Journal of Computer-Aided Molecular Design, 9 (1995) 160–170
ESCOM

J-CAMD 284

Structure determination of a tetradecapeptide mimicking the RXVRG consensus sequence recognized by a *Xenopus laevis* skin endoprotease: An approach based on simulated annealing and ^{1}H NMR

Sonja Meddeb[a], François-Regis Chalaoux[b], Jean-Pierre Ballini[a], Daniel Baron[b,c], Paul Vigny[c,d] and Jean-Philippe Demaret[a,*]

[a]*L.P.C.B., Institut Curie et Université Paris VI, 11 rue Pierre et Marie Curie, F-75231 Paris Cedex 05, France*
[b]*L.A.S.I.R., C.N.R.S., 2 rue Henry Dunant, F-94320 Thiais, France*
[c]*Université Paris VI, F-75252 Paris Cedex 05, France*
[d]*C.B.M., C.N.R.S., 1 rue Charles Sadron, F-45071 Orléans Cedex 2, France*

Received 4 August 1994
Accepted 14 November 1994

Keywords: Conformational search; Endoprotease; Maturation; Peptide

Summary

The tetradecapeptide of sequence H-Asp-Val-Asp-Glu-**Arg^5-Asp-Val-Arg-Gly^9**-Phe-Ala-Ser-Phe-Leu-NH_2 is recognized by a putative maturation endoprotease of the *Xenopus laevis* skin, which cleaves between Arg^8 and Gly^9. A conformational search has been performed on this peptide by simulated annealing calculations. Two different models in agreement with the NMR data were found. The conformational difference between the two types of model is located in the consensus sequence, i.e., from Arg^5 to Gly^9.

Introduction

Some maturation proteases require a relatively long specific sequence at the recognition site, such as RXXR or RXRR, involving two or three arginine residues [1,2]. It has been suggested that a structural organization in the environment of the cleavage site may also be required [3]. In most cases, the cleavage site was included in a β-turn, or very close to a β-turn, as shown by secondary structure predictions [3,4] and spectroscopic studies [5]. Moreover, other highly ordered domains (β-sheet, α-helix) were very often found to flank the cleavage signal [3,5]. This phenomenon has been reviewed by Darby and Smyth [6].

Consequently, when the recognition sequence is a long one (such as Leu-Leu-Ala-Gly [7] or Arg-Xaa-Val-Arg-Gly [8]) the question arises whether the processing enzyme recognizes the primary sequence of the fragment and/or a secondary structure generated by this sequence [6].

Maturation of antibacterial precursors in the *Xenopus laevis* skin liberates approximately 20 peptides at the level of a consensus sequence, Arg-Xaa-Val-Arg-Gly [9,10] (Xaa being any single amino acid), by hydrolysis between Arg and Gly [8]. These peptides, beginning with a glycine residue, have been identified in the *Xenopus laevis* exudate and have been shown to possess antibacterial properties. A single endoprotease was expected to be involved in the above cleavages [9]. A synthetic tetradecapeptide, H-Asp-Val-Asp-Glu-**Arg^5-Asp-Val-Arg-Gly^9**-Phe-Ala-Ser-Phe-Leu-NH_2, mimicking the environment of the recognition site, allowed the isolation of the putative maturation enzyme named RXVRG endoprotease [9].

In order to obtain further information about the recognition process, an NMR structure determination of this synthetic tetradecapeptide was achieved by 2D NMR studies (COSY and NOESY) [11]. This type of information could be a very helpful tool in a further protein-engineering approach to these problems. The low solubility of the peptide in water led to a study in DMSO solution. Furthermore, it has been emphasized by Saulitis et al. [12] that protein–ligand association generally involves both lipophilic and polar moieties, and DMSO could be better suited than water to mimic the environment of the interaction site between the enzyme and the propeptide.

The main results from the DMSO-d_6 solution study

*To whom correspondence should be addressed.

0920-654X/$ 6.00 + 1.00 © 1995 ESCOM Science Publishers B.V.

showed several structural trends in the consensus domain: (i) low values of the thermal coefficient of the NH chemical shift in Asp[6], Val[7] and Arg[8]; (ii) ψ angles around −20° for Val[7] and Arg[8]; (iii) 71 NOE distances: two medium-range NOEs of the $H_i^{\alpha}\cdots NH_{i+2}$ type, involving the pairs 5–7 and 8–6, 28 distances between adjacent residues and 41 intraresidual distances. With regard to the dihedral restraints, realistic bounds were obtained for most of the φ and ψ domains. These have been determined, for each residue, from the NMR data which depend on these angles, i.e., the scalar coupling constant $J_{N\alpha}$ ($HNC^{\alpha}H$) and five types of NOE involving at least one NH. Two dihedral restraints have a domain up to 80° (±30°), 18 have a domain up to 60° (±30°), while 11 restraints were better defined, with a more limited domain up to 30° (±15°).

Here, we report the results of an extensive exploration of the conformational space accessible to this peptide within the limits of NMR restraints, using a simulated annealing procedure.

Materials and Methods

Structure generations, simulated annealing calculations and energy minimizations were performed with the program XPLOR, v. 3.1 [13], on a Silicon Graphics Personal Iris 4D-35 workstation.

The starting structures used for the simulated annealing protocol consisted of a template set of coordinates in which the peptide backbone was extended along the x-axis, with randomized y- and z-coordinates. The structure was regularized to give good local geometry with no bad contacts, using a short period of dynamics and subsequent energy minimization (200 steps of the Powell algorithm [14]).

The basis of simulated annealing involves a temperature increase of the system, followed by slow cooling in order to overcome local minima and to locate the global minimum region of the target function [15].

Target function

For the structure determination based on NMR distances, the following target function was used:

$$F_{total} = F_{bond} + F_{angle} + F_{imp} + F_{repel} + F_{NOE} + F_{tor} \qquad (1)$$

where F_{bond}, F_{angle} and F_{imp} are harmonic functions which describe the covalent energy terms for bond lengths, bond angles, planarity and chirality, respectively (F_{imp} also constrains the backbone dihedral angles). In order to maintain the correct stereochemistry during the calculations at high temperature, high uniform constant force values were used, set to 1000 kcal mol^{-1} $Å^{-2}$ for bond lengths and 500 kcal mol^{-1} rad^{-2} for bond angles and improper torsions. (It should be noted that these values are somewhat arbitrary.)

