Journal of Computer-Aided Molecular Design, 11 (1997) 463–478.
KLUWER/ESCOM

J-CAMD 410

Fine specificity of antigen binding to two class I major histocompatibility proteins (B*2705 and B*2703) differing in a single amino acid residue*

Didier Rognan[a,**], Stefan Krebs[a], Oliver Kuonen[a], José R. Lamas[b], José A. López de Castro[b] and Gerd Folkers[a]

[a]*Department of Pharmacy, Swiss Federal Institute of Technology, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[b]*Centro de Biologia Molecular Severo Ochoa, Universidad Autonoma de Madrid, Facultad de Ciencas, E-28049 Madrid, Spain*

Received 10 March 1997
Accepted 26 May 1997

Keywords: MHC; HLA-B27; Drug design; Molecular dynamics simulations

Summary

Starting from the X-ray structure of a class I major histocompatibility complex (MHC)-encoded protein (HLA-B*2705), a naturally presented self-nonapeptide and two synthetic analogues were simulated in the binding groove of two human leukocyte antigen (HLA) alleles (B*2703 and B*2705) differing in a single amino acid residue. After 200 ps molecular dynamics simulations of the solvated HLA–peptide pairs, some molecular properties of the complexes (distances between ligand and protein center of masses, atomic fluctuations, buried versus accessible surface areas, hydrogen-bond frequencies) allow a clear discrimination of potent from weak MHC binders. The binding specificity of the three nonapeptides for the two HLA alleles could be explained by the disruption of one hydrogen-bonding network in the binding pocket of the HLA-B*2705 protein where the single mutation occurs. Rearrangements of interactions in the B pocket, which binds the side chain of peptidic residue 2, and a weakening of interactions involving the C-terminal end of the peptide also took place. In addition, extension of the peptide backbone using a β-Ala analogue did not abolish binding to any of the two HLA-B27 subtypes, but increased the selectivity for B*2703, as expected from the larger peptide binding groove in this subtype. A better understanding of the atomic details involved in peptide selection by closely related HLA alleles is of crucial importance for unraveling the molecular features linking particular HLA alleles to autoimmune diseases, and for the identification of antigenic peptides triggering such pathologies.

Introduction

Major histocompatibility complex (MHC)-encoded class I proteins play a major role in the immune surveillance of intracellular pathogens by presenting antigenic peptides to cytotoxic T-lymphocytes (CTLs) at the surface of infected cells [1]. In the last decade, tremendous research efforts have been made to delineate the molecular aspects of antigen presentation to CTLs. Three major breakthroughs in this field were the description of the first class I MHC crystal structure [2], the identification of allele-specific motifs for naturally bound peptides [3], and the prominent role played by MHC-encoded TAP heterodimers for antigen transport [4]. Starting from these three major observations, experimental data on antigen presentation to MHC proteins are presently accumulating at an incredible pace. The huge amount of molecular details have made human leukocyte antigen (HLA) proteins very interesting drug design targets for two reasons. First, it is likely that the three-dimensional (3D) structures of all class I and class II alleles are very similar, and that the MHC binding mode is rather conserved for most of the presented peptides, within each HLA class [5,6]. Reliable 3D pictures of MHC–peptide complexes are now accessible and can be interpreted with respect to biological data [7–10]. Second, the expression of certain MHC alleles is associated with either resistance or susceptibility to human immunological diseases like ankylosing spondylitis [11], diabetes [12], rheumatoid arthritis [13] or malaria [14].

One of the strongest linkages known to date between

*Dedicated to Prof. D. Seebach on the occasion of his 60th birthday.
**To whom correspondence should be addressed.

the expression of one class I HLA allele and susceptibility to a pathology is that of HLA-B27 to inflammatory diseases of the joints called spondyloarthropathies [15,16]. For example, more than 95% of patients suffering from ankylosing spondylitis are HLA-B27 positive, while this MHC type is only expressed in 7% of the healthy population [11]. Interestingly, among the 11 HLA-B27 alleles (HLA-B*2701 to HLA-B*2711) reported to date, at least two are not associated with spondyloarthropathies: HLA-B*2706 [17,18] and HLA-B*2709 [19]. All subtypes differ among each other only at a few residues, mostly located in the peptide binding groove.

Deciphering the molecular parameters responsible for the binding of antigenic peptides to the different HLA-B27 subtypes is an absolute prerequisite for better understanding their differential association with susceptibility to spondyloarthropathy, and to identify the sequence of potential arthritogenic peptides [15] that may trigger the disease. Sequencing of peptides naturally presented by the different alleles [20–26] shows that Arg at position 2 (P2) is a conserved motif for peptides binding to all subtypes. Gln^{2} is an additional motif among B*2701-bound peptides [25], but is a suboptimal residue (not found in vivo) for other subtypes [27–30]. Two other anchoring positions could be disclosed from the frequency of occurrence of amino acids in peptides naturally bound to B*2705. Hydrophobic amino acids are preferred at P3, and hydrophobic as well as positively charged residues can be found at P9. Other positions are more variable and probably indicate a less important binding role. These observations are compatible with the crystal structure of one allele (HLA-B*2705) in complex with a peptide pool [31,32]. As for all other class I alleles crystallized to date, bound ligands are mainly nonapeptides strongly hydrogen-bonded in a sequence-independent manner at their N- and C-termini to both ends of the MHC binding groove. The central part of the peptide (from P4 to P8) bulges out of the binding groove (Fig. 1) [32]. Some peptide side chains (P2, P3 and P9) are responsible for the allele specificity of the recognition process, by binding to complementary pockets of the MHC. Notably, the conserved Arg^{2} of B27-bound peptides is perfectly centered in a polar subsite (pocket B) composed of MHC polymorphic residues (Tyr^{7}, His^{9}, Thr^{24}, Glu^{45}, Cys^{67}). HLA-B*2705 is the only B27 subtype that has been crystallized up to now. Furthermore, although peptide binding motifs have been described for various HLA-B27 subtypes, the exact peptide repertoire selected by each allele is still under investigation, and the exact contribution of individual amino acids at the nine peptide positions to the various subtypes remains largely unknown.

HLA-B*2703 contains a single point mutation (Tyr^{59}→His) in a subsite (pocket A) responsible for the binding of the peptide N-terminal residue. HLA-B*2703 probably selects a subset of peptides presented by the common HLA-B*2705 allele [33,34]. Positively charged residues (Lys, Arg and His) at P1 have been proposed to be one of the main characteristics of this subset common to both subtypes as the replacement of Ser^{1} by Arg restored the binding of a B*2705-restricted viral epitope to HLA-B*2703 [35]. In addition, basic residues were predominant among self-peptides naturally bound to B*2703 [23]. Additional interactions given by basic side chains could compensate for the weaker interaction of the peptide N-terminus to B*2703 pocket A. However, this is not a stringent requirement because amino acids other than Arg/Lys at P1 (Ala for example) are compatible with a good binding to this allele [28].

To rationalize these data, molecular dynamics (MD) simulations of both alleles in complex with three different peptides were undertaken to delineate similarities and differences in MHC binding, which may explain specificity variations.

Materials and Methods

Coordinates setup

Starting coordinates were taken from the crystal structure of HLA-B*2705, solved at 2.1 Å resolution [32] and deposited in the Brookhaven Protein Databank [36] with the entry 1HSA. In order to save computational time, only the antigen-binding α1–α2 domains were taken into account in the study. This approximation was previously shown not to alter the accuracy of MD simulations [9,37,38] because only limited interactions exist between the α1-α2 part and the other two domains (α3 and β_2m) that do not significantly contact antigenic peptides in the binding groove. This observation has been validated by the crystal structure of one class I MHC protein lacking the membrane-proximal α3 domain, for which a conserved 3D fold of the α1–α2 antigen-binding domain was reported [39]. The C-terminal residue of the α2 domain (Thr^{182}) was protected by an N-methyl group to avoid unrealistic electrostatic interactions. The HLA-B*2703 subtype was obtained from the HLA-B*2705 crystal structure by mutating Tyr^{59} into His, without changing the direction of the side chain. All MHC-bound nonamers were built from the peptide (ARAAAAAAA) modeled in the original crystal structure [32] by substituting the corresponding residue for alanine without altering the direction of the side chains. Six crystal water molecules were explicitly taken into account, as they are located in the peptide binding cleft and bridge the binding of the peptide to the protein X-ray structure. Polar hydrogen atoms were then added and the complexes were centered in a 7.5 Å thick shell of TIP3P water molecules [40] without positional constraint on solvent atoms. Any water atom closer than 1.75 Å to any solute atom was discarded, so that approximately 1500 water molecules were added to each MHC–peptide binary complex.

Parametrization of the N-terminal β-alanine monomer

The N-terminal β-alanine was parametrized for AMBER 4.0 [41] using a previously described procedure [42]. Briefly, atomic coordinates for β-alanine were obtained from the SYBYL biopolymer dictionary [43] and optimized by semiempirical quantum mechanics (MOPAC 6.0) using the PM3 Hamiltonian [44]. Potential-derived atomic charges were then computed on this geometry by a single-point SCF calculation at the MNDO level [45]. As new atom types were not defined, nonexisting force constants have been assigned according to available AMBER values for closely related bond, angle and dihedral types. They were incorporated into the parm91 parameter set.

Molecular mechanics and dynamics simulations

All computations were performed on a CRAY J90 using the AMBER 4.0 program [41] and the united-atom representation of the parm91 parameter set. Since explicit water molecules were taken into account, a dielectric constant of 1 was used for all calculations. To avoid splitting dipoles, nonbonded interactions were calculated within a residue-based cutoff of 10 Å. The solvent atoms were first relaxed by 1000 steps of steepest descent energy minimization, the solute being held fixed. The solvated complex was then fully minimized by 1000 steps of steepest descent, followed by a conjugate gradient minimization procedure until the rms gradient of the potential energy was less than 0.25 kcal mol^{-1} $Å^{-1}$. The minimized coordinates were thereafter used as a starting point for an MD simulation at constant temperature. Initial velocities were taken from a Maxwellian distribution at 50 K and an integration step of 2 fs was used. The system was progressively heated from 50 to 297 K during the first picosecond, the temperature being held at 297 K for the rest of the simulation by coupling the system to a heat bath [46] using a temperature coupling constant of 0.05 ps. All bond lengths were constrained to their equilibrium values using the SHAKE algorithm [47] with a bond length tolerance of 2.5×10^{-4} Å. Coordinates, energies and velocities were collected and saved every 250 steps (0.5 ps) for 200 ps. The analyses of MD trajectories were achieved using in-house routines and the CARNAL module of AMBER [41].

Calculation of electrostatic interaction energies

Electrostatic free energies were computed by solving the linear form of the Poisson–Boltzmann equation using the finite-difference method [48,49] of the DelPhi program [50,51]. Peptides, MHC proteins and MHC–peptide complexes were centered in three-dimensional boxes with resolutions of 2.0, 1.20 and 1.10 grid points per Å, respectively. For each calculation, 90% of the box was filled with the corresponding molecule. Atomic radii and charges were taken from the AMBER 4.0 united-atom parameter set [41]. Inner and outer dielectrics were assigned values of 2.0 and 1.0 (vacuum) or 80 (water environment). An ionic strength of 0.145 M and an ion exclusion radius (Stern layer) of 2.0 Å were used according to previously reported solvent calculations [51]. A probe radius of 1.8 Å was utilized for computing the surface at which the electrostatic potential was extrapolated.

