Biochemistry

 Volume 33, Number 41 October 18, 1994

Articles

Sequential Assignments and Identification of Secondary Structure Elements of the Colicin E9 Immunity Protein in Solution by Homonuclear and Heteronuclear NMR†

Michael J. Osborne,‡ Lu-Yun Lian,§ Russell Wallis,‖ Ann Reilly,‖ Richard James,‖ Colin Kleanthous,‖ and Geoffrey R. Moore*,‡

Schools of Chemical and Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, U.K., and Biological NMR Centre, Medical Sciences Building, University Road, University of Leicester, Leicester, LE1 9HN, U.K.

Received March 4, 1994; Revised Manuscript Received May 25, 1994®

ABSTRACT: ^1H–^1H, ^1H–^{15}N, and ^1H–^1H–^{15}N multidimensional NMR spectroscopic studies of the 86 amino acid protein that provides immunity against the DNase action of colicin E9 are reported. Through a combination of 2D NOESY and TOCSY and 3D TOCSY-HMQC, NOESY-HMQC, and HMQC-NOESY-HMQC experiments, almost complete ^1H NMR and backbone ^{15}N NMR assignments have been obtained, and the secondary structure of the protein has been partially elucidated. Approximately 50% of the protein forms three helices. The specificity determining region of the DNase immunity protein, identified from previously reported biochemical studies to include residues 32–40, is helical, indicating that the protein–protein interaction involves residues from at least one helix.

Colicins are plasmid-encoded bacterial proteins which have antibiotic activity and are secreted as part of the stress response system of the producing organism (Luria & Suit, 1987). E-group colicins recognize the vitamin B_{12} extracellular receptor encoded by the *btuB* gene on target cells (DiMasi et al., 1973). Following binding to their receptor, the E-type colicin is translocated across the outer and cytoplasmic membranes, and its cytotoxic activity then leads to cell death. Three cytotoxic classes have been identified in the E-group colicins: pore-forming ionophores such as colicin E1 (Cramer et al., 1983; Pattus et al., 1990), RNases such as colicins E3 and E6 (Bowman et al., 1971; Senior & Holland, 1971; Akutsu et al., 1989), and DNases such as colicins E2 (Schaller & Nomura, 1976), E7 (Chak et al., 1991), E8 (Toba et al., 1988), and E9 (Eaton & James, 1989). Colicins are produced by host organisms in concert with the production of a small immunity protein that inhibits their cytotoxic action. The immunity proteins are completely specific for their coordinately produced colicin despite their high sequence similarities (James et al., 1992). We are studying the mode of action and inhibition of the DNase class of E-type colicins. Concentrating on colicin E9 and its 86 amino acid immunity protein Im9,[1] we have shown that the specificity determining region of Im9 is located between residues 16 and 43 (Wallis et al., 1992a), and the

† This work was supported by the Wellcome Trust and by the U.K. Science and Engineering Research Council via their establishment of both the Leicester Biological NMR Centre and the UEA research groups as Molecular Recognition Centres and via providing research assistantships to M.J.O. and R.W.

* Author to whom correspondence should be addressed; Fax: 0603-259396; e-mail address: G.Moore@cpc865.uea.ac.uk.

‡ School of Chemical Sciences, University of East Anglia.

§ University of Leicester.

‖ School of Biological Sciences, University of East Anglia.

® Abstract published in *Advance ACS Abstracts*, September 1, 1994.

[1] Abbreviations: Im9, immunity protein of colicin E9; Im3, immunity protein of colicin E3; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; 2D, two dimensional; 3D, three dimensional; HOHAHA, homonuclear Hartman–Hahn; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; COSY, correlated spectroscopy; DQF-COSY, double-quantum filtered COSY; INEPT, insensitive nuclei enhanced by polarization transfer.

0006-2960/94/0433-12347$04.50/0

DNase activity of colicin E9 resides in its C-terminal domain composed of 128 amino acids (Curtis & James, 1991; Wallis et al., 1994). Through a combination of gene fusion, chemical modification, and site-directed mutagenesis of the DNase colicins and their immunity proteins, we aim to define the mechanisms of toxicity and immunity. This requires that the structures of the colicins and immunity proteins be determined. We have used NMR to provide structural information on these proteins since attempts to obtain crystals suitable for X-ray structural determination have not been successful.

Detailed high-resolution NMR studies of the DNase colicins and their immunity proteins have not previously been reported, although we have reported a preliminary description of the ^{1}H NMR characteristics of Im9 (James et al., 1992). NMR studies of the 84 amino acid immunity protein of the RNase colicin E3 (Yajima et al., 1992) have been reported, but this protein bears no sequence identity to Im9 and, as we show in the present paper, has a markedly different structure from that of Im9. Thus, the data reported herein, which include the almost complete ^{1}H NMR assignments and complete backbone peptide ^{15}N NMR assignments of Im9 together with the identification of its secondary structure elements, contribute the first set of structural information for this class of colicin DNase inhibitors. This structural information, together with biochemical data, leads to a model for the mediation of specificity in DNase colicin–immunity protein complexes, and this model is described briefly in the present paper. A preliminary report of this model has been published elsewhere (Kleanthous et al., 1994).

EXPERIMENTAL PROCEDURES

Sample Preparation. Im9 was expressed in *Escherichia coli* and purified as previously described (Wallis et al., 1992b). All samples prepared for NMR were assayed for biological activity as described by Wallis et al. (1992b) and found to be fully active. Uniformly ^{15}N-labeled Im9 was obtained by growing the *E. coli* JM105 cells containing the expression system in minimal medium with $^{15}NH_4Cl$ (1 g/L). Concentrations of NMR samples in either 90% ^1H_2O/10% ^2H_2O or 99.9% ^2H_2O were typically 1–4 mM. The NMR solutions were prepared from Im9 lyophilized from either ^1H_2O or ^2H_2O solutions containing sodium phosphate at pH 7.0 or 6.2 to yield NMR samples with a final phosphate concentration of 25–50 mM.

NMR Spectroscopy. Preliminary NMR experiments were carried out on a JEOL GX-400 NMR spectrometer as previously described (James et al., 1992), but the majority of the data reported in the present paper were acquired on Bruker AM500 and AMX600 spectrometers with ^{1}H frequencies of 500.13 and 600.13 MHz, respectively, and ^{15}N frequencies of 50.664 and 60.79 MHz, respectively. Spectra measured with Bruker instruments were recorded in the pure phase absorption mode by the time-proportional phase incrementation method (Redfield & Kunz, 1975; Marion & Wüthrich, 1983). Proton chemical shifts were measured from internal dioxane at 3.77 ppm. Nitrogen chemical shifts were referenced using external $^{15}NH_4NO_3$ (5 M $^{15}NH_4NO_3$ in 2 M HNO_3) with the ammonium ^{15}N resonance taken as 0 ppm. Typical 90° pulse widths were as follows: AM500 [^{1}H, 8 μs; ^{1}H (homonuclear spin lock) 15 μs; ^{15}N (hard), 18 μs; ^{15}N (GARP (Shaka et al., 1985) or WALTZ-16 (Shaka et al., 1983) decoupling), 85 μs]; AMX600 [^{1}H, 8.5 μs; ^{1}H (homonuclear spin lock) with MLEV17 (Bax & Davies, 1985) 15 μs and DIPSI-II (Shaka et al., 1988) 32 μs; ^{15}N (hard), 18 μs; ^{15}N (GARP or WALTZ-16 decoupling), 100 μs]. Standard pulse sequences and phase cycling were employed for the 2D homonuclear experiments.

