Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 25 (2007) 605–615
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Wave function analysis of MHC–peptide interactions

Constanza Cárdenas [a,b,*], Mateo Obregón [a], Alejandro Balbín [a], José Luis Villaveces [b], Manuel E. Patarroyo [a,c]

[a] *Fundación Instituto de Inmunología, Kra 50 No. 26-00, Bogotá, Colombia*
[b] *Grupo de Química Teórica, Universidad Nacional de Colombia, Bogotá, Colombia*
[c] *Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia*

Received 19 September 2005; received in revised form 7 April 2006; accepted 10 April 2006
Available online 28 April 2006

Abstract

We have carried out an analysis of the wave function data for three MHC–peptide complexes: HLA-DRβ1*0101-HA, HLA-DRβ1*0401-HA and HLA-DRβ1*0401-Col. We used quantum chemistry computer programs to generate wave function coefficients for these complexes, from which we obtained both molecular and atomic orbital data for both pocket and peptide amino acids within each pocket region. From these discriminated data, interaction molecular orbitals (IMOs) were identified as those with large and similar atomic orbital coefficient contributions from both pocket and peptide amino acids. The present results correlate well with our previous research where only electrostatic moments were used to explore molecular component interactions. Furthermore, we show a quantum chemical methodology to produce more fine-grained results concerning amino acid behavior in the MHC–peptide interaction.

Keywords: Molecular interactions; Molecular orbital coefficients; MHC pockets; Theoretical study of MHC

1. Introduction

The bonding of major histocompatibility complex (MHC) molecules with peptidic fragments is the first step in the generation of a successful immune system response. Subsequently, this complex is presented and recognized by T cell receptor (TCR) molecules, which in turn activate the body's defense mechanism [1]. The study of complexes formed by MHC molecules and antigenic peptides has been conducted from different perspectives: the study of complexes by mean of X-ray crystallography [2–4], the analysis of the different behaviors of various alleles with respect to antigenic peptides [5–9], and the construction of virtual matrices and data bases for the prediction of epitopes for MHC complexes [10]. Based on experimental results, several rules of interaction have been formulated. One of these separates the MHC molecule into interaction regions, so-called ''pockets'', an approach used by us in previous works to analyze, from an electrostatic point of view, several complexes that had previously been crystallized [11–13]. These studies have provided interesting results that correlate well with experimental counterparts found in scientific literature.

An analysis of the molecule's wave function obtained from quantum chemical calculations, while using the same hypothesis that molecular interactions are electrostatic in essence, provides a different point of view of the interaction.

The description of the molecular system and its properties is contained in the system's wave function; it is through the analysis of the wave function that we can hope to reach the ''holy grail'' that contains an explicit and a priori description of molecular interactions.

1.1. Atomic and molecular orbital coefficients

The method of analysis used for this research depends on the fine-grain exploration of molecular and atomic orbital coefficients that are found in an ab initio manner from the wave function for the collection of nuclei in 3D space, of the complex pocket–peptide obtained. Each coefficient represents the contribution of the particular atomic and molecular orbital

* Corresponding author. Tel.: +57 1 3158919x122; fax: +57 1 3158919x108.
E-mail addresses: constanza_cardenas@fidic.org.co, conscc@yahoo.com (C. Cárdenas).
URL: http://www.fidic.org.co

1093-3263/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2006.04.005

to the wave function and thus gives an indication of the importance of the associated atom. For molecular orbitals that are shared among pocket and peptide, this is a measure of the contribution of the particular atom to the formation of the complex.

For the coefficients C in any molecular orbital, the following equations hold:

$$\sum_{i}^{\{\text{all atoms in complex}\}} C_i^2 = 1 \tag{1}$$

$$\sum_{k}^{\{\text{pocket atoms}\}} C_k^2 + \sum_{p}^{\{\text{peptide atoms}\}} C_p^2 = 1 \tag{2}$$

Since these coefficients are manipulated by simple additive processes in the wave function, we can calculate the relative contribution of each amino acid to a molecular orbital by summing the squares of the atomic orbital coefficients for each of its constituent atoms.

In our case we have two interacting fragments, namely, those amino acids participating in the HLA (human leucocite antigen[1]) pocket and those amino acids belonging to the peptide that face the pocket region. By selecting molecular orbitals that have high and numerically equivalent orbital coefficient contributions from each of these two fragments, peptide and HLA, we can focus on those molecular orbitals that have a similar involvement from both fragments, a necessary condition for molecular interactions. Those atoms whose atomic orbitals are of importance in these molecular orbitals will be the atoms involved in the formation of the complex.

The present study aims at comparing two human HLA-DR class II alleles (HLA-DRβ1*0101 [4] and HLA-DRβ1*0401 [3]) binding to the influenza hemaglutinin peptide (HA 306-318) and to the collagen II peptide (Col-II 1168-1179 [2]). The purpose of this research is to evaluate differences between these complexes and contribute to the understanding of its formation.

This study complements previous research using multipole moments by contributing a new and more profound vision into the workings of these molecular interactions. Furthermore, such knowledge will expand our understanding of the role these molecules have in immunity and autoimmune phenomena.

2. Methodology

Briefly, the methodology we applied is

(1) Obtain the coordinates in 3D space of the nuclei that are included in the region of the molecular complex to be studied: HLA-DRβ1*0401-Col II (PDB CODE:2SEB), HLA-DRβ1*0401-HA (PDB CODE: 1J8H) and HLA-DRβ1*0101-HA (PDB CODE:1DLH) (named DR4-Col, DR4-HA and DR1-HA, in the following article). These coordinates were extracted from the protein data bank (PDB) [14,15].

(2) Selection of nuclei participating in each interaction region or pocket for each of the complexes (see Table 1).

(3) Calculation of the electronic wave function for each of the complex with the nuclear geometry obtained above, using Gaussian'98 computer program [16]. We employed Hartree-Fock-MO-LCAO[2] method with a 3-21G* basis set [17,18] for these calculations.

(4) To obtain a survey of all possible electronic situations, we successively replaced the occupying amino acid that fits into the MHC pocket[3] with the remaining 19 genetically encoded amino acids and carried out the above calculation of the wave function for each of these plausible substituent amino acids.

(5) Extraction of the wave function coefficients for each possible atomic and molecular orbital from the Gaussian'98 output file. We developed the QuantumC2© computer program to carry out this task[4]. This program identifies also the amino acid corresponding to each atom, and hence each coefficient is labeled as either a member of the MHC pocket amino acids or a member of the peptide amino acids.

(6) The analysis of molecular orbitals begins by separating each molecular orbital into the constituent contributions parametrized by Eq. (2). The magnitudes of the pocket contribution K and the peptide contribution P are given by Eqs. (3) and (4), respectively:

$$K = \sqrt{\sum_{k}^{\{\text{pocket atoms}\}} C_k^2} \tag{3}$$

$$P = \sqrt{\sum_{p}^{\{\text{peptide atoms}\}} C_p^2} \tag{4}$$

The values of K and P relative to each other give rise to three cases for any molecular orbital:

- $K \gg P$, when the contribution of orbitals from the pocket to the molecular orbital is much larger than that from the peptide orbitals. These molecular orbitals are associated with the pocket alone (Fig. 1).
- $K \ll P$, in this case, the contribution to the molecular orbital from amino acids in the pocket is much less than the contribution from amino acids in the peptide. This is a molecular orbital associated with the peptide alone (Fig. 2).
- $K \approx P$, when the contribution from amino acids in the pocket and in the peptide to the molecular orbital is equivalent (''shared or interaction molecular orbitals''). In this case, atomic orbitals centered on both atoms from both MHC pocket amino acids and peptide amino acids

[1] This corresponds to human MHC.