Peptide synthesis

Peptides **1** and **3** were synthesized as previously described [28]. Peptides **2** and **4** were synthesized by automated, multiple solid-phase peptide synthesis with a robot system (Syro, MultiSynTech, Bochum, Germany) using an Fmoc/*t*Bu strategy. For side-chain protection, Tyr(*tert*-butyl), Ser(*tert*-butyl), Thr(*tert*-butyl), Glu(*tert*-butyl), Gln(trityl), Arg(2,2,3,5,5-pentamethyl-chromansulfonyl) and Lys(*tert*-butyloxy-carbonyl) were used. The N-terminal residues were obtained using single couplings with diisopropylcarbodiimide/1-hydroxy-benzotriazole activation, 10-fold excess and a coupling time of 1 h on the 2-chlorotritylchloride resin. The peptides were cleaved with trifluoroacetic acid/thioanisole/thiocresol (20:1:1) within 3 h, collected by centrifugation and lyophilized from water.

They were then purified by reversed-phase HPLC (Merck-Hitachi, Darmstadt, Germany) on a nucleosil 5 μM/C_{18} column (125×3 mm) at a flow rate of 600 μl/min. The absorbance was measured at 220 nm. The solvent system used consisted of 0.1% trifluoroacetic acid in water (A) and 0.1% trifluoroacetic acid in acetonitrile (B). A linear gradient from 10 to 60% B in 30 min was applied. The peptides were purified to homogeneity by a second HPLC on a versapack 10 μm C_{18} column (300× 7.8 mm) at a flow rate of 2 ml/min with the same buffer system, and a linear gradient from 0 to 40% B in 35 min followed by a 40–60% B linear gradient for 20 min. Furthermore, the peptides were analyzed by ion spray mass spectrometry on a triple-quadrupole mass spectrometer API III with a mass range of m/z = 10–2400 equipped with an ion spray interface (Sciex, Thornhill, ON, Canada), and quantified by amino acid analysis using a 6300 amino acid analyser (Beckman, Palo Alto, CA, U.S.A.).

Peptide binding assay

The quantitative assay used has been described previously [29]. Briefly, RMA-S transfectants expressing B*2705 or B*2703 were used. These are murine cells with impaired TAP-mediated peptide transport and low surface expression of (empty) class I MHC molecules, which can be induced at 26 °C [52] and stabilized at the cell surface through the binding of exogenously added peptides. These cells were incubated at 26 °C for 24 h. After this, they were incubated for 1 h at 26 °C with 10^{-4}–10^{-9} M peptides, transferred to 37 °C and collected for flow

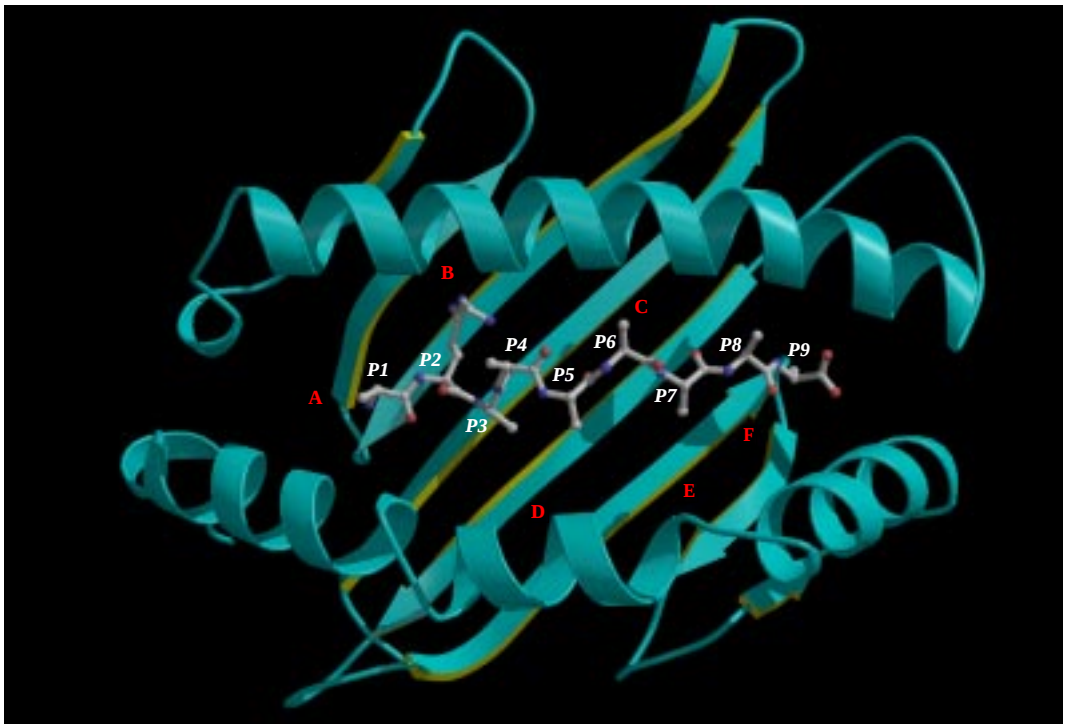

Fig. 1. Orientation of a canonical nonapeptide (ball-and-stick model) in the binding groove of HLA-B27 (cyan ribbons) [32]. Peptide positions are labeled at the C^{α} atoms from 1 (P1) to 9 (P9). MHC specificity pockets (A to F) are displayed according to the usual nomenclature [72]. The figure has been obtained with the MOLSCRIPT [73] and Raster3D [74] programs.

microcytometry (FMC) analysis with the ME1 mAb (IgG1, specific for HLA-B27, B7 and B22) [53] after 4 h for B*2705, or after 2 h for B*2703. The determinant recognized by ME1 is not affected by bound peptides or by HLA-B27 polymorphism (data not shown). The binding of a given peptide was measured as its C_{50}. This is its molar concentration at 50% of the fluorescence obtained with that peptide at 10^{-4} M. Peptides with $C_{50} \leq 5$ μM were considered to bind with high affinity, as these were the values obtained for most of the natural B27-bound ligands. C_{50} values between 5 and 50 μM were considered to reflect intermediate affinity. $C_{50} \geq 50$ μM indicated low affinity. Peptides having C_{50}>100 μM were assumed not to bind. The binding of peptide analogues was measured as the concentration of the peptide analogue required to obtain the fluorescence value at the C_{50} of the unchanged peptide. This was designated as EC_{50}. Relative binding was expressed as the ratio between the EC_{50} of the peptide analogue and the C_{50} of the corresponding unchanged peptide.

Results and Discussion

A naturally bound nonapeptide (**1**) from the human histone H3 protein was taken as reference for its equal binding to both subtypes (Table 1). The Ala^1 analogue (**2**) was chosen for its unexpected high affinity for both subtypes [28,35]. Finally, the Gln^2 analogue **3** of the human histone peptide was also studied, in order to investigate the influence of pocket B–P2 interactions that may also participate in the peptide discrimination [28].

Dynamical properties of MHC–peptide complexes

Monitoring instantaneous rms deviations (rmsd's) of protein atoms from the starting structure is usually performed to ascertain the reliability of MD simulations [54]. For both alleles complexed to peptides **1**–**3**, similar rmsd values were observed (about 1.7 Å for backbone atoms;

TABLE 1
RELATIVE BINDING OF THREE NONAPEPTIDES TO TWO HLA-B27 SUBTYPES

Peptide	Sequence	Relative binding[a]	
		B*2703	B*2705
1	R**R**YQKSTE**L**[b]	1 (2×10^{-6})	1 (2×10^{-6})
2	A**R**YQKSTE**L**	1.5	2
3	R**Q**YQKSTE**L**	>>100	10

[a] Data are expressed as the molar excess of peptide analogue, relative to the wild-type peptide **1**, at which HLA-B27 fluorescence (measured by FMC analysis with an anti-B27 monoclonal antibody) on RMA-S cells was half the maximum obtained with peptide **1**. The molar concentration of peptide **1** at 50% maximum fluorescence (C_{50}) is given in parentheses.

[b] Human histone H3 peptide: a self-peptide, naturally bound to HLA-B*2705 [20] and to B*2703 [23]. Dominant anchor residues are displayed in boldface.

data not shown), indicating that protein distortion upon ligand binding cannot account for the different binding affinities of these peptides. Stable rmsd values (1–1.5 Å) were also observed for bound peptides. Interestingly, they remain in the same range as that observed for different peptides co-crystallized with the same MHC host protein [55,56]. One exception concerns the weakest binder (peptide **3** to HLA-B*2703), for which larger distortions (up to 2.5 Å) were observed and still increasing after 200 ps simulation.

To follow a possible dissociation of peptides **1**–**3** from the two HLA-B27 alleles, the distance between ligand and protein center of mass (cmass) was monitored. For binding peptides (**1**, **2**), whatever the B27 subtype, this intermolecular distance increases during the warm-up phase from 8 to 8.5 Å and remains constant for the rest of the simulation (Fig. 2). The weak binder (peptide **3**) exhibits a slight but continuous increase of the intermolecular distance, suggesting a partial dissociation of the ligand from the binding groove. This phenomenon was significantly enhanced for the less stable complex (**3a**: peptide **3** in complex with HLA-B*2703). However, localization of the dissociating peptide amino acids is not possible with this analysis. For example, TcR-binding residues may bulge even more out of the peptide binding site and induce similar shifts in the intermolecular distance between center of masses.