F_{repel} is a single hard-sphere repulsion term which replaces all the nonbonded energy terms used in molecular dynamics (i.e., the Lennard-Jones van der Waals, hydrogen-bonding and electrostatic potentials):

$$F_{repel} = \begin{cases} 0 & \text{if } r > k_{rep}\, r_{min} \\ C_{rep}\,(k_{rep}^2\, r_{min}^2 - r^2)^2 & \text{if } r < k_{rep}\, r_{min} \end{cases} \qquad (2)$$

C_{rep} is an energy constant set to 1 in our simulation; k_{rep} specifies the factor by which the van der Waals radius r_{min} should be multiplied; it was set to different values during the simulations.

In a structure exhibiting a large number of close contacts, the Lennard-Jones potential would produce a very large gradient, while the repel function avoids this problem during the initial steps of the energy minimization [16].

F_{NOE} and F_{tor} introduce penalties when the interproton distances (d) or torsion angles (φ) deviate from the acceptable value ranges which were experimentally determined.

F_{NOE} was modelled by the 'soft square' potential for the simulated annealing calculation [13] and by the 'square' potential in the refinement protocol [13]. For the 'square' potential, the upper bound of the force constants was set to 1000, the scale factor S was set to 50, the additional scale factor C was set to 1 and the exponent was set to 2. For the 'soft square' potential, the switch between square well and asymptotic r_{sw} was set to 0.5, the asymptotic constant c was set to 0.1 and the soft exponent (softexp) was set to 1.

The restraints on the interproton distances involving methylene protons without stereospecific assignments were applied between the geometric centers of the atoms.

F_{tor} had the overall weight factor S set to different values during the simulations, the constant force C to 1 and the exponent of the restraining potential, ed, to 2.

A weight is used for each energy term of the Hamiltonian and can be modified during the calculations in order to obtain a good exploration of the conformational space and a good compromise between the experimental data and a correct internal geometry. At the beginning of the simulation at high temperature, a weak weight on the repel function and on the angles and improper potentials enables atoms to interpenetrate and thus facilitates exploration of conformational space within the limits imposed by the experimental restraints, for which a heavy weight is used. During the cooling and refinement steps, the weight on the internal geometry, and particularly on the repulsive term, must be increased to obtain a realistic structure with standard values for bonds, bond angles and impropers and in agreement with the experimental data.

It should be noticed that the target function to be minimized has no real physical meaning, since it is not complete (no Lennard-Jones van der Waals or hydrogen-bonding or electrostatic potential is taken into account) and the restraining potentials used are only mathematical implementations which have no physical sense.

Molecular dynamics was performed (by means of the Verlet algorithm [17]) using the T coupling method [18] for temperature control. This method is essentially a Langevin-type method with zero random forces and a scaled friction coefficient.

The m_i masses were all taken equal to 100 atomic mass units, to speed up the molecular dynamics calculations.

Simulated annealing and refinement protocol

In the simulated annealing protocol, identical weights were employed for all NOE constraints. The standard parameters of XPLOR, parallhdg.pro and topallhdg.pro, were used.

Initial minimization

First, 100 steps of minimization with the Powell algorithm were performed, using an asymptotic slope of 0.1 for the NOE soft square potential and a value of 1 for the k_{rep} factor which is multiplied with the van der Waals radius (R_{min}); the scale of dihedral restraints was set to 5.

High-temperature molecular dynamics

The MD calculations started with 80 ps at 1000 K, with a time step of 0.002 ps, using weights of 0.4, 0.1 and 0.002 for bond angles, improper torsion and repulsion energy terms, respectively (these weak values allow atoms to pass through each other); the initial velocities were chosen from a Maxwellian distribution.

Subsequently, 40 ps of MD were performed at 1000 K, with a time step of 0.002 ps, increasing weights on the geometry (all covalent energy terms having a weight of 1) and modifying the slope for the NOE potential to 1.

Cooling of the system

This was done with 60 ps MD with a time step of 0.002 ps, decreasing the temperature from 1000 K to 100 K with steps of 50 K in 18 cycles of 3.33 ps; the weight on the repel function was increased from 0.003 to 4, while decreasing the k_{rep} factor from 0.9 to 0.75; a scale factor of 200 was used for dihedral restraints.

Minimization

Five hundred steps of minimization with the Powell algorithm were performed, conserving a weight of 4 for the repel function.

Refinement and final minimization

Sixty ps of MD with a time step of 0.001 ps were carried out, immerging the previously minimized structure in a bath of 1000 K and decreasing the temperature to 100 K employing the characteristics and parameters used in the first cooling stage. Then, 1000 steps of minimization with the Powell algorithm were performed, followed by 200 steps of minimization with the Powell algorithm using a Lennard-Jones van der Waals potential to refine the nonbonded contacts; a scale factor of 250 was used for the dihedral restraints.

Energy calculation on the final structures

Thousand steps of minimization with the Powell algorithm were carried out, considering all internal conformational energy terms (bonds, bond angles, impropers and dihedrals) and the following nonbonded energy terms: explicit hydrogen-bond term, Lennard-Jones potential energy function for the van der Waals interactions, and electrostatic energy term (with a dielectric constant of r). The standard parameters of XPLOR, files parmallh3x.pro and topallh6x.pro, were used. A harmonic constraint potential was added to the atoms to restrain them as close as possible to their initial positions. The arginine, aspartate and glutamate residues were protonated, but a virtual counterion was added to every net charge, hence neutralizing it.

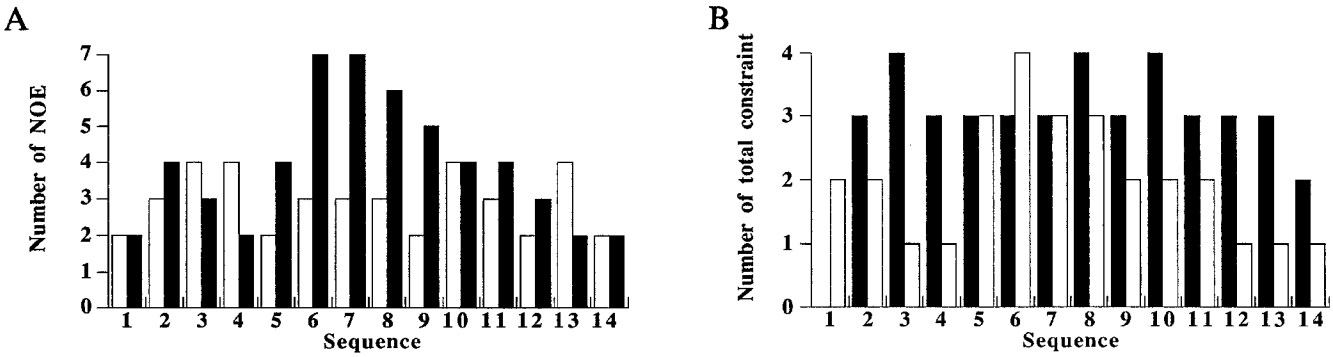

Fig. 1. (A) Number of intra- and interresidual NOEs. The intraresidual interactions consist of $H^{\alpha}N$, $H^{\beta}N$ and H^{α},H^{β} NOE restraints (white bars), while the interresidual ones consist of the $H^{\alpha}N$, $H^{\beta}N$ and NN restraints (black bars). (B) Number of restraints used for the φ and ψ angle determinations. The φ angles depend on the $^3J_{N\alpha}$ coupling constant and on the $H^{\alpha}N$, $H^{\beta}N$ and NN NOEs (black bars); the ψ angles depend on the $H^{\alpha}N$, NN and $H^{\beta}N$ NOEs (white bars).

