Journal of Biomolecular NMR, 3 (1993) 91–112
ESCOM

J-Bio NMR 091

Structural and dynamic studies of two antigenic loops from haemagglutinin: A relaxation matrix approach

Bruno Kieffer[a,*], Patrice Koehl[a], Serge Plaue[b] and Jean-François Lefèvre[a]

[a]*IBMC du CNRS, 15 rue Descartes, 67084 Strasbourg Cedex, France*
[b]*Neosystem, 7 rue de Boulogne, 67000 Strasbourg, France*

Received 16 January 1992
Accepted 22 September 1992

Keywords: NMR; Antigenic peptides; Structure determination; NOE back-calculation; Dynamics

SUMMARY

We have investigated the dynamics and structural behaviour of two antigenic peptides using ^{1}H NMR. The two cyclic peptides mimic the antigenic site A of influenza haemagglutinin protein; they only differ in the way they were cyclized and in the size of their respective linkers. Homonuclear relaxation parameters extracted from a complete NOE matrix were interpreted in terms of local dynamics. A set of distance constraints was deduced from these parameters which allowed 3D models to be constructed using distance geometry. NOE back-calculation was used to check the validity of the final models. Strong variations of internal motion amplitude have been found in both peptides along their backbone. Motions with high amplitudes have been localized in the Gly–Pro–Gly sequence which forms a β-turn in both structures.

INTRODUCTION

Although crystallographic studies of antigen–antibody complexes first led to an interpretation of the interaction within the framework of a 'key-lock' model (Mariuzza et al., 1987), there are now numerous examples where antigens adapt their structure to bind to the antibody (Colman et al., 1987; Glaudemans et al., 1990). A good correlation has been found between internal mobility in surface protein segments and localization of antigenic determinants (Westhof et al., 1984; Williams, 1989). This experimental evidence favours an 'induced-fit' model rather than a 'key-lock' model to explain protein surface–surface interaction mechanisms. In the 'induced-fit' model, each partner of the interaction undergoes structural rearrangements. Therefore, it is important

*To whom correspondence should be addressed.

Abbreviations: DSS, 3-(trimethylsilyl)-1-propanesulfonic acid; D-loop, aspartic acid loop; ELISA, enzyme-linked immunoabsorbent assay; f.i.d., free induction decay; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; HPLC, high pressure liquid chromatography; K-loop, lysine loop; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; r.m.s.d., root-mean-square deviation of atomic positions.

to gain insight into the flexibility of the antigen, which is related to its ability to adapt its conformation.

Haemagglutinin, a coat protein of influenza virus, provides an excellent model to investigate the structural and dynamic properties involved in the antigen recognition process. The 3D structure of haemagglutinin (Wilson et al., 1981) allows the localization of one of the major epitopes, i.e. the site, A, which forms a protruding loop on the native structure (Wiley et al., 1981). In an attempt to produce a synthetic vaccine which could stimulate an immunogenic response against haemagglutinin, two peptides were designed which mimic this antigenic loop site A (Muller et al., 1990). Both peptides were cyclized in order to constrain their structures, and they differ only in the manner of cyclization (Fig. 1): in the small loop (D-loop) the side-chain carboxyl group of Asp^{10} is linked to the amide of Cys^{1}; whereas in the large loop (K-loop) the side chain of Lys^{10} is linked to the amide of Cys^{1} through a bicarboxylic succinyl group. Despite this small difference in the primary sequence, only the D-loop is recognized by antibodies raised against haemagglutinin in an ELISA test (Muller et al., 1990).

Structural studies of biological macromolecules by interpretation of inter-proton NOE encounter three major difficulties: the magnetization transfer via alternative pathways, internal dynamics, and peak overlapping. The use of the relaxation matrix in interpretation of NOE values has proven to be useful in overcoming the problem of spin diffusion (Keepers and James, 1984; Lefèvre et al., 1987; Boelens et al., 1988; Borgias and James, 1990; Koehl and Lefèvre, 1990). Iterative structure refinement programs including the calculation of the relaxation matrix parameters have been used to solve DNA oligomer structure (Boelens et al., 1989) as well as protein structure (Nilges et al., 1991). In the case of small peptides, where resonances are well resolved, the relaxation matrix can be calculated from a single NOESY spectrum (Mirau, 1988). We have investigated the structural and dynamic properties of both the D-loop and the K-loop using a relaxation matrix approach to obtain true inter-proton relaxation parameters. The elements of the relaxation matrix were first interpreted in terms of local dynamics. They also provided us with a set of distance constraints which were used to build 3D models of both peptides using the distance geometry program FILMAN (Koehl et al., 1992), based on a statistical analysis.

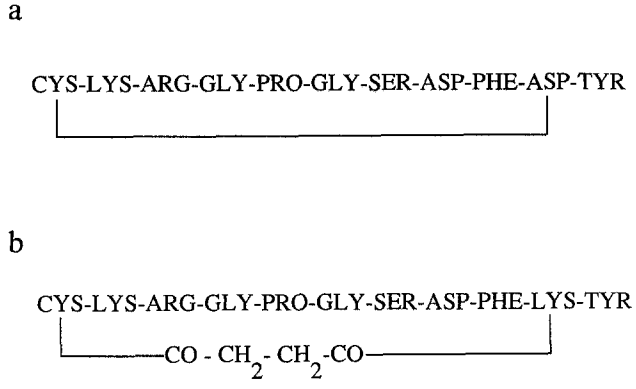

Fig. 1. Primary structure of synthetic peptides D-loop (a) and K-loop (b). The solid line represents the side chain of the aspartic acid and lysine for the D-loop and the K-loop, respectively. Linkage was via a peptidyl bond to the terminal amide of the cysteine.

MATERIALS AND METHODS

(a) Peptide synthesis

The synthesis of the D-loop and K-loop peptides has been described previously (Plaue, 1990). It was performed using the solid phase method (Merrifield, 1963). The cyclization was performed on the solid support prior to final cleavage of the peptides. The purity was verified by analytical HPLC and by consideration of the 1D NMR spectrum.

(b) NMR spectroscopy

NMR spectroscopy was performed on 500-µl samples containing 10 mg of each peptide (final concentration 15 mM) in 90% H_2O, 10% D_2O, 20 mM acetate buffer, pH 4, at 297 K and 303 K. DSS (0.1 mM final concentration) was added as an internal standard. The proton spectra were recorded at 400 MHz (on a Bruker AM-400 spectrometer) or at 500 MHz (on a Bruker AM-500 spectrometer).

2D spectra were acquired by recording 512 f.i.d. of 2048 points, in the f1 dimension. The relaxation delay was set to 3 s. All 2D spectra were acquired in the phased-sensitive mode with the time-proportional phase increment (TPPI) of the initial pulse (Wüthrich and Marion, 1983). The water signal was suppressed by continuous selective irradiation during the preparation period. The irradiation power was chosen as low as possible to avoid spill-over on the alpha proton resonances. The standard NOESY experiment (Macura and Ernst, 1980) was modified in order to control the signal phase and to flatten the baseline (Frenkiel et al., 1990).

For HOHAHA spectra (Davis and Bax, 1985), the radiofrequency field strength was set to 7800 Hz and the mixing time to 40 ms in order to allow observation of multiple bond correlations. Thirty-two scans were acquired for each point in the f1 dimension. For NOESY spectra, 40 scans were acquired for the D-loop and 64 for the K-loop for each f1 point.

2D spectra processing was performed on a micro Vax II computer using the program FTNMR (Hare Research Inc.). The final size of the matrices was 2K × 2K real points. Prior to Fourier transformation the signal was multiplied in both dimensions by a 45° shifted square sine-bell window function in the case of HOHAHA spectra, and by a 90° shifted square sine-bell in the case of NOESY spectra.

The volumes of the peaks were extracted from NOESY spectra using elliptical boxes provided by the integration routine in FTNMR. The volumes of one peak and its symmetric were averaged to calculate the corresponding NOE values. In the case of two overlapping cross peaks, we approximated the volume of each peak as half of the total cross-peak volume. For degenerate side-chain protons, only one pseudo-proton was considered.

(c) NOE measurements

The amount of magnetization exchanged between the protons, i and j, during a mixing time, τ_m, is proportional to the cross-peak volume, $V_{ij}(\tau_m)$. The absolute value of the Overhauser effect, $NOE_{ij}(\tau_m)$, can be calculated provided that the magnetization at zero mixing time ($M_i(0)$) for each proton, i, and the conversion factor (A) for estimating the net amount of magnetization represented by a cross peak are known:

$$NOE_{ij}(\tau_m) = \frac{A \cdot V_{ij}(\tau_m)}{M_i(0)} \quad (1)$$

$M_i(0)$ was calculated as described previously (Koehl et al., 1989). Briefly, the method is based on the fact that the area of a resonance line, in the first f1 1D spectrum of a NOESY, is equal to the sum of the volumes of the cross peaks lying in the 2D spectrum along the column corresponding to this line. The magnetization leakage, due to different sources of relaxation, can be monitored by a series of 1D spectra, recorded with the NOESY pulse sequence, in which the evolution time has been set to the first t1 of a 2D spectrum. Using these spectra, it was possible to compute the ratio between $M_k(0)$ and the total amount of magnetization still borne by a proton at any mixing time, $M_k(\tau_m)$, for all spins, k, which exhibit well-resolved resonances.

$$(R_{o,\tau_m}) = \frac{M_k(0)}{M_k(\tau_m)} \tag{2}$$

$M_k(\tau_m)$ is proportional to the area ($S_k(\tau_m)$) under the resonance line in the first t1 spectrum (with the same coefficient of proportionality, equal to A, defined above) and to the sum of the cross-peak volumes:

$$M_k(\tau_m) = A \cdot S_k(\tau_m) = A \cdot \sum_j V_{kj} \tag{3}$$

Using Eqs. 2 and 3, we obtain $M_k(0)$:

$$M_k(0) = R_{o,\tau_m} \cdot \sum_j A \cdot V_{kj} \tag{4}$$

Finally, considering that the initial magnetization, $M_i(0)$, is identical for all protons, the expression of any NOE value is:

$$NOE_{ij}(\tau_m) = \frac{A \cdot V_{ij}(\tau_m)}{M_k(0)} = \frac{V_{ij}(\tau_m)}{R_{o,\tau_m} \cdot \sum_j V_{kj}} \tag{5}$$

In practice, ($M_k(0) = R_{o,\tau_m} \cdot \sum_j V_{kj}$) was calculated for peaks which were isolated in the 1D spectrum, and used as reference for all other protons.

