Journal of Biomolecular NMR **30:** 81–96, 2004.

Characterisation of a mobile protein-binding epitope in the translocation domain of colicin E9

Colin J. Macdonald[1,*], Kaeko Tozawa[1,2,*], Emily S. Collins[1], Christopher N. Penfold[2], Richard James[2], Colin Kleanthous[3], Nigel J. Clayden[1,**] & Geoffrey R. Moore[1,**]

[1]*School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich NR4 7TJ, U.K.;* [2]*Division of Microbiology and Infectious Diseases, School of Molecular Medical Sciences, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, U.K.;* [3]*Department of Biology (Area 10), University of York, York YO10 5YW, U.K.*

Received 25 May 2004; Accepted 27 May 2004

Key words: backbone dynamics, colicin E9, ^{15}N relaxation, TolB, translocation

Abstract

The 61 kDa colicin E9 protein toxin enters the cytoplasm of susceptible cells by interacting with outer membrane and periplasmic helper proteins, and kills them by hydrolysing their DNA. The membrane translocation function is located in the N-terminal domain of the colicin, with a key signal sequence being a pentapeptide region that governs the interaction with the helper protein TolB (the TolB box). Previous NMR studies (Collins et al., 2002 *J. Mol. Biol.* **318**, 787-804) have shown that the N-terminal 83 residues of colicin E9, which includes the TolB box, is largely unstructured and highly flexible. In order to further define the properties of this region we have studied a fusion protein containing residues 1-61 of colicin E9 connected to the N-terminus of the E9 DNase by an eight-residue linking sequence. 53 of the expected 58 backbone NH resonances for the first 61 residues and all of the expected 7 backbone NH resonances of the linking sequence were assigned with 3D ^{1}H-^{13}C-^{15}N NMR experiments, and the backbone dynamics of these regions investigated through measurement of ^{1}H-^{15}N relaxation properties. Reduced spectral density mapping, extended Lipari-Szabo modelling, and fitting backbone R_2 relaxation rates to a polymer dynamics model identifies three clusters of interacting residues, each containing a tryptophan. Each of these clusters is perturbed by TolB binding to the intact colicin, showing that the significant region for TolB binding extends beyond the recognized five amino acids of the TolB box and demonstrating that the binding epitope for TolB involves a considerable degree of order within an otherwise disordered and flexible domain.

Abbreviations: Im9, the immunity protein for colicin E9; E9 DNase, the endonuclease domain of colicin E9; HSQC, heteronuclear single quantum coherence; ppm, parts per million; DSS, 2,2-(dimethylsilyl)propanesulfonic acid; TSP, sodium 3-trimethylsilypropionate; T_{1-61}-DNase fusion protein, residues 1-61 of colicin E9 connected to the N-terminus of the E9 DNase by an eight residue thrombin cleavage sequence.

Colicins are a heterogeneous group of anti-bacterial proteins that are produced by immune strains of *Escherichia coli* to kill competing bacteria (James et al., 1996, 2002). Cell killing normally occurs in three steps: passage of the colicin across the outer membrane of the target cell following binding of the colicin to one or more outer membrane proteins; translocation of the colicin across the periplasm of the susceptible cell; and subsequent cell killing by enzymatic cleavage of nucleic acids in the cytoplasm or pore-formation in the inner membrane (Cramer and Stauffacher, 1995;

*These workers contributed equally to the study.
**To whom correspondence should be addressed. E-mails: g.moore@uea.ac.uk and n.clayden@uea.ac.uk

Cao and Klebba, 2002; James et al., 2002; Zakharov and Cramer, 2002). They are classified into groups on the basis of the outer membrane cell surface receptor of the target cells to which they bind: E-type colicins bind to BtuB (Di Masi et al., 1973), an essential component of the high-affinity transport system for vitamin B_{12} in *E. coli* (Taylor et al., 1998). In common with most colicins, the E-type colicins consist of three functional domains; the killing activity is contained in the C-terminal domain, while the central section contains the receptor-binding domain, and the N-terminal region is responsible for translocation of the cytotoxic domain into the target cell (Masaki et al., 1992; Bénédetti et al., 1992; Pilsl and Braun, 1995; Garinot-Schneider et al., 1997; James et al., 2002). The E-type colicins fall into one of three cytotoxic classes: the pore-forming colicin E1, DNases such as colicins E2, E7, E8 and E9, and RNases such as colicins E3, E4, E5 and E6 (James et al., 2002). Despite the difference in killing action, these colicins all share a common translocation mechanism, interacting with TolB through possession of an N-terminal binding epitope known as the TolB box. This is a sequential 5-residue epitope that is essential for interaction with the periplasmic TolB, and for toxicity of the colicin (Pilsl and Braun, 1995; Garinot-Schneider et al., 1997; Bouveret et al., 1998; Carr et al., 2000; James et al., 2002). Many other types of colicin exist that are distinguished from the E-type colicins by the outer-membrane receptors they recognise and by the translocation pathway that mediates their toxicity. Despite these differences between different types of colicins, there appears to be a common theme in their mechanisms of action in that they seem to require a dynamic translocation domain to enable their killing action to be realised (Vetter et al., 1998). The 42 kDa pore-forming colicin N is a good example of this. Its TolA-binding translocation domain was not detected by X-ray crystallography (Vetter et al., 1998) and in the absence of TolA, CD, fluorescence and NMR spectroscopies indicate it is intrinsically disordered (Raggett et al., 1998; Gokce et al., 2000; Anderluh et al., 2003). Similarly, the N-terminal 83 residues of colicin E3 that includes the TolB box was not detected by X-ray crystallography (Soelaiman et al., 2001) and the homologous region of colicin E9 was shown by NMR to be disordered in the absence of TolB (Collins et al., 2002). NMR studies of the translocation domain of the 61 kDa colicin E9 were complicated by the presence of a glycine-rich region from residues 62 to 83, which contains 11 glycines, and to further characterise the properties of the TolB box region by NMR we have removed this glycine-rich region from a colicin E9 translocation domain construct. This construct also lacks residues 81-448 of the intact colicin (Figure 1), so that the final T_{1-61}-DNase fusion protein contains residues 1-61 of colicin E9 connected to the N-terminus of the E9 DNase by an eight-residue thrombin cleavage sequence. The simplification of the NMR spectra of the N-terminal region containing the TolB box provided by this construct has allowed the backbone dynamics of the TolB binding epitope to be explored in detail, and this indicates that clusters of interacting residues distributed throughout the N-terminal 61 residues are involved with TolB binding.

Materials and methods

Sample preparation. Uniformly ^{15}N-labelled and ^{13}C/^{15}N-labelled wild-type T_{1-61}-DNase fusion protein was obtained by growing *E. coli* ER2566 (*F^- lamda$^-$ fhuA2 [lon] ompT lacZ:: T7 gene1 gal sulA11* D*(mcrC-mrr)114::IS10 R(mcr-73::miniTn10– TetS)2 R(zgb-210::Tn10) (TetS) endA1 [dcm]*) (New England Biolabs Inc.) cells containing plasmid pNP330 (encoding wild-type T_{1-61}-DNase fusion protein and Im9, with a histidine tag attached to Im9) in minimal medium containing $^{15}NH_4Cl$ (1 g/l) and $^{13}C_6$-glucose (4 g/l), respectively. Purification and isolation of the T_{1-61}-DNase fusion protein from the His-tagged immunity protein was performed as previously described for preparation of colicin E9 and E9 DNase (Wallis et al., 1994) with the purified protein dialysed against EDTA and then buffer to remove any bound Ni^{2+} prior to concentration for NMR studies (Hannan et al., 2000).

NMR spectroscopy. All NMR samples contained 50 mM phosphate buffer in 90% H_2O/10% D_2O, pH 6.3, and 0.1% sodium azide. Protein concentrations for NMR measurements were 1.0–1.5 mM. All NMR spectra were acquired at 288 K with Varian Unity Inova 500 or 600 spectrometers equipped with triple resonance pulsed field gradient probes, operating at ^{1}H frequencies of 499.865 MHz and 599.162 MHz, and ^{15}N frequencies of 50.66 MHz and 60.72 MHz, respectively, using pulse sequences incorporated into the Varian (CA, U.S.A.) 'Protein Pack' suite of experiments. Resonance assignments were obtained from the well established triple resonance experiments HNCO, HNCA, CBCA(CO)NH, HN-