For experiments requiring homonuclear Hartman–Hahn spin-lock sequences [2D HOHAHA (Braunschweiler & Ernst, 1983; Bax et al., 1985), and 3D ^{15}N–^{1}H TOCSY-HMQC (Zuiderweg & Fesik, 1989; Marion et al., 1989)] total mixing times of 60 and 80 ms were used. NOESY (Jeener et al., 1979; Macura et al., 1981) data were collected using mixing times of 100 and 150 ms.

Typical data set sizes for 2D homonuclear experiments were 512 increments of 2K complex data points. These yielded, after zero-filling, spectra with digital resolution of about 2–3 Hz per point in each dimension. Prior to Fourier transformation, either a sine-bell squared ($\pi/3$ phase shift) or a Lorentzian-to-Gaussian window function was applied in both F_2 and F_1 dimensions.

For measurements in ^1H_2O, either a low-power phase-locked presaturation of the water signal or a combination of a soft presaturation and a $1 - \bar{1}$ pulse was used, the latter being employed in cases where the sample concentrations were low ($\sim$1–2 mM). To implement the latter method in a 2D homonuclear HOHAHA experiment, a soft presaturation pulse was applied during the preparation period, and after the spin-lock period, a three-pulse sequence was used that included a flip-back 90° pulse as the read pulse (Lian et al., 1991). Compared with a normal presaturation experiment, the quality of a 2D spectrum obtained in this manner is significantly improved in the fingerprint region of the spectrum with a superior baseline.

2D ^{1}H–^{15}N HSQC spectra were acquired by the method of Bax et al. (1990). ^{15}N decoupling was carried out using either the GARP or WALTZ-16 sequence. Solvent suppression for the ^{1}H–^{15}N spectra was achieved by low-power presaturation. The ^{1}H–^{15}N J-coupling evolution and refocussing delays of the INEPT portions of the pulse sequences were set to 2.7 ms. Typical data set sizes were 512 increments of 2K or 4K complex points with spectral widths of 14 ppm in the ^{1}H dimension and 36 ppm in the ^{15}N dimension. The ^{1}H carrier was placed on the water signal, and the ^{15}N carrier was placed on 98 ppm.

All heteronuclear 3D experiments were performed on the AMX600 spectrometer at 298K on a 4 mM sample at pH 6.2. The heteronuclear 3D TOCSY-HMQC, NOESY-HMQC (Zuiderweg & Fesik, 1989; Marion et al., 1989), and HMQC-NOESY-HMQC (Frenkiel et al., 1990) pulse schemes used were similar to those reported. Typical data sizes for the 3D TOCSY-HMQC and 3D NOESY-HMQC experiments were 1024(t_3) by 70(t_2) by 200(t_1). The data set size for the HMQC-NOESY-HMQC experiment was 1024 by 64 by 64. For all experiments, 16 scans per t_1 increment were acquired. The NOESY mixing time for the 3D experiments was 150 ms.

Measurements of the $^3J_{NH\alpha}$ coupling constants were obtained using the HMQC-J experiment (Kay & Bax, 1990) with a data set size 2212 t_1 increments of 2048 data points; zero-filling in F_1 gave a final digital resolution of 0.75 Hz per point in the F_1 dimension.

2D ^{15}N–^{1}H HSQC spectra of uniformly ^{15}N-labeled protein in ^2H_2O were used to monitor the exchange of backbone amide protons. It was possible to obtain good quality 2D spectra on a 3 mM ^{15}N-labeled sample in 10 min (256 t_1 increments). The NH solvent exchange experiment commenced after the addition of 99.9% ^2H_2O at 273K to a sample of ^{15}N-labeled Im9 previously lyophilized from 50 mM potassium phosphate in ^1H_2O at pH 6.2. The sample was inserted in the magnet at 293K and allowed to come to thermal equilibrium, when the lock stabilized. The first 2D ^{15}N–^{1}H HSQC experiment

