Journal of Biomolecular NMR, **12:** 523–534, 1998.
KLUWER/ESCOM

The NMR solution structure and characterization of pH dependent chemical shifts of the β-elicitin, cryptogein

Paul R. Gooley[a,*], Max A. Keniry[b], Roumen A. Dimitrov[c], Danny E. Marsh[a], David W. Keizer[a], Kenwyn R. Gayler[a] and Bruce R. Grant[a]
[a]*Department of Biochemistry & Molecular Biology, University of Melbourne, Parkville, VIC 3052, Australia;* [b]*Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia;* [c]*Department of Biochemistry, Wageningen Agricultural University, 6703 HA Wageningen, the Netherlands*

Received 29 May 1998; Accepted 8 July 1998

Key words: cryptogein, elicitin, electrostatics, pKa, structure

Abstract

The NMR structure of the 98 residue β-elicitin, cryptogein, which induces a defence response in tobacco, was determined using ^{15}N and ^{13}C/^{15}N labelled protein samples. In aqueous solution conditions in the millimolar range, the protein forms a discrete homodimer where the N-terminal helices of each monomer form an interface. The structure was calculated with 1047 intrasubunit and 40 intersubunit NOE derived distance constraints and 236 dihedral angle constraints for each subunit using the molecular dynamics program DYANA. The twenty best conformers were energy-minimized in OPAL to give a root-mean-square deviation to the mean structure of 0.82 Å for the backbone atoms and 1.03 Å for all heavy atoms. The monomeric structure is nearly identical to the recently derived X-ray crystal structure (backbone rmsd 0.86 Å for residues 2 to 97) and shows five helices, a two stranded antiparallel β-sheet and an Ω-loop. Using ^{1}H,^{15}N HSQC spectroscopy the pKa of the N- and C-termini, Tyr^{12}, Asp^{21}, Asp^{30}, Asp^{72}, and Tyr^{85} were determined and support the proposal of several stabilizing ionic interactions including a salt bridge between Asp^{21} and Lys^{62}. The hydroxyl hydrogens of Tyr^{33} and Ser^{78} are clearly observed indicating that these residues are buried and hydrogen bonded. Two other tyrosines, Tyr^{47} and Tyr^{87}, show pKa's > 12, however, there is no indication that their hydroxyls are hydrogen bonded. Calculations of theoretical pKa's show general agreement with the experimentally determined values and are similar for both the crystal and solution structures.

Introduction

A challenging question in the biology of the interactions between plants and pathogens, is the identification of the molecules that interact to initiate defence responses. Molecules from each partner in the interaction presumably bind, resulting in the induction of signal cascades that lead to the generation of an antimicrobial environment and successful repulsion of the attempted invasion. Generally, the complexity and heterogeneity of the molecules produced by many microorganisms which elicit such plant responses, has hindered the determination of their structures. However, members of a subgroup of such elicitors which are peptides or proteins have been purified, cloned and/or sequenced, thus enabling the examination of their structure/function properties (De Wit, 1992; Kooman-Gersmann et al., 1996; Jones et al., 1994).

One group of protein elicitors, the elicitins, has properties that make them particularly well-suited as model elicitors to use in the study of structure/function relationships in their interaction with plants. Elicitins are small proteins that are secreted by members of the genus *Phytophthora* and by a limited number of species in the genus *Pythium* (Huet et al., 1995; Gayler et al., 1997) and show no homology to other known families of proteins. The role of the elicitins in the bi-

*To whom correspondence should be addressed.
Supplementary material: A table with ^{1}H, ^{13}C and ^{15}N assignments of cryptogein is available from the authors.

ology of the microorganisms that secrete them remains in doubt (Grant et al., 1996). However, these proteins have the capacity to act as elicitors of both the hypersensitive response and systemic acquired resistance in tobacco (Bonnet et al., 1996) and in some members of the family *Cruciferae*, particularly radish (Kamoun et al., 1993). Their capacity to act as elicitors in these species, together with their small size and highly conserved amino acid sequences, make elicitins ideally suited as compounds in which the effect of substitution of specific amino acids on the three dimensional structures can be determined, and related to alterations in biological activity.

Towards understanding these structure/function relationships the crystal and solution structures of cryptogein have been reported (Boissy et al., 1996; Fefeu et al., 1997). The crystal structure showed a unique fold consisting of six helices, a two stranded β-sheet and an Ω-loop, where the loop and sheet formed a "beak motif". The protein also showed two salt bridges: the N-terminus to Asp^{72}, and Asp^{21} to Lys^{62}. The solution structure was similar, but showed some differences in the 64–70 and 83–88 region and the salt bridge between Asp^{21} and Lys^{62} was not apparent. Many of the ionizable residues of cryptogein are the least conserved residues and include the N-terminus, Lys^{13}, Asp^{21}, Lys^{39}, Lys^{48}, Lys^{61}, Asp^{72}, Tyr^{85} and Lys^{94}. At least some of these surface residues are expected to be functionally important, for example, substitution of Lys^{13} with Val reduces the activity of cryptogein in tobacco by 100-fold (O'Donohue et al., 1995). Characterization of the surface properties, such as the ionization states of surface residues, of elicitins will assist in understanding the molecular interactions which these proteins can undergo. NMR spectroscopy is the method of choice for determining pKa's of individual residues (Forman-Kay et al., 1992; Nakamura, 1996). As the N-H bond is more easily polarized than the C-H bond, the ^{15}N resonance shows large chemical shift changes with pH and therefore the 2D ^{1}H, ^{15}N HSQC experiment is an excellent method for analysing pH dependencies of chemical shift.

In this report we describe the refined solution structure of cryptogein, and show that under millimolar solution conditions the protein forms a dimer. The individual subunits, however, are nearly identical to the crystal structure. The experimental pKa's of a number of ionizable residues are determined and compared to the theoretical pKa's calculated with both the solution and crystal structures.

Materials and methods

Phytophthora growth conditions and cryptogein purification

Cryptogein was purified from liquid cultures of *P. cryptogea* Pethybr & Laff (Isolate P7407, supplied by Professor M. Coffey, University of California). For $^{13}C/^{15}N$ labelling, *P. cryptogea* was grown in 500 ml of modified High Phosphate Ribeiro medium (Fenn and Coffey, 1984) pH 6.2, which contained 50 mM ^{13}C-glucose as a carbon source (or ^{12}C-glucose if only ^{15}N-labelling was required), 3 mM ^{15}N-ammonium chloride and 1.5 mM potassium ^{15}N-nitrate, as nitrogen sources.

Cultures were grown in the light for 21 days at 26 °C. Culture medium was harvested by filtration through two GF/C filters (Whatman) to separate mycelia from the culture filtrate. Crude culture filtrate was concentrated and subjected to ammonium sulfate (55%) precipitation. Cryptogein remained in the supernatant, which was desalted by dialysis against 10 mM sodium acetate prior to separation by cation exchange chromatography using a S Hyper D column (Beckman).