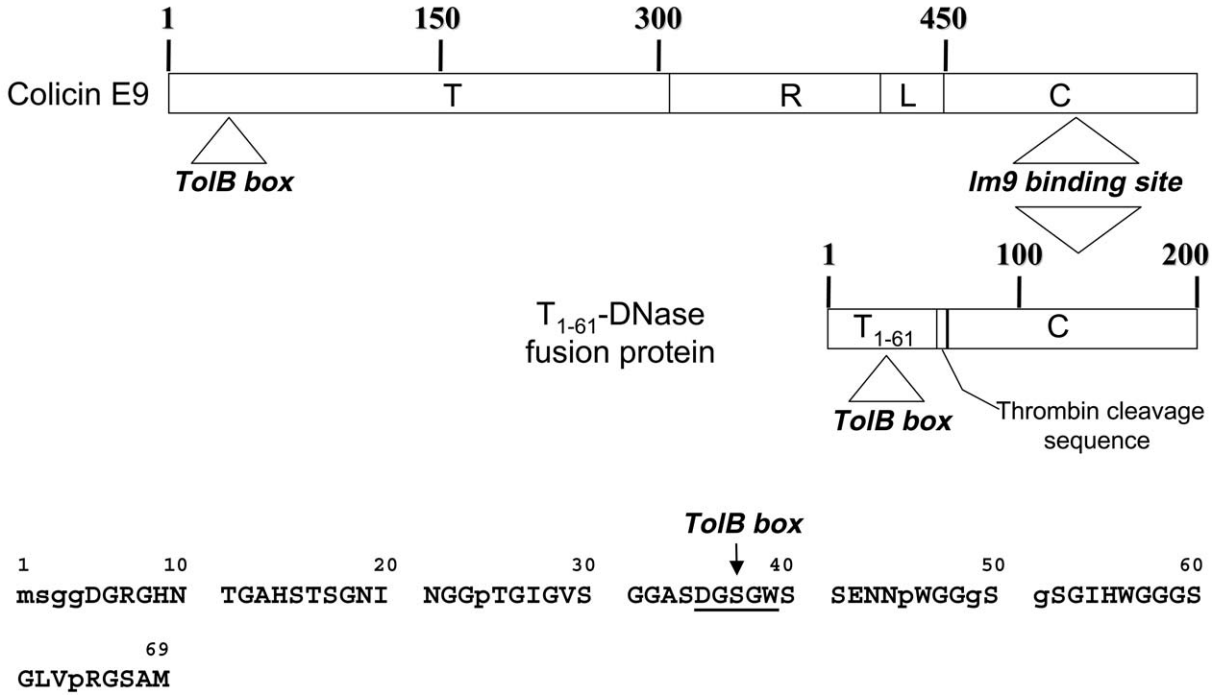

Figure 1. (Upper panel) Schematic diagram of intact colicin and the T_{1-61}-DNase fusion protein. T indicates the translocation-domain, R the receptor binding domain, L a short linking sequence between the R and C domains, and C the cytotoxic DNase domain. T_{1-61} is the first 61 residues of the translocation domain. The TolB box (see text) and the Im9 binding site on the colicin and fusion protein are indicated. (Lower panel) Primary sequence of the N-terminal 69 residues of the T_{1-61}-DNase fusion protein. Residues from the T_{1-61} region and the 8-residue thrombin cleavage sequence whose ^{1}H-^{15}N NH resonances have been assigned are indicated by uppercase lettering. The position of the TolB box is also indicated.

COCA, HNCACB, (HCA)CO(CA)NH, HCCONH, and C(CO)NH. However, owing to the unfolded nature of the T_{1-61}-DNase fusion protein and the sequences of repeating residues within the polypeptide chain, resonance overlap was a particular problem. To help overcome this problem the HNN experiment (Panchal et al., 2001) proved invaluable. The HNN spectrum simultaneously displays forward and backward ^{1}HN and ^{15}N correlations and utilises ^{15}N chemical shift dispersion in both indirect dimensions, which is generally larger than ^{13}C for unfolded proteins. Also, as the phases of the diagonal and cross peaks for any given residue are determined by the character of the neighbouring residues, additional amino acid specific information is also available. Parameters for triple-resonance experiments are given in Table 1 of the supplementary material. One-dimensional data were processed using Varian VNMR software and FELIX 95.0 (Biosym/MSI, California), and multidimensional data were processed using NMRPipe (Delaglio et al., 1995). Prior to Fourier transformation, a cosine-bell window function was applied to each dimension for apodization. The indirect dimensions were first linear predicted to double the number of data points and then zero-filled to round up the number of data points to the nearest power of 2. Proton chemical shifts were referenced to external DSS, the carbon chemical shifts to external TSP, and nitrogen chemical shifts indirectly to DSS as described by Wishart et al. (1995). Spectra were analysed with XEASY (Bartels et al., 1995) and FELIX 95.0 (Biosym/MSI, California), and CSI calculations were carried out in NMRView (Johnson and Blevin, 1994).

Backbone NH ^{15}N relaxation times were measured with the procedures described by Kay et al. (1992) and Farrow et al. (1994) using spectral widths of 8000 Hz (^{1}H) and 2500 Hz (^{15}N). Relaxation delays (τ) for the T_2 measurements were 10, 30, 50, 70, 90, 110, 150 and 250 ms with the experiments at 10, 50 and 150 ms repeated. It is the nature of these pulse sequences that sample heating may occur during the relaxation delay period, which precludes the use of delay times in excess of ~250 ms, and thus signal intensities of spins with T_2 times greater than ~50 ms will not have

Table 1. Parameters for the best-fit values of the 50.65 MHz ^{15}N R_2 relaxation times of the N-terminal regions of intact colicin E9 and the T_{1-61}-DNase fusion protein shown in Figure 11. The intrinsic relaxation rates ($R_{intrinsic} = 0.19\ s^{-1}$) and persistence lengths ($\lambda_0 = 7$) were the same for both proteins

Intact colicin				T_{1-61}-DNase fusion protein			
Cluster	Cluster centre	Persistence length (λ)	R_2 relaxation rate [s^{-1}]	Cluster	Cluster centre	Persistence length (λ)	R_2 relaxation rate [s^{-1}]
1	9	1	2.2	1	9	2	1.1
2	15	1	1.2	2	16	2	1
3	21	1	0.5	3	22	1	0.4
4	26	3	0.4	4	26	3	0.3
5	33	1	0.9	5	31	2	0.6
6	37	1	1.5	6	37	2	1.6
7	40	1	4.5	7	40	1	3.55
8	43	1	2.0	8	43	1	4.3
9	46	2	5.5	9	46	1	4.5
10	56	3	4.4	10	56	3	4.8

decayed to zero. Nonetheless, signals from the most flexible N terminal region (residues 5–8) decayed on average by 43% and residues 9–61 by >70%. The relaxation delays (τ) for the T_1 measurements were 10, 50, 80, 200, 500, 750, 2,000, and 3,500 ms, with the experiments at 10, 200 and 500 ms repeated. A recycle time of 3 s was used for each relaxation delay. Heteronuclear NOE spectra were measured with the procedure described by Farrow et al. (1994). The spectra were recorded as 1024 × 128 complex data points with 32 transients per point. Proton saturation was achieved with a pulse train of 120° pulses every 5 ms for 3 s. Steady-state NOE values were determined from spectra recorded in the presence and absence of proton saturation. For the spectra recorded with proton saturation a 2 s relaxation delay was followed by the period of saturation, whilst those recorded without proton saturation used a relaxation delay of 5 s. For the determination of peak height uncertainties three sets of the saturated/un-saturated experiments were run. Relaxation times and heteronuclear NOEs were determined from peak heights as described by Stone et al. (1992, 1993) using programs provided by Palmer: (http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software.html). In all cases relaxation data were analysed with a 2-parameter model (Viles et al., 2001a). Uncertainties in NOEs were calculated as described by Nicholson et al.(1992). Overlapping peaks were divided into three classes: those where overlap at 500 and 600 MHz prevented reliable measurement of their individual relaxation parameters, those which overlapped at 500 MHz but were sufficiently resolved at 600 MHz to allow their individual relaxation parameters to be measured, and those overlapped at 500 MHz but sufficiently resolved for their relaxation parameters to be determined. Resonances in these three categories were those for: Thr 16/Ser 50, Ser 37/Ser 67, Gly 53/Gly 61 and Gly 32/Gly 49/Gly 59; Gly 12/Gly 31, Gly 18/Gly 38 and Trp 39/Leu 62; and Ser 17/Ser 41 and Gly 22/Gly 36, respectively. Where two peaks overlapped at 500 MHz but were resolved at 600 MHz and had similar relaxation rates at this field, the average value of the 500 MHz rates was adopted for both residues. Model-free analysis was carried out in two stages using the in-house FORTRAN programs R2R1Win and LSWin with protocols adapted from Mandel et al. (1995). Both programs are based on a simulated annealing algorithm (Press et al., 1988). The program R2R1Win was used to determine the global rotational diffusion coefficient from the T_1/T_2 ratio with T_1 and T_2 given by Equations 1 and 2 in Mandel et al. (1995). Subsequently, the dynamics of the individual NH bonds were found using the program LSWin based on the Lipari–Szabo model (Lipari et al., 1982a,b) extended by Clore et al. (1990) using Equation 4 of Mandel et al. (1995) and with the global rotational diffusion coefficient determined in the first stage. Reduced spectral density mapping, assuming a single spectral density was performed using an in house FORTRAN program, RSAWin based on Equations 12–14 of Farrow et al. (1995b) which relate T_1, T_2 and the heteronuclear ^1H-^{15}N NOE to the spectral

densities J(0), J(ω_N) and J(0.87ω_H). J(ω) was assumed to be constant at $\omega \approx \omega_H$, following Method 1 of Farrow et al. (1995b). The ^{15}N chemical shift anisotropy for the amide NH was taken as -160 ppm and the NH bond length as 1.02 Å.

Results and discussion

Characteristics of the T_{1-61}-DNase fusion protein

In a previous NMR study, we showed that the first 83 residues of intact colicin E9 were largely unstructured and flexible (Collins et al., 2002). The backbone ^{1}H-^{15}N resonances of this N-terminal region had poor chemical shift dispersion, and were heavily overlapped, causing some difficulty for signal assignment. To simplify the spectrum and thus facilitate observation and signal assignment of the TolB box and neighbouring residues, we expressed the T_{1-61} region as a fusion protein with the E9 DNase to ensure its stability during expression in *E.coli* and purification (Figure 1). The first 61 residues of the intact colicin were directly linked to the DNase domain via an 8-residue thrombin cleavage sequence peptide, and thus this fusion protein lacked most of the T-domain, including the glycine-rich region (residues 62–83). The thrombin cleavage sequence was intended to enable proteolytic release of the T_{1-61} peptide. However, the fusion protein without cleavage was used for experiments reported herein owing to its increased stability and the poor yields of intact T_{1-61} peptide following proteolysis. The binding properties of the T_{1-61}-DNase fusion protein were checked by gel-permeation chromatography, isothermal titration calorimetry and surface plasmon resonance, and the results showed that its ability to bind TolB is similar to that of the intact colicin (unpublished data of R. James, C. Kleanthous, S. Loftus, S. Rendall and K. Tozawa). Therefore it can be concluded that the T_{1-61}-DNase fusion protein is a good alternative to the full-length colicin E9 for studying the properties of the T_{1-61} peptide region.