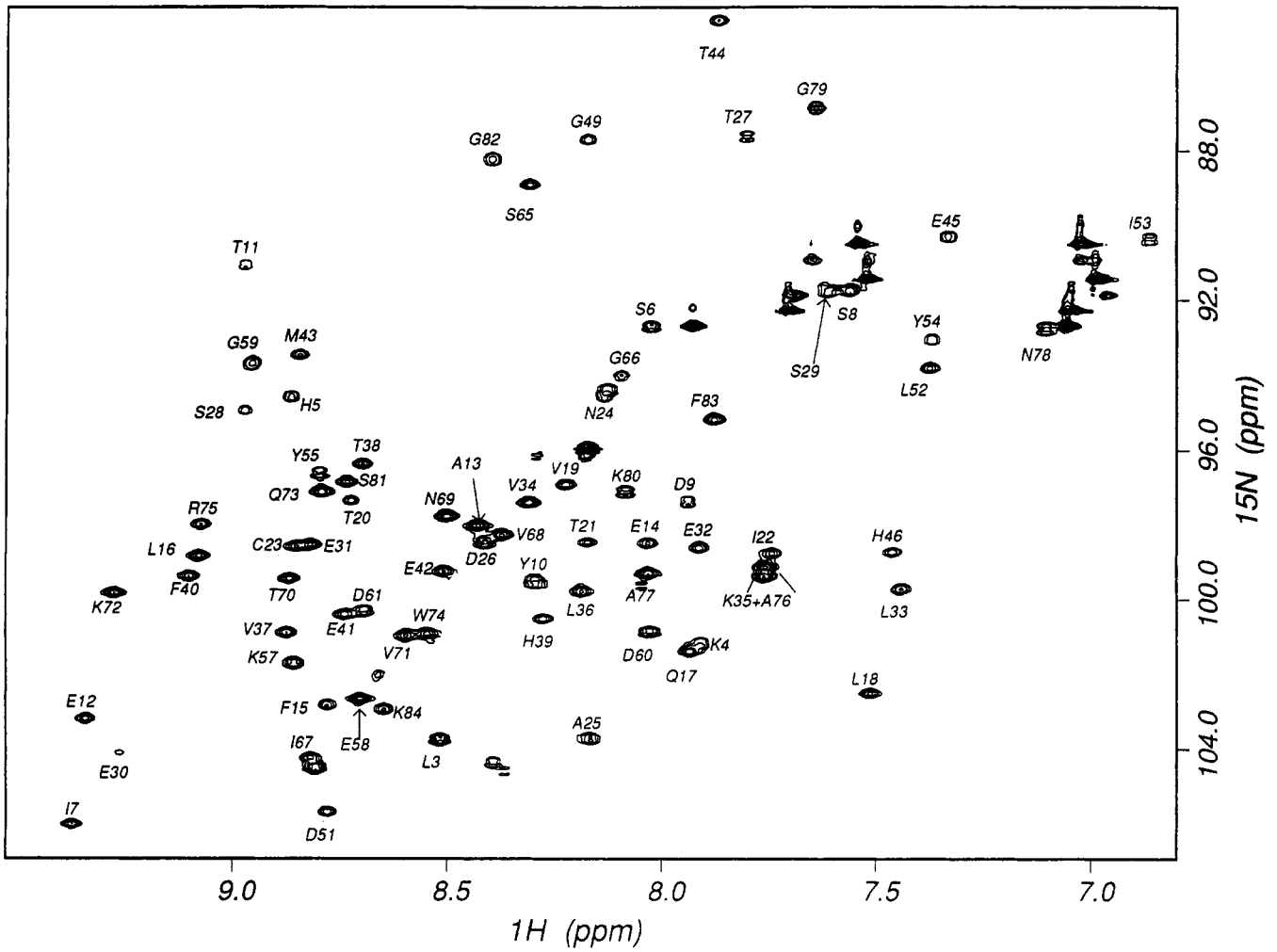

FIGURE 1: Selected region of the ^{15}N–^{1}H HMQC-J spectrum of uniformly ^{15}N-labeled Im9 recorded at pH 6.2 and 20 °C in a 3 mM 90% H_2O/10% D_2O solution. Assignments for backbone amide cross peaks are indicated.

was started 7 min after dissolution in ^2H_2O. A total of six spectra were recorded at the following times, measured from the addition of ^2H_2O to the lyophilized protein up to the midpoint of the acquisition time of each data set; 12, 24, 34 (all with two scans per t_1 increment), 50 (4 scans per t_1 increment), and 80 min (8 scans per t_1 increment).

2D spectra were processed on either the Bruker X32 data station using Bruker UXNMR software or on a Silicon Graphics Indigo R4000 workstation using FELIX (Hare research and BIOSYM). The 3D data sets were all processed on the Silicon Graphics Indigo R4000 workstation using FELIX. Typical processed 3D data sizes were $512(F_3)$ by $256(F_2)$ by $512(F_1)$. When it was necessary to improve the digital resolution a $1024(t_3) \times 512(t_2) \times 256(t_1)$ matrix was built and Fourier transformed in t_3 and t_2. t_1 fids were then loaded separately and zero-filled to 16K complex points prior to Fourier transformation, thus resulting in a series of 1D spectra. For all spectra, a sine-bell squared ($\pi/3$ shift) window function was applied.

RESULTS

Stability of Im9. Im9 was found to be appreciably soluble only above pH 6.0. At lower pH, solutions became turbid as the protein precipitated. At pH values of 6–7, samples of Im9 degraded if held for some hours at temperatures of 40 °C and above, and therefore most of the NMR data were collected at temperatures of 20–37 °C. A complicating factor for assignment was that, at pH values of 7.0, a number of amide proton resonances were not observed due to exchange with the bulk solvent. Accordingly, the majority of 3D experiments were performed at lower pH values (6.2) and temperatures (20 °C), so as to maximize the number of NH signals.

NMR Spectroscopy. Conventional 2D homonuclear ^1H NMR methods for identifying spin systems from scalar connectivities (COSY and TOCSY) and from dipolar connectivities (NOESY) produced limited information for the following reasons. First, there was severe overlap of peptide NH resonances. For example, 28 amide proton resonances have chemical shifts in the range 8.5–9.0 ppm of the ^1H NMR spectrum. Second, 30 of the 86 amino acids of Im9 were expected to exhibit AMX spin systems for their α/β-protons. These features, together with the presence of three prolines, considerably handicapped the identification of spin systems by the conventional sequential assignment procedure (Wüthrich, 1986). These problems were overcome by the use of HMQC experiments on uniformly enriched ^{15}N samples. Figure 1 shows an expanded region of the 2D ^1H–^{15}N HMQC-J spectrum of such a sample. A total of 78 of the expected 82 backbone amide connectivities are observed; a further two cross peaks occur at lower ^1H frequencies. Table 1 summarizes the assignments.

Sequential assignments are readily obtained by heteronuclear NMR, making use of the 3D ^1H–^{15}N HOHAHA-HMQC experiment to identify through-bond connectivities between peptide NH and side-chain protons and the 3D ^1H–^{15}N NOESY-HMQC experiment to identify through-space

Table 1: Assignments of Immunity Protein of Colicin E9 (Im9)[a]

