International Immunology, Vol. 10, No. 12, pp. 1753–1764

Molecular modeling of hen egg lysozyme HEL[52–61] peptide binding to I-A^k MHC class II molecule

Peter Weber, Isabelle Raynaud, Laurent Ettouati[1], Marie-Claude Trescol-Biémont[2], Pierre-Alain Carrupt, Joëlle Paris[1], Chantal Rabourdin-Combe[2], Denis Gerlier[3] and **Bernard Testa**

Institut de Chimie Thérapeutique, Ecole de Pharmacie, Université de Lausanne, 1015 Lausanne, Switzerland
[1]Laboratoire de Chimie Thérapeutique, Faculté de Pharmacie, 8 avenue Rockefeller, 69373 Lyon Cedex 08, France
[2]Immunobiologie Moléculaire, INSERM U98XX, ENS Lyon, 69364 Lyon Cedex 07, France
[3]Immunité et Infections Virales, IVMC, CNRS-UCBL UMR5537, 69372 Lyon Cedex 08, France

Keywords: antigen presentation, hen egg lysozyme, I-A^k, MHC class II, molecular modeling, mouse

Abstract

A bound conformation of the antigenic decapeptide hen egg lysozyme HEL[52–61] associated to the mouse MHC class II (MHC II) I-A^k was modeled by homology with the three-dimensional structure of hemagglutinin HA[306–318]–HLA-DR1 complex. HEL peptide Tyr53 could not be aligned with the HA peptide Tyr308 because this resulted in a buried Tyr53 side chain within the I-A^k peptide-binding groove and this conflicted with this side chain being recognized by T cells. Therefore, Asp52 of HEL was fixed as the P1 anchor and aligned on Tyr308 of HA. After molecular dynamics, the modeled complex was stable even in the absence of any constraint. The peptide backbone adopted a polyproline II-like conformation with canonical hydrogen bonding between the peptide backbone and MHC II molecule. Asp52, Ile55, Gln57 and Ser60 were predicted to be deeply buried into P1, P4, P6 and P9 MHC II pockets, and Tyr53, Leu56, Asn59 and Arg61 as TCR contacting residues. The modeling of 15 complexes associating I-A^k with peptides derived from HEL[52–61] by single amino acid substitution proved stable with conserved hydrogen bonds and side chain orientation compatible with their recognition by two T cell hybridomas. Moreover, comparison with the recently solved crystal structure of the related HEL[50–62]–I-A^k complex revealed striking similarities.

Introduction

MHC molecules are polymorphic cell-surface glycoproteins which play a key role in the development of an immune response by presenting antigen-derived peptides to specific T cells. The ability of each MHC allele to bind a large subset of different peptides explains how the expression of only a dozen of the different MHC alleles by a given organism enables the development of an immune response against every possible pathogen (1). The peptide–MHC complex can be described as resulting from a high-affinity but low-specificity interaction. This prevents exchange at the cell surface and possible generation of an inappropriate immune response. In the antigen-presenting cell, peptides are derived from the intracellular degradation of protein antigens and associated with maturating MHC molecules (for review, see 1,2). The stabilization of the MHC molecules resulting from their binding to peptide allows their stable expression at the cell surface where they can be recognized by the TCR expressed by specific T cells. MHC molecules are structurally adapted to bind peptides. They exhibit a peptide-binding groove made of eight antiparallel β sheets as a floor and two antiparallel α helices as sides (3). Hydrogen bonding between the peptide backbone and the MHC class II (MHC II) molecule is thought to be one of the major driving binding forces and selective binding of peptides is governed by the ability of several pockets to accommodate some side chains of the peptide.

Correspondence to: D. Gerlier

Transmitting editor: A. McMichael

Received 8 May 1998, *accepted* 7 August 1998

A knowledge of the rules governing the association of peptides with MHC molecules would allow the accurate prediction of the peptide sequences derived from a given self or non-self protein responsible for the immune response toward an autoantigen or a pathogen. The sequencing of peptides eluted from MHC class I molecules and tedious binding tests with series of synthetic peptides have resulted in the identification of several MHC allele-specific peptide-binding motifs which could be used to accurately predict the majority of good peptide binders to a given MHC I molecule (4). When the same strategy was applied to MHC II, peptide-binding motifs were hardly defined and their predictive value was low. Indeed, the three-dimensional (3-D) structure of peptide–MHC I complexes shows that the binding groove of MHC I molecules is closed at both ends and can accommodate peptides with a size limited to 8–11mers (5). The 3-D structure of MHC II molecules differs from that of MHC I by having a peptide-binding groove opened at both ends thus allowing binding of peptides with size varying from 10 to >24mer peptides (3,4). One way to be able to predict which peptides could bind to a given MHC II molecule is to know how a defined peptide actually binds to this MHC II molecule. This information would be useful for uncovering the MHC allele-specific peptide-binding motif. The polymorphism of MHC molecules implicates individual determination of every MHC allele-specific peptide-binding motif. A convenient method would be the accurate modeling of any peptide–MHC II complex.

A 3-D model of the murine MHC II I-A^k protein complexed with the hen egg white lysozyme HEL[52–61] peptide was built by its alignment on the 3-D structure of the human MHC II HLA-DR1 complexed with the influenza virus hemagglutinin HA[306–318] peptide (6). This model was chosen because it was recognized by two distinct T cells and no I-A^k-specific peptide-binding motifs could be inferred from natural peptides eluted from I-A^k molecules (7). The model was refined by modeling 15 peptides derived from HEL[52–61] by single substitution and evaluating the relevance of each model with the ability of these peptides to be recognized by two distinct T cell hybridomas. The model appeared consistent with the biological data and current knowledge of peptide–MHC II 3-D structures. Furthermore, it was in good agreement with the 3-D structure of the HEL[50–62]–I-A^k complex solved and published while writing this manuscript (8).

Methods

Synthesis

The synthesis of peptides listed in Table 2 was carried out by standard Fmoc, solid-phase chemistry as previously reported (9,10). All peptides were purified by HPLC, and their homogeneity assessed by analytical HPLC and fast atom bombardment or electrospray mass spectrometry.

Bio-assays

Anti-HEL I-A^k-restricted 2A11 and 3A9 mouse T cell hybridomas recognizing the minimal HEL[52–61] peptide were co-cultured with H-2^k mouse CH27 B cells in the absence or presence of serial dilution of peptides as reported (10). The level of T cell stimulation was determined by measuring the IL-2 release using the CTL-L2 bioassay. The results were expressed as stimulating peptide concentration (SPC), where SPC was the graphically determined peptide concentration giving 50% of the maximal T cell stimulation. For competition experiments, serial dilutions of the competing peptide were mixed with a concentration of the HEL[52–61] peptide giving a sub-optimal T cell stimulation. Biological data obtained with four other peptides were borrowed from a previous work performed by Allen *et al*. (11).

Molecular modeling

Molecular modeling was performed with the SYBYL package version 6.2 (Tripos, St Louis, MO). Starting from the HA–HLA-DR1 crystal structure (6) deposited in the Protein Data Bank, code 1DLH, we constructed a model of the I-A^k molecule. The Needleman and Wunsch algorithm (12), available in the BIOPOLYMER module of SYBYL, was the chosen method used to make the sequence alignment. The sequence identity between HLA-DR1 and I-A^k is 60%. Two insertions, Glyα9a located in the first β strand and Lysβ84a, and two deletions β65 and β66 located in the α2 helix were introduced to align the I-A^k sequence according to the numbering of DR1 (3,6).

All hydrogen atoms were explicitly taken into account throughout the calculations. The force field used was Koll_all, a clone of Amber (13) by Tripos. A dielectric function of 4d was used in combination with a non-bonded interactions cut-off of 8 Å. Molecular lipophilicity potential (MLP) was calculated with the program CLIP 1.0 (Institute of Medicinal Chemistry, University of Lausanne) (14) and with the program MOLCAD (15). The solvent-accessible surface was calculated with the program MOLSV (QCPE no. 509) using a 1.5 Å radius probe. All calculations were performed on Origin 2000 (R10000), O_2 (R5000) and Indy (R4400) Silicon Graphics workstations.