NMR assignments for the T_{1-61}-DNase fusion protein

The ^{1}H-^{15}N HSQC spectrum of the T_{1-61}-DNase fusion protein at a high threshold level (Figure 2A) has a relatively narrow ^{1}H chemical shift dispersion for the sharp intense backbone NH signals, typical of those for a largely unstructured protein (Dyson & Wright, 2001) and similar to the corresponding spectrum of intact colicin E9 (Collins et al., 2002), though with a reduced number of signals (Figure 3). At this high threshold level (Figure 2A) only the resonances of the T_{1-61} region, the 8-residue thrombin cleavage sequence and the two C-terminal residues of the DNase domain are clear. It has been possible to unambiguously assign 53 of the expected 58 backbone NH resonances for the first 61 residues, as well as all of the expected 7 backbone NH resonances of the thrombin cleavage sequence. The assignments are listed in Table 2 of the Supporting Information. It was difficult to assign some of the glycine and serine resonances through conventional triple-resonance spectra because of heavy overlap of signals (Figure 2A) and the repeated appearance of sequences such as GSG and GGG (Figure 1). We found the HNN experiment (Panchal et al., 2001) to be extremely useful for these difficult regions and it permitted the almost complete assignment of resonances of the T_{1-61} region. Figure 4 shows an example where the resonances of Gly 48 and Gly 58 were assigned through use of the HNN experiment.

At a lower threshold level, weaker resonances of the DNase domain of the T_{1-61}-DNase fusion protein become visible (Figure 2B). These were found to have similar chemical shifts to the resonances of E9 DNase on its own (Whittaker et al., 2000), allowing the assignments for the latter to be carried over to the fusion protein. The pattern of DNase chemical shifts show that this domain of the fusion protein does not contain bound metal ions at its HNH motif metal binding site (Hannan et al., 2000). To compare the line-widths of the resonances of the first 69 residues of the fusion protein and the DNase domain, four well-isolated signals from both regions were chosen (Figure 2, insets). These examples show that the resonances of the T_{1-61} region are generally sharper than those of the DNase domain, as is confirmed by the decay curves for the ^{15}N spin-spin relaxation times of selected residues from different parts of the T_{1-61}-DNase fusion protein (Figure 5).

Most of the signals of the T_{1-61} region of the fusion protein had similar chemical shifts to those of intact colicin E9, with chemical shift differences from the sequence-corrected random coil values indicating that it lacked any stable secondary or tertiary structure (Figure 6). No multiplicity of signals was detected, indicating there was no conformational heterogeneity of the T_{1-61} region, unlike for the intact colicin E9.

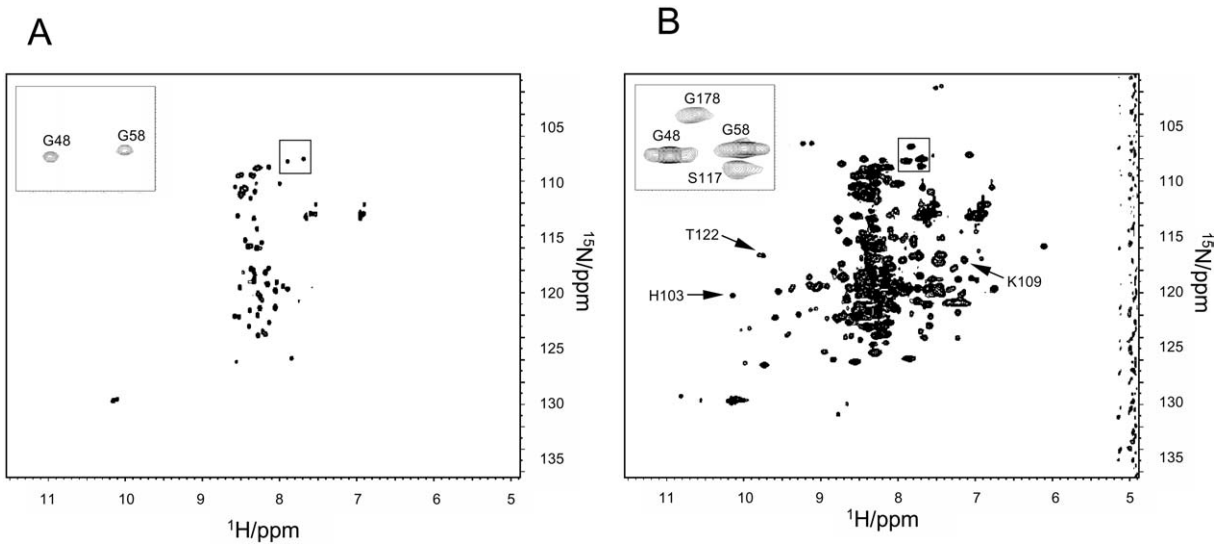

Figure 2. The ^{1}H-^{15}N HSQC spectrum of the T_{1-61}-DNase fusion protein at a high threshold (A) and at a lower threshold (B). Only intense signals belonging to the T_{1-61} region and the thrombin cleavage sequence are displayed in (A) but in (B) broader signals from the DNase domain are seen. *Insets*: enlargement of a region indicated in the main figures to show close-ups of resonances from the T_{1-61} region and the DNase domain. ^{1}H line-widths of these four resonances at half height were ~21 Hz, ~23 Hz, ~32 Hz and ~42 Hz for Gly 48, Gly 58, Gly 178 (corresponding to Gly 110 of E9 DNase) and Ser 117 (corresponding to Ser 49 of E9 DNase), respectively.

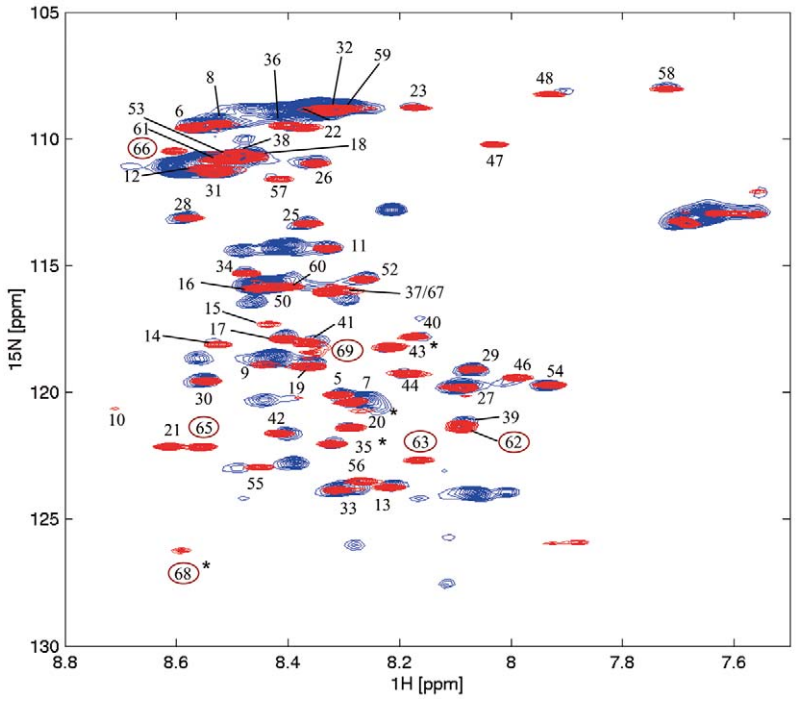

Figure 3. The overlaid ^{1}H-^{15}N HSQC spectra of the T_{1-61}-DNase fusion protein (red) and intact ColE9 with Im9 (blue). Assignments of selected resonances of the T_{1-61}-DNase fusion protein are shown. The residues of the thrombin cleavage sequence are circled. The spin-spin-relaxation decay curves for the resonances marked with asterisks are shown in Figure 5.

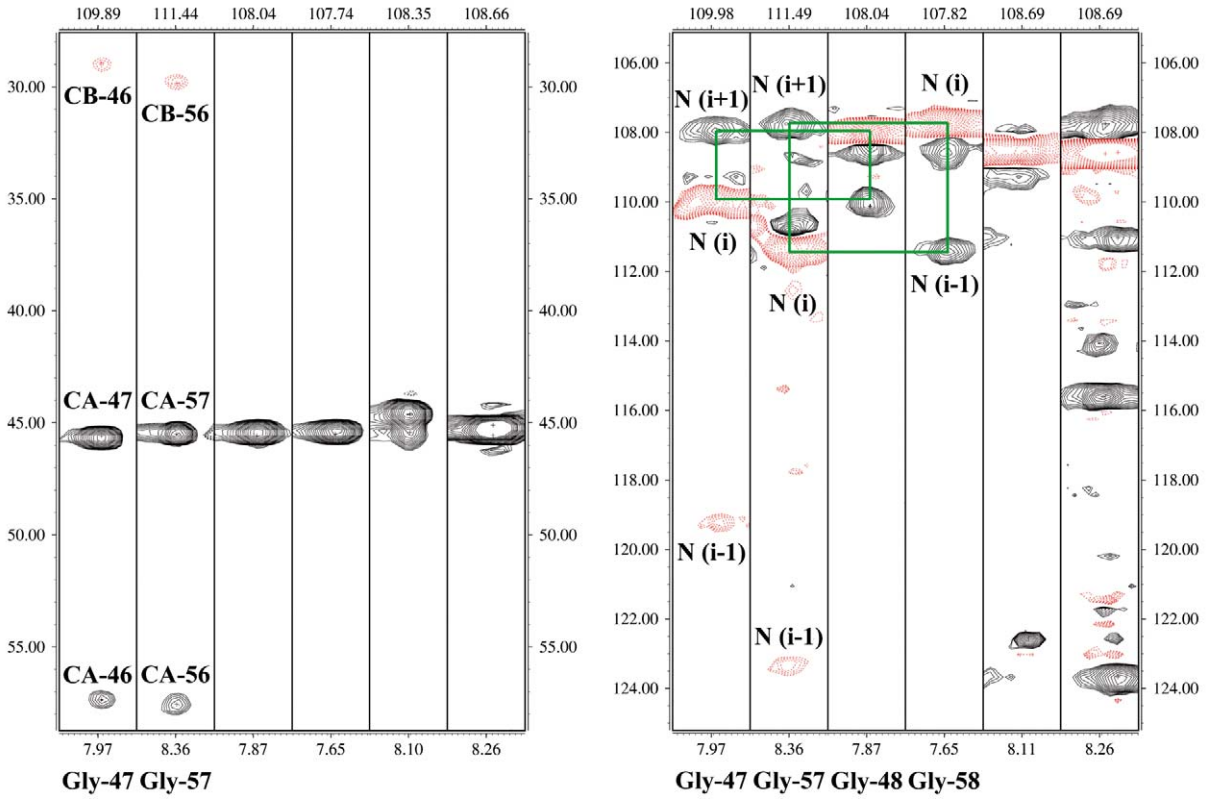

Figure 4. Comparison of selected strip plots of equivalent regions of HNCACB (left) and HNN (right) spectra of the T_{1-61}-DNase fusion protein; red and black contours are negative and positive respectively. Gly 47 and Gly 57 were unambiguously assigned from the standard triple resonance experiments and the strips for these residues are labeled. Alongside these are strips of potential neighboring (i+1) residues. From the HNCACB data no strip can be unambiguously assigned to either Gly 48 or Gly 58 however, the same strips from the HNN data clearly reveal the (i) to (i+1) and (i−1) cross peaks between both pairs of residues 47/48 and 57/58. Gly 47 and Gly 57 strips show the characteristic negative, negative, positive peak pattern for a X-Gly-Gly triplet and similarly the positive, negative, positive Gly 48 and Gly 58 peaks are characteristic of a Gly-Gly-Gly triplet.