residue	^{15}N	NH	αH	βH	γH	δH	others
Leu-3	103.92	8.52	4.55	1.55, 0.90	1.56		δCH_3 0.63, 0.55
Lys-4	101.11	7.92	4.35	1.20, 0.63	1.02	1.96, 1.30	ε 2.76
His-5	94.76	8.87	4.74	3.43			
Ser-6	92.89	8.03	4.75	3.95, 3.72			
Ile-7	106.18	9.37	4.12	1.65	0.90		γCH_3 0.49, δCH_3 –0.13
Ser-8	92.89	7.57	4.74	3.95			
Asp-9	97.51	7.94	4.79	2.83			
Tyr-10	99.74	8.30	5.01	3.59			
Thr-11	91.20	8.97	5.08	4.81	1.30		
Glu-12	103.41	9.34	2.45	1.86	2.05		
Ala-13	98.20	8.43	4.13	1.46			
Glu-14	98.77	8.03	4.16	2.79			2.56, 2.40
Phe-15	103.05	8.79	3.96	2.91, 2.74			
Leu-16	99.10	9.07	3.97	2.08	1.65, 1.47		δCH_3 0.93, 0.83
Gln-17	101.71	7.94	4.11	2.30	2.50		
Leu-18	102.68	7.52	4.07	1.81, 1.54	1.43		δCH_3 1.04, 0.68
Val-19	97.09	8.22	3.17	1.93			γCH_3 0.88, 0.77
Thr-20	97.47	8.73	4.35	3.56	1.22		γCH_3 0.88, 0.77
Thr-20	97.47	8.73	4.35	3.56	1.22		γCH_3 0.88, 0.77
Thr-21	98.57	8.17	3.89	4.50	1.13		
Ile-22	98.87	7.74	3.46	1.76	1.85		γCH_3 0.72 δCH_3 0.56
Cys-23	98.75	8.86	3.83	2.70			
Asn-24	94.66	8.14	4.67	3.00, 2.74			
Ala-25	103.92	8.15	3.92	1.52			
Asp-26	98.83	8.42	4.93	2.93, 2.46			
Thr-27	87.77	7.79	4.61	4.48	0.97		
Ser-28	95.12	8.95	4.49	4.03			
Ser-29	92.07	7.63	4.81	4.30, 4.06			
Glu-30	104.22	9.32	4.14	2.15	2.41		
Glu-31	98.75	8.83	4.08	2.01, 2.16	2.40		
Glu-32	98.77	7.93	4.10	2.16, 2.25	2.40		
Leu-33	99.93	7.48	4.12	2.20	1.75		δCH_3 1.04
Val-34	97.56	8.30	3.70	2.20			γCH_3 1.13, 1.02
Lys-35	99.71	7.75	4.14	1.68	1.75, 1.43	2.02	ε 2.96
Leu-36	100.02	8.19	4.21	2.16, 1.44	2.05		δCH_3 0.87
Val-37	101.13	8.87	3.59	2.39			γCH_3 1.22, 1.18
Thr-38	96.51	8.70	4.22	4.36	1.39		
His-39	100.67	8.27	4.62	3.70			
Phe-40	99.52	9.10	3.92	3.84			
Glu-41	100.55	8.74	3.67				3.06, 2.71, 2.15
Glu-42	99.45	8.51	4.05	2.11, 2.29	2.68		
Met-43	93.58	8.84	4.52				2.10, 1.47, ε 2.26
Thr-44	84.63	7.87	3.56				
Glu-45	90.42	7.34	3.90	2.50	2.28, 2.21		
His-46	98.85	7.48	3.01	2.69, 2.18			
Ser-48	97.31	11.07	4.59	3.94, 3.86			
Gly-49	87.80	8.17	4.15, 3.89				
Ser-50	99.55	10.07	4.13				
Asp-51	105.88	8.77	4.62	3.25, 2.88			
Leu-52	93.93	7.39	3.89	1.65, 1.37	1.52		δCH_3 0.50, 0.35
Ile-53	90.48	6.88	3.54	1.09	0.87, 0.35		δCH_3 –0.15 γCH_3 –0.08
Tyr-54	93.29	7.36	4.22	2.65, 2.57			
Tyr-55	96.89	8.79	5.02	3.19, 2.59			
Lys-57	101.95	8.86	4.31	1.74, 1.55		1.81	ε 3.04
Glu-58	102.85	8.70	4.10				2.34, 2.06
Gly-59	93.85	8.96	4.33, 3.75				
Asp-60	101.00	8.03	4.71	2.90, 2.64			
Asp-61	100.43	8.69	4.71	2.97, 2.72			
Ser-65	89.02	8.31	4.05	3.90			
Gly-66	94.08	8.10	3.90				
Ile-67	104.74	8.82	3.90	1.88	1.73		γCH_3 0.86, δCH_3, 0.45
Val-68	98.40	8.38	3.66	2.41			γCH_3 1.30, 1.26
Asn-69	97.96	8.50	4.63	3.09, 3.01			
Thr-70	99.59	8.87	4.07	4.55	1.23		
Val-71	101.16	8.61	3.66	2.55			γCH_3 1.14, 0.77
Lys-72	100.03	9.27	3.97	1.91	1.78	2.16	ε 3.13
Gln-73	97.35	8.79	4.25				2.73, 2.59, 2.32
Trp-74	101.10	8.55	4.26	3.68, 3.56			
Arg-75	98.17	9.07	3.56				
Ala-76	99.66	7.76	4.19	1.40			
Ala-77	99.48	8.03	4.21	1.46			
Asn-78	93.02	7.12	4.57	2.34, 1.33			
Gly-79	86.98	7.64	3.88				
Lys-80	97.28	8.09	4.58	1.66, 1.45	1.11	1.28, 1.99	ε 3.17, 2.87
Ser-81	97.06	8.74		3.98			
Gly-82	88.32	8.40	4.30, 3.59				
Phe-83	95.30	7.88	4.70	2.71			
Lys-84	103.11	8.65	4.25	1.70	1.48	1.93	ε 3.00

[a] Chemical shifts are in ppm referenced to external $^{15}NH_4NO_3$ at 0 ppm ^{15}N and to internal dioxane at 3.74 ppm for ^1H. Uncertainties are ±0.1 ppm, ^{15}N, and ±0.02 ppm, ^1H. Values are quoted for Im9 at 20 °C and pH 6.2.

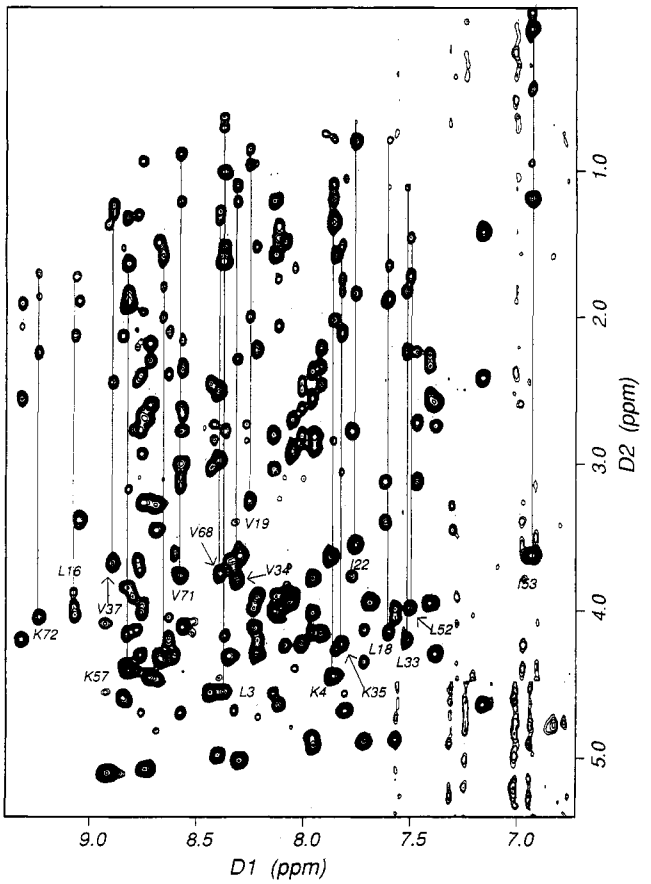

FIGURE 2: NH (D1 axis) aliphatic (D2 axis) region of the TOCSY spectrum (80-ms mixing time) in H_2O at pH 7.0 and 37 °C. Spin systems of various valine, leucine, isoleucine, and lysine residues are shown.

connectivities (Zuiderweg & Fesik, 1989; Marion et al., 1989). The 3D experiments overcome problems of overlap by spreading signals according to the frequency of the directly bound ^{15}N chemical shift. However, the HMQC-NOESY experiment does not afford sequential amide proton NOE connectivities if the proton chemical shifts are degenerate. In such cases, the HMQC-NOESY-HMQC experiment was used to alleviate this degeneracy (Frenkiel et al., 1990).

Identification of Spin Systems. Identification of spin systems was initiated by analysis of the 600-MHz 2D ^1H–^1H TOCSY spectrum of Im9 in 90% ^1H_2O/10% ^2H_2O (Figure 2). DQF-COSY spectra of Im9 in 90% ^1H_2O/10% ^2H_2O and in 99.9% ^2H_2O were used to distinguish between short-range coupled peaks (2-bond and 3-bond) and long-range coupled peaks in the TOCSY spectra. For some spin systems, this analysis was sufficient to identify them fully. For example, the five valine spin systems were readily identified in the TOCSY spectrum (Figure 2). However, for the longer chain aliphatic residues this approach was insufficient. Only one leucine spin system (Leu3) could be easily identified from the 2D TOCSY spectrum (Figure 2). For the remaining leucine residues, TOCSY experiments performed with mixing times up to 80 ms failed to show complete coupling patterns for these residues from the amide resonance. It was therefore necessary to bridge connectivities using the DQF-COSY spectrum, in particular making use of γCH and δCH_3 correlations. Using these complementary spectra, it was possible to identify five of the six leucine spin systems and two of the four isoleucine spin systems.