NMR spectroscopy

NMR measurements were performed at 40 °C on either ^{15}N-labelled or $^{15}N,^{13}C$ double labelled protein on Varian INOVA-400 and Varian INOVA-600 spectrometers. Cryptogein was dissolved to a concentration of 1 mM in 90% H_2O/10% $^{2}H_2O$, at pH 6.8. The following experiments were collected using the ^{15}N-labelled sample at 400 MHz: 2D $^{1}H,^{15}N$-HSQC, 3D $^{1}H,^{15}N$-TOCSY-HSQC (40 ms mixing time), NOESY-HSQC (100 ms mixing time) (Zhang et al., 1994), HSQC-NOESY-HSQC (100 ms mixing time) (Frenkiel et al., 1990), HNHA (Vuister and Bax, 1993) and HNHB (Archer et al., 1991); and the following experiments were collected with the $^{13}C,^{15}N$-double-labelled sample at 600 MHz: 3D CBCA(CO)NH (Muhandiram and Kay, 1994), ^{13}C-HCCH-TOCSY (Kay et al., 1993), ^{13}C HSQC-NOESY (100 ms mixing time) (Majumdar and Zuiderweg, 1993) and ^{13}C HSQC-ROESY (25 ms mixing time). 2D and 3D ^{13}C-edited, ^{12}C-filtered NOESY spectra (120 ms mixing time) (Vuister et al., 1994; Folmer et al., 1995) were acquired on samples of mixed labelled (1:1 of $^{13}C/^{15}N$ and $^{12}C/^{14}N$). All spectra were processed with NMRPipe (Delaglio et al., 1995). ^{1}H chemical shifts were referenced to residual water at 4.64 ppm

relative to TSP at 40 °C. The corresponding ^{13}C and ^{15}N reference frequencies were calculated from the ^{1}H spectrometer frequency. All spectra were analysed by NMRView (Johnson and Blevins, 1994).

Structure calculations

Structures were calculated with the torsion angle dynamics program, DYANA (Güntert et al., 1997). NOEs were converted to distance constraints with the macro CALIBA and angle constraints were derived from intraresidue NOE data and coupling constants with HABAS (Güntert et al., 1991). The calculations included the disulphide bonds previously determined for capsicein: Cys-^{3} to Cys^{71}, Cys^{27} to Cys^{56} and Cys^{51} to Cys^{95} (Bouaziz, et al., 1994). Typically, 100 structures were calculated with the 20 structures showing lowest residual target function being retained for further NOE assignment. The 20 structures from the final calculation were energy minimized with OPAL (Luginbühl et al., 1996) using the default values, except a dielectric constant of 4 and 2500 steps of minimization were used. Structure analysis and all figures were performed in MOLMOL (Koradi et al., 1996) or Insight (MSI, Inc).

pH titrations

The effect of pH on chemical shift was followed in 2D ^{1}H,^{15}N HSQC spectra over the range pH 1.5 to 11.2 in approximately 0.5 pH units. In addition, the tyrosine resonances were followed in 1D and 2D ^{15}N-filtered spectra (Ikura and Bax, 1992) and 2D NOESY spectra over the range pH 2 to 11.2. Appropriate buffers (ionic strength, 45 mM) were used for each pH: for pH 1.45, 2.03 and 2.66 (potassium chloride/hydrochloric acid), pH 3.36, 3.75, 4.24, 4.76, and 5.20 (d_3-acetate), pH 6.13, 6.80, 7.54 and 7.96 (phosphate), pH 8.52 and 9.03, (boric acid/borax), pH 9.47, 9.87 and 10.44 (carbonate/bicarbonate), pH 10.8 and 11.15 (phosphoric acid/sodium hydroxide).

pKa's were determined by fitting to the Henderson-Hasselbach equation expressed in the form:

$$\delta = [\delta_{acid} + \delta_{base} 10^{(pH-pKa)}]/[1 + 10^{(pH-pKa)}] \quad (1)$$

where δ is the observed chemical shift of a resonance, δ_{acid} and δ_{base} are the chemical shifts at extreme low and high pH, respectively. The program Microsoft Excel (Solver version) was used to perform non-linear least-squares fits of the data to (Equation 1). Data from all titration curves of nuclei that sensed the ionizable group and showed a reasonable shift were fitted.

Theoretical pKa calculations

Theoretical pKa values were determined using a novel method for treating the pH dependent properties in proteins with multiple interacting titratable groups in the frame of Macroscopic Continuum model (Dimitrov and Crichton, 1997). The method is based on the approximation of electrostatic interactions between the titratable groups with the interaction of each of them with a self-consistent determined molecular field. The field may differ for different titratable groups or for identical groups placed at different spatial coordinates. The use of molecular field approximation gives the possibility of replacing the complex dependence of electrostatic interactions on the space distribution of residue charges with a sum over separated energy terms, each of which depends only on the coordinates of the corresponding ionized residue. Calculations are based on the average charges of titratable groups (Dimitrov and Crichton, 1997), the distance of separation between these groups, their intrinsic pKa's (which define the free energy of charging the ionized groups when they are disconnected from each other and separated at a distance where they cannot interact and where there is no dissolved electrolyte), on residue volumes (Zamyatnin, 1972) and the local dielectric constant (Dimitrov and Crichton, 1997). The charge portions of residue groups are taken as spheres with charges placed at their centers and defined as a center of mass between the heavy polar atoms. Dielectric properties of protein molecules are described in terms of local dielectric constants determined by the space distribution of residue volume density around each ionized residue. The volume used in calculations are obtained by subtraction of the volume of the polar part of side groups as well as the volume of amino and carbonyl groups from the whole volume of residue groups. The residue center of mass is defined between the remaining heavy atoms after deletion of the polar parts. Local density around each residue charge is calculated in the sphere with radius 8 Å with the space coordinate of nonpolar parts. The internal dielectric constant of residue groups is 4, the dielectric constant of the solvent medium is 80 and the effective radius of the charge of ionized residues is 2 Å.

Results and discussion

Characterisation of aggregation

Line widths in the initial ^{1}H NMR spectra of cryptogein suggested that the protein may have a tendency to aggregate. Pulsed field gradient NMR diffusion experiments (Dingley et al., 1995) showed that 1 mM solutions of cryptogein have diffusion coefficients of about 0.9×10^{-6} cm^{2}s^{-1} compared to that of lysozyme, 1.04×10^{-6} cm^{2}s^{-1} and therefore showed an apparent molecular weight of ~20 kDa. Acquiring spectra, either at pH values of 4, 5 or 6.8, at low or high ionic strength (0, 200 and 400 mM NaCl in 10 mM phosphate, pH 6.8) and in the presence of 5, 10, 15 or 20 mM CHAPS (pH 6.8, 45 mM phosphate) did not substantially change the diffusion coefficient. Ultracentrifugation experiments on 20 μM samples at pH 5 or 7.5 showed that the protein had apparent molecular weights of 14 kDa or 11 kDa, respectively, indicating that dimerization is both pH and concentration dependent. We concluded that under most solution conditions and at 1 mM, cryptogein is predominantly a dimer. Collection of ^{1}H,^{15}N HSQC spectra at pH 4 or 6.8, and the eventual complete assignment, showed a single set of resonances was present (Figure 1), showing that if the predominant form is a dimer then it must be symmetric. In the analysis of ^{13}C-edited, ^{12}C-filtered spectra (Figure 2) of samples of mixed label (1:1 of ^{13}C/^{15}N and ^{12}C/^{14}N labelled sample) a number of intermolecular NOEs were assigned, proving that the protein dimerizes.

Structure calculation of the dimer of cryptogein

The sequence specific assignments, intramolecular and intermolecular NOEs of cryptogein were determined using a standard combination of three-dimensional spectra acquired on three samples: a ^{15}N labelled and a ^{13}C/^{15}N labelled sample in 90% H_2O/10% ^2H_2O, and a mixed sample of labelled and unlabelled protein. The complete assignment of all observable resonances is summarized in the 2D ^{1}H,^{15}N HSQC spectrum in Figure 1.