NMR assignments for intact colicin E9

The assignments for the T_{1-61} DNase fusion protein allowed equivalent peaks in the spectrum of intact colicin E9 to be identified. In total, assignments of 16 backbone NH resonances of the intact colicin E9 were obtained from such comparative analyses to add to the assignments reported by Collins et al. (2002). In addition to these, eight resonance assignments reported by Collins et al. (2002) were re-assigned. Most of these re-assignments arose because the samples of intact colicin E9 studied by Collins et al. (2002) exhibited a multiplicity of peaks for some residues with both a major and at least one minor form detected. For the T_{1-61} DNase fusion protein only a single form was detected and with few exceptions the chemical shifts of its backbone NH groups were the same as those of the major form of the intact protein. However, in eight cases this was not so and the T_{1-61} fusion protein signals were most similar to the previously identified minor signals of the intact colicin. Since the minor signals were not always identified by Collins et al. (2002) through scalar connectivities, where indicated by comparison with our data for the T_{1-61} DNase fusion protein we have re-assigned them to be the corresponding signals to those of the T_{1-61}-DNase fusion protein.

Backbone relaxation properties of the T_{1-61}-DNase fusion protein

Backbone ^{15}N T_1 and T_2 relaxation times were determined for 47 out of the first 61 residues of the T_{1-61}-DNase fusion protein, for 6 out of the 8 residues of its thrombin cleavage sequence (Figures 7A and C), and for 40 residues of its DNase domain (Figures 8A and C). There is a reasonable correspondence between the amide T_1 and T_2 times of the C-terminal regions

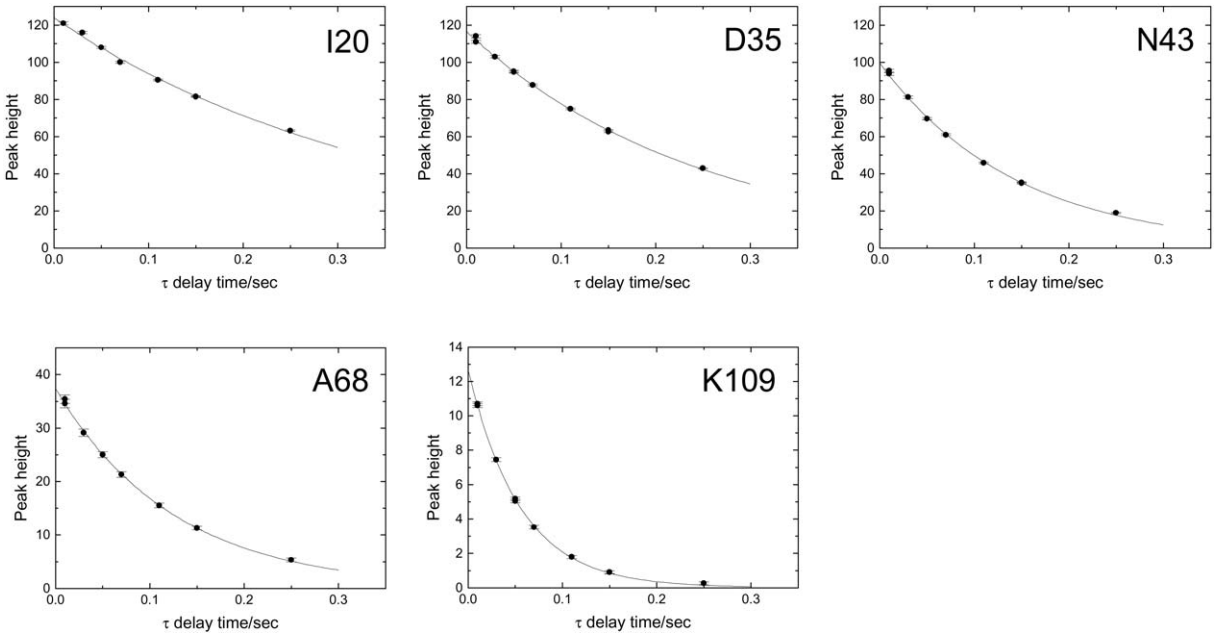

Figure 5. Typical examples of 50.65 MHz ^{15}N spin-spin relaxation decay curves for backbone NH resonances of the T_{1-61}-DNase fusion protein in 90% H_2O/10% D_2O and 50 mM sodium phosphate buffer, pH 6.3, and at 288 K. The position of each resonance in the ^{1}H-^{15}N HSQC spectrum is shown in Figure 3, except for Lys109 which is indicated in Figure 2B. T_2 values are: 0.36 ± 0.003 s (Ile20, pre-TolB box region), 0.25 ± 0.002 s (Asp35, TolB box), 0.14 ± 0.001 s (Asn43, post-TolB box region), 0.13 ± 0.002 s (Ala68, thrombin cleavage sequence), and 0.055 ± 0.001 s (Lys109, DNase domain).

of the fusion protein and the intact colicin (Figure 7), especially in the decrease in T_2 at residues forming the TolB box between Asp 35 and Trp 39. There is also good agreement between the T_1 and T_2 times of the C-terminal region of the fusion protein and those of the isolated DNase domain (Figure 8). The ^{1}H-^{15}N NH NOEs of the latter two proteins also showed a reasonably good correspondence (Figures 8E and F), but for the N-terminal regions of the fusion protein and the intact colicin there is a large difference in ^{1}H-^{15}N NOEs (Figures 7E and F). For the intact protein the NOEs of the N-terminal residues are negative apart from those of residues 40, 41, 47 and 57, while for the fusion protein the NOEs are negative before residue 37, apart from 27, and positive thereafter. As noted by Collins et al. (2002) in the earlier NMR study of intact colicin, the relaxation characteristics of the N-terminal region encompassing residues 1-83 indicates that there is motion of this stretch of amino acids that is more rapid than the tumbling motion of the globular domains of the complex, and this is also the case for the fusion protein.

Reduced spectral density analyses

Dynamics of a polypeptide chain can be deduced from the backbone NH relaxation parameters T_1, T_2, and ^{1}H-^{15}N NOE through the use of the reduced spectral density functions J(0), J(ω_N), and J(0.87ω_H). The magnitudes of the spectral density functions are sensitive to motions at the corresponding frequencies, zero, ω_N and 0.87ω_H. Thus J(0) reflects slow internal motions on the millisecond to microsecond time scale as well as slow global rotational diffusion. In contrast, J(0.87ω_H) reports on the presence of fast internal motions, on the picosecond timescale. Lying between these extremes is J(ω_N), which in the present case (Figure 9) is rather insensitive to the presence of internal motions and reflects mainly the global rotational diffusion.

Considering the T_{1-61}-DNase fusion protein, large values for J(0) and small values for J(0.87ω_H) are seen where the dynamics are dominated by nanosecond time scale motions, whilst for the more flexible regions the reverse is true. As an indication of the typical magnitudes expected for residues in a rigid 15 kDa globular protein showing only global rotational diffusion, J(0) would be greater than 5.0 ns · rad^{-1} and

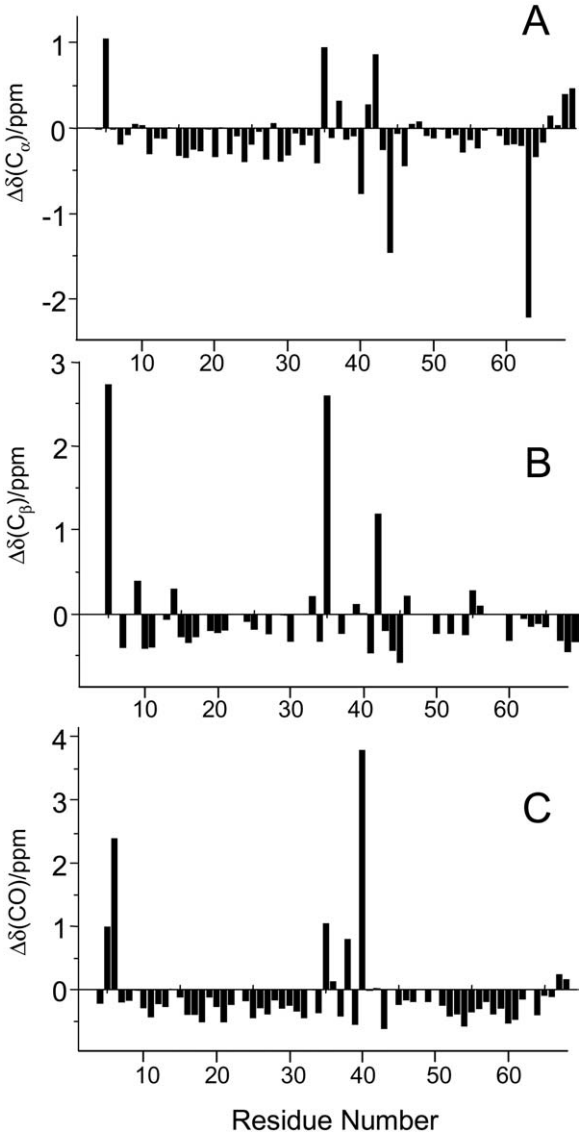

Figure 6. Conformation-dependent chemical shifts of the T_{1-61} component and thrombin cleavage sequence of the T_{1-61}-DNase fusion protein. The $\Delta\delta$ values for the Cα, Cβ and ^{13}CO resonances are the differences between the experimentally determined chemical shifts and the sequence-corrected random coil values reported by Schwarzinger et al. (2001).