Identification of all the glycine, alanine, and threonine spin systems was also not straight-forward using 2D TOCSY and DQF-COSY spectra. This was partly because of resonance overlap as described previously, but mainly it resulted from complications arising from the degeneracy of resonances. These complications were partially resolved by the use of a 3D ^1H–^{15}N HMQC-TOCSY spectrum with uniformly ^{15}N-labeled Im9. Four of the eight threonine spin systems were readily identified from this spectrum, and assignment of the α- and β-protons was afforded by the DQF-COSY spectrum. Six alanine-like spin systems were observed in the 3D spectrum. These arise from the four alanine residues and two of the remaining three threonine residues.

There are eight amino acid residues which give rise to AMX spin systems corresponding to connectivities between the α-proton and two β-protons. These are serine, cysteine, aspartate, asparagine, phenylalanine, tyrosine, histidine, and tryptophan. When the β-protons are degenerate, the amino acid appears as an AX_2 spin system. For Im9, 30 amino acids were expected to show this spin pattern, and a total of 42 AMX or AX_2 spin systems were identified, of which two were assigned to asparagine residues based on NOE connectivities between the side-chain amide and β-protons. The excess number of this type of spin system is a consequence of amino acid residues whose coupling patterns do not behave in a straight-forward manner. These include glycine, glutamate, and glutamine residues.

Scalar coupling experiments were used to assign the aromatic ring protons of the three histidine and sole tryptophan residues. NOESY experiments in ^2H_2O were used to connect the ring and side-chain spin systems. Assignment of the remaining tyrosine and phenylalanine residues was complicated by the absence of the characteristic coupling patterns of the phenylalanine residues, which gave rise to tyrosine-like spin systems. It was not possible to use the one-dimensional spectra to alleviate this problem due to substantial overlap in the regions of interest. Accordingly, assignment of these residues was postponed to the sequential assignment analysis.

Sequential Assignments. Many of the glutamate, lysine, arginine, asparagine, glycine, and aromatic spin systems were identified from the 3D spectra in combination with the 2D TOCSY and DQF-COSY spectra. However, only 39 spin systems were assigned to particular types of amino acids without sequence-dependent NMR data.

The sequence-specific assignment of the Im9 spectrum relied heavily on the observation of $d_{NN(i,i+1)}$, $d_{\alpha N(i,i+1)}$, and $d_{\beta N(i,i+1)}$ NOE connectivities in the 3D ^1H–^{15}N HMQC-NOESY spectrum of ^{15}N-labeled Im9. This stage of the analysis was made difficult by the large number of AMX and AX_2 spin systems not assigned to amino acid type. However, sufficient unique sequences with identified spin systems were present to provide useful starting points for the analysis. Additionally, the sequential assignment procedure afforded identification of the previously unassigned aromatic residues.

A large sequential stretch of $d_{\alpha N(i,i+1)}$, $d_{NN(i,i+1)}$, $d_{\beta N(i,i+1)}$ connectivities were observed incorporating residues 30–46 (Figure 3). Assignment of this stretch was initiated by observation of NOEs between the characteristic spin systems of leucine and valine. This dipeptide fragment occurs three times in the entire polypeptide sequence (Leu18–Val19, Leu33–Val34, and Leu36–Val37). Assignment to Leu33–Val34 was afforded by identification of NOE connectivities between the leucine and a glutamate/glutamine spin system (thus eliminating the Leu36–Val37 pair) and then connectivities from this spin system to a further glutamate–glutamine residue (thereby eliminating the Leu18–Val19 pair). Assignment of Lys35, whose amide proton and ^{15}N chemical

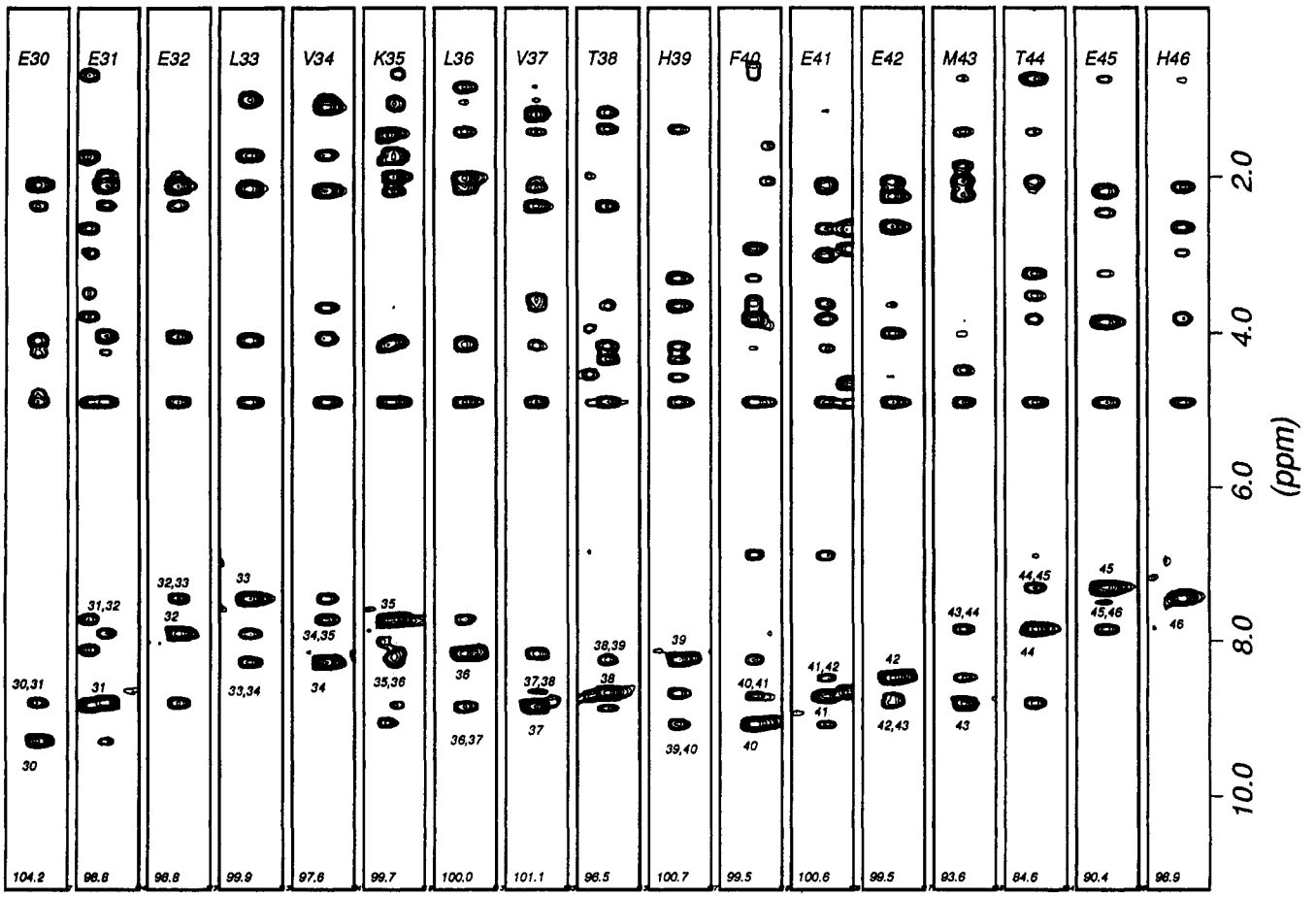

FIGURE 3: ^{1}H–^{1}H strips taken from the 3D ^{1}H–^{15}N NOESY-HMQC spectrum of uniformly ^{15}N-labeled Im9 at pH 6.2 and 20 °C in H_2O. Slices were taken at ^{15}N chemical shifts corresponding to the backbone amide resonances of residues E30–H46. Diagonal and sequential amide proton connectivities are shown.