Results and Discussion

The 71 NOE distances and 20 dihedral intervals experimentally determined were used as restraints in the simulations. Figure 1 shows the number of restraints for each residue. Computations using all the experimental data led to an incompatibility between the NOE distance restraints, since some of them were systematically severely violated. We have therefore considered two classes, each characterized by two specific NOE restraints. These two sets of NOE restraints are mutually incompatible: the NOEs $H^{\alpha}7$-HN8 and $H^{\alpha}5$-HN7 were considered for class #1 only (Fig. 2A), while the restraints HN8-$H^{\alpha}8$ and $H^{\beta}7$-HN8 were considered for class #2 only (Fig. 2B). Figures

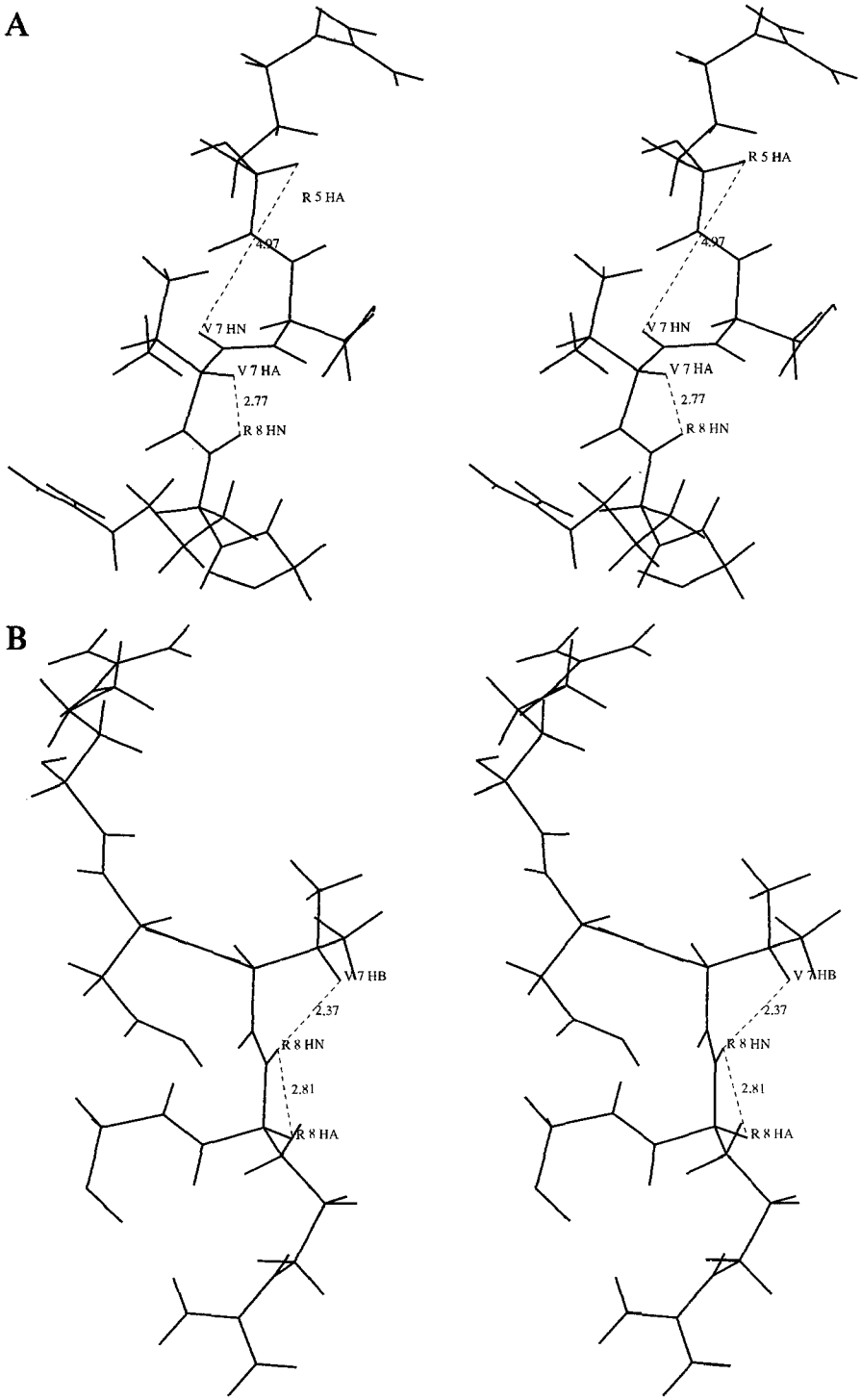

Fig. 2. NOE restraints (represented with dashed lines), specific to each class, in the stereoviews of residues 5 to 9 of the lowest energy structures of (A) class #1 and (B) class #2. The main chain is represented in bold.