The distance constraint data was accumulated from five spectra. The NH_i to NH_j ($j \geq i + 1$) NOE data were assigned in 3D ^{1}H,^{15}N HSQC-NOESY and HSQC-NOESY-HSQC spectra. All other interresidue NOEs, including all NH_i to CH_j ($j \geq i + 1$) were assigned in a 3D ^{13}C NOESY-HSQC recorded in 90% H_2O/10% ^2H_2O. The sequential assignment data are

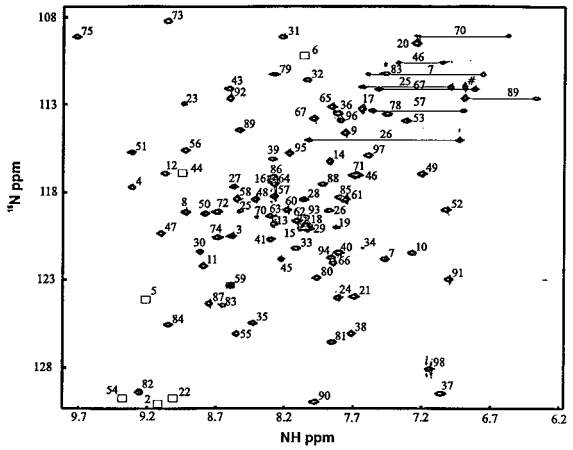

Figure 1. 2D ^{1}H, ^{15}N HSQC spectrum of ^{15}N labelled cryptogein at pH 6.8 and 40 °C. All resonances are assigned except for the peak (#) near 6.9 ppm (^{1}H) and 112 ppm (^{15}N) which would belong to the side chain amide resonances of either Gln^8 or Asn^{93}. These are the only ^{15}NH resonances not assigned. The correlations of Thr^{37}, Leu^{73}, Val^{75} and Gly^{90} are folded in this spectrum. Boxes indicate the chemical shift positions of the ^{15}NH resonances of Ala^2, Ala^5, Thr^6, Ala^{22}, Thr^{44} and Thr^{54}, which are observed at pH 4, but due to exchange with water are not observed at pH 6.8.

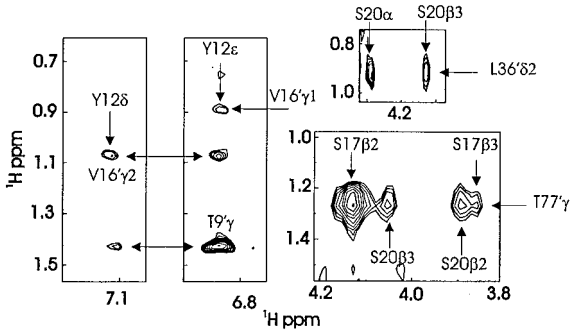

Figure 2. Sections of a 3D ^{13}C-edited, ^{12}C-flltered NOESY spectrum that show key intermolecular NOEs that define the dimer interface. The spectrum was acquired on a sample of 1 mM cryptogein in the ratio 1:1 of ^{13}C/^{15}N:^{12}C/^{14}N dissolved in 45 mM phosphate, pH 6.8, 100% ^2H_2O, at 40 °C.

summarized in Figure 3. Intraresidue NOEs were assigned in a 3D ^{13}C ROESY-HSQC recorded in 90% H_2O/10% ^2H_2O, but only ROEs that clearly indicated preferred rotamers were included. Intersubunit NOE data were assigned in either a 3D ^{13}C-separated, ^{12}C-filtered NOESY or 2D ^{13}C double-half filtered spectra (Figure 2). After trial structure calculations several NOE violations (Val^{16} to Pro^{76}) assigned to intrasubunit NOEs were clearly intermolecular and were reassigned. These peaks were not observable in the X-filtered spectra, presumably due to the lack of sensitivity of these experiments on mixed label samples.

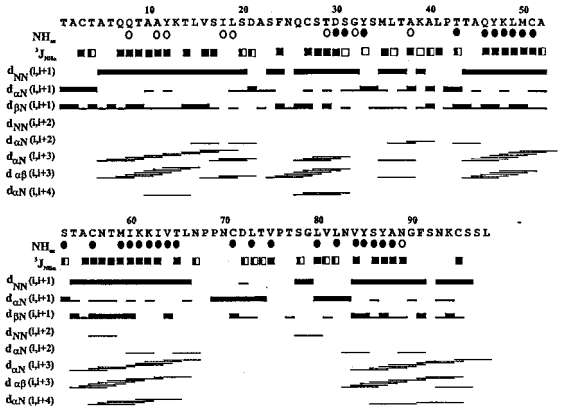

Figure 3. Summary of sequential and short range NOE data, $^3J_{HN\alpha}$ coupling constants and NH exchange of cryptogein. The thickness of the bar for the sequential NOE data indicates approximate intensity of the NOE with thick, medium or thin bars for strong, medium and weak NOEs. For short range data presence of the NOE, and not intensity, is indicated. $^3J_{HN\alpha}$ coupling constants determined in 3D HNHA experiments are classified as closed squares (<6 Hz), half-closed squares (6–8 Hz) and open squares (>8 Hz). NH exchange rates determined at pH 6.8, 25 °C are qualitatively indicated as closed circles (persisted after 3 h), open circles (present after 1 to 2 h, but no longer observed at 3 h).

The final 2134 NOEs (Table 1) were complemented with 236 angle constraints per subunit derived from sequential and intraresidue NOEs, 58 $^3J_{HNH\alpha}$ couplings (Figure 3), determined from 3D HNHA spectra, and 29 pairs of $^3J_{HNH\beta}$ couplings from 3D HNHB. The coupling constant data from the HNHA experiments were considered quantitative and were given the limits of ± 1 Hz, whereas the couplings from the HNHB experiment were treated qualitatively and given the limits of ± 2 Hz. Stereoassignments of $C\beta H_2$ were determined by considering intraresidue ROE, sequential NOE data and assignments of $C\beta H$ resonances in the 3D HNHB spectra. A total of 26 $C\beta H_2$ were unambiguously stereoassigned out of 62 spin systems. Similarly, nine pairs of Leu $C\delta H_3$ methyls and four pairs of Val $C\gamma H_3$ methyls were stereoassigned on the basis of intraresidue NOEs and ROEs (the methyl resonances of Leu^{82} could not be resolved from each other nor could the methyls of Val^{84}).

Structures were calculated using the molecular dynamics program DYANA (Güntert et al., 1997). An iterative process of calculation was employed, where successive rounds of structure calculations were used to further the NOE assignments. After the final round of calculation with DYANA the 20 structures with the lowest residual target function were energy minimized with OPAL. Structural statistics are presented before and after minimization in Table 1.

Description of the dimer solution structure of cryptogein and comparison to the X-ray crystal structure

The solution structure presented here was solved independently of the X-ray crystal structure (Boissy, et al, 1996) and the recently published solution structure (Fefeu et al., 1997). The solution structures of the homodimer show convergence to a single fold with the interface formed by the N-terminal helix (α1) and residues at the tip of the Ω-loop and β-sheet (Figure 4A). Superposition of the monomers of the solution structure with the crystal structure for the entire amino acid sequence (1 to 98) (Figure 4B) shows an rmsd of 0.94 for the backbone atoms and 1.38 for all heavy atoms, and for residues 2 to 97, an rmsd of 0.86 and 1.33 for backbone and heavy atoms, respectively. The global folds of the solution and crystal structure and analysis with PROCHECK (Laskowski et al., 1993; Laskowski et al., 1997) show that despite the solution structure forming a homodimer there are few differences observed between the two structures. Interestingly, Asn^{67}, which terminates helix-4, shows left-hand helix ϕ, ψ angles in both structures. One difference to be noted is that the crystal structure has a clear break in the C-terminal helix at residues Asn^{89} to Gly^{90} resulting in a sixth distinct helix. For the solution structure, the NH_i-NH_{i+1} and short range NOE data suggest (Figure 3) that this break is not observed, and that this helix is continuous to Ser^{97}.