Energy minimization was performed in two steps beginning with 1000 iterations using the steepest descent algorithm followed by complete minimization using the conjugate gradient algorithm until the r.m.s. gradient of the potential energy was <0.1 kcal/mol.Å^2. Molecular dynamics were performed at 300 K with a time step of 1 fs. Initial velocities were attributed according to the Bolzmann distribution and readjusted every 25 fs. Coordinates and velocities were collected every 0.5 ps.

For the constrained complexes, a 10 ps molecular dynamics simulation at constant temperature was performed on the minimized complex. Constraints were fixed on each of the 10 conserved hydrogen bonds by setting the force constant to 30 kcal/mol.Å^2. I-A^k backbone was set as an aggregate. Initial velocities were also taken from a Boltzmann distribution and readjusted every 25 fs. The system was progressively heated to 300 K during the first 1 ps. Then the temperature was being held at 300 K for the rest of the simulation by coupling the system to a heat bath (16) using a temperature coupling constant of 0.1 ps. The SHAKE algorithm (17) was used to constrain all bond lengths to their equilibrium values with a bond length tolerance of 2.5×10^{-4}. Coordinates, energies and velocities were collected and saved every 100 or 1000 fs for the longest dynamics. The energy minimization was performed the same way as above.

In order to save computational time, only the α_1 and β_1

Table 1. Two possible alignments of the HEL[52–61] peptide sequence (second and third lines) onto that of HA[306–318] peptide (first line) using the first anchor Tyr residue of the latter peptide bound to HLA-DR1

```
Pro–Lys–Tyr–Val–Lys–Gln–Asn–Thr–Leu–Lys–Leu–Ala–Thr
    Asp–Tyr–Gly–Ile–Leu–Gln–Ile–Asn–Ser–Arg
        Asp–Tyr–Gly–Ile–Leu–Gln–Ile–Asn–Ser–Arg
```

The residues used for the alignments are underlined.

domains of the MHC II molecule were taken into account. This approximation was shown not to alter the accuracy of molecular dynamic simulations because only limited interactions exist between the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ domains that do not significantly influence the shape of the peptide-binding groove (18).

Results

Modeling of HEL[52–61]–I-A^k complex

In the complex HA–HLA-DR1, the Tyr308 of the HA peptide is the primary anchor residue and is buried in a large hydrophobic pocket named P1 (6). In order to establish the existence of this pocket in the model of the I-A^k-binding cleft, the Connolly molecular surface (19) was calculated and confirmed the existence of the P1 pocket in the I-A^k binding site. The MLP was calculated on this surface with the CLIP package (14) in order to establish the physico-chemical properties of this pocket in our model. The MLP revealed that this pocket is globally lipophilic and could very well accept the first anchor of peptide HEL[52–61]. Among the first residues of HEL[52–61], Asp52 and Tyr53 have side chains large enough to act as an anchor and were examined as potential anchor into this P1 pocket. The sequence alignments of the two antigenic peptides HA and HEL suggested the superimposition of Tyr53 of HEL with Tyr308 of HA (Table 1). This resulted in a model in which the Tyr53 would be deeply buried into the MHC groove. However, both 3A9 and 2A11 TCR were exquisitely sensitive to minor changes in the side chain of residue 53 such as Phe to Tyr exchange (Table 2 and unpublished data) and diodo-Tyr or 3-NO_2-Tyr substitution (20), strongly indicating that the amino acid in position 53 should be pointing outside the binding cleft rather than being embedded into the P1 pocket of the I-A^k molecule. Furthermore, modeling of 15 HEL[52–61] peptides with single amino acid substitutions resulted in structures inconsistent with their biological properties (data not shown).

Therefore, we aligned the Asp52 of HEL with Tyr308 of HA (Table 1) and kept as a starting geometry of the peptide that of the HA peptide mutated on the adequate positions with the original orientations of the side chains. A working model was readily obtained after a 10 ps molecular dynamics of the complex with only the protein backbone fixed as an aggregate. The side chains residues of Asp52, Ile55, Gln57 and Ser60 stayed in P1, P4, P6 and P9 pockets respectively, where they were originally buried, and the most exposed ones Tyr53, Leu56, Asn59 and Arg61, kept pointing out. During the 10 ps molecular dynamics of the complex with HEL[52–61] or with any of the 15 other peptides, we noticed a shift of the peptide backbone inside the cavity compared to the starting one, particularly on the first three residues. We also noted that several original conserved hydrogen bonds existing between the peptide backbone and the protein side chains residues had disappeared. From the available 3-D structure of peptide–MHC II, 10 hydrogen bonds tend to be conserved (21,22). Then a 120 ps molecular dynamics with fixed constraints on these bonds by applying a minimal 30 kcal/mol.Å^2 force constant and with the protein backbone kept fixed as an aggregate was performed on the reference structure, HEL[52–61]–I-A^k. We collected some structures at different times and compared them to the starting one. The 10 hydrogen bonds were conserved and almost no shift of the peptide backbone was observed while the side chains positions were unchanged. It appears that the structure collected at 70 ps corresponds to an equilibrium one (Fig. 1) at which the r.m.s. distance of the α backbone of the peptide is 0.54 Å. This particular structure was submitted to another 20 ps molecular dynamics without constraints to verify that the complex remained stable and indeed only regular oscillations around positions 53, 54 and 55 of the peptide backbone were observed (r.m.s. distance is 0.67 Å).

Viewing the modeled peptide–MHC II complex from above, Arg61 was flanked by the α helix of the β_1 domain of I-A^k (Fig. 2). The polar head of the arginine went beyond the groove. The profile view with the P1, P4, P6 and P9 pockets filled with Asp52, Ile55, Gln57 and Ser60 respectively is shown on Fig. 2. Asp52 gave a hydrogen bond with Argα53. The other neighbors were Leuα31, Pheα32, Trpα43, Pheα54, Asnβ82 and Thrβ85, giving a lipophilic character to this pocket. Despite being charged, the Asp52 residue was able to fit inside this hydrophobic P1 pocket because of its strong interaction with Argα53 and its non-bulky side chain compared to a Tyr residue as observed in the HA[306–318]–HLA-DR1 complex. Ile55 filled the P4 pocket, and was in contact with the Pheβ11, Proβ13, Gluβ74 and Valβ78 residues. Gln57 filled the P6 pocket, and interacted with Asnα62, Asnα69, Hisβ9, Pheβ11, Tyrβ30 and Thrα65. Despite the presence of five residues able to form hydrogen bonds with the peptide, none was observed. The P9 pocket filled with Ser60 was formed by Asnα69, Ileα72, Leuα73, Argα76, Tyrβ37, Aspβ57 and Trpβ61 residues. Its bottom had a lipophilic character whereas its sides were rather polar. In addition to these four main pockets, the relatively hydrophobic P7 pocket was filled with the side chain of Ile58. This pocket pointed laterally, Ile58 interacting with Trpβ61, Tyrβ47 and Glnβ64.

The 10 conserved hydrogen bonds between I-A^k and the α backbone of the HEL peptide are represented in Fig. 2. The α backbone of Argα53 forms a bond with Asp52, Asnβ82 makes bidentate hydrogen bonds to Tyr53, Asnα62 forms a bond with Ile55 and another one with Gln57, Asnα69 forms a bond with Ile58 and another one with Ser60, Trpβ61 forms a bond with Asn59, Aspβ57 forms a bond with Arg61, and Argα76 makes a bond to Arg61. Nine of these bonds involve I-A^k residues that are conserved in most of human or mouse MHC II and engaged in hydrogen bonding with peptides in known 3-D structures (6,8,21–26) Aspβ57 is the only polymorphic residue of I-A^k to form a bond with the peptide backbone. Gly54 and Leu56 are the only residues that do

