Protein Science (1996), 5:969-972. Cambridge University Press. Printed in the USA.

FOR THE RECORD

^{1}H and ^{15}N nuclear magnetic resonance assignment and secondary structure of the cytotoxic ribonuclease α-Sarcin

RAMÓN CAMPOS-OLIVAS,[1] MARTA BRUIX,[1] JORGE SANTORO,[1]
ALVARO MARTÍNEZ DEL POZO,[2] JAVIER LACADENA,[2]
JOSÉ G. GAVILANES,[2] AND MANUEL RICO[1]

[1] Instituto de Estructura de la Materia, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain
[2] Departamento de Bioquímica y Biología Molecular I, Facultad de Química, Universidad Complutense, 28040 Madrid, Spain

(RECEIVED December 27, 1995; ACCEPTED February 16, 1996)

Abstract: The ribosome-inactivating protein α-Sarcin (αS) is a 150-residue fungal ribonuclease that, after entering sensitive cells, selectively cleaves a single phosphodiester bond in an universally conserved sequence of the major rRNA to inactivate the ribosome and thus exert its cytotoxic action. As a first step toward establishing the structure-dynamics-function relationships in this system, we have carried out the assignment of the ^{1}H and ^{15}N NMR spectrum of αS on the basis of homonuclear (^{1}H-^{1}H) and heteronuclear (^{1}H-^{15}N) two-dimensional correlation spectra of a uniformly ^{15}N-labeled sample, and two selectively ^{15}N-labeled (Tyr and Phe) samples, as well as a single three-dimensional experiment. The secondary structure of αS, as derived from the characteristic patterns of dipolar connectivities between backbone protons, conformational chemical shifts, and the protection of backbone amide protons against exchange, consists of a long N-terminal β-hairpin, a short α-helical segment, and a C-terminal β-sheet of five short strands arranged in a +1,+1,+1,+1 topology, connected by long loops in which the 13 Pro residues are located.

Keywords: α-Sarcin; nuclear magnetic resonance; protein structure; ribonuclease; ribosome-inactivating protein

α-Sarcin is a cytotoxin produced by the mold *Aspergillus giganteus*; it is a ribosome-inactivating protein and inhibits protein synthesis by cleaving a single phosphodiester bond in a strictly conserved sequence in the largest rRNA (Wool, 1984), which has been implicated in translocation (Wool et al., 1992). Additionally, the protein must enter cells to exert its ribonucleolytic action on the ribosome, translocating across certain types of membranes (for a review see Gasset et al., 1994). Therefore, this protein interacts with membrane lipids and hydrolyzes ribosomal RNA with high specificity.

The amino acid sequence of αS (150 residues) is almost coincident with those of three other proteins produced by different *Aspergillus* strains: mitogillin and restrictocin, both from *A. restrictus* (85% identity), and allergen I from *A. fumigatus* (86% identity). These constitute a clearly defined family of fungal extracellular ribonucleases, more closely related to the group of the smaller fungal ribonucleases of the T_1 family (Sacco et al., 1983; Mancheño et al., 1995) than to the other well known families, bacterial and mammalian, of extracellular ribonucleases (Hill et al., 1983). However, the larger size, the high specificity, and the ability to interact with biological membranes are unique to this group of ribonucleases.

Despite extensive studies on the ribonucleolytic activity (Endo et al., 1983, 1990), the protein-lipid interaction, and the cytotoxic activity of αS (Gasset et al., 1994 and references therein), little is known about the structure of this protein. In this respect, only a spectroscopic characterization of the protein conformation (Martínez del Pozo et al., 1988), and a sequence homology-based predictive study have been reported (Mancheño et al., 1995). However, interest in the elucidation of the three-dimensional (3D) structure of any member of this protein family has been put forward in the literature, and preliminary crystallographic reports (Martínez & Smith, 1991) and attempts (Wool, 1984) on this line were reported initially, without further confirmation or success.

