Journal of Biomolecular NMR, 1 (1991) 191–204
ESCOM

J-Bio NMR 019

Assignments of ^{1}H-^{15}N magnetic resonances and identification of secondary structure elements of the λ-cro repressor

H. Matsuo[a], M. Shirakawa[a], T. Ohkubo[b], T. Yamazaki[c] and Y. Kyogoku[a,*]

[a] *Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan*
[b] *Protein Engineering Research Institute, Suita, Osaka 565, Japan*
[c] *Biometrology Laboratory, JEOL Ltd., Akishima, Tokyo 196, Japan*

Received 6 February 1991
Accepted 15 April 1991

Keywords: Assignment; λ-cro; SQC; HMQC; 3D NMR; Isotope labeling; Secondary structure

SUMMARY

The assignments of ^{1}H-^{15}N magnetic resonances of the λ-cro repressor are presented. Individual ^{15}N-amino acids were incorporated into the protein, or it was uniformly labeled with ^{15}N. For the ^{13}C-^{15}N double-labeling experiments, ^{13}C-amino acids were incorporated into the uniformly ^{15}N-labeled protein. All the amide ^{1}H-^{15}N resonances could be assigned with such specific labeling, and sequential connectivities obtained by two-dimensional (2D) ^{1}H-^{15}N reverse correlation spectroscopies and three-dimensional (3D) $^{1}H/^{15}N$ NOESY-HMQC spectroscopy. Conventional 2D ^{1}H-^{1}H correlation spectroscopies were applied to the assignment of the side-chain protons. Some of the ^{1}H resonance assignments are inconsistent with those previously reported [Weber, P.L., Wemmer, D.E. and Reid, B.R. (1985) *Biochemistry*, **24**, 4553–4562]. The sequential NOE connectivities and H–D exchange rates indicate several elements of the secondary structure, including α-helices consisting of residues 8–15, 19–25 and 28–37, and three extended strands consisting of residues 4–7, 39–45 and 49–55. Based on several long-range NOEs, the three extended strands could be combined to form an antiparallel β-sheet. The amide proton resonances of the C-terminal residues except Ala^{66} (residues 60–65) were hardly observed at neutral pH, indicating that the arm is flexible. The identified secondary structure elements in solution show good agreement with those in the crystal structure of the cro protein [Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. (1981) *Nature*, **290**, 754–758].

INTRODUCTION

The λ-cro protein is the repressor protein which is responsible for the lytic growth of bacteriophage λ. It binds to the right (O_R) and left operator (O_L) regions of the λ phage genome competi-

* To whom correspondence should be addressed.

tively with the cI repressor (Ptashne et al., 1980). The cro protein consists of 66 amino acids (Hsiang et al., 1977) and is thought to be active as a dimer (dimer MW = 14 700). The crystal structure of the cro protein has been determined (Anderson et al., 1981). Based on the structure, a binding model for the cro protein with operator DNA O_R3 was proposed (Ohlendorf et al., 1982). In this model, the so-called α3 helical region from Gln^{27} to Ala^{36} is assumed to fit in the major groove of the operator DNA in the B form. A recent X-ray study on the cro-O_R3 complex demonstrated that the helical region fits in the major groove and the cro dimer undergoes a substantial conformational change, though the details were not revealed on low-resolution analysis (Brennan et al., 1990).

In our previous photo CIDNP NMR study of the interaction between the cro protein and O_R3 we deduced that the α helix part was unequivocally involved in the interaction, and that distortion of the cro dimer took place on binding to the operator DNA (Shirakawa et al., 1985a,b; Lee et al., 1987). In that experiment, we assigned the side-chain protons of tyrosine residues by means of nitration followed by partial digestion. The sequential resonance assignments of most of the ^1H resonances of the cro protein have been reported (Weber et al., 1985). Although the specific deuteration at Leu and Val residues partly confirmed their assignments, we found that some of them were inconsistent with our data (Shirakawa et al., 1987, 1991a). Leighton and Lu (1987) also raised doubt as to the assignments of the amide proton resonances of Lys residues made on the basis of the results of $^1H-^{15}N$ correlation experiments on ^{15}N-lysine-labeled cro. The strategy used for NMR signal assignments has changed recently. The combination of stable isotope labeling and multi-dimensional NMR spectroscopy enables unambiguous assignments due to the well-resolved resonance signals (Fesik and Zeiderweg, 1988; Marion et al., 1989; Ikura et al., 1990; Nagayama et al., 1990).

In this experiment, we made $^1H-^{15}N$ resonance assignments of the cro protein in a less acidic solution. At first, we performed cloning and overproduction of the protein. The protein was specifically labeled with ^{15}N amino acids, or doubly labeled with ^{15}N and ^{13}C. $^1H-^{15}N$ heteronuclear multiple-quantum coherence (HMQC; Müller, 1979; Bax et al., 1983; Redfield, 1983) and single quantum correlation (SQC; Bodenhausen and Ruben, 1980) spectra of the specifically ^{15}N- and/or ^{13}C-labeled protein allowed unambiguous $^1H-^{15}N$ resonance assignments. Starting from the assigned NH resonances, we used a sequential assignment procedure (Wüthrich, 1986) involving two-dimensional ^1H-^1H DQF-COSY, ^1H-^1H NOESY, ^1H-^1H TOCSY, two-dimensional (2D) $^1H-^{15}N$ HMQC-COSY (Clore et al., 1988), $^1H-^{15}N$ HMQC-NOESY (Gronenborn et al., 1989) and three-dimensional (3D) $^1H-^{15}N$ NOESY-HMQC spectroscopy (Nagayama et al., 1990). The pattern of the obtained sequential NOEs showed the presence of some secondary structure elements which are comparable to those in the crystal structure.

Abbreviations: NOE, nuclear Overhauser effect; HMQC, heteronuclear multiple-quantum coherence; SQC, single quantum correlation; COSY, correlated spectroscopy; DQF-COSY, double-quantum-filtered correlated spectroscopy; NOESY, NOE correlated spectroscopy; TOCSY, total correlated spectroscopy; CIDNP, chemically induced dynamic nuclear polarization; DSS, 2,2'-dimethyl-2-silapentane-5'-sulfonate.

MATERIALS AND METHODS

Sample preparation

The cro protein samples were mainly derived from transformed *Escherichia coli* (strain TG1) cells containing the expression plasmid, pKH1500 (Shirakawa et al., 1991a). In order to achieve essentially complete ^{15}N labeling of the protein, the bacteria were grown in M9 minimal medium with $^{15}NH_4Cl$ as the sole nitrogen source. The cro protein was also doubly-labeled specifically with ^{13}C by adding [1-^{13}C]-labeled amino acids to the M9 medium with $^{15}NH_4Cl$. A typical composition of 1 l of culture medium was 1 × M9 medium containing 0.5 g $^{15}NH_4Cl$, 50–100 mg ^{13}C-labeled amino acid, 4 g glucose, vitamins and a trace of metal ions. All ^{15}N- and ^{13}C-labeled amino acids were obtained from CIL, except for ^{15}N-Lys and ^{13}C-Arg, which were kindly supplied by Professor M. Kainosho, Tokyo Metropolitan University. $^{15}NH_4Cl$ was obtained from Isotec. The isotope enrichment was 95% or higher for each compound. The cro protein was purified by the procedures described previously by Shirakawa et al. (1985a,b).