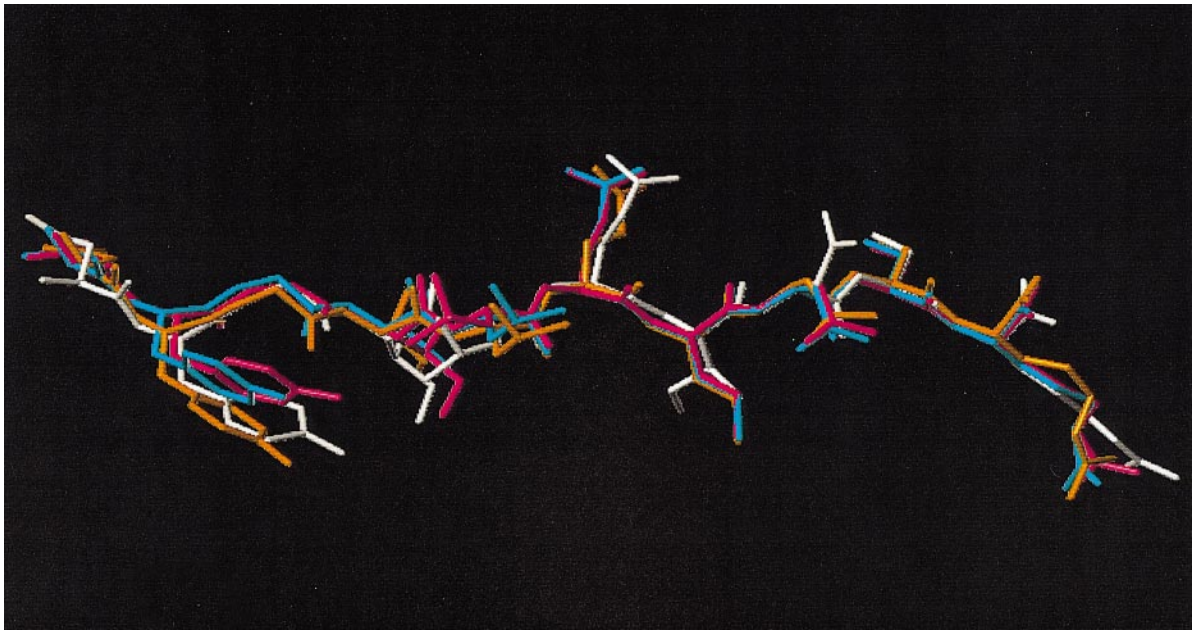

Fig. 1. Comparison of the HEL[52–61] peptide geometry at 0 ps (white), 30 ps (magenta), 70 ps (cyan) and 120 ps (orange).

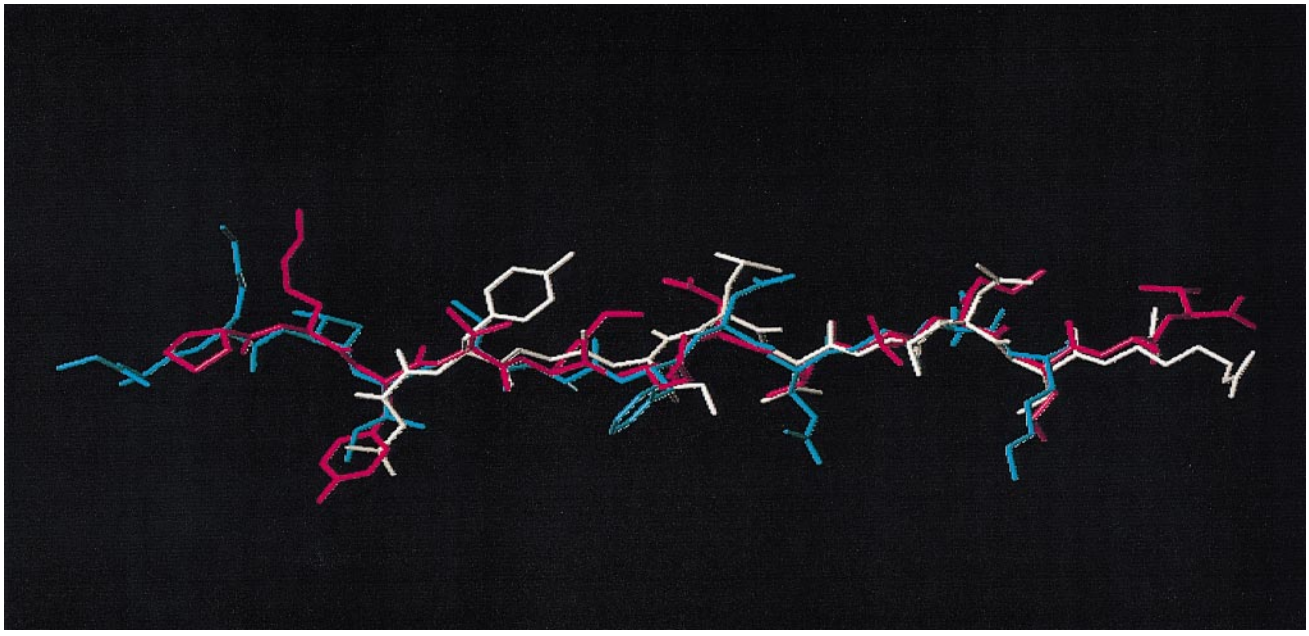

Fig. 3. Comparison of three peptide–MHC II complexes: the α backbone of HA (magenta) from HA–HLA-DR1 is superimposed on Hb (cyan) from Hb–I-E^k and HEL(white) from HEL[52–61]–I-A^k complexes

not make bonds with I-A^k. The fact that these hydrogen bonds do not involve HEL[52–61] side chains may indicate that this bonding network could be common to all peptides able to bind to I-A^k.

In order to check that the applied force constants on the 10 conserved hydrogen bonds are weak and do not interfere or force the peptide to adopt a specific position, we finally performed two 100 ps molecular dynamics on the 70 ps reference structure: one without constraining the hydrogen bonds but keeping the protein backbone set as an aggregate, the second one without any constraints and any aggregate. After the molecular dynamics with the I-A^k backbone kept as an aggregate, no major change in the structure was observed with the HEL peptide staying bound to the groove with the 10 conserved hydrogen bonds. Surprisingly, after the molecular dynamics on the unconstrained and relaxed HEL-I-A^k complex, the complex did not fall apart, but the β sheet and α helices of I-A^k were conserved as above with the HEL

Table 2. Structure and T cell-stimulating activity of HEL[52–61]-related peptides

Peptide	HEL residue number										2A11 SPC[a] (nM)	3A9 SPC[a] (nM)	Reference
	52	53	54	55	56	57	58	59	60	61			
52–61	Asp	Tyr	Gly	Ile	Leu	Gln	Ile	Asn	Ser	Arg	5	600	10
53-Phe	Asp	**Phe**	Gly	Ile	Leu	Gln	Ile	Asn	Ser	Arg	NS[b]	1000	this paper
53-Ala	Asp	**Ala**	Gly	Ile	Leu	Gln	Ile	Asn	Ser	Arg	NS	NS	this paper
53-Gln	Asp	**Gln**	Gly	Ile	Leu	Gln	Ile	Asn	Ser	Arg	NS	NS	this paper
54-Ala	Asp	Tyr	**Ala**	Ile	Leu	Gln	Ile	Asn	Ser	Arg	5	NS	11
56-Nva	Asp	Tyr	Gly	Ile	**Nva**	Gln	Ile	Asn	Ser	Arg	1000	NS	10
56-Val	Asp	Tyr	Gly	Ile	**Val**	Gln	Ile	Asn	Ser	Arg	NS	NS	11
56-Nle	Asp	Tyr	Gly	Ile	**Nle**	Gln	Ile	Asn	Ser	Arg	NS	NS	11
56-Phe	Asp	Tyr	Gly	Ile	**Phe**	Gln	Ile	Asn	Ser	Arg	NS	NS	this paper
57-Glu	Asp	Tyr	Gly	Ile	Leu	**Glu**	Ile	Asn	Ser	Arg	1000	800	10
57-Ala	Asp	Tyr	Gly	Ile	Leu	**Ala**	Ile	Asn	Ser	Arg	NS	NS	11
58-Val	Asp	Tyr	Gly	Ile	Leu	Gln	**Val**	Asn	Ser	Arg	150	6000	10
58-Nva	Asp	Tyr	Gly	Ile	Leu	Gln	**Nva**	Asn	Ser	Arg	30	NS	10
58-Leu	Asp	Tyr	Gly	Ile	Leu	Gln	**Leu**	Asn	Ser	Arg	200	NS	10
58-Phe	Asp	Tyr	Gly	Ile	Leu	Gln	**Phe**	Asn	Ser	Arg	NS	NS	10
52–60	Asp	Tyr	Gly	Ile	Leu	Gln	Ile	Asn	Ser		15	NS	this paper

[a]Half maximal stimulating peptide concentration for 2A11 or 3A9 T cell hybridomas
[b]Non-stimulating at the highest peptide concentration tested <100 μM. All peptides unable to stimulate both T cell hybridomas have been found to bind to I-A^k in competition experiments with HEL[52–61] peptide.

peptide similarly bound to the groove (r.m.s. distance of the peptide α backbone = 1.17 Å). Only some little modifications on the backbone looking like unfolding occurred in the loops that tie the β sheets and link them to the α helices. Also, the peptide backbone has slightly shifted within the groove.