[2] Hartree-Fock, with molecular orbitals as a linear combination of atomic orbitals.

[3] This occupying amino acid is identified as that which is most encased by the pocket amino acids as seen in the experimentally obtained crystallized structure.

[4] The Quantum C2© source code written in Common Lisp can be obtained by e-mail request to mateo_obregon@fidic.org.co.

Table 1
Amino acid composition of the interaction or pocket regions for the complexes studied

Pocket	Chain	Complex	Residues
P1	α	DR4-Col	I_7 F_{24} *D_{25}* F_{26} *D_{27}* F_{32} W_{43} A_{52} S_{53} F_{54} E_{55}
		DR4-HA	I_7 F_{24} *D_{25}* F_{26} *D_{27}* F_{32} W_{43} A_{52} S_{53} F_{54} E_{55}
		DR1-HA	I_7 F_{24} *D_{25}* F_{26} *D_{27}* F_{32} W_{43} A_{52} S_{53} F_{54} E_{55}
	β	DR4-Col	H_{81} Y_{83} *G_{84}* V_{85} G_{86} F_{89} T_{90} V_{91}
		DR4-HA	H_{81} Y_{83} *G_{84}* V_{85} G_{86} F_{89} T_{90} V_{91}
		DR1-HA	H_{81} Y_{83} V_{85} G_{86} F_{89} T_{90} V_{91}
	Peptide	DR4-Col	A_{1168} Y_{1169} Oc_{1170} (P1) R_{1171}
		DR4-HA	P_{306} K_{307} Oc_{308} (P1) V_{309}
		DR1-HA	P_{306} K_{307} Oc_{308} (P1) V_{309}
P4	α	DR4-Col	Q_9 A_{10} E_{11} F_{24} *D_{25}* N_{62}
		DR4-HA	Q_9 A_{10} E_{11} F_{24} *D_{25}* N_{62}
		DR1-HA	Q_9 A_{10} E_{11} F_{24} *D_{25}* N_{62}
	β	DR4-Col	**H_{13}** E_{14} C_{15} *R_{25}* **F_{26}** L_{27} **D_{28}** F_{40} Q_{70} **K_{71}** R_{72} A_{73} A_{74} Y_{78} C_{79}
		DR4-HA	**H_{13}** E_{14} C_{15} **F_{26}** L_{27} **D_{28}** F_{40} Q_{70} **K_{71}** R_{72} A_{73} A_{74} *D_{76}* Y_{78} C_{79}
		DR1-HA	**F_{13}** E_{14} C_{15} *R_{25}* **L_{26}** L_{27} **E_{28}** F_{40} Q_{70} **R_{71}** R_{72} A_{73} A_{74} Y_{78} C_{79}
	Peptide	DR4-Col	A_{1172} Oc_{1173} (P4) A_{1174}
		DR4-HA	K_{310} Oc_{311} (P4) N_{312}
		DR1-HA	K_{310} Oc_{311} (P4) N_{312}
P6	α	DR4-Col	Q_9 A_{10} E_{11} A_{61} N_{62} I_{63} A_{64} V_{65} D_{66} *K_{67}* N_{69}
		DR4-HA	Q_9 A_{10} E_{11} A_{61} N_{62} I_{63} A_{64} V_{65} D_{66} *K_{67}* N_{69}
		DR1-HA	Q_9 A_{10} E_{11} A_{61} N_{62} I_{63} A_{64} V_{65} D_{66} *K_{67}* N_{69}
	β	DR4-Col	**E_9** Q_{10} **V_{11}** K_{12} **H_{13}** **D_{28}** R_{29} **Y_{30}** *D_{66}* **K_{71}**
		DR4-HA	**E_9** Q_{10} **V_{11}** K_{12} **H_{13}** **D_{28}** R_{29} **Y_{30}** *D_{66}* **K_{71}**
		DR1-HA	**W_9** Q_{10} **L_{11}** K_{12} **F_{13}** **E_{28}** R_{29} **C_{30}** *D_{66}* **R_{71}**
	Peptide	DR4-Col	A_{1174} Oc_{1175} (P6) A_{1176}
		DR4-HA	N_{312} Oc_{313} (P6) L_{314}
		DR1-HA	N_{312} Oc_{313} (P6) L_{314}
P7	α	DR4-Col	V_{65} *D_{66}* *K_{67}* N_{69}
		DR4-HA	V_{65} *D_{66}* N_{69}
		DR1-HA	V_{65} *D_{66}* N_{69}
	β	DR4-Col	**V_{11}** **H_{13}** **D_{28}** **Y_{30}** V_{38} Y_{47} W_{61} Q_{64} *D_{66}* L_{67} **K_{71}**
		DR4-HA	**V_{11}** **H_{13}** **D_{28}** **Y_{30}** V_{38} Y_{47} W_{61} Q_{64} *D_{66}* L_{67} *E_{69}* **K_{71}**
		DR1-HA	**L_{11}** **F_{13}** **E_{28}** **C_{30}** V_{38} *E_{46}* Y_{47} W_{61} Q_{64} L_{67} **R_{71}**
	Peptide	DR4-Col	A_{1175} Oc_{1176} (P7) G_{1177}
		DR4-HA	T_{313} Oc_{314} (P7) K_{315}
		DR1-HA	T_{313} Oc_{314} (P7) K_{315}
P9	α	DR4-Col	*D_{66}* *K_{67}* A_{68} N_{69} L_{70} E_{71} I_{72} M_{73} *K_{75}* R_{76}
		DR4-HA	*D_{66}* *K_{67}* A_{68} N_{69} L_{70} E_{71} I_{72} M_{73} R_{76}
		DR1-HA	A_{68} N_{69} L_{70} E_{71} I_{72} M_{73} R_{76}
	β	DR4-Col	**E_9** **Y_{30}** **Y_{37}** V_{38} R_{55} D_{57} Y_{60} W_{61}
		DR4-HA	**E_9** **Y_{30}** **Y_{37}** V_{38} *E_{46}* R_{55} D_{57} Y_{60} W_{61}
		DR1-HA	**W_9** *R_{29}* **C_{30}** **S_{37}** V_{38} D_{57} *E_{59}* Y_{60} W_{61}
	Peptide	DR4-Col	G_{1177} Oc_{1178} (P9) G_{1179}
		DR4-HA	K_{315} Oc_{316} (P9) A_{317} T_{318}
		DR1-HA	K_{315} Oc_{316} (P9) A_{317} T_{318}

Residues that are different between MHC alleles are shown in bold, while residues that are incorporated into the calculations so as to maintain an overall neutrally charged complex are shown in italic.

contribute with similar $\sqrt{\sum C^2}$ magnitude to the molecular orbital (Fig. 3).