(d) Determination of the relaxation parameters

The strategy used to analyse the conformation and the dynamics of the peptides starts with the extraction of the longitudinal and the cross-relaxation rate constants from the NOESY spectra. These relaxation parameters are arranged in the so-called relaxation matrix, Γ:

$$\Gamma = \begin{pmatrix} \rho_1 & \sigma_{12} & \cdots & \cdots & \sigma_{1N} \\ \sigma_{21} & \rho_2 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{N1} & \cdots & \cdots & \cdots & \rho_N \end{pmatrix}$$

For a dominating dipolar interaction:

$$\rho_i = \alpha \cdot \sum_{j=1}^{N} \frac{1}{r_{ij}^6} \cdot [J(0) + 3J(\omega) + 6J(2\omega)] \tag{7}$$

$$\sigma_{ij} = \alpha \cdot \frac{1}{r_{ij}^6} \cdot [6J(2\omega) - J(0)] \tag{8}$$

For proton–proton relaxation, the constant, α, is equal to 56.92 10^9 Å$^6 \cdot s^{-2}$ · $J(\omega)$ is the spectral density function, which depends on the resonance frequency (ω) and the correlation time of the inter-proton vector motion (τ_c):

$$J(\omega) = \frac{\tau_c}{(1 + \omega^2 \tau_c^2)} \tag{9}$$

A NOE matrix was constructed with the off diagonal elements equal to the NOE values (see above) and the diagonal elements obtained from the integration of the diagonal peaks of the NOESY spectrum. This NOE matrix is related to the relaxation matrix, Γ, through the modified Bloch equation (Bloch, 1946; Solomon, 1955):

$$\frac{dNOE(t)}{dt} = -\Gamma \cdot NOE(t) \tag{10}$$

which can be solved:

$$NOE(t) = \exp(-\Gamma \cdot \tau_m) \cdot NOE(0) \tag{11}$$

The matrix NOE(0) contains only diagonal elements of value −1 for non-equivalent protons. If the full NOE matrix is known, Γ is derived from Eq. 11 by:

$$\Gamma = -(\tau_m)^{-1} \cdot \ln[NOE(t) \cdot NOE(0)^{-1}] \tag{12}$$

The calculation described by Eq. 12 was performed with the program MALAX, written in the laboratory.

The errors made on the σ_{ij} values were approximated, in the following, by the error made on NOEs. A calculation performed on a model (unpublished results), where the NOEs where randomly varied within their allowed limits, showed that this assumption is valid for the kind of σ_{ij} values we have obtained.

(e) 3D structure refinement

The structure calculations in this paper follow the new strategy based on optimal filtering (Koehl et al., 1992). This new procedure, encoded in FILMAN, makes use of a new implementation of the Kalman filter for distance geometry (Altman and Jardetzky, 1986, 1989), using dihedral angles as parameters. It differs from currently proposed methods in that it directly produces estimates of errors on the parameters that are refined, hence providing an image of the minimum that has been found.

Dihedral angles of the main chain of a cyclic molecule are not independent (Go and Sheraga, 1970), therefore a linear analogue of the peptide was considered and geometric cyclization was achieved using nine distance constraints with low initial variances in order to define the pseudo peptide bond with a trans geometry between the side chain of residue 10 and residue 1, in both the D-loop and the K-loop. In addition, a special residue was built in the case of the K-loop in which the succinyl linker group was added to the lysine side chain.

The distances where calculated directly (Eq. 8) from the values of the cross-relaxation rate constants (Eq. 12), using a mean correlation time for each peptide. The errors on the distances were deduced from the errors made on the σ_{ij} values (vide supra).

The quality of the convergence can be evaluated with a khi2 value, which is defined by:

$$\text{khi2} = \sum^{N} \left(\frac{d_{\text{observed}} - d_{\text{calculated}}}{\text{error}} \right)^2 \tag{13}$$

where d_{observed} and $d_{\text{calculated}}$ are the input distance constraints and the distance calculated on the final model, respectively, and error is the error on the calculated distance.

The input for D-loop includes 41 distance constraints related to the backbone geometry, one of which concerns non-adjacent residues. Seven $J_{NH\text{-}H\alpha}$ constraints were also used during the refinement. The input for K-loop was based on 42 short-range distance constraints.

For both peptides, the calculation was started from a set of 400 random initial dihedral angle files. The initial standard deviation on each angle was set to 90°, except for the phi angle of the proline residue, which was kept to 60° with a standard deviation of 5°. The final structures were sorted according to their khi2 values, their van der Waals contact violations, and their maximum error values. Calculations were done on an IBM RS6000 computer and required 30 seconds per conformation.

(f) NOE back-calculations

We used NOE back-calculation as an additional check of our set of final structures. This approach (Banks et al., 1989) consists of building a theoretical relaxation matrix from the coordinates of each proton in the molecule. The NOE were calculated for each structure using PROTNOE, a program developed in our laboratory. Magnetization leakage was introduced by substituting the calculated ρ_i values (see Eq. 6) by the experimental values. The current version of PROTNOE works under rigid body approximation, i.e. the same correlation time is used for all vectors.

A NMR R factor was introduced to quantify the difference between experimental and calculated NOE. This factor is defined by Gonzalez et al. (1991):

$$R = \frac{\Sigma \left(|\,|NOE_{cal}| - |NOE_{obs}|\,| \right)}{\Sigma |NOE_{obs}|} \tag{14}$$

A theoretical NOESY map was then calculated using a lorentzian line shape and the same line width in both dimensions. The map can be used for visual inspection of the quality of the structure with the following qualifications. (1) The overestimation of an experimental NOE or the presence of additional NOE on the calculated map could not be used to invalidate a structure as the absence of an experimental NOE could be due to the effect of fast internal motions. For example, a rigid model such as the one we used in the present approach may produce NOE cross peaks

TABLE 1

A. AMIDE PROTON TEMPERATURE COEFFICIENTS, $^3J_{NH-\alpha}$ COUPLING CONSTANTS AND 500-MHz ASSIGNMENTS FOR D-LOOP IN 90% H_2O (pH 4.0) AT 303 K

Amino acid	H_N dppm/dt ($\times 10^{-3}$)	$^3J_{NH-\alpha}$	Chemical shift (ppm)					
			H_N	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	Others
Cys[1]	6.81	6.1	8.45	4.57	3.20, 3.02			
Lys[2]	6.07	6.0	8.32	4.36	1.88, 1.75	1.43, 1.37	1.65[b]	2.96ε[b]
Arg[3]	5.69	7.7	8.20	4.46	1.92, 1.74	1.58[b]	3.16[b]	
Gly[4]	6.67	ND[a]	8.01	4.09				
Pro[5]				4.39	1.96, 2.22	2.22[b]	3.58[b]	
Gly[6]	8.00	5.9/6.1	8.54	3.91, 4·15				
Ser[7]	7.33	6.0	8.04	4.42	3.82, 3.87			
Asp[8]	7.33	9.4	8.43	4.59	2.56, 2.64			
Phe[9]	4.0	7.0	8.01	4.51	3.03[b]		7.14[b]	7.21ε[b]
Asp[10]	4.45	6.9	8.16	4.62	2.72, 2.82			
Tyr[11]	5.36	7.5	7.70	4.46	2.90, 3.11		7.13[b]	6.53ε[b]

B. AMIDE PROTON TEMPERATURE COEFFICIENTS, $^3J_{NH-\alpha}$ COUPLING CONSTANTS AND 500-MHz ASSIGNMENTS FOR K-LOOP IN 90% H_2O (pH 4.0) AT 303 K

Amino acid	H_N dppm/dt ($\times 10^{-3}$)	$^3J_{NH-\alpha}$	Chemical shift (ppm)					
			H_N	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	Others
Cys[1]	6.29	6.9	8.48	4.77	2.93, 3.23			
Lys[2]	ND[a]	7.4	8.41	4.35	1.75, 1.86	1.40, 1.45	1.66[b]	2.98ε[b]
Arg[3]	6.42	7.2	8.25	4.42	1.74, 1.87	1.60[b]	3.19[b]	
Gly[4]	6.90		8.20	4.12				
Pro[5]				4.44	1.94, 2.30	2.06[b]	3.63[b]	
Gly[6]	7.34	5.7/6.1	8.63	3.96, 4·09				
Ser[7]	3.39	7.2	8.09	4.44	3.82, 3.88			
Asp[8]	ND[a]	6.9	8.39	4.59	2.57, 2.65			
Phe[9]	6.17	7.1	8.09	4.59	2.99, 3.17		7.20[b]	7.33ε[b]
Lys[10]	6.53	6.6	7.95	4.16	1.41, 1.56	1.12[b]	1.40[b]	3.11ε[b], 7.86Nε linker: 2.54[b], 2.56[b]
Tyr[11]	7.09	7.6	7.88	4.53	2.90, 3.12		7.17[b]	6.84ε[b]

[a] Not determined.