On the profile view of the complex, the peptide backbone adopted an extended polyproline II-like conformation which superimposed nicely with the peptide backbone of HA[306–318] (r.m.s. distance of the peptide α backbone = 0.77 Å) (6) and mouse hemoglobin Hb[68–76] (r.m.s. distance of the peptide α backbone = 0.82 Å) (21) complexed to human HLA-DR1 and the mouse I-E^k MHC II molecules respectively (Fig. 3).

The solvent-accessible area of each peptide side chain either buried or surface exposed was calculated (19) and the corresponding graphic representation is shown on Fig. 4(A). Notably, Asp52, Gln57 and Ser60 were completely buried into the P1, P6 and P9 pockets, 92,7% of Ile55 and 80% of Ile58 were buried into the P4 and P7 pockets respectively. Most (44–61%) of the side chains of Tyr53, Leu56, Asn59, Arg61 and 20% of Ile58 were exposed and should be available for recognition by the TCR.

Modeling of HEL[52–61] peptide variants and relation to their biological properties

From the reference structure taken at 70 ps, we built by mutation the other 15 peptides listed on Table 2. After a short minimization, we achieved a 10 ps molecular dynamics on each complex, keeping the hydrogen bonds slightly constrained. For all peptides, we noticed the conservation of the side chains position of residues 52, 55, 57 and 60 in the four anchor pockets and those of residues 53, 56, 59 and 61 exposed at the complex surface and available for interaction with a TCR. The individual peptide residue surface areas exposed at the surface of the molecule and buried into the groove were distributed over a rather narrow range (Fig. 4B). We noted that none of the three fully buried residues at P1, P6 and P9 were ever exposed, and only a slight variation in the exposure of the residues embedded in P4 was observed. Comparison of the model of each substituted peptide with HEL[52–61] (Figs 5 and 6) and their respective recognition by a TCR (Table 2) revealed no major inconsistencies.

The models of the 10 substituted peptides showing poor or no recognition by both 2A11 and 3A9 TCR fitted with the expected major change in the exposed side chains. The non-conservative Tyr53Ala and Tyr53Gln substitution resulted in a single major change in the exposure of this residue, indicating that both TCR interact somehow with Tyr53 (53-Ala and 53-Gln, Fig. 5). The side chain of every substitution of the Leu56 residue tested (56-Nva, 56-Nle, 56-Phe and 56-Val) remained similarly exposed (Fig. 5) indicating that the side chain of Leu56 comes in direct contact with both TCR and plays a critical role in TCR recognition. We noted that 56-Nva, 56-Nle and 56-Val substitutions resulted also in a downward exposure of the Tyr53 (–43, –31 and –28 Å^2 respectively) and Ile58 (–18, –18 and –23 Å^2 respectively) side chains, and in the case of 56-Val, a 43 Å^2 upward exposure of the Arg61 side chain. These changes in other residues accessible to the TCR might also participate in the lost of recognition of these peptides by 2A11 and 3A9 T cells. When substituting Gln by Ala at position 57, the side chain of Ala was modeled to remain fully buried. The only major change in side chain exposure was the 25 Å^2 upward shift of the Ile58 side chain, suggesting that although this residue remained mostly buried, its small exposed area would play a major role in TCR recognition. Strikingly, changing the nature of the Ile58 side chain to that of Phe also resulted in the loss of T cell recognition with an exposure of the side chain of Phe as limited as that of the wild Ile residue but in a –32 Å^2 downward shift of the exposed Tyr53 and +22 Å^2 upward shift of the

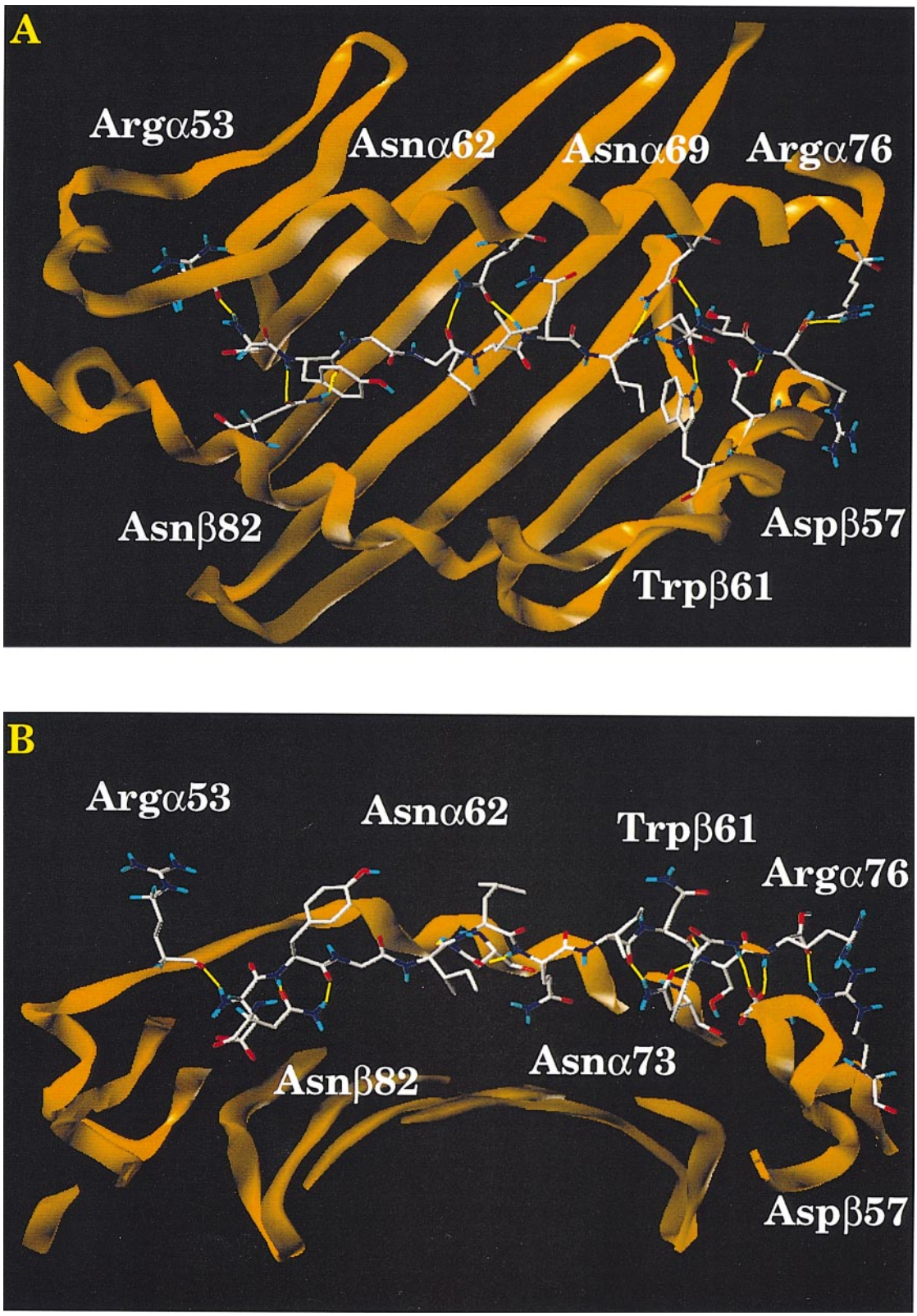

Fig. 2. View of the HEL[52–61] peptide in the conformation of the final model. Hydrogen bonds between the α backbone of the peptide and the protein are indicated in yellow. The α backbone of I-A^k is represented by the orange ribbon. (A) View from above. (B) Side view with the orange ribbon being cut (Z-clipping) allowing HEL[52–61] to appear on the foreground.