$J(0.87\omega_H)$ less than 10 $ps \cdot rad^{-1}$. Consequently, no part of the T_{1-61} region behaves as though its global dynamics are determined by the DNase domain of the fusion protein (Figure 9). This is particularly clear for the residues from 5 to 34 where the average J(0) is 0.63 $ns \cdot rad^{-1}$ and $J(0.87\omega_H)$ is 32 $ps \cdot rad^{-1}$ giving clear evidence for fast local segmental motion at the N-terminus. Yet even within this highly flexible region there are indications of a different pattern of spectral densities, for example: residues 9 and 15 have J(0) >0.75 $ns \cdot rad^{-1}$ and $J(0.87\omega_H) \sim 30$ $ps \cdot rad^{-1}$.

Two regions of structure for the T_{1-61} fusion protein are indicated by the spectral densities from residue 34 to the end of the T_{1-61} region at residue 61, the first lies between residues 35 and 52 and the second between residues 52 and 61. Thus there is a steady increase in the value for J(0) from residues 34 to 46, with J(0) equal to 1.93 $ns \cdot rad^{-1}$ at residue 46, whilst at the same time $J(0.87\omega_H)$ decreases to 21.2 $ps \cdot rad^{-1}$. Both J(0) and $J(0.87\omega_H)$ point to slower dynamics in this region, which includes the TolB box. Slower dynamics implies a less flexible structure and thus the presence of local structure, or order. However, because of the nature of the reduced spectral density analysis it is not possible to define the degree of structure present.

After residue 46 of the T_{1-61} fusion protein, J(0) starts to fall to a minium at residue 52 and in so doing defines the end of the first region of structure. However, J(0) does not fall back to the values typical of the highly flexible N-terminus, instead a minimum value of 1.22 $ns \cdot rad^{-1}$ is seen indicating that although some faster dynamics have been restored there must still be some relative degree of structure. Thereafter J(0) increases again to a second maximum of 2.01 $ns \cdot rad^{-1}$ at residue 56 before falling away to residue 61, defining the second region of structure. Given the caveat of the difficulty of interpreting J(0) quantitatively, providing similar motional models describe the dynamics in the structured regions the similar values for the J(0) maximum suggests a similar degree of ordering in these two regions. $J(0.87\omega_H)$ mirrors these changes, showing an increase from residues 46 to 52 and a decrease from residues 52 to 61. The pattern of an increase in J(0) and decrease in $J(0.87\omega_H)$ associated with slower dynamics and a less flexible polypeptide chain is seen again after residue 61, this time extending through the thrombin cleavage sequence and into the DNase domain itself where J(0) and $J(0.87\omega_H)$ are of the order of 5.0 $ns.rad^{-1}$ and 6.0 $ps.rad^{-1}$ respectively. In this case the residues increasingly reflect the slower global rotational diffusion of a structured protein, rather than the local segmental motion of a disordered chain. Given that a conformationally flexible chain might be expected to show a more drastic step change in J(0) and $J(0.87\omega_H)$ instead of the gradual one observed, some structure must be present in the thrombin cleavage sequence. However, the strong downward trend in J(0) from residue 69 to residue 61 means that the increase in these spectral

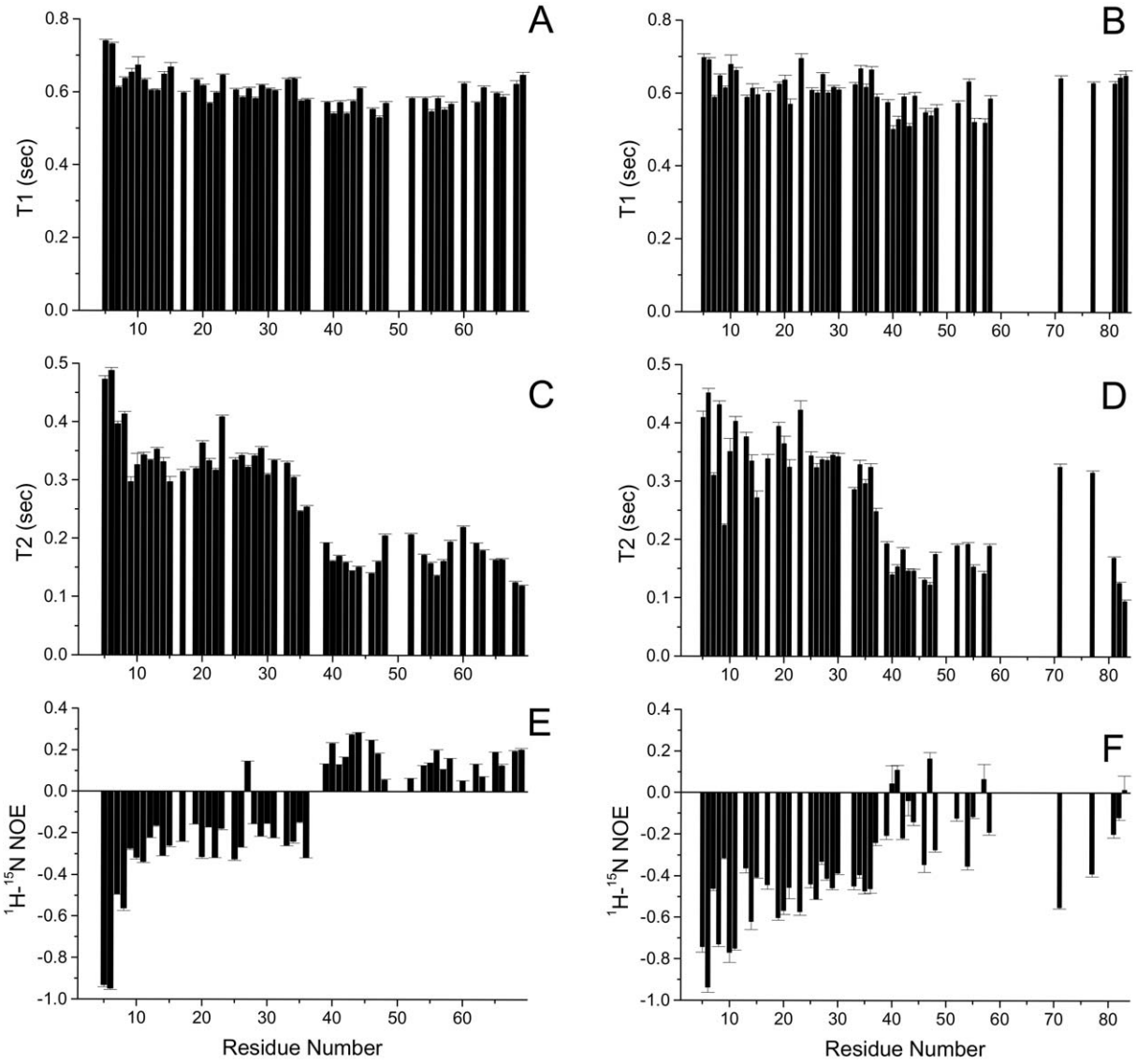

Figure 7. Backbone 50.65 MHz ^{15}N relaxation times and ^{1}H-^{15}N NOEs of the T_{1-61}-DNase fusion protein in 90% H_2O/10% D_2O and 50 mM sodium phosphate buffer, pH 6.3, and at 288 K, and the major conformer of intact colicin E9 with Im9 in 90% H_2O/10% D_2O and 50 mM potassium phosphate buffer, pH 6.2, and at 288 K: (A) and (B) T_1 times; (C) and (D) T_2 times; and (E) and (F) NOEs of NH resonances. (A), (C) and (E) are for T_{1-61}-DNase fusion protein and (B), (D) and (F) are for the major conformer of the intact ColE9 with Im9.

densities around residue 56 cannot be attributed to residual structure from the thrombin cleavage sequence but must stem from some ordering element within this region.

A similar pattern for J(0) was found for the intact colicin, with a difference from the T_{1-61}-DNase fusion protein being that after residue 60, J(0) reduces rather than increases. The J(ω_N) and J($0.87\omega_H$) plots also reveal that as with the fusion protein the 39–45 and 54–56 regions are different to the 1–36 region, though the difference is not as marked as for the J(0) plots. The reduction in J(0) after residues 54–56 is connected to the presence of the glycine-rich region from residues 61 to 80, which acts as a hinge similar to glycines of urea-denatured apomyoglobin (Schwarzinger et al., 2002), as the T_{1-61}-DNase fusion protein lacks the glycine-rich region and J(0) increases after residues 54–56 (Figure 9).