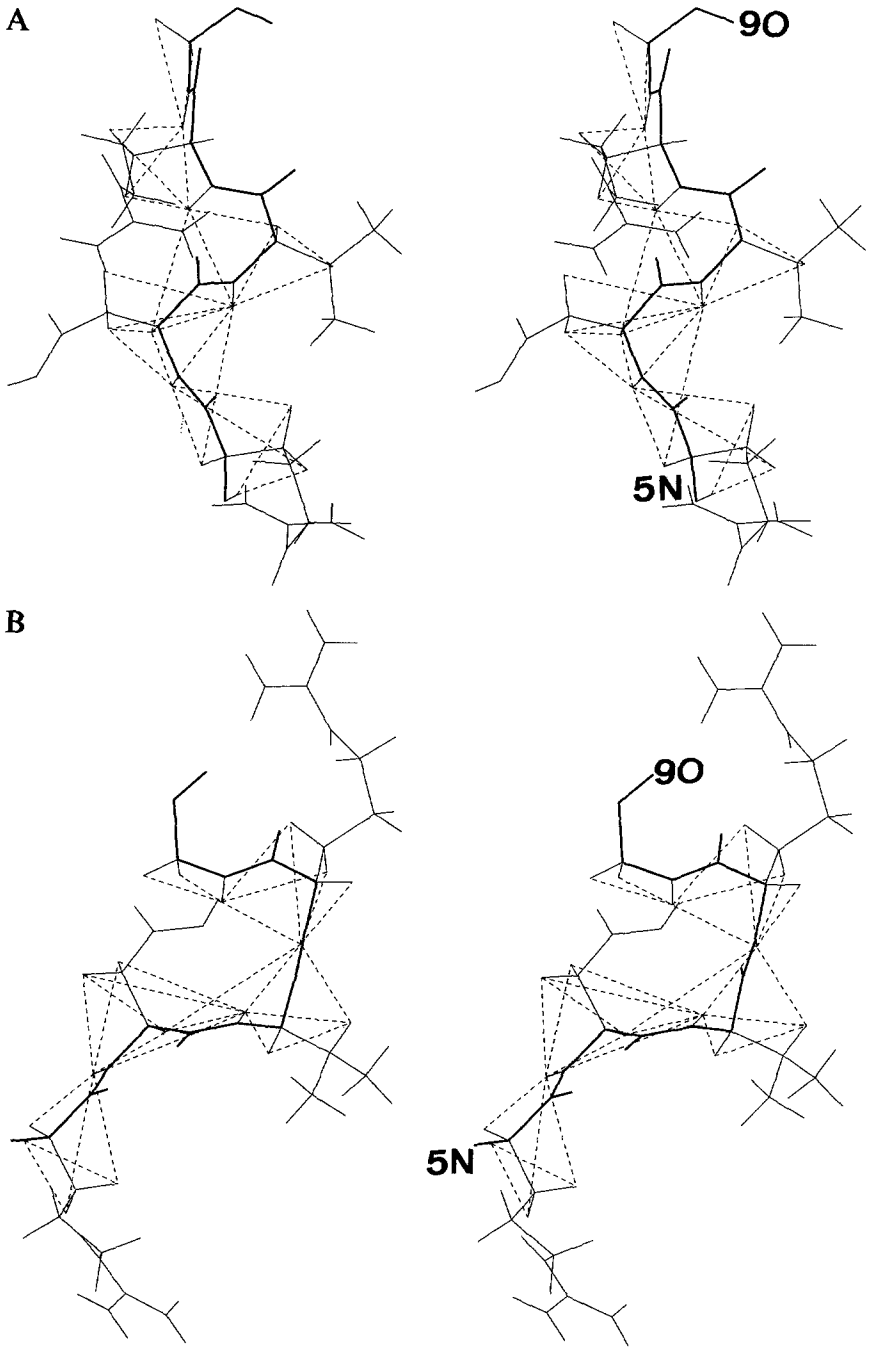

Fig. 3. The NOE restraints (represented with dashed lines), responsible for the structuration, in the stereoviews of residues 5 to 9 of the lowest energy structures of (A) class #1 and (B) class #2. The main chain is represented in bold.

3A and B represent the NOE restraints of classes #1 and #2, respectively, that are responsible for the structuration in the consensus sequence.

The simulations performed with each group of data led to no distance restraint violations greater than 0.5 Å and no dihedral angle restraint violations greater than 5°, with an almost perfect internal geometry: the rms difference for bond deviations from the ideal geometry was less than 0.01 Å and the rms difference for angle and improper deviations was less than 2°.

To fully explore the conformational space for each class, 200 structures were calculated by submitting the initial random conformations to simulated annealing and refinement.

Analysis of the computed structures

Energetic aspects

The mean values, over the 200 structures of the two classes, of each energy term considered during the simulations are given in Table 1.

In each class, the obtained structures are energetically close. For class #1, the mean value and the standard deviation of the total energy are 49.0 and 3.2 kcal mol^{-1}, respectively, while for class #2 these values are 37.6 and 3.4 kcal mol^{-1}, respectively. The deviation from idealized geometry is very small, as evidenced by the low values of F_{bond}, F_{angle} and F_{imp}, which are homogeneous over each class as shown by the low values of the rms difference for these terms. The absence of bad nonbonded contacts is illustrated by the negative value obtained for the van der Waals energy term. The large value of F_{NOE} is probably due to the compromise which has to be made with respect to the internal ideal geometry and the NOE restraints within 0.5 Å, so that the distances are often close to the limits of the allowed values. The restraints over the dihedral angles are easier to satisfy, as shown by the low value of F_{tor}.

When calculating and averaging the 'true' energy over the 200 structures (i.e., using the parameters of the standard XPLOR force field, comprising full van der Waals and electrostatic interactions as well as hydrogen bonds), no significant difference between the two classes can be observed. As shown in Table 1, the difference is less than 2.2 kcal mol^{-1} for each energy term; to be compared with the standard deviation within each class, which ranges from 0.1 to 5.7 kcal mol^{-1}. Accordingly, for energy computations performed in vacuo, both classes are identical in terms of stability.

Stereoviews of the lowest energy structure for classes #1 and #2 are represented in Figs. 4A and B, respectively. The conformations are more or less extended, so that the attractive part of the van der Waals energy term is not optimized. The reason for this may be the presence of some electrostatic repulsion due to the number of positive charges (although the charges borne by the side chains of the peptide are partially shielded in solution).

Geometrical analysis

To compare the structures, an average over the 200 structures was computed within each class and each structure was subsequently fitted to the average. The results are reported in Table 2. When fitting over the backbone heavy atoms, rms differences larger than 2.0 Å were observed: the mean rms values are 2.07 and 2.67 Å for classes #1 and #2, respectively. Nevertheless, an obvious structural similarity was observed at the center of the molecules. In fact, the fit over the consensus sequence (residues 5 to 9) gives much lower values of the averaged atomic rms difference from the mean structure: 0.30 Å in class #1 and 0.35 Å in class #2. The structure of the peptide is therefore well defined in this section of the molecule. Likewise, when fitting and averaging over the head section (residues 1 to 5) or over the tail section (residues 9 to 14), the average atomic rms difference with the mean structure was smaller than 0.7 Å in both classes. Thus, the structure of the peptide is completely defined as three blocks, with flexibility around residues 4 and 9, perhaps due to a lack of experimental data (there is no restraining range for $\psi 4$ and $\psi 9$, and the range for $\varphi 9$ is 80°) (Fig. 1); this flexibility explains the bad fit over the entire molecule. Figures 5A, B and C represent a view of the three partial fits for 30 structures of class #1 (only the backbones are represented).

