491.

[16] S. Wan, D.R. Flower, and P.V. Coveney, *Toward an atomistic understanding of the immune synapse: Large-scale molecular dynamics simulation of a membrane-embedded TCR-pMHC-CD4 complex*, Mol. Immunol. 45 (2008), pp. 1221–1230.

[17] D.R. Flower, K. Phadwal, I.K. Macdonald, P. Coveney, M.N. Davies, and S. Wan, *T-cell epitope prediction and immune complex simulation using molecular dynamics: State of the art and persisting challenges*, Immun. Res. 6 (2010), p. S4. Available at http://www.immunome-research.com/content/6/S2/S4

[18] M. Zacharias and S. Springer, *Conformational flexibility of the MHC class I a1-a2 domain in peptide bound and free states: A molecular dynamics simulation study*, Biophys. J. 87 (2004), pp. 1–12.

[19] V. Zoete, M.B. Irving, and O. Michielin, *MM-GBSA binding free energy decomposition and T cell receptor engineering*, J. Mol. Recognit. 23 (2009), pp. 142–152.

[20] U. Omasits, B. Knapp, M. Neumann, O. Steinhauser, H. Stockinger, R. Kobler, and W. Schreiner, *Analysis of key parameters for molecular dynamics of pMHC molecules*, Mol. Simulat. 34 (2008), pp. 781–793.

[21] C.A. Painter, A. Cruz, G.E. Lopez, L.J. Stern, and Z. Zavala-Ruiz, *Model for the peptide-free conformation of class II MHC proteins*, PLoS ONE 3(6) (2008), p. e2403. Available at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2408972

[22] R. Yaneva, S. Springer, and M. Zacharias, *Flexibility of the MHC class II peptide binding cleft in the bound, partially filled, and empty states: A molecular dynamics simulation study*, Biopolymers 91 (2009), pp. 14–27.

[23] W.F. van Gunsteren, J. Dolenc, and A.E. Mark, *Molecular simulation as an aid to experimentalists*, Curr. Opin. Struct. Biol. 18 (2008), pp. 149–153.

[24] H.M. Berman, T. Battistuz, T.N. Bhat, W.F. Bluhm, P.E. Bourne, K. Burkhardt, Z. Feng, G.L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J.D. Westbrook, and C. Zardecki, *The Protein Data Bank*, Acta Crystall. Sec. D: Biol. Crystall. 58 (2002), pp. 899–907.

[25] S. Wan, P. Coveney, and D.R. Flower, *Large-scale molecular dynamics simulations of HLA-A*0201 complexed with a tumor-*

specific antigenic peptide: Can the α3 and β2 m domains be neglected? J. Comput. Chem. 25 (2004), pp. 1803–1813.
[26] A.D. MacKerell, Jr, M. Feig, and C.L. Brooks, *Improved treatment of the protein backbone in empirical force fields*, J. Am. Chem. Soc. 126 (2004), pp. 698–699.
[27] A.D. MacKerell, Jr, M. Feig, and C.L. Brooks, *Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations*, J. Comput. Chem. 25 (2004), pp. 1400–1415.
[28] W. Humphrey, A. Dalke, and K. Schulten, *VMD: Visual molecular dynamics*, J. Mol. Graph. 14 (1996), pp. 33–38.
[29] M. Buck, S. Bouguet-Bonnet, R.W. Pastor, and A.D. MacKerell, Jr, *Importance of the CMAP correction to the CHARMM22 protein force field: Dynamics of hen lysozyme*, Biophys. J. 90 (2006), pp. 36–38.
[30] A. Stavrakoudis, *Molecular dynamics simulations of an apoliprotein derived peptide*, Chem. Phys. Lett. 461 (2008), pp. 294–299.
[31] P. Mark and L. Nilsson, *Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K*, J. Phys. Chem. A 105 (2001), pp. 9954–9960.
[32] W. Weber, P.H. Hunenberger, and J.A. McCammon, *Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: Influence of artificial periodicity on peptide conformation*, J. Phys. Chem. B 104 (2000), pp. 3668–3675.
[33] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, and K. Schulten, *Scalable molecular dynamics with NAMD*, J. Comput. Chem. 26 (2005), pp. 1781–1802.
[34] T. Darden, D. York, and L. Pedersen, *Particle mesh Ewald: An* Nlog(N) *method for Ewald sums in large systems*, J. Chem. Phys. 98 (1993), pp. 10089–11092.
[35] A.Y. Toukmaji and J.A. Board, Jr, *Ewald summation techniques in perspective: A survey, Comput*, Phys. Commun. 95 (1996), pp. 73–92.
[36] W.G. Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A 31 (1985), pp. 1695–1697.
[37] S.E. Feller, Y. Zhang, R.W. Pastor, and B.R. Brooks, *Constant pressure molecular dynamics simulation: The Langevin piston method*, J. Chem. Phys. B 103 (1995), pp. 4613–4621.
[38] N.M. Glykos, *Carma: A molecular dynamics analysis program*, J. Comput. Chem. 27 (2006), pp. 1765–1768.
[39] I.G. Tsoulos and A. Stavrakoudis, *Eucb: A C++ program for molecular dynamics trajectory analysis*, Comput. Phys. Commun. 182 (2011), pp. 834–841.
[40] D. Frishman and P. Argos, *Knowledge-based protein secondary structure assignment*, Proteins 23 (1995), pp. 566–579.
[41] I.T. Jolliffe, *Principal Component Analysis*, Springer, New York, 2002.
[42] Y. Mu, P.H. Nguyen, and G. Stock, *Energy landscape of a small peptide revealed by dihedral angle principal component analysis*, Proteins 58 (2005), pp. 45–52.
[43] G.G. Maisuradze and D.M. Leitner, *Free energy landscape of a biomolecule in dihedral principal component space: Sampling convergence and correspondence between structures and minima*, Proteins 67 (2007), pp. 569–578.
[44] A. Altis, M. Otten, P.H. Nguyen, R. Hegger, and G. Stock, *Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis*, J. Chem. Phys. 128 (2008), 245102.
[45] A. Lee, S. Kinnear, and A. Wand, *Redistribution and loss of side chain entropy upon formation of a calmodulin–peptide complex*, Nat. Struct. Biol. 7 (2000), pp. 72–77.