A loop region that has been classified as an Ω-loop is observed between helices-2 and -3 encompassing Tyr^{33} to Pro^{42} (Figure 4). The neck of this loop is characterised by NOEs between the side chains of Tyr^{33} and Pro^{42}. Sequential NOEs show that the region Met^{35} to Ala^{38} is helical-like, and the region encompassing Ala^{38} to Leu^{41} has alternating $d_{\alpha N}(i, i+1)$ and $d_{NN}(i, i+1)$ NOEs (Figure 3), and several short range NOEs suggesting that the loop consists of several bends. The Ω-loop is positioned over the two-stranded β-sheet to form what has been termed the beak motif (Boissy et al., 1996). As observed in the crystal structure, Tyr^{33} which is at the neck of the Ω-loop, is buried in the hydrophobic core. The hydroxyl proton of this residue can be readily observed in homonuclear ^1H spectra between pH 4 and 7. The ring of Tyr^{33} is stacked against the ring of Pro^{42} in both the solution and crystal structures. The hydrogen bond acceptor

Table 1. Structural statistics for cryptogein before and after energy minimization

Residual target function (Å^2)	3.22 ± 0.14	n.a
Total Energy (kcal/mol)	1141 ± 65	−1013 ± 254
Van der Waals (kcal/mol)	713 ± 48	−769 ± 102
Distance violations [a]		
RMSD	0.031 ± 0.002	0.030 ± 0.002
Sum (Å)	20.5 ± 0.6	23.3 ± 1.8
Maximum (Å)	0.32 ± 0.01	0.10 ± 0.00
Angle violation [b]		
RMSD	1.19 ± 0.58	0.44 ± 0.02
Sum (°)	46.9 ± 6.0	66.2 ± 7.2
Maximum (°)	4.2 ± 0.7	2.6 ± 0.4
Rmsd deviations from ideality		
Bond angles (°)	n.a	1.660 ± 0.101
Bond length (Å)	n.a	0.0054 ± 0.0004
Rmsd of atom coordinates of dimer [c]		
All heavy backbone atoms (Cα,C′,N,O) (Å)	0.80	0.82
All heavy atoms (Å)	0.99	1.03
PROCHECK [d]		
Residues in most favoured region of Ramachandran Plot A, B and L (%)	87.9 ± 1.3	89.3 ± 1.5
H-bond energy std dev	1.1 ± 0.0	0.7 ± 0.0
Bad contacts/100 residues	8.5 ± 1.5	0.0

[a] The final 2134 NOEs are categorised as: 1047 intrasubunit NOEs (105 intraresidue, 321 sequential, 305 short range and 323 long range) and 40 intersubunit NOEs.
[b] 236 dihedral angles.
[c] For each monomer the rmsd of the backbone and all heavy atoms are 0.49 and 0.79 Å, respectively.
[d] PROCHECK parameters determined with PROCHECK and PROCHECK-NMR (Laskowski et al., 1993, 1996).

of the OH proton of Tyr^{33} is the carbonyl of Ala-40 in the crystal structure, which is also observed in the family of solution structures. As previously noted (Fefeu et al., 1997) the hydroxyl of Ser^{78} is observed in NMR spectra and its acceptor appears to be the carbonyl of Ala^{38}. These hydrogen bonds may contribute to the stability of the Ω-loop and the overall protein. Recently, the solution structure of P14a, a protein isolated from tomato, has been reported (Fernández et al., 1997). This protein is induced in response to pathogen invasion and displays antifungal activity. P14a shows a network of internal side chain hydroxyl hydrogen bonds and it is suggested that such bonds would each contribute about 5 kJ mol^{-1} to the free energy of folding. These hydrogen bonds would make a significant contribution to the stability of the protein, which is important for the integrity of the protein in the harsh extracellular environment. While the elicitins do not appear to be structurally related to P14a it is noteworthy that a similar pattern of side chain hydroxyl hydrogen bonds has been identified and these may have similar roles for structural stability. Another tyrosine residue, Tyr^{87}, is also buried in the protein and near to Tyr^{33}, with its hydroxyl withdrawn from the surface of the protein, however no hydrogen bond acceptor is observed.

In the family of solution structures, the ring of Tyr^{47} appears to be either stacked or edge on to the opposite face of Pro^{42} with respect to Tyr^{33}. In the crystal structure Tyr^{47} is stacked against Pro^{42}. This stacking and proximity of aromatic rings to Pro^{42} results in significant upfield shifts of the Pro resonances and should contribute to the stability of the protein. Another unusual feature of the hydrophobic core of