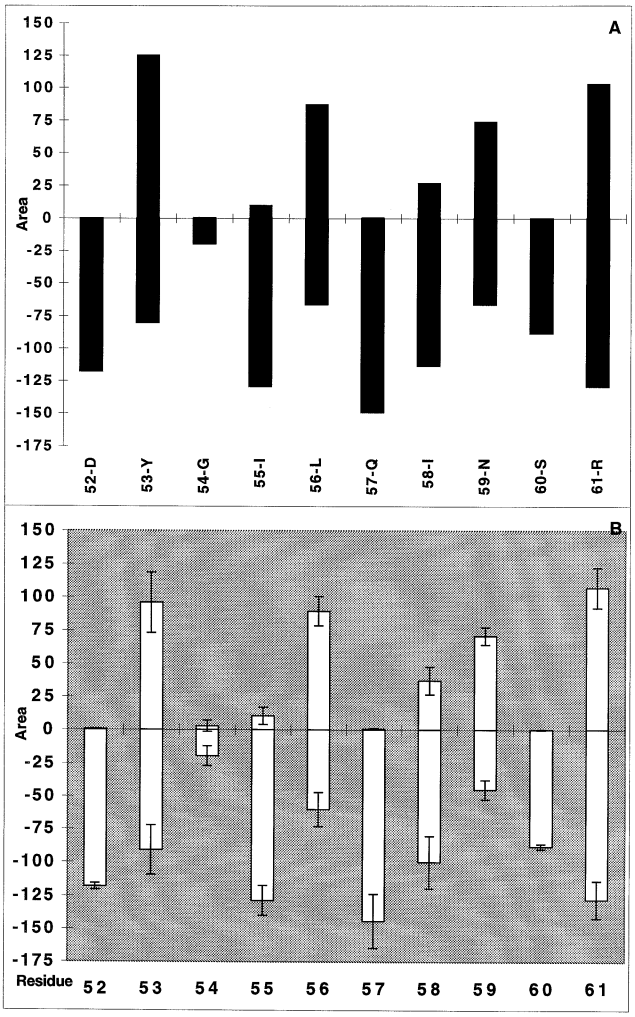

Fig. 4. Exposed and buried areas in Å^2 of peptide residues. (A) Exposed and buried areas of HEL[52–61] peptide associated to I-A^k molecule. (B) Mean and SD values of exposed and buried areas of each residue from 16 HEL[52–61]-related peptides

normally buried Ile55 (58-Phe, Fig. 5). The deleterious effect for TCR recognition of the substitution of Ile58 by the closely related Leu was correlated with modeled downwards shift of the side chains of the exposed Tyr53 (–26 Å^2) and upward shift of the exposed Arg61 (+22 Å^2) (58-Leu, Fig. 5).

Five single amino acid substitutions of HEL[52–61] resulted in the specific alteration of the recognition by only one T cell hybridoma (53-Phe, 54-Ala, 57-Glu, 58-Nva and 52–60, Table 1). Two of these directly affected the nature of one of the three most exposed residues. A change of Tyr53 to Phe corresponded to the loss of a single hydroxyl which was predicted to be exposed by our model and therefore to account for the specific loss of recognition by the 2A11 TCR (53-Phe, Fig. 6). However, some changes in the exposure of other side chains were observed with a major upward shift of the Ile58 (+32 Å^2) residue which might contribute to the recognition process. The lack of the strongly exposed Arg61 was likely responsible for the loss of the recognition by 3A9 TCR, though associated changes in the exposure of Tyr53 (–37 Å^2) and Ile58 (+18 Å^2) might also contribute to this effect (52–60, Fig. 6). As not being affected by the loss of the exposed and positively charged Arg61, the 2A11 TCR should not interact with the C-end of the HEL[52–61] peptide. When a CH_3 side chain was introduced at the 54 position, it was exposed and downward shifts of the three most exposed residues, Tyr53 (–20Å^2), Leu56 (–19 Å^2) and Arg61 (–19 Å^2), were predicted by the modeling (54-Ala, Fig. 6). The exquisite loss of recognition by 3A9 but not by 2A11 TCR indicated that those changes are subtly affecting the interaction of the HEL[52–61]–I-A^k complex with a TCR. The substitution of the fully buried Gln57 residue by its close relative Glu was predicted to result in a single –50 Å^2 downward shift of the Tyr53 (57-Gln, Fig. 6). This change could well account for the specific loss of recognition by the 2A11 TCR in agreement with its exquisite sensitivity to the Tyr53–Phe substitution (see above). The exchange of the mostly buried Ile58 by Nva was modeled to induced an upward shift of Leu56 (+21 Å^2) and a limited downward shift of Tyr53 (–16 Å^2) without apparent change in the exposure of Nva58 side chain (58-Nva, Fig. 6). These differences went undetected by the 2A11 TCR but preclude recognition by the 3A9 TCR. Together with the deleterious effect on the recognition by both TCR of substituting Ile58 by Leu or Phe and the deleterious upward shift of Ile58 in 57-Ala peptide (see above), this suggests that the limited exposed area of Ile58 could come into contact with the TCR.

Therefore when taking into account the exposed surface area of HEL[52–61]-derived peptides in this model, consistent relationships with the biological data could be made, though this parameter is perhaps not sufficiently discriminative to predict the precise interaction of a TCR with a peptide embedded in MHC II.

Comparison of a model of HEL[52–61]–I-A^k and the 3-D structure of HEL[50–62]–I-A^k

After completion of our modeling of the HEL[52–61]–I-A^k complex, the crystal structure of the closely related HEL[50–62]–I-A^k complex was published by Fremont *et al.* (8). When comparing the two structures, the correspondence was remarkably high (r.m.s. distance of the complex α backbone = 1.5 Å) (Fig. 7A). One discrepancy was seen in the β_1 strand of the I-A^k α chain, where the crystal structure revealed an upward small bulge due to the insertion of an α9a residue. Accordingly, and as predicted by Fremont et al., the hydrogen bonds between Aspα4 and Asnβ19 and between Hisα5 and Aspα27 were absent in our model. On the view from the top of the bound peptides (Fig. 7B), their backbones were nicely superimposed and, the side chains in position P1, P4, P6 and P9 were superimposed very closely with little short angle rotations. The other side chains have a different torsion angle. On the side view (Fig. 7C), one observes a good superimposition on the backbone except on the N- and C-ends of the shorter HEL[52–61]. This difference may be due to the presence of extra Ser–Thr and Trp residues in N- and C-ends of HEL[50–62] respectively. Our model correctly predicts the exposed and buried residues with close areas values for each residue (Fig. 8). Moreover, six out of 10 hydrogen bonds (Argα53–Asp52, bidentate Asnβ82–Tyr53,