[b] One value is given when both protons are equivalent.

between side chain and backbone protons which may not be observable because of internal motion of the side chain. (2) The absence of an observed NOE on the calculated map could be a criterion to reject the structure if the modelled distance is thought to be too large to allow any dynamic NOE.

RESULTS AND DISCUSSION

(a) ^1H resonance assignments

^1H resonance assignments of both the D-loop and the K-loop are shown in Tables 1A and 1B, respectively. The spin system assignments based on HOHAHA spectra were confirmed by sequential assignments on NOESY spectra. Figure 6 shows the sequential connectivities between amide and alpha protons in the finger-print region for both peptides. Small resonance shifts were observed between protons in related positions in the D-loop and the K-loop, which indicate that conformational changes have occurred. Only one set of resonances was found for the proline spin system, suggesting that only one isomer is present in both peptides. This was confirmed by ^{13}C NMR spectra where only one set of resonances has been found for the proline carbons (data not shown). The NOE intensity analysis revealed a close proximity between the alpha protons of Gly^4 and the delta protons of the proline, showing that the proline in both peptides is in the trans conformation.

(b) Temperature coefficients of amide proton resonances

The temperature dependence of amide proton resonances was determined over the range of 278–318 K for both peptides. The chemical shifts of all amide protons varied linearly with temperature, allowing calculation of the temperature coefficients given in Table 1. Most of the slopes found were similar, except for Ser^7 in the K-loop and for Phe^9 and Asp^{10} in the D-loop, which exhibited lower temperature coefficients. The amide proton exchange rate of these residues was lower, indicating that these protected protons might be involved in an intramolecular hydro-

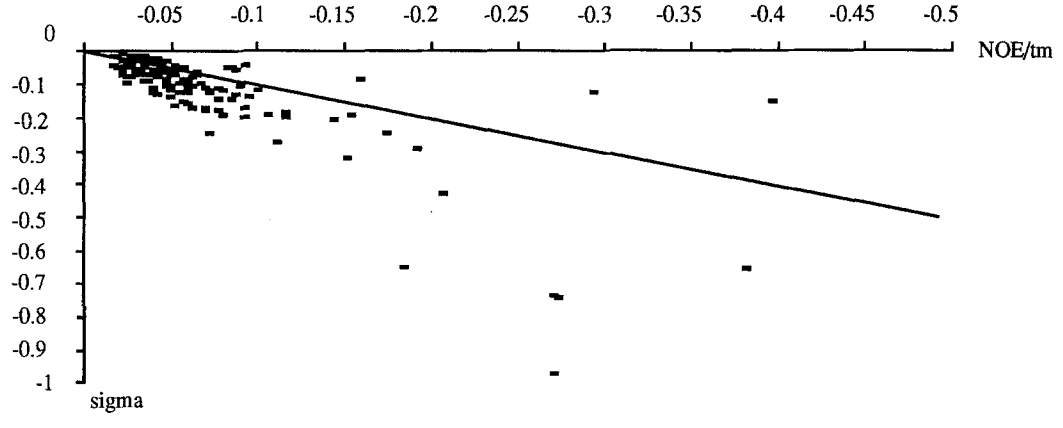

Fig. 2. Plot of sigma values versus NOE/τ_m. The sigma values were extracted according to Eq. 11 from a NOESY experiment performed with a mixing time of τ_m = 400 ms. The solid line represents the expected values of sigma with no spin diffusion (initial slope approximation).

gen bond. In the case of the K-loop, the hydrogen bond involving the Ser^7 amide proton could suggest that the sequence Gly–Pro–Gly–Ser forms a β-turn. The strong NOE (indicated by an arrow in Fig. 6) between the H_α of Pro^5 and H_N of Gly^6 supports this hypothesis since it favours a β-turn structure of type II (Wüthrich, 1986).

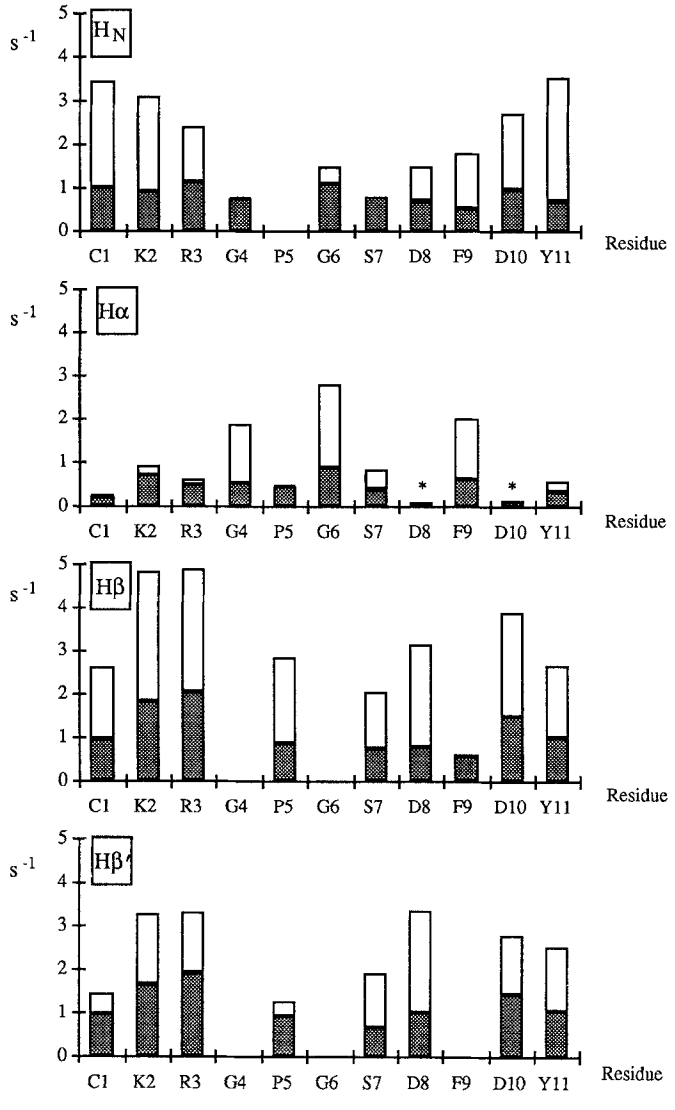

Fig. 3. Superimposition of the longitudinal relaxation rate (in white) and the negative sum of the cross-relaxation rates (in grey) for amide, alpha and beta protons for the D-loop relaxation matrix. The longitudinal relaxation rates are not reported for protons noted with (*), because these values were incorrect. This is probably due to the fact that the resonances of these protons were very close to the water resonance, so that the integration of the diagonal peaks is not very accurate.

(c) Relaxation matrix determination

Despite some overlapping peaks it was possible to determine the full NOE matrices and to calculate directly the relaxation matrices for both peptides using Eq. 12.

Figure 2 shows the cross-relaxation rate constants for both the D-loop and the K-loop versus the corresponding NOE values divided by the mixing time (400 ms in this case). This presentation gives an idea of the strength of the spin diffusion in the molecule. Indeed, if all the NOE built up were in the linear regime (that it, if the magnetization transfers could be modelled by isolated spin pairs), all the points in Fig. 2 would have been on the diagonal. As expected, most of the points were under the diagonal, and correspond to NOE that deviate from the linear regime because of longitudinal relaxation. Some points were above the diagonal, indicating that the corresponding NOE are essentially built by spin diffusion (Keepers and James, 1984; Lefèvre et al., 1987).

The longitudinal relaxation rate (ρ_i) of a given proton was generally larger than the sum of the related cross-relaxation (σ_{ij}) rates in absolute values (Fig. 3). This is partly due to internal motion. The correlation times of the overall motion for both peptides were in the range where the spectral density function at zero frequency is still dominant. In this range of correlation time and above, the non-zero frequency components of the spectral density function can be neglected and ρ_i is equal to the sum of $-\sigma_{ij}$ (see Eqs. 7 and 8). In the region where internal motions decrease the correlation time, as occurs for beta protons, ρ_i becomes larger than the sum of $-\sigma_{ij}$. However, in the case of H_N or beta protons, the longitudinal relaxation rates were much larger than expected. This suggests that relaxation processes other than intramolecular dipolar interactions occur efficiently, such as relaxation by the solvent molecules, quadrupolar relaxation or exchange (for amide protons).