Eqs. (1) and (2) are not equal to one if the originating wave function is not normalized, but the difference is a constant scaling factor and the partition of coefficients given by Eq. (2), which is the important point for us here, will still hold. The important issue here is to identify which molecular orbitals are formed by similar contributions from pocket and peptide amino acids in the MHC–peptide complex. Such molecular orbitals are called ''interaction molecular orbitals'' (IMOs), and are the orbitals used in this

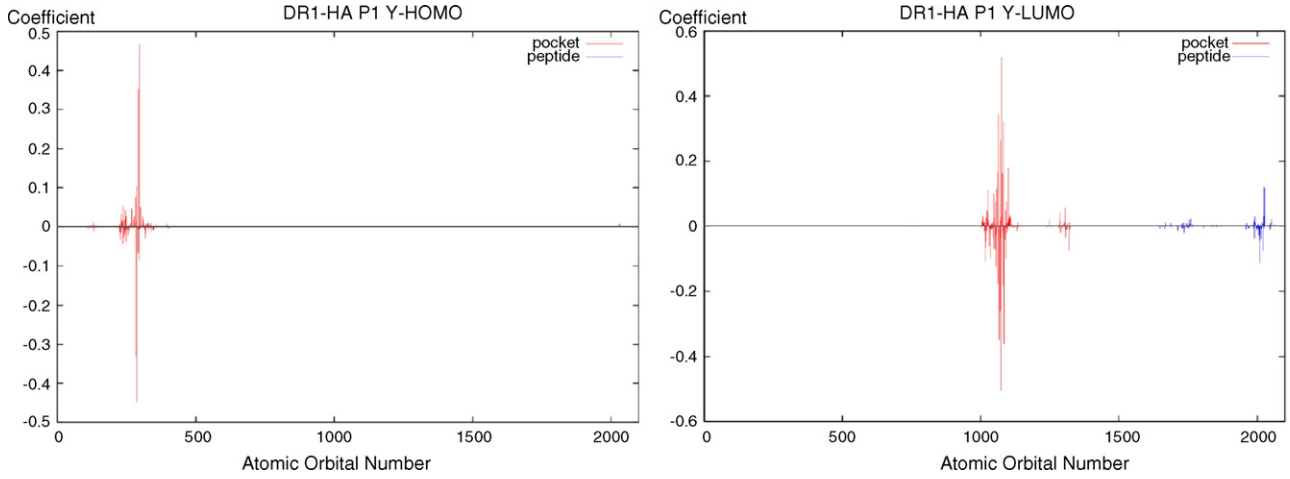

Fig. 1. Molecular orbitals that are almost exclusively made up of atomic orbitals which are centered on MHC amino acid atoms. As an example, the HOMO and LUMO orbitals that correspond to Pocket 1 in the DR1-HA filled with the occupying Tyr are shown. The *X*-axis correspond to the ordinal number of the atomic orbitals that conforms the molecular orbital according to the basis set used in the calculation, and the *Y*-axis is the value of the coefficient for each atomic orbital in the linear combination for the molecular orbital (according to LCAO approximation).

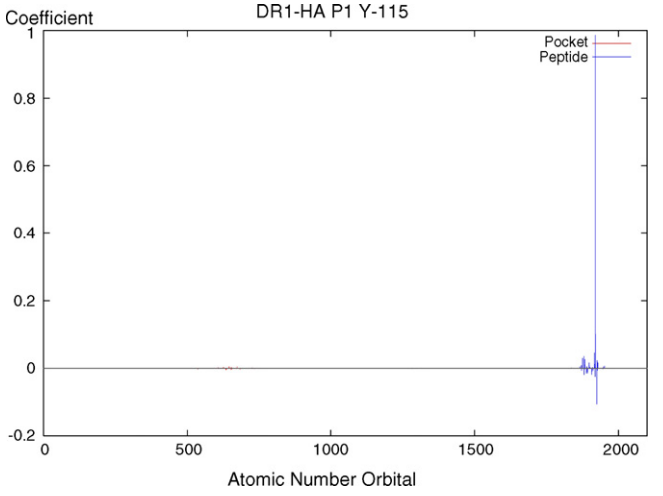

Fig. 2. Molecular orbitals that are almost exclusively formed by atomic orbitals centered on atoms of the peptide amino acids. As an example, molecular orbital number 115, corresponding to Pocket 1 in the DR1-HA filled with the occupying Tyr is shown (see Fig. 1 for the meaning of axes).

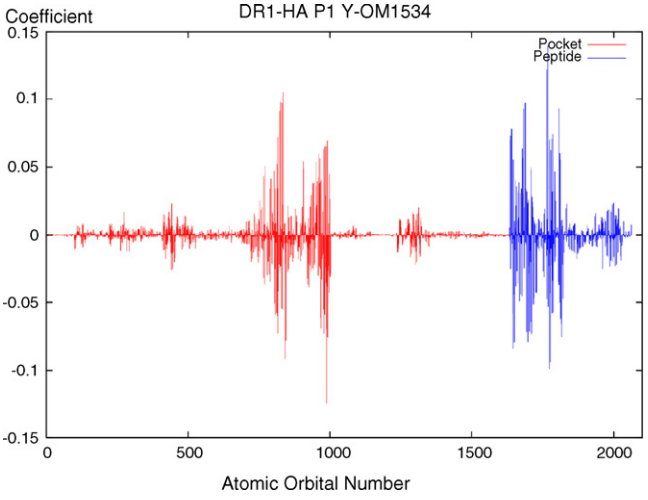

Fig. 3. Molecular orbitals that have equivalent contributions from both MHC and peptide atomic orbitals. As an example, molecular orbital 534 with similar contributions from both MHC and peptide atoms that correspond to Pocket 1 in the DR1-HA filled with the occupying Tyr is shown (see Fig. 1 for the axes descriptions).

analysis of the molecular wave function. The selection criteria that we used was to choose those molecular orbitals for which the difference between the pocket (K) contribution and the peptide (P) contribution is less than 10%:

$$\frac{|K-P|}{K} \leq 0.1 \tag{5}$$

(7) Since each atomic orbital may be assigned to a particular amino acid, we further employed principal component analysis (PCA) on the IMOs to select the principal components in the IMOs space. We built a matrix, where each amino acid in the complex occupied a column (both MHC and peptide amino acids included).

We took the first seven principal components (PCs) with eigenvalues greater than 1, which accounted for more than 70% of the variance[5]. In this new space we took the contributions of the amino acids by calculating the factor loadings summed over the IMOs for each variable. The larger these values are, the more important the amino acid is in the interaction between MHC and peptide. This procedure also showed which pocket amino acids were more highly correlated with peptide amino acids. We performed hierarchical clustering analyses taking the factor loadings as variables, to show how the amino acids grouped together in relation to their comparative importance in the pocket–peptide interaction. By locating amino acids that group together across different complexes, we were then able to identify generalities concerning the amino acids.

[5] We used the SPSS computer program [19] for this part of the analysis.