When considering heavy atoms, the average of the atomic rms difference from the mean structure, as described above, is around 2.5 Å, with higher values for the long side chains, since the NOE distances only constrain the H^{β} atoms. Figure 5D represents 30 structures of class #1, comprising side chains, fitted over the consensus sequence.

Figure 6 gives the distribution of the φ and ψ angles in the Ramachandran map. It can be seen that the φ and ψ angles cluster in a part of the allowed zones of the map. This is a consequence of the limitation of the angle values to intervals defined by experimental data (coupling constants and NOE distances) [11]. In particular, the vertical and horizontal lines correspond to either one φ or one ψ angle reaching one limit of the interval, while the related

TABLE 1
MEAN VALUE AND STANDARD DEVIATION OF EACH ENERGY TERM COMPUTED OVER 200 STRUCTURES FOR CLASSES #1 AND #2

		F_{total}	F_{bond}	F_{angle}	F_{imp}	F_{vdW}	F_{NOE}	F_{tor}	F_{elec}
Simulated annealing									
Class #1	Mean	49.0	5.7	12.8	1.3	−27.1	52.4	4.0	
	SD	3.2	0.2	1.2	0.1	3.1	0.9	0.4	
Class #2	Mean	37.6	5.1	9.2	1.1	−29.4	49.6	1.9	
	SD	3.4	0.3	0.7	0.1	2.9	1.7	0.4	
Complete force field									
Class #1	Mean	−136.3	6.6	15.2	0.6	−26.7			−131.1
	SD	5.5	0.1	0.3	0.1	2.7			4.0
Class #2	Mean	−137.5	6.4	15.0	0.5	−28.9			−129.5
	SD	5.7	0.1	0.2	0.1	2.6			4.4

Energy values are given in kcal mol^{-1}.

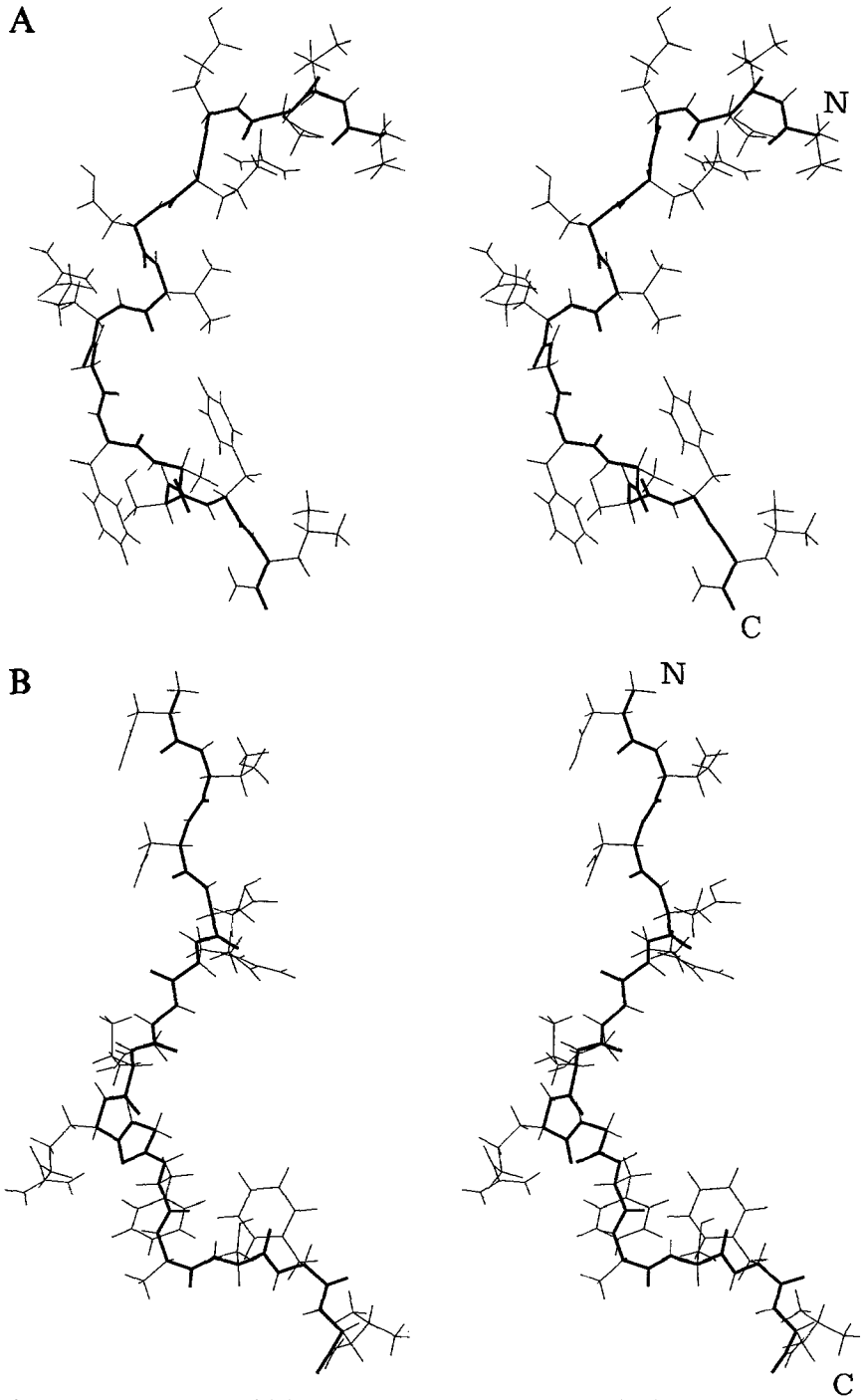

Fig. 4. Stereoviews of the lowest energy structures of (A) class #1 and (B) class #2. The main chain is represented in bold. 'N' and 'C' indicate the positions of the N- and C-termini, respectively.

φ or ψ angles range over the whole allowed interval. This occurs for Val[7], Arg[8], Phe[10] and Ala[11] in both classes.