shifts are degenerate with the Ala76 spin system was afforded by the identification of sequential NOEs to this residue from Val34. Throughout this polypeptide region, all but four of the residues exhibited their characteristic spin systems as identified from COSY and TOCSY spectra.

Sequential assignments for the sequence of 20 residues, running from Ser65 to Lys84, was obtained from $d_{NN(i,i+1)}$ connectivities starting from the unique tryptophan residue at position 74. Confirmation of the sequential assignments was obtained from the observation that nearly all the residues display their characteristic spin systems in *J*-correlation spectra. The spin systems of three residues (Lys80, Arg75, and Ile67) did not exhibit complete connectivities in the TOCSY and COSY spectra and were assigned based on sequential NOE connectivities to residues on either side of them, which afforded a unique tripeptide fragment in each case. The identity of Asn78 was confirmed by observation of NOE connectivities between its side-chain amide and β-protons. The first stage of assignment had tentatively identified three glycine spin systems which show only one NH–CαH connectivity. Two of these spin systems are assigned to Gly66 and Gly79 based on the above sequential assignments.

Sequential NOEs observed between the backbone atoms of isoleucine and threonine spin systems afforded assignment of the unique dipeptide Thr21–Ile22 and initiated further sequential assignments. Assignment of residues 12–25 was obtained on the basis of strong sequential $d_{NN(i,i+1)}$ connectivities and the observation of a number of $d_{\alpha N(i,i+1)}$ and $d_{\beta N(i,i+1)}$ NOEs within this segment.

The presence of a strong $d_{\alpha N(i,i+1)}$ NOE between a threonine spin system and the previously assigned Glu12 initiated sequential assignment in the region Leu3 to Thr11. A further strong $d_{\alpha N(i,i+1)}$ connectivity between Thr11 and an AX_2 spin system yielded assignment of Tyr10—the presence of NOEs between the aromatic protons of Tyr10 and the side-chain protons of Thr11 is further confirmation of this assignment. From Tyr10 to Ile7, sequential assignments were based initially on the presence of strong $d_{NN(i,i+1)}$ NOEs and substantiated by the observation of further sequential NOEs. Between residues Ile7 and Leu3, assignment was afforded by the observation of one or more sequential NOE connectivities. Throughout this region, residues show characteristic spin systems apart from Ile7 for which no connectivities from the α proton could be observed in COSY and TOCSY spectra.

The segment of amino acids flanked by two proline residues stretching from 48 to 56 was assigned after the identification of sequential NOEs between an isoleucine and leucine spin system, which forms the unique dipeptide fragment Leu52–Ile53. Sequential assignment from Leu52 to Ser48 was relatively straight-forward, with only Ser50 exhibiting an unusual spin system (it was not possible to detect connectivities from the α- to β-protons) in the TOCSY and COSY spectra. Due to chemical shift degeneracy of the amide protons of Asp51 with Tyr55 and Leu52 with Tyr54, unambiguous identification of sequential $d_{NN(i,i+1)}$ connectivities for residues 52–55 required the 3D HMQC-NOESY-HMQC experiment.

Residues 57–61 were assigned based on the observation of strong sequential $d_{\alpha N(i,i+1)}$ connectivities and weak $d_{NN(i,i+1)}$ NOEs. Assignments were initiated by observation of NOEs to and from the remaining lysine spin system that is otherwise unassigned. No NOE connectivities could be observed past Asp61.

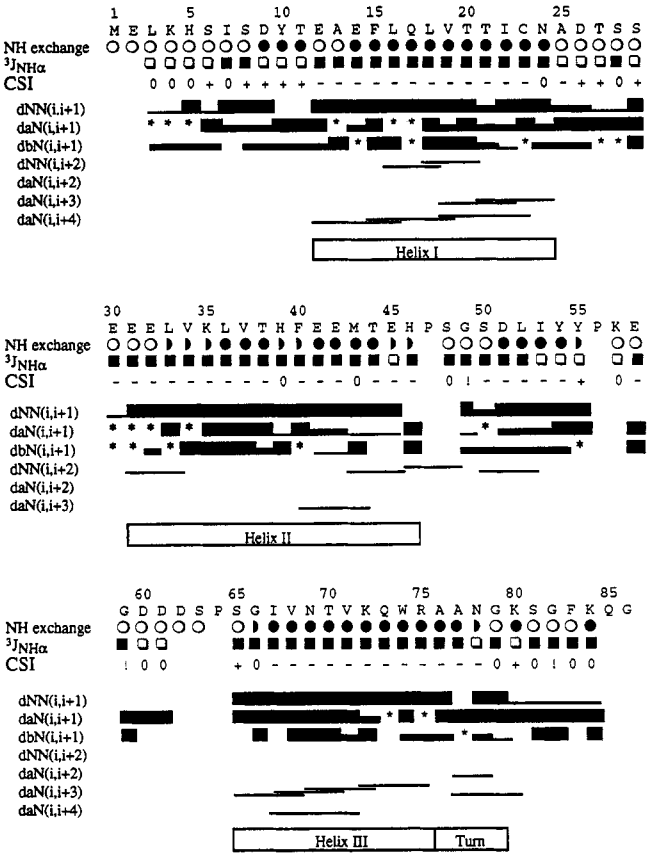

FIGURE 4: Summary of sequential and medium-range NOE connectivities observed in the Im9 protein. The intensities of NOEs are categorized as strong, medium, and weak and are represented by the thickness of the lines. Filled circles below the amino acid sequence (single letter code) represent slowly exchanging amide protons that were observed in the 2D ^{15}N–^1H HSQC spectrum recorded 80 min after dissolving the protein in D_2O. Filled semicircles and open circles indicate intermediate and fast exchanging amide protons, respectively. Asterisks represent sequential connectivities that could not be observed due to the degeneracy of α-proton chemical shifts. Filled and open boxes indicate respectively $^3J_{NH_\alpha}$ coupling constants of less than 6 Hz or larger than 8 Hz. CSI refers to the chemical shift index (see text); (+) – and (0) refer to αH chemical shifts which are greater than, less than, and within 0.10 ppm of the random coil value, respectively; (!) indicates Gly residues whose αH resonances are not degenerate and that exhibited shifts both greater and less than its random coil value.

