Journal of Biomolecular NMR, 8 (1996) 429–444
ESCOM

J-Bio NMR 404

^{1}H, ^{15}N and ^{13}C resonance assignments and secondary structure determination of the RNA-binding domain of *E. coli* rho protein

Deborah M. Briercheck[a], Timothy J. Allison[a], John P. Richardson[b], Jeffery F. Ellena[c], Todd C. Wood[a] and Gordon S. Rule[a,*]

[a]*Department of Biochemistry, Box 440, University of Virginia School of Medicine, Charlottesville, VA 22908, U.S.A.*
[b]*Department of Chemistry, Indiana University, Bloomington, IN 47405, U.S.A.*
[c]*Department of Chemistry, University of Virginia, Charlottesville, VA 22908, U.S.A.*

Received 18 June 1996
Accepted 15 August 1996

Keywords: Secondary structure; Rho factor; RNA-binding protein; Transcription

Summary

Protein fragments containing the RNA-binding domain of *Escherichia coli* rho protein have been overexpressed in *E. coli*. NMR spectra of the fragment containing residues 1–116 of rho protein (Rho116) show that a region of this protein is unfolded in solution. Addition of $(dC)_{10}$ to this fragment stabilizes the folded form of the protein. The fragment comprising residues 1–130 of rho protein (Rho130) is found to be stably folded, both in the absence and presence of nucleic acid. NMR studies of the complex of Rho130 with RNA and DNA oligonucleotides indicate that the binding-site size, affinity, and specificity of Rho130 are similar to those of intact rho protein; therefore, Rho130 is a suitable model of the RNA-binding domain of rho protein. NMR line widths as well as titration experiments of Rho130 complexed with oligonucleotides of various lengths suggest that Rho130 forms oligomers in the presence of longer oligonucleotides. ^{1}H, ^{15}N and ^{13}C resonance assignments were facilitated by the utilization of two pulse sequences, CN-NOESY and CCH-TOCSY. The secondary structure of unliganded Rho130 has been determined by NMR techniques, and it is clear that the RNA-binding domain of rho is more structurally similar to the cold shock domain than to the RNA recognition motif.

Introduction

Rho protein is a transcription termination factor, required for the release of nascent RNA transcripts from RNA polymerase and template DNA at rho-dependent terminators in both gram-negative and gram-positive bacteria (for reviews see Platt and Richardson, 1992; Platt, 1994; Richardson, 1996). Rho is a 47 kDa hexamer consisting of a ring-shaped 'trimer of dimers' (Geiselmann et al., 1992b). Hexameric rho protein binds unstructured regions of RNA of at least 80 nucleotides long and appears to dissociate the transcription complex by coupling its RNA–DNA helicase activity to its RNA-dependent ATPase activity (Brennan et al., 1987). Although there is no consensus sequence for binding, a relatively high proportion of cytidine residues is required for rho activation (Morgan et al., 1983). Thus, if rho binds unstructured regions of RNA that contain cytidine, rho becomes an active ATPase. This ATPase activity provides energy for 5'→3' translocation along the RNA. If transcription is paused or slowed, rho reaches the DNA–RNA duplex and effects release of the RNA transcript (Jin et al., 1992). A current model suggests that rho works via a tethered tracking mechanism in which rho binds RNA and moves 5'→3' while remaining bound to the initial binding site (Faus and Richardson, 1990; Steinmetz and Platt, 1994).

*To whom correspondence should be addressed at: Department of Biological Sciences, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15217, U.S.A.

Abbreviations: Rho116, Rho130, protein containing the first 116 (130) residues of rho; CSD, cold shock domain; RRM, RNA recognition motif; RBD, RNA-binding domain; IPTG, isopropyl β-D-thiogalactopyranoside; EDTA, ethylenediaminetetraacetic acid; NOE, nuclear Overhauser enhancement.

Supplementary Material: One table, containing the chemical shift assignments of Rho130, and two figures, illustrating the pulse sequences for the CCH-TOCSY and CN-NOESY experiments, are available on request from the authors.

Since rho monomers oligomerize at concentrations > 1 μM, most binding studies have been conducted on the rho hexamer. To simplify the analysis, the allosteric interactions between monomers have generally been ignored. Each rho hexamer possesses three high-affinity and three low-affinity RNA binding sites (Geiselmann and Von Hippel, 1992). Binding studies performed with long oligonucleotides (n > 20) identified the presence of the three high-affinity binding sites ($K_a = 3 \times 10^6\ M^{-1}$) and suggested a binding-site size of approximately 11–13 nucleotides. Studies performed with shorter oligonucleotides (n ≤ 13), using ultrafiltration to monitor RNA binding, identified both the three high-affinity sites and the three low-affinity sites ($K_a \sim 3 \times 10^5\ M^{-1}$), but with a binding-site size of 5–6 bases (Wang and Von Hippel, 1993b). Other studies, utilizing circular dichroism spectroscopy to monitor binding, showed only three high-affinity binding sites, each with a binding-site size of 5–6 bases (Geiselmann et al., 1992a). It was not possible to identify the lower affinity sites in the circular dichroism studies, because light absorbance prevented the use of the high concentrations of protein necessary to detect weak binding sites. In summary, the rho hexamer has been shown to have three high-affinity ($K_a = 3 \times 10^6\ M^{-1}$) and three low-affinity ($K_a \sim 3 \times 10^5\ M^{-1}$) RNA binding sites, with each site binding about six nucleotides. For all six RNA binding sites of hexameric rho to be filled, however, a total strand length of approximately 80 nucleotides is required. Presumably, rho requires a long RNA strand because the RNA wraps around the exterior of the hexamer. Von Hippel and coworkers (Geiselmann et al., 1993) have proposed a model for wrapping of the RNA around the rho hexamer, such that the high- and low-affinity RNA binding sites differ in polarity with respect to the 5'→3' direction of the RNA relative to the protein.

The binding specificity of rho has also been evaluated. First, it has been shown that rho protein has nearly equal affinity for RNA or DNA, although its ATPase is only activated by RNA binding. Further, poly(dC) is a competitive inhibitor of ATPase activation (Richardson, 1982; Seifried et al., 1992). A single cytidine can also compete for binding with rho-dependent terminators and inhibit ATPase activation (Richardson and Richardson, 1992). Based on oligonucleotide off-rates, rho protein binds poly(rC) with an approximately 100-fold higher affinity than it binds poly(rA) (Galluppi and Richardson, 1980). The affinity of rho for poly(rU) is more similar to that for poly(rC); its affinity for poly(rC) is only 20-fold greater than for poly(rU) (McSwiggen et al., 1988). However, the binding of rho to poly(rU) is primarily an ionic interaction, as increasing the salt concentration from 0.025 M KCl to 0.5 M KCl reduces binding to below detectable levels (Galluppi and Richardson, 1980; McSwiggen et al., 1988). In contrast, the binding of poly(rC) is less dependent on ionic strength. Studies with oligo(dC) have shown that the affinity of rho for $(dC)_{11}$ is reduced approximately threefold as the ionic strength is raised from 0.1 M KCl to 0.56 M KCl (Geiselmann et al., 1992a). Therefore, binding of oligo(rC) to rho is driven by contributions from both ionic and nonionic interactions, although the latter may play a more important role.

Each rho monomer is composed of three domains, as determined by limited proteolysis. The N-terminal domain, i.e., the first 130 residues, forms the RNA-binding domain. The middle domain, 153 residues, comprises the ATP-binding domain. The C-terminal domain, the last 136 residues, has no well-defined role (Dolan et al., 1990). Rho protein deletion studies have been performed to further define the RNA-binding domain of rho protein. Modrak and Richardson (1994) produced a number of fragments of the N-terminal domain of rho. A fragment containing the 116 N-terminal residues (Rho116) shows the same binding specificity as intact rho. However, the Rho116 fragment binds RNA with 10-fold less affinity than intact rho, suggesting that the full RNA-binding domain of rho is larger than 116 residues. Modrak and Richardson also showed that up to the first 22 N-terminal residues of Rho116 could be removed with no further decrease in affinity. Deletions at the C-terminus of Rho116 were much more deleterious for function; deletion of even two residues greatly reduced RNA binding.

The atomic structure of rho protein is unknown, and to date the structures of only a few other RNA-binding proteins are known. These proteins contain primarily β-sheet, and most are classified as possessing either the RNA recognition motif (RRM, also referred to as RNA-binding domain or RBD) or the cold shock domain (CSD). The RRM comprises a βαβ-βαβ motif in which the β-strands form an antiparallel sheet, with the helices located against one face of the sheet (Nagai et al., 1990; Hoffman et al., 1991; Wittekind et al., 1992). In contrast, the CSD comprises five β-strands, which form an antiparallel β-barrel (Schindelin et al., 1993; Schnuchel et al., 1993). Both folds possess an eight-residue sequence motif, RNP1, which is located on a central β-strand that is thought to be involved in binding of nucleic acid (Landsman, 1992).

Rho130 (the 130-residue N-terminal fragment of rho) possesses an RNP1-like sequence starting at residue 60, DGFGFLRS (Haynes, 1992; Burd and Dreyfuss, 1994). Mutagenesis studies have shown that this sequence is critical for binding. Specifically, Brennan and Platt (1991) demonstrated that changing residues Phe^{62} and Phe^{64} to alanines reduced RNA binding. Further, Martinez et al. (1996a,b) showed that changing residues Asp^{60}, Phe^{62}, or Arg^{66} had a large deleterious effect on the termination efficiency of rho protein. These studies, along with recent sequence alignments of rho homologues, led Martinez et al. (1996a) to propose a model for the secondary structure of residues 49–83 of rho protein. In this model, these residues form three strands of a β-sheet with the RNP1

sequence being on the second β-strand. Asp^{60} and Gly^{61} are proposed to be in a turn and the side chains of Phe^{62}, Phe^{64}, and Arg^{66} are proposed to be on the face of the β-sheet, which is suggested to interact with RNA. The location of the RNP1-like sequence on a central β-strand is in agreement with both the RRM and CSD folds. However, in rho the short linkers between strands disallow the presence of helices and suggest a fold that is more consistent with the CSD.

In this report we show that Rho130 is amenable to high-resolution NMR studies of either the unbound protein or the protein–nucleic acid complex. Further, we show that this Rho130 fragment has binding properties that are similar to those of intact rho protein, indicating that Rho130 is a suitable model of the RNA-binding domain of rho protein. The ^{1}H, ^{15}N and ^{13}C resonance assignments of unliganded Rho130 have been obtained, and the secondary structure of Rho130 has been determined using NMR techniques. The secondary structure of the C-terminus of Rho130 contains a five-stranded β-sheet rather than a βαβ-βαβ binding motif; therefore, it is clear that the RNA-binding domain of rho is structurally related to the CSD-containing RNA/DNA-binding proteins.

Materials and Methods

Expression vectors

The expression vector for Rho116 has been described by Modrak and Richardson (1994). Since NMR studies on Rho116 showed that this rho fragment is partially unfolded (see below), and therefore unsuitable for structural studies, a larger fragment of rho was expressed. Dolan et al. (1990) have shown that rho protein possesses a natural trypsinolysis site at Arg^{128} or Lys^{130}, and that this 130-residue trypsinolysis fragment binds RNA with characteristics similar to those of intact rho. To express this protein, the gene encoding the entire protein was inserted between the XbaI and EcoRI sites of the M13 phage, M13mp18 (Yanisch-Perron et al., 1985). Standard site-directed mutagenesis techniques (Kunkel et al., 1991) were used to introduce a stop codon and a BglII site after residue 130 (the sequence of the mutagenic oligonucleotide was: CGCCCGCAACAAATAGATCTTTGAGAACTT). The Rho130 gene was then removed from M13mp18 by cutting with both XbaI and BglII and it was subsequently inserted into the same sites on the pET-11a vector (Studier et al., 1990).

