

# Molecular Modeling of an Antigenic Complex Between a Viral Peptide and a Class I Major Histocompatibility Glycoprotein

Didier Rognan,<sup>1</sup> Matthias J. Reddehase,<sup>2</sup> Ulrich H. Koszinowski,<sup>2</sup> and Gerd Folkers<sup>1</sup>

<sup>1</sup>*Institute for Pharmacy, University of Tübingen, 7400 Tübingen, Federal Republic of Germany, and*

<sup>2</sup>*Department of Virology, Institute for Microbiology, University of Ulm, 7900 Ulm, Federal Republic of Germany*

**ABSTRACT** Computer simulation of the conformations of short antigenic peptides (5–10 residues) either free or bound to their receptor, the major histocompatibility complex (MHC)-encoded glycoprotein H-2 L<sup>d</sup>, was employed to explain experimentally determined differences in the antigenic activities within a set of related peptides. Starting for each sequence from the most probable conformations disclosed by a pattern-recognition technique, several energy-minimized structures were subjected to molecular dynamics simulations (MD) either in vacuo or solvated by water molecules. Notably, antigenic potencies were found to correlate to the peptides propensity to form and maintain an overall  $\alpha$ -helical conformation through regular  $i, i+4$  hydrogen bonds. Accordingly, less active or inactive peptides showed a strong tendency to form  $i, i+3$  hydrogen bonds at their N-terminal end. Experimental data documented that the C-terminal residue is critical for interaction of the peptide with H-2 L<sup>d</sup>. This finding could be satisfactorily explained by a 3-D Q.S.A.R. analysis postulating interactions between ligand and receptor by hydrophobic forces. A 3-D model is proposed for the complex between a high-affinity nonapeptide and the H-2 L<sup>d</sup> receptor. First, the H-2 L<sup>d</sup> molecule was built from X-ray coordinates of two homologous proteins: HLA-A2 and HLA-Aw68, energy-minimized and studied by MD simulations. With HLA-A2 as template, the only realistic simulation was achieved for a solvated model with minor deviations of the MD mean structure from the X-ray conformation. Water simulation of the H-2 L<sup>d</sup> protein in complex with the antigenic nonapeptide was then achieved with the template-derived optimal parameters. The bound peptide retains mainly its  $\alpha$ -helical conformation and binds to hydrophobic residues of H-2 L<sup>d</sup> that correspond to highly polymorphic positions of MHC proteins. The orientation of the nonapeptide in the binding cleft is in accordance with the experimentally determined distribution of its MHC receptor-binding residues (agretope residues). Thus, computer simulation

was successfully employed to explain functional data and predicts  $\alpha$ -helical conformation for the bound peptide. © 1992 Wiley-Liss, Inc.

**Key words:** major histocompatibility complex, antigenic peptide, molecular dynamics

## INTRODUCTION

As opposed to B lymphocytes, T lymphocytes do not recognize a protein antigen in its native conformation, but recognize short peptides derived thereof by intracellular processing and bound to cell surface glycoproteins encoded by the major histocompatibility complex (MHC).<sup>1–5</sup> Thus, in molecular terms, antigen recognition by T lymphocytes can be described as binding of the T cell receptor (TCR) to a molecular complex composed of an antigenic peptide and an MHC molecule on the surface of a target cell. This concept is supported by experimental evidence: (1) Intracellular antigen processing can be bypassed by extracellular addition of synthetic peptides.<sup>6,7</sup> (2) Synthetic antigenic peptides bind directly to MHC molecules.<sup>8,9</sup> (3) Peptide ligand binding is involved in assembly and intracellular transport of MHC class I molecules<sup>10</sup> and (4) naturally processed antigenic peptides can be isolated biochemically from cells synthesizing the respective antigenic proteins.<sup>11,12</sup>

Much less is known, however, of the structural features that determine the antigenic quality of a free peptide and about the geometry of the peptide when bound to its MHC receptor. Crystal structures are now available for two human MHC (HLA) molecules and have revealed a conformational groove at the surface as a putative peptide binding site with

Received December 29, 1990; revision accepted June 24, 1991.

Address reprint requests to Gerd Folkers, Department Pharmazie, ETH Zentrum, Clausiusstrasse 25, CH-8092 Zürich, Switzerland.

Present address of D. Rognan and G. Folkers: Departement Pharmazie, ETH Zentrum, Clausiusstrasse 25, CH-8092 Zürich, Switzerland.

**TABLE I. Antigenic Properties for Cytolytic T Lymphocytes (Clone IE1)**

Name	Sequence	Recognition	Peptide concentration*
Nona	YPHFMPNTL	+	$10^{-9}$
Nonar	PHFMPNTLG	+	$10^{-3}$
Nonal	MYPHFMPNTN	-	
Octar	PHFMPNTL	+	$10^{-4}$
Octal	YPHFMPNTN	-	
Heptar	PHFMPNTN	+	$10^{-7}$
Heptal	YPHFMPNT	-	
Penta	HFMPNT	+	$10^{-3}$

\*Detection limit concentration (M) leading to lysis of H-2L<sup>d</sup> expressing target cells by CTL clone IE1.

specificity pockets for the interaction with amino acid side chains of the antigenic peptide.<sup>13,14</sup> Yet, to date, a peptide-MHC complex could not be visualized. There is currently a debate about the most likely conformation of antigenic peptides: arguments have been collected for proposing an  $\alpha$ -helix,<sup>15-17</sup> a  $\beta$ -sheet,<sup>18</sup> or an extended conformation incorporating a  $3_{10}$  helical turn.<sup>19</sup>

The present study combines data from immunological measurements with molecular mechanics and dynamics simulations as well as 3-D Q.S.A.R. studies in order to explain observed functional properties of an antigenic sequence and to propose a likely 3-D model of a peptide-MHC complex. The particular peptide used for this study is the nonapeptide <sup>168</sup>YPHFMPNTL<sup>176</sup> (one-letter code) that is derived from the 595 aa immediate-early (IE) protein pp89 of murine cytomegalovirus.<sup>20,21</sup> This peptide is the optimal ligand bound to murine class I MHC molecule L<sup>d</sup> (H-2L<sup>d</sup>) and recognized by the TCR of the cytolytic T lymphocyte (CTL) clone IE1.<sup>22</sup>

A valid structural model must explain known functional characteristics of an antigenic peptide. We show here that a  $\alpha$ -helical conformation of both the free and the bound IE1 nonapeptide could explain the different antigenic potencies of a set of peptide analogs as well as the experimentally determined distribution of MHC-binding (agretope) and TCR-binding (epitope) residues.

## COMPUTATIONAL METHODS

### Generation of Starting Conformers for Peptides

Since no experimental structures were available for the peptides, we used a pattern-recognition technique to generate a set of starting conformations for each sequence in Table I. The PRISM (Pattern-Recognition Importance Sampling-Minimization) procedure<sup>23-25</sup> was chosen because of its advantage to propose 2-D conformational probabilities in the Ramachandran map for a peptide chain that can be

converted into 3-D conformations using an adequate set of  $\phi, \psi$  angles.

Ten parameters were taken into account for each amino acid,<sup>26</sup> resulting in 64 possible states occurring from 6,773 different tripeptide fragments of the Protein Brookhaven Databank.<sup>23</sup> Starting from the primary structure of each sequence, 2-D conformational probabilities were first computed for tripeptides (each residue and its neighbours) from the N-terminal to the C-terminal end, so that at the end of the process 2-D chain state probabilities were disclosed. The 15 first conformations with the highest probability of existence were then extracted and transformed into 3-D structures using a state-related  $\phi, \psi$  dataset.<sup>25</sup> All conformations were energy-minimized and those three with the lowest energy were submitted to molecular dynamics simulations.

### In Vacuo Computations (Free Peptides)

Free antigens were simulated using the AMBER all atom force field.<sup>27</sup> To avoid splitting dipoles, nonbonded interactions were calculated within a residue-based cutoff of 10 Å and the pair list updated every 50 steps. A distance-dependent dielectric constant (function of R<sub>ij</sub>) was chosen without scaling 1,4 nonbonded interactions. After 10 steps of steepest descent, a conjugate gradient energy minimization (EM) was achieved until convergence within a 0.01 kcal/mol·Å value for the root mean square gradient.

Minimized coordinates were used as starting point for a molecular dynamics (MD) simulation in vacuo. Initial velocities were taken from a Maxwellian distribution at a temperature of  $T_0 = 300$  K. To prevent abrupt changes in temperature, the system was weakly coupled to a heat bath at 300 K with a temperature relaxation time of  $\tau_T = 0.4$  psec. The MD simulation covered 40 psec of time with a time step of 1 fsec. The values of the potential energy and root mean square (rms) deviation from the initial structures were monitored to follow equilibration of the system. Coordinates, energies and velocities were saved every 100 steps (0.1 psec) and analyzed after 10 psec. The simulation of a nonapeptide took 1 central processing unit (CPU) hour on a CONVEX C220.

### Water Simulations (Proteins and Protein-Antigen Complex)

AMBER united atom force field parameters were used for the protein in addition to the TIP3P model<sup>28</sup> for simulating water. Since explicit water solvent was taken into account, the dielectric constant was kept fixed to 1.0 and 1,4 nonbonded interactions were scaled (factor 8 for Lennard-Jones and 2 for electrostatic interactions) in order to reproduce "Big Amber" parameters.<sup>29</sup> The solute was then solvated by a bag of 1764 Monte Carlo water molecules. Any water molecule closer than 1.5 Å or farther than 5 Å

TABLE II. Energy Minimization of Nona

Conformer	Probability*	State <sup>†</sup>	AE <sup>‡</sup>	rms <sup>§</sup>
1	0.163	022222220	25.01	3.42
2	0.140	022222110	23.79	4.05
3	0.052	022221110	14.01	3.37
4	0.049	011222220	29.28	3.80
5	0.042	011222110	17.19	2.84
6	0.042	022222210	35.12	4.64
7	0.039	022222120	18.32	4.21
8	0.031	021222220	35.78	4.01
9	0.026	021222110	23.87	3.40
10	0.025	022221120	24.66	3.19
11	0.023	021121120	29.50	4.05
12	0.021	011112210	2.98	1.36
13	0.020	011111110	—	—
14	0.018	011111220	5.58	1.80
15	0.017	012211130	13.38	3.50

\*2-D probability order.