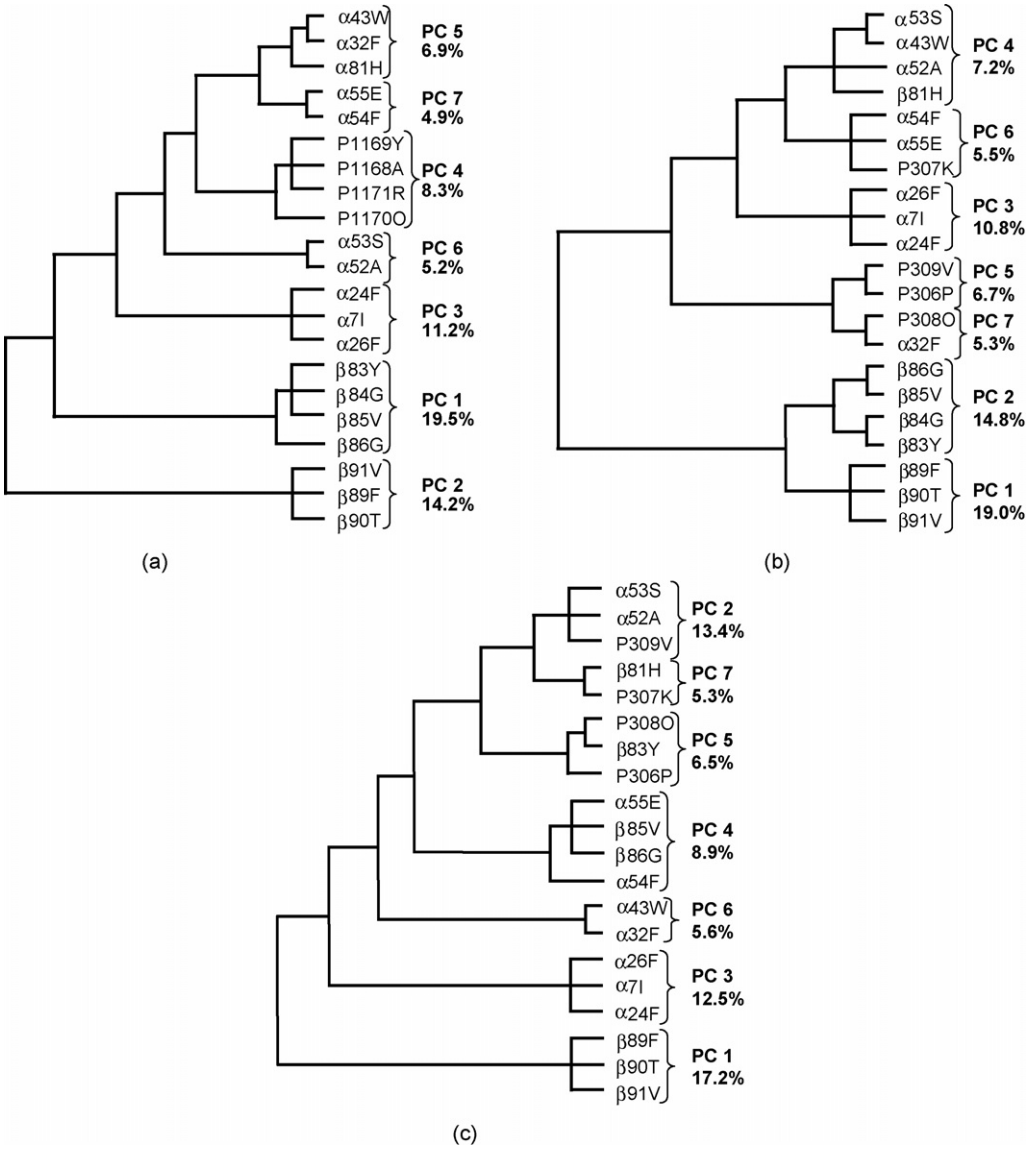

Fig. 4. Dendrograms for Pocket 1, highlighting amino acid cluster groups for the three complexes studied: (a) DR4-Col, (b) DR4-HA and (c) DR1-HA. α and β chains of the pocket amino acids, and *P* for peptide amino acids is shown in the item labels. Also shown are the PC number and percent of variance explained for each group, thus, highlighting groups that are more similar in PCA analysis.

3. Results and discussion

The first result to note is that IMOs have eigenvalues smaller than the eigenvalues for frontier molecular orbitals HOMO and LUMO, showing that the linking between the pockets and the peptides are second-order interactions as was to be expected (since the amino acid interactions we study do not imply either the creation or the dissolution of molecular bonds between the MHC and the peptide). In the cases we studied (20 amino acids per pocket for each of three complexes, totaling 300 cases) the frontier orbitals were exclusively made up of either pocket amino acids or peptide amino acids. These cases correspond to those represented by Figs. 1 and 2, respectively.

In the ensuing discussion, the following aspects will be considered relative to our principal component analysis and clustering analysis:

A. Which amino acids have the largest factor loadings according to the PCA data, and hence account for the greatest percentage of variance in the IMOs. These are the most affected amino acids in the pocket–peptide interaction.
B. Whether peptide amino acids have large factor loadings that are among the first principal components (of the seven extracted). This would highlight the importance of these amino acids regarding the interaction of peptide and pocket.
C. Whether a group of amino acids is found across the three complexes studied here.
D. Whether a group found in the clustering analysis includes both pocket and peptide amino acids, highlighting that these pocket amino acids are correlated with the occupying peptide.