The samples for 2D NMR typically comprised 1–2 mM protein in 90% H_2O/10% D_2O, or pure D_2O containing 20 mM phosphate buffer, 300 mM KCl and 0.01 mM EDTA, pH 5.0–6.7. The 3D NMR experiment was performed with a 3.3 mM sample dissolved in 90% H_2O/10% D_2O, 20 mM phosphate buffer, 300 mM KCl and 0.01 mM EDTA, pH 5.0, at 25°C.

NMR spectroscopy

All 2D ^1H–^1H spectra were acquired on a Bruker AM-600 spectrometer. The ^1H DQF-COSY, TOCSY and NOESY spectra were recorded at 30°C with a spectral width of 8333 Hz in both dimensions, the digital resolution being 16.3 Hz/point in $F1$ and 4.1 Hz/point in $F2$ axes.

The ^1H–^{15}N SQC, ^1H–^{15}N HMQC-NOESY and ^1H–^{15}N HMQC-COSY spectra were taken on a Bruker AM-500 spectrometer equipped with a reversed-mode ^1H probe. The spectra were recorded at 30°C, with spectral widths of 2800 and 7575 Hz in the ^{15}N and ^1H dimensions, respectively. Water suppression was achieved by presaturation during the relaxation delay for the solution at pH 5–6.7, 1–1 echo-type excitation (Sklenar and Bax, 1987) being also employed for some specified experiments. The digital resolution for the SQC spectrum was 10.9 Hz/point in $F1$ and 7.4 Hz/point in $F2$ axes. All 2D spectra were recorded in pure absorption with time-proportional phase incrementation (Redfield and Kuntz, 1975) and the data were processed with an ASPECT 3000 computer.

The 3D ^1H/^{15}N NOESY-HMQC spectrum of uniformly ^{15}N-labeled cro was recorded on a JEOL GSX400 spectrometer. The pulse train used in the experiment was described by Nagayama et al. (1990). The acquired data matrix was $128(t1) \times 32(t2) \times 250(t3)$ points, and the spectral widths were 4500 Hz ($F1$), 1600 Hz ($F2$) and 6000 Hz ($F3$). The mixing time was 150 ms. The total measurement time for the experiment was approximately 60 h. The 3D data were processed with a micro VAX computer using a program developed from TOOLKIT (Rawland Institute). The data matrix was zero-filled to $256(F1) \times 64(F2) \times 500(F3)$, and the Fourier-transformed data were plotted as 64 series of $F2$ slices.

^1H chemical shifts are given relative to external DSS (2,2′-dimethyl-2-silapentane-5′-sulfonate) and ^{15}N chemical shifts relative to external 3M $HCO^{15}NH_2$.

RESULTS AND DISCUSSION

Our assignment strategy was as follows. First, we labeled the cro protein with ^{15}N specifically or uniformly, and also doubly labeled it with ^{13}C amino acids to identify specific residues or the types of amino acids in HMQC or SQC spectra. Then we used the standard sequential assignment method using both homonuclear 2D ^{1}H DQF-COSY, NOESY and TOCSY spectra of the non-labeled cro protein, and heteronuclear 2D ^{1}H-^{15}N HMQC-COSY and HMQC-NOESY spectra of the ^{15}N-labeled cro protein. Although the heteronuclear experiments gave better resolution than the homonuclear ones, it was still difficult to obtain sequential connectivities for some parts of the main chain of the protein because of signal degeneracy. To get better resolution we performed 3D $^{1}H/^{15}N$ NOESY-HMQC.

For the sequential assignment of proteins, acidic solutions of around pH 4 are generally employed. The previous assignments of the cro protein were performed at that pH (Weber et al., 1985). However, cro is unstable at pH 4 and easily precipitates. Also, it does not bind to DNA in an acidic solution. Thus we performed the assignments in less acidic solutions. This meant that some amide proton resonances were hard to detect due to rapid exchange. In such a case, selective excitation, i.e., 1–1 echo, was employed for water suppression.

Double-labeling experiments, and 2D HMQC and SQC spectroscopy

Identification of the amide proton resonances of Lys and Leu was achieved by amino-acid-specific ^{15}N labeling (Shirakawa et al., 1991a), and identification of CH proton resonances of Leu and Phe was achieved by specific deuteration (Shirakawa et al., 1987). We then employed the double-labeling technique (Kainosho and Tsuji, 1982). We labeled the cro protein amino acids specifically with ^{13}C and uniformly with ^{15}N, by adding [1-^{13}C]amino acids and $^{15}NH_4Cl$ to the culture media, respectively. Figure 1 shows ^{1}H-^{15}N SQC spectra of the cro protein doubly labeled with ^{13}C and ^{15}N. $^{1}H/^{15}N$ cross peaks of the amino-acid residues next to the residues of which the carbonyl carbons were labeled with ^{13}C nuclei gave doublets along the ^{15}N axis due to $^{13}C-^{15}N$ coupling. Figure 1a shows a section of the SQC spectrum of the ^{15}N uniformly labeled cro protein, and Fig. 1b the same section of the SQC spectrum of the ^{15}N uniformly and [1-^{13}C] valine-labeled cro protein. On comparing these spectra, it was found that the signals designated as Tyr^{26}, Tyr^{51} and Lys^{56} split into doublets, which indicates that these peaks are due to the residues adjacent to three Val residues, Val^{25}, Val^{50} and Val^{55}. The split cross peak (^{1}H, 9.3; ^{15}N, 125 ppm) appeared in the ^{1}H-^{15}N HMQC spectrum of ^{15}N-lysine-labeled cro (Shirakawa et al., 1991a) and thus it was unambiguously assigned to Lys^{56}, which is consistent with the assignment by Leighton and Lu (1987). Sequence-specific assignments of Tyr^{26} and Tyr^{51} could be performed sequentially by 3D $^{1}H/^{15}N$ NOESY-HMQC spectroscopy, as is described later.

Identification of NH signals was also occasionally possible with the disappearance of the cross peaks on the addition of ^{13}C-labeled amino acids, since the added [1-^{13}C] amino acids carry ^{14}N at their amide positions. In Fig. 1c, the signals designated as Ala in parentheses disappeared. However, as shown in Fig. 1b, valine amide peaks were still observed on the addition of [1-^{13}C, ^{14}N] valine. This is due to scrambling of ^{15}N through transamination, i.e., ^{14}N of valine was diluted by ^{15}N in the culture media. In the case of [1-^{13}C, ^{14}N] arginine, $^{1}H/^{15}N$ resonances were weakly observed due to partial scrambling (Fig. 1d). Double-labeling experiments were also