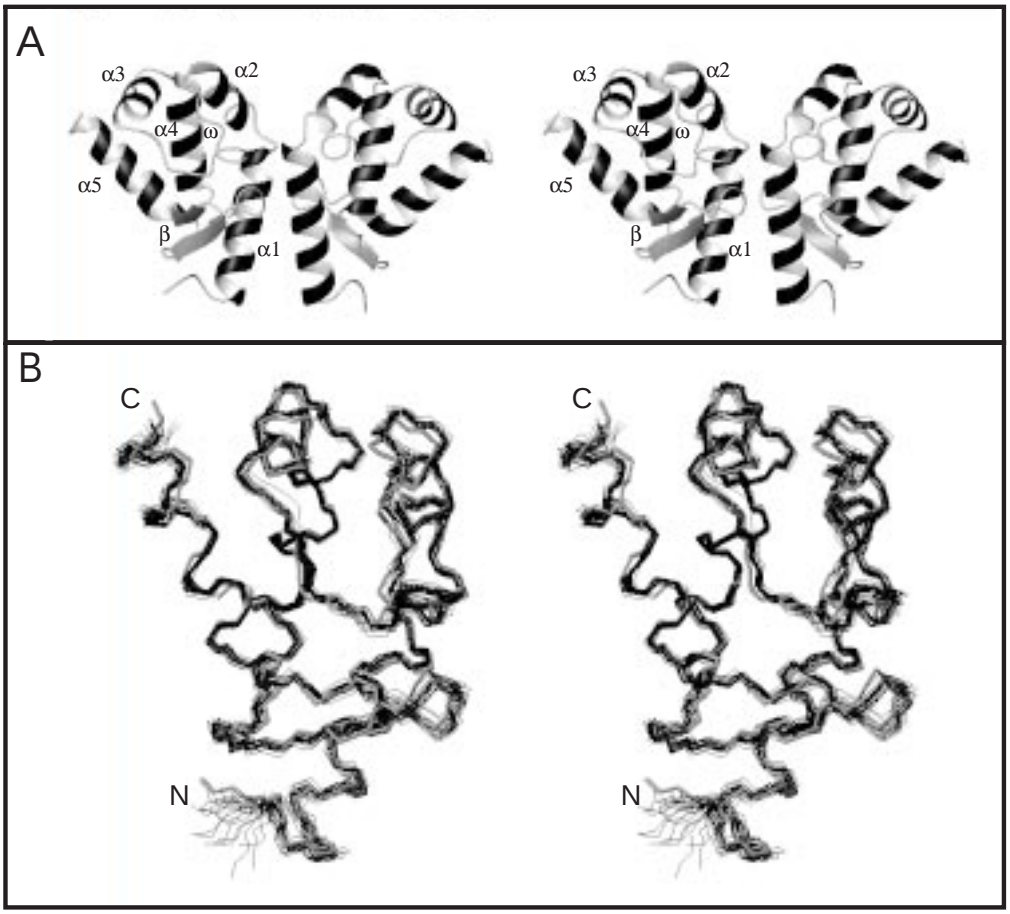

Figure 4. A. Ribbon diagram of the dimer solution structure of cryptogein. Helices span the following regions: α-1, Ala^{5} to Ser^{20}; α-2, Ala^{22} to Asp^{30}; α-3, Thr^{44} to Cys^{51}; α-4, Thr^{54} to Thr^{65}; α-5, Val^{84} to Ser^{97}. The two-stranded β-sheet encompasses Asp^{72} to Thr^{74} and Val^{81} to Asn^{83} and the Ω-loop is located between Tyr^{33} to Pro^{42}. The interface is characterized by helix α1 from one monomer running approximately parallel to the same helix of the other monomer. These helices are in contact with residues of the β-sheet and Ω-loop of the opposing monomer. B. Superposition of the first monomer of the family of 20 solution structures (thin black lines) with the X-ray crystal structure (1beo, thick grey lines) for minimal rmsd (0.86 Å for backbone atoms) for residues 2 to 97.

the protein is that the three methionines form a cluster. These residues show a substantial number of NOEs to various residues including those between Met^{59} to both Met^{35} and Met^{50}, which supports the formation of this cluster. The presence of buried tyrosine residues, the side chain hydroxyl hydrogen bonds and the clustered methionine residues may place unusual constraints on the packing and stability of the protein, and consequently the high conservation of these residues.

The N-terminal helix (helix-1) is associated with the edge of the sheet and Ω-loop, whereas the other helices form a globular unit. We have qualitatively used amide exchange rates at pH 6.8 and 25 °C to analyse the stability of the overall protein, in particular helix-1, and to determine if Gly^{90} breaks the C-terminal helix into two independent helices. Analysis of amide exchange rates (Figure 3) shows that the NH protons of helix-1, and those of the region that encompass helix-6 of the crystal structure (or the Gly^{90} to Ser^{96} of helix-5 of the solution structure), exchange markedly faster than the NH protons of the other secondary structure elements of cryptogein. Consequently, the N-terminal helix and the region Gly^{90} to Ser^{96} are less stable compared to the other helices and β-sheet of cryptogein. The Ω-loop does not appear to be stabilized by any main chain hydrogen bonds whereas the NH hydrogens internal to the β-sheet show the expected slow exchange.

Measurement of pKa's

The crystal structure shows two salt bridges: Asp^{21} to Lys^{62}, and Asp^{72} to the N-terminus, which may be structurally and/or functionally significant. Visual inspection of the solution structures suggests that the former bridge may be present, but the disorder of the N-terminus makes it difficult to support the presence of the latter (Figure 4). Several NOEs were assigned between the N-terminal region and Asp^{72}, including the $C\beta H_2$ of Asp^{72} to the $C\beta H_3$ of Ala^2, which supports an interaction between the N-terminus and Asp^{72}. Amongst the elicitins, the N-terminal residue and position 72 are homologously conserved, where the N-terminal is either Ala or Thr and position 72 is either Asp or Glu. Position 21 and 62, however, are not conserved, and although these residues are frequently Asp and Lys, respectively, position 21 may also be Glu, Lys or Thr and position 62 may also be Asn, Glu or Thr. In cryptogein, the salt bridge between Asp^{21} and Lys^{62} would present a stabilizing interaction between helices-2 and -4. These salt bridges and the ionization states of other residues are important elements of the surface of the protein which may play significant roles in receptor recognition.

The pKa values of the N- and C-termini, Asp^{21}, Asp^{30}, Asp^{72}, and Tyr^{85} were determined by collecting a series of 2D ^1H,^{15}N HSQC spectra at approximately half pH units over the pH range 1.5 to 10.8, at the same ionic strength, and fitting the chemical shifts to the Henderson-Hasselbach equation (Equation 1). Collecting 1D ^{15}N-filtered and 2D ^{15}N-filtered NOESY spectra over the pH range 2 to 11.2, the effect of pH on Tyr^{12}, Tyr^{33}, Tyr^{47}, Tyr^{85} and Tyr^{87} could also be observed. Representative plots of nuclei that show significant shifts ($\Delta\delta > 0.1$ ppm for ^{15}N, $\Delta\delta > 0.05$ ppm for ^1H) with pH are plotted in Figure 5 and all experimental pKa's are described in Table 2. The ionizable residues in cryptogein are shown in Figure 6. While many resonances showed pH shifts, several could not be unambiguously assigned to a particular ionizable group. For example, Cys^{56}, Asn^{57} and Met^{59} all show apparent pKa's of ~2.7, but these residues are approximately 7 to 11 Å from either Asp^{21} or Asp^{30} and may reflect the ionization of either or both residues. The pKa estimation for an ionizable group included apparent pKa's from residues that were near only one ionizable group (less than approximately 8 Å), but distant from other similar ionizable groups (greater than approximately 18 Å). These limits reflect the distances between the three Asp residues: Asp^{72} is 20 Å from both Asp^{21} and Asp^{30}, and Asp^{21} is 16 Å from Asp^{30}.

Table 2. Experimental and calculated pKa values for cryptogein

Residue	Exp pKa[a]	X-ray[b]	DYANA[c]	OPAL[c]
		Calculated pKa		
N-terminal, Thr^1	7.43	8.05	7.87	7.90
Tyr^{12}	~11.5	10.47	11.75	11.82
Lys^{13}		10.93	10.94	10.90
Asp^{21}	2.49±0.05	2.83	3.31	3.18
Asp^{30}	2.51±0.09	3.90	3.86	3.88
Tyr^{33}	>12	11.68	11.27	11.25
Lys^{39}		10.91	10.94	10.93
Tyr^{47}	>12	12.38	12.24	12.34
Lys^{48}		10.88	11.02	11.10
Lys^{61}	10.1[d]	10.81	10.87	10.87
Lys^{62}		12.05	11.59	11.67
Asp^{72}	2.61±0.15	3.25	3.54	3.50
Tyr^{85}	10.35±0.15	10.40	10.41	10.23
Tyr^{87}	>12	12.39	12.40	11.92
Lys^{94}	9.4[d]	10.85	11.09	11.44
C-terminal, Leu^{98}	3.51±0.01	3.31	3.33	3.20

[a] Experimental pKa values determined at 40 °C. Mean and standard deviations are given for pKa's determined by following the pH dependence of a number of chemical shifts ($\Delta\delta > 0.1$ ppm for ^{15}N; $\Delta\delta > 0.05$ ppm ^1H): N-terminus (peptide ^{15}N of D72); Asp^{21} (peptide ^{15}N of S20, D21, S23, T65 and peptide NH of A22, S23); Asp^{30} (peptide ^{15}N of D30, S31, G32); Lys^{61} (peptide ^{15}N of I63); Asp^{72} (peptide ^{15}N of D72, T74, peptide NH of V84 and side chain $^{15}N\delta^2$ of N70); Tyr^{85} (peptide ^{15}N of Y87, N89 and ring Hϵ of Y85); Lys^{94} (peptide ^{15}N of T43); C-terminus (peptide ^{15}N of S96, S97, L98).

[b] pKa's calculated for the X-ray crystal structure (1beo).

[c] pKa calculated for the 20 lowest energy conformers determined with DYANA and after energy minimization with OPAL using a dielectric constant of 4.