Rho116 and Rho130 expression and purification

Both Rho116 and Rho130 were expressed using an *E. coli* BL21(DE3)pLysS (Studier et al., 1990) cell line in minimal medium (Penington and Rule, 1992). Labeled Rho130 (^{13}C and/or ^{15}N) was prepared by growing the bacteria in minimal medium supplemented with ^{13}C glucose (3 g/l) and/or ^{15}N ammonium sulfate (1 g/l). When the cell density reached an A_{600} of 1.0, IPTG was added to a concentration of 1 mM. Improved yields were obtained with the addition of 50 mg/l of guanosine, cytosine, and uridine at induction. The cells were induced for a total of 3.5 h. Upon harvesting, the cells were spun down and resuspended in lysis buffer (50 mM Tris, 10 mM EDTA, pH 7.0). Lysozyme (5 mg/l) was added and the cells were chilled and sonicated to solubilize the protein. After the cellular debris was removed by centrifugation (33 000 rpm, Beckman 50.2 Ti rotor), the supernatant was bound to a CM Sephadex column and eluted using a gradient of 0.0 to 0.8 M NaCl in 50 mM Tris, 2 mM EDTA buffer, pH 7.8. The protein was precipitated by addition of a saturating amount of ammonium sulfate, after which it was spun down and resuspended in a buffer of 10 mM potassium phosphate, 150 mM potassium sulfate and 2 mM EDTA, pH 7.0. After chromatographing the protein through a G-50 Sephadex gel filtration column, it was exchanged into a buffer of 50 mM Tris, 2 mM EDTA, pH 7.8, and bound to a DEAE–agarose column. The protein was eluted with a gradient of 0.0 to 0.4 M K_2SO_4 in the same buffer that was used for loading. The protein was then dialyzed into the buffer used for NMR spectroscopy (10 mM potassium phosphate, 150 mM potassium sulfate, 2 mM EDTA, pH 7.0). The yields for this expression system were about 30 mg/l for both Rho116 and Rho130. The protein is stable in NMR buffer to concentrations of 2 mM for at least one month at room temperature. Rho130 can also be lyophilized in NMR buffer for long-term storage at −80 °C.

Oligonucleotide preparation and purification

DNA oligonucleotides were made at the on-site oligonucleotide sequencing facility. RNA oligonucleotides were purchased from Oligos, Etc. (Wilsonville, OR, U.S.A.). Both DNA and RNA oligonucleotides were purified over a G-25 Sephadex column in 30 mM ammonium acetate. The fractions were then lyophilized, resolubilized in water, and relyophilized to remove any remaining salts. The dry oligonucleotides were stored at −20 °C. The concentrations of oligonucleotide solutions were determined spectrophotometrically at $\lambda = 260$ nm using the following extinction coefficients (per residue) (Cantor and Tinoco, 1965; Stanley and Bock, 1965; Cantor et al., 1970; Wang and Von Hippel, 1993a): $(dC)_{6\text{-}16}$, 7.22×10^3 M cm/l; $(rC)_6$, 7.30×10^3 M cm/l; $(rA)_{12}$, 12.65×10^3 M cm/l.

DNA and RNA titrations

Rho130 titrations of DNA and RNA were performed to evaluate the binding affinity of nucleic acid to Rho130. An initial $^{15}N/^{1}H$-HSQC was obtained for 100 μM Rho130 in the presence of a large excess of oligonucleotide (5:1, oligonucleotide:protein). Concentrated protein was then

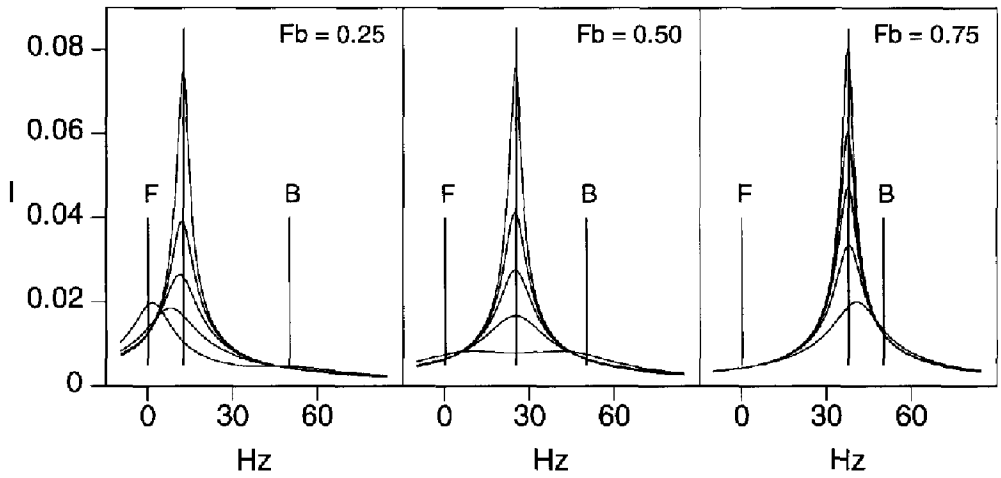

Fig. 1. The effect of chemical exchange on line shape. Line shapes were calculated using the formula for general chemical exchange given by Carrington and McLachlan (1967). The curves in each panel represent simulated line shapes obtained for affinity constants of 10^4 (sharpest line), 10^5, 2×10^5, 4×10^5, and 10^6 M^{-1}. In all cases the kinetic on-rate was 10^8 s^{-1} and the off-rate was adjusted to give the desired affinity constant. An intrinsic line width of 4 Hz was used. The left, middle, and right panels show simulations where the fraction of bound protein (Fb) is 0.25, 0.50, and 0.75, respectively. The positions of the resonance line in the unbound and fully bound states are indicated by the vertical lines labeled 'F' and 'B', respectively. The position of the resonance line under conditions of fast exchange is indicated by the tall vertical line.

added and another $^{15}N/^{1}H$-HSQC was acquired. The titration was continued in this manner until the concentration of Rho130 exceeded that of the oligonucleotide by about 2.5:1. Since Rho130 exchanges in the intermediate to fast exchange regime, the percent of bound Rho130 was determined by measuring the change in the chemical shifts of individual resonance lines during the course of the titration. The binding constants were determined by obtaining the minimum χ^2 between the data and the binding model by sampling the parameters as a discrete grid.

Determining binding constants in this manner assumes that the change in the position of the resonance line at different oligonucleotide concentrations is directly proportional to the fraction of protein bound to the oligonucleotide. This is the so-called 'fast-exchange limit' (Jardetzky and Roberts, 1981). However, the line widths of many of the resonances from residues within the β-sheet region increased three- to fivefold during the titration, placing these residues in the 'intermediate' to 'fast-exchange limit'. To include data from these residues in the binding constant determination, it is important to show that even at a relatively slow exchange rate the position of the peak maximum can be used to obtain the fraction of bound protein.

Figure 1 shows theoretical line shapes as a function of the fraction of bound protein and the exchange rate for a two-state system. In these calculations the kinetic on-rate was 1×10^8 s^{-1}, which is close to the diffusion-limited rate. The kinetic off-rate was varied to simulate the effect of the affinity constant on the line shape. Note that in these calculations the frequency separation between the bound and free forms (50 Hz) represents one of the largest frequency changes observed in this study. Thus, these simulations illustrate the largest effect of a reduced exchange rate on obtaining the fraction of bound protein from chemical shift changes.

The calculations presented in Fig. 1 show that, for affinity constants as high as 2×10^5 M^{-1}, the position of the peak maximum is essentially the same as that which would be obtained in the 'fast-exchange limit'. At an affinity of 4×10^5 M^{-1}, some error will be introduced for data obtained when there is excess protein. However, when the fraction of protein bound exceeds 0.5, the peak position of the line accurately reflects the true fraction of protein bound. Therefore, it is possible to obtain binding constants from chemical shift changes, even though some of the resonance lines show considerable (three- to fivefold) exchange broadening and are, therefore, not strictly in the 'fast-exchange limit'.

NMR spectroscopy

NMR spectra were recorded on a 500 MHz Varian UnityPlus spectrometer equipped with a Nalorac 5 mm triple resonance probe with a shielded z-gradient coil. Quadrature detection was obtained using the States–TPPI method (Marion et al., 1989). Table 1 provides a list of the triple resonance NMR experiments used to obtain the resonance- and sequence-specific assignments, along with basic experimental conditions. Pulsed field gradients were applied both to suppress water and to remove undesired coherences (Bax and Pochapsky, 1992). When necessary, gradients were added as z- and zz-filters to the pulse sequences in Table 1. Spin-lock purge pulses (Messerle et al., 1989) and spectral deconvolution in the time domain (Marion et al., 1989) were used to further suppress the water signal.

Two of the pulse sequences listed in Table 1, the CCH-TOCSY and the CN-NOESY, have not been described in

the literature at this time. The CCH-TOCSY sequence was designed to complement information obtained from the HCCH-TOCSY experiment. The degeneracies of aliphatic proton and carbon chemical shifts in Rho130 made side-chain assignments particularly difficult using data solely from the HCCH-TOCSY. The CCH-TOCSY experiment labels the magnetization with the ^{13}C chemical shift (in t_1). This magnetization is transferred to other ^{13}C nuclei within the side chain by isotropic mixing. The ^{13}C and ^{1}H chemical shifts of the recipient nuclei are recorded in t_2 and t_3, respectively. The CN-NOESY sequence was obtained by modifying the ^{15}N-NOESY-HSQC (Grzesiek et al., 1992) to record the carbon shift instead of the proton shift in t_1. This was accomplished by transferring magnetization to the carbon using a standard INEPT transfer (Morris and Freeman, 1979) and returning the magnetization to the attached proton in the same manner and is thus a projection of a 4D NOESY onto a ^{1}H plane. The combination of the CN-NOESY and ^{15}N-NOESY-HSQC spectra provided the carbon and proton shifts of protons displaying NOE connectivities to amide protons. These pulse sequences and the parameters used for these experiments are included in the Supplementary Material.

All spectra of Rho130 (liganded and unliganded) were obtained at pH 7.0 and 25 °C. A sweep width of 1600 Hz was used for the ^{15}N dimension in all experiments. A sweep width of 8400 Hz was used for the ^{13}C dimension, except for the HCCH-TOCSY and CCH-TOCSY experiments where the spectra were aliased twice (sw = 2800) to achieve higher resolution in the ^{13}C dimension. The ^{1}H and ^{13}C chemical shifts are given relative to 3-(trimethylsilyl)-propionic acid (TSP). The ^{15}N chemical shifts were indirectly referenced by multiplying the frequency of the standard with the ratio of the zero-point frequencies of ^{15}N with ^{1}H.

Data transformation and processing were done on Silicon Graphics workstations (Indigo R3000 and Indy R4600) using the FELIX software (Biosym Technologies, San Diego, CA, U.S.A.). Each dimension was zero-filled once and approximately 25% additional points were linearly predicted (Olejniczak and Eaton, 1990) in most of the indirectly detected heteronuclear dimensions. In some experiments, the first point was also modified by linear prediction to improve the flatness of the base plane.

Determining $J_{HN\text{-}HA}$ coupling constants

The three-bond coupling constant, $J_{HN\text{-}HA}$, was obtained from the HNHA spectrum as described by Vuister and Bax (1993). The coupling constant was calculated from the fraction of magnetization transferred from the H^N to the H^α proton. The ϕ dihedral angle was obtained from these coupling constants using the relationships described by Vuister and Bax (1993).