<sup>†</sup>PRISM states in the Ramachandran map for the Nona residues (YPHFMPNTL), 1:  $160 < \Phi < 360^\circ$ ,  $-120 < \Psi < 40^\circ$ ; 2:  $160 < \Phi < 360^\circ$ ,  $40 < \Psi < 240^\circ$ ; 3:  $0 < \Phi < 160$ ,  $-90 < \Psi < 110^\circ$ ; 4:  $0 < \Phi < 160$ ,  $110 < \Psi < 270^\circ$ . Probabilities are not calculated on the N- and C-terminal residues. See ref. 25 for attribution of  $\Phi, \Psi$  angles depending on tripeptide conformational states (e.g., conformation #13, residue Met-5: state 111,  $\Phi = -66^\circ$ ,  $\Psi = -38^\circ$ ).

<sup>‡</sup>Energy deviation (in kcal/mol) from the minimum #13.

<sup>§</sup>Root mean square deviation (in Å) from the minimum #13 (on backbone atoms).

away from any protein atom was discarded from the calculation. After 100 steps of steepest descent, 1,000 steps of conjugate gradient minimization were performed to provide minimized coordinates used for water MD simulation. The other MD parameters were the same as previously described for the in vacuo minimization and were used for a 30 psec simulation (15 psec of equilibrium and 15 psec of analysis). It covered 60 CPU hours on a CONVEX C220 for each computation.

### 3-D Q.S.A.R. Analysis

The CoMFA (Comparative Molecular Field Analysis)<sup>30</sup> module of the SYBYL molecular modeling package<sup>31</sup> was used, running on a micro VAX 3500 with graphic display on an E&S PS390 terminal. The analysis was carried out on 9 peptides (see Table IV) with AMBER all atom charges. A three-dimensional lattice with a 2 Å grid resolution was defined by overlapping the overall molecular volumes of tabulated molecules by 4 Å along all axes. Steric and electrostatic fields exerted by peptides on a probe atom (sp<sup>3</sup> carbon with a charge of +2 esu) were computed at each point of the lattice with a 30 kcal/mol cutoff for splitting each field contribution. They were then fit to a reference field (see text) using a field-fit process<sup>30</sup> and compared by using a partial least-squares (PLS) analysis.<sup>32</sup> Any CoMFA column whose standard deviation (minimum sigma factor) was less than 2.0 was excluded from the analysis to

save computational time, for a calculation that took 20 CPU minutes.

## RESULTS AND DISCUSSION

### Simulation of Antigens

#### Conformational analysis

A recent study has identified the nonapeptide YPHFMPTNL (from now on referred to as Nona) as the peptide with the highest biological activity in the assay used, that is the peptide for which the lowest concentration in solution is needed to get sufficient binding to H-2 L<sup>d</sup> on the target cell surface for measuring target cell lysis with the CTL clone IE1 as probe.<sup>22</sup> Furthermore, from biochemical peptide isolation,<sup>12</sup> Nona appears to be the naturally processed antigenic peptide presented by cells synthesizing pp89 (U.H. Koszinowski, unpublished observations). Activity is dramatically reduced and even lost when the tested sequence is shifted to the right (Nonar) and to the left (Nonal) along the natural sequence, respectively, and essentially the same is true after N-terminal (Octar) and C-terminal (Octal) truncation. Notably, antigenic activity is largely retained after biterminal truncation (Heptar) and even the core pentapeptide HFMP is more active than Nonal and Octal, although both encompass this pentapeptide sequence entirely (Table I).<sup>22</sup> Until now, there was no structural model that could explain these empirical findings. This problem therefore provided a good test for the validity of our conformational analysis. After energy minimization of 15 starting coordinates for each of the peptides in Table I,  $\alpha$ -helical structures were preferred, in particular in the case of sequences with high functional activity. When Nona was taken as template, the  $\alpha$ -helical conformation (#13) proved to be more stable than any other of the sampled conformations with energy deviations  $\Delta E$  ranging from 2.98 to 35.78 kcal/mol (Table II). Conformers #12 and #14 differed from the  $\alpha$ -helix only by some N-terminal backbone modifications as reflected by their small rms deviation from the minimized structure #13. In fact, after molecular dynamics simulation, these three conformers converged toward a stable mean  $\alpha$ -helical structure. Based on this experience, the relative stability of the  $\alpha$ -helical conformation of the sequences was considered as a reference parameter. Our prime observation was fully confirmed by molecular dynamics. Only antigenic sequences were observed to maintain and stabilize during the 40 psec simulation time an overall  $\alpha$ -helical conformation through regular hydrogen bonds involving residues  $i, i + 4$ . As can be seen in Figure 1a, four  $\alpha$ -helical hydrogen bonds were disclosed during the 40 psec dynamics simulation of Nona. The reported average mean hydrogen bond distances between residues  $i, i + 4$  fluctuate around a 2 Å value illustrating the permanent ex-

**TABLE III. Hydrogen Bonding Pattern of Free Peptides\***

Name	Antigenic potency <sup>†</sup> (M)	Number <i>i,i+4</i> H-bonds	Number <i>i,i+3</i> H-bonds	rms <sup>‡</sup>
Nona	10 <sup>-9</sup>	4	0	—
Nonar	10 <sup>-3</sup>	3	1	0.23
Octar	10 <sup>-4</sup>	3	1	0.26
Heptar	10 <sup>-7</sup>	3	1	0.17
Penta	10 <sup>-3</sup>	1	0	0.16
Nonal	0	2	1	0.45
Octar	0	1	2	0.81
Heptal	0	0	1	1.80

\*Results of 40 psec molecular dynamics runs, averaged over 300 conformations.

<sup>†</sup>See Table I.

<sup>‡</sup>Root mean square deviation (in Å) of mean conformations from the Nona mean structure.

**TABLE IV. Observed and Predicted Activities of a Series of Nonapeptides\***

Peptide number	Xaa	Observed activity <sup>†</sup>	Predicted activity <sup>†</sup>
1	Leu	9.0	8.9
2	Ala	3.0	3.0
3	Nle	10.0	9.9
4	Ile	5.0	5.1
5	Phe	9.0	8.9
6	Met	8.0	8.0
7	Val	3.0	3.1
8	Asn <sup>‡</sup>	<3.0	0.0
9	—	<3.0	0.0
10	Phg <sup>§</sup>	7.0	5.1
11	Hpa <sup>§</sup>	8.0	7.6
12	Tyr	5.0	7.5
13	Trp	<3.0	7.9

\*Sequence: Tyr-Pro-His-Phe-Met-Pro-Thr-Asn-Xaa.

<sup>†</sup>Expressed in  $-\log M$  (see Table I).

<sup>‡</sup>No residue (corresponding octapeptide: Octal).

<sup>§</sup>Phg, phenylglycine; Hpa, homophenylalanine.

istance of 4 hydrogen bonds during the simulation. The dynamics trajectories could be analyzed using our own program ADAPTU<sup>33</sup> on an E&S PS 390 graphic terminal. Shown in Figure 1b are 10 frames corresponding to 10 conformations at 10 different simulation times. They all share an  $\alpha$ -helical backbone with only expected minor modifications occurring on side chains. For less antigenic right-shifted peptides, only one *i,i+3* hydrogen bond (Pro-1-Met-4) was found to disrupt the  $\alpha$ -helix (Table III). For the inactive left-shifted peptides less *i,i+4* and in the case of Octal two *i,i+3* hydrogen bonds were disclosed during dynamics simulation. Their mean structure (averaged over 300 conformations) differed markedly from the optimal  $\alpha$ -helix bound for Nona, as can be seen clearly from the rms deviation of all dynamics mean structures to the Nona mean

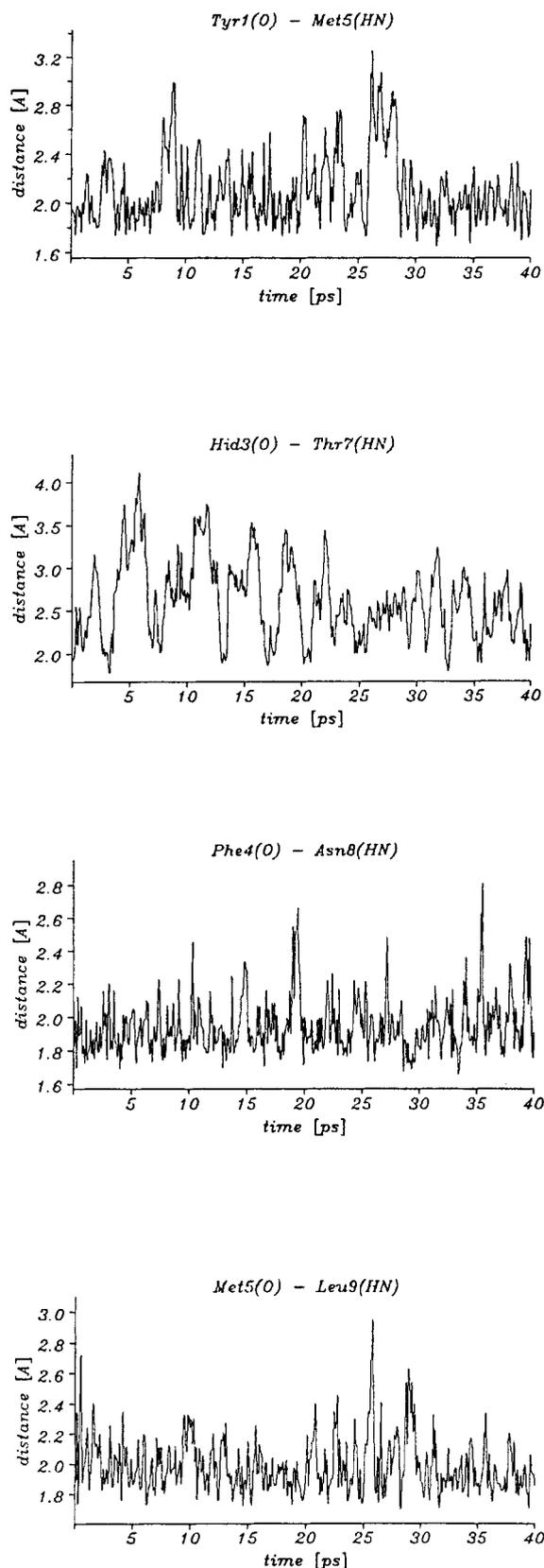


Fig. 1a. Legend appears on page 74.