Therefore, this analysis was extended by computing the cmass of bound peptide substructures (MHC anchors: P1, P2, P3, P9; TcR anchors: P4, P5, P6, P7, P8) for each MD conformation and the distance relating it to the protein cmass (Table 2). An examination of d2 and d3 intermolecular distances for the weak binder **3** clearly shows that MHC-anchoring amino acids only are progressively expelled from the binding groove. The dissociation is more significant when compound **3** is complexed to HLA-B*2703, which relates well with the observed binding data (Table 1). The TcR-binding region (P4–P8) is similarly bulging out (at least quantitatively) of the binding cleft for all the six studied complexes (d3 distance, Table 2). An even more precise analysis has been done by evaluating the inter-cmass distance between individual MHC anchor residues and their complementary pocket (A, B, D and F; d4–d7 distances, Table 2). Surprisingly, the repulsion noticed for peptide **3** in complex with both alleles seems to be located at the P9–pocket F interaction level (d7 = 7.3 Å), far from where protein and peptide single-point mutations occur. Interestingly, the critical d5 distance, illustrating the strength of the most important interaction between the invariant Arg^2 and pocket B, is higher (5.1 ± 0.4 Å) for the less stable complex (**3a**) than for other pairs (Table 2). The different d4 distances reported here, especially for peptides **1** and **3** and peptide **2**, are related to the size of the corresponding P1 side chain (Arg versus Ala). It indicates that the Arg^1 side chain is pointing away from its binding subsite (pocket A) and therefore induces higher d4 inter-cmass distances. However, for a given P1 side chain, higher d4 distances are always observed for HLA-B*2703. This indicates that the Tyr^{59}→His mutation found in the latter allele is detrimental for a stable and strong binding of P1 side chains to pocket A of the protein. The highest stan-

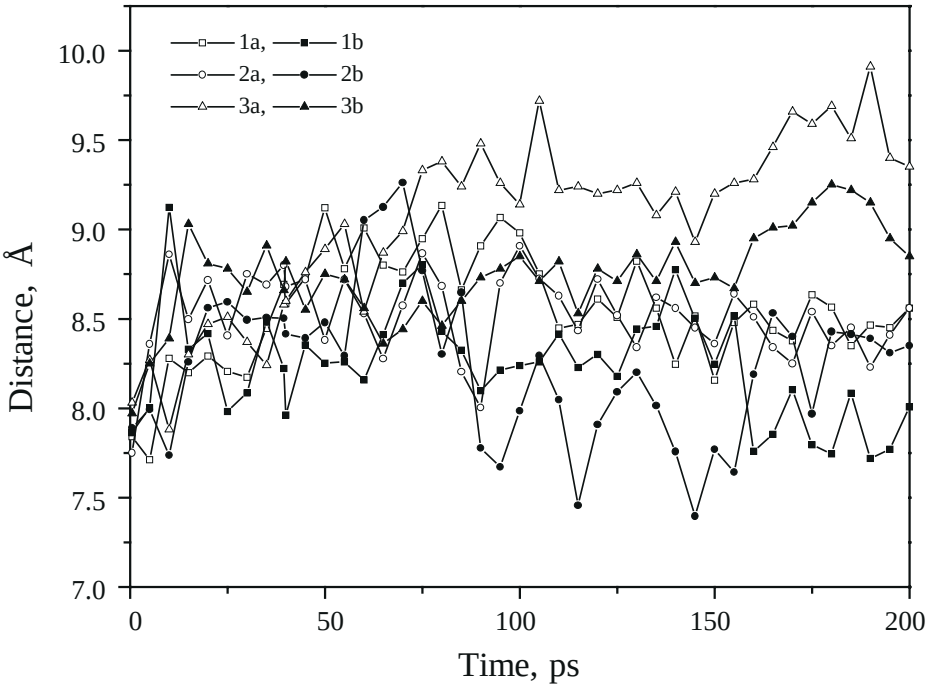

Fig. 2. Time course of the distance (in Å) between the center of mass of peptides **1**–**3** and that of the host HLA-B27 subtype (**a**: B*2703; **b**: B*2705).

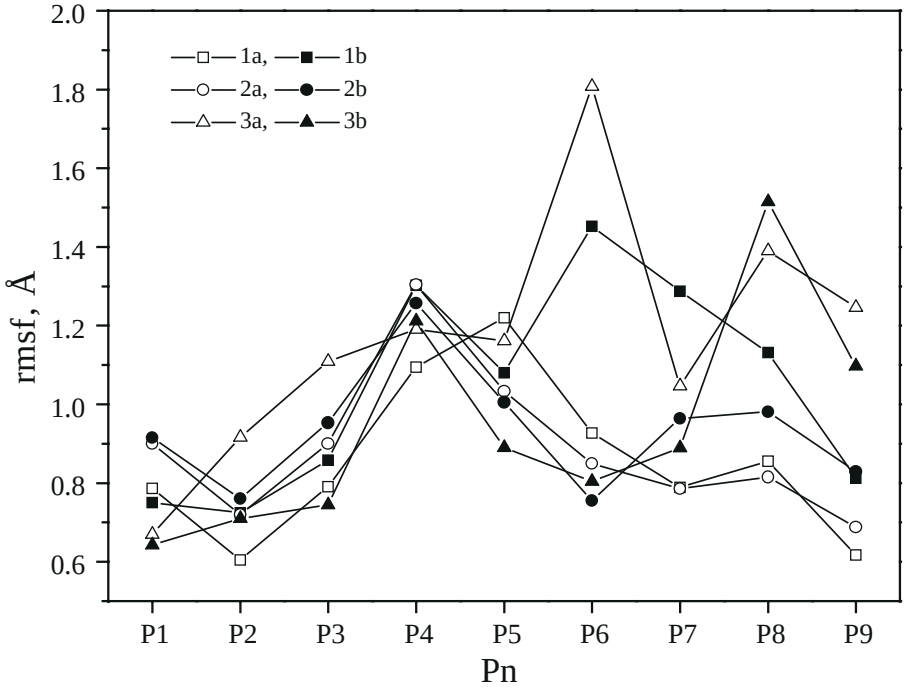

Fig. 3. Rms atomic fluctuations of peptides **1**–**3** when bound to an HLA-B27 protein (**a**: B*2703; **b**: B*2705). Pn represents the peptide position (from 1 to 9).

dard deviations are found for the P3–pocket D interaction (d6, Table 2). The 0.6 Å variations, observed in such nonbonded distances, correspond to the alternative establishment of strong and weak hydrophobic contacts, which may be explained by the topology of the binding groove. Pocket D is a hydrophobic subsite open to the central part of the binding cleft, and partially filled by hydrophobic side chains [20]. The significantly higher variance of the d6 distance (Table 2) suggests that this interaction is the most flexible one and that different local conformations at P3 are compatible with a good occupancy of pocket D. Retrospectively, it explains why the P3–pocket D interaction may be important for an optimal peptide binding to HLA-B*2705. Bulky hydrophobic nonnatural side chains (α- and β-naphthylalanine, cyclohexylalanine) have been shown to significantly enhance binding to the HLA-B*2705 protein [38,57]. This positive effect probably results from two correlated components: (i) the existence of additional contacts to pocket D, a pure enthalpic effect; and (ii) a reduced flexibility of the MHC-bound P3 side chain, a favorable entropic effect.

Atomic fluctuations

Atomic fluctuations of MHC-bound peptides were computed from mean conformations, time-averaged over the last 50 ps (Fig. 3). As expected, MHC anchors are much less flexible than the TcR-binding middle part. If atomic mobility of the bound ligand is considered, it is not possible to depict real differences in the binding of peptides **1** and **2** to the two subtypes. However, the C-terminal anchor residue clearly tends to be more flexible when the corresponding ligand does not strongly bind to the B27

TABLE 2
DISTANCE BETWEEN PROTEIN AND PEPTIDE CENTER OF MASSES

Distance (Å)	Peptide					
	1a	**1b**	**2a**	**2b**	**3a**	**3b**
d1	8.2 ± 0.3	8.5 ± 0.3	8.3 ± 0.3	8.2 ± 0.4	9.3 ± 0.9	9.1 ± 0.7
d2	5.9 ± 0.3	5.9 ± 0.3	5.5 ± 0.3	5.6 ± 0.4	7.3 ± 0.5	6.6 ± 0.5
d3	11.5 ± 0.4	12.1 ± 0.5	12.3 ± 0.5	10.7 ± 0.5	12.0 ± 0.6	12.3 ± 0.6
d4	3.6 ± 0.3	2.9 ± 0.4	2.3 ± 0.3	1.5 ± 0.5	3.4 ± 0.4	2.8 ± 0.3
d5	4.7 ± 0.2	4.7 ± 0.3	4.7 ± 0.3	4.4 ± 0.3	5.1 ± 0.4	4.7 ± 0.3
d6	4.9 ± 0.6	4.8 ± 0.5	5.3 ± 0.5	5.0 ± 0.7	4.6 ± 0.5	5.2 ± 0.6
d7	2.5 ± 0.4	2.6 ± 0.3	2.9 ± 0.3	3.4 ± 0.4	7.3 ± 0.6	4.8 ± 0.5

d1: protein–peptide; d2: protein–MHC anchors (P1–P3, P9); d3: protein–TcR anchors (P4–P8); d4: pocket A–P1; d5: pocket B–P2; d6: pocket D–P3; d7: pocket F–P9.

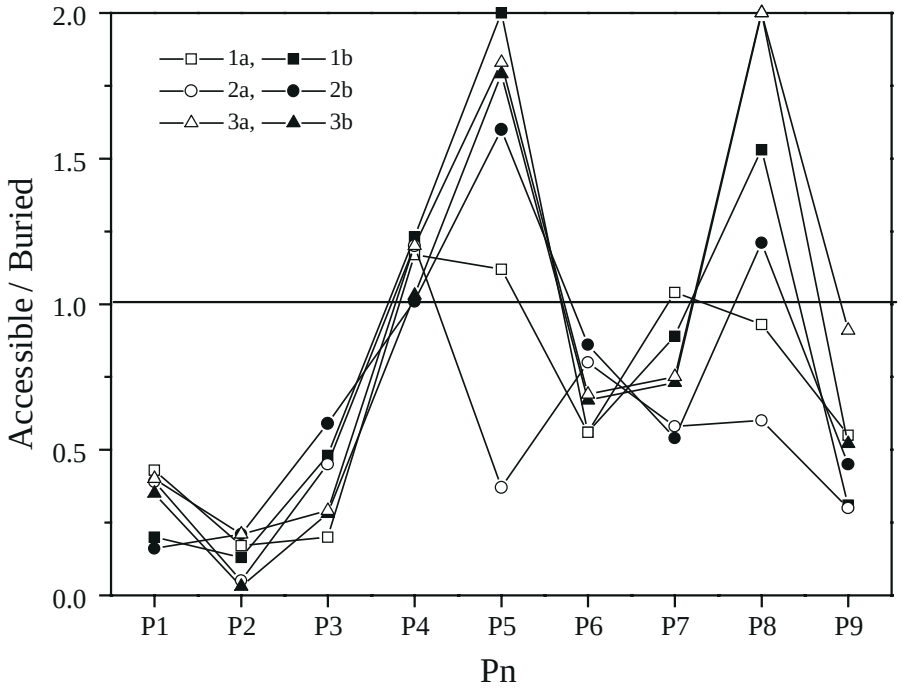

Fig. 4. Ratio of accessible to buried surface area for relaxed MD time-averaged conformations. Surface areas were calculated with the MS program [75] using a 1.4 Å probe radius. High ratios were truncated to a value of 2.0. Pn represents the peptide position (from 1 to 9).

subtype. The highest flexibilities are observed for the weakest binder (peptide **3** to B*2703), especially at the important anchoring positions (P2, P3 and P9). It is logical to find that the mutation at P2 (for peptide **3**) weakens the interactions to MHC pocket B that has been designed to accommodate an arginine side chain [32]. However, the P2 amino acid is more flexible when the corresponding peptide is complexed to HLA-B*2703