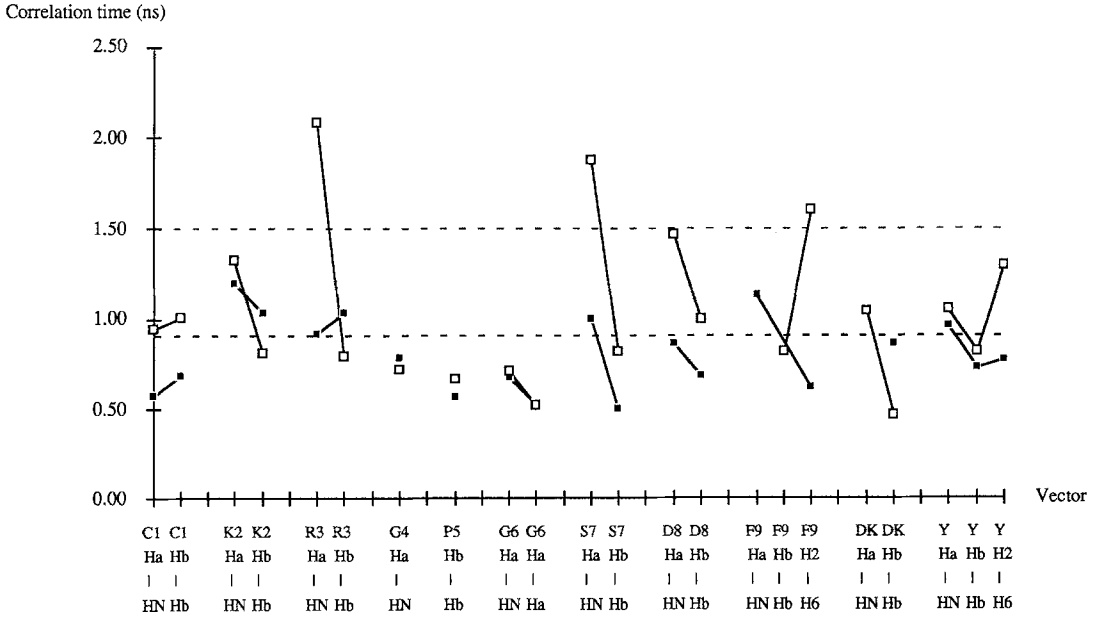

Fig. 4. Correlation times observed along the sequence of D-loop (■) and K-loop (□) for backbone vectors: H_N–H_α or $H_{\alpha 1}$–$H_{\alpha 2}$ of glycines and side-chain vectors H_β–H_β'. The points corresponding to vectors within an individual residue are linked by a line.

(d) Dynamics study

The cross-relaxation rate constants corresponding to vectors with known modules were extracted from the relaxation matrices and the related correlation times were calculated using Eqs. 8 and 9. This was done for the following vectors: the $H_{\alpha 1}$–$H_{\alpha 2}$ of glycines, the $H_{\beta 1}$–$H_{\beta 2}$ and the H_N–H_α. The vectors $H_{\alpha 1}$–$H_{\alpha 2}$ and $H_{\beta 1}$–$H_{\beta 2}$ are pairs of geminal protons, hence they can be considered as fixed in the NMR time scale. The length of the H_N–H_α vectors can be calculated from the coupling constants between these protons. Though this distance is not fixed, it should be noted that a fluctuation of 30° around the actual mean value of the torsion angle between these two protons leads to a maximum distance fluctuation of 0.2 Å. The correlation times obtained by this calculation (Fig. 4) provide insight into the backbone motions ($H_{\alpha 1}$–$H_{\alpha 2}$ and H_N–H_α) as well as the side-chain motions ($H_{\beta 1}$–$H_{\beta 2}$). Large variations in the backbone correlation time were observed along the sequence for both peptides. The mean value was 1 ns for the D-loop and 1.5 ns for the K-loop.

The comparison between these mean correlation times suggests a difference in the overall conformation of the two peptides. The larger correlation time reflects a larger hydrodynamic volume. Thus, compared to the overall conformation of the D-loop, which will be referred to here as compact, the K-loop has a more open-loop conformation. This is supported by the fact that no long-range NOE could be observed between two opposite parts of the K-loop, while the D-loop did exhibit some long-range NOE.

The 200% internal correlation time fluctuation along the backbone observed for both peptides can be considered as significant compared to the 20% estimated error on NOE values.

A more precise analysis of the internal dynamics of these vectors for both peptides was performed using the 'model-free' approach proposed by Lipari and Szabo (1982). In this model, internal motions are defined by a generalized order parameter, S^2, which describes the spatial restriction of the internal motion and by an additional internal correlation time $\tau_i \cdot S^2$ is equal to 1 when no internal motions are present, and 0 when the internal motions are isotropic. If we assume that these motions are fast compared to the overall tumbling of the molecule, this approach consists of scaling all sigma values, calculated for a rigid body (σ_{ij}^0), by the order parameter (Lane and Forster, 1989):

$$\sigma_{ij} = S^2 \cdot \sigma_{ij}^0 \tag{15}$$

We were able to check by ^{13}C relaxation analysis that all the internal correlation times were smaller than 0.1 ns in the case of the D-loop (unpublished results). Hence the values of the correlation times verify the conditions for Eq. 15.

In order to calculate σ_{ij}^0, it is necessary to know the correlation time, τ_c, for the overall tumbling of the molecule. τ_c was taken to be equal to the longest correlation time observed in each molecule. The assumption made here is that the longest correlation time observed corresponds to the overall motion of the molecule, thus the order parameters calculated below should be considered as relative order parameters, the reference being the less mobile vector in the molecule. The order parameters for the backbone vectors are plotted versus the residue number in Fig. 5a. Both peptides showed a region of high flexibility in the sequence Gly–Pro–Gly whereas they seem to be clamped at the level of residues Arg^3 and Phe^9.

Figure 5b shows the variations in the order parameters of the $H_{\beta 1}$–$H_{\beta 2}$ vectors along the

sequence, indicating the motion within each residue. For the D-loop, the side-chain order parameters followed those of the backbone quite closely. This supports the notion that the D-loop has a compact conformation. Such compactness would not permit the side chains to explore a large conformational space. In contrast, the order parameters of the $H_{\beta 1}$–$H_{\beta 2}$ vectors in the K-loop

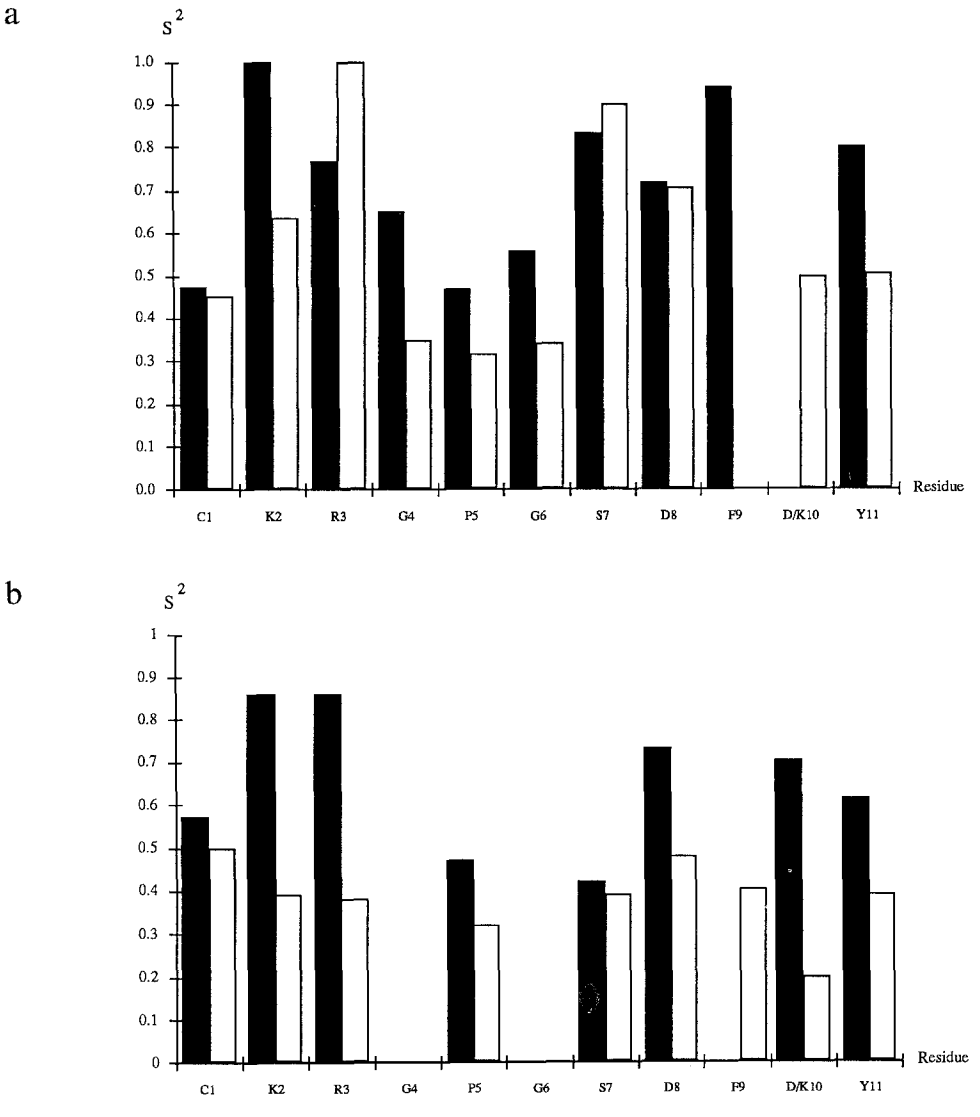

Fig. 5. (a) Relative order parameters, S^2, calculated for backbone vectors for D-loop (■) and K-loop (□). The longest correlation time observed was taken as the overall tumbling correlation time (1.2 ns for the D-loop and 2 ns for the K-loop) and internal motions were assumed to be very fast (see text). The value for the K-loop Phe^9 residue was not determined due to a peak overlapping between intra-residual Phe^9 $H_{N-H\alpha}$ NOE and inter-residual Phe^9 H_N–Lys^{10} H_α NOE. No S^2 value is reported for the Asp^{10} residue of D-loop because this value was expected to be erroneous (see legend of Fig. 3). (b) Order parameters, S^2, calculated for side-chain vectors H_β–H_β' in D-loop (■) and K-loop (□). The missing value for D-loop residue Phe^9 was due to the frequency degeneracy of H_β resonances.