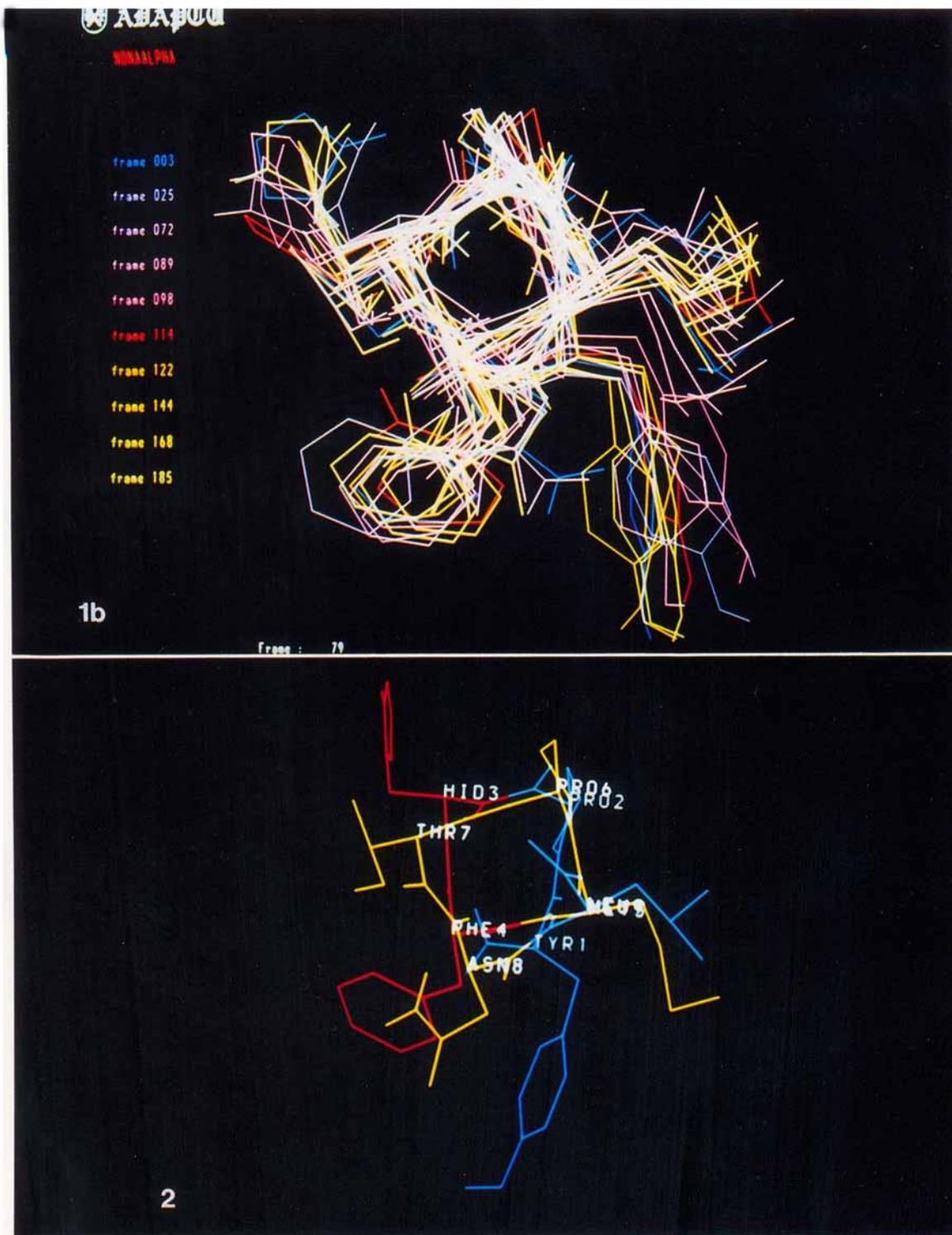


Fig. 1. (a) Hydrogen bonding pattern of free conformations of Nona, during a 40 psec in vacuo molecular dynamics simulation; (b) representation of 10 conformations (frames) of Nona at 10 different simulation times. The ADAPTU program<sup>39</sup> was used to analyze and visualize AMBER<sup>27</sup> atomic trajectories.

Fig. 2. Relaxed dynamics minimum of Nona. Agrotopes are colored in blue, epitopes in red, and spacer residues in yellow.

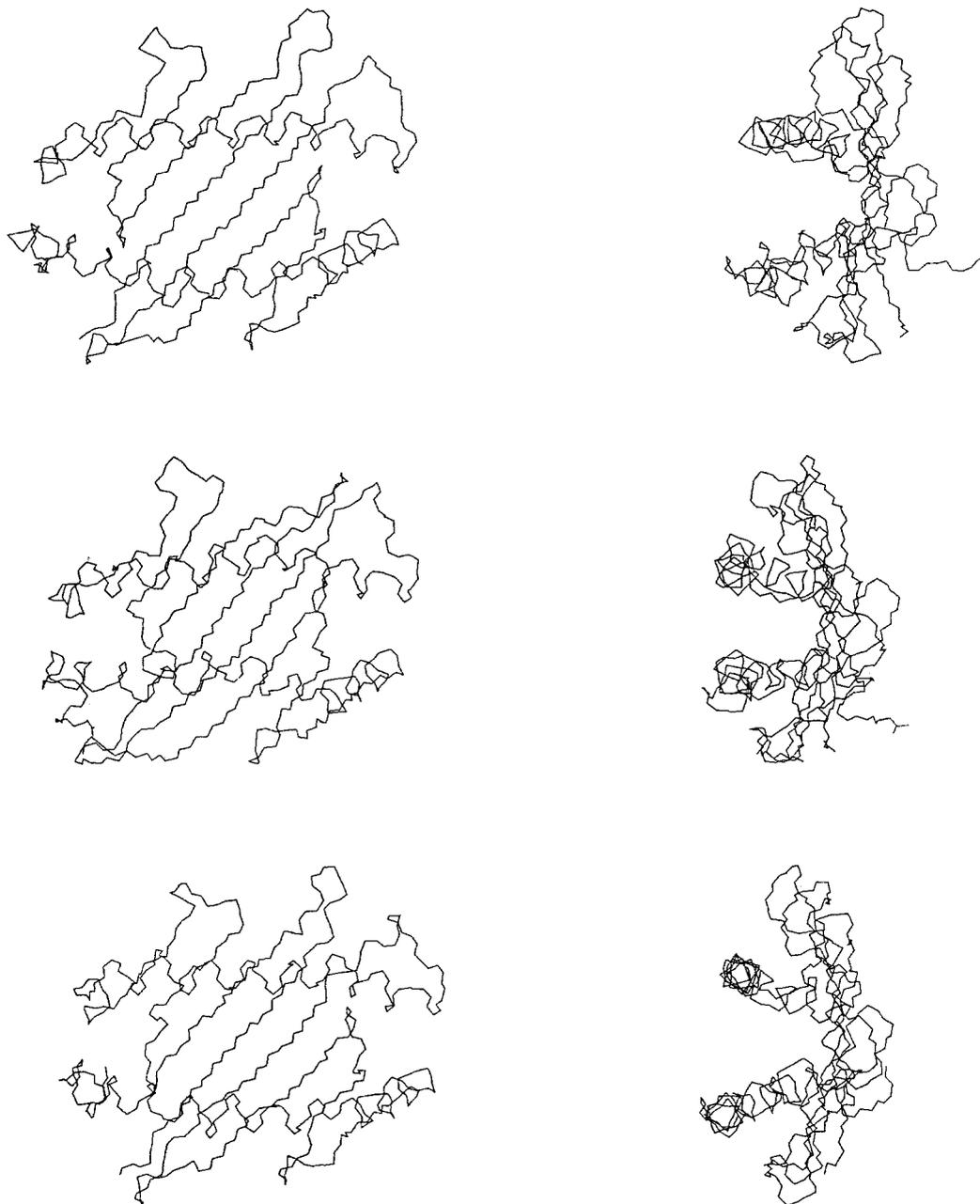


Fig. 3. Orthogonal views of three HLA-A2 backbone conformations: X-ray (**top**), D50 mean conformation (**middle**), DW mean conformation (**bottom**).

**TABLE V. Averaged Topology of the HLA-A2 Binding Site\***

Residue A	Residue B	Helices	X-Ray	<MD50> <sup>†</sup>	<MDW> <sup>‡</sup>
Lys-66	Tyr-159	$\alpha_1$ - $\alpha_2$	14.11	7.58	13.67
Trp-51	Tyr-84	$\alpha_1$ - $\alpha_1$	42.84	39.82	42.38
Met-138	Leu-179	$\alpha_2$ - $\alpha_2$	41.64	35.01	40.43

\*Inter-C <sub>$\alpha$</sub>  distances (in Å) over 150 conformations.

<sup>†</sup>Mean structure after molecular dynamics in vacuo.

<sup>‡</sup>Mean structure after molecular dynamics in water.

conformation. Notably, high rms deviations correspond to low antigenic activity (Table III). The fact that only right-shifted peptides retain some activity might be explained in part by the nature of the N-terminal residue (proline), which has limited conformational flexibility so that structural modifications occurring at the N-terminal end are less pronounced. Furthermore, the new Pro-1-Met-4 hydrogen bond found for right-shifted peptides substitutes in part for the precluded bond with the  $i + 4$  residue,

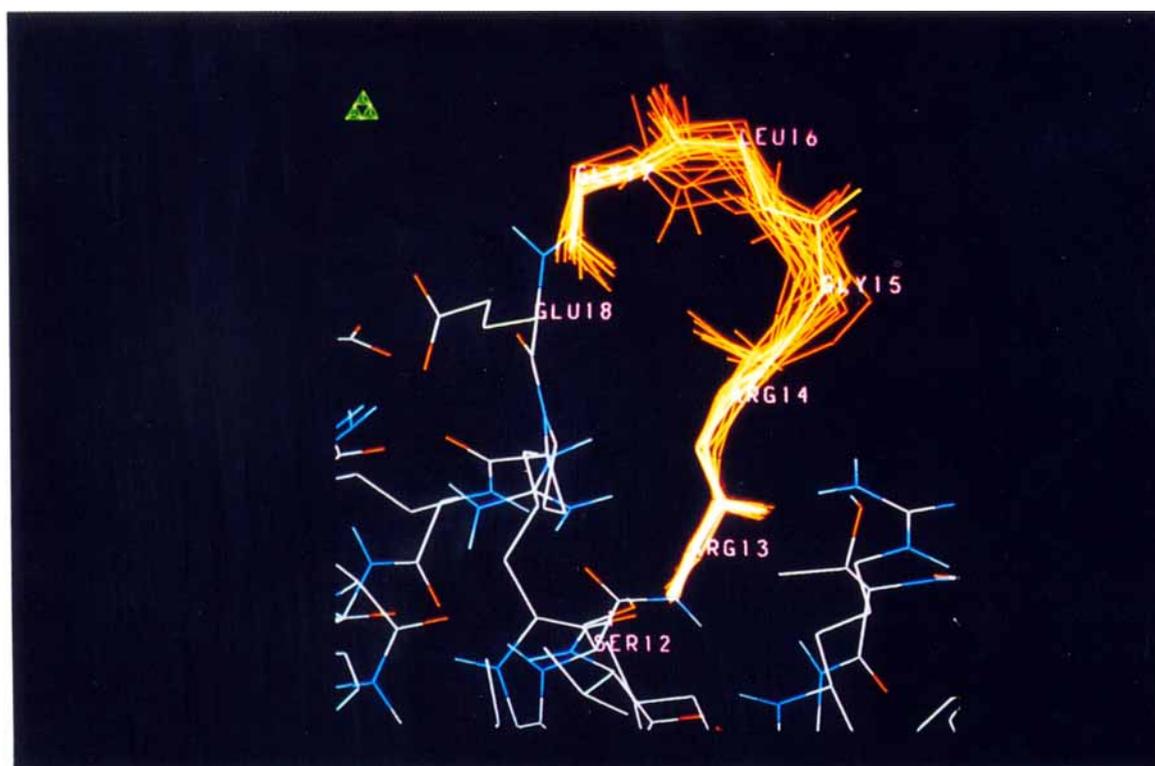


Fig. 4. Building a loop using the SYBYL<sup>31</sup> Loop Search algorithm. Proposed loop backbone conformations are represented in red. The one selected, depicted in green, showed the minimum rms deviation from the loop-defining inter- $C_{\alpha}$  distance (Ser-12-Glu-18).