TABLE 1
PULSE SEQUENCES USED TO ASSIGN RHO130

Experiment	Acquired data matrix: size (nucleus)			Spectral widths (Hz)			No. of transients
	t_1	t_2	t_3	ω_1	ω_2	ω_3	
HSQC[a]	128 (N)	1024 (H)		1600	7000		8–32
HNCA[b]	32 (N)	48 (C)	1024 (H)	1600	3500	3000	16
HN(CO)CA[b]	32 (N)	48 (C)	1024 (H)	1600	3500	3000	16
CBCA(CO)NH[c]	48 (C)	28 (N)	512 (H)	8400	1600	3000	32
HNCACB[d]	48 (C)	32 (N)	512 (H)	8400	1600	3000	32
HBHA(CBCACO)NH[e]	48 (H)	32 (N)	512 (H)	4000	1600	3000	16
HA(CACO)NH[f]	48 (H)	32 (N)	512 (H)	2000	1600	3000	8
HN(CA)HA[g]	32 (N)	48 (C)	512 (H)	1600	2000	3000	16
HNCO[b]	32 (N)	48 (C)	512 (H)	1600	2500	3000	8
HCACO[h]	24 (C)	48 (CO)	512 (H)	3000	2500	2000	16
CBACOHA[i]	48 (C)	64 (CO)	256 (H)	8400	2500	2000	16
HNHA[j]	32 (N)	64 (H)	512 (H)	1600	5500	3000	16
HNHB[k]	32 (N)	64 (H)	512 (H)	1600	5500	3000	16
^{15}N-TOCSY-HMQC[l]	128 (H)	32 (N)	512 (H)	5500	1600	3000	16
HCCH-TOCSY[m]	96 (H)	48 (C)	512 (H)	3400	2800	3400	8
CCH-TOCSY[n]	64 (C)	32 (C)	1024 (H)	8400	2800	5000	8
^{15}N-NOESY-HSQC[o]	128 (H)	32 (N)	512 (H)	1600	5500	3000	16
NHNH-NOESY[p]	32 (N)	32 (N)	512 (H)	1600	1600	3000	32
CHH-NOESY[q]	48 (C)	128 (H)	512 (H)	8400	3400	6000	8
CN-NOESY[r]	48 (C)	32 (N)	256 (H)	8400	1600	3000	48

[a] Bodenhausen and Ruben, 1980.
[b] Grzesiek and Bax, 1992 (constant-time versions).
[c] Grzesiek and Bax, 1992.
[d] Wittekind and Mueller, 1993.
[e] Grzesiek and Bax, 1993.
[f] Jerala and Rule, 1996.
[g] Clubb et al., 1992.
[h] Kay et al., 1990.
[i] Kay, 1993.
[j] Vuister and Bax, 1993; Bax et al., 1994.
[k] Archer et al., 1991; Bax et al., 1994.
[l] Driscoll et al., 1990; τ_m = 45, 70, 90 ms.
[m] Bax et al., 1990a; τ_m = 8.9, 17.8, 26.7 ms.
[n] τ_m = 8.9, 17.8, 26.7 ms.
[o] Bax et al., 1990b; τ_m = 120 ms.
[p] Frenkiel et al., 1990; τ_m = 120 ms.
[q] Majumdar and Zuiderweg, 1993; τ_m = 120 ms.
[r] τ_m = 120 ms.

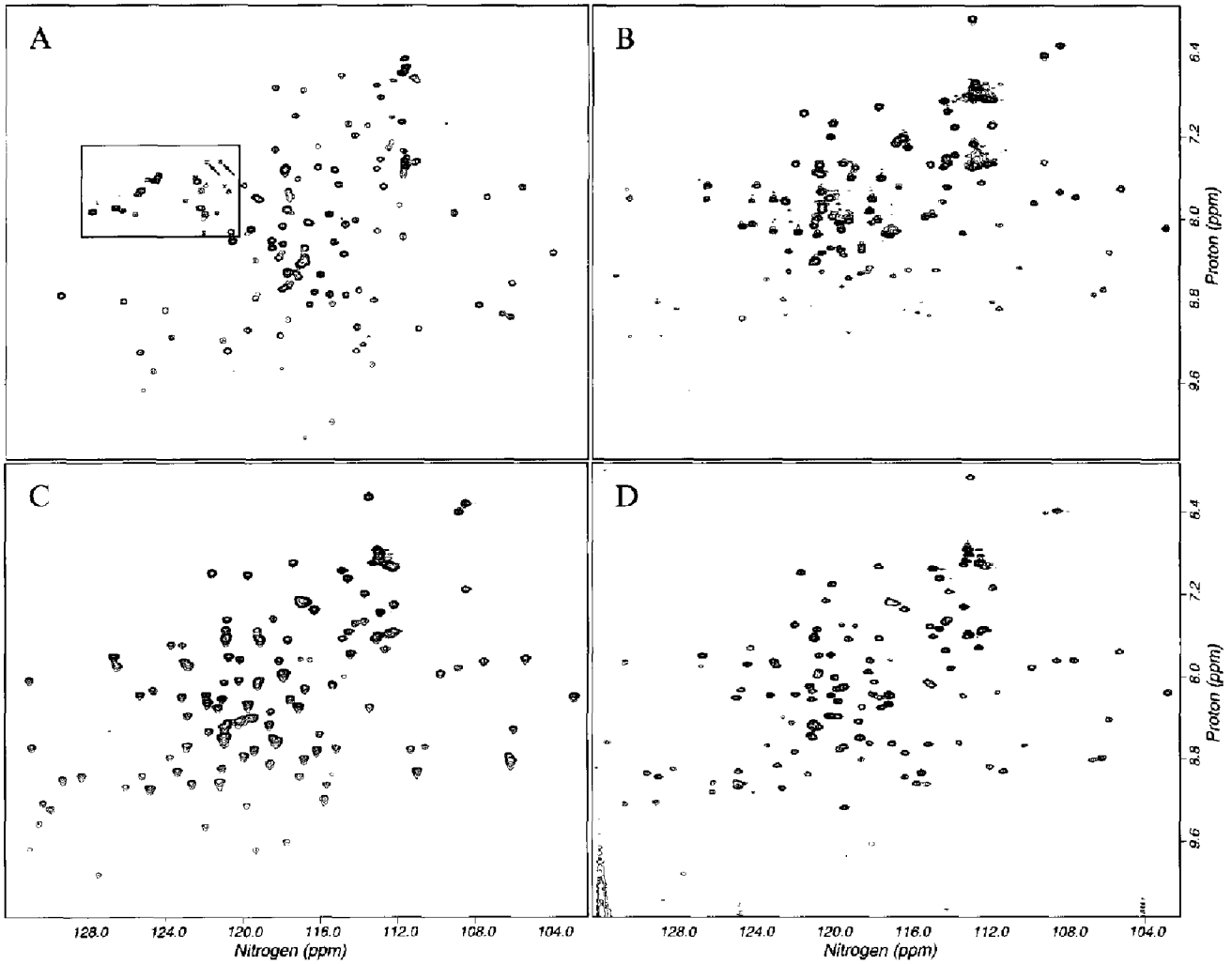

Fig. 2. HSQC spectra of Rho116 and Rho130, with and without DNA present. (A) Unliganded Rho116; (B) $(dC)_{10}$:Rho116; (C) unliganded Rho130; and (D) $(dC)_{10}$:Rho130. The protein concentration is approximately 1.5 mM, and the oligonucleotide concentrations were two- to threefold in excess of the protein concentration. The Rho116 spectra were obtained at pH 7.8, t = 25 °C, and the Rho130 spectra at pH 7.0, t = 25 °C.

Amide exchange

To determine whether the amide of each residue was protected, ^{15}N-labeled Rho130 was dissolved into D_2O buffer, and cross-peak intensities were measured from ^{15}N/^{1}H-HSQC spectra obtained over a time period of seven days. Since it took 9 h to exchange the solvent and to obtain the first spectrum, amides with exchange rates of greater than 0.11 h^{-1} could not be detected.

Residues whose amide protons exchange rapidly with solvent were identified by the presence of an intense cross peak at the chemical shift of water in the ^{15}N-NOESY-HSQC spectrum. Since these exchange peaks developed within the 120 ms mixing time, the exchange rate of these amide protons is on the order of 5 s^{-1}. This rate is comparable to the exchange rate for amide protons in unstructured peptides (Richarz et al., 1979; Wagner and Wüthrich, 1982); therefore, these residues are exposed to the surface and do not participate in stable hydrogen bonds.

^{15}N/^{1}H-heteronuclear NOEs

^{15}N/^{1}H heteronuclear dipolar couplings were obtained by recording two ^{15}N/^{1}H-HSQC spectra, one with amide-proton saturation and the other without amide-proton saturation. Heteronuclear dipolar coupling constants were determined by dividing the cross-sectional areas of each peak of the proton-saturated HSQC by the area of the corresponding peak in the unsaturated HSQC spectrum (Farrow et al., 1994).

Results

Rho130 as a model of RNA–rho interactions

Initially, Rho116 was chosen as a model of the RNA-binding domain of rho protein, as it had been previously shown to bind RNA with only 10-fold lower affinity than intact rho protein (Modrak and Richardson, 1994). Unliganded Rho116, however, is only soluble above pH 7.5 and between 10 and 30 °C. Figures 2A and B show ^{15}N/

^{1}H-HSQC NMR spectra of Rho116 in the absence and presence of $(dC)_{10}$, respectively. From the spectrum in Fig. 2A, it is clear that unliganded Rho116 is partially unfolded. Note, in particular, the presence of many narrow peaks in the boxed portion of the spectrum ($\delta_H \sim 8$ ppm, $\delta_N \sim 126$ ppm). Rho116 is stabilized in the presence of excess $(dC)_{10}$ and is soluble over a broader pH range (6.5–8.0). Since the ultimate goal of the structural studies of the RNA-binding domain of rho protein is to obtain an understanding of how rho protein binds with specificity for cytosine-containing nucleic acid, it is necessary to use a model of the rho RNA-binding domain that is stable in both liganded and unliganded states. Due to the conformational state of unliganded Rho116, we were prompted to search for a more structurally stable model of the N-terminal domain of rho protein.

The instability of unliganded Rho116 is probably a result of the high degree of hydrophobicity of its C-terminal residues. In intact rho, these residues are likely involved in important interactions that stabilize the folded form of the protein, but in Rho116 the same residues are likely solvent accessible. Since either Arg[128] or Lys[130] is a natural trypsinolysis site for rho protein and since adding residues 117–130 would increase the hydrophilicity of the C-terminus, we chose to clone and express the first 130 residues of rho protein (see Materials and Methods). Unliganded Rho130 is stable over wider temperature (10–40 °C) and pH (6.5–8.0) ranges than unliganded Rho116. The ^{15}N/^{1}H-HSQC NMR spectra of Rho130 in the absence and presence of DNA (Figs. 2C and D, respectively) show that Rho130 appears to be folded whether or not DNA is present, and while a number of cross peaks shift upon addition of $(dC)_{10}$ to Rho130, many cross peaks also remain at their original position, indicating that the core protein structure does not undergo large structural changes upon oligonucleotide binding.

The majority of the cross peaks in the unliganded Rho116 spectrum that are identical to peaks in the unliganded Rho130 spectrum correspond to residues in the N-terminal half of the protein. This supports the notion that Rho116 is unfolded in the C-terminal portion of the protein due to the hydrophobicity of the C-terminal residues. Also note that the spectra of liganded Rho116 and liganded Rho130 are quite similar, suggesting that the structure of the Rho116–DNA complex is similar to that of the Rho130–DNA complex.