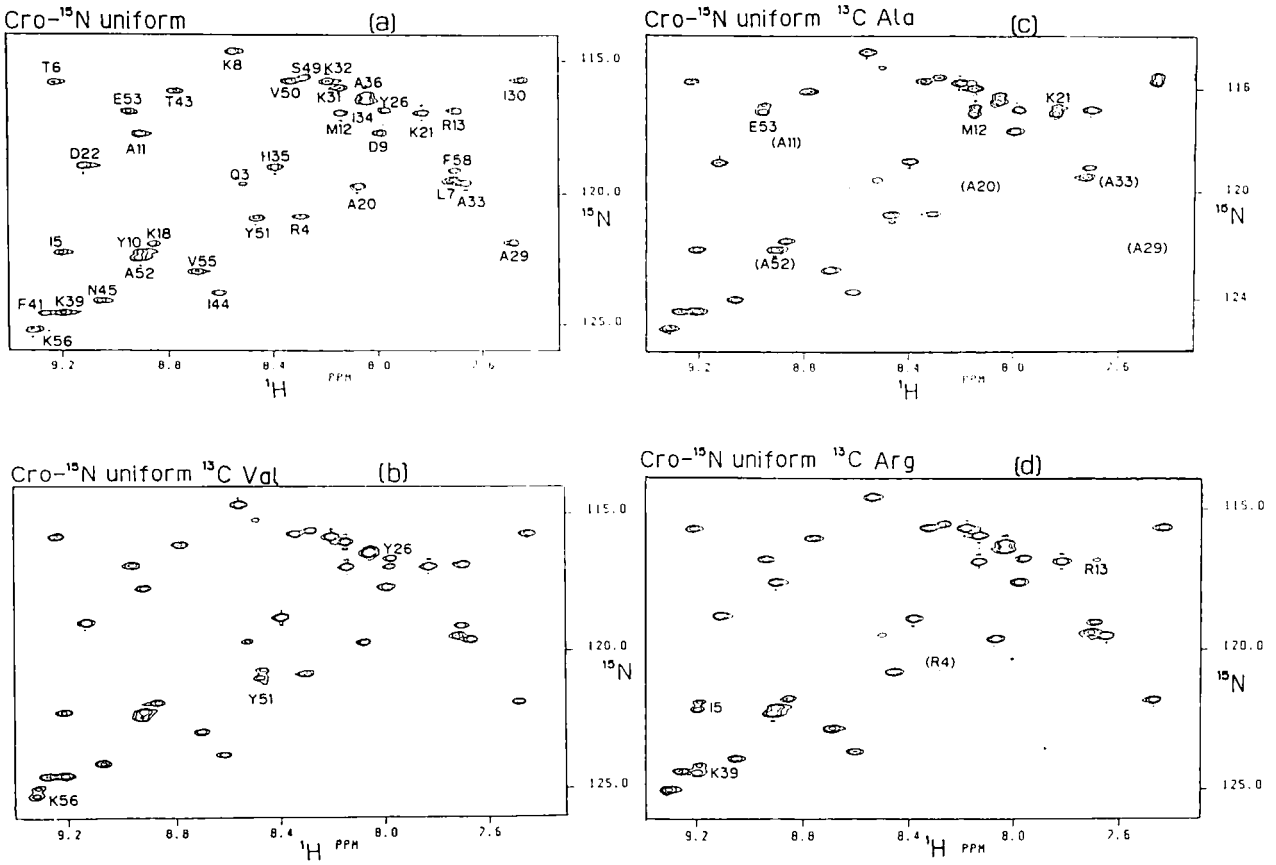

Fig. 1. Sections of 500-MHz ^{1}H-^{15}N SQC spectra of the λ-cro protein labeled with (a) ^{15}N, (b) ^{15}N and [1-^{13}C]Val, (c) ^{15}N and [1-^{13}C]Ala and (d) ^{15}N and [1-^{13}C]Arg, at 30°C and pH 6.4, with presaturation for water suppression. Residue numbers in parentheses indicate the positions of signals that have disappeared.

performed for Leu, Ile, Ser and Thr residues (data not shown). Ser and Thr did not give $^{1}H/^{15}N$ signals, but Leu and Ile showed weak $^{1}H/^{15}N$ peaks. In the case of Ser, the Gly peaks also disappeared by scrambling. Thus, we could specify the types of amino-acid residues, Arg, Lys, Gly, Ala, Ser, Thr, Leu and Ile, in the $^{1}H-^{15}N$ SQC spectrum.

Although the splitting patterns seen in the $^{1}H-^{15}N$ HMQC spectra of the doubly labeled cro proteins were ambiguous because of the poor *F*1 resolution (data not shown) (Bax et al., 1990), SQC spectra gave clear coupling constants of adjacent ^{13}C and ^{15}N, $^{1}J(^{13}C\text{-}^{15}N)$, as 11–15 Hz, and geminal coupling constants between ^{15}N and ^{1}H, $^{2}J(^{13}C\text{-}^{1}H)$, as 7–8 Hz.

Now we could make unambiguous assignments of some of the NH resonances of the cro protein by means of effective combination of selective ^{15}N-labeling and double-labeling with ^{13}C and ^{15}N. They were Leu^{7}, Lys^{8}, Lys^{18}, Lys^{21}, Lys^{36} and Lys^{56}. The $^{1}H/^{15}N$ peak of Lys^{32} was identified on the disappearance of the signal in the spectrum of ^{15}N-enriched mutant cro, where Lys^{32} was replaced by Gln (Shirakawa et al., 1990). Lys^{62} and Lys^{63} were assigned by Leighton and Lu (1987) through the combination of ^{15}N labeling and partial digestion. Thus, all $^{1}H/^{15}N$

signals of lysine residues have been assigned perfectly now (we have to revise the previous assignments for Lys[8], Lys[18] and Lys[21] (Shirakawa et al., 1990)).

Starting from the definite assignments, we could extend the sequential assignment by means of $^1H-^1H$ COSY, $^1H-^1H$ NOESY, $^1H-^1H$ TOCSY, $^1H-^{15}N$ HMQC-COSY and $^1H-^{15}N$ HMQC-NOESY. There are several advantages to the use of $^1H-^{15}N$ HMQC-NOESY and $^1H-^{15}N$ HMQC-COSY. Although the presence of diagonal peaks made it difficult or impossible to trace $^1HN-^1HN$ connectivities for the helical parts of the protein on homonuclear NOESY, the resolution was much better for $^1H-^{15}N$ HMQC-NOESY. A trace of the d_{NN} (i, i+1) connectivities for the α2 helical segment from residues 19 to 25 is given in Fig. 2.

3D $^1H/^{15}N$ NOESY-HMQC spectrum

Since peak separation was still not good enough for some parts of the spectra, we employed 3D $^1H/^{15}N$ NOESY-HMQC for the ^{15}N uniformly labeled cro protein. The resulting spectra gave series of ^1H homonuclear 2D spectra edited as to chemical shifts of ^{15}N nuclei, and almost complete separation of the resonances of the main-chain amide protons and C_α protons. Figures 3b and c show two representative slices at different ^{15}N chemical shifts (*F*2) of the 3D NOESY-HMQC spectrum. Two sets of NOE connectivities originating from the amide proton resonances of Ile[5] and Lys[39] can be dramatically separated due to their main-chain amide ^{15}N frequencies, although the corresponding region of the 2D NOESY spectrum (Fig. 3a) shows two overlapped connectivities. Thus, almost all amide proton resonances of both the main chain and

