Journal of Biomolecular NMR, 2 (1992) 235–256
ESCOM

J-Bio NMR 055

Two-dimensional ^{1}H NMR study of recombinant insect defensin A in water: Resonance assignments, secondary structure and global folding

Jean-Marc Bonmatin[a,*], Jean-Luc Bonnat[a], Xavier Gallet[a], Françoise Vovelle[a,b], Marius Ptak[a,b], Jean-Marc Reichhart[c], Jules A. Hoffmann[c], Elisabeth Keppi[d], Michèle Legrain[d] and Tilman Achstetter[d]

[a]*Centre de Biophysique Moléculaire (CNRS) and*
[b]*Université d'Orléans, 1A, Avenue de la Recherche Scientifique, F-45071 Orléans Cedex 02, France*
[c]*Laboratoire de Biologie Générale and URA 672 (CNRS), Endocrinologie et Immunologie des Insectes, Université Louis Pasteur, 12, Rue de l'Université, F-67000 Strasbourg, France*
[d]*Transgène, 11, Rue de Molsheim, F-67082 Strasbourg Cedex, France*

Received 9 October 1991
Accepted 4 December 1991

Keywords: Insect immunity; Defensins; Antibacterial proteins; 2D ^{1}H NMR; Proton assignments; Biomolecular structure; Molecular modeling

SUMMARY

A 500 MHz 2D ^{1}H NMR study of recombinant insect defensin A is reported. This defense protein of 40 residues contains 3 disulfide bridges, is positively charged and exhibits antibacterial properties. 2D NMR maps of recombinant defensin A were fully assigned and secondary structure elements were localized. The set of NOE connectivities, $^3J_{NH-\alpha H}$ coupling constants as well as $^1H/^2H$ exchange rates and $\Delta\delta/\Delta T$ temperature coefficients of NH protons strongly support the existence of an α-helix (residues 14–24) and of an antiparallel β-sheet (residues 27–40). Models of the backbone folding were generated by using the DISMAN program and energy refined by using the AMBER program. This was done on the basis of: (i) 133 selected NOEs, (ii) 21 dihedral restraints from $^3J_{NH-\alpha H}$ coupling constants, (iii) 12 hydrogen bonds mostly deduced from $^1H/^2H$ exchange rates or temperature coefficients, in addition to 9 initial disulfide bridge covalent constraints. The two secondary structure elements and the two bends connecting them involve approximately 70% of the total number of residues, which impose some stability in the C-terminal part of the molecule. The remaining N-terminal fragment forms a less well defined loop. This spatial organization, in which a β-sheet is linked to an α-helix by two disulfide bridges and to a large loop by a third disulfide bridge, is rather similar to that found in scorpion charybdotoxin and seems to be partly present in several invertebrate toxins.

* To whom correspondence should be addressed.

Abbreviations: SCUBA, Stimulated Cross peaks Under Bleached Alphas; MCD analysis, Main Chain Directed analysis; CSH motif, Cysteine Stabilized α-Helix motif.

INTRODUCTION

The term defensins was coined by Lehrer and associates for small cationic proteins which participate in the nonoxidative microbicidal mechanisms in phagocytic blood cells in mammals (for a review, see Lehrer et al., 1991). Defensins are particularly abundant in the azurophil granules of polymorphonuclear neutrophil leucocytes. To date, the primary amino acid sequences of 15 defensins have been established. They are variably cationic, relatively arginine-rich nonglycosylated proteins comprised of 29–34 amino acid residues and all contain a characteristic cysteine motif with 3 intramolecular disulfide bonds. It was suggested that the antimicrobial properties of mammalian defensins are related to their capacity to induce voltage-dependent ion-permeable channels in membranes (Kagan et al., 1990).

One of our groups has recently isolated from the immune blood of larvae of the fleshfly *Phormia terranovae* (Insecta, Diptera) potent antibacterial proteins which exhibit structural similarities to mammalian defensins and were tentatively named insect defensins (Lambert et al., 1989). These proteins are positively charged and comprise 40 residues with 6 cysteines engaged in 3 intramolecular disulfide bridges. Two defensin isoforms (A and B; MW $\approx$ 4.06 kDa) were isolated from the blood of *Phormia*: they differ by one single residue (Gly^{32} in A, Arg^{32} in B) and show some sequence similarity with rabbit defensin NP-1 (Selsted et al., 1983). Independent studies with an embryonic cell line N1H-Sape-4 derived from another fleshfly, *Sarcophaga peregrina*, resulted in the isolation from the culture medium of an insect defensin isoform which differs from *Phormia* defensins by one single substitution at position 34. This molecule, termed sapecin (Matsuyama and Natori, 1988a,b) is constitutively expressed by the cell line and its synthesis is induced in fleshfly larvae by injection of bacteria. It exhibits a strong antibacterial activity and shows affinity for cardiolipin (Kuzuhara et al., 1990; Matsuyama and Natori, 1990).

Recent results indicate that insect defensins are probably largely distributed among the huge class of the Insecta. We have now evidence for their presence in immune blood of Coleoptera (Bulet et al., 1991), Hemiptera and Odonata (unpublished results from the Strasbourg laboratory). As a rule, these molecules are synthesized by the fat body cells within a few hours following the injection of bacteria or a septic injury, together with other inducible antibacterial proteins e.g. cecropins (Hultmark et al., 1982), attacins (Engström et al., 1984), diptericins (Dimarcq et al., 1988), apideacin (Casteels et al., 1989) and coleoptericin (Bulet et al., 1991).

Mammalian defensins have attracted interest of NMR spectroscopists over the past 4 years and the 3-dimensional (3D) structure in solution of rabbit defensin NP-5 was determined by 2-dimensional (2D) ^{1}H NMR spectroscopy and molecular modeling techniques (Bach et al., 1987; Bassolino et al., 1988; Pardi et al., 1988). More recently, the crystal structure of human defensin HNP-3 was established by Hill et al. (1991) and revealed a particular dimeric β-sheet architecture differing noticeably from other lytic peptides such as melittin (Terwilliger and Eisenberg, 1982; Bazzo et al., 1988) and δ-haemolysin (Tappin et al., 1988). In the field of insect immune proteins, the first studies were centered around cecropin A, a major inducible antibacterial peptide initially isolated from the *cecropia* moth. The solution conformation of this molecule was investigated by ^{1}H NMR spectroscopy in a mixed solvent where it adopts a fully ordered structure (Holak et al., 1988). As regards insect defensin-like molecules, only a preliminary ^{1}H NMR study of the conformation of sapecin in methanol has been published (Hanzawa et al., 1990).

Two defensin isoforms, A and B, have been isolated from the immune hemolymph of fleshfly

larvae which differ by a single amino acid substitution (see sequence in Fig. 4). The disulfide array of natural as well as recombinant defensin A was established by Lepage et al. (1991), but their role in the overall conformation of the molecule requires tertiary structural analysis. Thus, in order to progress in the understanding of its biological role, the 3D structure of recombinant defensin A from *Phormia terranovae* has been investigated using 2D ^{1}H NMR techniques. During their excretion and their biological action, these defensins are probably in contact with both aqueous and nonaqueous environments. In the present paper our 2D NMR study of recombinant defensin A in water in which the protein is readily soluble, includes the complete assignment of proton resonances, the determination of the secondary structure elements and a model of the global folding.

MATERIALS AND METHODS

Recombinant insect defensin A was produced in *Saccharomyces cerevisae* as a fusion protein carrying at its N-terminus leader sequences derived from the precursor of the yeast pheromone mating factor α. The leader (or prepro) sequences allowed the biologically active protein to be secreted at high levels in a correctly processed form (Reichhart and Achstetter, 1990). A two-step procedure was developed to purify recombinant defensin A secreted into the yeast culture medium. An initial diafiltration step of the clarified supernatant through a spiral cross-flow filtration system S10Y3 (Amicon, Danvers, MA, U.S.A.) removed the majority of colored material from the fermentation medium, concentrated the protein and allowed the medium exchange for the starting buffer of the following chromatography step. This solution was loaded onto an S-Sepharose Fast Flow (Pharmacia, Uppsala, Sweden) column (13 × 3.2 cm ID, Amicon) equilibrated with 50 mM sodium acetate (pH 5.5) at a flow rate of 5.5 ml/min. Recombinant defensin A was eluted using a gradient from 0–1 M NaCl in the same buffer. The pure product was pooled according to the results of analytical reversed-phase HPLC as previously described (Lambert et al., 1989). Desalting of the material was achieved in an ultrafiltration cell on a YM3 membrane (Amicon, Danvers, MA, U.S.A.) and the product was lyophilized from water. This product was also active against *Micrococcus luteus* A270 as previously described (Lambert et al., 1989).

The recombinant defensin A was dissolved in a mixture of 10% ^2H_2O/90% ^1H_2O except for ^1H/^2H exchange experiments where a maximum deuterium content was desired. Typically, samples contained 10–18 mg of protein which leads to concentrations of about 4.9 and 8.9 mM, respectively. The noncorrected pH-meter reading was adjusted to pH 4.9 ± 0.1 for ^1H_2O and ^2H_2O solutions. Oxygen was removed from the NMR tube by carefully flushing the solution with argon for about 15 min before sealing.

1D and 2D ^{1}H NMR spectra were acquired at 283, 294, 300, 306 and 313 K on a Bruker AM-X 500 spectrometer operating at 500 MHz. Temperature was regulated within ± 0.1 K by using a Bruker special device. 2D spectra were collected with the carrier frequency in the middle of the spectrum, coinciding with the water resonance. This resonance was used to calibrate chemical shifts, taking into account the well-documented shift of the water resonance with temperature. Data processing was performed on a Bruker X32 station using the Bruker UXNMR program.

Double Quantum Filtered (DQF)-COSY spectra (Marion and Wüthrich, 1983) using SCUBA (Stimulated Cross peaks under Bleached Alphas; Brown et al., 1988) were recorded with a 90° pulse of 6.5 μs and a SCUBA delay of 50 ms. Homonuclear Hartmann–Hahn transfer (HOHAHA) experiments were carried out using an MLEV-17 composite pulse cycling sandwiched be-

tween two trim pulses of 2 ms (Bax and Davis, 1985). These data were collected with 90° pulses of 7 μs (high power) and 28.2 μs (attenuation of 14 dB) at mixing times of 50 and 81 ms. Nuclear Overhauser enhancement spectroscopy (NOESY; Macura and Ernst, 1980) was performed with presaturation during both relaxation delay and mixing time. Spectra were obtained with a 90° pulse of 6.8 μs and mixing times of 150 and 300 ms. Most of the 2D spectra were recorded using time-proportional phase increments (TPPI; Marion and Wüthrich, 1983) and acquisition parameters kept constant were: spectral width: 6098 Hz, recycle time: 1.1 s, dwell time: 82 ms, presaturation delay: 1.3 s (attenuation of 57 dB in ^1H_2O, 77 dB in ^2H_2O), 4 dummy scans, 8–104 transients. For DQF-COSY, 4 K data points were used in the F2 direction and 800 increments were acquired in F1. This led to a matrix size of 8 K × 1 K after zero-filling (digital resolution of 0.7 Hz/point in F2 and 5.9 Hz/point in F1). For HOHAHA and NOESY experiments, 512 increments were used in F1 to give a matrix size of 2 K × 1 K after zero-filling (digital resolution of 2.9 Hz/point in F2). Prior to Fourier transform, data were digitally filtered by using a Gaussian window in F2 and a shifted sine-bell window in F1. Finally a baseline correction with a third-order polynomial was successively applied in both frequency directions (the water resonance area being excluded in F2).