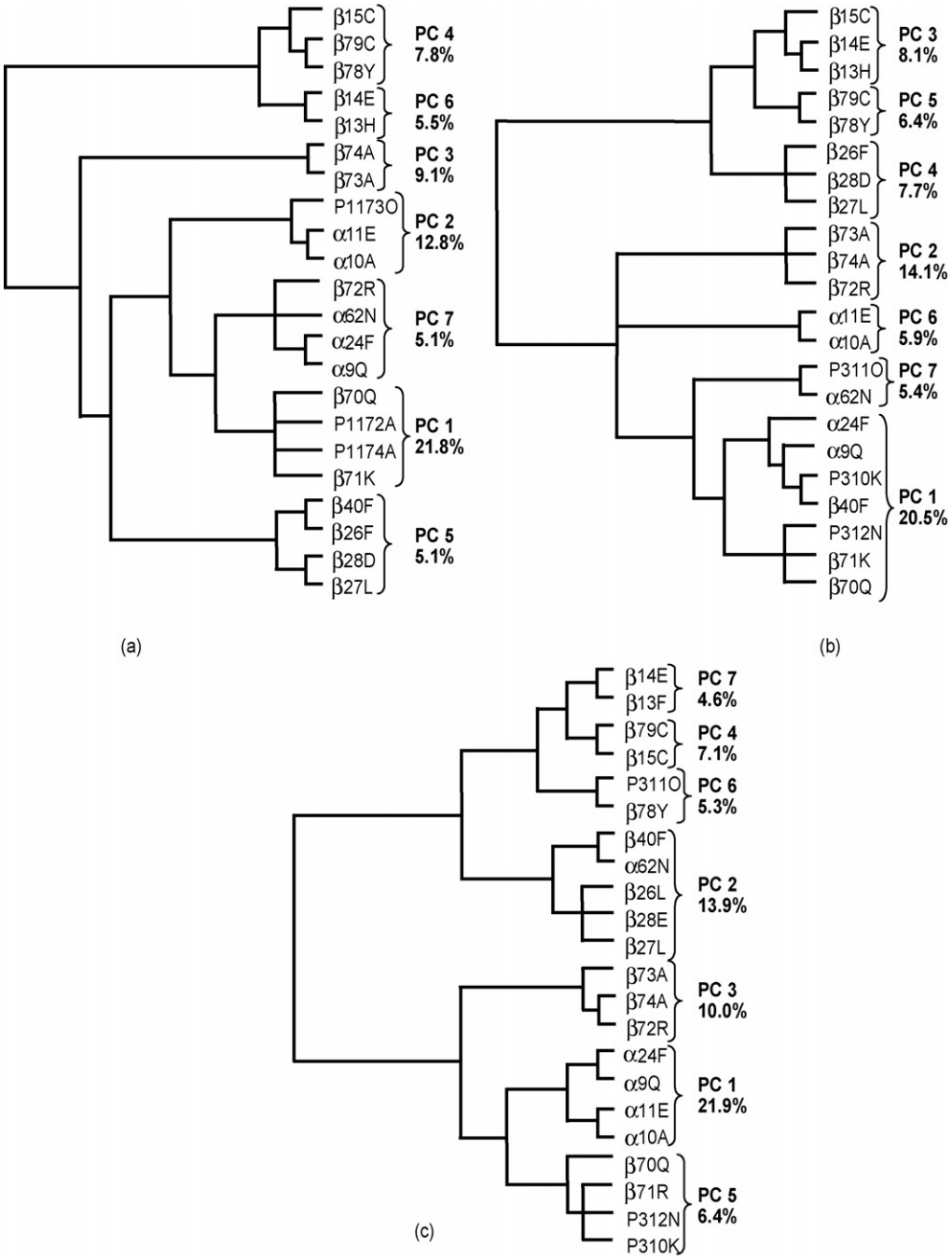

Fig. 5. Dendrograms for Pocket 4, highlighting amino acid cluster groups for the three complexes studied: (a) DR4-Col, (b) DR4-HA and (c) DR1-HA. See Fig. 4 for details.

Our principal component analysis[6] data together with the hierarchical clustering of amino acids are presented in the dendrograms shown in Figs. 4–8, and will be discussed for each pocket involved.

3.1. Pocket 1

Those target peptide amino acids (both the original and the 19 remaining substitute amino acids) whose contribution varies most are almost the same for the three complexes studied. For the first principal component (PC1) from the variance table used in the clustering, β89F, β90T and β91V in DR4-HA and DR1-HA, which corresponds to PC2 in DR4-Col. For PC2, β84G, β85V and β86G in DR4-HA, which is the PC1 for DR4-Col and PC4 for DR1-HA. For PC3 α7I, α24F and α26F in the three complexes. These amino acids are in consecutive order and they also form the base of the pocket and hence any variation in the occupying amino acid's side chain will directly alter their electronic configuration in space. The variance they account for is 19.5%, 14.2% and 11.2% for DR4-Col; 19%, 14.8% and

[6] The tables of PCA data can be obtained upon request.

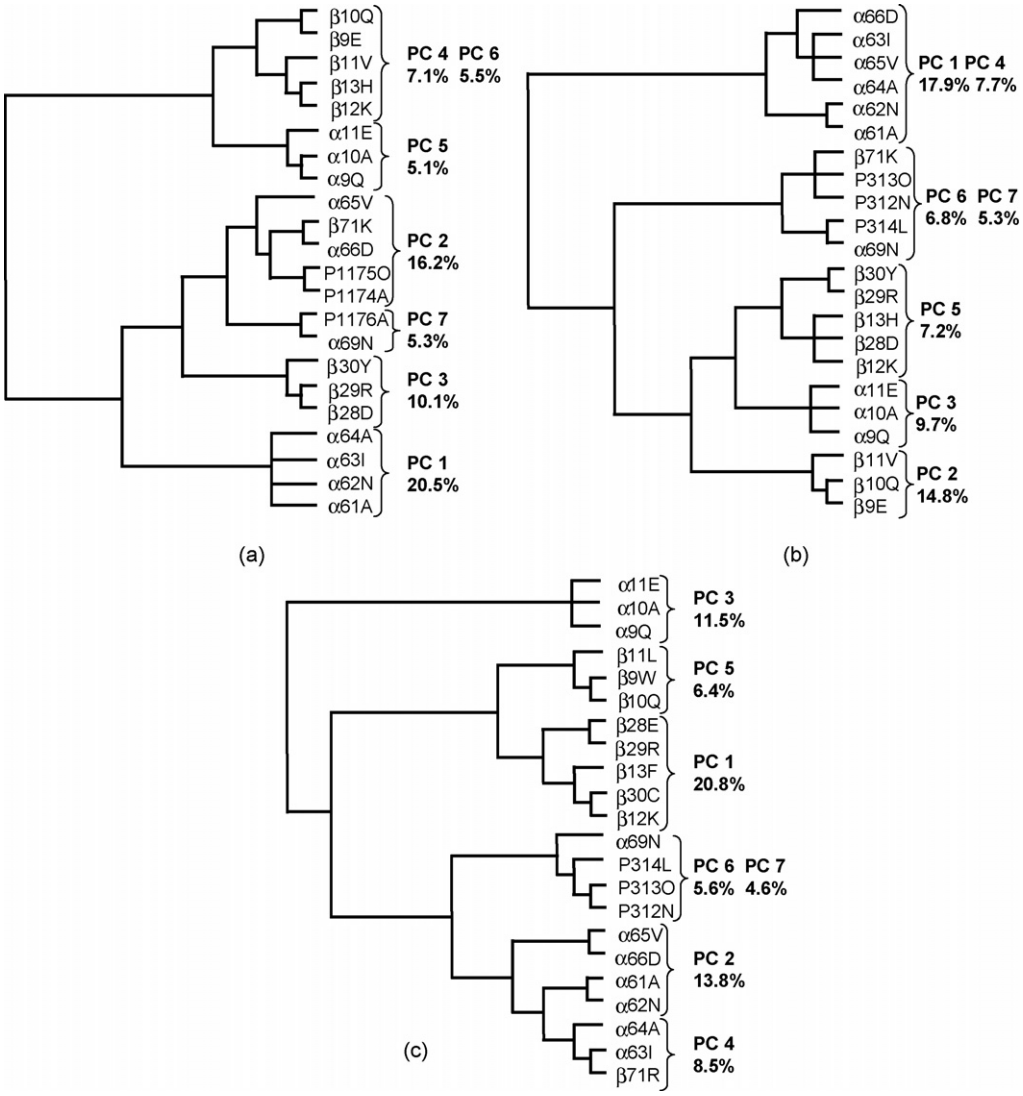

Fig. 6. Dendrograms for Pocket 6, highlighting amino acid cluster groups for the three complexes studied: (a) DR4-Col, (b) DR4-HA and (c) DR1-HA. See Fig. 4 for details.