Although Rho116 has been shown to be monomeric (Modrak and Richardson, 1994), it is possible that Rho130 contains residues which are involved in forming either a dimer or a hexamer. However, the hydrodynamic properties of Rho130 as determined by gel filtration chromatography (data not shown) suggest that Rho130 is mono-

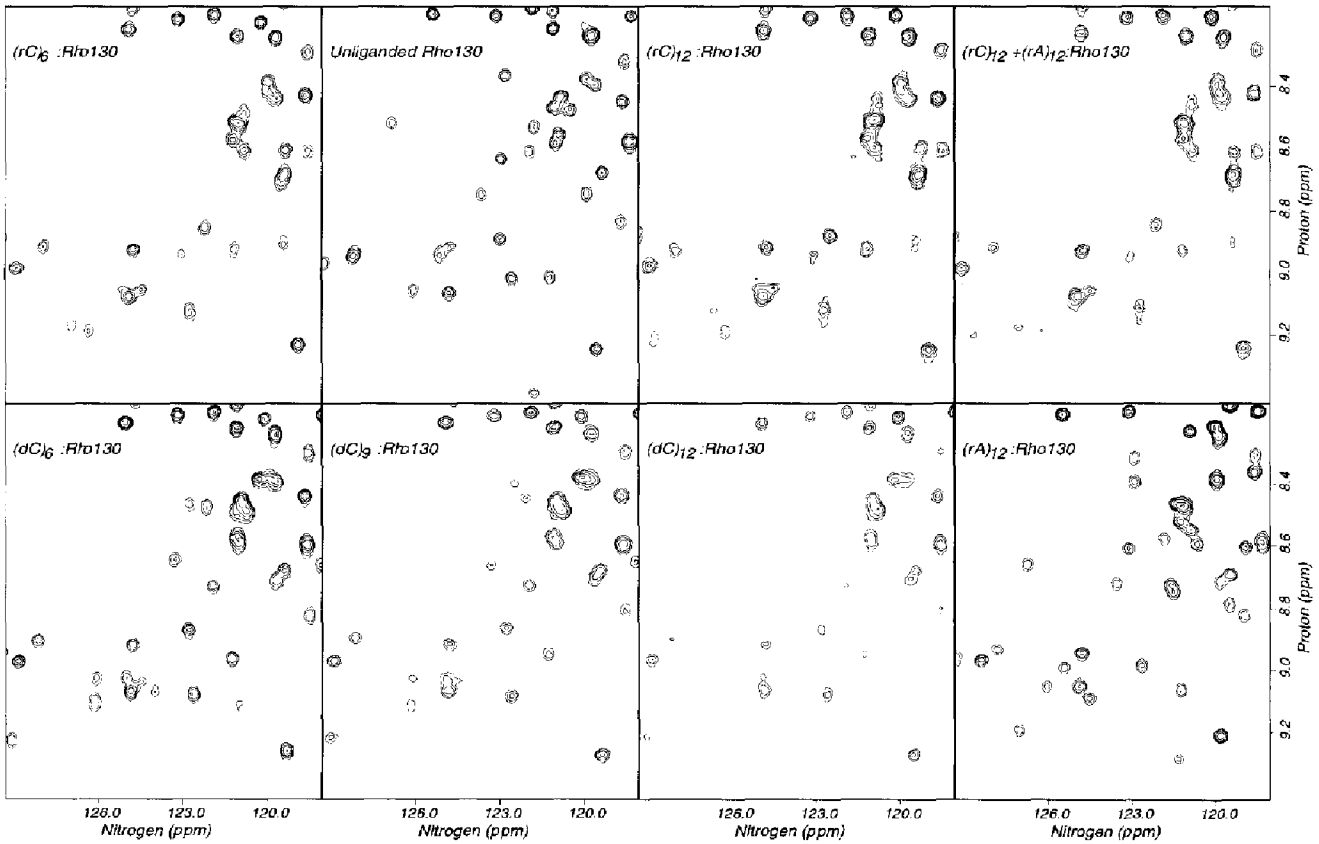

Fig. 3. HSQC spectra of rho130–nucleic acid complexes. An expanded region of the ^{15}N/^{1}H-HSQC spectrum is shown for 100 μM Rho130 complexed with the indicated nucleic acid. The nucleic acid concentration is in fivefold excess of the protein concentration.

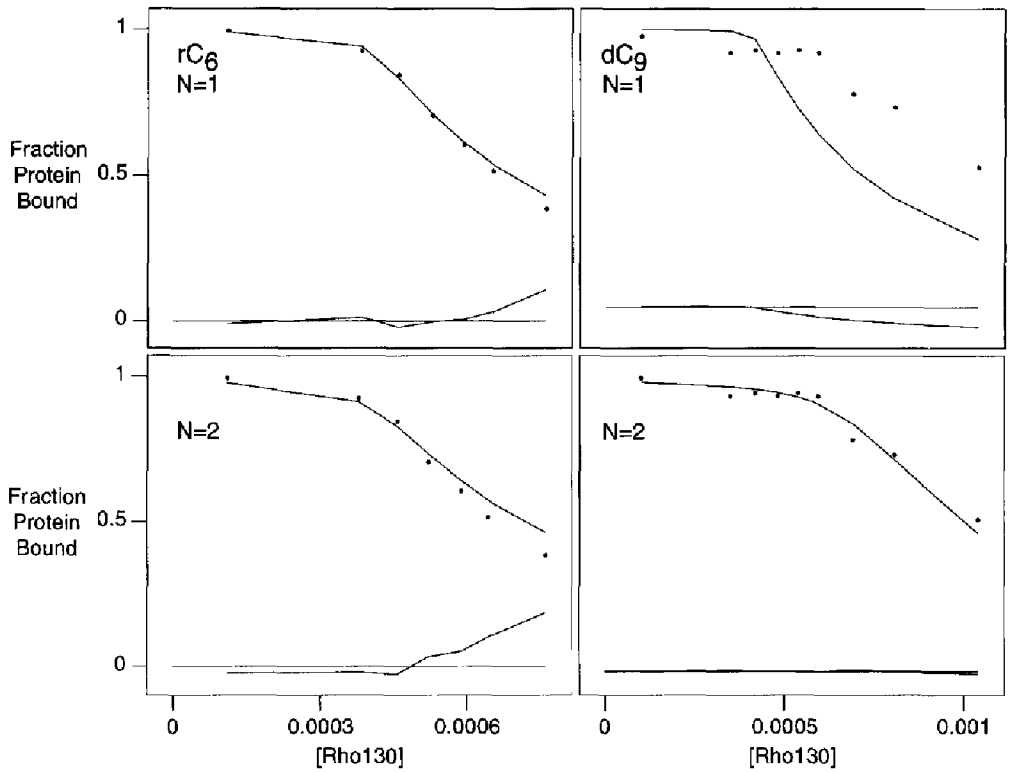

Fig. 4. Rho130–oligonucleotide binding constant determination using NMR. The binding curves were obtained from chemical shift changes of Asn^{117} during the addition of Rho130 to $(dC)_9$ and $(rC)_6$. Bullets represent experimental points and lines are the best fit to the experimental data. The lower solid lines within each panel show the residuals. The top two panels represent fits to the experimental data assuming a single-site binding model; the lower two panels are fits to the experimental data assuming a two-site binding model. Note that these binding curves do not resemble typical binding curves, as both ligand and protein concentrations are changing through the course of the titration. (The oligonucleotide concentration decreased from 0.5 to 0.25 mM during the titration.) While only data for Asn^{117} are shown, the chemical shift changes and binding constants of the other residues greatly affected by oligonucleotide binding were similar, whether the residues were undergoing fast or intermediate-to-fast exchange.

meric in solution. This conclusion is supported by the fact that the ^{15}N line widths (4–6 Hz) are of an appropriate size for a monomeric protein.

For Rho130 to be a reasonable model of the RNA-binding domain of rho protein, it must bind both RNA and DNA with similar affinity and specificity as compared to rho protein. The binding-site size, affinity, and specificity of Rho130 were investigated using NMR spectroscopy. The approximate binding-site size of Rho130 was determined by acquiring $^{15}N/^{1}H$-HSQCs of Rho130 fully saturated with dC oligonucleotides of varying length. If the oligonucleotide is of equal or shorter length than the binding site, then the oligonucleotide may bind in more than one way, resulting in the formation of different microstates. Under conditions of slow to intermediate exchange, line broadening due to chemical exchange between the microstates can occur. The exchange rate depends on the rate of conversion between the microstates. If the oligonucleotide is longer than the binding site, then more than one Rho130 molecule can bind a single oligonucleotide, resulting in line broadening due to an increase in the rotational correlation time. The narrowest spectral lines, therefore, may be obtained when the bound oligonucleotide has the same length or is slightly larger than the Rho130 binding site. The site size of intact rho protein is predicted to be about six nucleotides (Geiselmann et al., 1992a; Wang and Von Hippel, 1993b); therefore, oligonucleotide lengths of 6, 9, 10, 12, 14, and 16 were used to estimate the Rho130 site size. Several of the spectra are shown in Fig. 3. For oligonucleotides with a length of nine residues or longer, oligomerization is clearly occurring, as peaks progressively disappear from the spectra due to line broadening as the length of the oligonucleotide is increased (Fig. 3, second and third spectra of the second row). In fact, the HSQC spectrum of Rho130 complexed with $(dC)_{16}$ (data not shown) contains only peaks from mobile side chains of asparagines and glutamines, suggesting extensive aggregation in the presence of this oligonucleotide. In contrast, many of the cross peaks in the Rho130–$(rC)_6$ complex have line widths that are similar to those found in the spectrum of unliganded Rho130. On the basis of these data, the site size of Rho130 is less than nine and, therefore, similar to that of rho protein.

The binding constants for Rho130 to DNA and RNA oligonucleotides were determined by titrating oligonucleotides with Rho130 (see Materials and Methods). Since the bound and unbound species of Rho130 are in fast or intermediate-to-fast exchange, the amount of bound oligo-

nucleotide can be monitored by quantifying changes in the chemical shifts of various amide resonances as the protein goes from the bound state to the unbound state (see Materials and Methods and Fig. 1). Residues in the α-helical (N-terminal) portion of the protein are undergoing fast exchange, as there is almost no increase in line width through the course of the titration. However, residues in the β-sheet (C-terminal) portion of the protein are under conditions of intermediate-to-fast exchange, as their line widths increased three- to fivefold through the course of the titration. As shown in Fig. 1, the error in determining the fraction of bound protein under these conditions is fairly small, and in fact, the same binding constants were obtained for residues under fast or intermediate-to-fast exchange.

Figure 4 shows titration data for $(dC)_9$ and $(rC)_6$. For oligonucleotides with a length of six residues, a single-site binding model produced the best fit. The binding constant for $(rC)_6$ was found to be $5.0 \pm 3.0 \times 10^5$ M^{-1}. For oligonucleotides with lengths of nine or more, a two-site binding model fits the data far more accurately (which is in support of the premise that Rho130 oligomers form on longer oligonucleotides). The binding constants for $(dC)_9$ were found to be $1.1 \pm 0.5 \times 10^5$ M^{-1} for the first binding site and $2.3 \pm 0.4 \times 10^4$ M^{-1} for the second binding site.

To demonstrate that the specificity of RNA binding to Rho130 is similar to that of intact rho protein, the affinity of Rho130 to $(rA)_{12}$ was compared to the binding of Rho130 to $(rC)_{12}$ by a competition experiment. The spectra in Fig. 3 include $^{15}N/^1H$-HSQCs of Rho130 in the presence of an excess of $(rA)_{12}$, $(rC)_{12}$, and a 50:50 mixture of $(rA)_{12}$:$(rC)_{12}$. It is clear from these spectra that Rho130 binds both $(rC)_{12}$ and $(rA)_{12}$. There are numerous differences in chemical shifts between the two spectra, suggesting that these two oligonucleotides interact differently with Rho130. When equimolar amounts of both $(rC)_{12}$ and $(rA)_{12}$ are present, the HSQC spectrum is essentially identical to that of the Rho130–$(rC)_{12}$ complex. Therefore, the $(rC)_{12}$ has displaced most of the $(rA)_{12}$ from the binding site. Based on the known rho protein binding constants for $(rA)_{12}$ and $(rC)_{12}$, it would be expected that 95% of Rho130 be bound to $(rC)_{12}$ and only 5% of Rho130 be bound to $(rA)_{12}$ in an excess of an equimolar mixture of the oligonucleotides. These calculations are in agreement with the experimental data. Therefore, Rho130 exhibits a binding specificity similar to that of rho protein.