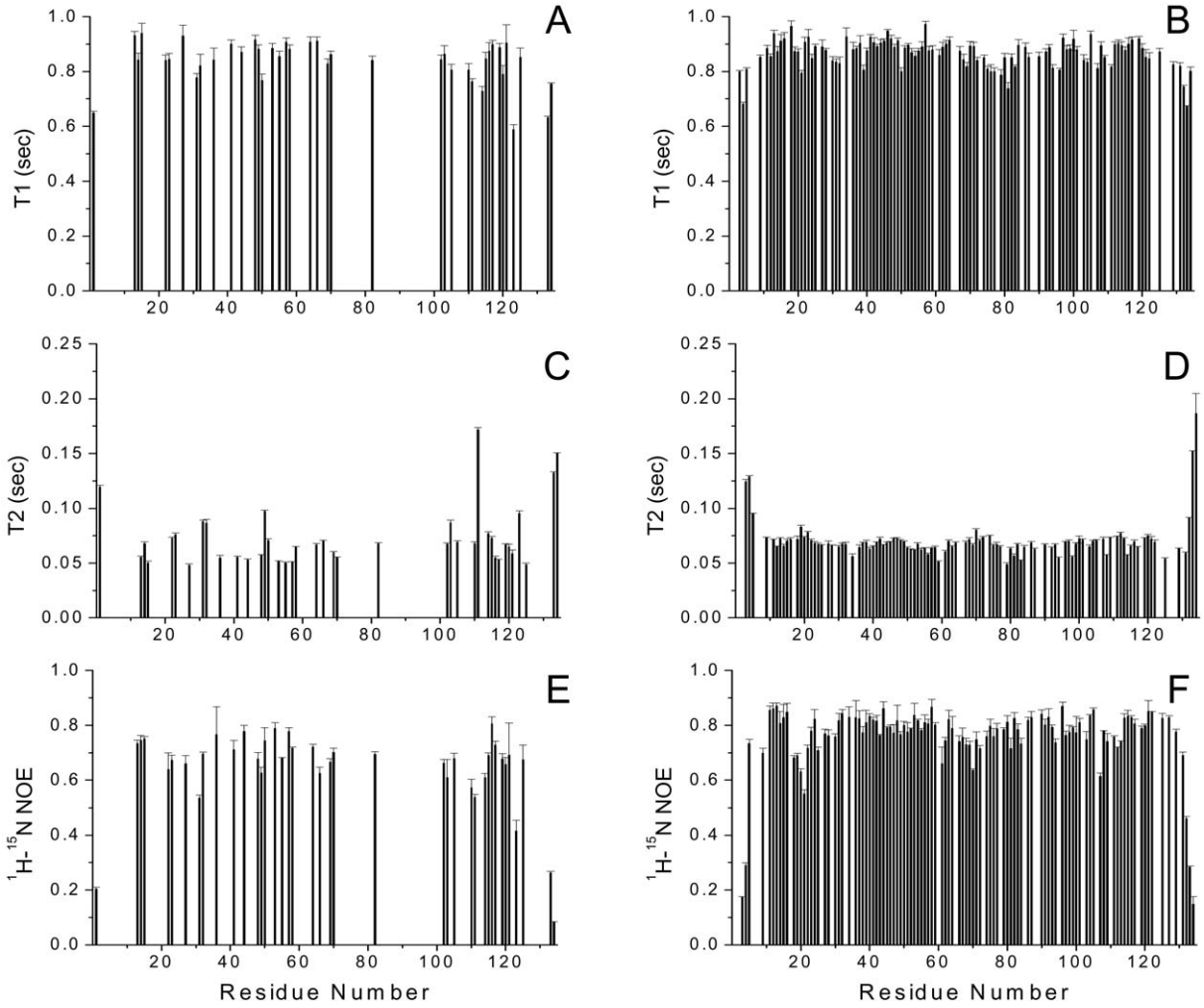

Figure 8. Backbone 50.65 MHz ^{15}N relaxation times and ^{1}H-^{15}N NOEs of the DNase domain of the T_{1-61}-DNase fusion protein (left) and the major conformer of the DNase domain on its own (right) in 90% H_2O/10% D_2O and 50 mM sodium phosphate buffer, pH 6.3, and at 288 K. (A) and (B) T_1 times; (C) and (D) T_2 times; and (E) and (F) NOEs of NH resonances. (A), (C) and (E) are for T_{1-61}-DNase fusion protein and (B), (D) and (F) for the isolated metal-free DNase domain.

Lipari–Szabo model-free formalism

Analysis of ^{15}N relaxation times is commonly carried out using the Lipari–Szabo model-free formalism (Lipari and Szabo, 1982a,b). This methodology is applicable when the motions of the backbone N-H vectors can be understood in terms of fast internal motions superimposed on an anisotropic global diffusion. These are represented by the fast-motion order parameter, S_f^2, with a motional frequency fixed at 1 ps, and the generalised order parameter, S^2, relative to the rotational diffusion of the protein. In the present case the aim was to determine the extent to which the relaxation characteristics of the T_{1-61} tethered domain of the fusion protein were influenced by global rotational diffusion of the DNase. Effective decoupling of the dynamics of the T_{1-61} tail would be evident through small values of the generalised order parameter S^2. For simplicity, the rotational diffusion was treated as isotropic, and from the T_1/T_2 ratio of residues meeting the rigid backbone criteria of a ^{1}H-^{15}N NOE > 0.65 (Tjandra et al., 1995), and in the absence of conformational exchange, the isotropic diffusion coefficient of the DNase domain was determined to be 1.53×10^7 s^{-1}, compared to the value of 1.54×10^7 s^{-1} for the apo-DNase itself (N.J. Clayden, E.S. Collins and G.R. Moore, unpublished data). Lipari–Szabo fitting was performed using standard model selection pro-

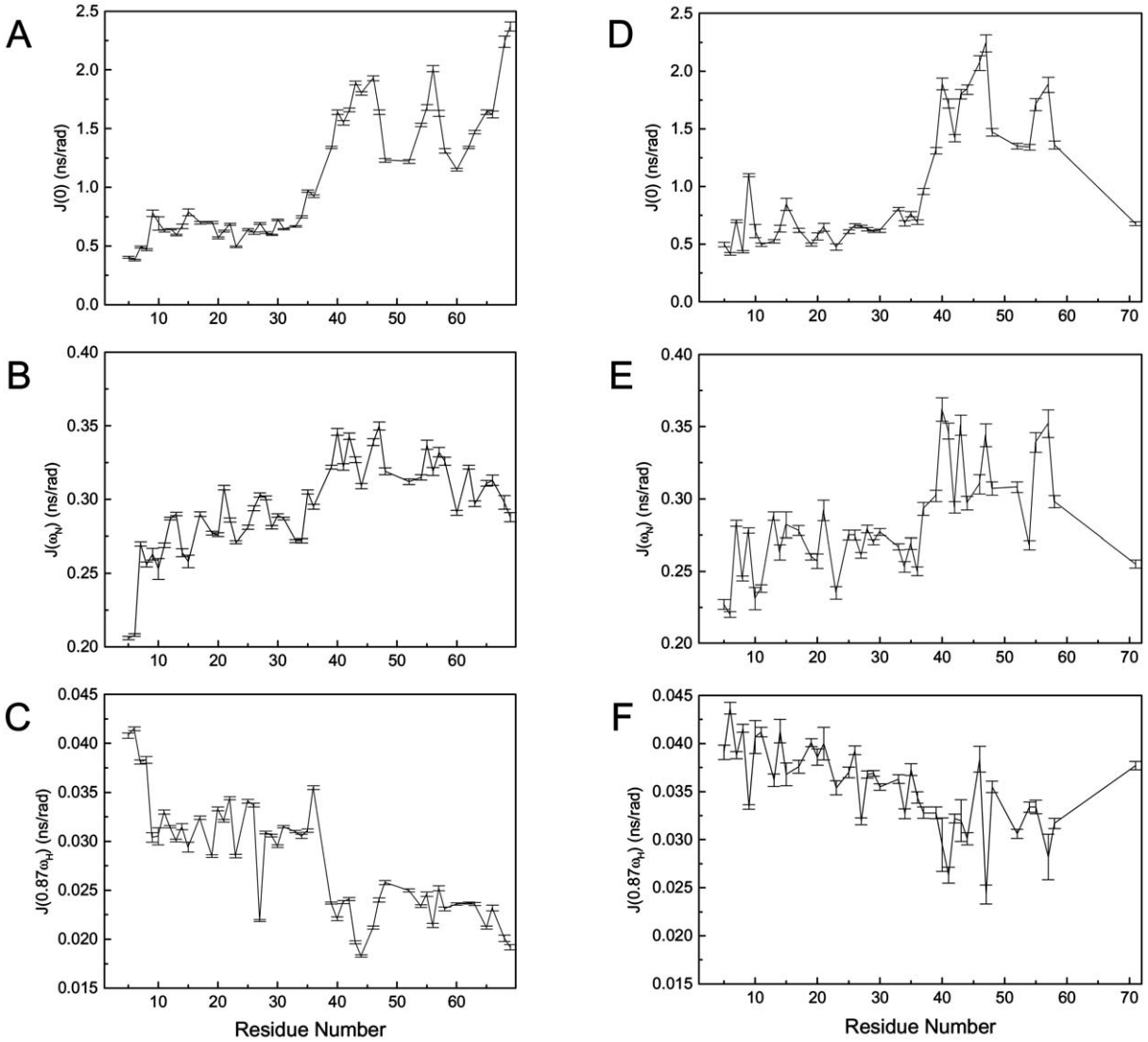

Figure 9. Plots of the J(0), J(ω_N) and J(0.87ω_H) spectral density values of backbone NH groups calculated from the relaxation data in Figure 8 versus residue number for the T_{1-61} DNase fusion protein (A–C) and intact colicin E9 (D–F).

cedures (Mandel et al., 1995), and the variation in S^2 is shown in Figure 10A for the T_{1-61} region. All residues in this region required model 5 fits (Mandel et al. 1995) and had S^2 values of <0.5, demonstrating the importance of local segmental motions in their ^{15}N relaxation. The sequence variation of S^2 indicates that the region between residues 1 and 36 has different motional characteristics from those of the region between residues 39 and 61, though this difference was not reflected by the fast-motion order parameter, S_f^2, (Figure 10B). Since $S^2 = S_s^2 S_f^2$ (where S_s^2 is the slow-motion order parameter) the data indicate that the T_{1-61} region has a common fast motional behaviour reflected by S_f^2 with a slower and sequence-dependent motion represented by S_s^2 superimposed on it. The fast motion is the independent segmental motion of the T_{1-61} region decoupled from the global diffusion of the DNase domain. The slower motion does not appear to be a gradual decoupling of the T_{1-61} region from the motional behaviour of the thrombin cleavage sequence or the DNase domain because the sequence dependence of S^2 (Figure 10A) is reflected by the ^{15}N T_2 times (Figure 7C), amide ^1H-^{15}N NOEs (Figure 7E) and J(0) spectral densities (Figure 9A) for the T_{1-61} region of the fusion protein, and a sim-

ilar pattern of sequence variation is observed for these parameters with the intact colicin (Figures 7D, 7F and 9D). Thus the slower motional behaviour affecting S^2 (Figure 10A) must arise from local interactions of the residues and therefore they are a fundamental property of the amino acid sequence.