Knowledge of the 3D structure of αS is a first step to achieve our final goal of interpreting the cytotoxic action of αS at the molecular level. We are studying this molecule by NMR spectroscopy with the aim of obtaining sufficient data for a high-resolution structure determination. As a prerequisite, however, the assignment of all resonances and their correlations is necessary to allow the identification of structural constraints to be used in the calculation of the 3D structure. In this report, we present the assignment of the proton and nitrogen spectra of all

Reprint requests to: Manuel Rico, Instituto de Estructura de la Materia, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain; e-mail: emcampos@roca.csic.es.

non-proline residues of αS, and outline the elements of secondary structure as deduced from the analysis of the NMR data.

Purification and overexpression of αS in bacteria was achieved as described (Lacadena et al., 1994). ^{15}N uniformly labeled protein was produced by growing the bacteria in M9 minimal medium containing $^{15}NH_4Cl$ as the only source of nitrogen. Appropriate *Escherichia coli* strains were used to produce recombinant αS labeled selectively with ^{15}N-Tyr or ^{15}N-Phe, as described (McIntosh & Dahlquist, 1990; Yamakasi et al., 1992). The isolated recombinant isotopically labeled αSs, as well as that purified from the extracellular media of the mold, were homogeneous and retained their native conformation according to their HPLC and PAGE-SDS behavior, amino acid composition, spectroscopic features, and ribonucleolytic activity. This characterization of the purified proteins was performed routinely with all the samples used in this study. Most NMR spectra were obtained at 600 MHz, and some others at 750 MHz, with ~2 mM concentration protein samples, 50 mM NaCl, in 90/10 (v/v) H_2O/D_2O mixtures. Two-dimensional (2D) spectra at a variety of experimental conditions of pH (4.0 and 6.0) and temperature (25, 30, 33, 35, and 38 °C) were recorded and analyzed to resolve for spectral overlap and to check the hypothesis formulated along the assignment process in different experimental data sets. A sample in D_2O (pH 6.0, 35 °C) was used for obtaining information about the chemical exchange of backbone amide protons.

The assignment process was conducted by using the classical sequence-specific strategy (Wüthrich, 1986) and was based on COSY (Aue et al., 1976), TOCSY (Bax & Davis, 1985), and NOESY (Kumar et al., 1980) homonuclear 2D spectra, HSQC (Bodenhausen & Ruben, 1980) 2D heteronuclear spectra, and NOESY-HSQC 3D spectra (Palmer et al., 1992). HSQC and ^{15}N-filtered NOESY (Otting et al., 1986) were recorded for the selectively labeled samples. On the basis of the sequence of αS (150 residues, 13 Pro), 136 independent spin systems are expected in the H_N to H_α fingerprint region of 2D correlation spectra. In a first stage of the assignment process, about one half of the intraresidual H_N to H_α correlations were classified into different spin systems (and sometimes particular amino acids) according to the chemical shifts of their associated side-chain protons, as obtained from COSY and TOCSY spectra, and backbone nitrogen resonances present in the HSQC. Extensive overlap and poor TOCSY transfer (even when a variety of mixing times of 30, 60, 70, and 90 ms were used) precluded the classification and even detection of the rest of the expected spin systems, which were only identified after their sequence-specific assignment. The severe spin diffusion present in the 3D NOESY-HSQC spectrum at 150 ms mixing time facilitated the establishment of pairs of consecutive systems by their symmetric and very often detectable H_N to H_N NOE. These were complemented by the observation of H_α to H_N and side-chain protons to H_N contacts in the set of NOE cross peaks to the H_N resonances at each corresponding ^{15}N plane in the 3D spectrum, thus defining the direction of the pair in the protein sequence. Unambiguous pairs were then assembled in stretches of residues. The location of the 13 Pro residues in the protein sequence determines 12 polypeptide fragments within which H_α to H_N and H_N to H_N sequential connectivities could in principle be established. The sequential assignment of a whole stretch of spin systems was then achieved when the nature and order of the spin systems could only match a unique sequence in one of the 12 non-prolyl fragments. The identification in the HSQC spectrum of the N_H to H_N cross peaks corresponding to the eight Tyr and the six Phe residues in the spectra of the two selectively labeled samples were of great help on applying the outlined procedure. All assignments made on the basis of dipolar correlations in NOESY spectra (mixing times of 50, 100, 120, and 150 ms) were checked subsequently and in most cases confirmed in the scalar connectivities present in COSY and TOCSY experiments. The complete HSQC spectrum is shown in Figure 1. All peaks have been assigned as indicated in the labels attached to them. It was also possible to assign at least some of the side-chain proton resonances of most residues and all side-chain nitrogen resonances of Trp, Asn, Gln, and Arg. Despite the severe overlap and broadening observed for several aromatic proton resonances, the connection of them to the corresponding N_H-H_N-H_α-$H_{\beta\beta'}$ spin system was unambiguously performed in most cases. For 11 Pro residues of the 13 present, it was possible to identify the conformation of the corresponding X-Pro peptide bond based on the characteristic sequential NOEs between protons. These were Pro 13 (*trans*), Pro 38 (*trans*), Pro 49 (*cis*), Pro 63 (*trans*), Pro 68 (*trans*), Pro 79 (trans), Pro 98 (*trans*), Pro 101 (*trans*), Pro 113 (cis), Pro 117 (*trans*), and Pro 127 (*trans*). Assignment of some Pro resonances is yet to be accomplished.