TABLE 2
MAXIMUM ERRORS OBSERVED ON RESULTING STRUCTURES AFTER A FILMAN RUN

D-loop					K-loop				
Structure	Constraints		S.D.	Dist.	Structure	Constraints		S.D.	Dist.
D226	D10 C_γ	C1 N	1.71	0.38[a]	K335	D8 H_α	F9 H_N	1.21	0.60
	D10 H_N	Y11 $H_\beta 1$	1.45	0.80					
	D10 H_N	Y11 $H_\beta 2$	1.38	0.75	K150	D8 H_α	F9 H_N	1.21	0.60
	D10 H_N	D10 $H_\beta 1$	1.20	0.66	K240	G4 O	S7 H_N	1.34	0.60[b]
D129	D10 C_γ	C1 N	2.23	0.50[a]		D8 H_α	F9 H_N	1.21	0.60
	D10 C_γ	C1 C_α	1.58	0.47[a]					
	D10 H_N	Y11 $H_\beta 1$	1.44	0.79	K155	D8 H_α	F9 H_N	1.21	0.60
	D10 H_N	Y11 $H_\beta 2$	1.43	0.78					
	D10 H_N	D10 $H_\beta 1$	1.24	0.68	K183	D8 H_α	F9 H_N	1.20	0.59
D260	C1 H_N	C1 $H_\beta 2$	2.34	1.05	K372	D8 H_α	F9 H_N	1.21	0.60
	D10 H_N	Y11 $H_\beta 1$	1.47	0.81					
	D10 H_N	Y11 $H_\beta 2$	1.45	0.80	K118	D8 H_α	F9 H_N	1.21	0.60
	D10 H_N	D10 $H_\beta 1$	1.26	0.69					
D303	C1 H_N	C1 $H_\beta 2$	2.34	1.05					
	D10 H_N	D10 $H_\beta 1$	2.21	1.21					
	D10 H_N	Y11 $H_\beta 2$	1.40	0.77					
	D10 H_N	Y11 $H_\beta 1$	1.31	0.72					
D174	D10 H_N	D10 $H_\beta 1$	2.24	1.22					
	D10 C_γ	C1 N	1.51	0.34[a]					
	D10 C_γ	C1 C_α	1.44	0.43[a]					
	P5 $H_\beta 2$	G6 H_N	1.43	0.79					
	D10 O_δ	C1 H_N	1.41	0.35[a]					
	D10 H_N	Y11 $H_\beta 1$	1.36	0.75					
	D10 H_N	Y11 $H_\beta 2$	1.35	0.74					
	D8 H_N	F9 H_N	1.33	0.68					
	R3 $H_\beta 1$	G4 H_N	1.27	0.81					
D185	D10 C_γ	C1 N	2.55	0.57[a]					
	D10 H_N	Y11 $H_\beta 1$	2.49	1.37					
	D10 C_γ	C1 C_α	1.81	0.54[a]					
	D8 H_N	F9 H_N	1.26	0.64					
D154	D10 H_N	D10 $H_\beta 1$	2.45	1.34					
	D10 C_γ	C1 N	1.84	0.41[a]					
	D8 H_N	F9 H_N	1.56	0.80					
	P5 $H_\beta 2$	G6 H_N	1.55	0.85					
	D10 C_γ	C1 C_α	1.41	0.42[a]					
	D10 H_N	Y11 $H_\beta 1$	1.39	0.76					

Shown are cases where the error exceeded the estimated standard deviation on measurements by 20% or more. The errors were sorted according to the standard deviation (S.D.). The corresponding distance error is also given (dist.) in Å.

[a] Cyclization constraint.

[b] Hydrogen bonding constraint.

were smaller than the order parameters calculated along the backbone. This is consistent with an open structure of the loop, which would provide more conformational space to each residue.

(e) Computer modelling of 3D structures

For the D-loop, we selected the best seven structures with khi2 values ranging from 0.35 to 0.53 (Fig. 8a and Table 2). Since the pairwise r.m.s.d.-value calculated for the backbone atom positions between these structures was 1.2 Å, we conclude that our dataset is sufficient to define a single conformational family. However, resonance degeneracy for the Hα atoms of Gly^4 leads to uncertainty about the phi angle of this residue. Our dataset is also insufficient to define accurately the position of the C-terminal tyrosine residue, which is located outside the loop and does not belong to the antigenic recognition site.

For the K-loop, two distance constraints were used to model the geometry of a hydrogen bond (Schulz and Schirmer, 1979) between the Ser^7 amide proton and the Gly^4 carboxyl group. The distance between these two atoms was set to 1.9 Å, and an additional constraint of 4.4 Å between the nitrogen atom and the carbon atom was used to ensure the planarity of the bond. Ten structures were selected with khi2 ranging from 0.21 to 0.24 (Fig. 8b and Table 2). The superimposition of these 10 structures also gave a r.m.s.d.-value of 1.2 Å. Thus, despite great fluctuations in the dihedral angle values of Gly^4 and Ser^7, we were able to define the overall shape of the molecule.

The mean dihedral values and the standard deviation calculated over the selected structures of both peptides are given in Tables 3A and B. The standard deviation extracted from the final covariance matrix given by FILMAN is also given. Comparison of the two standard deviation values gives some insight into the significance of these values. In general, both values agreed reasonably well, with the exception of the following cases. (1) The covariance matrix standard deviation was much higher than the one calculated for resulting structures. This was mainly so for the phi angle of Cys^1. Here, the observed discrepancy was due to the use of the linear analogue to mimic a cyclic molecule. The high standard deviation given by FILMAN (53° for the D-loop) reflects that FILMAN does not handle the problem of peptide cyclization globally. Nevertheless, only a single phi value of Cys^1 residue was obtained, which was compatible with a cyclic geometry. This selection gave rise to a very low standarddeviation for the average value of this angle (0.5°). (2) The covariance matrix standard deviation given by FILMAN was much lower than the value calculated over the selected structures. This was observed for phi angles of residue Gly^4 for both peptides and arises when two different values of the dihedral angle are compatible with the input data set. The standard deviation given by FILMAN describes only the minimum near each compatible structure. It is interesting to note that this effect was observed for residues which were affected by high amplitude internal motions, as shown by the dynamics analysis.

(f) NOE back-calculations

For each selected structure, we calculated a set of NOE. The finger-print region of back-calculated and experimental NOESY maps are compared for D-loop structure D226 and K-loop structure K335 in Figs. 6 and 7, respectively. There was a good agreement between the two maps. Additional NOE which appeared on the back-calculated NOESY maps were not used as criteria to reject these structures (see Materials and Methods).

The quantitative difference factors, R, are summarized in Table 2. Calculation of the R factor,

which involves some approximations (vide infra), was restricted to the proton belonging to the backbone. The overall correlation times used to compute the relaxation parameters were the longest correlation times observed for each molecule, i.e., 1.2 ns for the D-loop and 2ns for the K-loop. The R factor lies around 40% for both the D-loop and the K-loop structures. This shows that the quality of the resulting structures was similar in both peptides despite the better khi2 values obtained for the K-loop final structures. Analysis of Table 2 also shows that special care

TABLE 3

A. COMPARISON BETWEEN STANDARD DEVIATION (S.D.) ON THE DIHEDRAL ANGLE VALUES OF THE 7 SELECTED D-LOOP STRUCTURES AND THE STANDARD DEVIATION EXTRACTED FROM THE FINAL FILMAN COVARIANCE MATRIX AFTER 10 FURTHER REFINEMENT STEPS ON STRUCTURE D266

Residue name	PHI angle			PSI angle		
	Average value	Average S.D.	FILMAN S.D.	Average value	Average S.D.	FILMAN S.D.
Cys	−75	0.5	35	159	48	27
Lys	−102	37	36	104	10	27
Arg	−135	31	31	175	16	18
Gly	−51	77	29	−152	22	20
Pro	−71	10	5	154	2	14
Gly	74	9	20	34	23	20
Ser	−155	37	16	59	16	19
Asp	−119	12	25	52	19	16
Phe	−158	3	17	68	5	12
Asp	−159	4	16	67	6	24
Tyr	−154	2	35	−167	87	90

B. COMPARISON BETWEEN STANDARD DEVIATION ON THE DIHEDRAL ANGLE VALUES OF THE 10 SELECTED K-LOOP STRUCTURES AND THE STANDARD DEVIATION EXTRACTED FROM THE FINAL FILMAN COVARIANCE MATRIX OF THE STRUCTURE K335