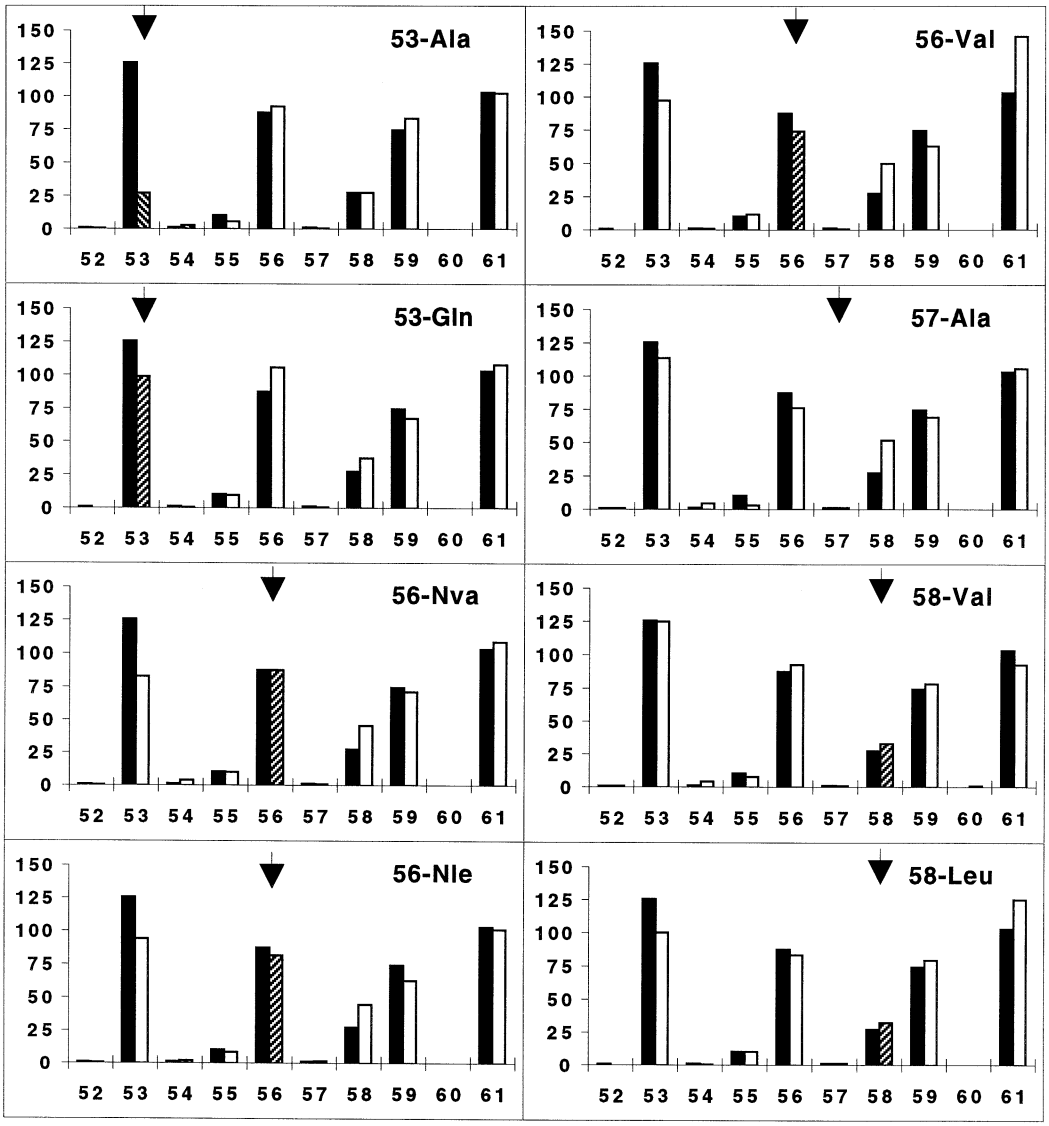

Fig. 5. Comparison of the surface exposed areas of HEL[52–61] (black columns) with those of HEL peptide with single amino acid substitution (white columns) deleterious for their recognition by both 2A11 and 3A9 T cells but still able to bind to the I-A^k molecule. The substitution is indicated by an arrow and shading of the corresponding column.

Asnα62–Gln57, Asnα69–Ile58 and Asnα69–Ser60) were correctly predicted.

Discussion

How do previous reports dealing with the identification of anchor and TCR contact residues of HEL[52–61]-related peptides bound to I-A^k molecules fit with our model? In their initial work based on T cell recognition and peptide competition bioassays using Ala-substituted HEL[52–61], Allen *et al.* have proposed Asp52, Ile58 and Arg61 as anchors residues, and Tyr53, Leu56 and Gln57 as T cell contact residues (11). Our model yields a fit with four residues, Asp52, Tyr53 and Leu56 and Ile58. The discrepancy with the other two positions could be explained by an indirect detrimental effect of Ala61 substitution on the binding of the peptide to I-A^k and, for Ala57, by the change of the exposure of side chains in other positions (see the dramatic change in Tyr53 exposure of the Glu57 substituted peptide) as previously proposed (10). Curiously, the substitution of the slightly less buried P7 Ile58 with Ala strongly decreased peptide binding to I-A^k, while the substitution of the deeply buried P4 Ile55 by Ala affected neither the MHC II binding nor the recognition by the 2A11 TCR (10), underlining an ususpected major role of P7 anchoring in HEL[52–61] binding to I-A^k. Using poly-Ala peptides, Asp52 has been described as a potent anchor residue (27) and fits with our model. Interestingly, Sant' Angelo *et al.* have identified P1, P4, P6 and P9 as anchor residues for the HRGAIEWEG peptide bonded to I-A^k, and P2, P5 and P8 as TCR contact residues (28). By an alignment procedure, they have proposed Asp52, Ile55, Gln57 and Ser60 as anchor residues for the HEL[46–61] peptide, which nicely agrees

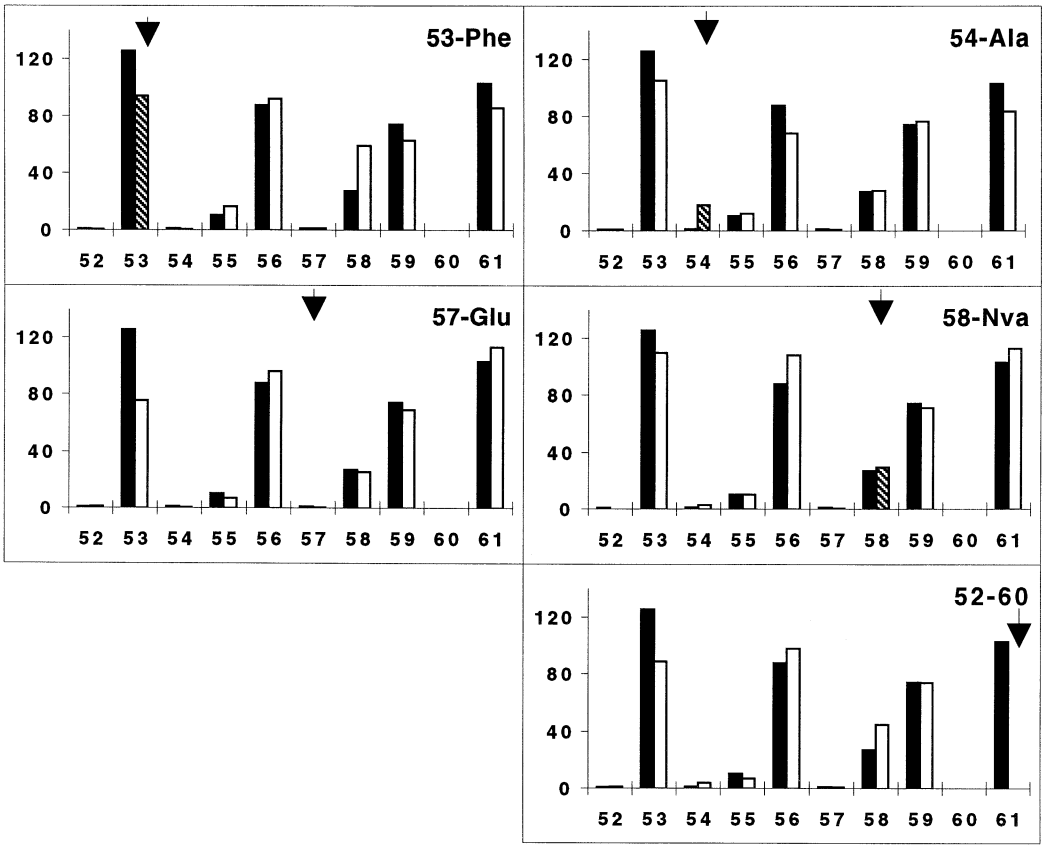

Fig. 6. Comparison of the surface exposed areas of HEL[52–61] (black columns) with those of HEL peptide with single amino acid substitution (white columns) deleterious for their recognition by the 2A11 (left histograms) or 3A9 (right histograms) T cells. The substitution is indicated by an arrow and shading of the corresponding column.

with our model. Using a mAb specific for the HEL[48–62]–I-A^k complex, Asp52 was confirmed as a main anchor residue and Tyr53 as being surface exposed (29). The structure of I-A^k associated with CLIP peptide derived from the invariant chain has been modeled, and anchor residues were predicted to be in P1, P3, P6, P7 and P9 position with a crucial involvement of CLIP Met91 as peptide primary anchor in P1. The exposed residues were located in P2, P5 and P8 positions (30). The proposed P7 anchor for CLIP is in agreement with having the corresponding HEL Asn58 mostly buried in our model. The major discrepancy with our model is the second anchor residue we located in the P4 and not the P3 position. Since the corresponding sequence at P3–P4 for HEL and CLIP are opposite in the order of minimal and bulky side chains within their sequence (Gly54–Ile55 for HEL and Met93–Ala94 for CLIP) this could have biased one of the modeling procedures. Alternatively, a local plasticity of the MHC II molecules might allow alternative P3–P4 anchoring. Indeed, the 3-D structure of the CLIP–HLA-DR3 revealed that CLIP residues interact not only with the pockets P1, P3, P6, P7 and P9 as predicted (30), but also with the non-predicted pocket P4 (26).