$^1H/^2H$ exchange was observed at 283 K (HOHAHA, mixing time of 81 ms) and 294 K (DQF-COSY; NOESY, mixing time of 300 ms). In ^2H_2O, the residual ^1H_2O resonance was correctly saturated with an attenuation 100-fold higher than in ^1H_2O solutions. After dissolution in ^2H_2O, 2D spectra were collected with various delays ranging from 20 min to 2 days. For the first exchange acquisition, only 256 increments of 8 transients each were acquired in F1 in order to detect protons in fast exchange (experiment duration shorter than 1 h). In addition, a DQF-COSY experiment was performed after amide proton exchanges, at 313 K, to determine $^3J_{\alpha H\text{-}\beta H}$ coupling constants for further stereospecific assignments (Güntert et al., 1989).

The DISMAN distance geometry (DG) algorithm (Braun and Gō, 1985) was used to generate a large number of structures consistent with the experimental data. Observed NOEs at 150 ms were classified into 3 distance ranges: 2–3 Å, 3–4 Å and 4–5 Å, corresponding to strong, medium and weak NOE intensities, respectively. A larger range of 2–5 Å was chosen for NOE intensities estimated from the volume of cross peaks detected only with a larger mixing time (300 ms). Three additional restraints were added for each disulfide bridge, viz: $d_{Si\text{-}Sj} = 2.02 \pm 0.05$ Å, $d_{Si\text{-}Cj}$ and $d_{Sj\text{-}Ci} = 2.99 \pm 0.5$ Å which ensure the right geometry of the disulfide bond. Another source of input constraints came from the secondary structures. For β-sheet and α-helical conformations, 12 hydrogen bond constraints were used: $d_{Oi\text{-}Nj} = 2.9 \pm 0.3$ Å. Fifty-four initial structures were randomly generated in the (Φ, Ψ) space. In a first step, 9 S-S bridge constraints were introduced in addition to experimental secondary structure restraints: (i) 13 NH-NH or NH-αH long-range distances from NOEs for the β-sheet, (ii) 11 Φ angle constraints derived from $^3J_{NH\text{-}\alpha H}$ coupling constants for the α-helix. In a second step, 6 hydrogen bond constraints and 28 intrahelix distance restraints from NOEs were added for the helical fragment, in addition to 6 other hydrogen bond constraints and 10 Φ angle restraints for the β-sheet. In the last step, 92 intraresidual or sequential distance constraints were finally added. At the end of the process 162 NMR constraints, compared to a total of 175 constraints, were used, most of them involving backbone protons only. Four selected (vide infra) distance geometry structures were energy-refined by using the AMBER molecular mechanics program (Weiner and Kollmann, 1981). NOE constraints were introduced in the form of a pseudoenergy function. For this purpose, instead of a classical harmonic potential function, we used the function $V = \Sigma K(d\text{–}d_{NOE})^4$ which seems more appropriate to reflect the

precision of NOE data. The weighting factor, K, was set at a rather low value of 15 kcal mol^{-1} $Å^{-4}$ in order to avoid distortions of the standard geometry of the model.

RESULTS

Judging from circular dichroism spectra (not shown), no major structural change of recombinant defensin A in water (20 μM) occurred between pH 4.9 and pH 7.4, at room temperature. No modification was detected in NMR spectra after several series of experiments recorded at pH 4.9 and at temperatures ranging from 283 to 313 K. A careful analysis of the complete network of NOE connectivities does not provide evidence for the existence of intermolecular interactions. Furthermore, the absence of any concentration effect on the occurrence of NOEs in the range of 1–9 mM suggests that there is no aggregation of recombinant defensin A in water.

Assignment of the ^1H NMR spectra of recombinant defensin A

Two strategies are generally used to assign 2D ^1H NMR spectra of peptides and proteins. The first one is based on the identification of scalar coupling networks, or spin systems, corresponding to different types of amino acids (Wagner and Wüthrich, 1982). For recombinant defensin A, the scalar connectivities between NH, αH and side-chain protons were mainly established on HOHAHA maps recorded in ^1H_2O at acidic pH. In our experimental conditions (mixing time of 81 ms), the transfer of polarization was complete within all amino acid residues, even those carrying a long side chain, such as Leu, Lys or Arg residues. Some side-chain assignments were confirmed by considering COSY and HOHAHA maps in ^2H_2O. The second procedure, termed MCD analysis (Main Chain Directed analysis), involves the identification of repeating NH-αH-βH NOE patterns which characterize secondary structure elements such as α-helices and β-sheets (Englander and Wand, 1987). The recognition of a few amino acid residues is then generally sufficient to set the secondary structure elements within the polypeptide sequence. At several temperatures, from 283 to 313 K, these two strategies were combined with a continuous feedback procedure to assign the COSY, HOHAHA and NOESY maps. As explained below, MCD analysis shows that one α-helical fragment and one β-sheet fragment are present in the structure. For the rest of the molecule, the assignment was mainly based on the identification of spin systems and on sequential assignment from dipolar connectivities.

Spin system identification. In the fingerprint region of a DQF-COSY map 47 cross peaks which correspond to connectivities of the NH-αH type should be observed for recombinant defensin A. Note that Gly residues give rise to both NH-αH and NH-α′H cross peaks. Figure 1 shows that 28 spin systems are clearly observed at 294 K. Five Gly residues among 7 are easily identified by their NH-αH and NH-α′H couplings. By varying the temperature, 8 other cross peaks which nearly overlap at 294 K can be separated: (Ser^7, Thr^9), (His^{13}, Cys^{36}) and (Leu^6, Ala^{15}, Cys^{16}, Arg^{26}). In addition, another Gly residue (Gly^{28}) appears with one of its αH resonances close to the water line. Figure 1 also reveals 3 connectivities involving the side-chain $N^εH$ of the 3 Arg residues and the NH_3^+-ε,ε′H peak of the single Lys^{33} ($^3J \approx 25$ Hz). Among the 3 spin systems which still remain difficult to detect (Ala^1, Cys^{30}, Gly^{32}), the NH resonance of Cys^{30} was found at a rather low-field value (10.04 ppm) on the HOHAHA map. This assignment is supported by sequential NOEs involving Tyr^{29} and Asn^{31} as reported in Fig. 2. It should be noted that resonances correlated with the amide proton of Cys^{30} appear particularly weak and broad compared to other amide proton

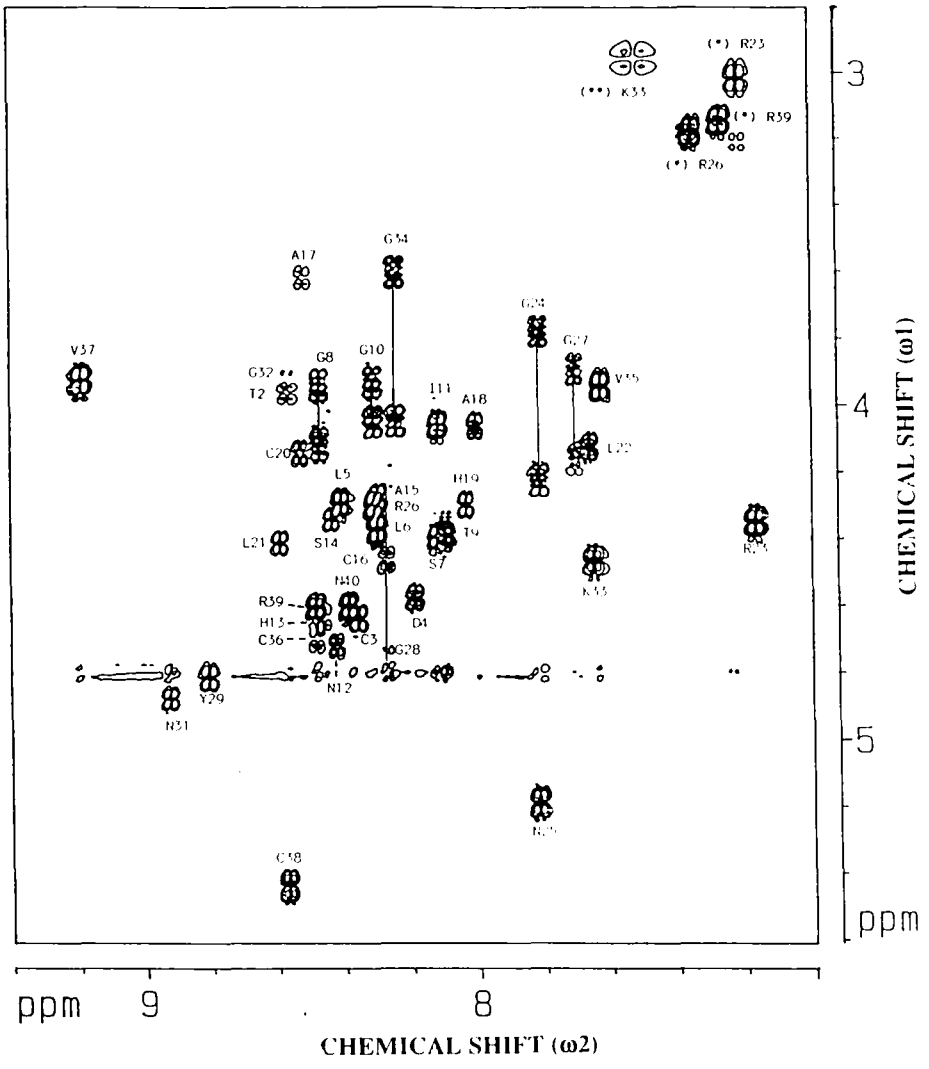

Fig. 1. Portion (fingerprint region) of a 500 MHz phase-sensitive (with SCUBA) DQF-COSY spectrum of recombinant defensin A in ^1H_2O (pH 4.9, 8.9 mM, 294 K, see Materials and Methods section for acquisition and processing parameters). Cross peaks appear in dispersion with positive and negative levels both displayed in black. All NH-αH couplings are sequence specifically labeled except for Ala^1 and Cys^{30} (see text). Solid lines indicate pairs of peaks for Gly residues. $N^{\varepsilon}H$-δ,δ'H cross peaks for Arg residues (*) and NH_3^+-ε,ε'H cross peaks for the single Lys residue (**) are also indicated.