10.9% for DR4-HA; and 17.2%, 13.4% and 8.9% for DR1-HA.[7]

The occupying peptide does not appear to be largely influencing the pocket–peptide interaction, as the principal components that contain it occur after the PC5 (implying that the level of variance explained by it is low) for the two DR4 complexes studied here. Yet, the DR1-HA complex has a PC2 (explaining a significant amount of the variance) that contains several of the peptide amino acids that flank the occupying amino acid. This indicates that only in the last case does the peptide have an important role in the interaction, with PC2 accounting for 13.4% of the variance. Considering that the Pocket 1 is conserved, the interaction differences between these would reflect the structural differences of crystallized molecules. An exploration of the RMSDs of the heavy atoms for these three structures shows that the two DR4 molecules are similar (0.46 Å) while the RMSD between DR4-HA and DR1-HA is 0.71 Å, highlighting the larger side-chain displacements between the two. The groups formed by these amino acids can be seen in Fig. 4.

3.2. Pocket 4

We have pocket amino acids that appear in the same principal component across the three complexes, but in different positions, and hence the PCs of each of these three cases accounts for a different amount of variance[8]. Specifically, PC3 for DR4-Col corresponds with PC2 for DR4-HA and PC3 for DR1-HA. In this latter case the amino acids with greater factor loadings are β72R, β73A and β74A, with PC variances of 9.1%, 14.1% and 10%, respectively.

[7] The amino acids that are mentioned here and in all pockets' analysis are the most important in each PC taking a cutoff of 0.5 in the factor loading values for each principal component.

[8] The principal components (PC) are sorted in decreasing order, according to the percentage of variance.

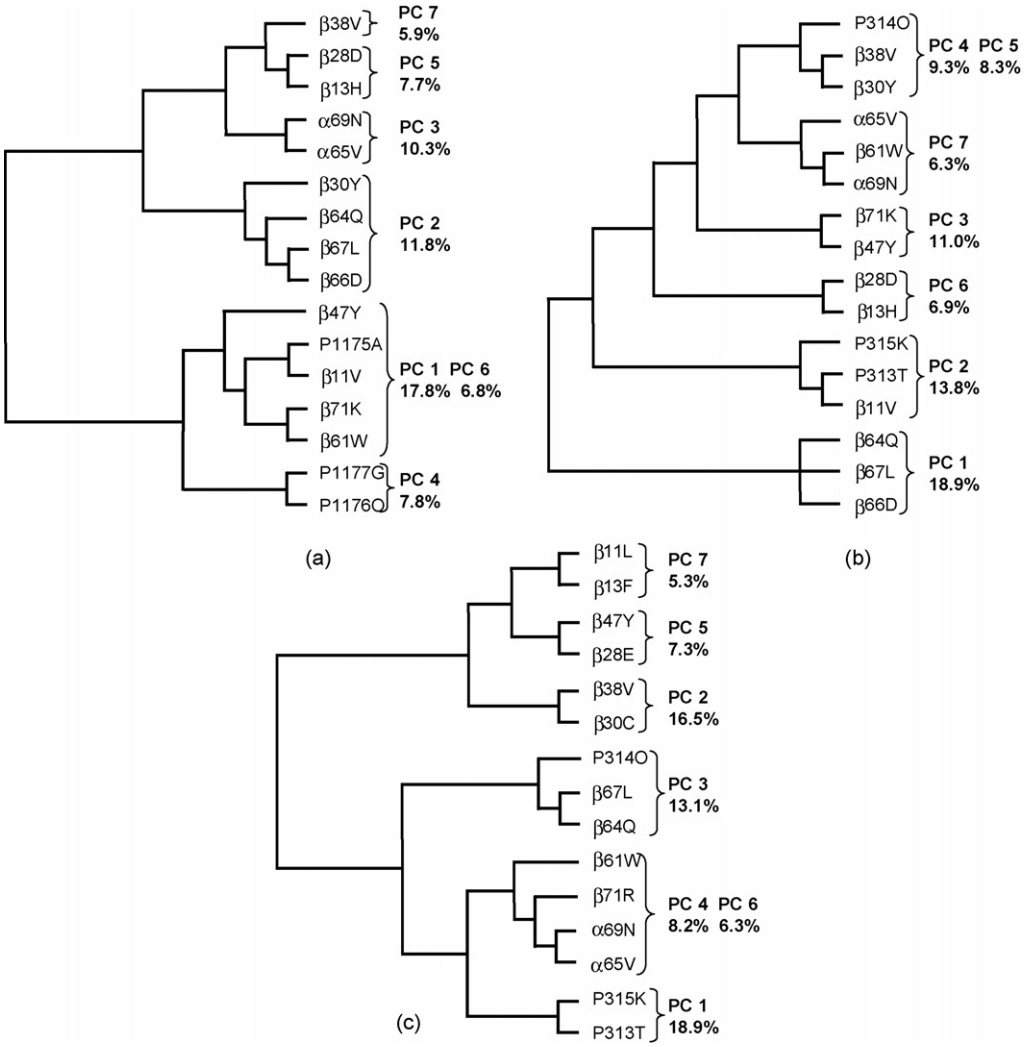

Fig. 7. Dendrograms for Pocket 7, highlighting amino acid cluster groups for the three complexes studied: (a) DR4-Col, (b) DR4-HA and (c) DR1-HA. See Fig. 4 for details.

For principal component PC2 for DR4-Col, PC6 for DR4-HA and PC1 for DR1-HA, pocket amino acids α9Q, α10A and α11E (variances of 12.8%, 5.9% and 21.9%, respectively) all belong to the conserved α chain and are most likely implicated in the pocket–peptide anchoring interaction. Generally, we see that complexes DR4-Col and DR4-HA are more similar in terms of the amino acids that group together in the principal components, indicating that for this pocket the allele effect predominates over the peptide effect. With respect to the occupying peptide, the most correlated pocket amino acids are α9Q, β70Q and β71K, with PC variances of 21.8% for DR4-Col and 20.5% for DR4-HA. The principal component similar in amino acid group to these for DR1-HA is PC5 with only a 6.4% variance level of explanation. These residues are geometrically directed towards the side-chains of the flanking peptide amino acids, which could explain their association with the former amino acids groups. Furthermore, this highlights that the overall pocket behavior is similar to the complexes formed with DR4, emphasizing that there is a predominating allele effect for the pocket.

The principal components for DR1-HA that present mixed pocket and peptide components are below the PC5 level, indicating that, for this complex, the peptide amino acids do not lead the interaction for this pocket. Apart from this, the two first PCs are made up of amino acids that are geometrically pointing towards the peptide base chain (PC1) and towards the side-chain of the occupying amino acid (PC2), emphasizing the anchoring and modulating nature of this pocket. This can be associated with the fact that PC1 explains the greatest amount of variance and is made up of conserved α chain amino acids. The mentioned groups can be seen in Fig. 5.

3.3. Pocket 6

While the three complexes have similar principal components, equivalences among them are not so clear.