Side chain cluster analyses

As discussed above, the sequence variation of the backbone relaxation properties of the N-terminal regions of the T_{1-61}-DNase fusion protein and intact colicin indicate that neither are completely disordered. Transverse relaxation rates have been used to detect regions of preferential structure in otherwise disordered proteins and to this end we adopted the approach of Schwalbe et al. (1997), which has been used to describe unfolded states of lysozyme (Klein-Seetharaman et al., 2002) and apo-myoglobin (Schwarzinger et al., 2002). This approach starts with a model of polypeptide motion in which the nature of neighbouring residues has little effect on the backbone ^{15}N relaxation rates, which are largely determined by the segmental motion of the polypeptide chain, and which are given by Equation 1 (Klein-Seetharaman et al., 2002; H. Schwalbe, personal communication). This equation describes the relaxation rate of the backbone amide of residue i in a flexible and extended polypeptide as a function of chain position.

$$R(i) = R_{\text{intrinsic}} \sum_{j=1}^{N} e^{-\frac{|i-j|}{\lambda_0}}, \qquad (1)$$

where, $R_{\text{intrinsic}}$ is the intrinsic relaxation rate of a chain of length N, and λ_0 is the persistence length in number of residues of the chain. Equation 1 was used to fit the two parameters to the experimental R_2 values of the intact colicin and T_{1-61}-DNase fusion protein, and the resulting plots are shown in Figure 11 as a line of + symbols. Although these plots follow the underlying trend in experimental R_2 values, being essentially constant throughout the extended polypeptides and reducing at their termini, large positive deviations are seen around residues 40, 46 and 56. These deviations can be attributed to localized non-random structural features and can be incorporated into the polymer dynamics model and simulated using Equation 2 (Klein-Seetharaman et al., 2002; H. Schwalbe, personal communication). This modification superimposes upon the segmental motion further gaussian distributions of deviations that are centred

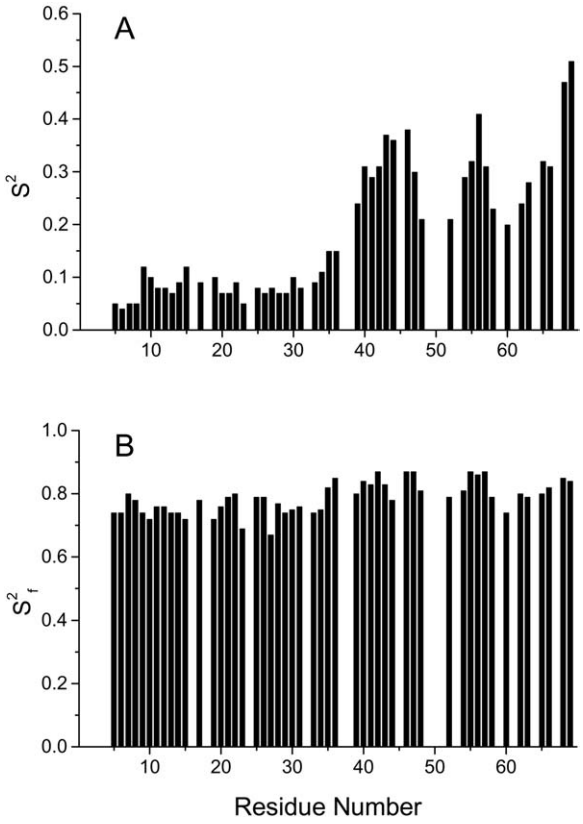

Figure 10. Order parameters for T_{1-61}-DNase fusion protein. Bar graphs of (a) the generalized order parameter S^2 and (B) the fast-motion order parameter S_f^2. Values were obtained by fitting the relaxation data presented in Figure 8 to the extended Lipari-Szabo model.

about clusters of residues at χ_{cluster}, with each cluster having a characteristic relaxation rate R_{cluster} and persistence length λ_{cluster}. The persistence length indicates the number of amino acids in each cluster whose relaxation is significantly affected by being part of the cluster and thus it defines the size of the cluster.

$$R(i) = R_{\text{intrinsic}} \sum_{j=1}^{N} e^{-\frac{|i-j|}{\lambda_0}} + \sum_{\text{cluster}} R_{\text{cluster}} \; e^{-0.5\left[\frac{|i-\chi_{\text{cluster}}|}{\lambda_{\text{cluster}}}\right]^2} \qquad (2)$$

We made no attempt to fit data to the model for regions beyond where R_2 values begin to indicate the presence of highly structured domains.

Modeling with Equation 2 for the intact colicin was only carried out for residues 1–83, which largely resembles an unrestrained polypeptide (Collins et al., 2002). The R_2 values for this rise rapidly at residues

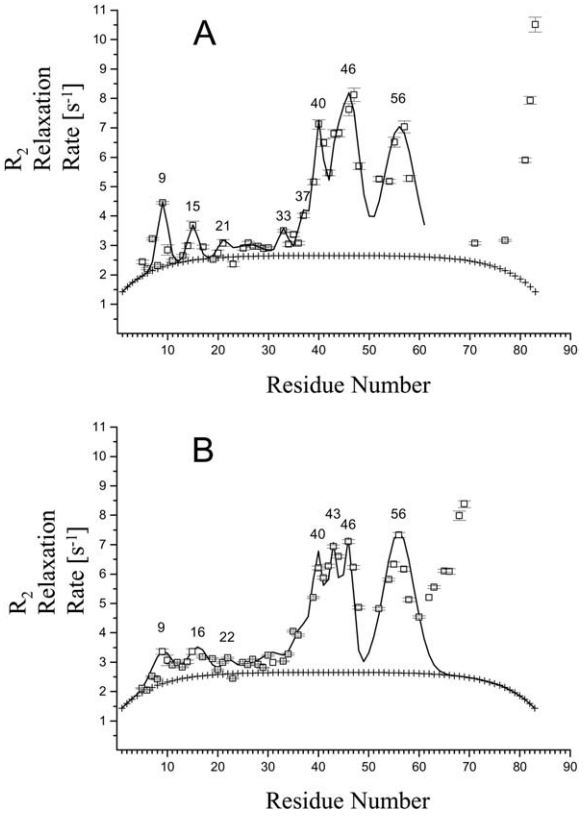

Figure 11. Comparison of 50.66 MHz transverse ^{15}N relaxation rates, R_2 s^{-1}, for backbone amide groups of the intact colicin (A) and the T_{1-61}-DNase fusion protein (B). The experimental data are shown as scatter plots of open squares with error bars, the simulation of the rates assuming only unrestricted segmental motion according to Equation 1 by +, and additional slow dynamics according to Equation 2 using the data in Table 1 by the solid line. RMSDs between the real and simulated data are 0.452 and 0.48 for (A) and (B), respectively. For clarity some of the cluster center residues are indicated.

81–83 as the effects of the globular part of the colicin become apparent. Between residues 62 and 77 it appears that the relaxation rates approach the underlying rates of those of the individual amide groups as described by Equation 1, and although overlap in the NMR spectra severely limit the number of residues in this region for which R_2 values have been obtained, those that are available are unambiguous. This region between residues 62–77 is the 'glycine rich' region, which is absent in the T_{1-61} construct (Figure 1). From residues 1 to 61 several clusters of structural features are visible and these can be modeled using Equation 2. Clearly two regions stand out in particular: residues 1–35, containing minor features where the increase in R_2 due to clustering is less than double the intrinsic relaxation rate, and residues 36–61 containing major perturbations of local dynamics. Although the minor clusters are of lower intensity than those between residues 36 and 61, they nevertheless represent real deviations from the intrinsic relaxation rate of over four times the experimental error and are therefore significant ordered features. The enhanced relaxation observed between residues 36 and 61 can be modeled assuming three major clusters; two centred in and immediately after the TolB box at residues 40 and 46, and a third centred at residue 56, with persistence values of 1, 1 and 3 respectively (Table 1). These clusters are augmented by two further clusters that are centred at positions 37 and 43 and have persistence values of 2 and 1, respectively. It would seem from this cluster analysis therefore, that the tryptophan residues at positions 39, 46 and 56 are central to the foundation of regions of enhanced structure, possibly because they act as hydrophobic cores around which other side chains pack.

For the T_{1-61}-DNase fusion protein it was initially assumed that the underlying dynamics of its 1–61 region would be characterized in a similar way to that of the intact colicin, with a λ_0 of 7 and $R_{\text{intrinsic}}$ of 0.19 s^{-1} and this was indeed borne out by the simulation (Table 1). The simulated relaxation rate curves of the two proteins have generally similar overall features, reflected in the root mean squared deviation between them of 0.679. So despite the apparent stiffness of the thrombin cleavage sequence, as illustrated by the rising R_2 values for residues 62–69, there are only small observable differences in cluster formation and dynamics of the T_{1-61} region between the fusion protein and intact colicin (Table 1 and Figure 11). Thus, the thrombin cleavage sequence is not sufficient to significantly couple the dynamics of this region to the more rigid DNase domain, as also indicated by the extended Lipari–Szabo analysis (Figure 10).

General discussion

The NMR data reported herein are consistent with other biophysical data concerning binding characteristics of the T_{1-61}-DNase fusion protein (unpublished data of R. James, C. Kleanthous, S. Loftus, S. Rendall and K. Tozawa) in showing that it represents a valuable model construct to study properties of the translocation domain of the intact colicin. The spectral simplification following removal of the glycine-rich region between residues 61–80 allowed a more com-

plete set of resonance assignments to be obtained for the ^{1}H-^{15}N HSQC spectrum of the T_{1-61} region of the colicin, and also revealed that this region acts as a hinge that leads to greater motional behaviour of the N-terminal region of intact colicin compared to the fusion protein. The consequences of this are most notable in the ^{1}H-^{15}N NOE data (Figures 7E and 7F).