Residue name	PHI angle			PSI angle		
	Average value	Average S.D.	FILMAN S.D.	Average value	Average S.D.	FILMAN S.D.
Cys	−134	24	53	177	10	16
Lys	−161	14	32	94	7	22
Arg	−114	13	34	168	1	16
Gly	−127	113	17	−144	67	40
Pro	−62	2	5	152	93	18
Gly	71	16	36	−4	14	16
Ser	−42	59	21	169	13	14
Asp	−107	38	26	−60	11	14
Phe	96	18	17	91	6	23
Lys	−60	25	31	87	3	17
Tyr	142	71	42	86	67	90

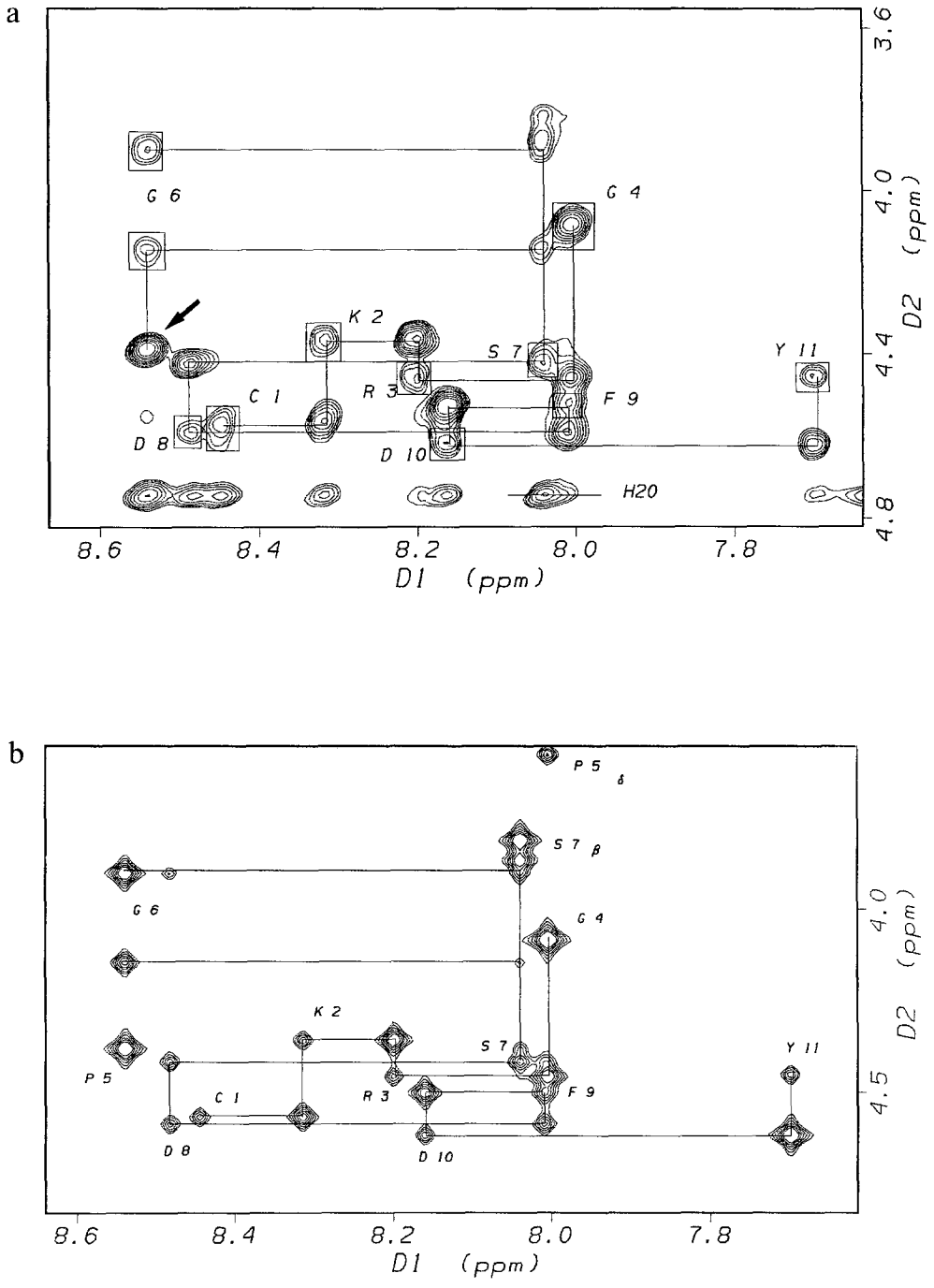

Fig. 6. (a) 400 ms NOESY finger-print region of the D-loop taken at 500 MHz (the NOEs are negative). Assignment pathway is shown by solid lines. Squares identify H_N–H_α connectivities found in a HOHAHA spectrum. The arrow identifies the Gly^6 H_N–Pro^5 H_α connectivity (see text). (b) Back-calculated NOESY spectra of the D-loop structure D226. The overall correlation time was set to 1.2 ns, the line width was 1.5 Hz and the mixing time was 400 ms.

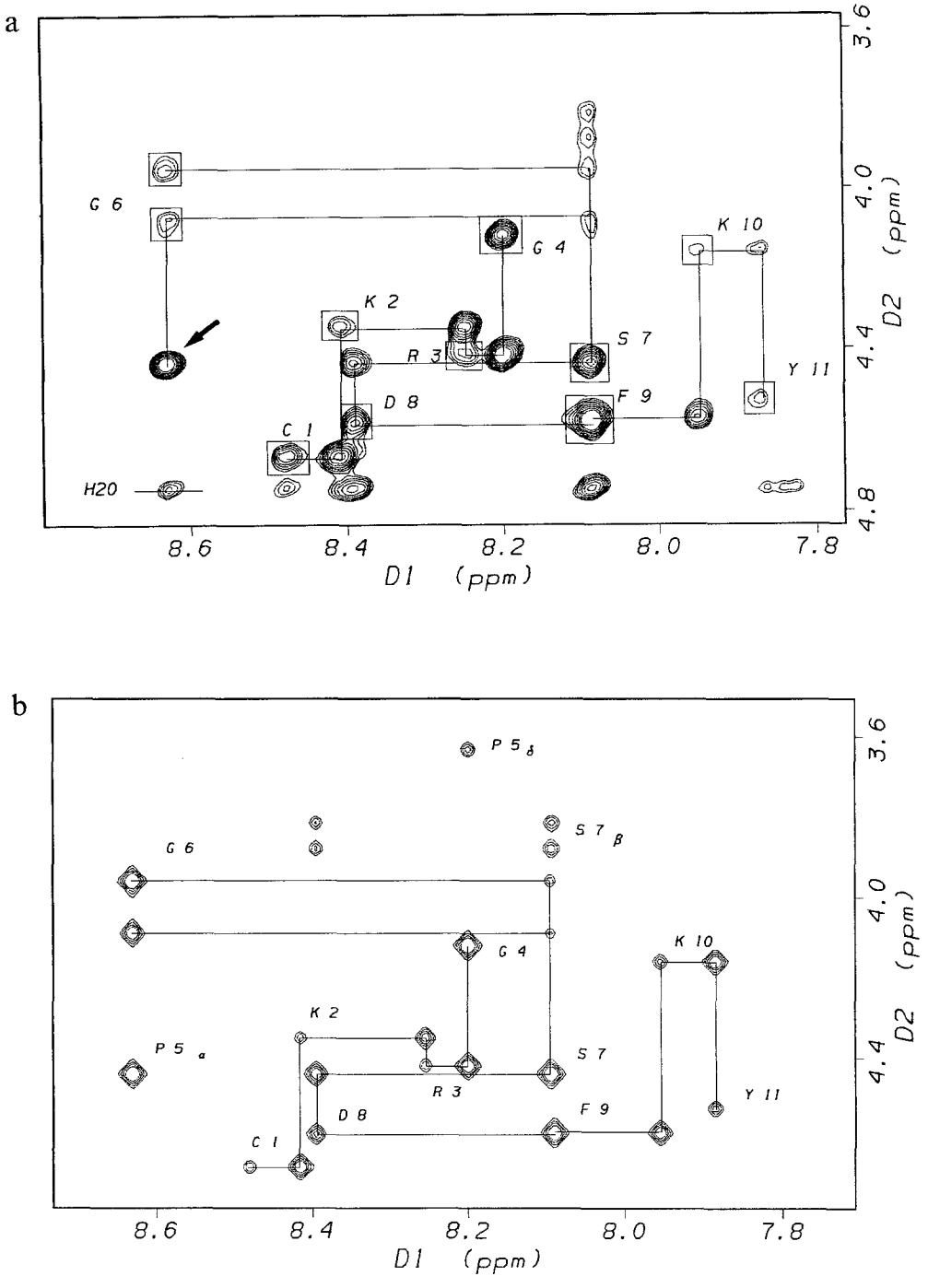

Fig. 7. (a) 400 ms NOESY finger-print region of the K-loop taken at 500 MHz (the NOEs were negative). Assignment pathway is shown by solid lines. Squares identify H_N–H_α connectivities found in a HOHAHA spectrum. The arrow identifies the Gly^6 H_N–Pro^5 H_α connectivity (see text). (b) Back-calculated NOESY spectra of the K-loop structure K335. The overall correlation time was set to 2 ns, the line width was 1.5 Hz and the mixing time was 400 ms.

should be taken when estimating the validity of a structure on the basis of a khi2 value, or the validity of a given structure over alternative structures using a r.m.s.d.-value. This is because these values validate the consistency between the structures found and the set of input constraints but do not validate the structures themselves. Therefore the back-calculation represents a valuable verification that the resultant structure is consistent with the measured NOE. We note however that refinements accounting for the dynamic factors in the calculated sigma values would further improve the accuracy of this procedure.