While PC2 in DR4-Col contains the occupying amino acid as well as one of the flanking peptide amino acids, in the other two complexes the peptide amino acids appear only at the PC6

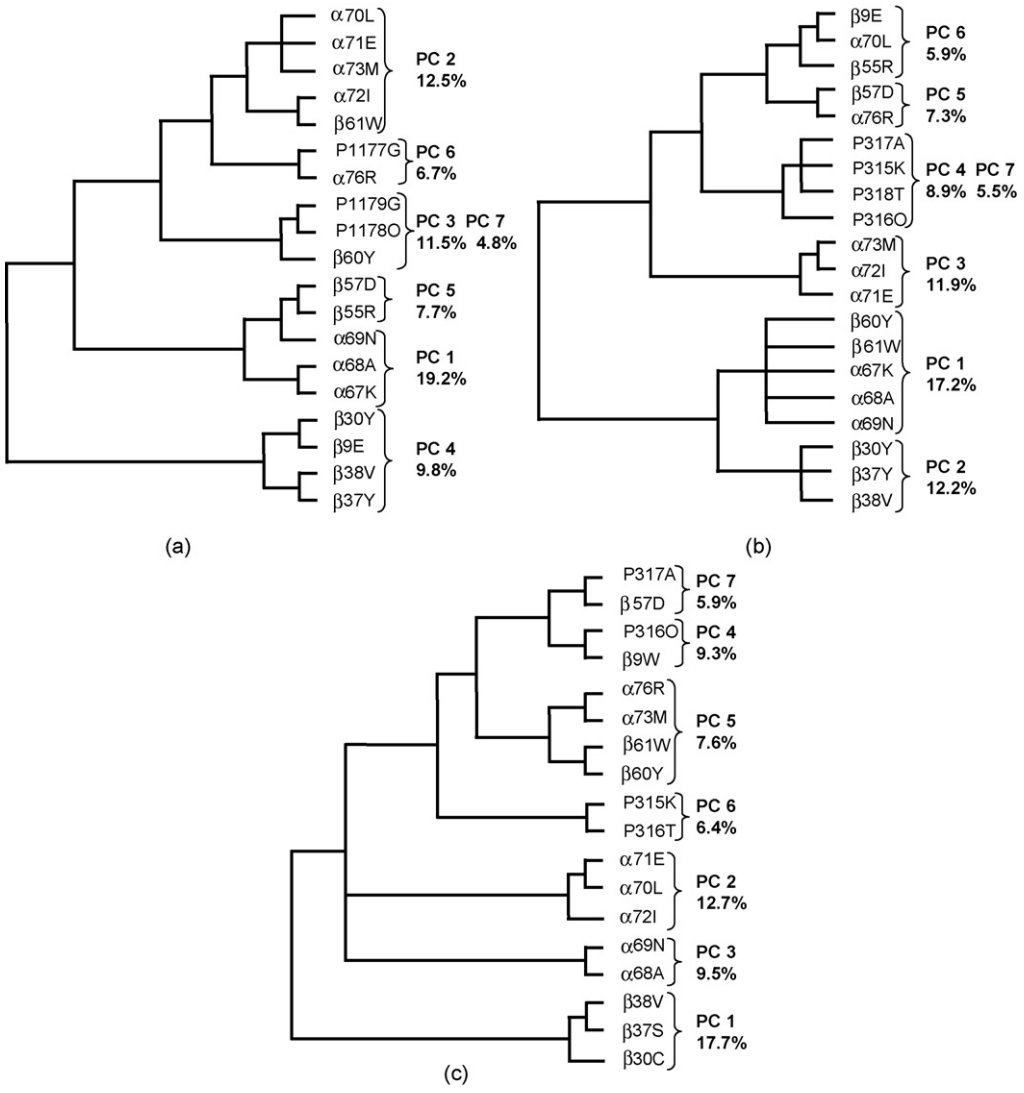

Fig. 8. Dendrograms for Pocket 9, highlighting amino acid cluster groups for the three complexes studied: (a) DR4-Col, (b) DR4-HA and (c) DR1-HA. See Fig. 4 for details.

level, indicating that for this complex the behavior of the pocket depends heavily on the peptide sequence, especially considering that the flanking amino acids are Alanines.

For both DR4 complexes, the first principal component is made up of α61A, α62N, α63I and α64A (with variances of 20.5% for DR4-Col and 17.9% for DR4-HA). Considering that these residues point toward the peptide main chain, together with the fact that they all belong to the conserved α chain, this PC group can be seen to be important in the anchoring of the peptide.

For the DR1-HA complex, PC1 explains 20.8% of the variance and is made up of amino acids that are geometrically found on the ''floor'' of the pocket. Again, this difference with the DR4 complexes highlights the latter predominance for anchoring interactions.

By inspecting this general behavior of the pockets, we can say that interactions in Pocket 6 depend on both allele and peptide. The mentioned groups can be seen in Fig. 6.

3.4. Pocket 7

The three complexes studied here behave differently, according to an analysis of their principal components. There are similar PCs for the two DR4 complexes, but in general no equivalences can be readily established.

One general characteristic is that the occupying peptide appears in the first few PCs in all three cases, indicating that interactions in this pocket are directed by this occupying peptide and hence this pocket is the one that depends most on specificity or modulation for the three cases.

The Lβ11V[9] amino acid appears to be important for the interaction behavior because it appears together with peptide amino acids in the main PC for the three complexes.

[9] When amino acids differ between alleles, the two amino acids correspond to DR1 and DR4, respectively.

Within the DR4-HA complex the interactions with the flanking peptide amino acids appear to be more important than with the occupying amino acid, since this amino acid only appears with a high factor loading value in PC4. Furthermore, it is associated with α65V (with a 0.43 factor loading value) and this amino acid is directed toward the main chain of the peptide. In the other two complexes the occupying amino acid is found in PCs mixed with β67L which is a pocket amino acid located on the pocket floor.

Pockets 6 and 7 present the largest diversity in the composition of their principal components and in their contributions to the different complexes. In both of them the behavior in interaction depend on both allele and peptide, corroborating that these two pockets determine specificity and modulation interactions for the three complexes studied here. The above-mentioned groups can be seen in Fig. 7.

3.5. Pocket 9

The behavior in interactions of this pocket is similar to the one observed for Pocket 4.

In DR4 complexes, PC1 is formed up by amino acids α67K, α68A, α69N, α70L and β61W. This same group of amino acids corresponds to PC3 in DR1 complex. The explained variances for the PCs are 19.2%, 17.2% and 9.5%, respectively. These amino acids are geometrically near to the occupying peptide, a situation which agrees with the fact that they are the amino acids with the largest variations in behavior towards the occupying peptide. This would be an allele dependent effect and the two complexes formed by DR4 allele show a more similar behavior in this respect. Another important principal component for the pocket region is made up of Cβ30Y, β37S and β38V, and is found as PC1 for DR1-HA (variance explained: 17.7%), as PC2 for DR4-HA (variance explained: 12.2%), and as PC4 for DR4-Col (variance explained: 9.8%). These amino acids are found in the base of the pocket region and are principally involved in interactions with the peptide.

The amino acids that form mixed PCs with peptide amino acids are different for the three complexes. According to general behavior patterns, there are some characteristics that seem dependent upon allele effects while others seem to correspond to peptide effects. We also have PCs that are similar across the three complexes but have different levels of explanation or different percentage of variance in the set of PCs.

This pocket apparently presents both anchoring and modulation interaction behaviors: we found amino acids belonging to the conserved α chain that are highly important in the first few PCs. The peptide amino acids are found combined with different pocket amino acids is each of the three complexes, which favors a modulation of the pocket–peptide interaction. The groups addressed here can be seen in Fig. 8.