Assignment of residues 26–29 was initiated by the presence of a strong $d_{\alpha N(i,i+1)}$ connectivity between an AMX spin system and the previously assigned Glu30 spin system. With the exception of Thr27, which exhibits an alanine-like spin system, characteristic spin systems were shown by these residues. The complete set of assignments obtained in this work are given in Table 1.

Identification of Secondary Structure. A description of the secondary structure of Im9 together with a summary of the NMR parameters identifying the secondary structure is given in Figure 4. Numerous studies (Wüthrich, 1986; Grzesiek et al., 1992; Ikura et al., 1991; Omichinski et al., 1992; Stockman et al., 1992; Weir et al., 1993) have established that a qualitative interpretation of NOEs, $^3J_{NH\alpha}$ coupling constants, and hydrogen exchange rates accurately define the location of regular secondary structure elements. We derived NOE data from homo- and heteronuclear 2D and 3D NOESY spectra with mixing times of 150 ms. Approximate $^3J_{NH\alpha}$ values were extracted from HMQC-J experiments. Hydrogen exchange rates were obtained as described in the methods section. Additionally, Wishart et al. (1992) have devised a method known as the chemical shift index that allows the identity, location, and extent of regular secondary structure elements to be obtained by inspection of the α-proton chemical shifts. Sequential α-protons with chemical shifts more than 0.10 ppm lower than the random coil value are designated as being part of an α-helix, whereas sequential α-proton chemical shifts more than 0.1 ppm greater than the random coil value are designated as β-sheet. α-protons exhibiting shifts within 0.1 ppm of their random coil values are designated as coil. The chemical shift index for Im9 is given in Figure 4 and essentially concurs with NOE, coupling constant, and amide exchange data that shows the structure to be comprised of three helices spanning residues Glu12–Ala25, Glu30–His46, and Ser65–Ala76.

A major problem with the secondary structure analysis was the relatively poor dispersion of α-proton resonances which prevented a thorough interpretation of the helical regions. In particular, it was difficult to unambiguously assign a sufficient number of connectivities which differentiate α-helices and 3_{10}-helices. We have therefore assigned these regions as helices only.

The sequential NOE connectivities for residues 12–25 indicate that this segment is helical. Despite the degeneracy of a number of α-proton resonances for this segment, five $d_{\alpha N(i,i+3)}$ and $d_{\alpha N(i,i+4)}$ NOE connectivities were observed, consistent with a helix. The continuous stretches of strong $d_{NN(i,i+1)}$ NOE connectivities for residues 30–46 and for residues 65–76 indicate that these regions are helical, consistent with the relatively small $^3J_{NH\alpha}$ coupling constants and slowly exchanging amide protons for residues toward the center of the two helical regions. The large number of $d_{\alpha N(i,i+1)}$, $d_{\beta N(i,i+1)}$, and $d_{\alpha N(i,i+3)}$ connectivities are consistent with this.

The conformation of other regions of the protein is not so clear. The strong sequential interresidue $d_{\alpha N}$ connectivities plus the large $^3J_{NH\alpha}$ coupling constant value for Thr11 and Tyr10 indicate that these residues have an extended conformation. For residues 3–8, the presence of large $^3J_{NH\alpha}$ coupling constants also suggests an extended conformation, and the fast rates of amide exchange is consistent with these residues being removed from the protein interior. Slow exchange for residues Asp9 to Thr11 suggests that regular structure commences at this point of the sequence and is substantiated by the observation of an NOE connectivity between Thr11 and Lys84. Residues 26–29 probably form a loop between the 12–25 and 30–46 helices, consistent with the relatively large coupling constants and variable strength $d_{NN(i,i+1)}$ NOEs. NOE data for residues 48–55 suggest that a small helix may be present in this region, although coupling constant and amide exchange data (Figure 4) show no obvious patterns. Nevertheless, the presence of slowly exchanging amide protons for residues Asp51–Tyr55 indicates that these residues may be buried within the protein surface and/or are involved in hydrogen bonding networks.

There is evidence that residues Ala76–Gly79, which connect the third helix and C-terminal segments, form a type I turn. This is characterized by the presence of a $d_{\alpha N(i,i+2)}$ connectivity between residues Ala77 and Gly79, a $d_{\alpha N(i,i+3)}$ NOE between Ala77 and Lys80, and a large $^3J_{NH\alpha}$ coupling constant exhibited by Asn78.

Residues 80–84 show weak $d_{NN(i,i+1)}$ connectivities, but strong sequential $d_{\alpha N}$ connectivities, which is characteristic of an extended conformation. The interresidue $d_{\alpha N}$ NOEs are stronger than intraresidue $d_{\alpha N}$ NOEs, confirming the extended conformation. Additionally, residues Lys80 and Gly82 exhibit large $^3J_{NH\alpha}$ coupling constants (>8 Hz). Extended backbone conformation is often associated with a

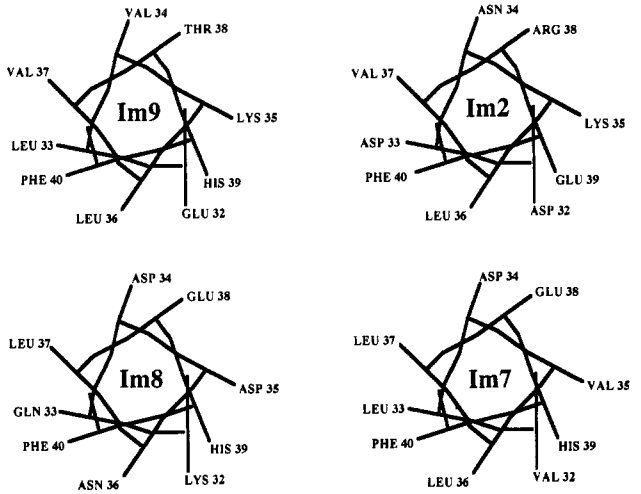

FIGURE 5: Helical wheel representations of residues 32–40 for all the known nuclease-specific immunity proteins of the E-group colicins: Im9, Im2, Im8, and Im7. Numbering scheme is according to alignment with the Im9 sequence.

β-sheet, characterized by typical interstrand NOEs, but no such NOEs could be observed involving these residues. The amide exchange rate for residues 80–84 might be expected to be fast considering their position toward the end of the C-terminus; however, both amides of Lys80 and Lys84 exhibit relatively slow rates of exchange, indicating they may be involved in hydrogen bonding.

DISCUSSION

We have obtained an almost complete set of ^{1}H NMR and peptide ^{15}N NMR assignments for the colicin E9 immunity protein (Table 1) and determined the location of regular secondary structure elements (Figure 4). Overall, Im9 contains three helical segments, which comprise ~50% of its structure. This is very different from the structure of the 84 amino acid immunity protein to the RNase colicin E3 (Im3) proposed by Yajima et al. (1992, 1993) based on ^{1}H and ^{15}N NMR data. The lack of apparent structural similarity is consistent with amino acid sequence comparisons which show no homology between Im3 and Im9 or the cytotoxic domains of colicins E3 and E9, even though other domains of the colicins are highly homologous.