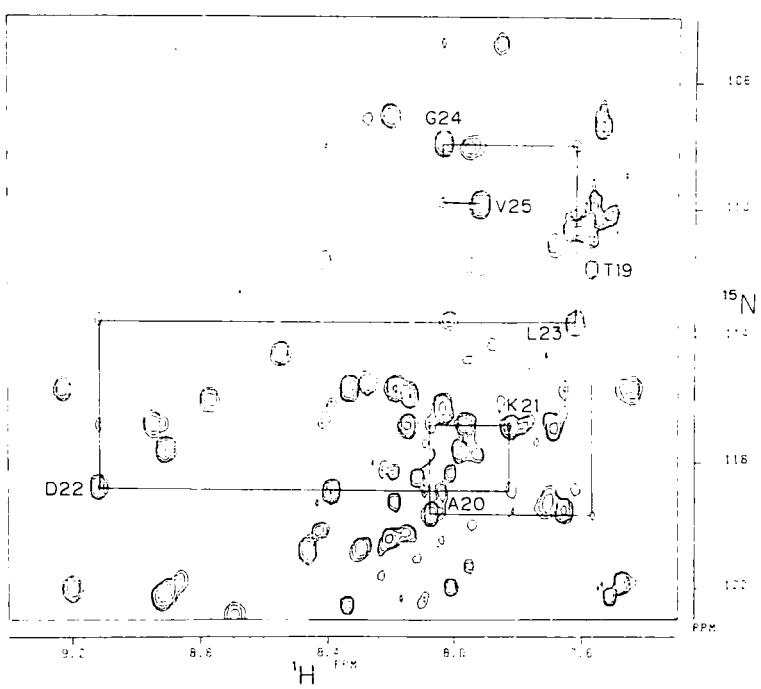

Fig. 2. Sections of a $^1H-^{15}N$ HMQC-NOESY spectrum of the uniformly ^{15}N-labeled λ-cro protein. The spectrum was acquired with a mixing time of 120 ms in H_2O at 30°C at pH 6.4, with presaturation for water suppression. d_{NN}(i, i+1) connectivities for the segment from residues 19 to 25 are shown.

Fig. 3. Comparison of 2D NOESY and slices of 3D ^{1}H/^{15}N NOESY-HMQC. (a), Low-field region of a conventional 2D NOESY spectrum of the λ-cro protein. (b, c), Cross sections through the 3D ^{1}H/^{15}N NOESY-HMQC data for the ^{15}N-enriched λ-cro protein. The ($F1$ and $F3$) ^{1}H planes were taken at different $F2$ ^{15}N frequencies. NOE cross peaks of the amide proton resonances of Ile^5 and Lys^{39} are well separated in the 3D data set.

side chains were separated in 64 ^{1}H-NOESY spectra edited as to ^{15}N frequency $F2$. The sequence connectivities obtained in the present 3D experiment for residues 52–55, which form a part of a β-sheet, are demonstrated in Fig. 4a. The sequential connectivities for three extended β-strand regions, using $d_{\alpha N}(i, i+1)$, were quite easily obtained through the present 3D NMR with little ambiguity.

The sequential connectivities for the helical part of the protein, through $d_{NN}(i, i+1)$, could also be easily traced. In Fig. 4b, the trace for the α1 helical part from residues 9–13 is given. Because NOE cross peaks tend to be close to strong diagonal peaks, $d_{NN}(i, i+1)$ connectivities are occasionally difficult to trace in homonuclear 2D NOESY. However, with the introduction of the third dimension to the 2D NOESY, the NN(i, i+1) cross peaks could be efficiently separated in the F 2 slices in the present 3D experiment. The quite good separation also enabled the detection of $d_{NN}(i, i+2)$ connectivities for the helical region of the protein (Fig. 4b).

The assignments of $^{1}H-^{15}N$ resonances of amide groups are summarized in Fig. 5 and listed in Table 1. Twelve amide proton signals of the six side chains of Asn and Gln were also identified on the basis of sequential NOE from C_β protons, and are indicated in parentheses in Fig. 5. Short- and medium-range connectivities are shown in Fig. 6.

Assignments of the side-chain protons

Most of the side-chain protons could be assigned through a combination of DQF-COSY, TOCSY, NOESY, $^{1}H-^{15}N$ HMQC-NOESY and 3D $^{1}H/^{15}N$ NOESY-HMQC spectroscopies. There are several long side-chain residues, three Arg, three Gln, three Glu and eight Lys. The unambiguous assignment of these side-chain protons was difficult to achieve. Because of the relatively large molecular weight of the cro protein (14.7 kDa as a dimer), the magnetization transfer for the spin systems of the long side-chain residues was not efficient enough in the TOCSY experiments. Therefore, we could not get cross peaks for some spin systems of the side chains of glutamine, glutamic acid, lysine and arginine residues. The obtained assignments are given in Table 1.

The ^{1}H resonance assignments of the cro protein were previously reported by Weber et al. (1985). However, about a quarter of the amide proton resonance assignments reported in the literature are inconsistent with ours, especially those for the helical regions of the protein. The in-

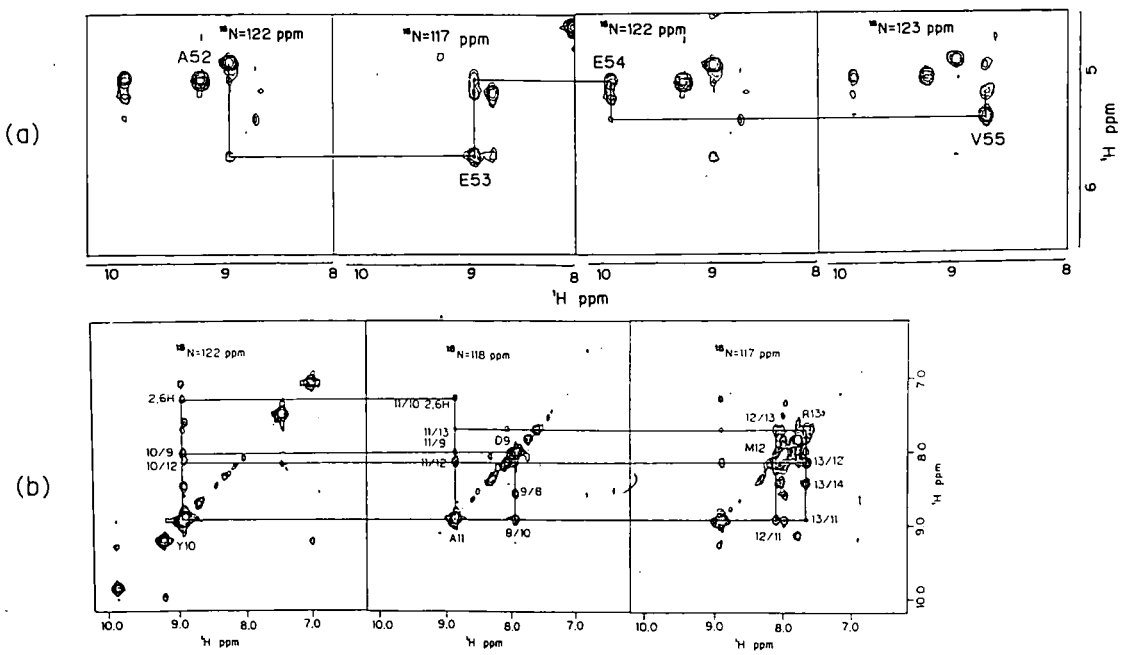

Fig. 4. Selected regions of slices of the 3D $^{1}H/^{15}N$ NOESY-HMQC data set at different ^{15}N frequencies. (a), $d_{\alpha N}(i, i+1)$ connectivities for the segment from residues 52 to 55 are presented. Vertical lines indicate the interconnections between the NH_i-$C_\alpha H_{i-1}$ and the NH_i-$C_\alpha H_i$ connectivities, and horizontal lines the interconnections between the NH_i-$C_\alpha H_i$ and the NH_{i+1}-$C_\alpha H_i$ connectivities. (b), $d_{NN}(i, i+1)$ and $d_{NN}(i, i+2)$ connectivities for the segment from residues 9 to 13 are shown.