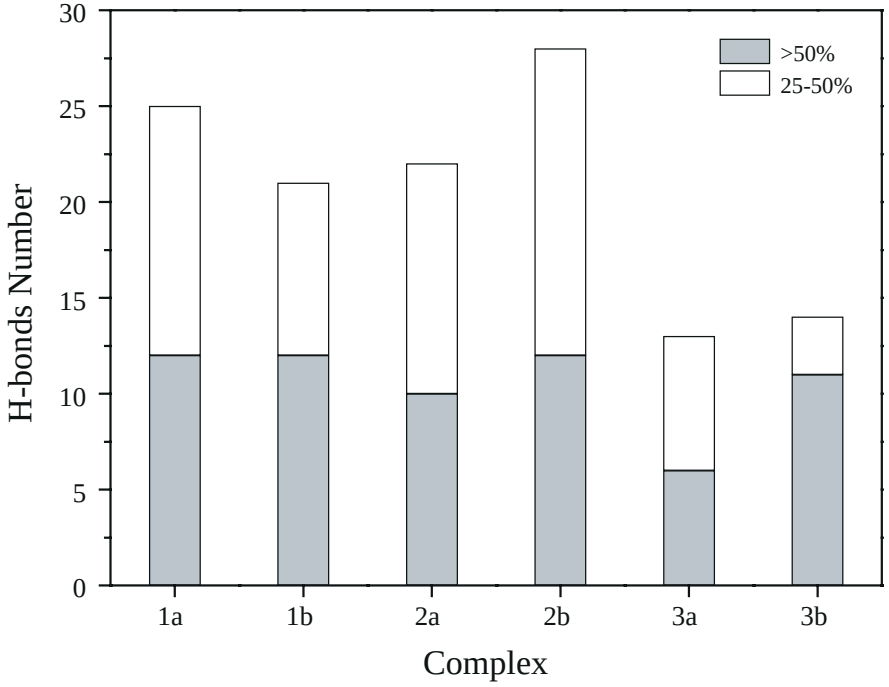

Fig. 5. MHC–peptide H-bonding frequency for peptides **1**–**3** in complex with HLA-B27 subtype (**a**:B*2703; **b**: B*2705). H-bonds have been geometrically defined by an acceptor (A) to donor (D) distance less than 3.25 Å and a D-H..A angle greater than 120°. Interactions were statistically monitored throughout the simulations for a total of 400 conformations per MHC–peptide complex. Two categories of H-bonds were defined: strong ones with frequencies higher than 50% and medium ones with occurrences between 25 and 50%.

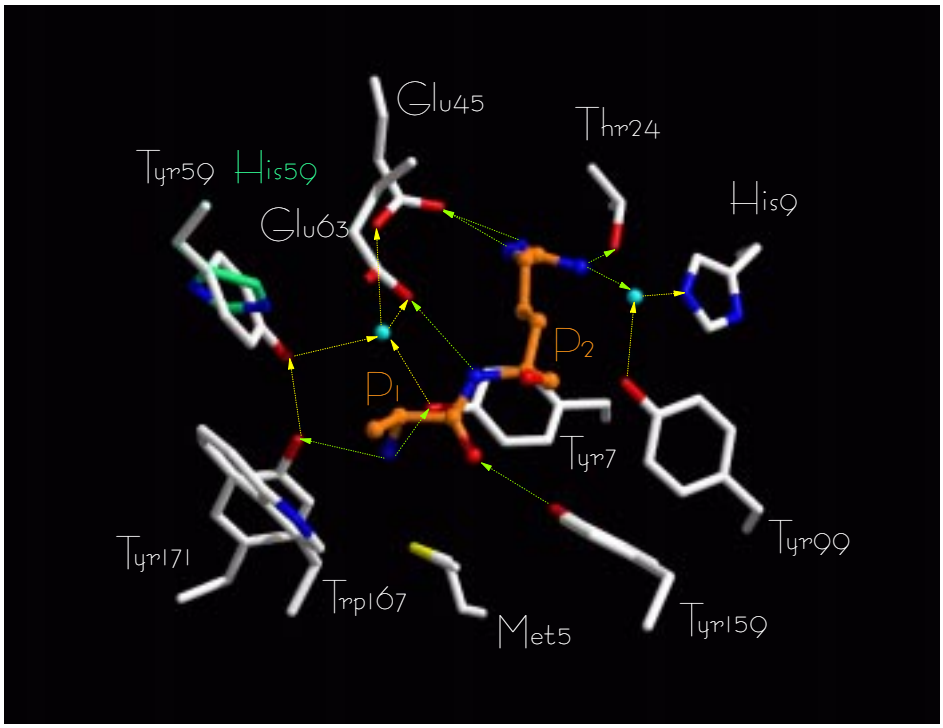

Fig. 6. Crystal structure of HLA-B*2705 [32]. The view is focused on MHC pocket A (Met^{5}, Tyr^{7}, Tyr^{59}, Glu^{63}, Tyr^{159}, Trp^{167}, Tyr^{171}) and pocket B (His^{9}, Thr^{24}, Glu^{45}, Tyr^{99}) side chains interacting via H-bonds (direct H-bonds: green broken lines; water-mediated H-bonds: yellow broken lines) with the P1–P2 positions of a bound peptide. The following color coding has been used: carbon, white (protein) or orange (peptide); nitrogen, blue; oxygen, red; sulfur, yellow. Bound water molecules are shown as cyan balls. Arrows indicate the direction of the H-bonds (from the donor to the acceptor). HLA-B*2703 was obtained by mutating Tyr^{59} into His (SYBYL Biopolymer module) [43]. The side-chain χ^2 dihedral was just modified here in order to bring the N$^{\varepsilon}$ atom as close as possible to the bound water molecule. However, whatever the rotamer chosen, a direct H-bond to a water molecule is not possible. The interatomic distance between the peptide N-terminus and Tyr^{59} (OH) or His^{59} (N$^{\varepsilon 2}$) atoms is 4.31 and 5.35 Å, respectively. Figures 6–8 have been obtained by using the rendering program Raster 3D [74].

(compare **3a** and **3b**, Fig. 3). Again, the most striking differences are not observed at the variable amino acids of the peptides (P1, P2) but at the C-terminal anchor (P9), a feature already noticed for other MHC–peptide complexes [9]. One may hypothesize that the N-terminal tripeptide (P1-P2-P3) determines the stability of peptide–MHC interactions over the whole length of the binding groove by controlling the conformational space accessible to the bulging middle part and, consequently, the binding capacity of the following C-terminus. However, atomic flexibilities of the P1–P3 and P4–P8 parts are not interrelated. It is also possible that these differences are related to the short time scales (200 ps) used for simulating the complexes and that much longer simulations are needed to see significant molecular differences at positions P1 and P2 of the peptide ligand. Nanosecond MD simulations of macromolecules are nowadays feasible [58,59], but still remain unrealistic as a structure–activity relationship tool for comparing a series of ligands in their protein-bound state.

Accessible versus buried surface areas

Whether peptide flexibility correlates with dissociation from the binding cleft was addressed by looking at accessible and buried surface areas of each HLA-bound peptide residue (Fig. 4). Only position 9 of peptide **3** in complex with B*2703 was much more accessible than the others. Otherwise, the main anchor P2 was similarly buried whatever the peptide and the host HLA protein. This means that a partial dissociation was only observed for one position (P9) in one complex (**3a**) and that the previously reported higher flexibility of Gln^{2} for the same peptide–MHC complex did not correspond with a release of the P2 side chain from pocket B. It may however influence the quality and frequency of hydrogen bonds between the Gln^{2} side chain and pocket B of both B27 alleles. The influence of secondary anchor positions is more difficult to ascertain. P3 is similarly buried for all MHC–peptide complexes. P6 (Ser) and P7 (Thr) positions are probably accessory anchor positions that marginally bind, in some conformations, to the central part of the peptide binding groove (pockets C/E). This observation is not incompatible with the high atomic fluctuations of P6-P7 amino acids (Fig. 3), as their side chains are directed towards the binding groove but without reaching its floor. P4 (Gln), P5 (Lys) and P8 (Glu) residues are potential candidates for TcR recognition because of their concomitant atomic flexibility and surface accessibility.

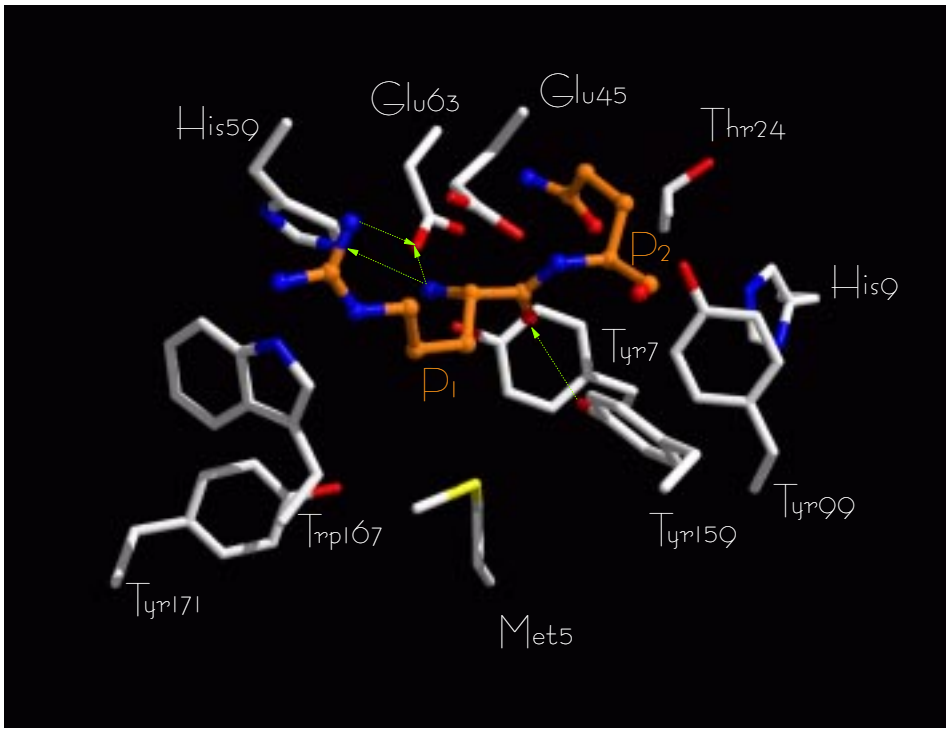

Fig. 7. MD model of complex **3a**, focused on MHC pockets A–B (see the legend to Fig. 6). A mean conformation was averaged from the last 100 conformers and submitted to 500 steps of steepest descent, followed by 1500 steps of conjugate-gradient energy relaxation.