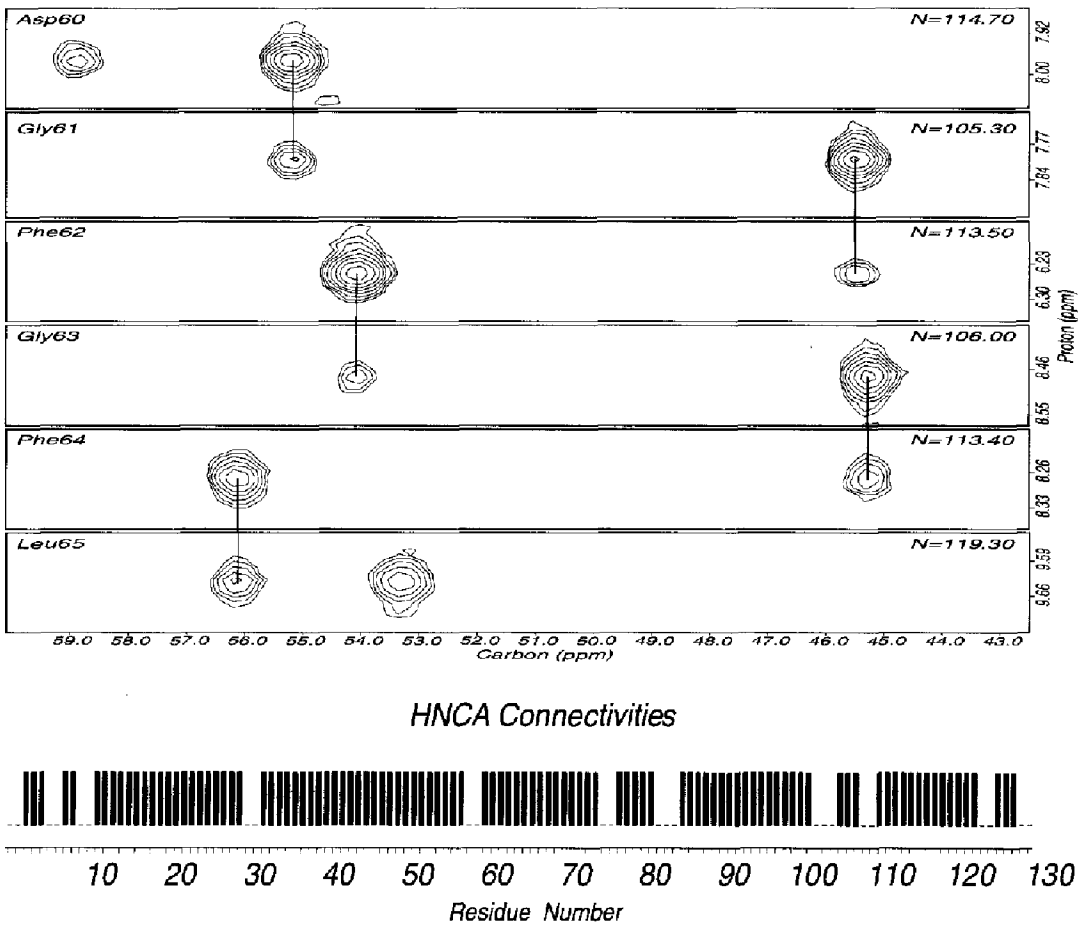

Fig. 5. Slices from the HNCA spectrum of Rho130. The spectra show C^{α}_{i-1} to C^{α}_i connectivities from Asp^{60} to Leu^{65}. The bars in the graph below the HNCA spectra indicate the residues for which a C^{α}_{i-1} to C^{α}_i connectivity was found.

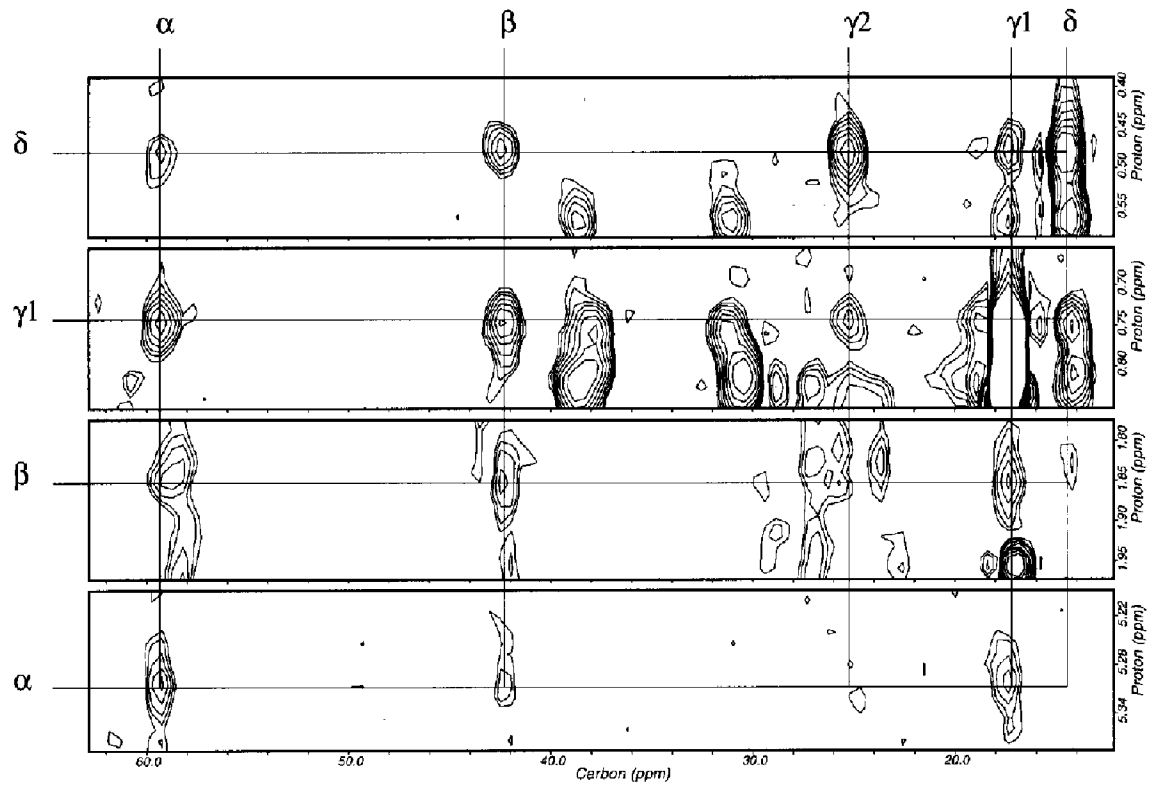

Fig. 6. Representative slices of the ω_1-ω_3 plane in a CCH-TOCSY spectrum. Spectra were recorded with $64 \times 32 \times 1024$ complex points and were processed to give a final data matrix of $128 \times 64 \times 1024$. The sample concentration was about 1.5 mM and the experimental time was approximately 36 h. The four slices show how the CCH-TOCSY was used to obtain assignments of an isoleucine side chain. Spectra were obtained with an isotropic mixing time of 17.8 ms.

Resonance assignments

The assignment of the main-chain atoms (H^N, N, C^α, H^α, CO) was accomplished using a number of standard triple resonance experiments (see Table 1). However, before acquiring triple resonance spectra, a number of residue- and sequence-specific assignments were obtained using selective labeling. These assignments were extremely useful as assignment starting points, because of H^N, N, C^α, and H^α chemical shift degeneracies in Rho130. Residue-specific assignments were obtained by ^{15}N-labeling a single residue. Sequence-specific assignments were made by simultaneously ^{15}N-labeling one residue type and ^{13}C-labeling the carbonyl of another residue type. If a ^{13}C-labeled residue directly preceded a ^{15}N-labeled residue, splitting of that amide resonance was observed in the ^{15}N/^{1}H-HSQC (Kainosho et al., 1987). This specific labeling technique was used for several residues (Phe, Tyr, Leu, Val, and Ile) and yielded 34 residue-specific assignments and six sequence-specific assignments.

Established triple resonance experiments were then implemented to obtain the rest of the assignments (Table 1). The first set of experiments to be used included HNCA, HN(CO)CA, and CBCA(CO)NH. The HNCA provided the C_i^α and C_{i-1}^α chemical shifts for each residue. The HN(CO)CA confirmed the C_{i-1}^α cross peaks and resolved C^α degeneracies of adjacent residues. The CBCA(CO)NH provided the C^β shift of the preceding residue, yielding residue-specific information. These C^β assignments were confirmed with an HNCACB experiment. Further residue-specific information, including the H^α and H^β chemical shifts, was obtained by acquiring HNHA, HNHB, ^{15}N-TOCSY-HMQC, and HBHA(CBCACO)NH spectra. The majority of the N, H^N, C^α, C^β, H^α, and H^β assignments were obtained from these experiments. Using this residue-specific information along with the previously determined sequence-specific assignments from the labeling experiments, assignments were obtained for the majority of the Rho130 main-chain nuclei. Figure 5 shows representative HNCA slices connecting the C^α resonances of Asp^{60} to Leu^{65}. Approximately 82% of the Rho130 residues can be successfully traced by C^α connectivities.

Gaps in the sequential assignment of Rho130, primarily due to the presence of prolines, were resolved using the HNCO, HCACO, and CBACOHA experiments. The HNCO correlates the carbonyl shift of a residue with the amide group (H^N, N) of the following residue. The HCACO then provided C^α and H^α frequencies. The HNCO, HCACO, and CBACOHA experiments also served to confirm the sequence-specific assignments obtained from the HNCA. The assignments of the Rho130 main-chain atoms are listed in the Supplementary Material.

Side-chain assignments

Side-chain assignments were difficult, since many H^{α} and C^{α} chemical shifts were degenerate. For residues with simple spin systems, such as glycine, serine, and alanine, complete resonance assignments were obtained while assigning the main-chain atoms. To obtain long-chain ^{1}H and ^{13}C assignments, several ^{15}N-TOCSY-HMQCs and HCCH-TOCSYs using different isotropic mixing times were obtained. To facilitate the assignment of unresolvable long-chain atoms, a 3D CCH-TOCSY was implemented. This experiment complements the HCCH-TOCSY, in that the CCH-TOCSY provides the carbon shift in t_1 and the HCCH-TOCSY provides the proton shift in t_1. Figure 6 shows an example of how the CCH-TOCSY experiment was used to obtain the resonance assignments for most of an isoleucine side chain (only $\gamma 2$ is missing). Using all three of the TOCSY experiments together, about 80% of the protons beyond H^{β} were assigned. Further, assignment of many of the glutamine and asparagine side-chain atoms was obtained from the HNCA and CBCA(CO)NH spectra. Since these experiments transfer magnetization from the amide through the carbonyl to the α-carbon, magnetization was also transferred from the side-chain amide protons to the aliphatic protons and carbons of the glutamine and asparagine side chains. Assignments for the Rho130 side chains may be found in the Supplementary Material.