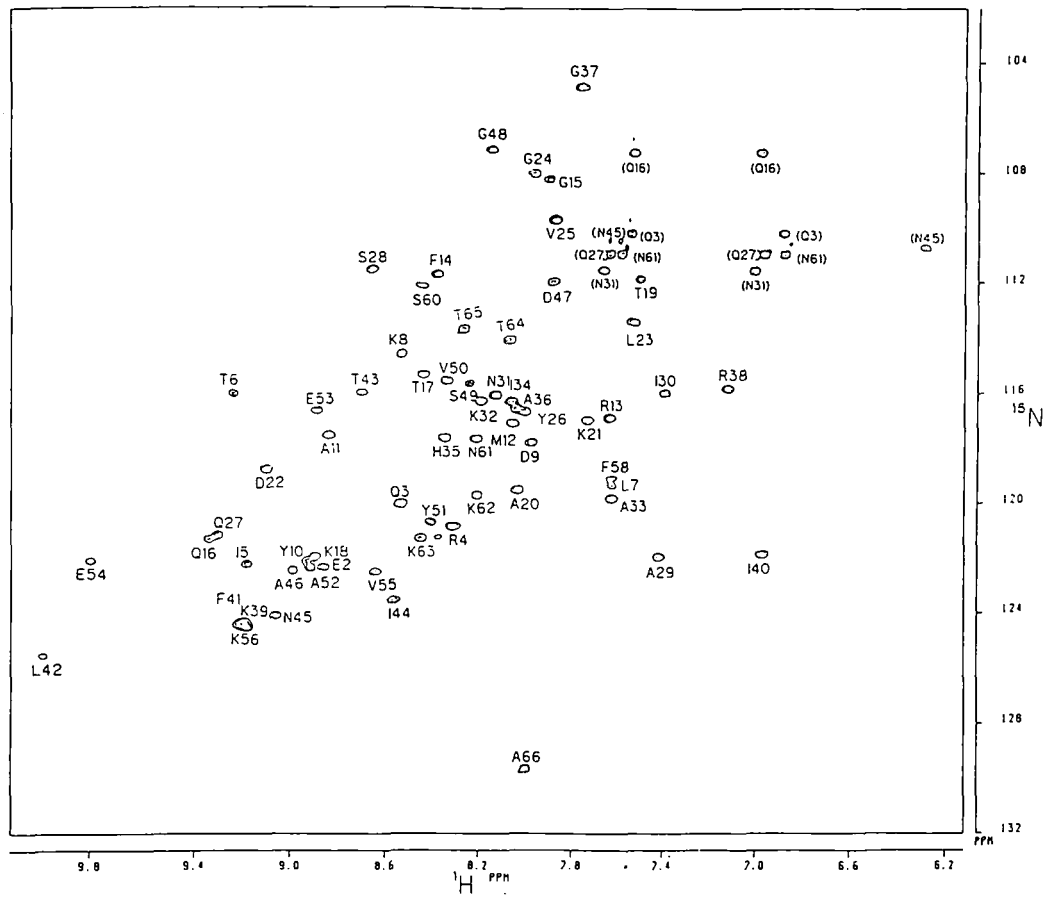

Fig. 5. ^{1}H-^{15}N SQC spectrum of the λ-cro protein taken with presaturation for water suppression at 20°C and pH 5, where the signal assignments are indicated by amino-acid residue numbers. Residue numbers in parentheses show the side-chain amide proton signals.

consistent chemical shifts are those of Leu^{7}–Tyr^{10}, Met^{12}, Arg^{13}, Gln^{16}–Lys^{18}, Lys^{21}–Leu^{23}, Ala^{36} and Ser^{49}. For instance, the HMQC spectrum of the ^{15}N-leucine-substituted cro protein gave three signals at ^{1}H chemical shifts of 10.1 ppm, 7.65 ppm and 7.75 ppm due to Leu^{42}, Leu^{23} and Leu^{7}, respectively. The 10.1-ppm peak could be definitely assigned to Leu^{42} from the results of the double-labeling experiment with ^{13}C-Thr. In previous work, we could not assign the peaks at 7.65 and 7.75 ppm to either Leu^{23} or Leu^{7} because of the close chemical shifts (Shirakawa et al., 1991a). In this experiment, we could definitely assign the 7.65-ppm peak to Leu^{23} from the sequential NOE typical for the helical region in the ^{1}H-^{15}N HMQC-NOESY spectrum (Fig. 2). Two of the three resonance assignments for Leu residues reported by Weber et al. (1985) are 9.97 ppm for Leu^{42} and 7.64 ppm for Leu^{7}, which agree with ours, but their 9.19-ppm peak for Leu^{23} is inconsistent with our observation.

TABLE I
^1H AND ^{15}N CHEMICAL SHIFTS OF λ PHAGE CRO REPRESSOR AT PH 6. 4, 30°C[a]