Qualitative and quantitative analysis of peptide–MHC H-bonds

Reporting the number of MHC–peptide H-bonds as well as their frequencies during the simulation allows a clear distinction between peptides **1** and **2** and the Gln^2 analogue (peptide **3**, Fig. 5). The stability of protein–peptide H-bonds was assessed by computing the frequency of occurrence of the interaction throughout the MD tra-

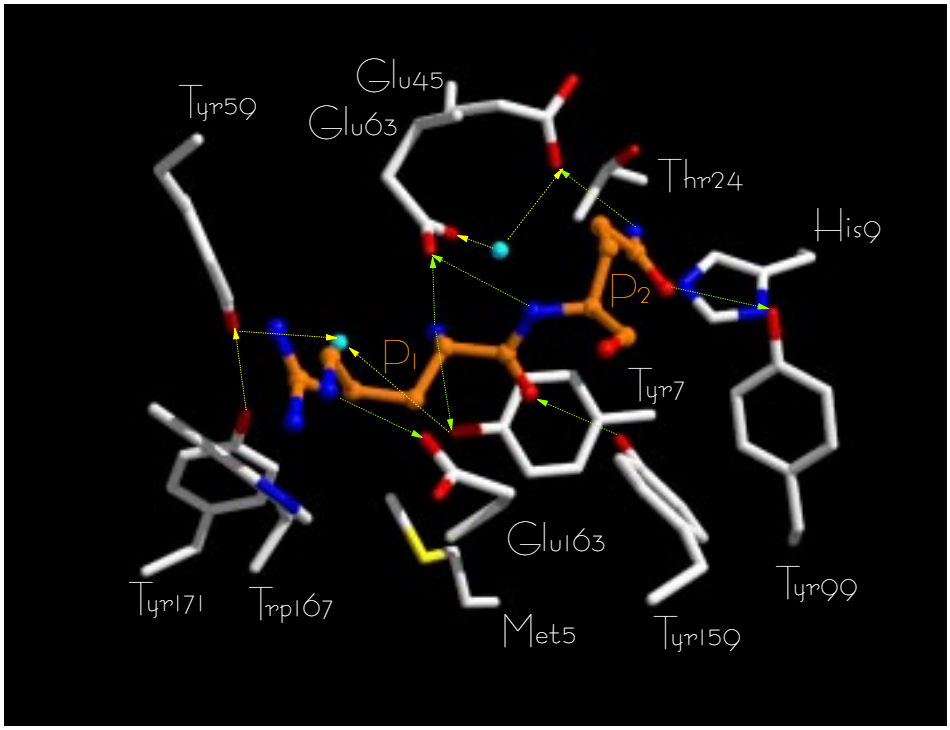

Fig. 8. MD model of complex **3b**, focused on MHC pockets A–B (see the legend to Fig. 6). The mean conformation was obtained as described in Fig. 7.

TABLE 3
MHC–PEPTIDE H-BONDS WITH A FREQUENCY HIGHER THAN 50%

Pn	atom	HLA-B27	1a	1b	2a	2b	3a	3b
P1	N	Tyr^{7} (OH)		□		□		□
		His^{59} (NE2)	□		□		□	
		Glu^{45} (OE2)					■	
		Glu^{63} (OE1)	□	□				□
		Glu^{63} (OE2)			□	□		
		Glu^{163} (OE2)				■		
	NE	Glu^{163} (OE2)	□		×	×		□
	NH2	Glu^{63} (OE1)			×	×	■	
		Glu^{63} (OE2)			×	×	■	
	O	Tyr^{159} (OH)		□	□		□	□
P2	N	Glu^{63} (OE1)	□					□
		Glu^{63} (OE2)			□	□		
	NE	Glu^{45} (OE2)	■				×	×
		Glu^{63} (OE1)				■	×	×
	OE1	Tyr^{99} (OH)	×	×	×	×		■
	NE2	His^{9} (NE2)	×	×	×	×		■
		Glu^{45} (OE2)	×	×	×	×		■
	NH1	His^{9} (NE2)	□		□	□	×	×
		Thr^{24} (OG1)		■			×	×
	NH2	Thr^{24} (OG1)	□		□	□	×	×
		Glu^{45} (OE1)	□	□		□	×	×
		Glu^{45} (OE2)	□		□	□	×	×
		Glu^{63} (OE1)		■			×	×
P3	N	Tyr^{99} (OH)	□	□		□	□	
P6	OG	Ala^{69} (O)				□		
		Thr^{73} (OG1)				□		
P8	O	Trp^{147} (NE1)		□	□			□
	OE1	Lys^{146} (OE1)			□			
P9	N	Asp^{77} (OD1)	□	□	□	□		□
	OXT	Tyr^{84} (OH)		□				
		Thr^{143} (OG1)		□		□		□
		Lys^{146} (NZ)	□	□	□			

Empty boxes indicate interactions that are common to at least two complexes, whereas filled boxes represent unique MHC–peptide hydrogen bonds. The absence of a specific side chain is featured by a cross.

jectory (400 conformations). A frequency higher than 50% was chosen to characterize strong H-bonds. Medium interactions were assigned a frequency between 25 and 50%.

About 25 H-bonds have been identified for peptides **1** and **2** in complex with B*2703 and B*2705 while 50% less could be found for the Gln^{2} analogue with the two subtypes (Fig. 5). The distribution of strong and medium H-bonds correlates well with the binding potency of the peptide. A similar number of strong H-bonds were found for complexes **1a**, **1b**, **2a** and **2b**, consistent with the similar binding efficiencies of peptides **1** and **2** to both subtypes. On the other hand, a reduced number of medium and/or strong H-bonds (peptide **3** in complex with the two alleles) correlates with the decreased binding of this peptide. The weakest binding potency (peptide **3** to B*2703) could effectively be qualitatively and quantitatively related to the distribution of intermolecular H-bonds. Not only the number but also the quality of the MHC–ligand interactions correlates well with the binding potency.

To accurately localize the interactions that may explain peptide specificity variations, all H-bonds with frequencies higher than 50% were identified for the six complexes (Table 3). The first noticeable difference between the two HLA-B27 alleles is the H-bonding network between the MHC residues involved in binding to the peptide P1 position. In HLA-B*2705, two amino acid side chains are H-bonded to the peptide N-terminus (Tyr^{7}/Tyr^{171} in the crystal structure, Tyr^{7}/Glu^{63} in the MD models) (Table 4, Figs. 6–8). Both side chains are fixed by a subtle water-relayed H-bond network involving proximal MHC side chains (Tyr^{59}, Glu^{45}, Tyr^{171}). Tyr^{59} is directly bound to Tyr^{171}, and indirectly to Tyr^{7}, Glu^{45} and Glu^{63}. The single point mutation occurring for HLA-B*2703 (Tyr^{59}→His) perturbs this network. The bound water molecule disappears and the peptide N-terminus binds to His^{59} and no more to Tyr^{7} (Table 3, Fig. 7). The consequence on the H-bond balance is a loss of one direct MHC–MHC interaction (His^{59} cannot interact with Tyr^{171}) and five water-mediated interactions for complex **3a** (Fig. 7). The resulting conformational change may be well accommodated as far as P2 is strongly bound to pocket B (His^{9}, Thr^{24}, Glu^{45}) and the resulting H-bonds are strong enough to maintain the peptide in the binding groove (P2=Arg). If P2 is not an arginine (peptide **3**), the resulting interaction to pocket B (Thr^{24} and Glu^{45} notably) is much weaker and the conformational rearrangement at P1 is important (see the three new H-bonds for the P1 position in complex **3a**, Fig. 7). The Arg to Gln change at the P2 position of the bound peptide is better tolerated by HLA-B*2705 (complex **3b**, Fig. 8) as the Tyr^{59} side chain is still able to fix the position of Tyr^{7}. During the MD simulation, the N-terminal C^{α}-N bond of the bound peptide has rotated to gain a new H-bond to Glu^{63}. However, it is still bound to Tyr^{7} as in the reference structure. Importantly, the Gln^{2} side chain is bound to Glu^{45} and Tyr^{99}, thus providing additional interactions to pocket B when compared to complex **3a** (Fig. 8). The quality of the interaction between Gln^{2} and pocket B is, however, much inferior to that observed for peptide analogues **1** and **2** bearing an optimal Arg residue (Table 4), thus explaining the reduced binding affinity of peptide **3** for HLA-B*2705.

For the set of peptides studied here, the advantage of

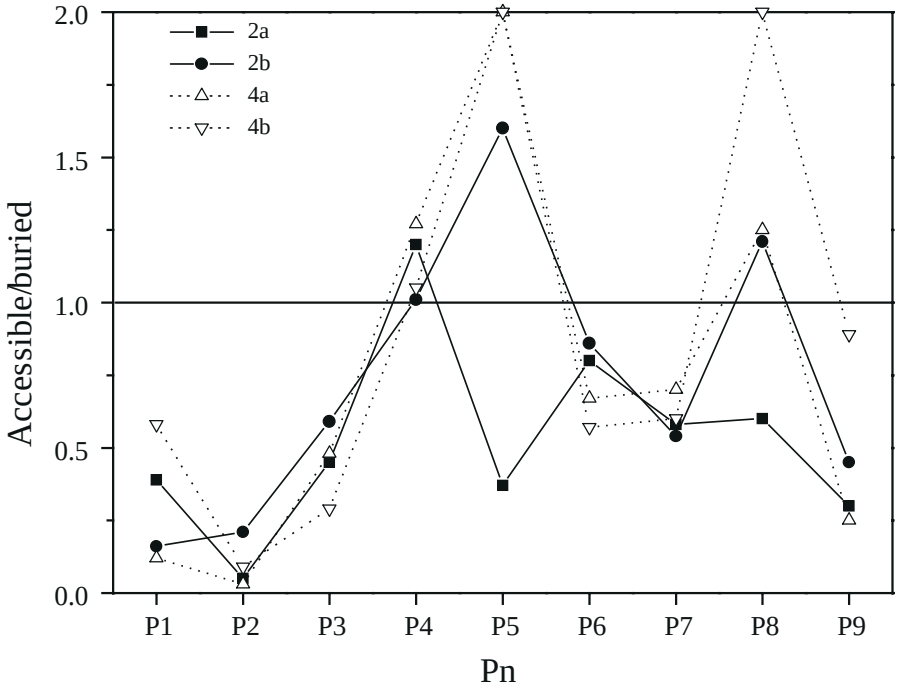

Fig. 9. Accessible versus buried surface areas of peptides **2** and **4** in complex with B*2703 (a) and B*2705 (b) alleles (see the legend to Fig. 4).

Arg over Ala at P1 could be quantified by the gain of two water accessible salt bridges to Glu^{63}/Glu^{163} (Table 3). However, this does not correspond to a higher binding affinity of the Arg analogue when compared to the Ala^{1} peptide. An Ala side chain is much easier to desolvate and optimally interacts with conserved apolar residues of pocket A (Met^{5}, Trp^{167}), thus explaining a rather similar binding affinity of peptides **1** and **2** to both subtypes. However, the present model cannot fully explain recent data, indicating that basic residues are overrepresented at the P1 position of B*2703-bound natural ligands [23]. From a purely statistical point of view, various rotamers of basic P1 side chains could develop a salt bridge with at least three negatively charged amino acids located at the rim of pocket A (Glu^{58}, Glu^{63}, Glu^{163}), and thus stabilize the MHC–peptide complex.