Secondary structure determination of Rho130

The secondary structure of Rho130 has been determined by analysis of short- and long-range NOEs, chemical shifts, and $J_{HN\text{-}H\alpha}$ coupling constants. First, the presence of certain short-range NOEs is indicative of a residue's secondary structure. In particular, d_{NN} NOEs are

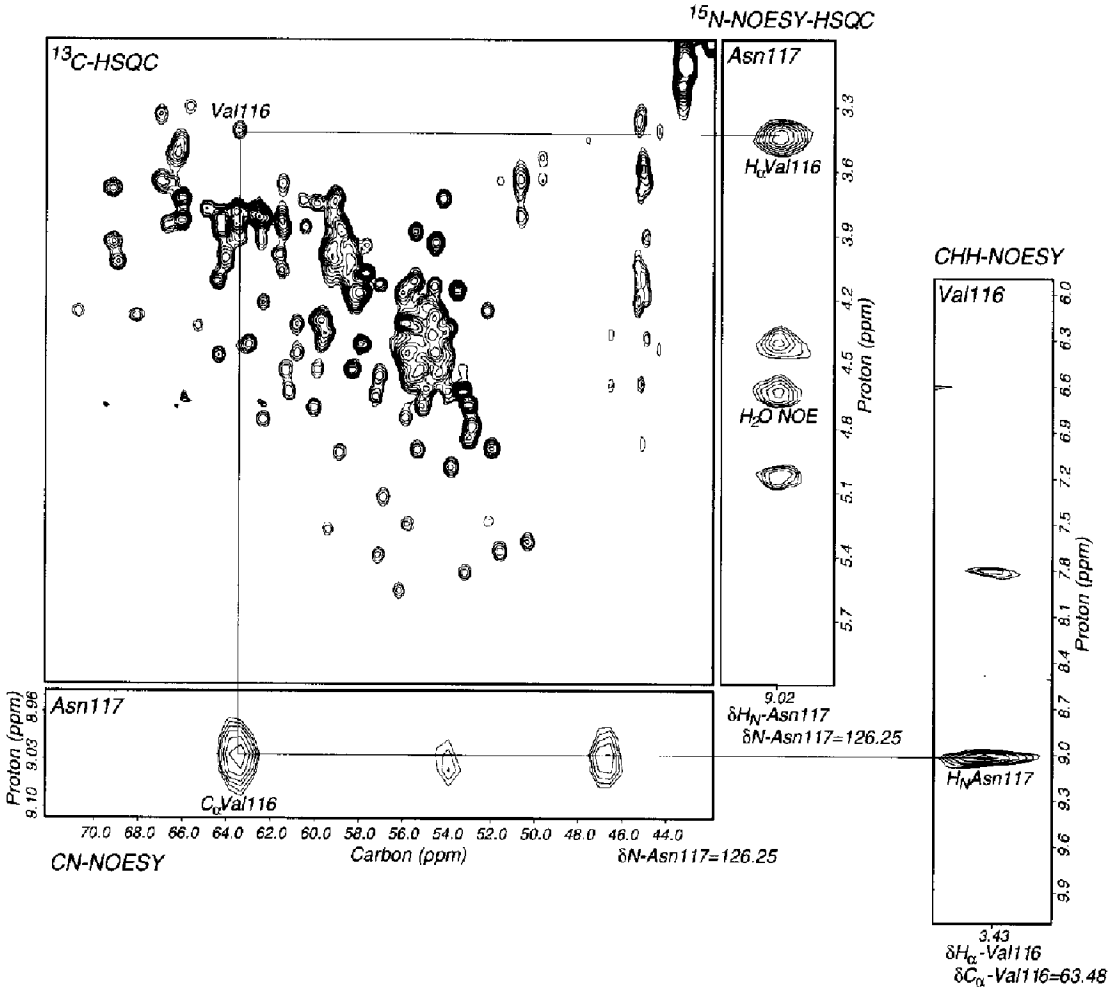

Fig. 7. A representative slice of the ω_1-ω_3 plane in a CN-NOESY spectrum, along with representative slices from ^{15}N-NOESY-HSQC and CHH-NOESY spectra. These slices are shown along with a portion of a $^{13}C/^{1}H$-HSQC spectrum to demonstrate how the CN-NOESY can be used along with other NOESY experiments to unambiguously assign NOEs. Here, an NOE assignment is made between the H^{N} of Asn^{117} and the H^{α} of Val^{116}. The ^{15}N-NOESY-HSQC provided the H^{α} shift of the proton that was dipolarly coupled to the H^{N} of Asn^{117}. The CN-NOESY provided the C^{α} shift of the carbon bound to the coupled proton. The possible number of C^{α} assignments for a given H^{α} NOE and vice versa is limited by referencing to the $^{13}C/^{1}H$-HSQC spectrum. The assignment was then confirmed by the presence of the Asn^{117} H^{N} shift at Val^{116} (C^{α}, H^{α}) in the CHH-NOESY. The CN-NOESY spectrum was recorded with $48 \times 32 \times 256$ complex points and was processed to give a final data matrix of $128 \times 64 \times 256$. The sample concentration was about 1.5 mM and the experimental time was approximately 48 h.

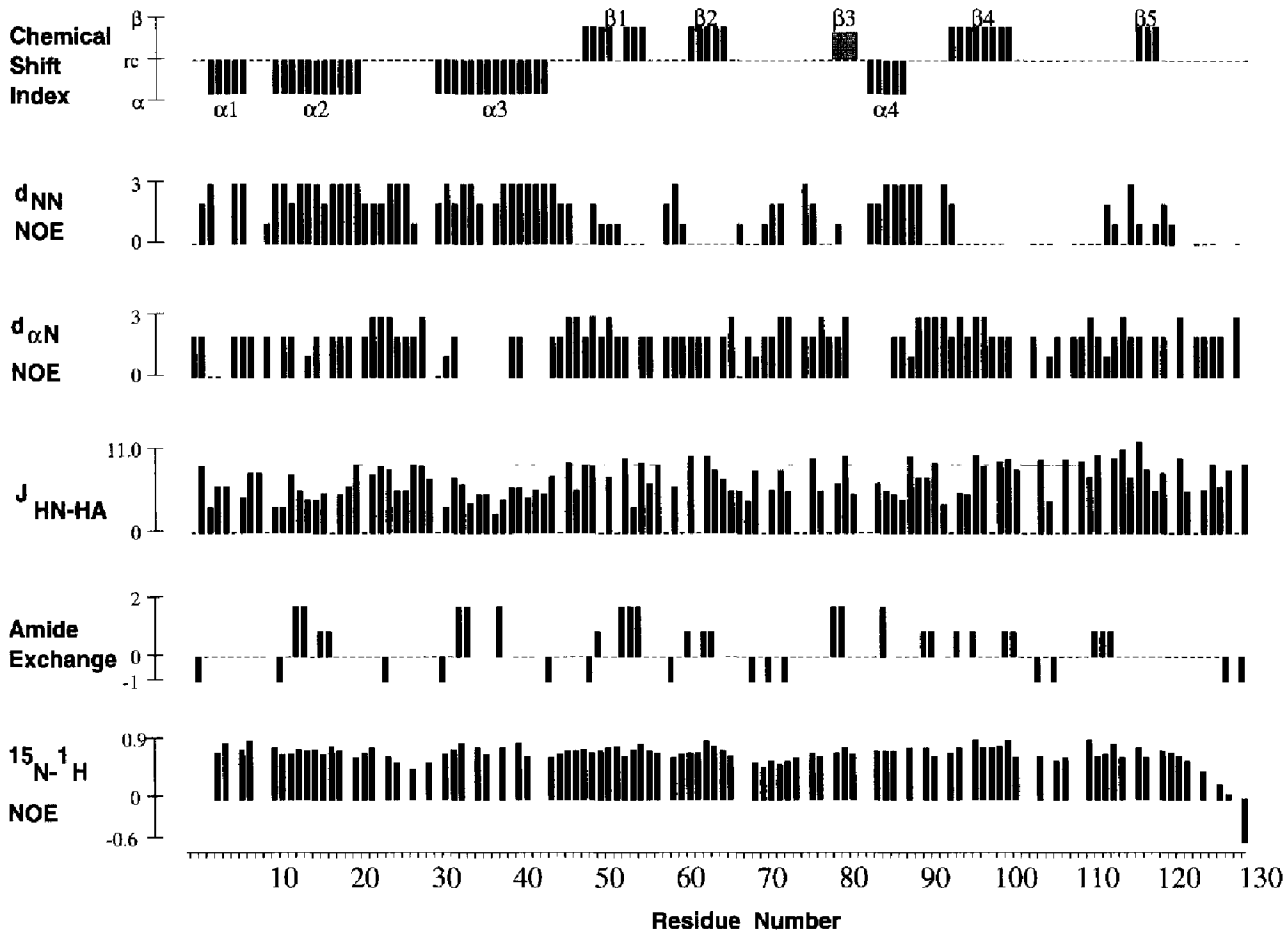

Fig. 8. The secondary structure of Rho130. Chemical Shift Index: the bars indicate the secondary structure predicted for each residue based on its C^{α}, C^{β}, CO, and H^{α} chemical shifts (Wishart and Sykes, 1994). The shaded box labeled β3 was not indicated by the CSI, but has been identified by both short- and long-range NOEs. d_{NN} NOEs: the height of the bar indicates the intensity, on an arbitrary scale, of the H^{N}_{i}-H^{N}_{i+1} NOE cross peak in ^{15}N-NOESY-HSQC and NHNH-NOESY spectra. Blanks indicate NOEs that cannot be determined because either one residue is absent in the HSQC or a residue is completely degenerate. $d_{\alpha N}$ NOEs: the height of the bar indicates the intensity, on an arbitrary scale, of the H^{α}_{i}-H^{N}_{i+1} cross peak in ^{15}N-NOESY-HSQC, CHH-NOESY, and CN-NOESY spectra. Blanks indicate NOEs that could not be determined because a residue was absent in one of the spectra or because a residue was completely degenerate. J_{HN-HA}: the height of the bar indicates the strength of the J_{HN-HA} coupling constant for a residue. α-Helical residues have coupling constants of approximately 4–7 Hz, while β-sheet residues have coupling constants of approximately 8–10 Hz. Residues with a coupling constant of '0 Hz' either could not be found in the HNHA or were completely degenerate with another residue (Vuister and Bax, 1993). Amide exchange: a value of −1 indicates that the amide is exposed to the protein surface and cannot participate in hydrogen bonds. A value of 0 indicates that the amide of the indicated residue is completely exchanged in 9 h and is, therefore, somewhat solvent exposed. A value of 1 indicates limited amide protection; complete amide exchange occurred in less than 25 h. A value of 2 indicates strong amide protection, with complete exchange not occurring in more than 25 h. A blank indicates that the degree of amide protection is undetermined due to degeneracy of the amide chemical shift. ^{15}N/^{1}H-heteronuclear NOE: the height of the bar indicates the intensity of the NOE; the highest bar is 0.92 and the lowest bar is −0.64. Blanks indicate residues for which the NOE could not be calculated, either because the amide of the residue is not present in an HSQC spectrum or because the amide is completely degenerate with that of another residue (Farrow et al., 1994).

much stronger for α-helical residues than for β-sheet residues. The high degree of aliphatic resonance degeneracies, however, made NOE assignments difficult. To definitively assign more NOEs, a CN-NOESY pulse sequence was implemented. This pulse sequence complements ^{15}N-NOESY-HSQC data by providing the carbon shifts of attached protons that are dipolarly coupled to amide protons, providing a strategy for assignments (see Fig. 7). The advantages of combining two 3D pulse sequences, the CN-NOESY and the ^{15}N-HSQC-NOESY, rather than implementing a 4D ^{13}C- and ^{15}N-separated NOESY are that higher digital resolution and sensitivity are achieved, as well as improved suppression of artifacts through additional phase cycling. All of the assignable d_{NN} NOEs for Rho130 and their respective intensities are shown in Fig. 8. Note that the d_{NN} connectivities are particularly strong in the first half of the Rho130 sequence, suggesting a significant degree of α-helicity in this region of the protein.

Secondary structural information can also be obtained from the chemical shifts of the main-chain atoms. In particular, the chemical shifts of C^{α}, C^{β}, CO, and H^{α} atoms depend on the type of secondary structure. Using the Chemical Shift Index (CSI; Wishart and Sykes, 1994),

the deviations of the four resonances were combined for each residue of Rho130, yielding a predicted secondary structure for the protein. The predicted secondary structure based on the combined CSI for assigned residues is shown in Fig. 8.

Information on secondary structure was also obtained from J_{HN-HA} coupling constants. J_{HN-HA} coupling constants of approximately 4–7 Hz correspond to a ϕ range of $-50°$ to $-90°$, which is indicative of α-helical residues. J_{HN-HA} coupling constants of 8–10 Hz correspond to a ϕ range of $-90°$ to $-150°$, which is indicative of β-sheet residues. Figure 8 shows the J_{HN-HA} coupling constants obtained for Rho130.

Combining these secondary structure determination techniques yields a fairly definitive secondary structure for Rho130. There are three α-helices, followed by a five-stranded β-sheet with another short helix present between the third and fourth strands of the sheet. The β-sheet region has been confirmed and further refined by analyzing long-range NOEs, including H^N-H^N, H^N-H^α, and H^α-H^α NOEs (Fig. 9). From these data it appears that the first β-strand is considerably longer than the others, leading to the formation of two three-stranded β-sheets with the first β-strand common to both sheets.