Residues	Chemical shifts (ppm)				Others
	^{15}N	NH	αH	βH	
Met[1]					
Glu[2]	(122.5)	(8.82)	(4.49)		
Gln[3]	119.6	8.49	4.48	2.01, 2.15	γH 2.34
					γNH_2 7.53, 6.89; $\gamma^{15}N$ 109.8
Arg[4]	120.8	8.26	5.00	1.61, 1.73	γH 1.46
Ile[5]	122.1	9.17	4.80	2.01	γH 1.38, 1.82
					γH_3 1.26
					δH_3 1.02
Thr[6]	115.8	9.20	4.30	4.53	γH_3 1.37
Leu[7]	119.5	7.66	4.57	1.95	γH 1.40; δH_3 0.61, 0.78
Lys[8]	114.6	8.51	3.80	1.70, 1.83	γH 1.40
Asp[9]	117.1	7.95	4.40	2.60, 2.88	
Tyr[10]	124.4	8.88	3.92	2.94, 3.18	H(2,6) 7.23
					H(3,5) 6.55
Ala[11]	117.7	8.87	4.12	1.35	
Met[12]	117.0	8.09	4.20	2.16, 2.24	γH 2.60, 2.74
Arg[13]	116.8	7.65	3.85	1.03, 1.38	γH 0.74
Phe[14]	111.7	8.38	4.63	2.98, 2.42	H(2,6) 6.49
					H(3,5) 6.97
					H(4) 7.13
Gly[15]	107.8	7.91	(4.08)		
			(4.68)		
Gln[16]	(121.5)	(9.34)	(4.68)	(2.14, 2.17)	γH 2.68
					γNH_2 7.52, 6.99; $\gamma^{15}N$ 107.4
Thr[17]	115.2	8.45	3.94	4.16	γH_3 1.30
Lys[18]	121.8	8.83	4.04	1.78, 1.90	γH 1.30
Thr[19]	111.8	7.53	3.86	4.20	γH_3 1.14
Ala[20]	119.7	8.03	3.80	1.56	
Lys[21]	116.9	7.78	4.10	2.00	
Asp[22]	118.9	9.09	4.30	2.36, 2.83	
Leu[23]	113.6	7.58	4.46	1.70	γH 1.74; δH_3 0.72, 0.74
Gly[24]	107.8	7.99	3.94		
			4.15		
Val[25]	109.8	7.88	4.57	2.15	γH_3 0.64, 0.73
Tyr[26]	116.7	7.94	4.78	2.92, 3.44	H(2,6) 7.29
					H(3,5) 6.90
Gln[27]	(121.1)	(9.28)	3.80	2.18, 2.22	γH 2.40
					γNH_2 7.64 6.96; $\gamma^{15}N$ 110.6
Ser[28]	111.3	8.60	4.24	3.81, 3.98	
Ala[29]	121.9	7.43	4.30	1.62	
Ile[30]	115.7	7.40	3.75	2.46	γH 1.37, 1.63
					γH_3 0.90
					δH_3 0.75
Asn[31]	116.0	8.11	4.30	2.80, 3.00	γNH_2 7.68 7.01; $\gamma^{15}N$ 111.2
Lys[32]	115.8	8.13	4.07	1.93	
Ala[33]	119.6	7.61	4.13	1.26	
Ile[34]	116.5	7.98	3.75	1.80	γH 1.10, 1.38
					γH_3 0.78
					δH_3 0.51
His[35]	118.7	8.35	4.31	3.17	H(2) 7.75
					H(4) 7.03
Ala[36]	116.5	7.98	4.33	1.52	
Gly[37]	104.7	7.81	3.90		
			3.98		
Arg[38]	115.9	7.18	4.09	1.75	
Lys[39]	124.4	9.16	4.29	2.30	
Ile[40]	122.0	6.99	4.90	1.02	γH 0.94, 0.30
					γH_3 0.30
					δH_3 −0.25
Phe[41]	124.4	9.23	5.00	2.66, 2.77	H(2,6) 6.92
					H(3,5) 7.23
					H(4) 7.27
Leu[42]	125.3	9.96	5.12	1.31	γH 1.33; δH_3 0.53, 0.00
Thr[43]	116.1	8.73	4.67	3.21	γH 1.03
Ile[44]	123.7	8.58	4.46	1.82	γH 1.15, 1.61
					γH_3 0.97
					δH_3 0.98
Asn[45]	124.0	9.02	4.86	2.91, 2.99	γNH_2 7.56, 6.33; $\gamma^{15}N$ 110.4
Ala[46]	(122.6)	(8.98)	4.14	1.48	
Asp[47]	111.9	7.88	4.51	2.65, 3.10	
Gly[48]	107.0	8.16	(3.54)		
			(4.32)		
Ser[49]	115.6	8.26	4.34	4.02, 3.97	
Val[50]	115.7	8.30	5.21	1.69	γH_3 0.66, 0.78
Tyr[51]	121.2	8.42	4.86	2.98, 3.10	H(2,6) 7.01
					7.03; H(3,5)
					6.79, 6.78
Ala[52]	122.4	8.88	5.63	1.55	
Glu[53]	116.8	8.90	5.00	1.98, 1.86	γH 2.20
Glu[54]	122.1	9.82	5.32	1.97	
Val[55]	122.9	8.39	5.12	2.05	γH_3 0.90
					1.02
Lys[56]	125.1	9.29	5.13	1.82	
Pro[57]			4.72	2.52, 2.57	γH 2.17; δH 3.76, 3.97
Phe[58]	119.1	7.66	4.72	2.53, 2.74	H(2,6) 7.07, 7.10
					H(3,5) 6.90
					H(4) 7.12
Pro[59]					
Ser[60]	(112.0)	(8.45)	(4.35)	(3.80)	
Asn[61]	(117.8)	(8.22)	(4.38)	(2.81)	γNH_2 7.58, 6.90; $\gamma^{15}N$ 110.8
Lys[62]	119.7	8.20	(4.28)		
Lys[63]	(121.4)	(8.48)	(4.40)		
Thr[64]	114.0	8.06	4.37	4.28	γH_3 1.20
Thr[65]	113.5	8.22	4.43	4.29	γH_3 1.21
Ala[66]	129.8	8.00	4.12	1.33	

[a] Uncertainties are ±0.02 ppm for ^1H and ± 0.1 ppm for ^{15}N chemical shifts.

[b] Chemical shifts in parentheses refer to pH 5.3, 30°C.

Rapidly exchanging amide protons

In order to determine the NH exchange rates of individual residues, we measured $^{1}H-^{15}N$ SQC spectra of the ^{15}N-enriched cro protein in a D_2O solution at pH 5 and 20°C. The ^{15}N-enriched protein was once lyophilized from H_2O at pH 5 and dissolved in D_2O. $^{1}H-^{15}N$ SQC spectra were measured at 12 min, 75 min and 143 min after dissolution. Each spectrum was acquired in 1 h. At 12 min after dissolution, 31 of 63 backbone amide proton signals had already disappeared, and 32 signals could be observed. Nine more peaks vanished in the spectrum of 75 min, and only a few correlation peaks were recognized in the spectrum measured at 143 min after dissolution. These results are given in Fig. 6. This experiment clearly indicated that the exchange rates of the NH protons of the cro protein are rather rapid. A mutant, where Val^{55} is replaced by Cys, forms a disulfide bonded dimer, and has more rigid structure than the wild cro protein (Shirakawa et al., 1991b). Some amide proton exchange rates of the Cys mutant are significantly slower than those of the wild cro protein and are in the order of days, showing that the rapid NH exchange of the wild protein reflects flexible motion of two subunits of the cro dimer.