Table 3 shows the mean value and the standard deviation of the φ, ψ and χ^1 angles for both classes. In the consensus sequence, as suggested by the analysis of the atomic rms difference, the φ and ψ angles lie in small intervals around the mean values (standard deviation below 26°), except for residues 4, 9 and 11. The ψ angle value of Glu[4] varies quite uniformly in the range −50° to 180°. Thus, a large part of the Ramachandran map is accessible to the φ and ψ angles of this residue, which could be due to the lack of experimental data (Fig. 1). In the case of Gly[9], the φ and ψ angles have a positive value in one half of the structures and are negative in the other half: ±50° for φ and ±175° for ψ. Thus, this residue can reach backbone conformations localized in four regions of the Ramachandran map.

Consequently, flexibility is present at Glu[4] and Gly[9],

TABLE 2
MEAN RMS DIFFERENCES (Å) BETWEEN THE MINIMIZED STRUCTURES AND THEIR AVERAGE

Fit over:	1–14		1–5		5–9		9–14	
	Main chain	Non-H atoms	Main chain	Non-H atoms	Main chain	Non-H atoms	Main chain	Non-H atoms
Class #1	2.07	3.28	0.66	8.56	0.30	5.10	0.68	6.95
Class #2	2.67	4.01	0.67	7.62	0.35	5.23	0.69	7.99

The rms differences are calculated both over the heavy backbone atoms and all heavy atoms belonging to the residues considered for the fit. The values are listed for classes #1 and #2 and for each partial fit considered in the text.

explaining the poor superposition when fitting over the entire molecule and justifying our division of the molecules in three sections.

When comparing the two classes, the φ and ψ mean angle values of all residues, except 6, 7 and 8, differ in the worst case by only ±10° from one class to the other. As expected, the difference is located on the residues involved in the NOE restraints specific to each class, and has no real effect on the other.

For residues 5, 6, 7 and 8 the following points should be mentioned:

Arg^5: the $H^{\alpha}5$-HN7 restraint particular to class #2 does not result in modification for this residue. The standard deviations from the mean values of φ and ψ are smaller than 0.5°; this low value is a consequence of both the narrowness of the allowed interval and the fact that the equilibrium values for both φ and ψ correspond to the boundary interval.

The standard deviation for the χ^1 angle is only 3°. The side chain therefore seems to be trapped in a deep local minimum. This could be an indication of hydrogen bonding between the side chains of Asp^3 and Arg^5. The orientation of the side chain is the opposite of that of Glu^4 and Asp^6 in class #1.

Asp^6: the φ angle has almost the same value in classes #1 and #2, while the ψ and χ^1 angles exhibit differences of around 30° (the standard deviation being more important in class #2).

Val^7: the φ angle has the same mean value in both classes, but the ψ angle has opposite values, the standard devi-

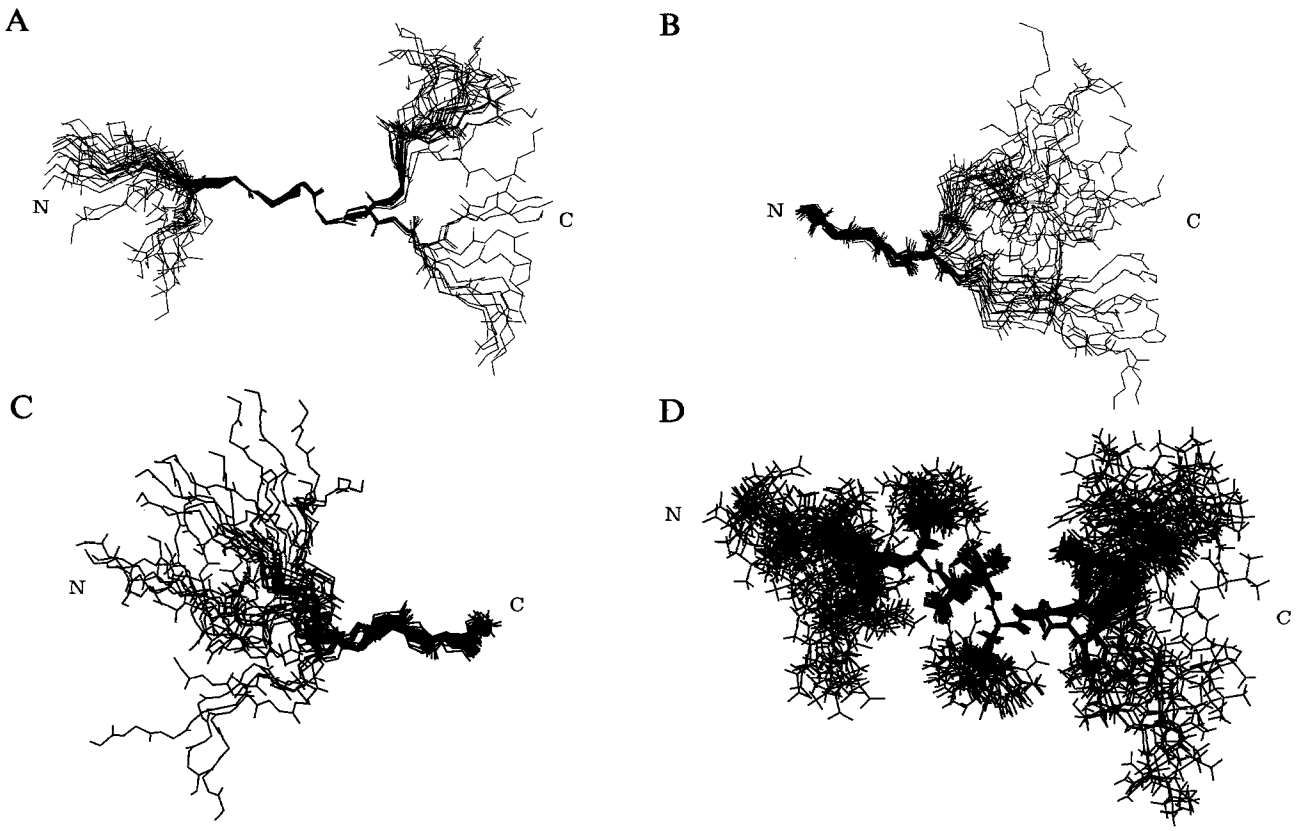

Fig. 5. Superposition of 30 structures of class #1 when fitting over (A) the consensus sequence (residues 5 to 9); (B) the head section (residues 1 to 5); and (C) the tail section (residues 9 to 14). Only main-chain atoms are represented. (D) shows the fit over the center section for the entire molecule, where all atoms are represented. 'N' and 'C' indicate the positions of the N- and C-termini, respectively.