resonances. In the sequence, Asp, Ile, Tyr and Lys residues are present only once. Their pattern can be unambiguously recognized as in the case of Asp^4 where the αH correlates with β and β' proton resonances, very close together (Table 1). Both aromatic low-field resonances of $C^{2,6}$ and $C^{3,5}$ protons of the single Tyr residue were used to identify Tyr^{29}. The unique NH_3^+-ε,ε'H cross peak described previously in the DQF-COSY map (Fig. 1) is characteristic for Lys^{33} for which all intra-residual connectivities are observed on the HOHAHA map. This assignment is supported by the

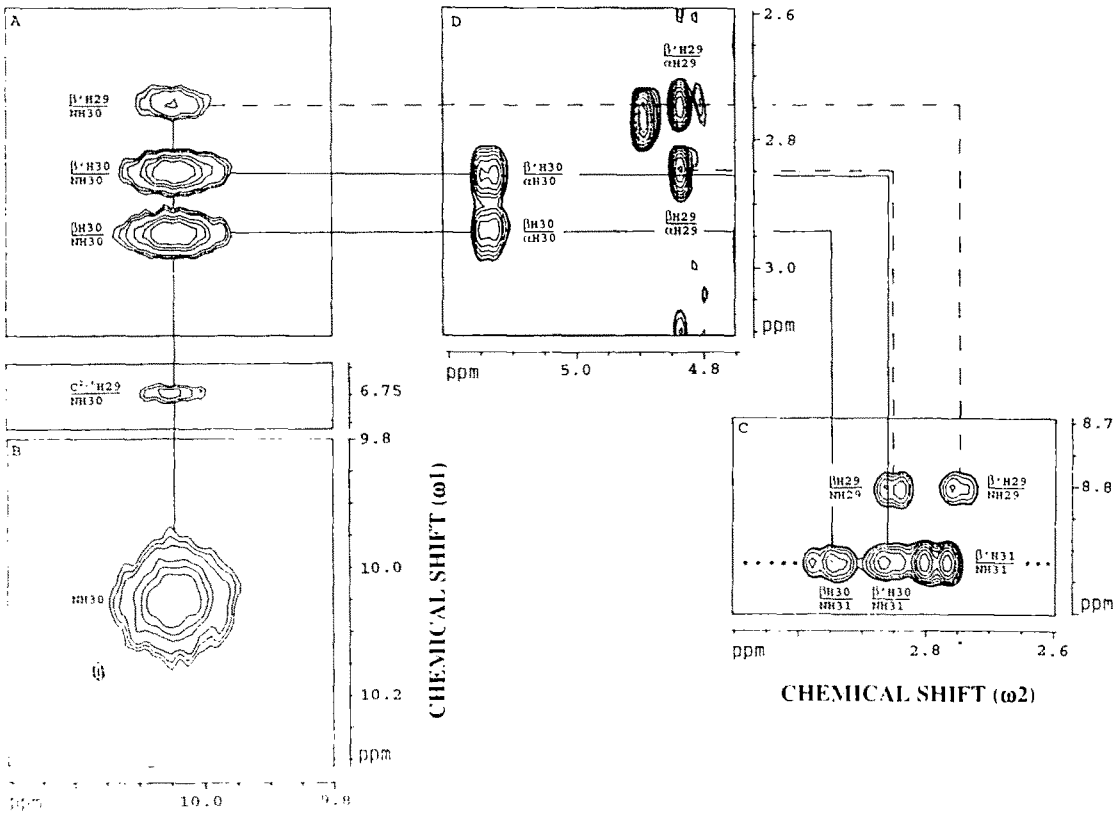

Fig. 2. Assignment of resonances observed for Cys[30]. Parts of 500 MHz NOESY spectra (294 K, 300 ms) of recombinant defensin A in ^1H_2O (A,B,C) and in ^2H_2O (D) are displayed with the labeled connectivities involving Tyr[29] (dashed lines), Cys[30] (solid lines) and Asn[31] (dotted line). Identical processing parameters were used for these NOESY experiments and 2D regions are displayed with the same chemical shift scale.

low-field displacement of ε,ε′ protons while cross-peak intensities confirm the chemical shift assignment of resonances along the side chain. The δH_3 resonance of Ile[11] appears at the lowest chemical shift value observed in the 2D map and correlates with intra-γ methyl, -γ,γ′ methylene, -β and -α protons. This rich pattern clearly separates this spin system from Val and Leu residues which have pairwise cross peaks in the same high-field spectral region. Considering remaining AMX spin systems (Ser, Cys, Asn), we based our assignment on the position of β and β′ proton resonances. For Ser[7] and Ser[14] residues, these resonances are shifted downfield as compared to other AMX spin systems. β and β′ protons of the 6 Cys residues are detected at higher field (2.5–3.1 ppm). αH resonances of Asn residues arise between 4.6 and 5.1 ppm. Moreover, detection of $N^{\delta},N^{\delta'}$ side-chain protons (6.9–7.7 ppm), coupled in HOHAHA maps, supports assignments of Asn[12], Asn[25], Asn[31] and Asn[40], since N^{δ}H-β,β′H intraresidual connectivities are also perceived. The two histidines, His[13] and His[19], were identified through their C^2H and C^4H resonances which are well separated in the amide proton region (Table 1). A cross peak between these two protons is present in HOHAHA spectra and scalar correlations involving C^4H with both α-, β- and β′-protons can be seen. Moreover, β- and β′-protons of His residues are the only ones to appear between

TABLE I

^1H CHEMICAL SHIFTS[a] OF RECOMBINANT INSECT DEFENSIN A IN ^1H_2O (500 MHz, pH 4.9, 294 K)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	Others
Ala^{1}	[b]	4.12	1.46			
Thr^{2}	8.48	3.96	3.96	1.04		
Cys^{3}	8.35	4.64	2.98,2.88			
Asp^{4}	8.18	4.57	2.65,2.62			
Leu^{5}	8.40	4.28	1.57,1.57	1.18	0.84,0.79	
Leu^{6}	8.29	4.37	1.67,1.67	1.56	0.86,0.80	
Ser^{7}	8.11	4.37	3.88,3.83			
Gly^{8}	8.46	4.11,3.92				
Thr^{9}	8.09	4.38	4.28	1.17		
Gly^{10}	8.30	4.04,3.92				
Ile^{11}	8.11	4.04	1.69	1.37,1.06 CH_3: 0.81	0.64	
Asn^{12}	8.42	4.72	2.99,2.64			NH_2: 7.69,7.05
His^{13}	8.44	4.64	3.38,3.09			C^2H: 8.37; C^4H: 7.21
Ser^{14}	8.43	4.33	4.03,3.98			
Ala^{15}	8.29	4.26	1.48			
Cys^{16}	8.26	4.46	3.07,2.61			
Ala^{17}	8.51	3.60	1.53			
Ala^{18}	8.00	4.05	1.47			
His^{19}	8.03	4.28	3.56,3.27			C^2H: 8.35; C^4H: 7.04
Cys^{20}	8.52	4.13	2.67,2.61			
Leu^{21}	8.58	4.40	1.79,1.79	1.59	0.87,0.84	
Leu^{22}	7.65	4.12	1.81,1.76	1.62	0.90,0.87	
Arg^{23}	7.16	4.34	2.10,1.68	1.53,1.53	3.01,3.01	$N^{\varepsilon}H$: 7.22
Gly^{24}	7.80	4.22,3.76				
Asn^{25}	7.81	5.18	3.15,2.25			NH_2: 7.13,7.05
Arg^{26}	8.30	4.29	1.88,1.76	1.71,1.65	3.17,3.17	$N^{\varepsilon}H$: 7.34
Gly^{27}	7.69	4.16,3.88				
Gly^{28}	8.26	4.76,4.21				
Tyr^{29}	8.81	4.82	2.84,2.74		$C^{2,6}H$: 6.74	$C^{3,5}H$: 6.59
Cys^{30}	10.04	5.10	2.94,2.86			
Asn^{31}	8.92	4.88	3.63,2.78			NH_2: 7.58,6.95
Gly^{32}	8.56	3.96,3.96				
Lys^{33}	7.63	4.46	2.06,1.66	1.42,1.42	1.35,1.35	$C^{\varepsilon}H_2$: 2.96; NH_3^+: 7.58
Gly^{34}	8.23	4.04,3.60				
Val^{35}	7.62	3.95	1.78	0.76,0.76		
Cys^{36}	8.47	4.70	2.86,2.51			
Val^{37}	9.18	3.91	1.32	0.80,0.70		
Cys^{38}	8.56	5.42	2.93,2.61			
Arg^{39}	8.48	4.60	1.80,1.60	1.54,1.54	3.16,3.14	$N^{\varepsilon}H$: 7.25
Asn^{40}	8.38	4.60	2.80,2.60			NH_2: 7.53,7.01

[a] Chemical shifts (in ppm) are relative to the signal of ^1H_2O at 4.81 ppm and are accurate to ± 0.01 ppm.

[b] Not observed.