[d] pKa could not be assigned unambiguously, for Lys^{61} this pKa may be assignable to Lys^{62} and for Lys^{94} this pKa may be assigned to Lys^{48}.

pKa's of C-terminus and Asp residues

The C-terminus shows a typical pKa of 3.5 (Figure 5A) indicating that it does not interact with any residues, which is in agreement with the effects of its ionization being limited to the region Ser^{96} to Leu^{98}. In contrast, all three Asp show low pKa's 2.5 to 2.6 (Figure 5A), suggesting that they interact with positive charges. These data support the presence of the two salt bridges observed in the crystal structure: Asp^{21} to Lys^{62} and Asp^{72} to the N-terminus. Indeed, the titration of Asp^{21} is reflected in the chemical shift of the ^{15}NH resonance of Thr^{65} (Figure 5B). Asp^{30} is

not near any positively charged residue. However, this residue which is at the C-terminal end of the short helix-2, is located near the N-terminal end of helix-4 and its position suggests that it may be aligned with the dipole of helix-4, thus effectively capping that helix. The titration of Asp^{30} is clearly observed on the chemical shifts of the ^{15}NH resonances of Thr^{54} and Ala^{55} which are the first residues of helix-4.

pKa's of N-terminus and Lys residues

The pKa of the N-terminus was determined by following the peptide ^{15}N resonance of Asp^{72} (Figure 5A) and is estimated to be 7.3, similar to a pKa expected for a free N-terminus. This pKa is incompatible with the low pKa (2.6) of the carboxylate of Asp^{72}, which supported the presence of a salt bridge between the carboxylate of Asp^{72} and the N-terminus, and suggests that the pKa of Asp^{72} is due to other interactions.

Several other shifts are observed in the alkaline pH range, which may be attributable to lysine residues (Figure 5B). The ^{15}N peptide resonance of Thr^{43} shows a small but significant pH dependence ($\Delta\delta$ = 0.12 ppm) with an apparent pKa of 9.4 and as Thr^{43} is near Lys^{94} (~8 Å) we assign the pKa to this residue. However, Lys^{48} is also relatively close to Thr^{43} (10 Å) and may have an effect and thus we can not assign this pKa unambiguously. The ^{15}N peptide resonance of Ile^{63} has an apparent pKa of 10.1. As this residue is near both Lys^{61} and Lys^{62} this pKa also can not be assigned unambiguously. The ^{15}N peptide resonance of Thr^{65} shows three pKa inflections, the first pKa is 2.4, and is assigned to Asp^{21} whose carboxylate forms a salt bridge with Lys^{62}. The next two inflections do not give reasonable fits, but are most likely due to either or both Lys^{61} and Lys^{62}.

pKa's of Tyr residues

The ring resonances of the tyrosine residues are substantially unaffected between pH 2 and 8.5. The CϵH resonances of Tyr^{85} show a large chemical shift dependence ($\Delta\delta$ = 0.26 ppm) which fits to a pKa 10.2. The pKa of Tyr^{85} appears to be reflected in the peptide ^{15}N resonances of Tyr^{87} and Asn^{89} with apparent pKa's of 10.5 and 10.4, respectively. The CϵH resonances of Tyr^{12} show a small pH dependence between pH 7 and 11 ($\Delta\delta$ = 0.09 ppm) that does not appear to fit to a single ionization. The difficulty with this titration is that it may reflect the ionization of both Tyr^{12} and Lys^{13} in both the monomeric and dimeric species (Figure 6). Several ^{15}N peptide resonances of

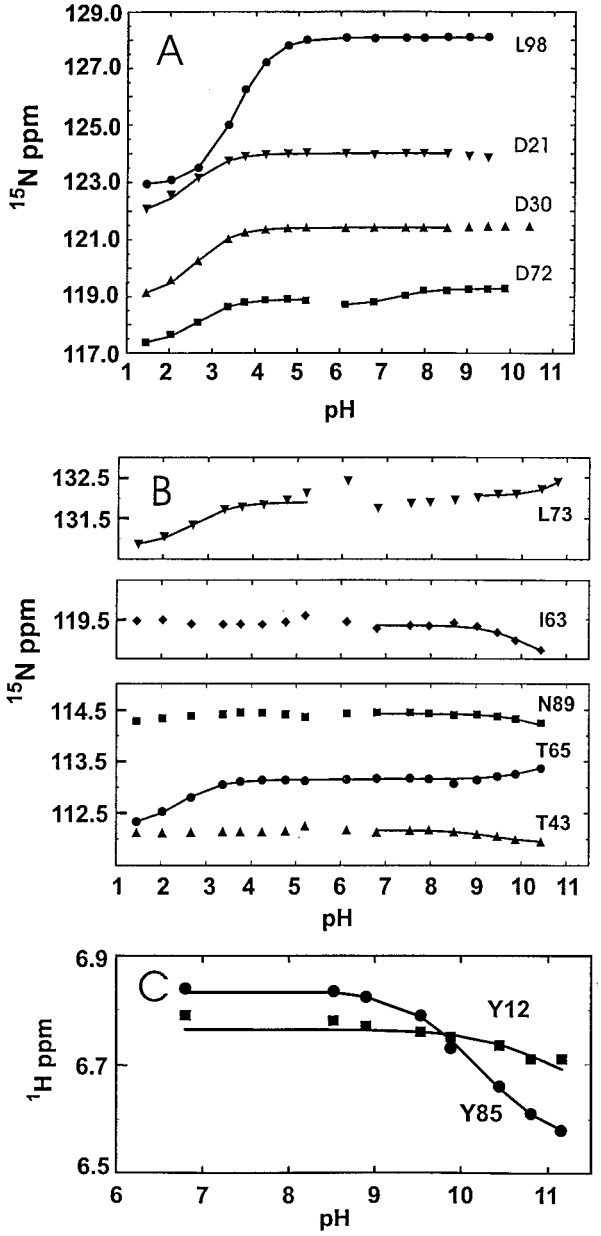

Figure 5. pH dependence of the chemical shifts of (A) the ^{15}N resonances of Asp^{21}, Asp^{30}, Asp^{72} and Leu^{98}; (B) the ^{15}N resonances of Thr^{43}, Ile^{63}, Thr^{65}, Leu^{73} and Asn^{89}; and (C) the CϵH resonances of Tyr^{12} and Tyr^{85}. Data for (A) and (B) were from 2D ^{1}H, ^{15}N HSQC spectra and for (C) from 1D and 2D ^{15}N-filtered NOESY and 2D ^{1}H, ^{1}H NOESY spectra. Curves are shown for only the portion of the data fitted to obtain single pKa's.

residues located in the β-sheet and near Tyr^{12} show pH dependence in both the acid and the alkaline regions. The ^{15}N resonance of Leu^{73} shows three pH inflections (Figure 5B). The first ($\Delta\delta = 1.09$ ppm) is readily fitted to a carboxylate ionization with a pKa of 2.7 and is thus assigned to Asp^{72}. The second inflection occurs between pH 5 and 8 ($\Delta\delta = 0.24$ ppm), and while it was not possible to fit this inflection it must be due to the N-terminus. The third inflection occurs above pH 9 ($\Delta\delta = 0.5$ ppm). If this latter shift is fitted, an estimate of an apparent pKa of ∼11.5 (with reasonable fits ranging from pKa of 11 to 12) is observed and thus we assign this pKa to Tyr^{12}. The CϵH resonances of Tyr^{33} show no shifts at all, while those of Tyr^{47} and Tyr^{87}, show small shifts above pH 9.0 ($\Delta\delta = 0.02$ and 0.03 ppm, respectively), suggesting that the pKa's of these three tyrosines are greater than 12.