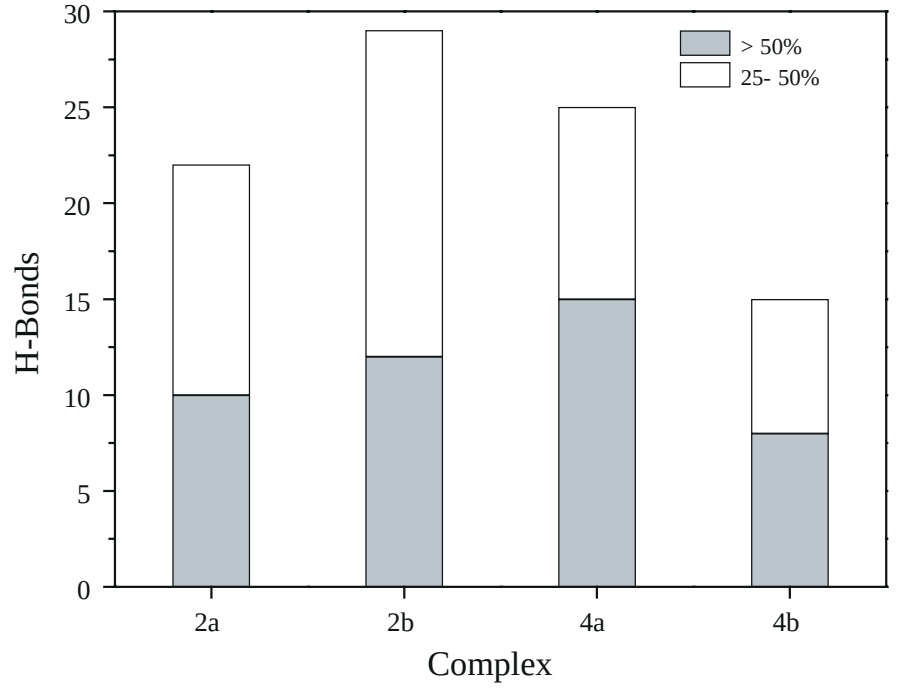

Fig. 10. Intermolecular hydrogen bonds for peptides **2** and **4** in complex with B*2703 (a) and B*2705 (b) alleles (see the legend to Fig. 5).

TABLE 4
MHC–PEPTIDE INTERACTION ENERGIES CALCULATED FROM THE LINEAR POISSON–BOLTZMANN EQUATION AND AMBER FORCE-FIELD CALCULATIONS

Peptide	ΔG^0_{coul} [a]	ΔG^0_{reac} [b]	ΔG^0_{elec} [c]	ΔH_{elec} [d]	ΔH_{vdw} [e]	ΔG_{tot} [f]	ΔH_{tot} [g]
1a	–499	487	–12	–87	–90	**–102**	**–177**
1b	–542	498	–44	–79	–97	**–141**	**–176**
2a	–359	340	–19	–72	–77	**–96**	**–149**
2b	–389	335	–54	–75	–80	**–134**	**–155**
3a	–306	334	28	–44	–68	**–40**	**–112**
3b	–371	380	9	–62	–67	**–58**	**–129**

[a] ΔG^0_{coul}: Coulombic component of MHC–peptide electrostatic interaction energy (charge–charge, charge–dipole, dipole–dipole interactions). $\Delta G^0_{coul} = \Delta G^0_{coul(P-L)} - \Delta G^0_{coul(P)} - \Delta G^0_{coul(L)}$ [50], where P–L describes the protein–ligand complex, P the protein and L the ligand.
[b] ΔG^0_{reac}: corrected self-reaction field component of MHC–peptide electrostatic interaction energies (energy required to transfer a molecule from a continuum dielectric (vacuum) to another (water). $\Delta G^0_{reac} = \Delta G^0_{reac(P-L)} - \Delta G^0_{reac(P)} - \Delta G^0_{reac(L)}$. As the contribution of the protein–ligand complexes ($\Delta G^0_{reac(P-L)}$) and of the isolated protein ($\Delta G^0_{reac(P)}$) could be omitted from the calculation without affecting the reliability of the results [61], this component corresponds here to the free energy of peptide desolvation ($-\Delta G^0_{reac(L)}$).
[c] ΔG^0_{elec}: total electrostatic interaction energy ($\Delta G^0_{coul} + \Delta G^0_{reac}$).
[d] ΔH_{elec}: AMBER electrostatic interaction energy ($\varepsilon = 4r_{ij}$).
[e] ΔH_{vdw}: AMBER van der Waals interaction energy. $\Delta H_{vdw} = A/r^{12} - B/r^6$, where r is the distance between atom pairs and A and B are atom-type-dependent parameters.
[f] ΔG_{tot}: total interaction energy ($\Delta G^0_{elec} + \Delta H_{vdw}$).
[g] DH_{tot}: total AMBER interaction enthalpy ($\Delta H_{elec} + \Delta H_{vdw}$).

MHC–peptide interaction energies

Interaction energies were extrapolated for all six energy-minimized time-averaged conformations (Table 4) by summing up the van der Waals nonbonded interaction energy (calculated with the AMBER 4.0 force field) and the electrostatic component (calculated by solving the linear form of the Poisson–Boltzmann equation), as recently described [60]. As both ligands and protein structures are very similar for all complexes, distortion energies as well as translational/rotational entropy losses upon binding were neglected here. Moreover, the self-reaction field energy component of the electrostatic interaction energy was limited to the contribution of the isolated peptide (free energy of desolvation) and calculated from the bound-peptide coordinates extracted from the MHC–peptide binary complexes. It has recently been shown that neglecting the protein contribution to the self-reaction field energy is indeed possible and does not alter the reliability of the obtained results [61].

Our computational protocol is able to properly rank the binding of the three peptides **1**–**3** to both MHC alleles. Peptide **3** clearly interacts much weaker than the other two peptides **1** and **2**, whatever the MHC allele. The weakest interaction energy was observed for binding of **3** to B*2703, and is thus in agreement with binding data (Table 1). Force-field interaction enthalpies (calculated by summing up both AMBER van der Waals and electrostatic components, using a dielectric permittivity of $4r_{ij}$) were much less related to the observed binding data, as peptide **2** was always disfavoured with regard to peptide **1** (Table 4). Notably, taking into account the peptide desolvation energy by the continuum eletrostatics method permits to compensate for the weakest Coulombic interactions provided by peptide **2** (P1=Ala) to both alleles, with respect to that observed for the Arg[1] analogue. It may be noticed that the free electrostatic interaction energies (ΔG^0_{elec}, Table 4) computed by the continuum electrostatics method were also in rather good qualitative agreement with the binding data reported in Table 1. Hence, the three peptides are highly polar and interact mainly via H-bonds and salt bridges.

For an even more realistic ranking of highly polar ligands than those presented here, free energy perturbation [62] is probably the method of choice. Unfortunately, the enormous amount of CPU time that would be necessary for this computation precludes its systematic use in fast screening of a set of congeneric molecules. Synthesis and in vitro binding assays in this case provided a faster and experimentally determined answer.

MD simulation of MHC–peptide complexes could relate observed binding potencies and allele specificity to simple molecular criteria (inter-cmass distances, atomic fluctuations, accessible surface areas, distribution and

TABLE 5
INFLUENCE OF A P1 β-AMINO ACID ON THE HLA-B27 SUBTYPE SELECTIVITY OF A MODIFIED HLA-B27 LIGAND

Peptide number	Sequence	EC_{50} (μM)[a]	
		B*2703	B*2705
	(P1-RYQKSTEL)		
2	P1=Ala	3.0	4.0
4	P1=Bal[b]	7.5	20

[a] Concentration of the peptide (in μM) at which HLA-B27 fluorescence (measured by FMC analysis with an anti-B27 monoclonal antibody) on RMA-S cells was half the maximum obtained with the wild-type peptide (peptide **1**, Table 1).
[b] Bal: β-alanine (H_2N-CH_2-CH_2-CO).

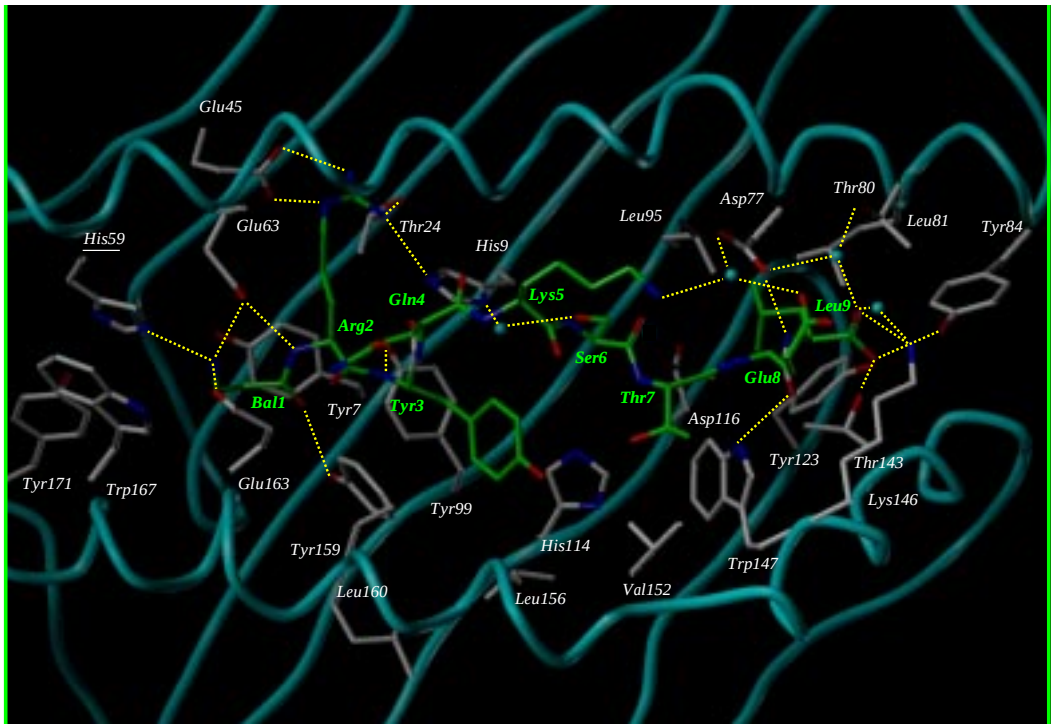

Fig. 11. Energy-minimized time-averaged MD model of complex **4a**. The MHC protein backbone is displayed as a solid cyan tube, with peptide-interacting side chains. The bound peptide **4** is represented by sticks. The following color coding has been used: carbon, white (protein) or green (peptide); nitrogen, blue; oxygen, red; sulfur, yellow. Bound water molecules are shown as cyan balls. Yellow broken lines indicate MHC–peptide H-bonds.

location of intermolecular H-bonds). A single point mutation in the HLA binding groove is sufficient to break an H-bond network in the vicinity of the peptide N-terminus. As previously suggested on the basis of peptide binding analyses [28], this minor change strengthens even more the binding role of the dominant anchor P2 side chain (Arg) for one allele (HLA-B*2703) and explains why changing P2 to Gln has more detrimental effects in pep-