Further structural information about Rho130 is provided by measuring the degree of amide protection and $^{15}N/^1H$ heteronuclear NOEs. Slowly exchanging amide protons were identified by following the loss of peaks from a $^{15}N/^1H$-HSQC spectrum after Rho 130 had been transferred to a D_2O NMR buffer. Figure 8 shows which residues had measurable intensity after 9 or 25 h in D_2O, corresponding to exchange rates of approximately 2×10^{-5} and 8×10^{-6} s^{-1}, respectively. Residues whose amide protons exchange rapidly with solvent (exchange rate ~5 s^{-1}) are also shown in Fig. 8. These data are in good agreement with the predicted secondary structure of Rho130. Residues that are involved in α-helical or β-sheet secondary structure appear to be protected from amide exchange, whereas those that are part of loops undergo rapid amide exchange.

The conformational dynamics of Rho130 were investigated by measurement of $^{15}N/^1H$-heteronuclear dipolar coupling (heteronuclear NOE). Heteronuclear NOEs for Rho130 are shown in Fig. 8. The heteronuclear NOE can vary from a maximum of 0.82 for an immobile residue to a minimum of -3.6 for a mobile residue (Kay et al., 1989). The heteronuclear NOEs of the Rho130 residues clearly decrease for residues in the interhelical (residues 23–28) and inter-β-strand (residues 67–75), confirming that these residues are somewhat more mobile than the residues

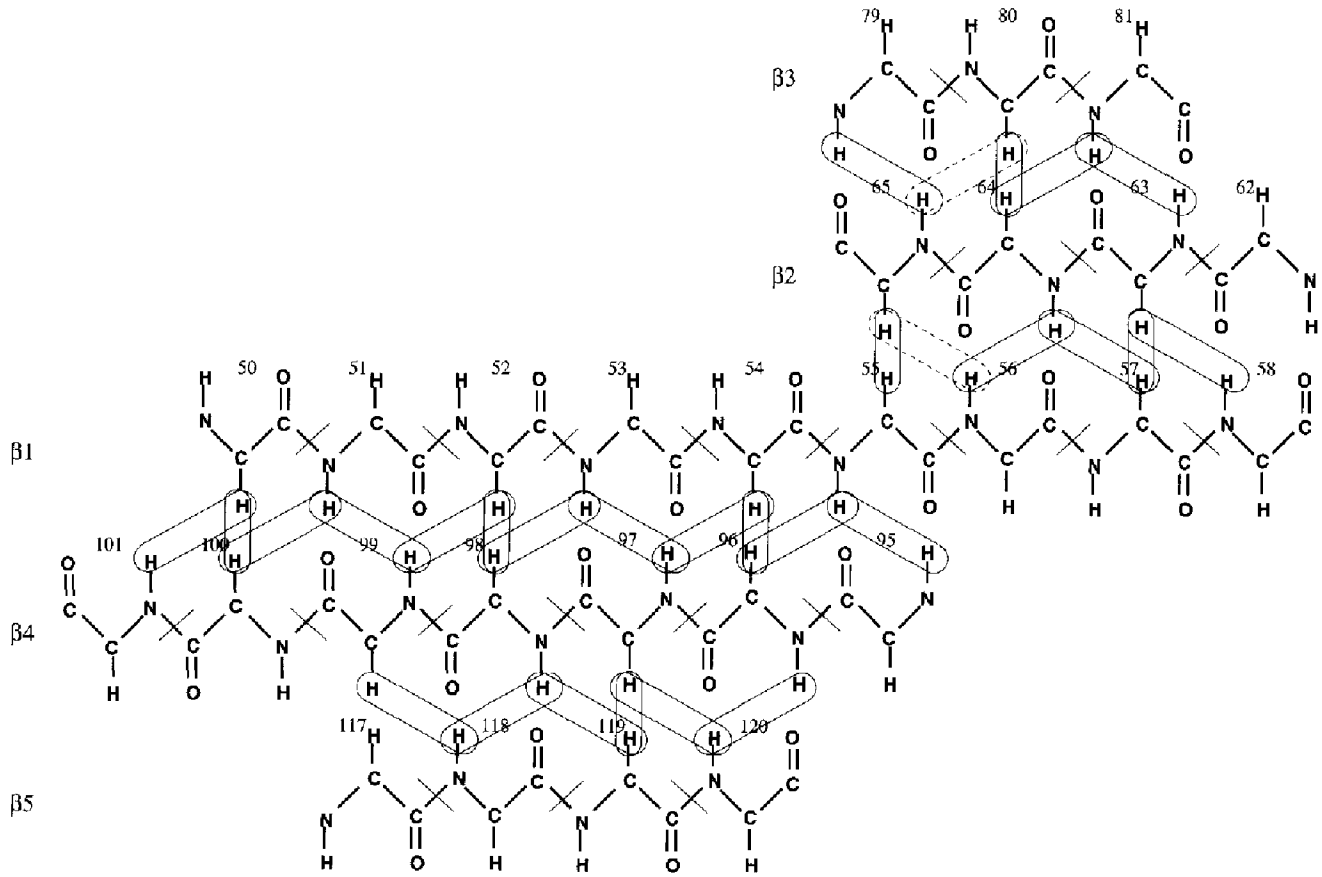

Fig. 9. Five-stranded β-sheet of Rho130. The five β-strands of Rho130 are shown with the assigned long-range H^N-H^N, H^N-H^α, and H^α-H^α NOEs indicated with solid ovals. The NOEs shown with dashed ovals cannot be confirmed unambiguously due to degeneracies.

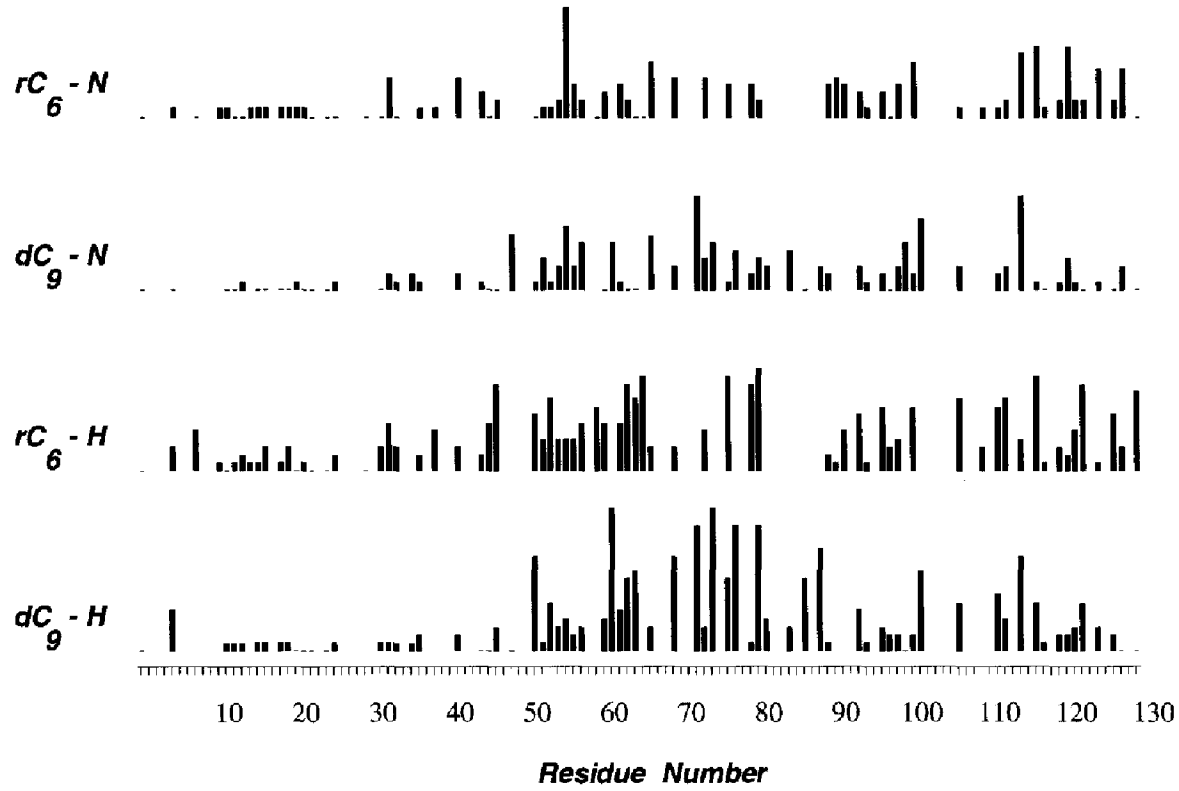

Fig. 10. Residues affected by oligonucleotide binding. The absolute value of the chemical shift change for the N and H^N of each residue due to oligonucleotide binding was determined by comparing $^{15}N/^1H$-HSQC spectra of the Rho130–$(dC)_9$ and Rho130–$(rC)_6$ complexes with the spectrum of unliganded Rho130. The highest bars represent changes of 0.18 and 1.4 ppm for ^1H and ^{15}N shifts, respectively. Residues with blanks are either degenerate in the $^{15}N/^1H$-HSQCs or could not be clearly followed through the course of the titration.

involved in secondary structure. The heteronuclear NOEs also show an increase in mobility for the last 6–7 residues of the C-terminus. Thus, the C-terminus of the rho RNA-binding domain appears to be connected by a flexible tether to the rest of the rho protein, likely forming a disordered and exposed loop that is accessible to trypsin.

Effects of oligonucleotide binding on Rho130 residues

Since the chemical shift changes of most Rho130 residues can be followed during an oligonucleotide titration, it was possible to determine the total chemical shift change of each residue that occurs due to oligonucleotide binding. In general, residues with large chemical shift changes are those residues which undergo the greatest environmental change upon oligonucleotide binding. Residues with little to no change in chemical shift are likely in a similar environment, whether or not oligonucleotide is present. Figure 10 shows the degree of N and H^N chemical shift changes for all nondegenerate residues of Rho130 that could be clearly followed through the course of the titration. From these data, it appears that residues 50–130 are involved in oligonucleotide binding, while residues 1–50 are not. Although the heteronuclear NOE data suggest that residues C-terminal to residue 120 are disordered in the unliganded protein, it is clear from the data presented in Fig. 10 that these residues are affected by the binding of nucleic acid and, therefore, may contribute to oligonucleotide binding.

Discussion

The initial goal of this study was to employ Rho116 as a model for the RNA-binding domain of rho protein. However, the $^{15}N/^1H$-HSQC spectrum of unliganded Rho116 indicated that it is partially unfolded in the absence of DNA. It is not clear at this time why Rho116 is less stable than Rho130. Figure 8 indicates that Rho130 has well-defined secondary structure up to residue 120. This statement is supported by the high values of the $^{15}N/^1H$ dipolar coupling and the existence of a β-strand between residues 117–120. Therefore, these residues are probably necessary for stabilization of the β-sheet in Rho130. From the unliganded Rho116 and Rho130 spectra, it is clear that the β-sheet of Rho116 is partially unfolded in the absence of oligonucleotides, but becomes more ordered in the presence of oligonucleotides. The reduced affinity of Rho116 for $(rC)_n$ may reflect the free energy required to constrain this region of the protein in the conformation necessary for RNA binding.