Our model compares favorably with the recently solved 3-D structure of HEL[50–62]–I-A^k complex (8) with accurate prediction of the exposed and buried areas of each peptide residue and six out of 10 hydrogen bonds linking the peptide backbone to the MHC molecule. To our knowledge, the CLIP–HLA-DR3 complex is the only other peptide–MHC II complex which has been modeled (30) before the crystal 3-D structure had been obtained (26). The model obtained using a similar molecular modeling strategy but different computer programs correctly predicted the anchor and exposed residues. No information on hydrogen bonds and relative exposed/embedded residue areas were given in this model, and no direct comparison with the crystal structure were reported. How can we explain the few discrepancies in our own study between the HEL[52–61]–I-A^k model and the HEL[50–62]–I-A^k 3-D structure ? The solvation was not taken into account due to prohibitive calculation time and we failed to model the specific bulge in the first strand of the peptide-binding platform of the I-A^k α chain. Moreover, the HEL peptide size present in the crystal and model structures differs (13 and 10mers respectively), and the presence of extra residues on both the N- and C-end of the HEL peptide could explain the shift observed in the extremities of the HEL[52–61] peptide backbone. Indeed, the elongation of this peptide enhances the I-A^k-restricted presentation to T lymphocytes (10). Moreover, the invariant chain Ii[81–90] sequence adjacent to the cleft-bound CLIP 91–99 peptide has a dual enhancing (31) and dissociating (32) effect on the peptide–MHC II complex, which

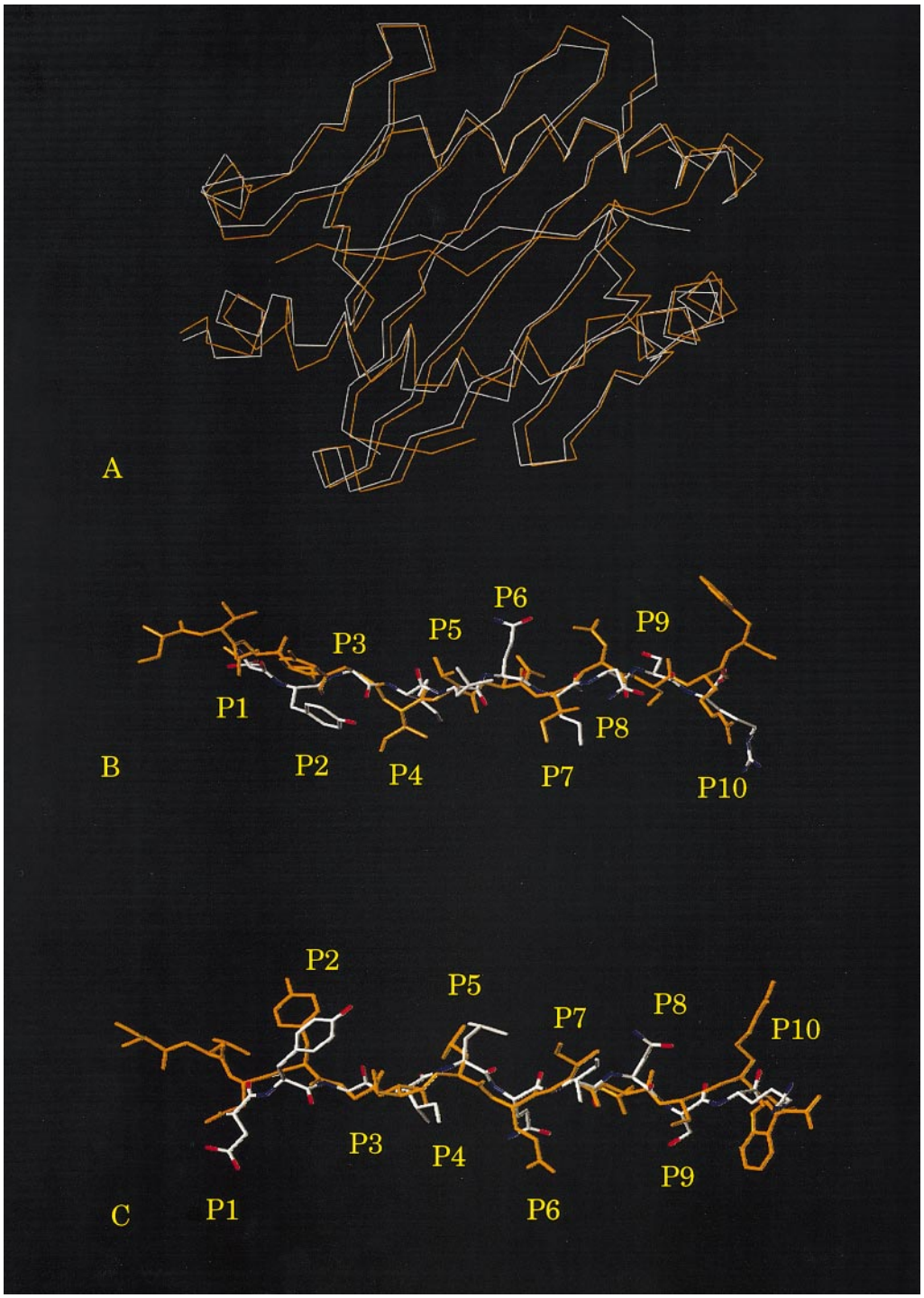

Fig. 7. Comparison of the HEL[52–61]–I-A^k model with the HEL[50–62]–I-A^k crystal structure. (A) View from the top of the HEL[52–61]–I-A^k (white) and HEL[50–62]–I-A^k (orange) α backbone. (B) View from the top of the HEL[52–61] (white for carbon, blue for nitrogen and red for oxygen) and the HEL[50–62] (orange) α backbone. (C) Side view of the two peptides.

suggests a second peptide-binding site being located close to the P1 pocket and having an allosteric effect on the peptide–MHC II groove interaction (33). Last but not least, homology modeling, which relies on the hypothesis that the general structure and folding remain the same inside one family of proteins, has a known limitation (34). In particular, the position of side chains may be trapped into local minima. However, the use of molecular dynamics simulation allows to overcome some of these difficulties. In our case, the comparison between the X-ray structure and the theorical model

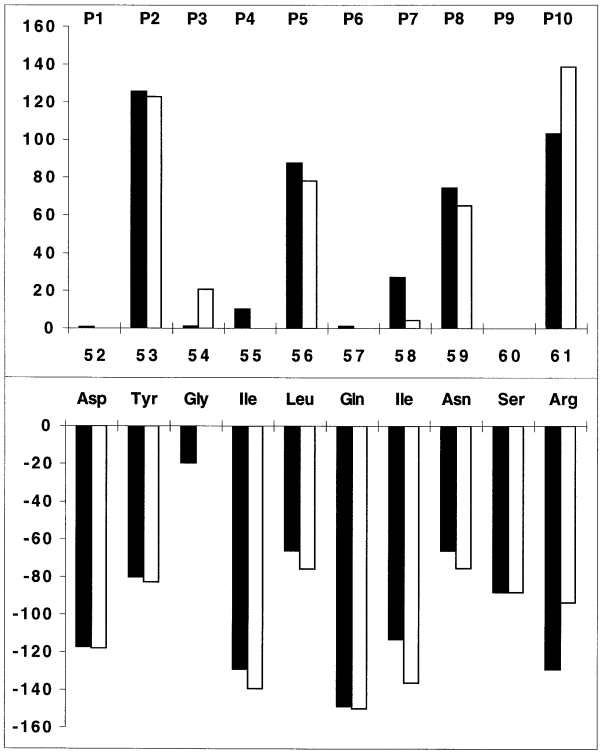

Fig. 8. Comparison of the modeled surface exposed (upper histogram) and buried (lower histogram) areas of HEL[52–61] peptide residues (close columns) with those observed in the 3-D structure of the HEL[50–62]–I-A^k complex (8) (open columns). Ordinate scale in Å^2.

remains good (r.m.s. = 1.5 Å), even if some local details cannot be predicted accurately. This discrepancy can result from the different environments leading to variable distortion in the modelization (gas phase) and in the crystal X-ray structure. Indeed, a crystal structure likely describes only one among the actual structures encountered during the molecular dynamics.