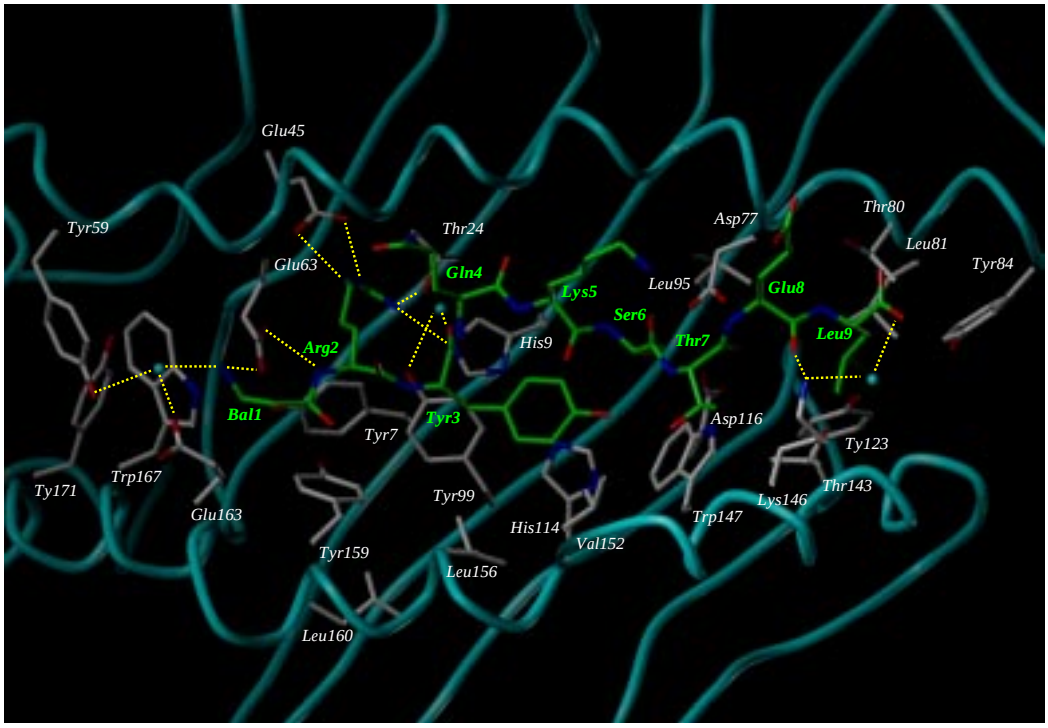

Fig. 12. Energy-minimized time-averaged MD model of complex **4b** (see the legend to Fig. 11).

tide binding for HLA-B*2703 than for HLA-B*2705, where compensatory stabilization of the MHC–peptide complex is still possible by H-bonded MHC side chains. Interestingly, the most spectacular consequence of protein/peptide variability affects an area far away (10 Å) from the location of the point mutations. It concerns the stability of the interaction between the peptidic C-terminal residue and its complementary pocket F, which has recently been shown to play a decisive role in linking particular B27 alleles to spondyloarthropathies [24,26].

Protein-based design

Relating the structure of B27 subtypes to the sequence of their naturally bound peptides is a crucial step in identifying potential immunodominant epitopes that may discriminate alleles and confer susceptibility or resistance to autoimmune diseases. One striking feature concerns the single point mutation (Tyr59His) distinguishing HLA-B*2705 from HLA-B*2703, which is unique among HLA proteins. It is believed that B*2703 selects a subset of the peptides presented by HLA-B*2705 [34]. Recent studies have identified some peptides that are naturally presented by both subtypes, and at least one natural B*2705 ligand (the undecamer RRYLENGKETL) is not presented by B*2703 [23,63]. The only difference between both alleles concerns the position 59 located in pocket A which interacts with the N-terminal amino acid of the bound peptide (Fig. 6). As pocket A is slightly wider for HLA-B*2703, extending the peptide backbone towards His^{59} by replacing the natural P1 residue by a β-amino acid should theoretically allow a better discrimination of both alleles. This structural change should be much better accommodated by HLA-B*2703 (H-bond between His^{59} and the N-terminus of the P1 β-amino acid) than by HLA-B*2705, for which a steric clash with the Tyr^{59} side chain may be expected.

Starting from the self-peptide **1** (RRYQKSTEL) naturally presented by HLA-B*2705 [20] and B*2703 [23], Ala and Bal (β-alanine) were substituted for the natural Arg at P1 (Table 5). The Ala^{1} peptide analogue was here taken as a reference for its strong binding to both subtypes. The two ligands were synthesized and tested for their binding to B*2703 and B*2705. As expected from the topology of the binding cleft, only the Bal analogue could discriminate between the two subtypes, with a better binding to HLA-B*2703 (Table 5).

To rationalize the experimental binding data, the non-natural ligand **4** was simulated a posteriori, in complex with both alleles, using exactly the same conditions as those employed for simulating the natural MHC–peptide complexes (see the section Computational procedures). Using two of the previously described molecular parameters (accessible versus buried surface area of the bound ligand, intermolecular H-bonds) as quality control of the complex stability, peptide **4** was indeed found to be much better accommodated by B*2703 than by B*2705 (Figs. 9 and 10). Interestingly, the weak binding of the β-Ala peptide to B*2705 could also be related to a partial dissociation of the C-terminus from its complementary pocket F (Fig. 9), far away from the peptide mutation site. The present data, in agreement with previous MD simulations of different MHC–peptide complexes [9], suggest that the expulsion of the C-terminus from pocket F could be the very first event in the dissociation of weak binding peptides from class I MHC binding grooves. The modified P1 position is, however, significantly more buried when the host protein is the B*2703 allele (Fig. 9). A qualitative and quantitative analysis of intermolecular hydrogen bonds also supports the reported binding data. A total of 15 H-bonds could be depicted for complex **4b** (peptide **4** in complex with B*2705), whereas 26 interactions have been found for complex **4a** (peptide **4** in complex with B*2703, Fig. 10). However, this analysis was unable to explain the reduced affinity of the β-Ala compound for B*2703, when compared to that of the natural Ala analogue **2** (Table 4). The slight differences seen in the epitope stabilization assay are certainly too subtle for the short MD runs reported here. They probably result from: (i) the absence of a side chain at position P1 of ligand **4**; and (ii) a weaker binding contribution of the bulging P4–P8 part, for which higher atomic fluctuations (data not shown) and less nonbonded contacts (see the high solvent accessibility of the P5 and P8 residues for ligand **4**, Fig. 9) have been noticed. Energy-minimized time-averaged conformations of both complexes (Figs. 11 and 12) clearly depict significant differences in the MHC pocket A (Tyr^{59}, Trp^{167}, Tyr^{171}), which deviates dramatically from the starting crystal coordinates for B*2705 only (rmsd values from all pocket A atoms of 1.5 and 2.5 Å for B*2703 and B*2705, respectively). The major conformational alterations upon Bal1 binding were observed for the Tyr^{59}-Trp^{167}-Tyr^{171} triad (rmsd values of 1.7 and 2.7 Å for B*2703 and B*2705, respectively)sl. As predicted, the Tyr^{59} side chain was shifted away from the peptide N-terminus and is now interacting via a water molecule with the β-alanine terminal ammonium (Fig. 12). In contrast, the β-amino acid can directly interact with the larger pocket A of B*2703 through three H-bonds to His^{59} (mutated position), Glu^{63} and Glu^{163} (Fig. 11). Another significant difference in the binding of peptide **4** to both alleles concerns the C-terminal amino acid, which has nearly lost, upon binding to B*2705, all H-bonds to the polar side chains of pocket F (Tyr^{84}, Thr^{143}, Lys^{146}; compare Figs. 11 and 12). The incorporation of a β-amino acid at P1 has modified the above described H-bond network between MHC side chains and the peptide N-terminus. The bound water molecule located in pocket A of HLA-B*2705 (recall Fig. 6) has either disappeared (B*2703, Fig. 11) or has been shifted towards the extreme left end of the binding groove (B*2705, Fig. 12).

More importantly, these results show that the incor-

poration of a β-amino acid in the peptide sequence does not abrogate binding to HLA-B27 subtypes. Peptide **4** is one of the very first ligands for which the backbone modification of an anchor residue does not abolish class I MHC binding. Up to now, only a retro-inverso (NH-CO instead of CO-NH) and a reduced peptide bond (CH_2-NH instead of CO-NH) pseudopeptide analogue of an HLA-A2-binding peptide have been proposed as successful P1 modifications [64,65]. However, a β-amino acid at P1 presents the advantage to preserve the backbone direction of the peptide ligand and the H-bonding capacity of the first peptide bond (to Glu^{63} and Tyr^{159}), so that less 3D conformational changes of the MHC binding cleft are necessary to accommodate the modified ligand. The recently described X-ray structure of an MHC–peptide–TcR ternary complex [66] suggests that the latter feature may be of particular importance for a proper recognition of the MHC–ligand pair by a TcR. Furthermore, it opens the door to the incorporation of β-amino acids at other anchor positions, notably P2, P3 and P9. Potential TcR-binding amino acids have already been replaced by various organic spacers without affecting the binding of the corresponding ligand to class I MHC proteins [38,67,68]. The present design study demonstrates that substituting a β-amino acid for a natural residue is a further solution for designing high-affinity MHC ligands with improved stability and pharmacokinetic properties. This is an absolute prerequisite for the therapeutic use of MHC ligands either as MHC blockers [69] or as T-cell receptor antagonists [70].

Conclusions

MD simulations have been used in the present study as a tool for explaining peculiar structure–activity relationships at the level of the protein–ligand interaction complexes. The current study is not aimed at quantitatively ranking MHC ligands and predicting their binding affinities. For that purpose, free energy calculations using much longer equilibration and conformational sampling would be necessary. More simply, dynamical properties of the modeled complexes can be qualitatively well related to known binding data. Notably, monitoring protein–ligand intramolecular distances, the atomic mobility of the bound ligands, the ratio of accessible versus buried surface areas, the history and the quality of peptide–protein H-bonds allow a clear discrimination of high-affinity from weak-binding peptides. This computational approach, based on the qualitative analysis of short MD trajectories, has already been used to succesfully predict the bound conformation of a natural HLA-A2-restricted epitope [37] prior to X-ray structure determination [55], to identify T-cell epitopes from the primary structure of potentially interesting proteins [9] and to design high-affinity nonnatural ligands [38,71]. Herewith, we propose its application to the rationalization of peptide specificity for closely related HLA alleles and the design of non-natural ligands with increased specificity for one HLA-B27 subtype. Identifying the molecular rules, fine tuning peptide selection by HLA alleles is a crucial step for better understanding the peptide–HLA interactions that may confer either susceptibility or resistance to immunological diseases associated with particular HLA alleles.

Acknowledgments

D.R. wishes to thank the computational center of the ETH Zürich for generous allocation of computer time on the CRAY J90 and PARAGON machines. This work was supported by the Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung (Project No. 31-45504.95) and by Grant SAF 94-0891 from the Plan Nacional de I+D to J.A.L.C. J.R.L. is a fellow of the Basque Government.