TABLE VI. Averaged Topology of the H-2L<sup>d</sup> Binding Site\*

Residue A	Residue B	Helices	MIN50 <sup>†</sup>	MINW <sup>‡</sup>	<MD50> <sup>§</sup>	<MDW> <sup>**</sup>
Ile-65	Tyr-158	$\alpha_1$ - $\alpha_2$	13.70	13.81	8.22	13.50
Trp-50	Tyr-83	$\alpha_1$ - $\alpha_1$	42.58	42.24	41.21	42.18
Met-137	Leu-178	$\alpha_2$ - $\alpha_2$	40.91	41.40	40.11	42.49

\*Distances between  $C_{\alpha}$  atoms (in Å) over 150 conformations.

<sup>†</sup>In vacuo energy minimization.

<sup>‡</sup>Energy minimization in water.

<sup>§</sup>Mean structure after molecular dynamics in vacuo.

<sup>\*\*</sup>Mean structure after molecular dynamics in water.

which is also a proline. Since Tyr-1 is missing, this appears the best way of stabilizing the N-terminal end.

The relaxed dynamics minimum conformation of Nona (Fig. 2) represented with a special color coding is fully compatible with competitive binding experiments<sup>34</sup> classifying residues as agretopes (binding to the MHC protein), epitopes (binding to the T-cell receptor), and spacers (not critically involved in binding). The conformation found after molecular dynamics locates agretopes (Tyr-1, Pro-2, and Leu-9; colored in blue) on the same side of the helix, and epitopes (Hid-3 and Phe-4; colored in red) on the opposite side so that both can interact with

their respective partner in the ternary complex. In this respect, our example resembles the model proposed by Allen et al.<sup>17</sup> and, since hydrophobic and hydrophilic residues segregate accordingly, it is also in agreement with antigenicity prediction algorithm based on amphiphatic sequences.<sup>15,16</sup> Regardless of whether regular structures of short peptides ever form in solution, our calculations have revealed that the propensity to adopt a  $\alpha$ -helical conformation might be higher for the antigenic forms within the set of peptides tested, and for a 40 psec MD simulation. Stronger statements on conformational tendencies for this series of peptides should demand much longer simulation times. However, the present

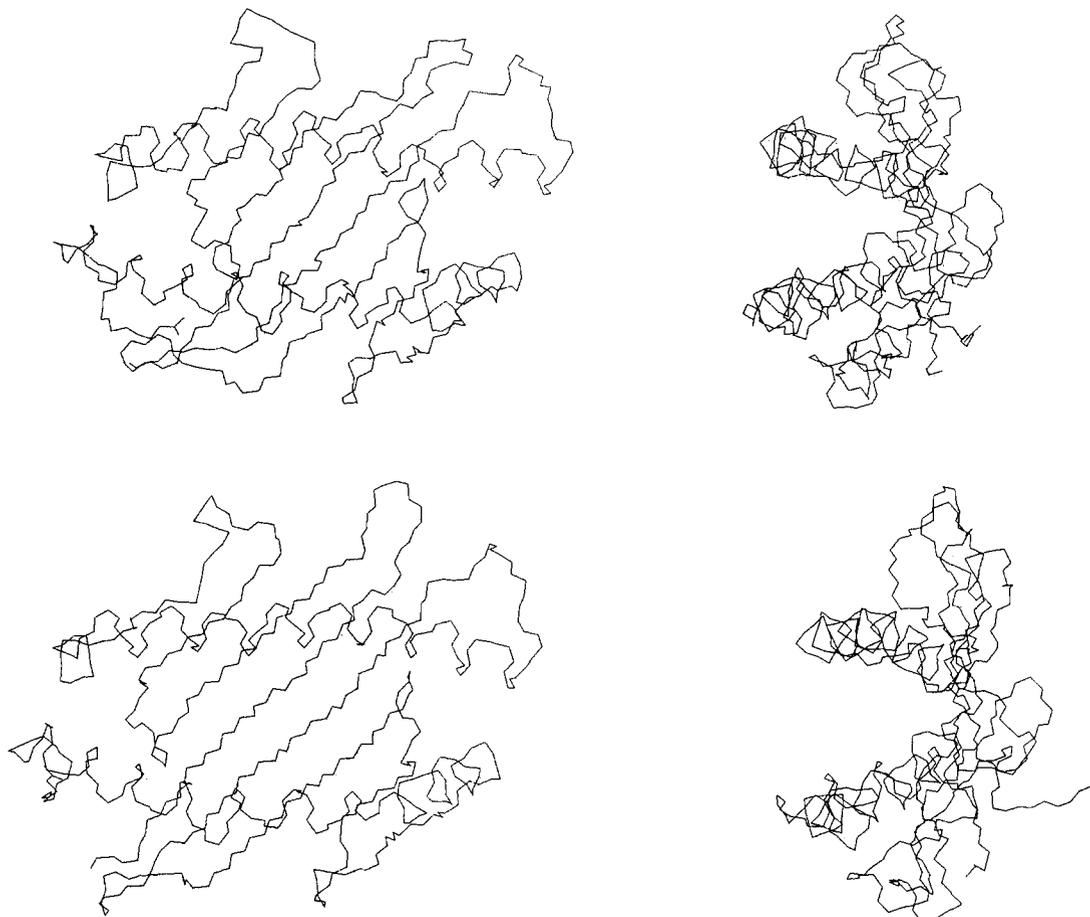


Fig. 5. Orthogonal views of two H-2L<sup>d</sup> mean backbone conformations observed after 30 psec of molecular dynamics simulation: D50 structure (**top**), DW structure (**bottom**).

model offers a satisfactory explanation for functional data that could not be explained before.<sup>2</sup> It not only explained the alternation of activity in the series of the nonapeptide to the pentapeptide, but even the differential effects of C- and N-terminal truncations.

**Importance of the C-terminal residue:  
3-D Q.S.A.R. analysis**

In addition to conformational properties, the nature of the C-terminal residue appears to be critical. Starting from the original Nona (Xaa = Leu, Table IV), modification of the last residue was shown to influence strongly the antigenic activity, going from more active (Xaa = Nle) to equipotent (Xaa = Phe, Met) or totally inactive (Xaa = Asn, H). Since none of the residue variation involved a charged amino acid that could weaken the stability of an  $\alpha$ -helix by interacting with the C-terminal end as well as the helix macrodipole,<sup>35</sup> it is possible that the experimentally observed structure-activity relationships can be explained solely by the nature of the C-terminal residue and its binding to the H-2L<sup>d</sup> molecule.

A 3-D Q.S.A.R. study was therefore performed to correlate these data by using the CoMFA module of the SYBYL modeling package. Steric and electrostatic fields were computed for all C-terminal Nona analogs. These fields were fit together and used for a PLS analysis. The reference field for field fitting was a combination of fields induced by the authentic and an equipotent peptide (#1 and #5, see Table IV). When cross-validation was applied over peptides #1 through #9 (with 9 cross-validation groups) a predictive  $r^2$  of 0.632 was obtained (press = 7.040). The corresponding final analysis without cross-validation and with an optimal number of components gave the following results:

$$r^2 = 0.999, s = 0.001$$

$$F \text{ values } (n_1 = 7, n_2 = 1): 1.42 e^{+06}$$

$$\text{Prob. of } r^2 = 0:0.001.$$

The derived model is able to almost perfectly explain *experimental data* within the original set of analogs tested (peptides #1 through #9). The predominance of the steric field contribution in the analysis (70%) is intriguing and suggests that the interaction of the

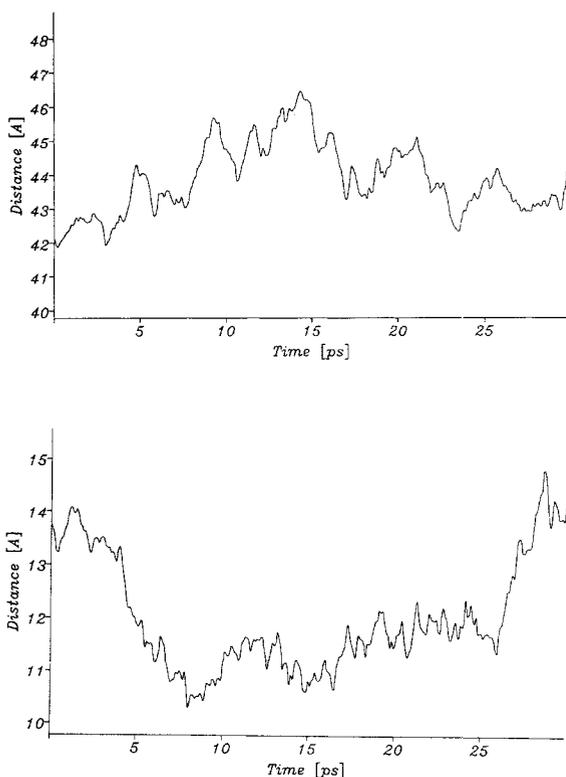


Fig. 6. Time course of two inter- $C_{\alpha}$  distances (**top**: Tyr-50-Tyr-83, **bottom**: Ile-65-Tyr-158) observed during molecular dynamics simulation of free H-2L<sup>d</sup>. The first distance represents the length of the  $\alpha$ -helical part of domain  $\alpha_1$ , the second the interhelices distance at their central part between  $\alpha_1$  and  $\alpha_2$  domains (antigen-binding site width).