The specificity determining region of the DNase immunity proteins, Im8 and Im9, have been identified by gene fusion, mutagenesis, and chemical modification experiments to be located in the region 16–43 (Wallis et al., 1992a). Residues 32–40, which form part of a helical structure in Im9, appear to be particularly important. We assume that all four DNase-specific immunity proteins (Im2, Im7, Im8, and Im9) have similar structures based on the sequence alignment reported by James et al. (1987) and Chak et al. (1991), which shows them to be ~60% homologous. Helical representations of residues 32–40 (Figure 5) are thus revealing. Such helices are amphipathic with a strongly polar side. All but Im2 have an extensive nonpolar surface, and depending upon the exact position of the Asp33 side chain of Im2, its helix may also behave with strong characteristic amphipathic properties. These include interactions of the nonpolar surface of the helix with the interior core of the protein and the exposure of the polar part of the helix to the solvent. This implies that the specificity determining groups include residues 34, 35, and 38, which is consistent with mutagenesis studies of Im9. In these Val34 was replaced by Asp34, with the result that some immunity to colicin E8 was produced in the variant which also retained its colicin E9 immunity. Determination of the 3D structure of Im9 and other DNase immunity proteins together with NMR studies of the colicin–immunity protein complexes will allow the model for specificity determination outlined above to be explored.

We note with interest that the recently reported X-ray structure of the complex formed by the bacterial ribonuclease barnase and its 89 amino acid inhibitor barstar shows that an α-helix of the inhibitor plays an important part in the protein–protein interaction (Guillet et al., 1993), a situation similar to the one we propose for the colicin E9 system. In addition, the predominant secondary structure in both barstar and Im9 is the α-helix, although barstar also contains short segments of the β-sheet. This observation is even more intriguing when it is put in the context of another RNase inhibitor, the immunity protein of colicin E3 (Im3) which is predominantly β-sheet (Yajima et al., 1992, 1993). Thus it seems that, at least in terms of secondary structure elements, barstar is more similar to Im9, a DNase inhibitor, than it is to Im3, another RNase inhibitor.

REFERENCES

Akutsu, A., Masaki, H., & Ohta, T. (1989) *J. Bacteriol. 171*, 6430–6436.
Bax, A., & Davies, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Bax, A., Davies, D. G., & Sarkar, S. K. (1985) *J. Magn. Reson. 63*, 230–234.
Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990) *J. Magn. Reson. 86*, 304–318.
Bodenhausen, G., & Reuben, D. J. (1980) *Chem. Phys. Lett. 69*, 185–189.
Bowman, C. M., Dahlaberg, J. E., Ikemura, T., Konisky, J., & Nomura, M. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 964–968.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Chak, K.-F., Kuo, W.-S., Lu, F.-M., & James, R. (1991) *J. Gen. Microbiol. 137*, 91–100.
Cramer, W. A., Dankert, J. R., & Uratani, Y. (1983) *Biochim. Biophys. Acta 737*, 173–179.
Curtis, M. D., & James, R. (1991) *Mol. Microbiol. 5*, 2727–2733.
DiMasi, D. R., White, D. C., Schnaitman, C. A., & Bradbeer, C. (1973) *J. Bacteriol. 115*, 506–513.
Eaton, T., & James, R. (1989) *Nucleic Acids Res. 17*, 1761.
Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B., & Feeney, J. (1990) *J. Magn. Reson. 90*, 420–425.
Grzesiek, S., Döbeli, H., Gentz, R., Garotta, G., Labhart, A. M., & Bax, A. (1992) *Biochemistry 31*, 8180–8190.
Guillet, V., Lapthorn, A., Hartley, R. W., & Maugen, Y. (1993) *Structure 1*, 165–176.
Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M., & Bax, A. (1991) *Biochemistry 30*, 9216–9228.
James, R., Curtis, M. D., Wallis, R., Osborne, M. J., Kleanthous, C., & Moore, G. R. (1992) in *Bacteriocins, microcins & lantibiotics* (James, R., Lazdunski, C., & Pattus, F., Eds.) pp 181–201, NATO ASI Series H, Springer, Heidelberg.
James, R., Jarvis, M., & Barker, D. F. (1987) *J. Gen. Microbiol. 133*, 1553–1562.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110–126.
Kleanthous, C., Wallis, R., Reilly, A., Osborne, M. J., Lian, L.-Y., Moore, G. R., & James, R. (1994) in *Bacterial Protein Toxins* (Freer et al., Eds.) Zbl. Bakt. Suppl. 24, pp 230–240, Gustav Fischer, Stüttgart.

Lian, L. Y., Yang, J. C., Derrick, J. P., Sutcliffe, M. J., Roberts, G. C. K., Murphy, J. P., Goward, C. R., & Atkinson, T. (1991) *Biochemistry 30*, 5335–5340.

Luria, S. E., & Suit, J. L. (1987) in *Escherichia coli & Salmonella typhimurium, cellular & molecular biology* (Neidhardt, F. C., Ed.) Vol. 2, pp 1615–1624, American Society of Microbiology, Washington, DC.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.

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

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 6150–6156.

Omichinski, J. G., Clore, G. M., Robien, M., Sakaguchi, K., Appella, E., & Gronenborn, A. M. (1992) *Biochemistry 31*, 3907–3917.

Pattus, F., Massotte, H. U., Wilmsen, J., Lakey, D., Tsernoglou, A., Tucker, A., & Parker, M. W. (1990) *Experientia 46*, 180–192.

Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.

Schaller, K., & Nomura, M. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3989–3993.

Senior, B. W., & Holland, I. B. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 959–963.

Shaka, A. J., Keeler, J., Frenkiel, T., & Freeman, R. J. (1983) *J. Magn. Reson. 52*, 335–338.

Shaka, A. J., Barker, O. B., & Freeman, R. J. (1985) *J. Magn. Reson. 64*, 547–552.

Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–293.

Stockman, B. J., Scahill, T. A., Roy, M., Ulrich, E. L., Strakalaitis, N. A., Brunner, D. P., Yem, A. W., & Deibel, M. R., Jr. (1992) *Biochemistry 31*, 5237–5245.

Toba, M., Masaki, H., & Ohta, T. (1988) *J. Bacteriol. 170*, 3237–3242.

Wallis, R., Moore, G. R., Kleanthous, C., & James, R. (1992a) *Eur. J. Biochem. 210*, 923–930.

Wallis, R., Reilly, A., Rowe, A., Moore, G. R., James, R., & Kleanthous, C. (1992b) *Eur. J. Biochem. 207*, 687–695.

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

Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D., & Thomas, J. O. (1993) *EMBO 12*, 1311–1319.

Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647–1651.

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

Yajima, S., Muto, Y., Yokoyama, S., Masaki, H., & Uozumi, T. (1992) *Biochemistry 32*, 5578–5586.

Yajima, S., Muto, Y., Morikawa, S., Nakamura, H., Yokoyama, S., Masaki, H., & Uozumi, T. (1993) *FEBS Lett. 333*, 257–260.

Zuiderweg, E. R. P., & Fesik, S. (1989) *Biochemistry 28*, 2387–2391.