(g) Analysis of the structure and biological implications

The D-loop and the K-loop exhibited the same overall shape. An extended chain from residues 1 to 4 is broken by the proline residue. The following more flexible Gly–Pro–Gly–Ser sequence is involved in a turn structure. This type IIβ-turn has been identified in both peptides on the basis of short-range NOE connectivities (Wüthrich, 1986; Dyson et al., 1988) between Pro^5 Hα and Gly^6 H_N, and between Gly^6 H_N and Ser^7 H_N, which are identical in both peptides and therefore yield the same dihedral angle values. However, the amide temperature coefficient studies suggest that the Ser^7–Gly^4 intermolecular hydrogen bond is found only in the K-loop. Although we cannot completely eliminate the possibility that this hydrogen bond is present in the D-loop Gly–Pro–Gly–Ser sequence, its life time should be shorter than in the K-loop.

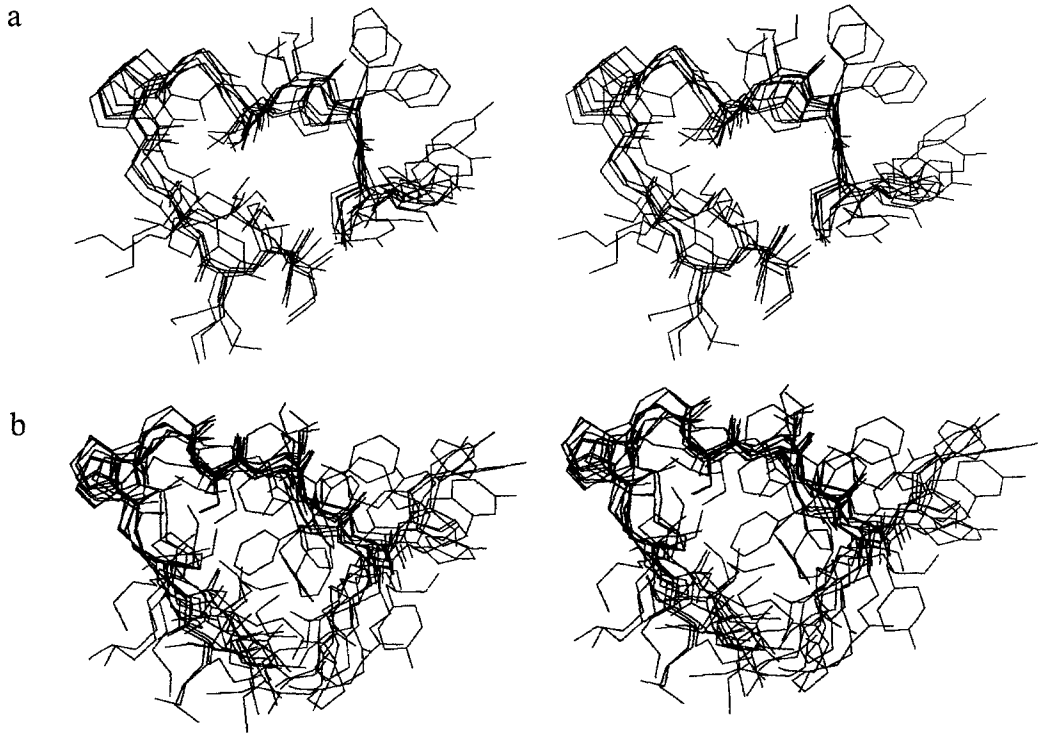

Fig. 8. (a) Superimposition of 7 D-loop structures: The average r.m.s.d. calculated over these 7 structures for the backbone atoms of residues Cys^1–Asp^{10} was 1.2 Å. (b) Superimposition of 10 K-loop structures: The average r.m.s.d. calculated over these 10 structures for the backbone atoms of residues Cys^1–Lys^{10} was 1.2 Å.

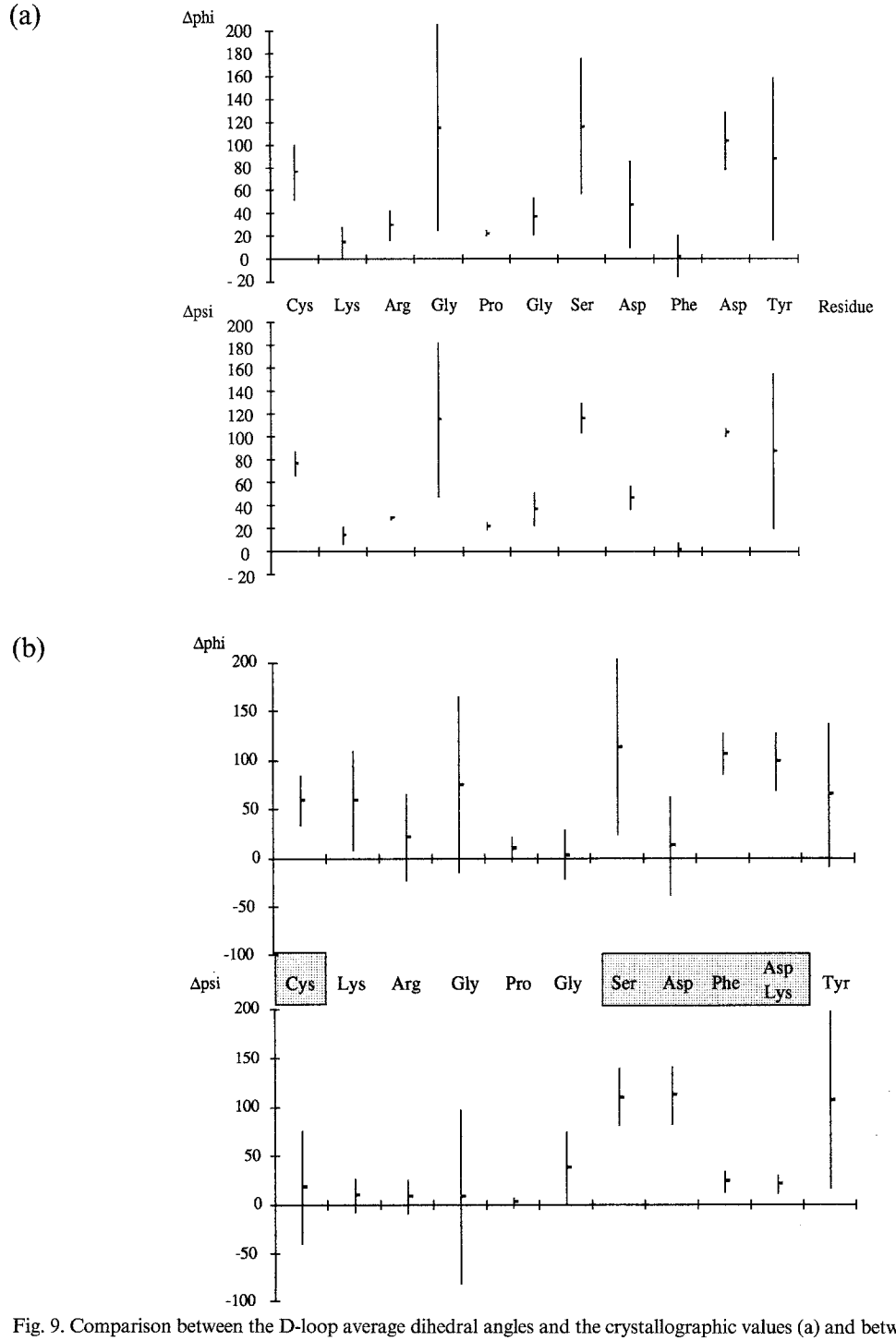

Fig. 9. Comparison between the D-loop average dihedral angles and the crystallographic values (a) and between the D-loop and the K-loop average values (b). Ordinate axes represent the absolute difference value between the average phi angles (top) and average psi angles (bottom) of the two structures. The vertical lines represent the standard deviation on NMR-derived values. The box indicates residues of which the conformation was noticebly affected by the replacement of Asp^{10} by Lys^{10}.

The slow exchange rate observed for the D-loop amide proton of Phe^9 could be explained by a low solvent accessibility since this proton is pointing inside the loop.

Comparison of the D-loop structure with the structure found for residues 139–148 in the crystallographic structure shows that most of the residues adopt a structure with nearly the same values for the phi and psi dihedral angles, except for Gly^4, Ser^7 and Asp^{10} (Fig. 9a). For Asp^{10} the difference could be explained by structural rearrangements caused by cyclization. For Gly^4 and Ser^7 the comparison is less significant, because of uncertainty about the position of these residues in the crystallographic structure as well as in our own structures.