HSQC spectra of Rho130 indicate that this rho fragment is fully folded in the absence and presence of oligonucleotides, and thus, Rho130 is amenable to structural

studies of both the bound and free states. To demonstrate that Rho130 is a relevant model for the RNA-binding domain of rho protein, the binding affinity, specificity, and site size of nucleic acid binding to Rho130 were evaluated. As is shown in Fig. 4, Rho130 binds $(rC)_6$ with an affinity of about 5×10^5 M^{-1}, approximately twofold less than intact rho. This decrease in binding can be explained by the use of higher ionic strength buffer in NMR studies. The effect of ionic strength on the binding of oligonucleotides to Rho130 can be estimated from studies that showed a threefold reduction in binding of rho to $(dC)_{11}$ if the salt concentration was raised from 100 to 560 mM KCl (Geiselmann et al., 1992a). Therefore, if the binding constants obtained for Rho130 are corrected for an ionic strength of 150 mM K_2SO_4, they should be on the order of 1×10^6 M^{-1} or higher. Thus, the binding constant of Rho130 to $(rC)_6$ is similar to that found for intact rho, which ranges from 0.6×10^6 M^{-1} to 4×10^6 M^{-1} for binding of $(rC)_{8\text{-}10}$ (Geiselmann et al., 1992a; Wang and Von Hippel, 1993a; Modrak and Richardson, 1994). The binding specificity of Rho130 was investigated by comparing Rho130 binding to $(rC)_{12}$ and $(rA)_{12}$. Like rho protein, Rho130 binds both $(rC)_{12}$ and $(rA)_{12}$, but also like rho protein, Rho130 binds $(rC)_{12}$ with approximately 100-fold higher affinity. The binding-site size of Rho130 is clearly less than nine nucleotides, which is expected, as the binding-site size of intact rho protein is five to six nucleotides (Geiselmann et al., 1992a; Wang and Von Hippel, 1993b). Therefore, these binding studies show that the binding affinity, specificity, and site size of Rho130 are similar to those of rho protein, making Rho130 an excellent model of the RNA-binding domain of rho protein.

These binding studies also indicated that Rho130 oligomerizes in the presence of longer oligonucleotides. While investigating the binding-site size of Rho130, oligomerization was observed in the presence of oligonucleotides that were nine bases or longer. Figure 3 clearly shows that as the length of (dC) oligonucleotides is increased, the line widths in the HSQC spectra broaden. These data could be explained by at least three phenomena, i.e., chemical exchange between microstates of the bound DNA, nonspecific aggregation, or oligomerization. Since the number of microstates should decrease as the length of the DNA increases, it is not likely that microstates are contributing to the line broadening. Although nonspecific aggregation could also explain these data, the formation of well-defined polymers is the more likely explanation for at least two reasons. First, extensive line broadening is only observed with long oligonucleotides; if DNA binding caused the protein to form nonspecific aggregates, then substantial effects would also be seen with shorter oligonucleotides (e.g. $(dC)_9$). Second, the binding data obtained for $(dC)_9$ can be well fit to a two-site binding model, suggesting the formation of a complex with a defined stoichiometry.

The formation of nucleic acid polymers due to protein–nucleic acid interactions has been observed in other systems as well. For example, the binding of $(dT)_{20}$ to recA filaments induces the formation of DNA–DNA contacts (Hockensmith et al., 1993). The Trp repressor, when bound to a 17-residue duplex oligonucleotide containing a half-site operator sequence, forms extended copolymers of the protein and nucleic acid (Lawson and Carey, 1993). It is not known at this time whether the oligomerization of Rho130 is related to the formation of dimers and hexamers in intact rho.

The oligonucleotide-induced chemical shift changes presented in Fig. 10 indicate that residues 50–120 probably represent the primary contact site for oligonucleotide binding. The secondary structure predicted for this region is largely β-sheet, with a single short helix between residues 84–88. Thus, Rho130 is not an RRM, as the two α-helices packed against the β-sheet in RRMs are missing in the rho RNA-binding motif. Based on assigned long-range H^N-H^N, H^N-H^α, and H^α-H^α NOEs, Rho130 appears to contain a pair of three-stranded β-sheets, β1-β2-β3 and β1-β4-β5 (Fig. 9). These sheets are in good agreement with the five-stranded β-sheet seen in CSD proteins, and the β1-β2-β3 strands are also in good agreement with those predicted by Martinez et al. (1996a). Thus, the secondary structure of Rho130 appears to be more like the CSD structure than like the RRM structure.

Conclusions

In summary, it has been shown that Rho130 is a good model for structural studies of the RNA-binding domain of rho protein. The main-chain assignments and a majority of the side-chain assignments of Rho130 have been obtained, and the secondary structure of Rho130 has been evaluated using various NMR techniques. The structure of Rho130 is clearly different from the RRM-containing RNA-binding proteins, but is similar to the CSD-containing RNA/DNA-binding proteins. Future reports will include the high-resolution structures of Rho130 with and without bound oligonucleotides. A detailed analysis of these structures will address the important questions of rho's binding affinity and specificity.

Acknowledgements

We gratefully acknowledge Dr. Joel Hockensmith for many insightful discussions concerning solution conditions for NMR spectroscopy and analysis of the binding data. This research was supported by the American Heart Association (94010790). The NMR spectrometer was purchased by a shared equipment grant from the National Science Foundation (BIR-9217013). Equipment for protein purification was obtained, in part, by a shared equipment grant from the National Science Foundation (BIR-9216996).

References

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

Bax, A., Clore, M. and Gronenborn, A.M. (1990a) *J. Magn. Reson.*, **88**, 425–431.

Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990b) *J. Magn. Reson.*, **36**, 304–318.

Bax, A. and Pochapsky, S.S. (1992) *J. Magn. Reson.*, **99**, 638–643.

Bax, A., Vuister, G.W., Grzesiek, S., Delaglio, F., Wang, A.C., Tschudin, R. and Zhu, G. (1994) *Methods Enzymol.*, **239**, 79–105.

Bodenhausen, G. and Ruben, D.G. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Brennan, C.A., Dombroski, A.J. and Platt, T. (1987) *Cell*, **48**, 945–952.

Brennan, C.A. and Platt, T. (1991) *J. Biol. Chem.*, **266**, 17296–17305.

Burd, C. and Dreyfuss, G. (1994) *Science*, **265**, 615–621.

Cantor, C. and Tinoco, I. (1965) *J. Mol. Biol.*, **13**, 65–77.

Cantor, C., Warshaw, M. and Shapiro, H. (1970) *Biopolymers*, **9**, 1059–1077.

Carrington, A. and McLachlan, A.D. (1967) *Introduction to Magnetic Resonance with Applications to Chemistry and Chemical Physics*, Harper and Row, New York, NY, U.S.A., pp. 204–208.

Clubb, R.T., Thanabal, V. and Wagner, G. (1992) *J. Biomol. NMR*, **2**, 203–210.

Dolan, J.W., Marshall, N.F. and Richardson, J.P. (1990) *J. Biol. Chem.*, **265**, 5747–5754.

Driscoll, P.C., Clore, G.M., Marion, D., Wingfield, P.T. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 3542–3556.

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

Faus, I. and Richardson, J.P. (1990) *J. Mol. Biol.*, **212**, 53–66.

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

Galluppi, G.R. and Richardson, J.P. (1980) *J. Mol. Biol.*, **138**, 513–539.

Geiselmann, J. and Von Hippel, P.H. (1992) *Protein Sci.*, **1**, 850–860.

Geiselmann, J., Yager, T.D., Gill, S.C., Calmettes, P. and Von Hippel, P.H. (1992a) *Biochemistry*, **31**, 111–121.

Geiselmann, J., Seifried, S.E., Yager, T.D., Liang, C. and Von Hippel, P.H. (1992b) *Biochemistry*, **31**, 121–132.

Geiselmann, J., Wang, Y., Seifried, S.E. and Von Hippel, P.H. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 7754–7758.

Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **99**, 201–207.

Grzesiek, S., Dobeli, H., Gentz, R., Garotta, G., Labhardt, A.M. and Bax, A. (1992) *Biochemistry*, **31**, 8180–8190.

Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 185–204.

Haynes, S. (1992) *New Biologist*, **4**, 421–429.

Hockensmith, J.W., Kubasek, W.L., Vorachek, W.R. and Von Hippel, P.H. (1993) *J. Biol. Chem.*, **268**, 15712–15720.

Hoffman, D.W., Query, C.C., Golden, B.L., White, S.W. and Keene, J.D. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 2495–2499.

Jardetzky, O. and Roberts, G. (1981) *NMR in Molecular Biology*, Academic Press, New York, NY, U.S.A., pp. 120–122.

Jerala, R. and Rule, G.S. (1996) *J. Magn. Reson.*, **B110**, 87–90.

Jin, D.J., Burgess, R.R., Richardson, J.P. and Gross, C.A. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1453–1457.

Kainosho, M., Nagao, H. and Tsuji, T. (1987) *Biochemistry*, **26**, 1068–1075.

Kay, L., Torchia, D. and Bax, A. (1989) *Biochemistry*, **28**, 8972–8979.

Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.

Kay, L.E. (1993) *J. Am. Chem. Soc.*, **115**, 2055–2057.

Kunkel, T.A., Bebenek, K. and McClary, J. (1991) *Methods Enzymol.*, **204**, 125–139.

Landsman, D. (1992) *Nucleic Acids Res.*, **20**, 2861–2864.

Lawson, C. and Carey, J. (1993) *Nature*, **366**, 178–182.

Majumdar, A. and Zuiderweg, E.R.P. (1993) *J. Magn. Reson.*, **B102**, 242–244.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Martinez, A., Opperman, T. and Richardson, J. (1996a) *J. Mol. Biol.*, **257**, 895–908.

Martinez, A., Burns, C. and Richardson, J. (1996b) *J. Mol. Biol.*, **257**, 909–918.

McSwiggen, J.A., Bear, D.G. and Von Hippel, P.H. (1988) *J. Mol. Biol.*, **199**, 609–622.

Messerle, B.A., Wider, G., Otting, G., Weber, C. and Wüthrich, K. (1989) *J. Magn. Reson.*, **85**, 608–613.

Modrak, D. and Richardson, J.P. (1994) *Biochemistry*, **33**, 8292–8299.

Morgan, W.D., Bear, D.G. and Von Hippel, P.H. (1983) *J. Biol. Chem.*, **258**, 9565–9574.

Morris, G. and Freeman, R. (1979) *J. Am. Chem. Soc.*, **101**, 760–762.

Nagai, K., Oubridge, C., Jessen, T.H., Li, J. and Evans, P.R. (1990) *Nature*, **348**, 515–520.

Olejniczak, E. and Eaton, H. (1990) *J. Magn. Reson.*, **87**, 628–632.

Penington, C.J. and Rule, G.S. (1992) *Biochemistry*, **31**, 2912–2920.

Platt, T. and Richardson, J.P. (1992) *Transcriptional Regulation*, Cold Spring Harbor Laboratory Press, Plainview, NY, U.S.A., pp. 365–388.

Platt, T. (1994) *Mol. Microbiol.*, **11**, 983–990.

Richardson, J.P. (1982) *J. Biol. Chem.*, **257**, 5760–5766.

Richardson, J.P. (1996) *J. Biol. Chem.*, **271**, 1251–1254.

Richardson, L.V. and Richardson, J.P. (1992) *Nucleic Acids Res.*, **20**, 5383–5387.

Richarz, R., Sehr, P., Wagner, G. and Wüthrich, K. (1979) *J. Mol. Biol.*, **130**, 19–30.

Schindelin, H., Marahiel, M.A. and Heinemann, U. (1993) *Nature*, **364**, 164–168.

Schnuchel, A., Wiltscheck, R., Czisch, M., Herrler, M., Willimsky, G., Graumann, P., Marahiel, M.A. and Holak, T.A. (1993) *Nature*, **364**, 169–171.

Seifried, S.E., Easton, J.B. and Von Hippel, P.H. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 10454–10458.

Stanley, W. and Bock, R. (1965) *Anal. Biochem.*, **13**, 43–65.

Steinmetz, E.J. and Platt, T. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 1401–1405.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) *Methods Enzymol.*, **185**, 60–89.

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

Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **160**, 343–361.

Wang, Y. and Von Hippel, P.H. (1993a) *J. Biol. Chem.*, **268**, 13940–13946.

Wang, Y. and Von Hippel, P.H. (1993b) *J. Biol. Chem.*, **268**, 13947–13955.

Wishart, D.S. and Sykes, B.D. (1994) *J. Biomol. NMR*, **4**, 171–180.

Wittekind, M., Gorlach, M., Friedrichs, M., Dreyfuss, G. and Mueller, L. (1992) *Biochemistry*, **31**, 6254–6265.

Wittekind, W. and Mueller, L. (1993) *J. Magn. Reson.*, **B101**, 201–205.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, **33**, 103–119.