C-terminal residue is mainly driven by hydrophobic forces. Yet, the predictive power of our 3-D model built from peptides #1 through #9 proved to be limited, as the prediction of the antigenic activity for a set of designed peptides (#10 to #13, Table IV) was less accurate. In particular, the complete inactivity of the tryptophan analog (peptide #13) was unexpected from our model and is probably due to steric hindrance within the active site. Thus, in this example, the 3-D model derived by Q.S.A.R. studies was not valid for extrapolating to the biological activity of molecules that differ sterically or electrostatically from the set used originally to build the three-dimensional box. In addition, adaptation of the ligand to its receptor site as well as induced-fit process cannot be taken into account in Q.S.A.R. studies so that modeling the peptide-MHC complex became necessary for explaining the observed discrepancies.

### Simulation of MHC Proteins

#### Use of HLA-A2 as template

Because the three-dimensional structure of the H-2L<sup>d</sup> molecule is unknown, a model was built from

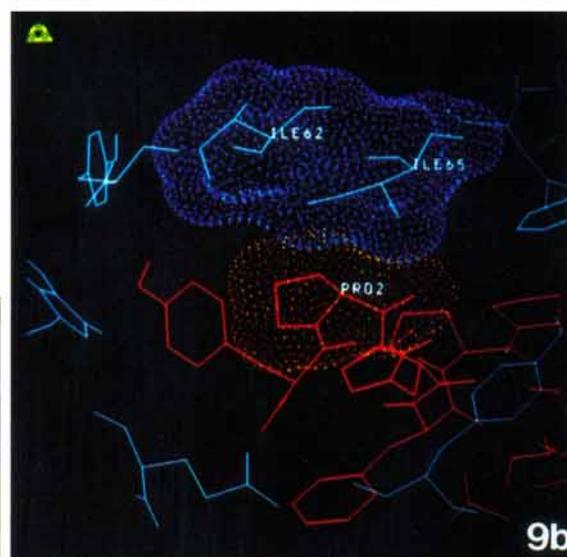
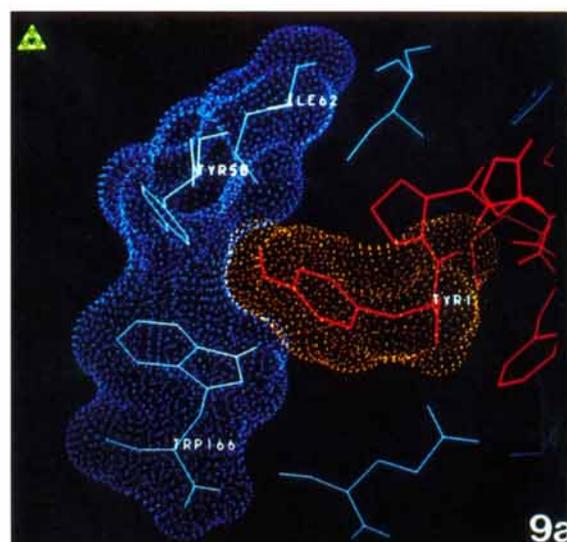
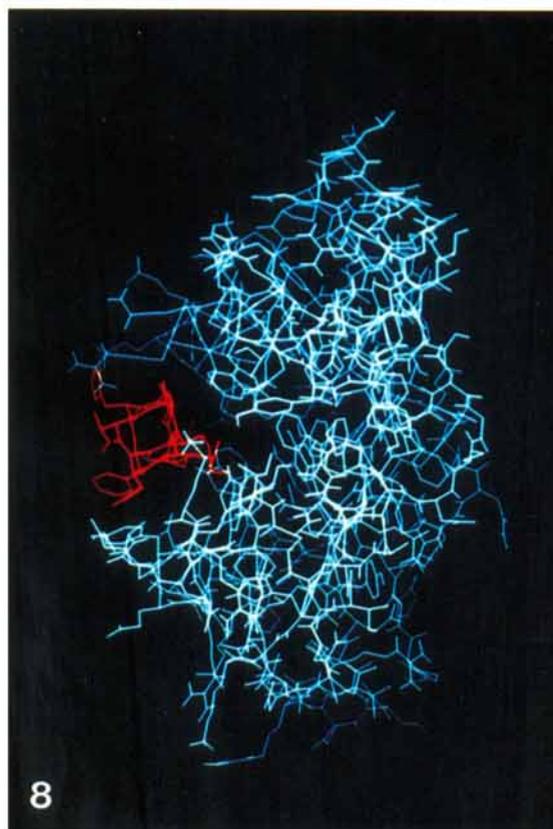
the X-ray coordinates of the human HLA-A2 MHC protein at 2.6 Å resolution<sup>36</sup> (entry in the Brookhaven Protein Databank: 3HLA), which shares 70% amino acid homology with the L<sup>d</sup> molecule in the  $\alpha_1$  and  $\alpha_2$  domains (182 residues) that form the peptide binding site.<sup>13</sup> HLA-A2 was considered as a good template since comparison of simulated structures with the experimental X-ray conformation could be achieved. The crystal coordinates of HLA-A2 were first refined to remove crystal packing forces. Using three different dielectric functions, particular attention was given to the treatment of electrostatic effects: (1) a low dielectric model (DR) with a distance-dependent dielectric function (of  $R_{ij}$ ), (2) a high dielectric model (D50) with a dielectric constant of 50, previously used to simulate implicit water solvation,<sup>37</sup> and (3) a high dielectric model (DW) with explicit water molecules and a dielectric constant set to 1.

The minimized structures closest to the averaged X-ray conformation were obtained by using high dielectric models. The rms deviations (on backbone atoms) from the crystal coordinates were 0.44 and 0.57 Å for the D50 and DW minimized structures, respectively. Use of a distance-dependent dielectric function caused more significant modifications (rms = 0.75 Å) occurring on nonregular structures imputable to overestimation of salt bridge formation at the protein surface. Therefore, the two high dielectric models were used for a 30 psec MD simulation. Surprisingly, major discrepancies were found in the two simulations, possibly attributable to influence of electrostatic interactions. After averaging 150 conformations, the rms deviation from the X-ray structure was much higher for the D50 (3.90 Å) than for the DW mean conformation (2.11 Å). Only the water simulation was found to be realistic with respect to the solid-state conformation. In fact, an artificial reorganization of the binding site (delimited by two long  $\alpha$ -helices) was found during the in vacuo D50 simulation (Table V). The distance between the two helices in their central parts was reduced by 50%. Moreover, both helices were found to be constricted significantly by 3 Å for helix 1 and 6.6 Å for

Fig. 7. CPK representations of the H-2L<sup>d</sup> molecule with the Fauchère-Pliska hydrophobic scores.<sup>36</sup> Magenta: Trp, Ile, Phe, Leu, Cys (scores from 2.85 to 1.54); cyan: Met, Val, Tyr, Pro (scores from 1.23 to 0.72); white: Ala, Thr, His, Gly, Ser (scores from 0.31 to -0.04); yellow: Gln, Asn, Glu (score from -0.22 to -0.64); red: Asp, Lys, Arg (scores from -0.77 to -1.01).

Fig. 8. Minimized conformation of a complex between H-2L<sup>d</sup> (in cyan) and Nona (in red). For the sake of clarity, water molecules are omitted.

Fig. 9. Interaction of Nona (red) with H-2L<sup>d</sup> residues (cyan). van der Waals volumes of Nona agretope residues (yellow) are represented embedded in the L<sup>d</sup> hydrophobic pockets (in blue): (a) Tyr-1, (b) Pro-2, (c) Leu-9.



Figs. 7-9. Legends appear on page 78.

helix 2 (Table V). In this case, the antigen binding site would be far too narrow to accommodate any antigenic peptide (Fig. 3). Such *in vacuo* behavior may be explained by hydrophobic packing forces tending to reduce the inter- $\alpha$ -helical distances, and illustrates the inability of *in vacuo* simulations to preserve the integrity of large hydrophobic cavities, which then tends to minimize their surface areas.

In contrast, by using explicit water molecules, motions were found to fluctuate around the crystal conformation, thus leading to averaged topologies close to the solid state situation (Table V). The mean DW conformation (Fig. 3), like the X-ray structure, should remain sterically able to accommodate an  $\alpha$ -helical peptide.

Alternatively, the observed binding site constriction could be explained by retarded conformational sampling, when explicit solvent is included. However, current work in progress has demonstrated that the HLA-A2 solvated structure was stable over a longer simulation time (D. Rognan, unpublished observations).

### **Building the H-2L<sup>d</sup> molecule**

The DW minimized coordinates of HLA-A2 were then used to build the L<sup>d</sup> molecule with the same dielectric parameters. Adequate residue modifications corresponding to the polymorphic positions were performed only on the side chains while maintaining the backbone conformation. The deletion occurring in a loop led us to rebuild this part of the protein by using a Loop-Search procedure.<sup>31</sup> The loop was defined from Ser-12 to Glu-18 (H-2L<sup>d</sup> numeration) as a Ser-Arg-Arg-Gly-Leu-Gly-Glu sequence (Fig. 4) 25 possible loop backbone conformations were disclosed by searching in the Brookhaven Databank for C $\alpha$  distances between defining-loop residues close to the examined Ser-12-Glu-18 inter-C $\alpha$  distance. The one selected (represented in green, Fig. 4) was extracted from a serine protease (PDB entry: 3RP2) for a Ser-Tyr-Asn-Ser-Val-Pro-Glu sequence whose backbone showed a 0.38 Å rms deviation from the original loop-defining inter-C $\alpha$  distance. Side chains were then added and the whole L<sup>d</sup> structure minimized. MD simulations were performed by using the previously described parameters starting from the D50 and DW minimized structures. Again, the same conformational rearrangement of the peptide binding cleft was observed during the *in vacuo* computation (Table VI) with the same spectacular narrowing of the binding site (Fig. 5). Water simulation yielded a 2.40 Å mean deviation from the initial conformation as compared to 3.60 Å for the D50 model. The DW mean structure was in fact analogous to that of HLA-A2 and showed periodic fluctuations around the starting conformation (Fig. 5). It is of particular interest that the distance between the two  $\alpha$ -helices that flank the putative binding cleft was found to fluctuate within

the 30 psec simulation time around an averaged value with a same periodicity as the motion of helix 1 (Fig. 6). When helix 1 is elongated (from 8 to 24 psec), the inter- $\alpha$ -helices distance (featuring the antigen binding width) is reduced and vice versa. Obviously, longer computations over many cycles should give more informations about such hypothetical periodic fluctuations of the L<sup>d</sup> structure.