More instructive is the comparison of the D-loop with the K-loop structures in terms of dihedral angle values (Fig. 9b). This comparison should take into account the flexibility of both molecules. The translation of the order parameter, S2, into dihedral angle fluctuations is not straightforward. A rough estimate of the amplitude of the motion experienced by a proton–proton vector in the molecule can be obtained through the model proposed by Woesner (1962), where the vector is supposed to move isotropically within a cone of semi-angle θ:

$$S = (1/2)\cdot\cos\theta\cdot(1+\cos\theta)$$

Thus, S2 = 0.8 corresponds to a value of θ of about 20° and S2 = 0.5 implies a θ value of about 40°. To a very rough approximation, these values of θ may be taken as maximum values for the dihedral angle directly related to the proton–proton vector (i.e. the angle phi for the $H\alpha$–H_N vector). The principal value of this approach is to provide some idea of the order of magnitude of the amplitude of the motion. It is by no means an attempt to accurately calculate the dihedral fluctuation angle. However, the result can be used to estimate whether two structures are actually different or might be superimposed as a result of internal motion. For instance, in the Gly–Pro–Gly sequence, where the observed S2 values were about 0.5, differences in dihedral angles of up to 80° between K-loop and D-loop might not represent divergence between the two peptide structures because they are within the flexibility of the molecules.

Differences in the dihedral angles plotted in Fig. 9b show that the substitution of Asp^{10} by a lysine residue affected the conformation of Cys^1 and the four residues going from Ser^7 to Asp^{10}, whereas the other residues were unaffected. The main structural changes occurred opposite to the turn and affected only one side of the loop. It is interesting to note that these changes affected the conformation of the two residues Cys^1 and Phe^9, which have been shown to be invariant in several haemagglutinin mutants (Wiley et al., 1981).

In Fig. 10, we present two structures, D266 and K335 (see Table 2), of the D-loop and the K-loop, respectively, in which the backbone atoms of residues Gly^4 to Ser^7 have been superimposed with a local r.m.s.d. value of 0.7 Å. In this case, the overall shape of the two structures of the D-loop and the K-loop, chosen from Table 2 (Fig. 10), shows that the latter is more open at the level of residues 8–10. It is possible to superimpose the backbone atoms of residues Gly^4 to Ser^7 with a low r.m.s.d. (0.7 Å). In this case, the overall backbone atom r.m.s.d.-value was equal to 5.3 Å. The difference in the behaviour of the two peptides could then be explained by steric hindrance. The flexibility of the antigen could play a role in the surface adaptation of the molecule to the antibody. At this level, both peptides are equivalent as they exhibit similar local flexibility. However, this adaptation step follows first a recognition process, during which the antigen site enters the antibody pocket. The larger volume occupied by the K-loop would impede the forma-

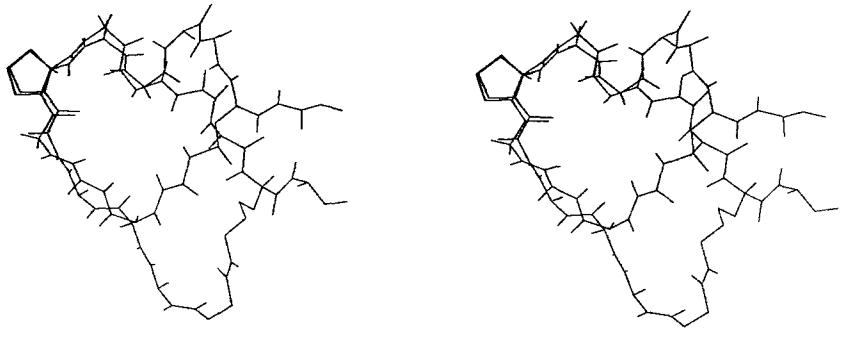

Fig. 10. Superimposition of backbone atoms of D-loop structure D226 and K-loop structure K335 (see Table 2) using backbone atoms of residues Gly^4–Ser^7. The r.m.s.d. calculated over all backbone atoms was 5.3 Å.

tion of the complex and lower its affinity constant. It should also be noted that the presence of three mutation points (Thr^{122}, Asn^{133} and Asn^{137}) located outside the variable loop on the haemagglutinin protein surface (Wiley et al., 1981) indicates that the epitope is larger than the loop alone, therefore the side chain of Lys^{10} may interfere with the antibody region that recognizes these residues.

In this study, we compared the structure of two cyclic peptides with respect to their internal motions. We showed that motions with different amplitudes occur along their backbone. This different dynamic behaviour is due to the intrinsic nature of amino acids and the overall geometry of the peptide. Gly–Pro–Gly–Ser appears to be a segment of high mobility and contains residues which have a high mutation frequency and which are likely to determine the antigenic properties of the different strains of influenza virus (Wiley et al., 1981). Studies on the dynamic and structural properties of other mutant loops are currently in progress. The D-loop structure could provide a 'framework' model to design the geometry of other types of antigens. These studies should give further insight into the role of flexibility in antigen–antibody recognition processes.

ACKNOWLEDGEMENTS

B.K. was supported by a fellowship from the Ministère Français de la Recherche et de la Technologie. We thank Dr. Iain Lamber for careful reading of the manuscript. We gratefully acknowledge the Merrell Dow Research Institute in Strasbourg for giving us access to their Bruker AM-500 MHz spectrometer.

REFERENCES

Altman, R. and Jardetzky, O. (1986) *J. Biochem. (Tokyo)*, **100,** 1403–1423.
Altman, R. and Jardetzky, O. (1989) *Methods Enzymol.*, **177,** 218–246.
Banks, K.M., Hare, D.R. and Reid, B.R. (1989) *Biochemistry,* **28,** 6996–7010.
Bloch, F. (1946) *Phys. Rev.*, **70,** 460–474.
Boelens, R., Koning, T.M.G. and Kaptein, R. (1988) *J. Mol. Struct.*, **173,** 299–311.
Boelens, R., Koning, T.M.G., Van der Marel, G.A., Van Boom, J.H. and Kaptein, R. (1989) *J. Magn. Reson.*, **82,** 290–308.

Borgias, B.A. and James, T. (1990) *J. Magn. Reson.*, **87,** 475–487.
Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M. and Webster, R.G. (1987) *Nature,* **326,** 358–363.
Davis, D.G. and Bax, A. (1985) *J. Am. Chem. Soc.*, **107,** 2820–2821.
Dyson, H.J., Rance, M., Houghten, R.A., Lerner, R.A. and Wright, P.E. (1988) *J. Mol. Biol.*, **201,** 161–200.
Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) *J. Magn. Reson.*, **90,** 420–425.
Glaudemans, C.P.J., Lerner, L., Daves, D.G., Kovàc, P., Venable, R. and Bax, A. (1990) *Biochemistry,* **29,** 10906–10911.
Go, N. and Sheraga, H.A. (1970) *Macromolecules,* **3,** 178–187.
Gonzalez, C., Rullmann, J.A.C., Bonvin, A.M.J.J., Boelens, R. and Kaptein, R. (1991) *J. Magn. Reson.*, **91,** 659–664.
Keepers, J.W. and James, T.L. (1984) *J. Magn. Reson.*, **75,** 404.
Koehl, P. and Lefèvre, J.F. (1990) *J. Magn. Reson.*, **86,** 565–583.
Koehl, P., Kieffer, B. and Lefèvre, J.F. (1989) *Protein Structure and Engineering,* Erice, Sicily, Italy (N. ASI) Plenum Press, New York, London, pp. 139–154.
Koehl, P., Lefèvre, J.F. and Jardetsky, O. (1992) *J. Mol. Biol.*, **223,** 299–315.
Lane, A.N. and Forster, M.J. (1989) *Eur. Biophys. J.*, **17,** 221–232.
Lefèvre, J.F., Lane, A.N. and Jardetzky, O. (1987) *Biochemistry,* **26,** 5076–5090.
Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104,** 4546–4570.
Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41,** 95–117.
Mariuzza, R.A., Philips, S.E.V. and Poljak, R.J. (1987) *Ann. Rev. Biophys. Chem.*, **16,** 139–159.
Merrifield, R.B. (1963) *J. Am. Chem. soc.*, **85,** 2149–2153.
Mirau, P.A. (1988) *J. Magn. Reson.*, **80,** 439–447.
Muller, S., Plaue, S., Samama, J.P., Valette, M., Briand, J.P. and Van Regenmortel, M.H.V. (1990) *Vaccine,* **8,** 308–314.
Nilges, M., Habazettl, J., Brünger, A.T. and Holack, T.A. (1991) *J. Mol. Biol.*, **219,** 499–510.
Plaue, S. (1990) *Int. J. Pept. Prot. Res.*, **35,** 510–517.
Schulz, G.E. and Shirmer, R.H. (1979) In *Principles in Protein Structure* (Ed, Cantor, C.R.) Springer–Verlag, New York, pp. 33–36.
Solomon, I. (1955) *Phys. Rev.*, **99,** 559–565.
Westhof, E., Altschuh, D., Moras, D., Bloomer, A.C., Mondragon, A., Klug, A. and Van Regenmortel, M.H.V. (1984) *Nature,* **311,** 123–126.
Wiley, D.C., Wilson, I.A. and Skehel, J.J. (1981) *Nature,* **289,** 373–378.
Williams, R.J.P. (1989) *Eur. J. Biochem.*, **183,** 479–497.
Wilson, I.A., Skehel, J.J. and Wiley, D.C. (1981) *Nature,* **289,** 366–373.
Woesner, D.E. (1962) *J. Chem. Phys.*, **37,** 647–656.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* Wiley, New York.
Wüthrich, K. and Marion, D. (1983) *Biochem. Biophys. Res. Commun.*, **113,** 967–974.