3 and 3.6 ppm. Side-chain $N^{\varepsilon}H$ resonances (7.2–7.4 ppm) of the 3 Arg residues (Arg^{23}, Arg^{26}, Arg^{39}) are identified from the 3 $N^{\varepsilon}H$-δ,δ'H cross peaks in Fig. 1. In HOHAHA maps, the pattern of Arg residues displays relayed correlations all along the side chain with a decreasing intensity from NH to α-, β-, γ-, δ- and N^{ε}-protons in succession. Concurrently the cross-peak intensity increases between $N^{\varepsilon}H$ and these consecutive protons. Defensin A contains two Thr residues (Thr^2, Thr^9) which have a unique spin pattern and give rise to a strong βH-γH_3 cross peak. In HOHAHA spectra, relay connectivities from the amide proton to αH-, βH- and γH_3-protons are effectively observed for Thr^9 thus distinguishing this residue from Ala residues. Note that methyl resonances of Thr residues generally appear at higher field (1–1.2 ppm) than those of Ala residues ($\approx$1.5 ppm). In the case of Thr^2, αH and βH resonances overlap. However, the analysis of sequential NOEs (mixing time of 300 ms) provides confidence for this assignment. The NH/side chain domain of 2D maps reveals 3 Ala residues (Ala^{15}, Ala^{17}, Ala^{18}) among the 4 of the sequence. This spin system pattern, with a strong αH-βH connectivity (COSY) as well as a strong NH-βH_3 cross peak (HOHAHA), is easily identified except for the Ala^1 residue for which rapid reorientation or fast exchange with solvent prevent the NH resonance detection. Nevertheless, this Ala^1 residue is recognized by means of its strong αH-βH_3 scalar coupling and by inspection of dipolar connectivities involving its αH- and βH_3-protons with protons belonging to Thr^2 and Cys^3 residues. Two Val (Val^{35}, Val^{37}) and 4 Leu residues (Leu^5, Leu^6, Leu^{21}, Leu^{22}) complete this search of spin systems in the spectra of recombinant defensin A. In their side chain, Leu residues have a supplementary CH_2 group which differentiates these two types of amino acids in HOHAHA maps. Furthermore, β- and β'-protons of Leu^6 and Leu^{22} do not degenerate (Table 1). Note that Leu as well as Val residues have slightly nonequivalent methyl group resonances except in the case of Val^{35}. For this residue the substantial chemical shift difference between β-proton resonances of Val and Thr residues is used to discriminate them (Groß and Kalbitzer, 1988). Two major difficulties involving Ala^1 (absence of the NH resonance) and Gly^{32} (overlapping of αH and α'H resonances) have been encountered for complete spin system identification, even varying the temperature from 283 to 313 K. At this stage, the sequential procedure resulting from NOESY spectra analysis was a powerful tool to check our assumptions and to identify missing assignments.

Sequential assignment. NOEs involved in this procedure are mainly d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ connectivities between consecutive residues (Wüthrich, 1986; Englander and Wand, 1987). Residues present only once served as sequential starting points. This leads to the recognition of more or less long fragments placed in a unique way in the sequence. Figure 3 shows an example of a sequential pathway from Cys^{30} to Asn^{40}. These data were acquired with a mixing time of 150 ms for which spin diffusion effects should be weak. Intraresidual $d_{\alpha N}$ and interresidual $d_{\alpha N}(i,i+1)$ connectivities are mainly used. For Asn^{31} and Cys^{36}, αH resonances are hidden by the water presaturation and $d_{\beta N}$ connectivities are used. αH-NH cross peaks are well spread out except for αH32-NH32/αH37-NH38 and αH32-NH33/αH35-NH35. In these cases further βH-NH and αH-βH connectivities appearing in other regions of the spectrum confirm both Gly^{32}-Lys^{33} and Val^{37}-Cys^{38} couplings. On the other hand, experiments performed below and above 294 K verify this sequential pathway. Information extracted from Fig. 3 is included in Fig. 4 where the main connectivities between consecutive residues are summarized. Using $d_{NN}(i,i+1)$, $d_{\alpha N}(i,i+1)$ and $d_{\beta N}(i,i+1)$ connectivities, 3 stretches from Cys^3 to Asn^{12}, His^{13} to Tyr^{29} and Cys^{30} to Asn^{40} were identified. Breaks in this pathway result from the disappearance of the αH resonance of Asn^{12} masked by the water signal at 294 K and from the fact that $d_{\beta N}(29,30)$ connectivities only arise for longer mixing times (Fig.

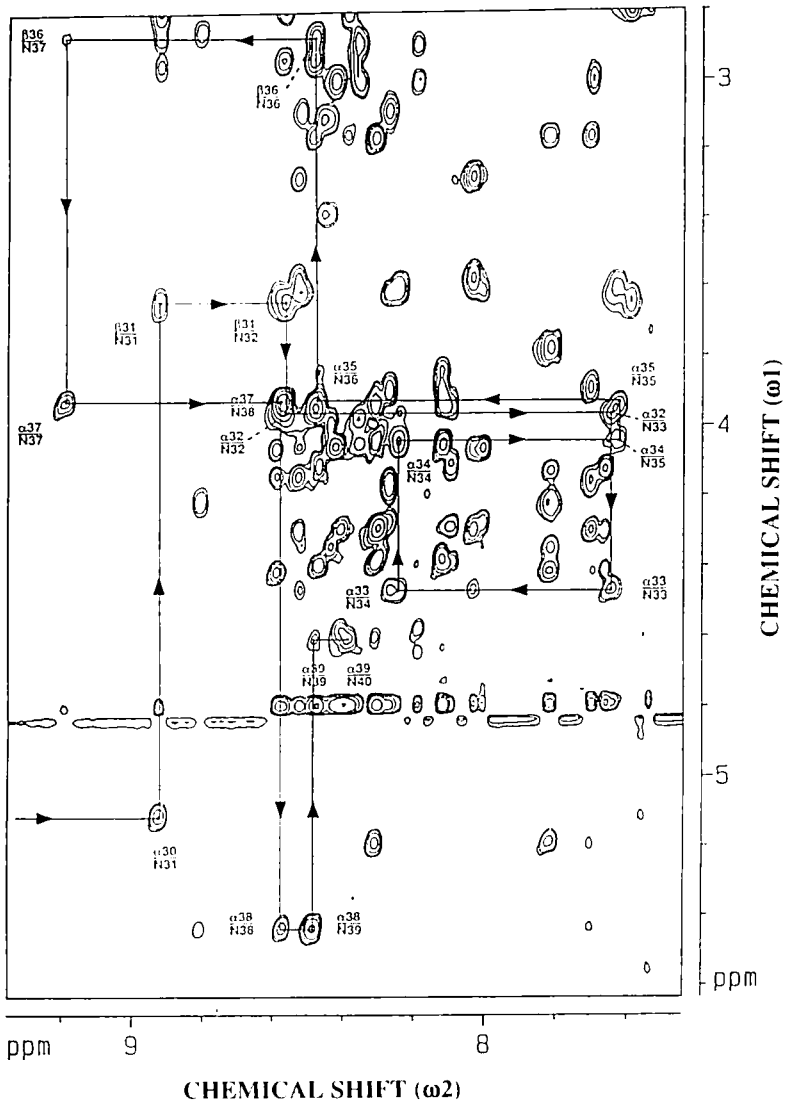

Fig. 3. Part of a 500 MHz NOESY (150 ms) spectrum of recombinant defensin A in ^1H_2O (pH 4.9, 8.9 mM, 294 K). The sequential pathway (i,i + 1), mainly involving $d_{\alpha N}$ connectivities (see text), is displayed from Cys^{30} to Asn^{40}.

2). Looking at other side-chain connectivities and examining NOESY spectra acquired at 300 ms mixing time allows the establishment of a continuous pathway from the N- to the C-terminal protons of recombinant defensin A (not shown).

Secondary structure elements

Helices are generally characterized by strong NH(i)-βH(i), βH(i)-NH(i + 1), NH(i)-NH(i + 1) and medium-range αH(i)-NH(i + 3) and αH(i)-βH(i + 3) connectivities (Wüthrich, 1986). Examination of Fig. 4 shows that strong $d_{\beta N}(i,i)$, $d_{\beta N}(i,i+1)$ and $d_{NN}(i,i+1)$ connectivities are principal-

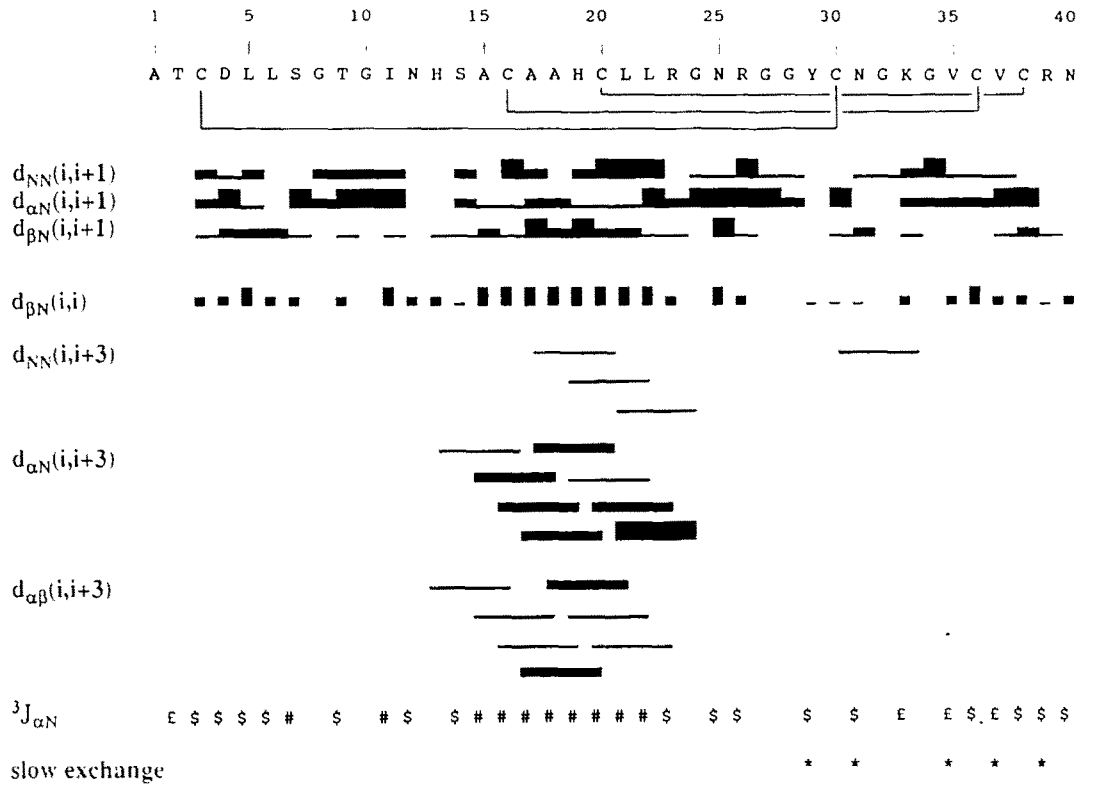

Fig. 4. Summary of some short and medium-range NOE connectivities observed in a 500 MHz NOESY (150 ms) spectrum of recombinant defensin A (pH 4.9, 8.9 mM, 294 K). The sequence is displayed with the one-letter code and the 3 disulfide bridges are enhanced by solid lines. The intensity of NOEs is classified as strong, medium and weak by the height of lines. Amide protons in slow exchange with ^2H_2O at 283 and 294 K are indicated by a star. $^3J_{NH-\alpha H}$ values measured from DQF-COSY experiments at 313 K are classified as large (£ ≥ 9 Hz), medium (7 Hz ≤ $ < 9 Hz) and small (# < 7 Hz).