### **Simulation of a Complex**

The more significant MD mean structure (DW model) of free H-2L<sup>d</sup> was used in order to model the interaction with the antigenic peptide Nona (see Table I) for which a large set of data is available.<sup>22,34</sup>

### **Docking statement**

Previous competitive binding experiments have already identified agretope, epitope, and spacer residues for Nona.<sup>34</sup> Taking these data into consideration, agretopes (Tyr-1, Pro-2, Leu-9) must bind to the H-2L<sup>d</sup> protein, epitopes (His-3, Phe-4) must face the TCR (or solvent, since the TCR as third partner cannot be taken into account as yet) and spacer residues (Met-5, Pro-6, Thr-7, Asn-8) should not be critically involved in binding to either partner. As anchoring residue, we chose the C-terminal amino acid (Leu-9), which may interact with a well-delimited hydrophobic pocket of the protein as suggested by our previous 3-D Q.S.A.R. studies. We had to find a pocket in the H-2L<sup>d</sup> molecule, large enough to accommodate Leu, Nle, or Phe residues but excluding Trp (see Table IV). As a graphical help, we used a color-coded representation of the target protein employing the Fauchère-Plisca hydrophobicity scale.<sup>38</sup> These scores were used mainly because they are experimentally derived and were successfully used previously for predicting antigenic sites on proteins.<sup>16</sup> Color coding went from red (hydrophilic residues) to yellow, white, cyan, and magenta for the more hydrophobic residues (Fig. 7). As to be expected, hydrophilic residues (in red and yellow) are mainly located at the protein surface exposed to solvent, whereas hydrophobic amino acids (cyan, magenta) are more buried in the three-dimensional structure of the receptor protein. Of particular interest was one hydrophobic pocket (in magenta, center of the protein structure) that fulfilled our expectations. It was indeed large enough for binding Leu or Nle side chains but too narrow (in such a conformation) for a Trp residue. This choice is also supported by the two X-ray studies done on HLA proteins (HLA-A2 and HLA-Aw68)<sup>13,14,36</sup> that have detected extra electron densities in this region (known as 74 pocket) supposedly resulting from co-crystallized peptide. For all these reasons, this peculiar pocket was selected to anchor Nona through its C-terminal end, using the minimized  $\alpha$ -helical conformation proposed by MD simulations on free peptides (Fig. 2).

**Molecular dynamics of the complex**

After energy minimization, MD simulations revealed stable conformations for the solvated complex. For the mean conformation (Fig. 8), epitopes (His-3, Phe-4) were indeed found to be orientated toward solvent and did not interact with H-2L<sup>d</sup> residues. Spacer residues were pointing toward solvent (Thr-7 and Asn-8) or toward the protein (Met-5), but in the latter case without any significant interaction, not even by the backbone atoms. Only agretope residues (Tyr-1, Pro-2, and Leu-9) could develop interactions with the receptor molecule by close van der Waals contacts, locking the peptide in two hydrophobic pockets of the L<sup>d</sup> protein. The first agretope (Tyr-1) was shown to bind closely to a binding site composed of Tyr-58, Ile-62, and Trp-166 (Fig. 9a). A hydrogen bond was disclosed for some conformations between Tyr-1 (OH) of the peptide and Trp-166 (HE1) of the protein. It remains however too versatile (only 20% of the analyzed conformers were concerned) to be really significant. Furthermore, the N-terminal end was also observed to develop a temporary salt bridge with Glu-162, at the vicinity of the binding pocket. The second agretope (Pro-2) is closely surrounded by two isoleucines (Ile-62 and Ile-65) in the continuity of the Tyr-1 binding site (Fig. 9b). The C-terminal residue (Leu-9) interacts with a second hydrophobic pocket of the receptor (Fig. 9c). This site correspond to the previously identified 74 pocket in HLA-Aw68<sup>14</sup> (73 pocket in L<sup>d</sup> numeration). It is composed of large hydrophobic amino acids (Trp-72, Phe-73, Trp-96, Phe-115, Trp-146, Tyr-155) that delimits a pocket of about 150 Å<sup>3</sup>. Notably, most of the H-2L<sup>d</sup> residues involved in Nona binding correspond to highly polymorphic positions of MHC proteins<sup>36</sup> (notably Ile-62 and Ile-65, that are unique among MHC molecules), which illustrates the highly specific nature of the antigenic recognition process.

The complex was shown to be stabilized mainly by nonbonded van der Waals interactions by at least 60 kcal/mol. Binding to the H-2L<sup>d</sup> protein did not seriously affect the antigenic peptide conformation (Table VII). Free and bound Nona mean conformations were in fact similar as shown by a rms of only 0.64 Å. An overall  $\alpha$ -helical conformation was maintained for the peptide during the 30 psec simulation of the complex with the same  $i, i + 4$  hydrogen bonds as found for the free peptide (compare Fig. 1a and Fig. 10). Some variations occurred at the more flexible C-terminal residue (Leu-9) corresponding to optimal locking into its binding site. The bound conformation of the L<sup>d</sup> molecule showed a 2.5 to 3 Å rms deviation from the free mean conformation, with equilibrated fluctuations after 15 psec (Fig. 11). No notable reorganization of the three postulated binding areas was observed, so that the free and complexed structures remained fairly similar over the

**TABLE VII. Dihedral Angles of Free and Bound Nona**

Residue	Free peptide		Bound peptide			
	$\Phi^*$	$\Psi^*$	$\Phi^\dagger$	$\Psi^\dagger$	$\Phi^\ddagger$	$\Psi^\ddagger$
Tyr-1		-55		-44		-42
Pro-2	-56	-25	-63	-42	-54	-58
His-3	-86	-46	-76	-42	-80	-18
Phe-4	-57	-56	-56	-40	-62	-41
Met-5	-50	-62	-60	-59	-51	-60
Pro-6	-59	-44	-64	-31	-75	-26
Thr-7	-62	-42	-69	-47	-76	-43
Asn-8	-72	-44	-71	-50	-83	-44
Leu-9	-75		-97		-126	

\*Mean values after 40 psec molecular dynamics in vacuo (over 300 conformations).

<sup>†</sup>Values after energy minimization in water of the complex ( $t = 0$  psec of a 30 psec molecular dynamics simulation).

<sup>‡</sup>Mean values after 30 psec molecular dynamics in water (over 150 conformations).

whole simulation (Fig. 12). Consideration of water proved to be important, since, when in vacuo simulations were performed on the complex (data not shown), narrowing of the antigen-binding cleft was observed as described previously for simulation of free MHC proteins. As a consequence, with increasing simulation time, the antigen was progressively expelled from the binding site with a concomitant loss of its  $\alpha$ -helical conformation. Only simulation with explicit water could keep the antigen locked in the L<sup>d</sup> binding sites although water molecules were not observed to be directly involved in peptide binding (Fig. 13). Tyr-1 side chain as well as His-3 and Thr-7 backbone oxygens contact by water-relayed hydrogen bonds networks three potential histotopes<sup>39</sup>: Glu-162, Glu-165, and Gln-148 (L<sup>d</sup> residues whose side chains are pointing away from the binding cleft). No hydrogen bonding to potential desotopes<sup>39</sup> (L<sup>d</sup> residues whose side chains are pointing toward the binding cleft) was observed. Most of these hydrogen bonds may be easily disrupted when a TCR is binding to the L<sup>d</sup>-Nona complex. Histotopes would become available for binding to TCR residues in competition with solvent. As expected, one epitope residue (His-3) was bound to a water molecule. Interestingly, Thr-7 and Asn-8 side chains did not develop any hydrogen bond with the solvent. They remained bound to the peptide main chain atoms (Table VIII) during the whole simulation. This feature could explain their classification as "spacer residues,"<sup>34</sup> which role could be to stabilize the  $\alpha$ -helical conformation of the peptide by additional side chain-backbone hydrogen bonding.

From this study, it appears that Nona is bound to the H-2L<sup>d</sup> protein in an  $\alpha$ -helical conformation. As an  $\alpha$ -helix, it is locked in two hydrophobic pockets, one each side of the binding cleft. The distance separating the two hydrophobic pockets in L<sup>d</sup> (pocket 1:

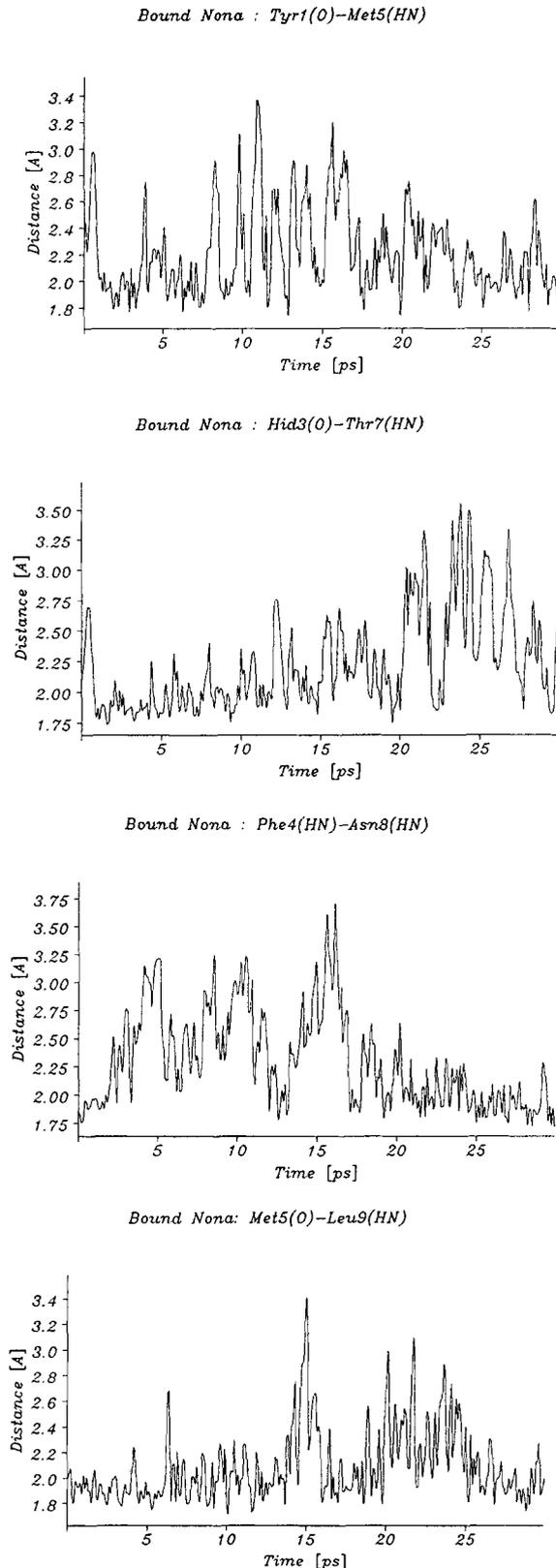


Fig. 10. Hydrogen-bonding pattern of Nona conformations bound to the H-2L<sup>d</sup> protein during a 30 psec molecular dynamics simulation in water.