In conclusion, using very simple computations without taking solvation into account, we have succeeded in modeling the I-A^k molecule complexed to the HEL[52–61] peptide and related HEL[52–61] peptides with single amino acid substitution in good qualitative agreement with its recognition by two specific T cells and the recently described 3-D crystalline structure of HEL[50–62]–I-A^k complex. Thus we should not have to wait too long before being able to accurately model any peptide–MHC II complex when further refined computational methods will be available. The prediction of the ability of a given peptide to bind to a given MHC II molecule may not be so close, owing to our inability to predict the initial steps required for the association of a peptide to MHC II molecules. Indeed, *in vitro*, long-lived peptide–MHC II complexes occurred after a preliminary step of a specific, rapidly formed, short-lived complexes and these two types of complexes differ in their structure (35). In addition, peptide binding to MHC II is optimized at acidic pH where it adopts a particular conformation (36). The recently described procedure of global minimization of energy interactions of an individual peptide residue for its accommodation into the MHC II pockets may help in circumventing this difficulty (37).

Acknowledgements

This work was supported in part by the Swiss National Science Foundation, the Ministère de l'Enseignement Supérieur et de la Recherche and the Association pour la Recherche contre le Cancer. The authors thank G. Deléage for his helpful critical reading.

Abbreviations

HA	hemagglutinin
HEL	hen egg lysozyme
Hb	hemoglobin
MHC II	MHC class II
MLP	molecular lipophilicity potential
SPC	stimulating peptide concentration
r.m.s.	root mean square

References

1 Germain, R. N. 1994. MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. *Cell* 76:287.
2 Watts, C. 1997. Capture and processing of exogenous antigens for presentation on MHC molecules. *Annu. Rev. Immunol.* 15:821.
3 Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L. and Wiley, D. C. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 364:33.
4 Rammensee, H. G. 1995. Chemistry of peptides associated with MHC class I and class II molecules. *Curr. Opin. Immunol.* 7:85.
5 Madden, D. R., Gorga, J. C., Strominger, J. L. and Wiley, D. C. 1991. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature* 353:321.
6 Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L. and Wiley, D. C. 1994. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 368:215.
7 Marrack, P., Ignatowicz, L., Kappler, J. W., Boymel, J. and Freed, J. H. 1993. Comparison of peptides bound to spleen and thymus class II. *J. Exp. Med.* 178:2173.
8 Fremont, D. H., Monnaie, D., Nelson, C. A., Hendrickson, W. A. and Unanue, E. R. 1998. Crystal structure of I-Ak in complex with a dominant epitope of lysozyme. *Immunity* 8:305.
9 Ettouati, L., Salvi, J. P., Trescol-Biémont, M. C., Walchshofer, N., Gerlier, D., Rabourdin-Combe, C. and Paris, J. 1996. Substitution of peptide bond 53–54 of HEL(52–61) with an ethylene bond rather than reduced peptide bond is tolerated by an MHC II restricted T cell. *Peptide Res.* 9:248.
10 Hernandez, J. F., Cretin, F., Lombard-Platet, S., Salvi, J. P., Walchshofer, N., Gerlier, D., Paris, J. and Rabourdin-Combe, C. 1994. Critical residue combinations dictate peptide presentation by MHC class II molecules. *Peptides* 15:583.
11 Allen, P. M., Matsueda, G. R., Evans, R. J., Dunbar, J. B., Marshall, G. R. and Unanue, E. R. 1987. Identification of the T-cell and Ia contact residues of a T-cell antigenic epitope. *Nature* 327:713.
12 Needleman, S. B. and Wunsch, C. D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443.
13 Weiner, S. J., Kollman, P. A., Nguyen, D. T. and Case, D. A. 1986. An all atoms force field for simulation of proteins and nucleic acids. *J. Comput. Chem.* 7:230.
14 Gaillard, P., Carrupt, P. A., Testa, B. and Boudon, A. 1994. Molecular lipophilicity potential, a tool in 3-D-QSAR. Method and applications. *J. Comput.-Aided Mol. Design* 8:83.
15 Heiden, W., Moeckel, G. and Brickmann, J. 1993. A new approach to analysis and display of local lipophilicity/hydrophilicity mapped on molecular surfaces. *J. Comput.-Aided Mol. Design* 7:503.

16 Ryckaert, J. P., Ciccotti, G. and Berendsen, H. J. C. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Phys.* 23:327.
17 van Gunsteren, W. F. and Berendsen, H. J. C. 1977. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* 5:1311.
18 Rognan, D., Scapozza, L., Folkers, G. and Daser, A. 1994. Molecular dynamics simulation of MHC-peptide complexes as a tool for predicting potential T cell epitopes. *Biochemistry* 33:11476.
19 Connolly, M. L. 1993. The molecular surface package. *J. Mol. Graphics* 11:139.
20 Allen, P. M., Matsueda, G. R., Haber, E. and Unanue, E. R. 1985. Specificity of the T cell receptor: two different determinants are generated by the same peptide and the I-A^k molecule *J. Immunol.* 135:368.
21 Fremont, D. H., Hendrickson, W. A., Marrack, P. and Kappler, J. 1996. Structures of an MHC class II molecule with covalently bound single peptides. *Science* 272:1001.
22 Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L. and Wiley, D. C. 1996. Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proc. Natl Acad. Sci. USA* 93:734.
23 Dessen, A., Lawrence, C. M., Cupo, S., Zaller, D. M. and Wiley, D. C. 1997. X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*04101) complexes with a peptide from human collagen II. *Immunity* 7:473.
24 Murthy, V. L. and Stern, L. J. 1997. The class II MHC protein HLA-DR1 in complex with an endogenous peptide: implications for the structural basis of the specificity of peptide binding. *Structure* 5:1385.
25 Scott, C. A., Peterson, P. A., Teyton, L. and Wilson, I. A. 1998. Crystal structure of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues *Immunity* 8:319.
26 Ghosh, P., Amaya, M., Mellins, E. and Wiley, D. C. 1995. The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. *Nature* 378:457.
27 Nelson, C. A., Viner, N. J., Young, S. P., Petzold, S. J. and Unanue, E. R. 1996. A negatively charged anchor residue promotes high affinity binding to the MHC class II molecule I-A(k). *J. Immunol.* 157:755.
28 Sant' Angelo, D. B., Waterbury, G., Preston-Hurlburt, P., Yoon, S. T., Medzhitov, R., Hong, S. C. and Janeway, C. A. 1996. The specificity and orientation of a TCR to its peptide–MHC class II ligands. *Immunity* 4:367.
29 Dadaglio, G., Nelson, C. A., Deck, M. B., Petzold, S. J. and Unanue, E. R. 1997. Characterization and quantitation of peptide–MHC complexes produced from hen egg lysozyme using a monoclonal antibody. *Immunity* 6:727.
30 Lee, C. and McConnell, H. M. 1995. A general model of invariant chain association with class II major histocompatibility complex proteins. *Proc. Natl Acad. Sci. USA* 92:8269.
31 Stumptner, P. and Benaroch, P. 1997. Interaction of MHC class II molecules with the invariant chain (81–90) region. *EMBO J.* 16:5807.
32 Kropshofer, H., Vogt, A. B., Stern, L. J. and Hammerling, G. J. 1995. Self-release of CLIP in peptide loading of HLA-DR molecules. *Science* 270:1357.
33 Gerlier, D., Trescol-Biémont, M. C., Ettouati, L., Paris, J. and Rabourdin-Combe, C. 1998. An accessory peptide binding site with allosteric effect on the formation of peptide–MHC II complexes. *C. R. Acad. Sci.* 321:19.
34 Whittle, P. J. and Blundell, T. L. 1994. Protein structure-based drug design. *Annu. Rev. Biophys. Biomol. Struct.* 23:349.
35 Sadegh-Nasseri, S., Stern, L. J., Wiley, D. C. and Germain, R. N. 1994. MHC class II function preserved by low-affinity peptide interactions preceding stable binding. *Nature* 370:647.
36 Boniface, J. J., Lyons, D. S., Wettstein, D. A., Allbritton, N. L. and Davis, M. M. 1996. Evidence for a conformational change in a class II major histocompatibility complex molecule occurring in the same pH range where antigen binding is enhanced. *J. Exp. Med.* 183:119.
37 Androulakis, I. P., NAyak, N. N., Ierapetritou, M. G., Monos, D. S. and Floudas, C. A. 1997. A predictive method for the evaluation of peptide binding in pocket 1 of HLA-DRB1 via global minimization of energy interactions. *Proteins* 29:87.