ly found from Ala[15] to Arg[23]. In Fig. 5, two portions of a NOESY spectrum which correspond to αH-NH and αH-βH NOE domains are presented. The digital resolution of such a spectrum is sufficient to observe the pair of doublets of Gly[24]. Each vertical solid line, related to one αH(i) resonance, intercepts both αH(i)-NH(i) and αH(i)-NH(i+3) cross peaks in the bottom part of Fig. 5, and the αN(i)-βH(i+3) cross peaks in the upper part. For His[19] and Cys[20] which have well-recognizable β- and β′-protons, $d_{\alpha\beta'}(i,i+3)$ connectivities are also shown. Medium-range connectivities including those of Fig. 5 appear principally from Ser[14] to Arg[23]. Helices are also characterized by small $^3J_{NH-\alpha H}$ coupling constants ($\approx$ 5 Hz) in contrast to higher values ($\approx$ 9 Hz) for β-sheets (Pardi et al., 1984). In order to measure $^3J_{NH-\alpha H}$ values, the DQF-COSY data set was acquired at 313 K and data were collected, then processed, with high digital resolution (0.7 Hz/point in F2). Coupling constants were classified in 3 groups as reported in Fig. 4. $^3J_{NH-\alpha H}$ smaller than 7 Hz were observed from Ala[15] to Leu[22]. Examination of αH resonance chemical shifts shows that the latter are shifted upfield (less than 4.6 ppm) from Ser[14] to Gly[24], which is in agreement with a helical portion between these residues (Pastore and Saudek, 1990). This whole set of results demonstrates

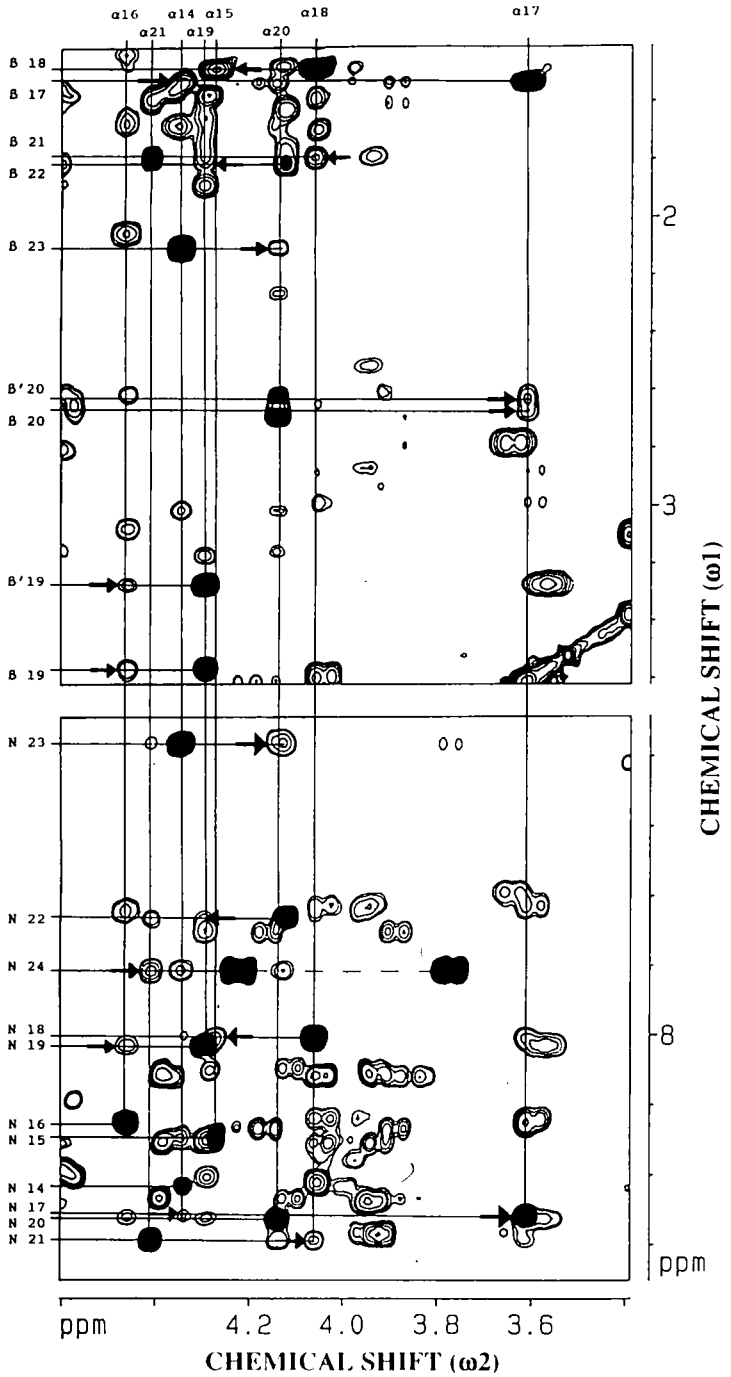

Fig. 5. Parts of a 500 MHz NOESY (150 ms) spectrum of recombinant defensin A in ^1H_2O (pH 4.9, 8.9 mM, 294 K). The αH-βH and αH-NH cross-peak regions covering residues from Ser^{14} to Gly^{24} are displayed at the top and bottom, respectively. NH, αH and βH chemical shifts are enhanced by solid lines. A dashed line connects the two doublets of Gly^{24} (bottom). Black filled cross peaks denote intraresidual NOEs from Ser^{14} to Gly^{24} and arrows indicate (i,i + 3) connectivities.

that the helix starts at Ser^{14} and continues as far as Gly^{24}. A search of $d_{\alpha N}(i,i+4)$ connectivities shows the presence of an α-helix type more than a 3_{10} helix for which αH(i)-NH(i+2) cross peaks are expected (Wüthrich et al., 1984). The strong intensity of $d_{\alpha\beta}(i,i+3)$ connectivities as described above supports this assumption.

Among the backbone protons, weak αH(i)-NH(i), strong αH(i)-NH(i+1) and weak NH(i)-NH(i+1) sequential NOEs are predicted in β-sheets (Wüthrich, 1986). These connectivities are observed from Gly^{24} to Tyr^{29} and from Lys^{33} to Arg^{39} (Fig. 4). Figure 4 also shows that high values of $^3J_{NH-\alpha H}$ mainly occur in the C-terminal part of the protein. In addition, αH chemical shifts of Asn^{25}, Tyr^{29}, Cys^{30}, Asn^{31} and Cys^{38} arise downfield from the ^1H_2O signal. This is in agreement with an extended structure for the residues concerned. Chemical $^1H/^2H$ amide proton exchange gives a wealth of information since protons involved in hydrogen bonds are very persistent. At 283 K exchange rates were too fast to permit any measurement of NH exchange rates for the helical portion but amide protons of Tyr^{29}, Asn^{31}, Val^{35}, Val^{37} and Arg^{39} were still detected after several hours (Fig. 4). We emphasize here that a hydrogen bond present in a β-sheet, between NH(i) on one strand and the carbonyl oxygen of residue (j) on the other strand, is consistent with NH(i)-NH(j), NH(i)-αH(j+1), and αH(i−1)-αH(j+1) connectivities. Such long-range NOEs were observed for residues 26–(39, 40), 28–(38, 39), 29–(37, 38), 30–(35, 36, 37) and 31–(34, 35, 37) as reported in Fig. 6 where all the medium and long-range connectivities are depicted. An illustration of hydrogen bonds and interstrand connectivities observed in the antiparallel β-sheet, from Arg^{26} to Asn^{40}, is given in Fig. 7. This representation requires a turn located up to Asn^{31}-Gly^{34}. Analysis of connectivities observed between Gly^{32} and Val^{35} shows that the conformation would be most likely a tight turn, still difficult to classify at this stage. In addition, a bend is requested at the end

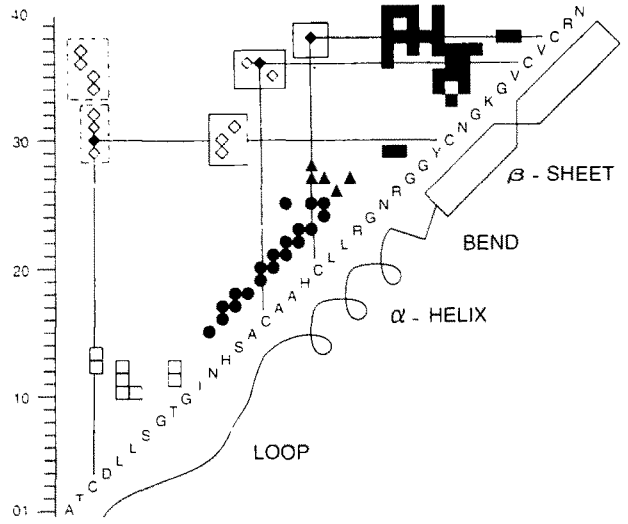

Fig. 6. Summary of medium and long-range connectivities (i, j≥i+3) observed in NOESY maps (500 MHz, 150 ms) of recombinant defensin A in ^1H_2O or ^2H_2O ($p^1H = p^2H$ 4.9, 294 K). The sequence is displayed on the diagonal (one-letter code) with elements of secondary structure in which residues are involved. Symbols may represent more than one NOE between two residues of the diagonal. (|), loop; (●), α-helix; (▲), bend; (■), β-sheet. Off-diagonal diamonds placed in dashed boxes and solid boxes connect the two strands of the β-sheet to the N-terminal part and to the helix, respectively. Connectivities presented with filled diamonds (◆) are related to S-S bridges (solid lines).

of the helix (Gly^{24}-Arg^{26}; see Fig. 6). The complete and rigorous assignment of all long-range NOEs was done and leads to the data of Fig. 7. In Fig. 6, very long-range connectivities are presented, especially those related to the position of (i) the N-terminal part in comparison of each β-strand (including the first S-S bridge); (ii) the helix facing the β-sheet (including the two other S-S bridges). All this information was introduced as constraints to model the global folding of recombinant defensin A.