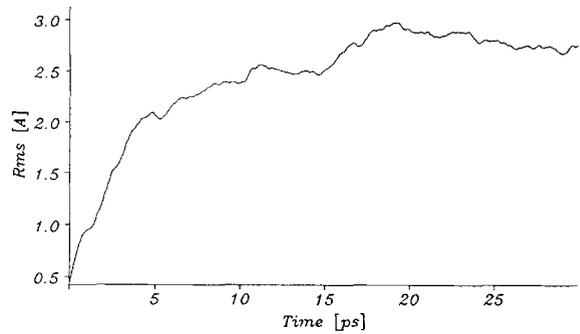


Fig. 11. Time course of rms fluctuations (in Å) of peptide-bound H-2L<sup>d</sup> conformations from the mean free structure.

Tyr-58, Ile-62, Ile-65, Trp-166; pocket 2: Trp-72, Phe-73, Trp-96, Phe-115, Trp-146, Tyr-155) was found to fluctuate around 20 Å. It allows an optimal binding of Nona, which, in an  $\alpha$ -helical conformation, spans 19.3 Å as measured between Tyr-1 (OH) and Leu-9 (CD2). In contrast, as a minimized  $\beta$ -sheet, Nona would have a length of 25.5 Å and would therefore have to span over one of these pockets, which would preclude proper interaction of Tyr-1, Leu-9, or both. The topology of the binding cleft thus makes it unlikely that Nona could interact with L<sup>d</sup> in any regular conformation other than  $\alpha$ -helical.

Interestingly, the model provides also an immediate explanation for the inactivity of the C-terminal Trp analog of Nona, that 3-D Q.S.A.R. studies have failed to explain (recall Table IV). The binding pocket into which the authentic Nona is proposed to be anchored by Leu-9 has an averaged diameter between Trp-96 and Tyr-155 of 4.5 Å, which is simply too narrow to accommodate the bulky Trp. Thus, the modeled MHC-peptide complex has resolved the discrepancies between 3-D Q.S.A.R. prediction and observed activity, and is in remarkable agreement with all functional data available so far for Nona and its C-terminally substituted analogs.

Our findings with a particular MHC-peptide pair do not exclude, of course, that other peptide conformations are favored in other MHC-peptide complexes. In particular, the nature of MHC binding pockets (hydrophilic or hydrophobic, neutral or charged) and the intersite distances, which may vary in different MHC molecules, are likely to dictate the optimal length and conformation of antigenic peptides. One may criticize that any model of antigen recognition must necessarily be incomplete as long as the contribution of the TCR in the ternary MHC-peptide-TCR complex cannot be included in the simulation. Yet, to this argument, the fact has to be recalled that formation of the ternary complex shows polarity. The peptide must first bind to the MHC molecule before the MHC-peptide binary complex is recognized by a TCR. Attempts to get peptide

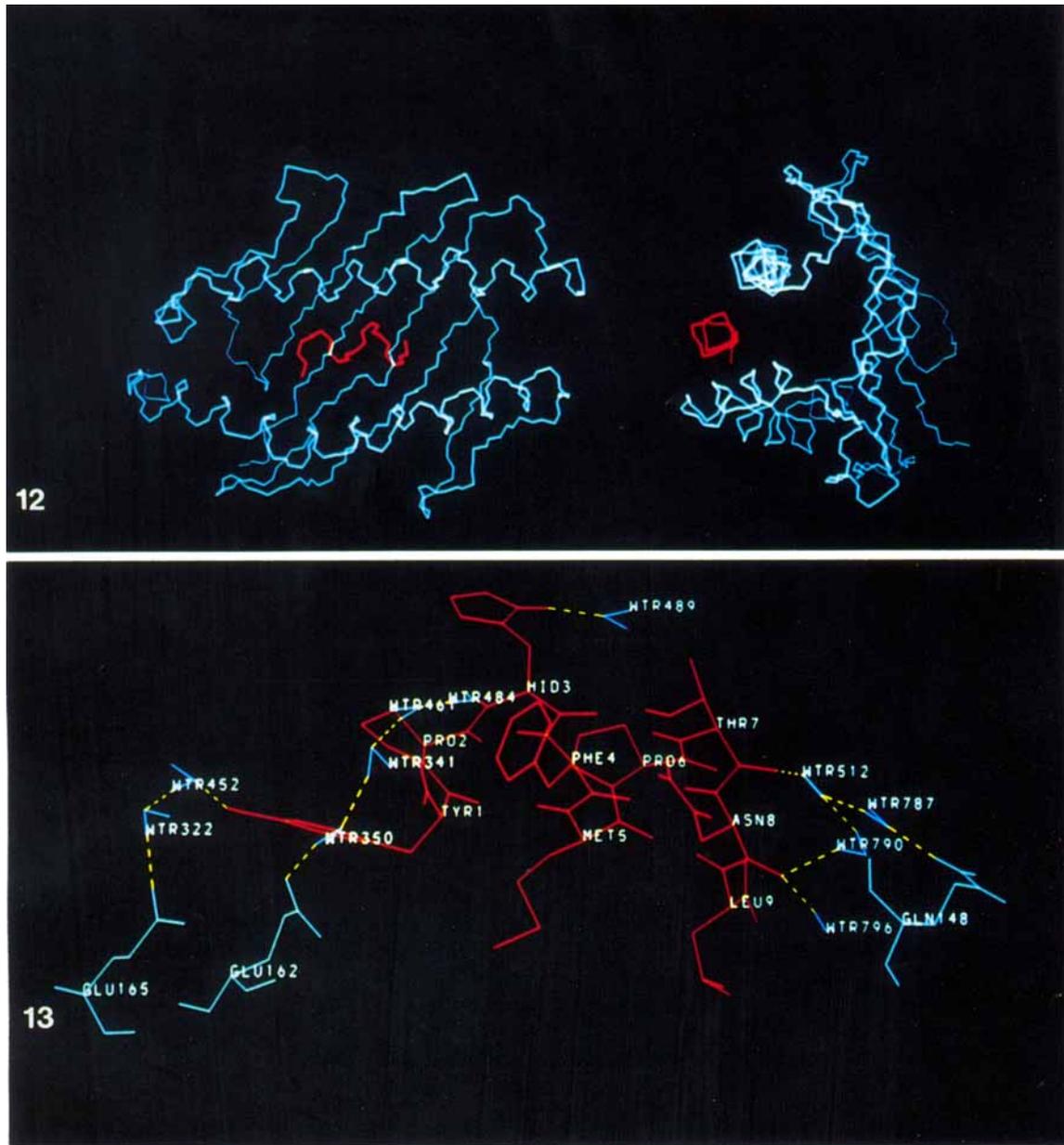


Fig. 12. Backbone mean conformation of the Nona- $L^d$  complex after 30 psec molecular dynamics simulation in water (nona is colored in red and H-2 $L^d$  in cyan).

Fig. 13. Hydrogen bonds (yellow broken lines) between Nona (in red), H-2 $L^d$  (in cyan), and water molecules (blue) at the end of the 30 psec dynamics simulation of the complex.

recognition by incubating the T cells with synthetic peptides before adding the MHC-expressing target cells to the assay failed. This sequence of events makes sense also in view of the physiologic formation of the complex: naturally processed antigenic peptides associate with the MHC molecule within the target cell and the binary complex is then transported to the target cell surface.<sup>10-12</sup> It is clear that the model proposed here will not be able to explain

functional data for Nona analogs with substitutions in epitope positions (His-3, Phe-4).<sup>34</sup> In fact, all substitutions or truncations of Nona considered in this report were shown by functional assays to affect primary binding to H-2 $L^d$ .<sup>34</sup> For this reason, the antigenic potency of the peptides used in this study reflected predominantly, if not exclusively, their capability of binding to H-2 $L^d$ , which reduced the three-molecular problem to a two-molecular inter-

**TABLE VIII. Nona-Water-L<sup>d</sup> Hydrogen Bonds Network\***

Donor	Acceptor		D . . . A distance <sup>†</sup>	D-H . . . A angle <sup>‡</sup>
Met-5	(NH)	Tyr-1 (O)	2.89	168.7
Thr-7	(NH)	His-3 (O)	2.99	176.4
Thr-7	(OG1)	His-3 (O)	2.78	165.0
Asn-8	(NH)	Phe-4 (O)	2.78	158.4
Asn-8	(ND2)	Phe-4 (O)	2.85	147.3
Leu-9	(NH)	Met-5 (O)	2.81	150.1
Tyr-1	(OH)	Wat-452 (OW)	2.80	159.3
Wat-452	(OW)	Wat-332 (OW)	2.77	163.9
Wat-332	(OW)	Glu-165 (OE2)	2.69	159.7
His-3	(NH)	Wat-484 (OW)	2.86	154.5
Wat-484	(OW)	Wat-461 (OW)	2.61	168.8
Wat-461	(OW)	Wat-341 (OW)	3.28	162.0
Wat-341	(OW)	Wat-350 (OW)	2.85	158.9
Wat-350	(OW)	Glue-162 (OE1)	2.58	146.4
His-3	(ND1)	Wat-489 (OW)	2.86	171.0
Gln-148	(NE2)	Wat-787 (OW)	2.89	172.4
Wat-787	(OW)	Wat-512 (OW)	2.89	157.3
Wat-512	(OW)	Thr-7 (O)	2.80	132.3
Wat-790	(OW)	Wat-512 (OW)	2.76	161.3
Wat-790	(OW)	Leu-9 (O)	2.83	136.5
Wat-796	(OW)	Leu-9 (O)	2.77	159.1

\*Situation after 30 psec of molecular dynamics of the complex in water.

<sup>†</sup>Donnor (D)-acceptor (A) distance in Å.