References

1 Heemels, M.T. and Ploegh, H.L., Annu. Rev. Biochem., 64 (1995) 643.
2 Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennet, W.S., Strominger, J.L. and Wiley, D.C., Nature, 329 (1987) 506.
3 Falk, K., Rötzschke, O., Stevanovic, S., Jung, G. and Rammensee, H.-G., Nature, 351 (1991) 290.
4 Spies, T., Bresnahan, M., Bahram, S., Arnold, D., Blank, G., Mellins, E., Pious, D. and DeMars, R., Nature, 348 (1990) 744.
5 Stern, L.J. and Wiley, D.C., Structure, 2 (1994) 245.
6 Madden, D.R., Annu. Rev. Immunol., 13 (1995) 587.
7 Corr, M., Boyd, L.F., Frankel, S.R., Kozlowski, S., Padlan, E.A. and Margulies, D.H., J. Exp. Med., 176 (1992) 1681.
8 Huczo, E.L., Bodnar, W.M., Benjamin, D., Sakaguchi, K., Zhu, N.Z., Shabanowitz, J., Henderson, R.A., Appella, E., Hunt, D.F. and Engelhard, D., J. Immunol., 151 (1993) 2572.
9 Rognan, D., Scapozza, L., Folkers, G. and Daser, A., Biochemistry, 33 (1994) 11476.
10 Chelvanayagam, G., Jakobsen, I.B., Gao, X. and Easteal, S., Protein Eng., 9 (1996) 1151.
11 Brewerton, D.A., Hart, F.D., Nicholls, A. and Sturrock, R.D., Lancet 2, (1973) 994.
12 Todd, J.A., Bell, J.I. and McDevitt, H.O., Nature, 329 (1987) 599.
13 Wordsworth, B.P., Lanchbury, J.S., Sakkas, L.I., Welsh, K.I., Panayi, G.S. and Bell, J.I., Proc. Natl. Acad. Sci. USA, 86 (1989) 10049.
14 Hill, A.V.S., Elvin, J., Willis, A.C., Aidoo, M., Allsopp, C.E.M., Gotch, F.M., Gao, X.M., Takiguchi, M., Greenwood, B.M., Townsend, A.R.M., McMichael, A.J. and Whittle, H.C., Nature, 360 (1992) 434.
15 Benjamin, R. and Parham, P., Immunol. Today, 11 (1990) 137.
16 Kingsley, G. and Sieper, J., Immunol. Today, 14 (1993) 387.
17 López-Larrea, C., Sujirachato, K., Mehra, N.K., Chiewsilp, P., Isarangkura, D., Kanga, O., Dominguez, O., Coto, E., Peña, M., Setién, F. and Gonzales-Roces, S., Tissue Antigens, 45 (1995) 169.
18 Nasution, A.R., Mardjuadi, A., Kunmartini, S., Suryadhana, N.G., Setyohadi, B., Sudarsono, D., Lardy, N.M. and Feltkamp, T.E.W., J. Rheumatol., 24 (1997) 1111.

19 D'Amato, M., Fiorillo, M.T., Carcassi, C., Mathieu, A., Zuccarelli, A., Bitti, P.P., Tosi, R. and Sorrentino, R., Eur. J. Immunol., 25 (1995) 3199.
20 Jardetzky, T.S., Lane, W.S., Robinson, R.A., Madden, D.R. and Wiley, D.C., Nature, 353 (1991) 326.
21 Rojo, S., Garcia, F., Villadangos, J.A. and López de Castro, J.A., J. Exp. Med., 177 (1993) 613.
22 Rötzchke, O., Falk, K., Stevanovic, S., Gnau, V., Jung, G. and Rammensee, H.G., Immunogenetics, 39 (1994) 74.
23 Boisgérault, F., Tieng, V., Stolzenberg, M.C., Dulphy, N., Khalil, I., Tamouza, R., Charron, D. and Toubert, A., J. Clin. Invest., 98 (1996) 2764.
24 Garcia, F., Marina, A. and López de Castro, J.A., Tissue Antigens, 49 (1997) 215.
25 Garcia, F., Galocha, B., Villadangos, J.A., Lamas, J.R., Albar, J.P., Marina, A. and López de Castro, J.A., Tissue Antigens, 49 (1997) 580.
26 Fiorillo, M.T., Meadows, L., D'Amato, M., Shabanowitz, J., Hunt, D.F., Apella, E. and Sorrentino, R., Eur. J. Immunol., 27 (1997) 368.
27 Parker, K.C., Biddison, W.E. and Coligan, J.E., Biochemistry, 33 (1994) 7736.
28 Villadangos, J.A., Galocha, B., Garcia, F., Albar, J.P. and López de Castro, J.A., Eur. J. Immunol., 25 (1995) 2370.
29 Galocha, B., Lamas, J.R., Villadangos, J.A., Albar, J.P. and López de Castro, J.A., Tissue Antigens, 48 (1996) 509.
30 Raghavan, M., Lebron, J., Johnson, J. and Bjorkman, P., Protein Sci., 5 (1996) 2080.
31 Madden, D.R., Gorga, J.C., Strominger, J.L. and Wiley, D.C., Nature, 353 (1991) 321.
32 Madden, D.R., Gorga, J.C., Strominger, J.L. and Wiley, D.C., Cell, 70 (1992) 1035.
33 López, D., García-Hoyo, R. and López de Castro, J.A., J. Immunol., 152 (1994) 5557.
34 Villadangos, J.A., Galocha, B., García-Hoyo, R., López, D., Garcia, F. and López de Castro, J.A., Eur. J. Immunol., 24 (1994) 2548.
35 Colbert, R.A., Rowland-Jones, S.L., McMichael, A.J. and Frelinger, J.A., Immunity, 1 (1994) 121.
36 Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.M., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
37 Rognan, D., Zimmermann, N., Jung, G. and Folkers, G., Eur. J. Biochem., 208 (1992) 101.
38 Rognan, D., Scapozza, L., Folkers, G. and Daser, A., Proc. Natl. Acad. Sci. USA, 92 (1995) 753.
39 Collins, E.J., Garboczi, D.N., Karpusas, M.N. and Wiley, D.C., Proc. Natl. Acad. Sci. USA, 92 (1995) 1218.
40 Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L., J. Chem. Phys., 70 (1983) 926.
41 Pearlman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A., AMBER 4.0, University of California, San Francisco, CA, U.S.A., 1992.
42 Besler, B.H., Merz, K.M. and Kollman, P.A., J. Comput. Chem., 11 (1990) 431.
43 SYBYL, release 6.2, TRIPOS Associates Inc., St. Louis, MO, U.S.A.
44 Stewart, J.J.P., J. Comput.-Aided Mol. Design, 4 (1990) 1.
45 Dewar, M.J.S. and Thiel, W., J. Am. Chem. Soc., 99 (1977) 4899.
46 Ryckaert, J.P., Cicotti, G. and Berendsen, H.J.C., J. Comput. Phys., 23 (1977) 327.
47 Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684.
48 Klapper, I., Hagstrom, R., Fine, R., Sharp, K. and Honig, B., Proteins Struct. Funct. Genet., (1986) 47.
49 Warwicker, J. and Watson, H.C., J. Mol. Biol., 157 (1982) 671.
50 Gilson, M.K. and Honig, B., Proteins Struct. Funct. Genet., 4 (1988) 7.
51 Gilson, M.K., Sharp, K. and Honig, B., J. Comput. Chem., 9 (1988) 327.
52 Ljunggren, H.G., Stam, N.J., Ohlen, C., Neefjes, J.J., Hoglund, P., Heemels, M.T., Bastin, J., Schumacher, T.N., Townsend, A., Karre, K. and Ploegh, H.L., Nature, 46 (1990) 476.
53 Ellis, S.A., Taylor, C. and McMichael, A., Hum. Immunol., 5 (1982) 49.
54 van Gunsteren, W.F. and Berendsen, H.J.C., Angew. Chem., Int. Ed. Engl., 29 (1990) 992.
55 Madden, D.R., Garboczi, D.N. and Wiley, D.C., Cell, 75 (1993) 693.
56 Reid, S.W., McAdam, S., Smith, K.J., Klenerman, P., O'Callaghan, C.A., Harlos, K., Jakobsen, B.K., McMichael, A.J., Bell, J.I., Stuart, D.I. and Jones, E.Y., J. Exp. Med., 184 (1996) 2279.
57 Rovero, P., Vigano, S., Pegorado, S., Revoltella, R., Riganelli, D., Fruci, D., Greco, G., Butler, R. and Tanigaki, N., J. Pept. Sci., 1 (1995) 266.
58 Brunne, R.M., Berndt, K.D., Güntert, P., Wüthrich, K. and van Gunsteren, W.F., Proteins Struct. Funct. Genet., 23 (1995) 49.
59 Fox, T. and Kollman, P.A., Proteins Struct. Funct. Genet., 25 (1996) 315.
60 Gallego, J., Ortiz, A.R., de Pascual-Teresa, B. and Gago, F., J. Comput.-Aided Mol. Design, 11 (1997) 114.
61 Taylor, N.R. and von Itzstein, M., J. Comput.-Aided Mol. Design, 10 (1996) 233.
62 Bash, P.A., Singh, U.C., Langridge, R. and Kollman, P.A., Science, 236 (1987) 564.
63 Garcia, F., Marina, A., Albar, J.P. and López de Castro, J.A., Tissue Antigens, 49 (1997) 23.
64 Guichard, G., Calbo, S., Muller, S., Kourilsky, P., Briand, J.-P. and Abastado, J.-P., J. Biol. Chem., 270 (1995) 26057.
65 Guichard, G., Connan, F., Graff, R., Ostankovitch, M., Muller, S., Guillet, J.-G., Choppin, F. and Briand, J.-P., J. Med. Chem., 39 (1996) 2030.
66 Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E. and Wiley, D.C., Nature, 384 (1996) 134.
67 Weiss, G.A., Collins, E.J., Garboczi, D.N., Wiley, D.C. and Schreiber, S.L., Chem. Biol., 2 (1995) 401.
68 Bouvier, M. and Wiley, D.C., Proc. Natl. Acad. Sci. USA, 93 (1996) 4583.
69 Adorini, L., Muller, S., Cardinaux, F., Lehmann, P.V., Falcioni, F. and Nagy, Z.A., Nature, 334 (1988) 623.
70 De Magistris, M.T., Alexander, J., Coggeshall, M., Altman, A., Gaeta, F.C.A., Grey, H.M. and Sette, A., Cell, 68 (1992) 625.
71 Rognan, D., Habilitationsschrift, ETH Zürich, Zürich, Switzerland, 1997.
72 Saper, M.A., Bjorkman, P.J. and Wiley, D.C., J. Mol. Biol., 219 (1991) 277.
74 Kraulis, P.J., J. Appl. Crystallogr., 24 (1991) 846.
75 Merritt, E.A. and Murphy, M.E.P., Acta Crystallogr., D50 (1994) 869.
76 Connolly, M.J., J. Appl. Crystallogr., 16 (1983) 548.