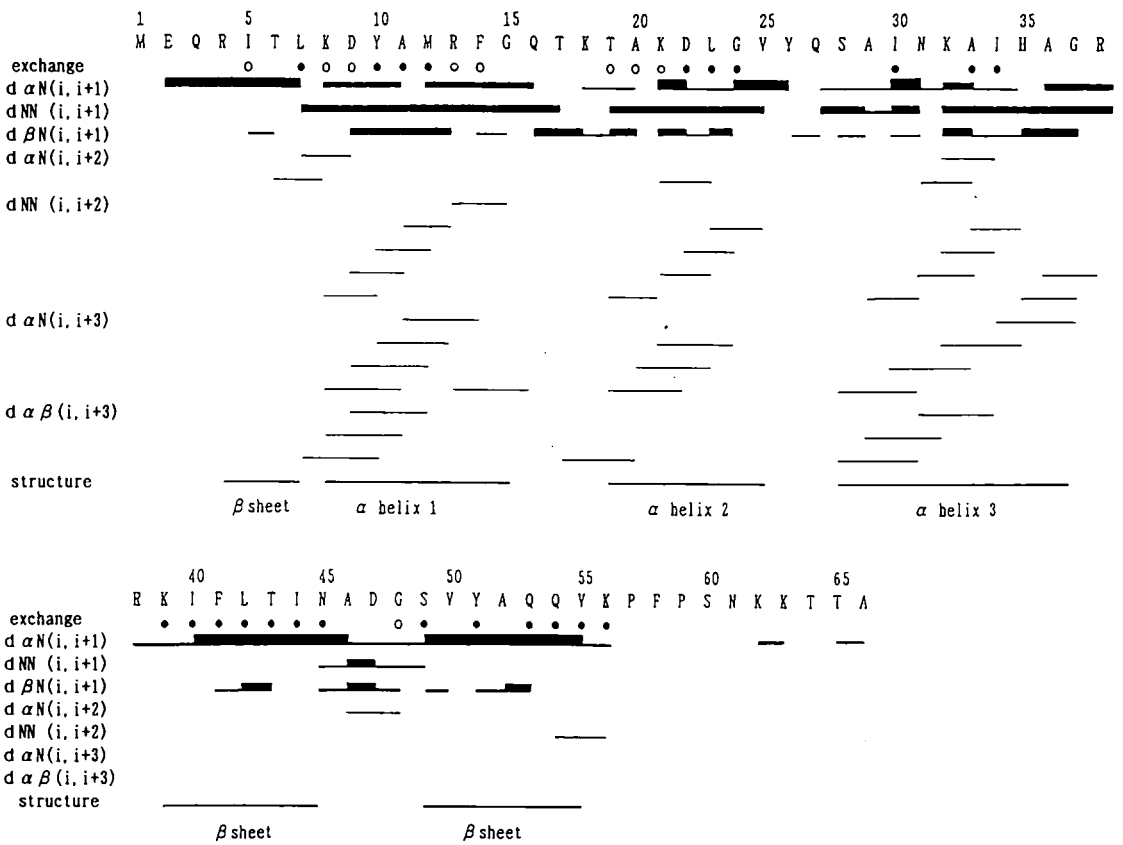

Fig. 6. Summary of the sequential NOE connectivities observed for the λ-cro protein. Sequential NOEs are represented by boxes, the height of each box indicating the NOE intensity (strong, medium or weak). Protons that show slow and intermediate exchange rate are indicated: ●, the lifetime longer than 1 h and ○, the lifetime longer than 10 min but shorter than 1 h.

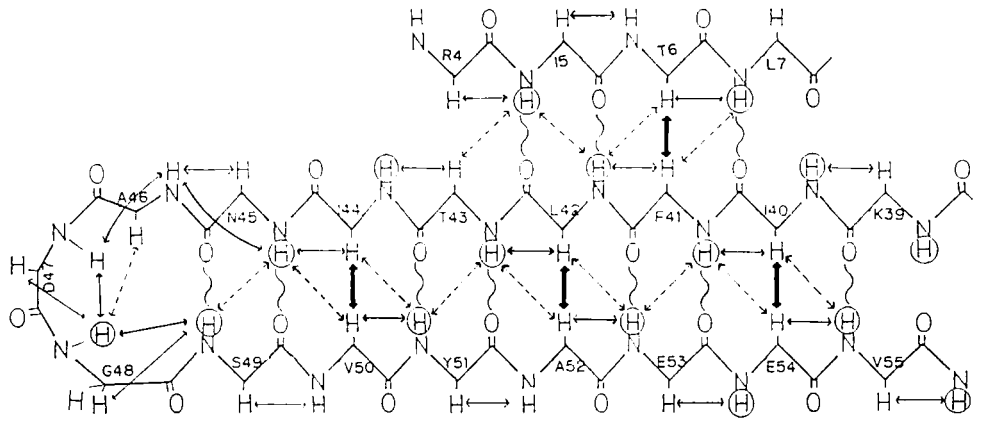

Fig. 7. Short- and long-range backbone NOE pattern indicating the presence of an antiparallel β-sheet involving the residues from Arg^4 to Leu^7 and from Lys^{39} to Val^{55}. Thick, thin and dotted double-ended arrows indicate interstrand $d_{\alpha\alpha}$, nearest-neighbor sequential d(i, i+1) and long-range $d_{\alpha N}$ or d_{NN} connectivities, respectively. The protons in circles show slower exchange. Wave marks indicate the assumed positions of hydrogen bonds.

Above pH 6.4, 12 backbone amide proton resonances, Glu^2, Gln^{16}, Thr^{17}, Gln^{27}, Ser^{28}, Ala^{46} and those of residues 60–65, are hardly observed in H_2O when water suppression is achieved by presaturation of the water resonance (compare Fig. 1a and Fig. 5). When we applied 1–1 echo-type excitation instead of presaturation, most of them appeared at this pH except for Glu^2, Gln^{16}, Gln^{27}, Ala^{46}, Ser^{60} and Asn^{61}.

Secondary structure elements

The sequential NOEs shown in Fig. 6 provide a basis for identification of several secondary structure elements. Stretches of strong $d_{NN}(i, i+1)$ and weak $d_{\alpha N}(i, i+1)$ connectivities in combination with $d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$ and $d_{\alpha,\beta}(i, i+3)$ were found for residues Lys^8–Gly^{15}, Thr^{19}–Val^{25} and Ser^{28}–Gly^{37}, indicating that these regions are α-helical (Wüthrich, 1986). These segments are referred to as α1, α2 and α3, respectively. The slow exchange rate of residues Lys^8–Phe^{14} and Thr^{19}–Gly^{24} is consistent with the α1 and α2 helices, but the α3 helical region seems to be exposed to the solvent. Three amide protons which may be located on the hydrophobic side of the helix are exchanged slowly with the solvent, but those of the other residues in the helix are quite easily accessible to water. The α3 helix is thought to be the helix which is directly involved in the binding with DNA. The flexible property of the α3 helix is interesting in relation to its binding ability. The segment from Gln^{16} to Lys^{18} forms a turn between the α1 and α2 helices. The amide proton resonances of Gln^{16} and Thr^{17} are hardly observed at neutral pH, indicating their exposure to the solvent. Tyr^{26} and Gln^{27} participate in a turn structure between the α2 and α3 helices which corresponds to the turn part of the 'helix-turn-helix' motif. Gln^{27} and Ser^{28} are easily accessible to water.

Regions of strong $d_{\alpha N}(i, i+1)$ and weak $d_{\alpha N}(i, i)$ connectivities, indicative of an extended peptide chain conformation, are found for residues Arg^4–Leu^7, Lys^{39}–Asn^{45} and Ser^{49}–Val^{55} (they are referred to as β1, β2 and β3, respectively). Extended peptide segments are associated in β-sheets characterized by typical interstrand NOEs, especially interstrand $d_{\alpha\alpha}$ and $d_{\alpha N}$ connectivi-

ties (Wüthrich, 1986). This pattern can be seen among these three extended segments, indicating that they form an antiparallel β-sheet, as shown in Fig. 7. The observed connectivities and the rates of exchange imply the presence of three hydrogen bonds between β1 and β2, and seven hydrogen bonds between β2 and β3. Glu^2 and Gln^3 seem to fray out at the N-terminal and not to be involved in the β-sheet. The NH proton of Leu^7 may participate in the formation of the β-sheet rather than the α1 helix, judging from intra- and interstrand NOEs. Three residues, Ala^{46}, Asp^{47} and Gly^{48}, between the β2 and β3 helices form a loop rather than a β-turn. The amide proton of Ala^{46}, whose resonance is hardly observed at neutral pH, seems to protrude into the solvent, resulting in rapid exchange with water.

The NH resonances of the C-terminal seven residues (Phe^{58}, Ser^{60}–Thr^{65}) were quite sensitive to a change in pH, and were hardly seen above pH 6.4. Not so many NOE connectivities could be detected for this region. This evidence indicates that the C-terminal region is flexible in solution. The presence of two proline residues in the region made sequential assignments difficult. The amide proton of Ala^{66} behaves peculiarly. The signal is insensitive to pH, but very sharp. Its C_α proton signal shows a fine structure. The residue is located in a flexible part, but the amide proton may interact with some group inter- or intramolecularly.