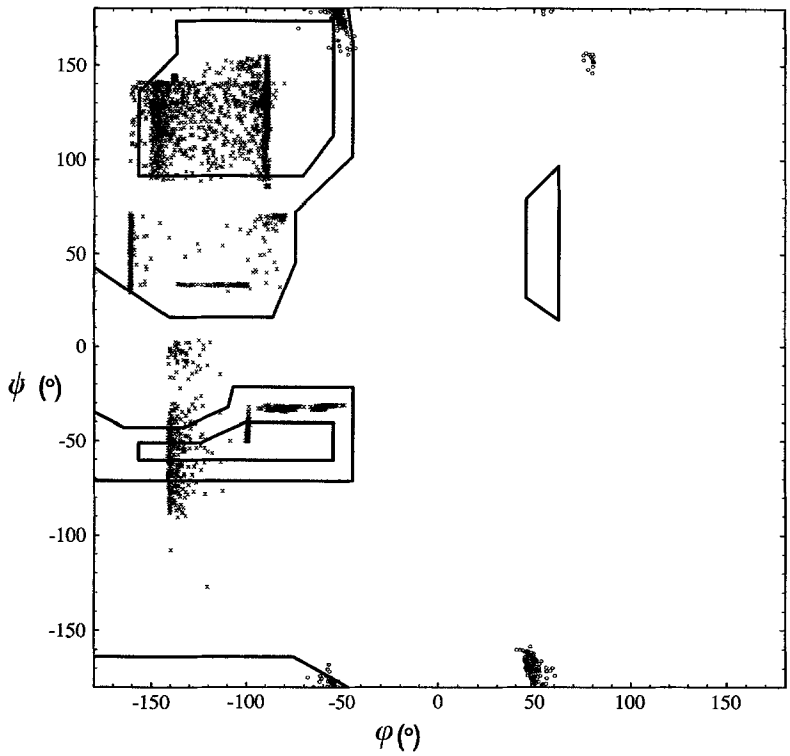

Fig. 6. Ramachandran plot for the 200 structures of classes #1 and #2.

ation being less than 1° in class #1. The side chain is also well located, despite the large value of the standard deviation. In fact, in class #1 χ^1 has a mean value of −166° (standard deviation 6°) for 60% of the structures and of 75° (standard deviation around 3°) for the others. In class #2, χ^1 ranges around ±170° (standard deviation less than 10°).

Arg[8]: the φ angle differs about 20° from one class to the other. As in the case of Val[7], the ψ angle is particularly constrained, with a standard deviation less than 1° from the mean value of −32°. The first part of the side chain is well located, as in the case of Arg[5]. The volume of the side chain of the valine contributes to limit the accessible conformational space.

Figures 7A and B show stereoviews of the superposition of the lowest energy structures of the two classes, fitted over the entire molecule or over the consensus sequence. Table 4 gives the rms deviations between these structures for each residue. As expected, the differences between the two classes are mainly located at the consensus sequence: the mutually excluding constraints are located on this section of the sequence. A very good fit over the head and tail sections can also be observed (Table 4).

TABLE 3
MEAN VALUES AND STANDARD DEVIATIONS FOR THE φ, ψ AND χ^1 ANGLES (°)

Residue number		2	3	4	5	6	7	8	9	10	11	12	13
Class #1													
φ	Mean	−127.8	−122.4	−126.6	−137.9	−88.8	−108.5	−80.0	−4.6	−91.4	−141.0	−125.0	−136.8
	SD	25.4	24.7	18.6	0.2	0.3	10.9	4.9	51.3	4.8	32.8	21.7	5.0
ψ	Mean	122.9	116.1	58.8	142.7	85.9	33.6	−32.7	−177.3	134.8	59.1	125.4	−48.7
	SD	7.3	15.6	73.1	0.3	0.2	0.3	0.9	8.2	10.9	11.5	13.6	24.5
χ	Mean	62.6	−96.9	−134.1	−105.2	−89.5	−91.7	−133.6	–	101.8	173.8	5.4	−74.4
	SD	44.6	40.3	42.7	3.0	3.4	59.4	2.9	–	46.0	100.9	43.4	19.4
Class #2													
φ	Mean	−131.5	−123.8	−125.0	−138.0	−92.4	−102.6	−59.7	3.2	−90.9	−139.2	−130.6	−135.8
	SD	22.8	24.6	20.6	0.3	10.5	11.2	3.1	53.7	4.8	33.7	20.0	6.0
ψ	Mean	123.6	114.0	50.9	143.6	126.1	−45.9	−32.0	180.0	137.4	56.5	126.2	−54.6
	SD	6.7	16.4	66.6	0.4	13.0	4.2	0.8	12.8	10.1	13.2	14.9	22.7
χ	Mean	59.4	−133.6	−97.4	−106.5	−131.3	56.0	−131.1	–	106.3	178.2	9.3	−74.4
	SD	44.5	39.4	49.5	3.1	26.8	8.7	2.9	–	49.1	99.8	48.8	19.9

Comparison of experimental and computational data

Experimentally, the low values of the thermal coefficient of NH chemical shifts for Asp[6], Val[7] and Arg[8] suggested that these protons are either involved in a hydrogen bond or poorly accessible. In order to check whether our models are in agreement with these experimental data, we have determined both the number of NH protons involved in a hydrogen bond and the maximum and mean accessible surfaces to a water molecule for the NH protons. These values are given as a ratio with the theoretically maximum accessible surface, which is calculated by considering an isolated residue.

In class #1, the maximum accessible fractions of the total surfaces are 7 and 13% for the NH protons of Asp[6] and Val[7], respectively (the mean value is less than 3% in both cases), while 16% of the NH7 are involved in hydrogen bonds. For Arg[8], the maximum accessible fraction of the total surface is less than 2%.

In class #2, the maximum accessible fractions of the surfaces are 19 and 11% for the NH protons of Asp[6] and Val[7], respectively (the mean value is 6% in both cases). For Arg[8], the maximum accessible fraction of the surface is almost zero and 16% of the NH are implied in hydrogen bonds.