<sup>‡</sup>Donnor (D)-hydrogen (H)-acceptor (A) angle in degrees.

action. As Nona represents the bona fide naturally processed peptide, it is worth emphasizing that its N- and C-terminal amino acids are both critically involved in binding to H-2L<sup>d</sup> and, according to the model presented here, these terminal residues are buried within hydrophobic pockets of the receptor. An hypothesis that proposes trimming of longer peptides to the optimal length by peptidases after binding to the MHC molecule<sup>40</sup> is therefore difficult to envisage. Thus, our model predicts that antigenic peptides should be finally processed at least at either of their termini before interaction with the MHC molecule.

### CONCLUSION

From molecular dynamics simulations based on biological data, we propose a mode of binding of a set of related antigenic peptides to the H-2L<sup>d</sup> major histocompatibility protein. Free peptides and free MHC molecule as well as the complex between both were studied and revealed the following features: (1) For free peptides, antigenicity was observed to correspond to the aptitude to maintain an  $\alpha$ -helical conformation through  $i, i + 4$  hydrogen bonds. Less active and inactive peptides showed a disruption of the  $\alpha$ -helical conformation at the N-terminal end. In addition, the nature of the C-terminal residue was shown to be determinative for an optimal binding to the H-2L<sup>d</sup> protein, driven mainly by hydrophobic forces. (2) A free conformation of the murine MHC

molecule H-2L<sup>d</sup> is proposed, built from X-ray coordinates of the homologous human MHC molecule HLA-A2. It revealed the crucial importance of electrostatics in molecular dynamics simulations since realistic computations were achieved only by using a model fully solvated with water. In contrast, in vacuo simulations, regardless of which dielectric functions were employed, were unable to maintain the integrity of the large antigen-binding cleft and led to dramatic narrowing of the binding site. (3) Binding of an antigenic peptide to H-2L<sup>d</sup> is achieved in an  $\alpha$ -helical conformation by different van der Waals interactions with two distinct hydrophobic pockets of the protein. Polymorphic residues of MHC proteins were involved in the binding process, which illustrates the specific nature of antigenic peptide association with its receptor.

The proposed model is not only fully compatible with functional data but also offers a reasonable explanation for previously unexplained experimental findings: alternation of antigenic potency during systematic biterminal truncation of a highly active nonapeptide,<sup>22</sup> differential effects of N- and C-terminal truncations,<sup>22</sup> distribution of agretope, epitope, and spacer residues within the sequence,<sup>34</sup> and effects of amino acid side chain substitutions at the C-terminus (this report). The model highlights the largely hydrophobic nature of peptide-L<sup>d</sup> interaction and will aid the design of new experiments for studying peptide presentation through the H-2L<sup>d</sup> molecule.

### ACKNOWLEDGMENTS

The authors thank Dr. K. Burke for critical reading of the manuscript, and Drs. U. Weber and H. Kalbacher for peptide synthesis. Gerd Helms is gratefully acknowledged for its technical assistance. This work was supported by FOURNIER Laboratories (Dijon, France) and by the Deutsche Forschungsgemeinschaft SFB 322.

### REFERENCES

1. Zinkernagel, R.M., Doherty, P.C. MHC-restricted cytotoxic T cells: Studies on the biological role of polymorphic major transplantation antigens determining T cell restriction specificity, function and responsiveness. *Adv. Immunol.* 27:52-177, 1979.
2. Schwartz, R.H. T lymphocyte recognition of antigen in association with gene products of the major histocompatibility complex. *Annu. Rev. Immunol.* 3:237-261, 1985.
3. Buus, S., Sette, A., Grey, H.M. The interaction between protein-derived immunogenic peptides and Ia. *Immunol. Rev.* 98:115-141, 1987.
4. Kourilsky, P., Claverie, J.M. MHC-antigen interactions: What does the T cell receptor see? *Ad. Immunol.* 45:107-193, 1989.
5. Townsend, A.R.M., Bodmer, H. Antigen recognition by class I-restricted T lymphocytes. *Annu. Rev. Immunol.* 7: 601-624, 1989.
6. Unanue, E.R., Beller, D.I., Lu, C.Y., Allen, P.M. Antigen presentation: Comments on its regulation and mechanism. *J. Immunol.* 132:1-5, 1984.
7. Townsend, A.R.M., Rothbard, J., Gotch, F.M., Bahadur, G., Wraith, D., McMichael, A.J. The epitopes of influenza nucleoprotein recognized by cytolytic T lymphocytes can be

- defined with short synthetic peptides. *Cell* 44:959–968, 1986.
8. Babbitt, B.P., Allen, P.M., Matsueda, G., Haber, E., Unanue, E.R. Binding of immunogenic peptides to Ia histocompatibility molecules. *Nature (London)* 317:359–361, 1985.
  9. Chen, B.P., Parham, P. Direct binding of influenza peptides to class I HLA molecules. *Nature (London)* 337:743–745, 1989.
  10. Townsend, A.R.M., Ohlen, C., Bastin, J., Ljunggreen, H.G., Foster, L., Käre, K. Association of class I major histocompatibility heavy and light chains induced by viral peptides. *Nature (London)* 340:443–448, 1989.
  11. Van Bleek, G.M., Nathanson, S.J. Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K<sup>b</sup> molecule. *Nature (London)* 348:213–216, 1990.
  12. Rötzschke, O., Falk, K., Deres, K., Schild, H., Norda, M., Metzger, J., Jung, G., Rammensee, H.G. Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. *Nature (London)* 348:252–253, 1990.
  13. Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., Wiley, D.C. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature (London)* 329:506–512, 1987.
  14. Garrett, T.P.J., Saper, M.A., Bjorkman, P.J., Strominger, J.L., Wiley, D.C. Specificity pockets for the side chains of peptide antigens in HLA-AW68. *Nature (London)* 342:692–696, 1989.
  15. DeLisi, C., Berzofsky, J.A. T-cell antigenic sites tend to be amphipatic structures. *Proc. Natl. Acad. Sci. U.S.A.* 82:7048–7052, 1985.
  16. Margalit, H., Spouge, J.L., Cornette, J.L., Cease, K.B., DeLisi, C., Berzofsky, J.A. Prediction of immunodominant helper T cell antigenic sites from the primary structure. *J. Immunol.* 138:2213–2229, 1987.
  17. Allen, P.M., Matsueda, G.R., Evans, R.J., Dunbar, J.B., Jr., Marshall, G.R., Unanue, E.M. Identification of the T-cell and Ia contact residues of a T-cell antigenic epitope. *Nature (London)* 327:713–715, 1987.
  18. Sette, A., Buus, S., Colon, S., Smith, J.A., Miles, C., Grey, H.M. Structural characteristics of an antigen required for its interaction with Ia and recognition by T cells. *Nature (London)* 328:395–399, 1987.
  19. Maryanski, J.L., Verdini, A.S., Weber, P.C., Salemme, F.R., Corradin, G. Competitor analogs for defined T cell antigens: Peptides incorporating a putative binding profile and polyproline or polyglycine spacers. *Cell* 60:63–72, 1990.
  20. Reddehase, M.J., Koszinowski, U.H. Significance of herpes virus immediate early gene expression in cellular immunity to cytomegalovirus infection. *Nature (London)* 312:369–371, 1984.
  21. Del Val, M., Volkmer, H., Rothbard, J.B., Jonijic, S., Messerle, M., Schickendanz, J., Reddehase, M.J., Koszinowski, U.H. Molecular basis for cytolytic T-lymphocyte recognition of the murine cytomegalovirus immediate-early protein pp89. *J. Virol.* 62:3965–3972, 1988.
  22. Reddehase, M.J., Rothbard, J.B., Koszinowski, U.H. A pentapeptide as minimal antigenic determinant for MHC class I-restricted T lymphocytes. *Nature (London)* 337:651–653, 1989.
  23. Lambert, M.H., Sheraga, H.A. Pattern-recognition in the prediction of protein structure. I. Tripeptide conformational probabilities calculated from the amino acid sequence. *J. Comput. Chem.* 10:770–797, 1989.
  24. Lambert, M.H., Sheraga, H.A., Pattern-recognition in the prediction of protein structure. II. Chain conformation from a probability-directed search procedure. *J. Comput. Chem.* 10:798–816, 1989.
  25. Lambert, M.H., Sheraga, H.A., Pattern-recognition in the prediction of protein structure. III. An importance-sampling minimization procedure. *J. Comput. Chem.* 10:817–831, 1989.
  26. Kidera, A., Konishi, A., Oka, M., Ooi, T., Sheraga, H.A. Statistical analysis of the physical properties of the 20 naturally occurring amino acids. *J. Protein. Chem.* 4:23–55, 1985.
  27. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr., Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784, 1984.
  28. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 70:926–935, 1983.
  29. Tirado-Rives, J., Jorgensen, W.L. Molecular dynamics of proteins with the OPLS potential function. Simulations of the third domain of silver pheasant ovomucoid in water. *J. Am. Chem. Soc.* 112:2773–2781, 1990.
  30. Cramer, R.D., III, Patterson, D.E., Bunce, J.D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 110:5959–5967, 1988.
  31. SYBYL 5.3 release, Tripos Assoc., Inc., St-Louis, MO 63144, U.S.A.
  32. Wold, S., Albano, C., Dunn, W.J., III, Esbensen, K., Hellberg, S., Johansson, E., Lindberg, W., Sjostrom, M. Modeling data files by principal component and PLS: Class patterns and quantitative predictive relations. *Analisis* 12:477–485, 1984.
  33. Krug, M., Folkers, G. ADAPTU: Animated Dynamics Analysis Program at the Tübingen University. *J. Mol. Graph.* 9:119–121, 1991.
  34. Reddehase, M.J., Koszinowski, U.H. Redistribution of critical major histocompatibility complex and T cell receptor-binding functions of residues in an antigenic sequence after biterminal substitution. *Eur. J. Immunol.* 21:1697–1701.
  35. Serrano, L., Fersht, A. Capping and alpha helix stability. *Nature (London)* 342:296–299, 1989.
  36. Saper, M.A., Bjorkman, P.J., Wiley, D.C. Refinement of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J. Mol. Biol.* 219:277–319, 1991.
  37. Wendoloski, J.J., Matthew, J.B. Molecular dynamics effects on protein electrostatics. *Proteins* 5:313–321, 1989.
  38. Fauchère, J.L., Pliska, V. Hydrophobic parameters of amino acids side chains from the partitioning of N-acetyl-amino-acid amides. *Eur. J. Med. Chem.* 18:369–375, 1983.
  39. Owen, M.J., Lamb, J.R. General principles of recognition. In: "Immune Recognition." D.E. Male, ed. Oxford: IRL Press, 1988: pp. 1–16.
  40. Falk, K., Rötzschke, O., Rammensee, H.G. Cellular peptide composition governed by major histocompatibility complex class I molecules. *Nature (London)* 348:248–251, 1990.