Temperature variations

We took advantage of the experiments recorded between 283 and 313 K to approach qualitatively the internal dynamics of recombinant defensin A. It is widely accepted that variations of NH chemical shifts with temperature are related to screening from the solvent of these protons, or to their participation to hydrogen bonds. The temperature coefficients: $\Delta\delta_{NH}/\Delta T$, are generally of the order of 10^{-3} ppm K^{-1} and the largest negative values (high-field shift with increasing temperature) generally correspond to protons exposed to the solvent or involved in conformational changes increasing their exposure. Except for the Ala^{1} residue, we have measured all NH chemical shifts in a temperature range of 30 K. They vary linearly as illustrated for Leu^{5}, His^{19} and Arg^{23} in Fig. 8. Residues have been classified in 4 groups depending on their temperature coefficient val-

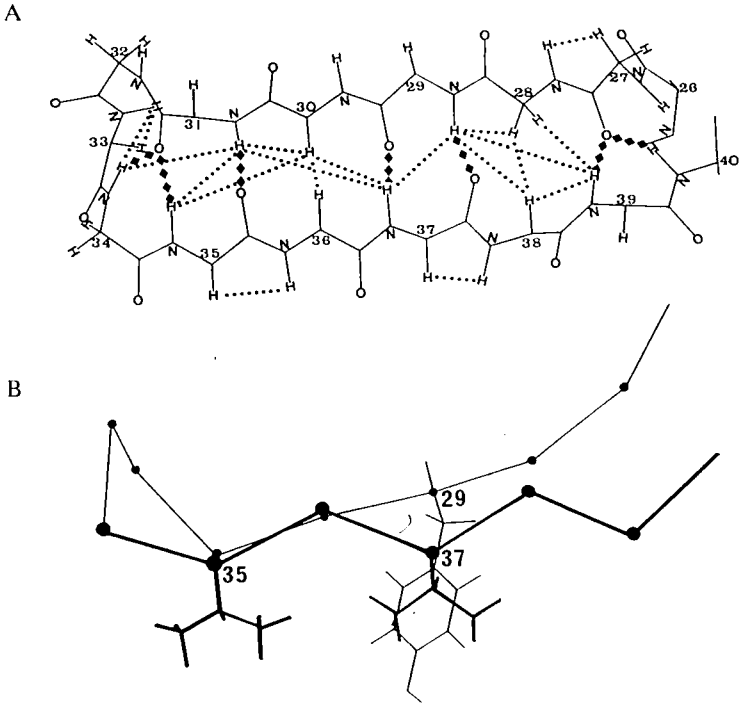

Fig. 7. The β-sheet structure (model A_4; see text) of the C-terminal fragment of recombinant defensin A. In (A) the backbone is displayed from Arg^{26} to Asn^{40} and C^{α} atoms are numbered. Hydrogen bonds are represented by filled diamonds (◆) and only a selection of typical (i,i + 1) and interstrand NOEs is shown by dotted lines. (B) represents a profile view of (A) with respect to a horizontal axis. Among the backbone, only C^{α} atoms (●) are shown in (B) and bold lines enhance the foreground strand. Side chains of Tyr^{29}, Val^{35} and Val^{37} are also displayed.

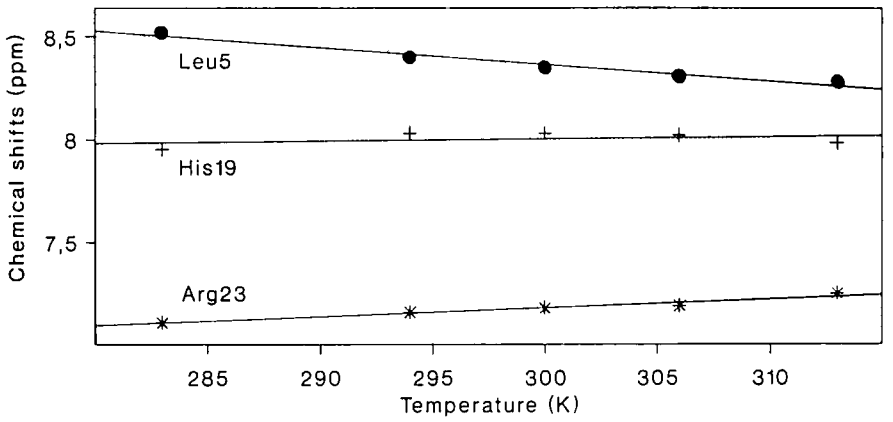

Fig. 8. Amide proton chemical shifts of Leu^5 (●), His^{19} (+) and Arg^{23} (*) extracted from NOESY and HOHAHA experiments of recombinant defensin A in water (pH 4.9) at several temperatures. Solid lines fit the data within experimental errors.

ues (Table 2). NH chemical shifts of Leu^{22}, Arg^{23}, Gly^{24}, Asn^{25}, Lys^{33} and Gly^{34} singularly increase with temperature which leads to positive coefficients. The first 4 residues are located at the end of the helix and precede the β-structure deduced from the network of NOE connectivities. It is then suggested that some special rearrangement can occur in this region, modifying the environment of NH protons. Lys^{33} and Gly^{34} are involved in a tight turn favoring the interactions between two antiparallel chains. A second group of residues is characterized by small absolute values of $\Delta\delta_{NH}/\Delta T$. Tyr^{29}, Asn^{31}, Val^{35}, Val^{37} and Arg^{39} amide protons exhibit low $^1H/^2H$ exchange rates and they are in a region corresponding to a β-structure. Therefore, these 5 protons are certainly engaged in hydrogen bonds stabilizing this structure. Ala^{15}, Ala^{18} and His^{19} are located at the beginning of the helix. Four other residues (Cys^3, Asp^4, Ser^7, Gly^{10}) found in this group suggest that the N-terminal part is not fully statistical as attested by few long-range NOEs shown in Fig. 6. The two next groups include amide protons more exposed to the solvent or located in more flexible domains. Eight residues belong to the N-terminal fragment whereas most of the other resi-

TABLE 2
TEMPERATURE COEFFICIENTS[a] OF NH CHEMICAL SHIFTS[b] OF RECOMBINANT INSECT DEFENSIN A IN ^1H_2O (pH 4.9)

$\Delta\delta_{NH}/\Delta T$[c]	Residue[d]
$+2 \leq \Delta\delta/\Delta T < +5$	L22, R23, G24, N25, K33, G34
$-2 < \Delta\delta/\Delta T < +2$	C3, D4, S7, G10, A15, A18, H19, L21, R26, G27, Y29, N31, V35, V37, R39
$-5 < \Delta\delta/\Delta T \leq -2$	T2, L6, G8, T9, I11, N12, H13, S14, C20, G28, C36
$-9 < \Delta\delta/\Delta T \leq -5$	L5, C16, A17, C30, G32, C38, N40

[a] Except for A1.
[b] Measured on NOESY and HOHAHA maps recorded at 283, 294, 300, 306 and 313 K.
[c] Values in ppm 10^{-3} K^{-1}.
[d] One-letter code.

dues are found in secondary structure elements. The hydrophobic pair of residues Leu^5-Leu^6, adjacent to Cys^3 and to the negatively charged Asp^4 residue, is especially sensitive to temperature variations. Cys^{16} and Ala^{17} amide protons, enclosing a Cys residue and located in the first turn of the helix, are theoretically not engaged in hydrogen bonds. Those of Cys^{36} and Cys^{38} should be exposed to the solvent (see the β-sheet in Fig. 7). We noted previously the particular shape of cross peaks involving the Cys^{30} amide proton, as well as its particular chemical shift (Table 1 and Fig. 2). This proton should be exposed to water and its large temperature coefficient suggests the existence of a special local dynamics in this region of the β-sheet.

Molecular modeling

A total of 54 structures was generated by using the DISMAN program. During the distance geometry (DG) optimization process, violations of constraints were analyzed and led to a re-examination of NOE data. We ended up with an input file (vide supra) including: 12 hydrogen bond constraints, 21 Φ angle restraints, 133 NOE constraints (32 intraresidual, 50 sequential, 42 medium and long-range backbone constraints and 9 constraints involving side-chain atoms). Additionally, 9 constraints ensured the formation of S-S bridges. Four structures (E_1, E_2, E_3, E_4) were finally selected, these structures correspond to the 4 lowest values obtained for the target function used in DISMAN, i.e. less than 40 $Å^2$. Note that for these selected structures, no S-S bridge constraints are violated and that the number of NOE violations is satisfactorily low (Table 3). The RMSD values (Root-Mean-Square Deviation on atomic coordinates) of these 4 structures were calculated by using successively each structure as a reference (Table 4). Structures E_2 and E_3 are much more similar to E_1 than structure E_4 and the large RMSD values between E_4 and E_1, E_2 and E_3 are due to the poorly determined conformation of the fragment between Cys^3 and Asn^{12}. The RMSD values calculated for backbone atoms from residues 14 to 40 (α-helix and β-sheet regions) are much lower for the 4 structures indicating that this fragment is well defined. These 4 structures were then subjected to energy minimization (EM). During this step, the constraints involving S-S bridges were removed and 4 structures: A_1, A_2, A_3 and A_4 were generated from E_1, E_2, E_3 and E_4, respectively. The RMSD values between DG and EM structures are relatively low (1.32, 1.78, 0.74 and 1.09 Å, respectively). After the minimization process the global folding of the backbone was preserved. However, there was a reorientation of side chains and of the turn involved in the

TABLE 3
RESIDUAL ERRORS OF 4 DISMAN STRUCTURES: E_1, E_2, E_3, and E_4

	Number of NOE violations				Final value of the target function ($Å^2$)
Range:	0.5 Å 1.0 Å		> 1.0 Å		
Violation of the:	lower limit	upper limit	lower limit	upper limit	
E_1	1	13	1	2	29
E_2	3	10	0	3	40
E_3	1	8	1	0	22
E_4	1	7	0	0	19

TABLE 4
RMS DEVIATIONS (in Å) OF 4 DISMAN STRUCTURES BEFORE AND AFTER (in bold) ENERGY REFINEMENT (AMBER)

Structure	E_1 (A_1)	E_2 (A_2)	E_3 (A_3)	E_4 (A_4)	Footnotes
		2.40	2.24	4.14	[a]
E_1 (A_1)	**ref.**	1.78	1.65	1.99	[b]
		3.21	4.00	5.30	[c]
	2.48		2.60	5.01	[a]
E_2 (A_2)	**1.17**	**ref.**	2.40	3.02	[b]
	3.27		3.72	5.80	[c]
	2.46	**2.00**		4.52	[a]
E_3 (A_3)	**1.63**	**1.30**	**ref.**	1.57	[b]
	3.41	**3.20**		5.26	[c]
	3.83	**4.40**	**4.36**		[a]
E_4 (A_4)	**1.97**	**2.06**	**2.00**	**ref.**	[b]
	4.90	**5.05**	**5.07**		[c]
Energy (total)	**−298**	**−244**	**−210**	**−312**	[d]
Energy (constraints)	**41**	**66**	**71**	**54**	[d]

[a] For C^{α} atoms.
[b] For C^{α} atoms of the fragment (14–40).
[c] For all atoms.
[d] In kcal mol^{-1}.