The crystal structure of a cro monomer consists of three α-helices (residues 7–14, 15–23 and 27–36) and three strands of antiparallel β-sheet (residues 2–6, 39–45 and 48–55). The C-terminal region (residues 63–66) is disordered, and is not seen in the crystal structure. This feature of the crystal structure is quite similar to that in solution deduced from the present NMR data. A little difference is seen at the starting residue of the α2 helix; the helix is shorter than that in the crystal.

In a crystal, four monomers of the cro protein are related in pairs by mutually perpendicular intersecting 2-fold axes (Anderson et al., 1981). They chose a pair of monomers related by R symmetry axis as a dimer which may persist in solution. In the pair the sequence Glu^{54}–Val^{55}–Lys^{56}, close to the symmetry axis, of two monomers forms a pair of antiparallel β-sheet strands. A hydrodynamic study of a solution of the cro protein showed that cro exists as a dimer in solution at room temperature, and dissociates into monomers at around 45°C, accompanied by denaturation of the protein (Iwahashi et al., 1982). When we changed Val^{55} to a Cys residue, the mutant protein formed a disulfide bond-bridged dimer spontaneously (Shirakawa et al., 1991b). This may indicate that the Cys^{55} residues are located close to each other in the dimer. In the course of analyses of NMR spectra of the cro protein, we expected to observe intermolecular NOEs which might disturb sequential assignments. However, we did not observe many such NOEs. One candidate for the intermonomer cross peak is $d_{NN}(Glu^{54}, Lys^{56})$, which cannot be explained as an intrastrand NOE when we assume the stretching of the β3 strand to Lys^{56}. The slow exchange rates of the amide protons of Glu^{54} and Lys^{56} support the presence of hydrogen bonding. $d_{\alpha N}(Val^{55}, Lys^{56})$ can be explained by either an intra- or interstrand NOE. Therefore, it is quite probable that the pair of the sequence Glu^{54}–Val^{55}–Lys^{56} form an intermonomer antiparallel β-sheet in solution. Since cro is a dimer in solution, we should interpret NOEs in terms of their intra- or intersubunit nature. The analysis is somehow based on the X-ray structure by Anderson et al. (1981).

We have identified several long-range NOEs in addition to the sequential connectivities shown in Fig. 6. Some of them were used for the arrangement of the β-strands, but the others have not yet been employed because of the lack of assignments of some side-chain protons. When we perform further analyses of the spectra, we can use these data for determination of the tertiary structure of the cro monomer and the arrangement of two monomers in solution.

ACKNOWLEDGEMENTS

The authors wish to express their sincere thanks to Prof. M. Kainosho for the supply of valuable ^{15}N- and ^{13}C-labeled amino acids. Thanks are also due to Dr. W. Yoshikawa, Protein Engineering Research Institute, and Dr. M. Waelchli, Bruker Japan Co. Ltd., for their help in the practical use of an AM500 spectrometer in heteronuclear correlation spectroscopy, and to Dr. K. Nagayama, JEOL Ltd., for his support in the measurement of 3D NMR. This work was partly supported by a Grant-in-Aid for Special Project Research from the Ministry of Education, Science and Culture of Japan (02263103), and partly by Special Coordination Funds from the Science Technology Agency.

REFERENCES

Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. (1981) *Nature*, **290**, 754–758.
Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) *J. Magn. Reson.*, **55**, 301–315.
Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990) *J. Magn. Reson.*, **86**, 304–318.
Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.
Brennan, R.G., Roderick, S.L., Takeda, Y., and Matthews, B.W. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8165–8169.
Clore, G.M., Bax, A., Wingfield, P. and Gronenborn, A.M. (1988) *FEBS Lett.*, **238**, 17–21.
Fesik, S.W. and Zuiderweg, E.R.P. (1988) *J. Magn. Reson.*, **78**, 588–593.
Gronenborn, A.M., Bax, A., Wingfield, P.T. and Clore, G.M. (1989) *FEBS Lett.*, **243**, 93–98.
Hsiang, M.W., Cole, R.D., Takeda, Y. and Echols, H. (1977) *Nature*, **270**, 275–277.
Ikura, M., Kay, L.E. and Bax, A. (1990) *Biochemistry*, **29**, 4659–4667.
Iwahashi, H., Akutsu, H., Kobayashi, Y., Kyogoku, Y., Ono, T., Koga, H. and Horiuchi, T. (1982) *J. Biochem.*, **91**, 1213–1221.
Kainosho, M. and Tsuji, T. (1982) *Biochemistry*, **21**, 6273–6299.
Lee, S.J., Shirakawa, M., Akutsu, H., Kyogoku, Y., Shiraishi, M., Kitano, K., Shin, M., Ohtsuka, E. and Ikehara, M. (1987) *EMBO J.*, **6**, 1129–1135.
Leighton, P. and Lu, P. (1987) *Biochemistry*, **26**, 7262–7271.
Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989) *J. Am. Chem. Soc.*, **111**, 1515–1517.
Müller, L. (1979) *J. Am. Chem. Soc.*, **101**, 4481–4484.
Nagayama, K., Yamazaki, T., Yoshida, M., Kanaya, S. and Nakamura, H. (1990) *J. Biochem.*, **108**, 149–152.
Ohlendorf, D.H., Anderson, W.F., Fisher, R.G., Takeda, Y. and Matthews, B.W. (1982) *Nature*, **298**, 718–723.
Ptashne, M., Jefferey, A., Johnson, A.D., Maureu, R., Meyer, B.J., Pabo, C.O., Roberts, T.M. and Sauer, R.T. (1980) *Cell*, **19**, 1–11.
Redfield, A.G. (1983) *Chem. Phys. Lett.*, **96**, 537–540.
Redfield, A.G. and Kuntz, S.D. (1975) *J. Magn. Reson.*, **19**, 250–254.
Shirakawa, M., Kawata, Y., Lee, S.J., Akutsu, H., Sakiyama, F. and Kyogoku, Y. (1985a) *J. Biochem.*, **98**, 799–805.
Shirakawa, M., Lee, S.J., Akutsu, H., Kyogoku, Y., Kitano, K., Shin, M., Ohtsuka, E. and Ikehara, M. (1985b) *FEBS Lett.*, **181**, 286–290.
Shirakawa, M., Lee, S.J., Yamamoto, K., Takimoto, M., Akutsu, H. and Kyogoku, Y. (1987) in *Structure and Expression, Vol. 1: From Proteins to Ribosomes* (Eds, Sarma, R.H. and Sarma, H.H.), Adenine Press, New York, pp. 167–179.
Shirakawa, M., Serikawa, K., Matsuo, H., Lee, S.J., Takimoto, M. and Kyogoku, Y. (1990) in *Prot. Engineering* (Ed, Ikehara, M.), Springer, Tokyo, pp. 73–81.
Shirakawa, M., Lee, S.J., Takimoto, M., Matsuo, H., Akutsu, H. and Kyogoku, Y. (1991a) *J. Mol. Struct.*, **242**, 355–366.
Shirakawa, M., Matsuo, H. and Kyogoku, Y. (1991b) *Prot. Engineering*, **4**, No. 5, in press.
Sklenar, V. and Bax, A. (1987) *J. Magn. Reson.*, **74**, 469–479.
Weber, P.L., Wemmer, D.E. and Reid, B.R. (1985) *Biochemistry*, **24**, 4553–4562.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.