The reduced spectral density (Figure 9) and extended Lipari–Szabo (Figure 10) analyses indicate that there is some ordering within the T_{1-61}region of the sequence, most clearly involving the TolB region between residues 35 and 39 and the sequence downstream of this between residues 40 and 60. Similar indications of order have been reported from ^{1}H-^{15}N NMR relaxation data for regions of recombinant prion proteins (López Garcia et al., 2000; Viles et al., 2001b) and numerous proteins denatured with agents such as urea (Schwalbe et al., 1997; Dyson & Wright, 2001; Klein-Seetharaman et al., 2002; Schwarzinger et al., 2002). Insight into the kind of order present in the otherwise disordered N-terminal region of the colicin comes from modeling the R_2 relaxation times following the procedure of Klein-Seetharaman et al. (2002) (Table 1 and Figure 11). This reveals three major clusters, each involving a tryptophan residue, and two minor clusters further toward the N-terminus. Strikingly, these clusters appear to be involved in binding TolB. Residues previously reported from NMR studies (Collins et al., 2002) to be affected by TolB binding include Ala 33, Ser 34, Asp 35, Trp 39, Ser 40, Ser 41, Glu 42, Asn 43 and Asn 44, correlating well with the TolB box identified by mutagenesis (Garinot-Schneider et al., 1997) and simulated cluster centres at residues 33, 37 and 43 (Table 1). The backbone NH resonances of Asn 10, His 14, His 55 and Trp 56 were also reported to be perturbed when colicin E9 binds TolB (Collins et al., 2002), and these positions correspond to two of the simulated minor clusters centred at residues 9 and 15 and the major cluster centred at residue 56 (Table 1). Thus the picture that emerges from these NMR studies is of a disordered protein containing linear runs of amino acids whose side chains interact, at least some of the time, with the resulting clusters contributing to a protein binding epitope. On binding its TolB partner protein, however, the entire N-terminal 61 residues do not fold into a defined structure (Collins et al., 2002) in contrast to many other disordered proteins that fold on binding partners (Dyson and Wright 2002). While further studies of site specific mutants may reveal details of the nature and relationships of the clusters, it is apparent that the significant region for TolB binding extends beyond the recognized five amino acids of the TolB box (Garinot-Schneider et al., 1997; Bouveret et al., 1998) and involves a considerable degree of order within an otherwise disordered region.

The occurrence of flexibility in the N-terminal regions of translocation domains of colicins N and Ia, which are not homologous with the TolB binding region of colicin E9, has been noted by others (Vetter et al., 1998; Zakharov and Cramer, 2002). The colicin N TolA binding epitope contains two essential tryptophans and on binding TolA NMR studies have shown that this binding epitope switches from a disordered to a structured unit (Anderluh et al., 2003). This suggests that the kind of clusters we have detected in the colicin E9 translocation domain may be present also in other colicin sequences.

Supporting Information Available at: http://kluweronline.com/issn/0925-2738

Acknowledgements

We gratefully acknowledge the BBSRC for its support of this work through a project grant to RJ, the Wellcome Trust for a 600 MHz NMR spectrometer and its support of the Colicin Research Group, HEFCE for a 500 MHz NMR spectrometer, and the CEC for support through contract number QLRT-1999-01003.

References

Anderluh, G., Hong, Q., Boetzel, R., MacDonald, C., Moore, G.R., Virden, R. and Lakey, J.H. (2003) *J. Biol. Chem.*, **278**, 21860–21868

Bartels, C., Xia, T.H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J Biomol NMR*, **6**, 1–10.

Bénédetti, H., Letellier, L., Lloubès, R., Géli, V., Baty, D. and Lazdunski, C. (1992) In *Bacteriocins, Microcins and Lantibiotics*, NATO ASI Series H65, James, R., Lazdunski, C. and Pattus, F. (Eds), Springer-Verlag, Berlin, pp. 215–223.

Bouveret, E., Rigal, A., Lazdunski, C. and Bénédetti, H. (1998) *Mol. Microbiol.*, **27**, 143–157.

Cao, Z. and Klebba, P.E. (2002) *Biochimie*, **84**, 399–412.

Carr, S. Penfold, C.N., Bamford, V., James, R. and Hemmings, A.M. (2000) *Structure*, **8**, 57–66.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. and Gronenborn, A.M. (1990) *J. Am. Chem. Soc.*, **112**, 4989–4991.

Collins, E.S., Whittaker, S.B-M., Tozawa, K., MacDonald, C., Boetzel, R., Penfold, C.N., Reilly, A., Clayden, N.J., Osborne, M.J., Hemmings, A.M., Kleanthous, C., James, R. and Moore, G.R. (2002) *J. Mol. Biol.*, **318**, 787–804.

Cramer, W.A. and Stauffacher, C.V. (1995) *Annu. Rev. Biophys. Biomol. Struct.*, **24**, 611–641.

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

Di Masi, R.D., White, J.C., Schnaitman, C.A. and Bradbeer, C. (1973) *J. Bacteriol.* **115**, 506–513.

Dyson, H.J. and Wright, P.E. (2001) *Meth. Enzymol.*, **339**, 258–270.

Dyson, H.J. and Wright, P.E. (2002) *Curr. Opin. in Struct. Biol.*, **12**, 54–60.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.E. and Kay, L.E. (1994) *Biochemistry*, **33,** 5984–6003.

Farrow, N.A., Zhang, O., Forman-Kay, J.D. and Kay, L.E. (1995a) *Biochemistry*, **34**, 868–878.

Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A. and L.E. Kay (1995b) *J. Biomol. NMR*, **6**, 153-162

Garinot-Schneider, C., Penfold, C.N., Moore, G.R., Kleanthous, C. and James, R. (1997) *Microbiology*, **143**, 2931–2938.

Gokce, I., Raggett, E.M., Hong, Q., Virden, R., Cooper, A. and Lakey, J.H. (2000) *J. Mol. Biol.*, **304**, 621–632.

Hannan, J.P., Whittaker, S.B-M., Hemmings, A.M., James, R., Kleanthous, C. and Moore, G.R. (2000) *J. Inorg. Biochem.*, **79**, 365–370

James, R., Kleanthous, C. and Moore, G.R. (1996) *Microbiology*, **142**, 1569–1580

James, R., Penfold, C.N., Moore, G.R. and Kleanthous, C. (2002) *Biochimie*, **84**, 381–389.

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

Kay, L.E., Nicholson, L.K., Delaglio, F., Bax, A. and Torchia, D.A. (1992) *J. Magn. Reson.*, **97**, 359–375.

Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M. and Schwalbe, H. (2002) *Science*, **295**, 1719–1722.

Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**, 4546–4559.

Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.

López Garcia, F., Zahn, R., Riek, R. and Wüthrich, K. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 8334–8339.

Mandel A. M., Akke, M. and Palmer III A.G. (1995) *J. Mol. Biol.*, **246**, 144–163.

Masaki, H., Yajima, S., Akutsu-Koide, A., Ohta, T. and Uozumi, T. (1992) In *Bacteriocins, Microcins and Lantibiotics*, NATO ASI Series H65, James, R., Lazdunski, C. and Pattus, F. (Eds.), Springer-Verlag, Berlin, pp. 379–395.

Nicholson, L.K., Kay, L.E., Baldisseri, D.M., Arango, J., Young, P.E., Bax, A. and Torchia, D.A. (1992) *Biochemistry*, **31**, 5253–5264.

Panchal, S.C., Bhavesh, N.S. and Hosur, R.V. (2001) *J. Biomol. NMR*, **20**, 135–147.

Pilsl, H. and Braun, V. (1995) *Mol. Microbiol.*, **16**, 57–67.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vettering, W.T. (1988) *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, New York.

Raggett, E.M., Bainbridge, G., Evans, L.J.A., Cooper, A. and Lakey, J.H. (1998) *Mol. Microbiol.*, **28**, 1335–1343.

Schwalbe, H., Fiebig, K.M., Buck, M., Jones, J.A., Grimshaw, S.B., Spencer, A., Glaser, S.J., Smith, L.J. and Dobson, C.M. (1997) *Biochemistry*, **36**, 8977–8991.

Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Chung, J., Wright, P.E. and Dyson, H.J. (2001) *J. Am. Chem. Soc.*, **123**, 2970–2978.

Schwarzinger, S., Wright, P.E. and Dyson, H.J. (2002) *Biochemistry*, **41**, 12681–12686.

Soelaiman, S., Jakes, K., Wu, N., Li, C.M. and Shoham, M. (2001) *Mol. Cell*, **8**, 1053–1062

Stone, M.J., Chandrasekhar, K., Holmgren, A., Wright, P.E. and Dyson, H.J. (1993) *Biochemistry*, **32**, 426–435.

Stone, M.J., Fairbrother, W.J., Palmer, III A.G., Reizer, J., Saier, Jr. M.H. Jr. and Wright, P.E. (1992) *Biochemistry,* **31,** 4394–4406.

Taylor, R., Burgner, J.W., Clifton, J. and Cramer, W.A. (1998) *J. Biol. Chem.*, **273**, 31113–31118.

Tjandra, N., Feller, S.E., Pastor, R.W. and Bax, A. (1995) *J. Am. Chem. Soc.* **117**, 12562–12566.

Vetter, I.R., Parker, M.W., Tucker, A.D., Lakey, J.H., Pattus, F. and Tsernoglou, D. (1998) *Structure*, **6**, 863–874.

Viles, J.H., Duggan, B.M., Zaborowski, E., Schwarzinger, S., Huntley, J.J.A., Kroon, G.J.A., Dyson, H.J. and Wright, P.E. (2001a) *J. Biomol. NMR*, **21**, 1–9.

Viles, J.H., Donne, D., Kroon, G., Prusiner, S.B., Cohen, F.E., Dyson, H.J. and Wright, P.E. (2001b) *Biochemistry*, **40**, 2743–2753.

Wallis, R., Reilly, A., Barnes, K., Abell, C., Campbell, D.G., Moore, G.R., James, R. and Kleanthous, C. (1994) *Eur. J. Biochem.*, **220**, 447–454.

Whittaker, S.B-M., Czisch, M., Wechselberger, R., Kaptein, R., Hemmings, A.M., James, R., Kleanthous, C. and Moore, G.R. (2000) *Protein Sci.*, **9**, 713–720.

Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Magn. Reson.*, **B101**, 63–71.

Zakharov, S.D. and Cramer, W.A. (2002) *Biochim. Biophys. Acta*, **1565**, 333–346.