The titration behaviour of all the tyrosine residues mostly agrees with the extent of their solvent exposure and their interactions. Tyr^{85} is completely exposed to the solvent giving a typical pKa, and Tyr^{12} is mostly exposed, but lies flat against the protein in both the solution and crystal state. Tyr^{33} and Tyr^{87} are all partially or completely buried. Tyr^{33} is buried with its OH hydrogen bonded. Tyr^{87} is also predominantly buried with its hydroxyl withdrawn from the solvent, however, we have not detected this hydroxyl hydrogen in any of our spectra. The hydroxyl of Tyr^{47} is exposed to the solvent, but is stacked against Pro^{42} and near to Phe^{91}, and therefore is in a hydrophobic environment. Further, this residue is near Tyr^{87} and thus the deprotonation of these two residues may affect each other. All experimentally determined pKa's of the tyrosines, except the pKa of Tyr^{87}, are in reasonable agreement with those determined by UV difference spectroscopy (Nespoulous et al., 1994). Tyr^{87} was found to have a typical pKa (10.5) by this latter method.

Calculations of the pKa's based on the crystal and solution structures

The theoretical pKa's were determined by using the crystal structure (Boissy et al., 1996), the unrefined and refined solution structures that are described above. In Table 2 the calculated pKa's of the individual ionizable groups are reported along with the experimental pKa's. In most cases, the experimental and calculated values agree well, and the data for the crystal structure is similar to both the refined and unrefined solution structures. The five tyrosine residues have values that in good agreement and consistent with their interactions and/or their exposure to the solvent as discussed above. Tyr^{12} is on the protein surface, but lies flat against its own monomer. This residue is at the interface of the dimer, and is near the Lys of the opposing dimer (Figure 6). However, the pKa of Lys^{13} is calculated to have a typical value and thus we do not expect an interaction between these residues. The calculated pKa of Tyr^{12} for the solution structure, but not the crystal structure, is in agreement with the experimental pKa, suggesting that this difference is due to dimerization.

The calculated pKa for Lys^{94} is markedly lower to the experimental pKa of 9.4, however, the assignment of this pKa is ambiguous (Table 2). Recent calculations of the pKa's of Lys^{94} and Tyr^{47} suggest that these residues interact to raise the pKa of Lys^{94} to a value of 13 and lower the pKa of Tyr^{47} to a value of 9 (Vogel and Juffer, 1998). This interaction is not apparent in our calculations or in the experimental data for Tyr^{47}. The remaining Lys residues have typical pKa's (∼11) except for Lys^{62} which is expected to participate in a salt bridge with Asp^{21}. The experimental and/or theoretical pKa's for Asp^{21} and Lys^{62} agree with the presence of the salt bridge (Table 2). The salt bridge between Asp^{72} and the N-terminus is less apparent with both a normal experimental and theoretical pKa for the N-terminus. The low experimental pKa for Asp^{72} may reflect other interactions that remain to be determined. These conclusions are supported by the calculations of the pKa's of Asp^{21} and Asp^{72} reported by others (Vogel and Juffer, 1998) which are in better agreement with our experimental data. The theoretical pKa of Asp^{30} is typical for an Asp residue in small peptides, which is in contrast to the experimental pKa. The method (Dimitrov and Crichton, 1997) that we have used here does not account for aligned partial charges that would occur in helix dipoles. Calculations by others using a different approach (Juffer et al., 1997; Vogel and Juffer, 1998) also show that Asp^{30} has a typical pKa, and these methods do include partial charges (Juffer, personal commun.) and should therefore consider interactions with helix dipoles. While it is clear from the experimental data that this residue is not interacting with other ionizable residues, we can not conclude on the actual nature or importance of the interactions that lower the pKa of Asp^{30}.

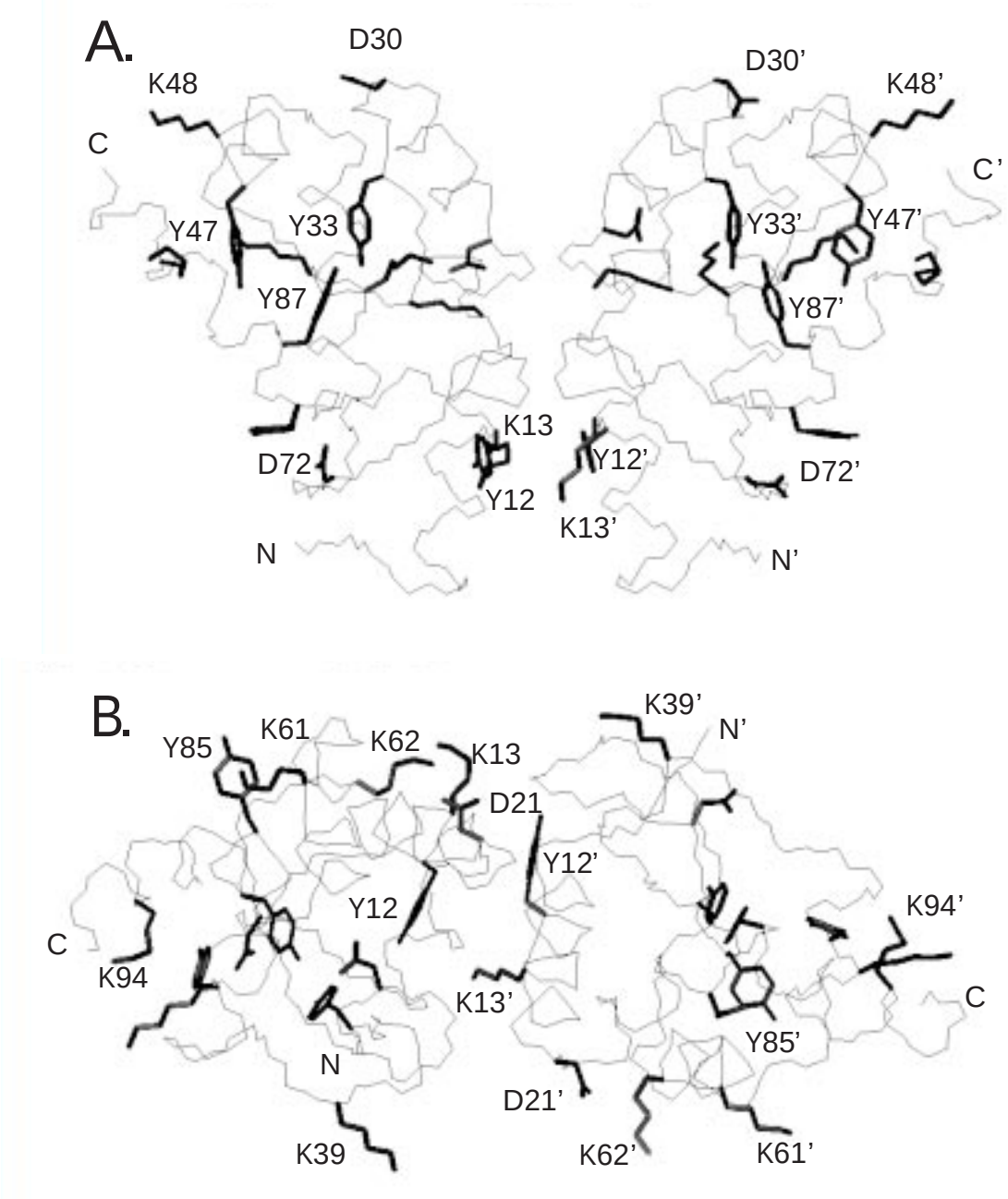

Figure 6. The dimer solution structure highlighting the ionizable residues. A. The dimer is oriented similarly to Figure 4A. B. The dimer is rotated approximately 90° around the x-axis compared to (A). Tyr^{12} and Lys^{13} are the only ionizable residues at the interface of the dimer. Tyr^{12} lies parallel to the protein surface of the monomer and the side chain of Lys^{13} is well exposed to the solvent. The pKa's of these residues do not appear affected by dimerization.