TABLE 4
INDIVIDUAL RMS DIFFERENCES (Å) BETWEEN THE TWO LOWEST ENERGY STRUCTURES OF EACH CLASS OVER THE HEAVY BACKBONE ATOMS FOR EACH RESIDUE

Residue number	Fit over residues		
	1–5	5–9	9–14
1	0.5		
2	0.8		
3	0.5		
4	0.4		
5	0.5	1.3	
6		0.9	
7		1.1	
8		1.2	
9		1.8	0.4
10			0.1
11			0.3
12			0.4
13			0.3
14			0.4

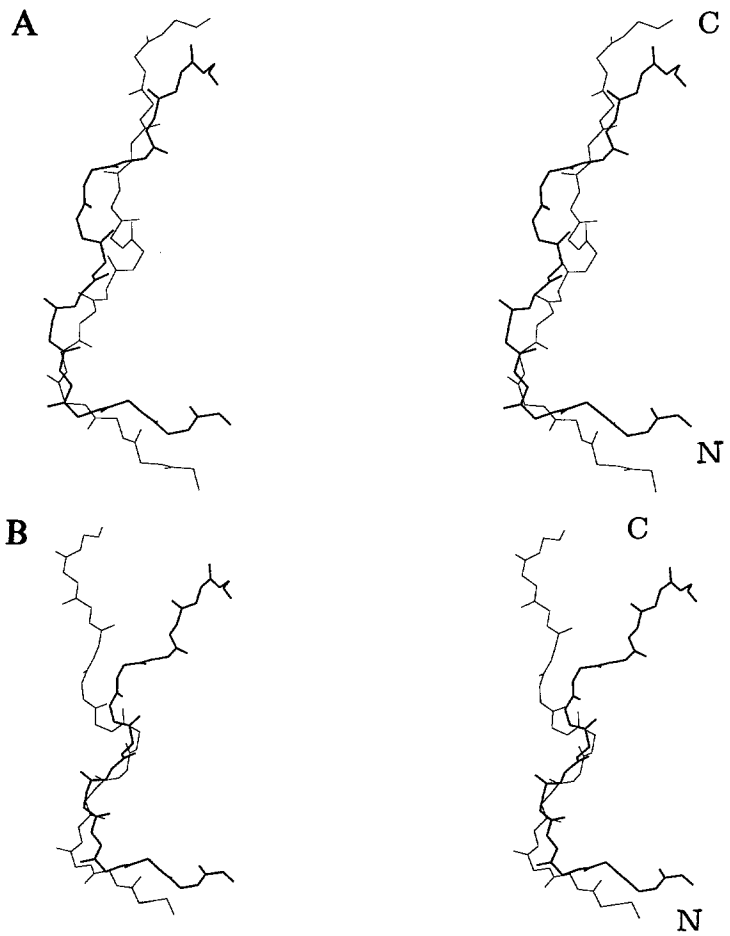

Fig. 7. Stereoviews of the superposition of the main chains of the optimized structures of classes #1 and #2 when fitting over (A) the entire sequence and (B) the consensus sequence. 'N' and 'C' indicate the positions of the N- and C-termini, respectively.

Thus, our models for the consensus sequence are in agreement with the experimental data concerning the low values of the thermal coefficient of the NH chemical shift. These low values result from burial of the NH protons rather than their involvement in hydrogen bonds. Moreover, concerning the side chain rotamers, most of the NMR data are in agreement with the χ^1 data obtained by molecular modelling. These include the scalar coupling constants $^3J_{H\alpha H\beta}$ which were not used as restraints in the simulations. Only for Val[7] a problem occurs: the available NOE restraints do not permit discrimination between t and g^- conformers, because (i) the same NH···H^β values were expected and (ii) the accuracy on H^α···H^β was not sufficient to induce a t conformation. Otherwise, the side chain NMR data for Val[2], Glu[4], Arg[5], Asp[6] and Arg[8] are in quantitative agreement with our mean χ^1 values.

Conclusions

Both NMR and molecular modelling results show a structural organization in the consensus sequence of the peptide under investigation. While NMR suggested equilibria between different structures (type II β-turns, α-helix), this hypothesis suffers from a lack of data, because of NOE overlaps or possible internal mobility. Type I and/or II β-turns and α-helix structures have not been revealed with the simulated annealing protocols and are furthermore incompatible with the available NOE restraints.

The two types of conformation obtained from the two sets of interproton distance restraints reveal a weak structural organization at the level of the consensus sequence. Both types of structure are also in agreement with the experimentally found low values of the thermal coefficient of NH chemical shifts. The two types of structure are quite similar and the existence of the two classes could be explained by a flexibility of the sequence, recognized by the enzyme RXVRG endoprotease.

These results should be compared with those to be obtained from the structure determination of two analogs: one comprising an alanine in position 12 (instead of a serine) and the other comprising only the last 11 amino acids of the sequence. These analogs are both recognized by the RXVRG endoprotease. A structure determination by ^{1}H NMR has been performed and we are proceeding to a conformational analysis by molecular modelling.

References

1 Barr, P.J., Cell, 66 (1991) 1.
2 Nakayama, K., Watanabe, T., Nakagawa, T., Kim, W.S., Nagahama, M., Hosata, M., Hatsuzawa, K., Kondoh-Hashiba, K. and Murakami, K., J. Biol. Chem., 267 (1992) 16335.
3 Rholam, M., Nicolas, P. and Cohen, P., FEBS Lett., 207 (1986) 1.
4 Duffaud, G. and Inouye, M., J. Biol. Chem., 264 (1988) 10224.
5 Paolillo, L., Simonetti, M., Brakch, N., D'Auria, G., Saviano, M., Dettin, M., Rholam, M., Scatturin, A., Di Bello, C. and Cohen, P., EMBO J., 11 (1988) 2399.
6 Darby, N.J. and Smyth, D.G., Biosci. Rep., 10 (1990) 1.
7 Inouye, S., Duffaud, G. and Inouye, M., J. Biol. Chem., 261 (1986) 10970.
8 Gibson, B.W., Poulter, L., Williams, D.H. and Maggio, J.E., J. Biol. Chem., 261 (1986) 5341.
9 Kuks, P.F.M., Creminion, C., Leseney, A.M., Bourdais, J., Morel, A. and Cohen, P., J. Biol. Chem., 264 (1989) 14609.
10 Bevins, C.L. and Zasloff, M., Annu. Rev. Biochem., 59 (1990) 395.
11 Chalaoux, F.-R. et al., Int. J. Pept. Protein Res., (1995) submitted for publication.
12 Sautilis, J., Mierke, D.F., Byk, G., Gilon, C. and Kessler, H., J. Am. Chem. Soc., 114 (1992) 4818.
13 Brünger, A.T., Kuriyan, J. and Karplus, M., Science, 235 (1987) 458.
14 Powell, M.J.D., Math. Program., 12 (1977) 241.
15 Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P., Science, 220 (1983) 671.
16 Nilges, M., Gronenborn, A.M., Brünger, A.T. and Clore, G.M., Protein Eng., 2 (1988) 27.
17 Verlet, L., Phys. Rev., 159 (1967) 98.
18 Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F., DiNola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684.