β-sheet so that the backbone geometry was modified to allow the formation of hydrogen bonds with a proper geometry. A superimposition of the 4 energy-refined structures (Fig. 9) shows that the global folding of the 4 structures is almost similar and comprises (i) a large loop (Asp^4-Asn^{12}); (ii) an α-helical region (Ser^{14}-Gly^{24}); (iii) an antiparallel β-sheet made of two strands (Gly^{27}-Asn^{31} and Val^{35}-Asn^{40}). The α-helix is connected by two loose turns (no hydrogen bond) on one side to the large loop (4–12) and on the other side to the upper strand of the β-sheet. A tight turn connects the two strands of the β-sheet. This turn cannot be classified as a standard γ- or β-turn (Némethy and Printz, 1972; Rose et al., 1985; Milner-White et al., 1988) and its conformation is slightly different for the 4 structures. Figure 10 reveals a poorly determined conformation for the large loop: Asp^4-Asn^{12}. This is confirmed by the RMSD values (Table 4) which dropped when only the C-terminal fragment (14–40) was used for the calculation since, for the N-terminal less well defined domain, few NOEs were available (Figs. 4 and 6) which led to rather contorted structures in which several turns appear. Incertitudes on the structure may reflect a large flexibility of this loop region. The 4 EM structures have negative total energies (Table 4) and NOE constraint energies range from 40 to 70 kcal mol^{-1}. The two lowest values are obtained for A_1 and A_4 structures which consequently lead to low RMSD values between calculated and experimentally derived interproton distances: 0.35 and 0.36 Å, respectively. For A_2 and A_3, they both reach 0.41 Å. Nevertheless, all these values are smaller than the incertitude on experimental data. The bond, an-

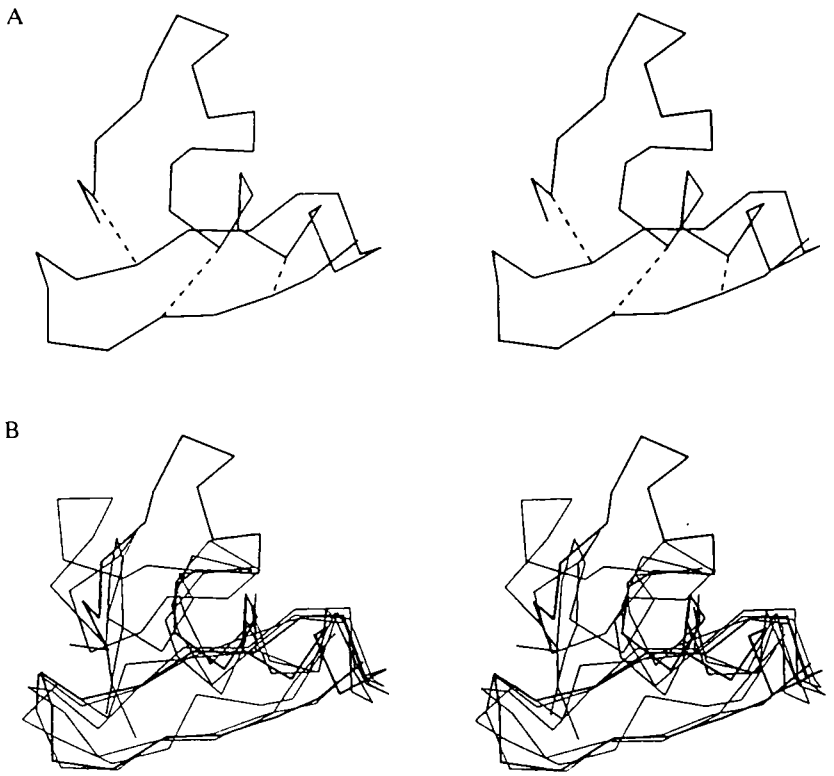

Fig. 9. Stereoviews of the backbone (C^{α} atoms only) of defensin A emerging from our NMR data. (A) gives a picture of structure A_4 (see text) with disulfide bridges symbolized by dashed lines. In (B), C^{α} atoms of structures A_1, A_2, A_3 and A_4 are best-fit superimposed on the basis of secondary structure elements (α-helix and β-sheet).

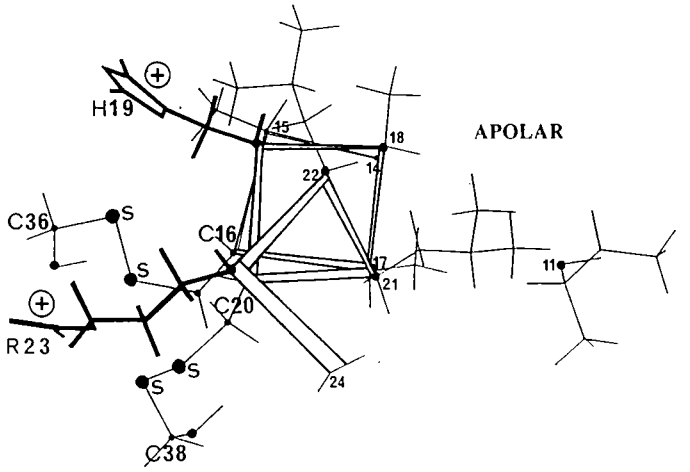

Fig. 10. The α-helical structure (model A_4, see text) of the fragment Ser^{14}-Gly^{24} of recombinant defensin A nearly viewed along the principal axis of the helix. The backbone is enhanced by double lines and C^{α} atoms are labeled. Disulfide bridges (16,36) and (20,38) indicate the orientation of the β-sheet with respect to the helix. Side chains of hydrophobic residues (Ile^{11}, Ala^{17}, Ala^{18}, Leu^{21}, Leu^{22}) and charged residues (His^{19}, Arg^{23}; in bold) are displayed. The helix presents a polar side and an apolar surface.

gle and torsion energy terms are reasonably low and show that distortions from the standard geometry are not important. In fact, only two peptide bonds in structures A_2 and A_3 deviate significantly from the standard planar geometry. These two peptide bonds are involved in the hairpin loop of the β-sheet thus leading to the poorer fit with experimental data for structures A_2 and A_3.

DISCUSSION AND CONCLUSION

Well-resolved 2D maps enabled the complete assignment of ^1H spectra of recombinant defensin A except for NH protons belonging to the N-terminal Ala^1 residue. The analysis of NOE connectivities and $^3J_{NH-\alpha H}$ coupling constants revealed the existence of a helical fragment and of an antiparallel β-structure connected by two Cys^{16}-Cys^{36} and Cys^{20}-Cys^{38} disulfide bridges. In the β-structure, 4 NH protons exposed to water and C-terminal NH protons have high $\Delta\delta/\Delta T$ coefficients whereas other hydrogen-bonded NH protons exhibit small coefficients and slow $^1H/^2H$ exchange rates. The conformational behavior of the helical fragment seems rather complex. As expected, the NH protons engaged in the first turn of the helix were found to be accessible to the solvent. The NH protons of the following turns exhibited positive temperature coefficients but intermediate $^1H/^2H$ exchange rates. This suggests that the end of the helix (Cys^{20}-Gly^{24}) which precedes the bend (24-25-26) is somewhat distorted and rearranges when temperature is increased. Geometrical modeling followed by energy minimization led to a global folding which is in excellent agreement with most of the experimental data. The 3 disulfide bonds certainly play an important role in the stabilization of the 3D structure of a small protein such as defensin A. Nevertheless, the progressive increase of the number of constraints during the modeling process (vide supra) clearly revealed the crucial role of secondary elements in the spatial organization. In the first structures generated with a reduced number of NOEs, some pseudomirror images (Pastore et al., 1991) of these structures appeared, which we eliminated by taking into account the full set of 133 NOEs. The complete set of intra- and interresidual NOEs (more than 4 times larger) will be used in a future refinement study which will include a dynamics simulation. It will be of especial interest to compare the results of such a simulation with the qualitative picture resulting from the present analysis of temperature effects. Schematically, the global folding emerging from the present study shows 3 regions: (i) a loop formed by the 1–13 N-terminal fragment; (ii) an α-helix (14–24); (iii) an antiparallel β-sheet (27–40) with a turn involving residues 31–34. Such a folding does not create a clear separation between an external charged region and an internal hydrophobic nucleus, but a 'mosaic' structure. This situation differs from that of human HNP-3 defensin, where a dimeric structure is stabilized by hydrophobic interactions (Hill et al., 1991).

Seven strongly hydrophobic residues (Leu, Ile, Val) are present in the sequence of defensin A; 6 of them appear almost as pairs of residues (Leu^5-Leu^6, Leu^{21}-Leu^{22} and Val^{35}-Val^{37}) and form small hydrophobic domains. These domains might be involved in the stability of the structure, as well as in possible interactions of defensin A with lipids or receptors. In our models of backbone folding, Leu^5 and Leu^6, present in the N-terminal loop, are oriented towards the exterior. Figure 10, presenting a view along the principal axis of the helix, shows the position of the Leu^{21}-Leu^{22} pair of residues. This helix possesses an amphipathic character with a hydrophobic (3 Ala and 2 Leu residues) and a positively charged surface (His^{19} and Arg^{23}). Interestingly, the orientation of Ile^{11} is close to the orientation of Ala^{17} and Leu^{21}, forming a hydrophobic cluster. Cys^{16} and Cys^{20}, involved in the two disulfide bridges linking the helix and the β-sheet, have a common

orientation opposite to the hydrophobic surface of the helix. This suggests that this surface is exposed to the solvent. Another hydrophobic cluster is observed in the β-sheet where the side chains of Val^{35} and Val^{37} are oriented as the side chain of Tyr^{29} (Fig. 7). In contrast to large proteins, recombinant defensin A contains no hydrophobic nucleus and the relative exposure to the solvent of the 3 small hydrophobic clusters seems to be balanced by the number of electrical charges of the protein.

From the model proposed by Hanzawa et al. (1990) for sapecin, there are no fundamental differences between the global folding of recombinant defensin A in water and sapecin in methanol. This is certainly due to the existence of two similar secondary-structure elements stabilized by two disulfide bridges. Nevertheless, one can expect differences in the local organization of these proteins, especially in the bent regions and for the conformation of disulfide bridges and side chains.

In defensin A, a fragment CXXXC (one-letter code where X describes any amino acid except a Cys residue) involved in the α-helix is linked through two S-S bridges to a fragment CXC located on a strand of the β-sheet. This kind of arrangement between a helix and a β-strand has been reported as the CSH motif (Cysteine Stabilized α-Helix motif) by Tamaoki et al. (1991). It is present in several peptides and proteins such as endothelins (Aumelas et al., 1991a; Tamaoki et al., 1991) and toxins from various venom (Pease and Wemmer, 1988; Aumelas et al., 1991b; Darbon et al., 1991; Kobayashi et al., 1991). A striking structural analogy is found with charybdotoxin (ChTX) from scorpion venom, recently studied by ^1H NMR (Bontems et al., 1991; Takashima et al., 1991). This is further substantiated by the third disulfide bridge which links, in both cases, the N-terminus fragment with a Cys residue located between the two elements of the CSH motif (Fig. 11). A more detailed comparison between the two 3D structures will be of especial interest once more refined models will be available.