Functionally important residues

The >70% identity of the sequences of the elicitins, and the absence of a characterized receptor reduces the conclusions that can be made with respect to which are functionally important residues. One would expect that the evolutionary pressure to conserve residues amongst the elicitins would be due to the unknown function of the elicitins within the genera

Phytophthora and *Pythium*, and not the advantageous induction of defence responses displayed within the resistant plant. In this report, a number of potentially important structural/dynamic features have been raised: firstly, the amide exchange rates of helix-1, that include the functionally significant Lys^{13}, highlights the relative instability of the first helix compared to the other helices and the β-sheet; secondly, the three Asp residues show low pKa's and may be involved in both salt bridges and other interactions that stabilize protein structure; thirdly, four of the five Tyr residues show shifted pKa's and are withdrawn completely or to some extent from the solvent; and finally the protein shows a tendency to dimerize at concentrations above 20 μM. While it is tempting to suggest that dimerization may be important for inducing a defence response, this activity is observed at concentrations <1 nM, where dimerization is less likely. Interestingly, the regions that appear to interact, ∼20s, ∼30s and the ∼70s, are on the same side of the protein as is the functionally important residue 13. Whether these residues at the dimer interface are important in protein-protein interactions in the plant defence response can only be proved through site-directed mutagenesis studies and the isolation of the receptor.

Acknowledgements

This work benefited from the use of NMR facilities at the University of Melbourne and the Australian National University and was funded by ARC grants S09711445 and A09801407, and a joint Institute of Advanced Studies/Australian Universities Collaborative Research Grant. We thank Ms. Sieu Cleland for preparing cultures, Dr. Geoff Howlett for the ultracentrifugation data and Prof. Hans Vogel and Dr. André Juffer for providing their manuscript prior to publication.

References

Archer, S.J., Ikura, M., Torchia, D.A. and Bax, A. (1991) *J. Magn. Reson.*, **95**, 636–641.

Boissy, G., de la Fortelle, E., Kahn, R., Huet, J.-C., Bricogne, G., Pernollet, J.-C. and Brunie, S. (1996) *Structure*, **4**, 1429–1439.

Bonnet, P., Bourdon, E., Ponchet, M., Blein, J.P. and Ricci, P. (1996) *Eur. J. Plant Pathol.*, **102**, 181–192.

Bouaziz, S., Van Heijenoort, C., Guittet, E., Huet, J.-C. and Pernollet, J.-C. (1994) *Eur. J. Biochem.*, **33**, 8188–8197.

De Wit, P.J.G.M. (1992) *Annu. Rev. Phytopathol.*, **30**, 391–418.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Dimitrov, R.A. and Chrichton, R.R. (1997) *Proteins*, **27**, 576–596.

Dingley, A.J., MacKay, J.P., Chapman, B.E., Morris, M.B., Kuchel, P.W., Hambly, B.D. and King, G.F. (1995) *J. Biomol. NMR*, **6**, 321–328.

Fefeu, S., Bouaziz, S., Huet, J.-C., Pernollet, J.-C. and Guittet, E. (1997) *Protein Sci.*, **6**, 2279–2284.

Fenn, M.E. and Coffey, M.D. (1984) *Phytopathology*, **74**, 606–611

Fernández, C., Szyperski, T., Bruyère, T., Ramage, P., Mösinger, E. and Wüthrich, K. (1997) *J. Mol. Biol.*, **266**, 576–593.

Folmer, R.H.A., Hilbers, C.W., Konings, R.N.H. and Hallenga, K. (1995) *J. Biomol. NMR*, **5**, 427–432.

Forman-Kay, J.D., Clore, G.M. and Gronenborn, A.M. (1992) *Biochemistry*, **31**, 3442–3452.

Frenkiel, T., Bauer, C., Carr, M.D., Birdsall, B. and Feeney, J. (1990) *J. Magn. Reson.*, **90**, 420–425.

Gayler, K.R., Popa, K.M., Maksel, D.M., Ebert, D.L. and Grant, B.R. (1997) *Mol. Plant Pathol. On-Line* http://www.bspp.org.uk/mppol/1997/0623gayler.

Grant, B.R., Ebert, D. and Gayler, K.R. (1996) *Aust. Plant Pathol.*, **25**, 148–157.

Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Huet, J.C., La Cear, J.P., Nespoulous, C. and Pernollet, J.C. (1995) *Mol. Plant-Microbe Interactions*, **8**, 302–310.

Ikura, M. and Bax, A. (1992) *J. Am.Chem. Soc.*, **114**, 2433–2440.

Johnson, B.A. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.

Juffer, A.H., Argos, P. and Vogel, H.J. (1997) *J. Phys. Chem. B*, **101**, 7664–7673.

Jones, D.A., Thomas, C.M., Hommon D'Kosack, K.E., Balint-Kurti, P.J. and Jones, J.D.G. (1994) *Science*, **266**, 789–793.

Kamoun, S., Young, M., Glascock, C.B. and Tyler, B.M. (1993) *Mol. Plant-Microbe Interactions*, **6**, 15–25

Kay, L.E., Xu, G.-Y., Singer, A.U., Muhandiram, D.R. and Forman-Kay, J.D. (1993) *J. Magn. Reson.*, **101**, 333–337.

Kooman-Gersmann, M., Honée, G., Bonnema, G. and De Wit, P.J.G.M. (1996) *Plant Cell*, **8**, 929–938.

Koradi, R., Billeter, M. and Wüthrich, K. (1996) *J. Mol. Graph*, **14**, 51–55.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.

Laskowski, R.A., Rullman, J.A.C., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) *J. Biomol. NMR*, **8**, 477–486.

Luginbühl, P., Güntert, P., Billeter, M. and Wüthrich, K. (1996) *J. Biomol. NMR*, **8**, 136–146.

Majumdar, A. and Zuiderweg, E.R.P. (1993) *J. Magn. Reson.*, **102**, 242–244.

Muhandiram, D.R. and Kay, L.E. (1994) *J. Magn. Reson.*, **103**, 203–216.

Nakamura, H. (1996) *Q. Rev. Biophys.*, **29**, 1–90.

Nespoulous, C. and Pernollet, J.-C. (1994) *Int. J. Pept. Protein Res.*, **43**, 154–159.

O'Donohue, M.J., Gousseau, H., Huet, J.-C., Tepfer, D. and Pernollet, J.-C. (1995) *Plant Mol. Biol.*, **27**, 577–586.

Vogel, H.J. and Juffer, A.H. (1998) *Theor. Chem. Acc.* (in press).

Vuister, G.W. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 7772–7777.

Vuister, G.W., Kim, S.-J., Wu, C. and Bax, A. (1994) *J. Am. Chem. Soc.*, **116**, 9206–9210.

Zamyatnin, A.A. (1972) *Prog. Biophys. Mol.*, **24**, 803–813.

Zhang, O., Kay, L.E., Olivier, J.P. and Forman-Kay, J.D. (1994) *J. Biomol. NMR*, **4**, 845–858.