4. Conclusions

(1) It is possible to establish the nature of the interactions between MHC molecules and antigenic peptides by analysis of the wave function.
(2) The interaction between DR alleles of MHC class II molecules and antigenic peptides is of a second-order kind, not involving the frontier molecular orbitals HOMO and LUMO.
(3) The interaction molecular orbitals (IMOs) for the complexes studied here are located just below the HOMO molecular orbital for the complex. These IMOs contain sufficient information to carry out a fine-grained analysis of the interactions occurring in the MHC pocket regions.
(4) Each one of the pockets studied here (numbered 1, 4, 6, 7 and 9) has its own particular interaction behavior. Specifically, an allele effect or a peptide effect can be identified by the use of pocket or peptide amino acid's factor loadings, in the principal component analysis (PCs), as clustering variables to grouping the amino acids. For each pocket:
   (a) Pocket 1 has an anchoring behavior, very similar for the three complexes studied here. It is a conserved pocket and the main differences obtained in our calculations are caused by structural differences in the pocket between alleles.
   (b) Pocket 4 is governed by an allele effect that shares similarities between the DR4-Col and DR4-HA complexes. Furthermore, this pocket also appears to have both an anchoring and a modulating effect on the interaction with the peptide, since the first few PCs contain residues that are directed towards the peptide's main chain (an anchoring effect), as well as directed in a discriminated manner towards the occupying residue (a modulating effect).
   (c) Pockets 6 and 7 have differing behaviors in the three complexes studied here, making it difficult to generalize about pocket–peptide interactions. Both allele and peptide effects can be seen, as well as modulation or specificity tendencies can found in the analysis of the amino acids participating the first few PCs.
   (d) Pocket 9 behaves in a similar fashion to Pocket 4, by showing an ''equilibrium'' between allele and peptide effects. Correlations between peptide and pocket amino acids are different for the three complexes, indicating a degree a modulation or specificity in the role this pocket has.
(5) All the above annotations clearly show that pocket–peptide interactions are not based on individual amino acid interactions, but rather that the occupying peptide (flanking amino acid 1–occupying amino acid–flanking amino acid 2) interact with almost all pocket amino acids. Moreover, these behaviors can be seen in all pockets, and no pocket acts in isolation from the others (as can be seen with Pockets 6 and 7). This calls into doubt the notion of isolated and independent pocket regions. Perhaps, for the sake of simplifying the analysis process, we are loosing valuable information regarding MHC–peptide interactions. This extra information, in the form of using all the amino acids that comprise the MHC interaction region together with the bound peptide, would help us better understand, and better predict, these complexes behaviors.

(6) The characteristic behavior of each pocket region were also found in our previous work using the three first multipolar moments (charge, dipole and quadrupole). That is, Pocket 1 is an anchoring region; Pockets 6 and 7 are the most variable regions, and depend on allele and peptide effects being responsible for modulating their interactions; and Pockets 4 and 9 have a double purpose, anchoring and modulating the interaction. Additionally the different amino acids found to be important in the interactions were also found in the multipolar study, having a role as global interaction residues or modulating the interaction (differential effect residues). These two different roles are well correlated and the wave function analysis provides a more detailed insight into MHC–peptide interactions.

Acknowledgments

The work carried out here would not have been possible without the support of the Presidency of Colombia and the Colombian Ministry of Health.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2006.04.005.

References

[1] W.E. Paul (Ed.), Fundamental Immunology, Williams & Wilkins, Philadelphia, 1998 (Lippincott Pwachu Awachu).
[2] A. Dessen, C.M. Lawrence, S. Cupo, D.M. Zaller, D.C. Wiley, X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II, Immunity 7 (1997) 473–481.
[3] J. Hennecke, D.C. Wiley, Structure of a complex of the human alpha/beta T cell receptor (TCR) HA 1.7, influenza heagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): insight into TCR cross-restriction and alloreactivity., J. Exp. Med. 195 (2002) 571–581.
[4] L.J. Stern, J.H. Brown, T.S. Jardetzky, J.C. Gorga, R.G. Urban, J.L. Strominger, D.C. Wiley, Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide, Nature 368 (1994) 215–221.
[5] H. Bian, J.F. Reidhaar-Olson, J. Hammer, The use of bioinformatics for identifying class II restricted T-cell epitopes, Methods 29 (2003) 299–309.
[6] I.A. Doytchinova, D.R. Flower, Quantitative approaches to computational vaccinology, Immunol. Cell. Biol. 80 (2002) 270–279.
[7] A. Logean, D. Rognan, Recovery of known T-cell epitopes by computational scanning o a viral genome, J. Comput. Aided Mol. Des. 16 (2002) 229–243.
[8] M. Schirle, T. Weinschenk, S. Stevanovic, Combining computer algorithms with experimental approaches permit the rapid and accurate identification of T cell epitopes from defined antigens, Methods 257 (2001) 1–16.
[9] G.R. Smith, M.J.E. Sternberg, Prediction of protein–protein interactions by docking methods, Curr. Opin. Struct. Biol. 12 (2002) 28–35.
[10] T. Sturniolo, E. Bono, J. Ding, L. Raddrizzani, O. Tuereci, U. Sahin, M. Braxenthaler, F. Gallazi, M.P. Protti, F. Sinigaglia, J. Hammer, Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices, Nat. Biotechnol. 17 (1999) 555–561.
[11] C. Cardenas, J.L. Villaveces, H. Bohorquez, E. Llanos, C. Suarez, M. Obregon, M.E. Patarroyo, Quantum chemical analysis explains hemagglutinin peptide–MHC class II molecule HLA-DRβ1*0101 interactions, Biochem. Biophys. Res. Commun. 323 (2004) 1265–1277.
[12] C. Cardenas, J.L. Villaveces, C. Suarez, M. Obregon, M. Ortiz, M.E. Patarroyo, A comparative study of MHC class II HLA-DRJ31*0401-Col II and HLA-DRβ1*0101-HA complexes: a theoretical point of view, J. Struct. Biol. 149 (2005) 38–52.
[13] C. Cardenas, M. Ortiz, A. Balbin, J.L. Villaveces, M.E. Patarroyo, Allele effects in MHC–peptide interactions: a theoretical analysis of HLA-DRβ1*0101-HA and HLA-DRβ1*0401-HA complexes, Biochem. Biophys. Res. Commun. 330 (2005) 1162–1167.
[14] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, P.E. Bourne, The protein data bank, Nucl. Acid Res. 28 (2000) 235–242.
[15] F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F.J. Meyer, M.D. Brice, J.R. Rogers, O. Kennard, T. Shimanouch, M. Tasumi, The protein data bank: a computer based archival file for macromolecular structures, J. Mol. Biol. 112 (1977) 535–542.
[16] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G.Z.J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, A.N. Kudin, M.C. Strain, O. Farkas, J. Tornasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petterson, P.Y. Ayala, Q. Cui, K.M.D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkarat, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian'98, Revision A.5, Gaussian Inc., Pittsburgh, PA, 1998.
[17] A. Szabo, N.S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, Dover Publications Inc., New York, 1996.
[18] R. McWeeny, Methods of Molecular Quantum Mechanics, Academic Press, New York, 1989.
[19] SPSS Inc., SPSS Base 12.0 for Windows, SPSS Inc., Chicago, IL, 2003.