In the B isoform of *Phormia* defensin, the Gly^{32} residue is replaced by an Arg residue. On the basis of the present model, it can be expected that such a substitution is more important for the interaction of defensin with other molecules than for the 3D structure itself. Defensin A has been known to act on the cytoplasmic membrane (Lambert et al., 1989) and sapecin binds to cardiolipin (Matsuyama and Natori, 1990). Our model, depicting the distribution of positive charges and apolar domains in defensin A, is in favor of both electrostatic and hydrophobic interactions between these proteins and membranes. This does not exclude the existence of specific receptors.

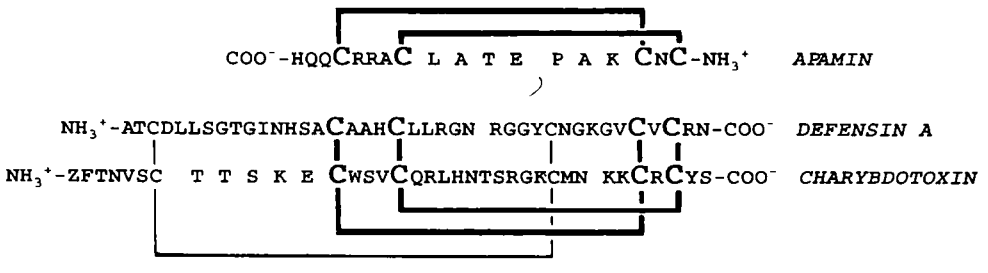

Fig. 11. Comparison of 3 sequences (one-letter code) exhibiting the CSH motif (in bold; Tamaoki et al., 1991). These sequences have been aligned according to the location of disulfide bridges and secondary structure elements that they rely (see text). Spaces between residues have been inserted for this purpose. Note that for apamin (from bee venom; Pease and Wemmer, 1988) the sequence is inverted. A third S-S bridge is also present in charybdotoxin (from scorpion venom; Bontems et al., 1991) and in defensin A.

The analysis at the molecular level of the mode of action of defensin A will obviously be rewarding.

In conclusion, this paper presents a first and complete analysis of ^{1}H NMR spectra of recombinant insect defensin A in water. In our experimental conditions, strong evidence is provided for the existence of two major secondary-structure elements consecutive to a large N-terminal loop. Geometrical modeling using the DISMAN and AMBER programs indicates a global folding which is in excellent agreement with our experimental NMR parameters. A basis is now established for further refinements of the 3D structure, including dynamics simulations.

ACKNOWLEDGEMENTS

This work was supported by the 'Centre National de la Recherche Scientifique', 'Université d'Orléans', 'Région Centre' and 'Ministère de la Recherche et de la Technologie'. Our appreciation goes to Catherine Fahrner for her technical assistance in the recombinant defensin A purification, to Jean Sabatié and Luc Caussin for fermentation of the recombinant yeast strain and to Henri Labbé for his helpful assistance in NOE data handling. We acknowledge Carolyn Roitsch for useful discussions and reading of the manuscript. The authors thank Pierre Chambon, Philippe Kourilsky and Jean-Pierre Lecocq for their interest in this project.

REFERENCES

Aumelas, A., Chiche, L., Mahe, E., Le-Nguyen, D., Sizun, P., Berthauld, P. and Perly, B. (1991a) *Int. J. Pept. Prot. Res.*, **37**, 315–324.

Aumelas, A., Chiche, L., Mahe, E., Le-Nguyen, D., Sizun, P., Berthauld, P. and Perly, B. (1991b) *Neurochem. Int.*, **18**, 471–475.

Bach, A.C., Selted, M.E. and Pardi, A. (1987) *Biochemistry*, **26**, 4389–4397.

Bassolino, D.A., Hirata, F., Kikchen, D.B., Kominos, D., Pardi, A. and Levy, R.M. (1988) *Int. J. Supercomp. Appl.*, **2**, 41–61.

Bax, A. and Davis, D.G. (1985) *J. Magn. Reson.*, **65**, 355–360.

Bazzo, R., Tappin, M.J., Pastore, A., Harvey, T.S., Carver, J.A. and Campbell, I.D. (1988) *Eur. J. Biochem.*, **173**, 139–146.

Bontems, F., Roumestand, C., Boyot, P., Gilquin, B., Doljansky, Y., Menez, A. and Toma, F. (1991) *Eur. J. Biochem.*, **196**, 19–28.

Braun, W. and Gō, N. (1985) *J. Mol. Biol.*, **186**, 611–626.

Brown, S.C., Weber, P.L. and Muller, L. (1988) *J. Magn. Reson.*, **77**, 166–169.

Bulet, P., Cociancich, S., Dimarcq, J.L., Lambert, J., Reichhart, J.M., Hoffmann, D., Hetru, C. and Hoffmann, J.A. (1991) *J. Biol. Chem.*, **266**, 24520–24525.

Casteels, P., Ampe, C., Jacobs, F., Vaeck, M. and Tempst, P. (1989) *EMBO J.*, **8**, 2367–2391.

Darbon, H., Weber, C. and Braun, W. (1991) *Biochemistry*, **30**, 1836–1845.

Dimarcq, J.L., Keppi, E., Dunbar, B., Lambert, J., Reichhart, J.M., Hoffmann, D., Rankine, S.M., Fothergill, J.E. and Hoffmann, J. (1988) *Eur. J. Biochem.*, **171**, 17–22.

Englander, S.W. and Wand, A.J. (1987) *Biochemistry*, **26**, 5953–5958.

Engström, Å., Engström, P., Tao, Z.J., Carlssen, A. and Bennich, H. (1984) *EMBO J.*, **3**, 2065–2070.

Groß, K.H. and Kalbitzer, H.R. (1988) *J. Magn. Reson.*, **76**, 87–99.

Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 3997–4004.

Hanzawa, H., Shimada, I., Kuzuhara, T., Komano, H., Kohda, D., Inagaki, F., Natori, S. and Arata, Y. (1990) *FEBS Lett.*, **269**, 413–420.

Hill, C.P., Yee, J. and Selsted, M.E. (1991) *Science*, **251**, 1481–1485.

Holak, T.A., Engström, Å., Kraulis, Per J., Lindeberg, G., Bennich, H., Jones, T.A., Gronenborn, A.M. and Clore, G.M. (1988) *Biochemistry*, **27**, 7620–7629.

Hultmark, D., Engström, Å., Bennich, H., Kapur, R. and Boman, H.G. (1982) *Eur. J. Biochem.*, **127**, 207–217.

Kagan, B.L., Selsted, M.E., Ganz, T. and Lehrer, R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 210–214.

Kobayashi, Y., Sato, A., Takashima, H., Tamaoki, H., Nishimura, S., Kyogoku, Y., Ikenaka, K., Kondo, T., Mikoshiba, K., Hojo, H., Aimoto, S. and Moroder, L. (1991) *Neurochem. Int.*, **18**, 525–534.

Kuzuhara, T., Nakajima, Y., Matsuyama, K. and Natori, S. (1990) *J. Biochem.*, **107**, 514–518.

Lambert, J., Keppi, E., Dimarcq, J.L., Wicker, C., Reichhart, J.M., Dunbar, B., Lepage, P., Van Dorsselaer, A., Hoffmann, J., Fothergill, J. and Hoffmann, D. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 262–266.

Lehrer, R.I., Ganz, T. and Selsted, M.E. (1991) *Cell*, **64**, 229–230.

Lepage, P., Bitsch, F., Roecklin, D., Keppi, E., Dimarcq, J.L., Reichhart, J.M., Hoffmann, J.A., Roitsch, C. and Van Dorsselaer, A. (1991) *Eur. J. Biochem.*, **196**, 735–742.

Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41**, 95–117.

Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.

Matsuyama, K. and Natori, S. (1988a) *J. Biol. Chem.*, **263**, 17112–17116.

Matsuyama, K. and Natori, S. (1988b) *J. Biol. Chem.*, **263**, 17117–17121.

Matsuyama, K. and Natori, S. (1990) *J. Biochem. (Tokyo)*, **108**, 128–132.

Milner-White, E.J., Ross, B.M., Ismail, R., Belhadj-Mostefa, K. and Poet, R. (1988) *J. Mol. Biol.*, **204**, 777–782.

Némethy, G. and Printz, M.P. (1972) *Macromolecules*, **5**, 755–758.

Pardi, A., Billeter, M. and Wüthrich, K. (1984) *J. Mol. Biol.*, **180**, 741–751.

Pardi, A., Hare, D.R., Selsted, M.E., Morrison, R.D., Bassolino, D.A. and Bach, A.C. (1988) *J. Mol. Biol.*, **201**, 625–636.

Pastore, A. and Saudek, W. (1990) *J. Magn. Reson.*, **90**, 165–176.

Pastore, A., Atkinson, R.A., Saudek, V. and Williams, R.J.P. (1991) *Proteins*, **10**, 22–32.

Pease, J.H.B. and Wemmer, D.E. (1988) *Biochemistry*, **27**, 8491–8498.

Reichhart, J.M. and Achstetter, T. (1990) *Res. Immunol.*, **141**, 943–946.

Rose, G.D., Gierasch, L.M. and Smith, J.A. (1985) *Adv. Prot. Chem.*, **37**, 1–109.

Selsted, M.E., Brown, D.M., DeLange, R.J. and Lehrer, R.I. (1983) *J. Biol. Chem.*, **258**, 14485–14489.

Takashima, H., Kobayashi, Y., Tamaoki, H., Kyogoku, Y., Lambert, P., Kuroda, H., Chino, N., Watanabe, T.X., Kimura, T. and Sakakibara, S. (1991) In *Peptides 1990* (Eds, Giralt, E. and Andreu, D.) ESCOM, Leiden, pp. 557–559.

Tamaoki, H., Kobayashi, Y., Nishimura, S., Ohkubo, T., Kyogoku, Y., Nakajima, K., Kumagaye, S., Kimura, T. and Sakakibara, S. (1991) *Prot. Eng.*, **4**, 509–518.

Tappin, M.J., Pastore, A., Norton, R.S., Freer, J.H. and Campbell, I.D. (1988) *Biochemistry*, **27**, 1643–1647.

Terwilliger, T.C. and Eisenberg, D. (1982) *J. Biol. Chem.*, **257**, 6016–6022.

Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347–366.

Weiner, P. and Kollmann, P.A. (1981) *J. Comput. Chem.*, **2**, 287–303.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Wüthrich, K., Billeter, M. and Braun, W. (1984) *J. Mol. Biol.*, **180**, 715–740.