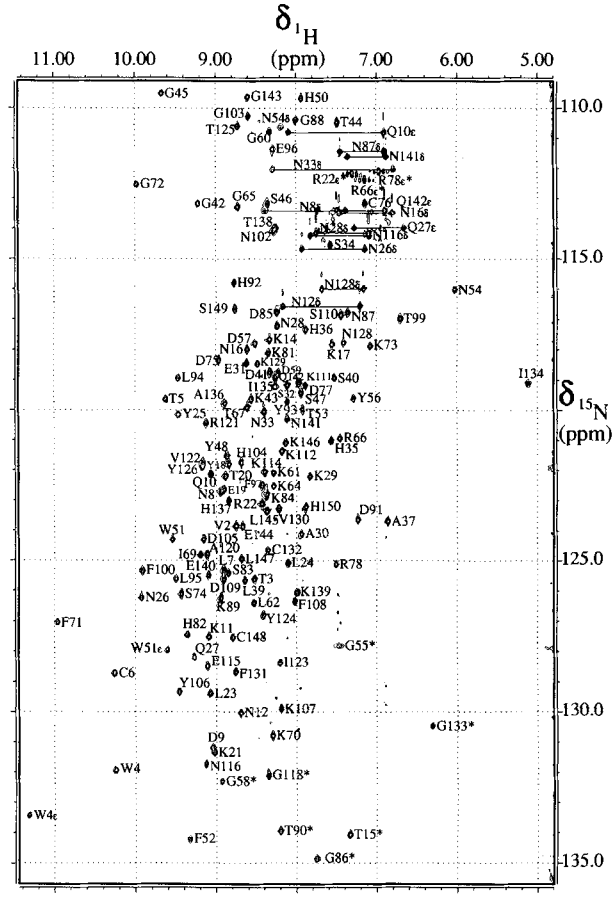

Fig. 1. 750 MHz ^1H-^{15}N HSQC spectrum of α-Sarcin (pH 4.0, 33 °C). All correlations have been assigned and labeled according to residue type and sequence number, and the atom type is also indicated for side-chain correlations (those from Asn and Gln joined by horizontal lines). Folded peaks (in the ^{15}N dimension) are indicated with the asterisk.

An initial characterization of the structure can be obtained considering a few well-resolved and unambiguously assigned NOEs that correspond to the characteristic patterns of secondary structure elements (Wüthrich, 1986). In particular, the relative intensities of sequential NOEs constitute an important source for the characterization of secondary structure elements and are shown schematically in Figure 2. These observations are further supported by other independent NMR parameters: conformational chemical shifts (Jiménez et al., 1987; Wishart et al., 1991) and amide proton exchange properties (Englander & Kallenback, 1984), which are represented against the protein sequence in Figure 3 and Figure 2, respectively. On this basis, three elements of regular secondary structure have been characterized in αS: a long N-terminal β-hairpin, a short α-helical segment, and a C-terminal antiparallel β-sheet. The first two β-strands defining the hairpin comprise the residues 3–12 (β1) and 17–26 (β2), the turn of the hairpin residing in residues 13–16. Residues 29–34 form a short α-helix with scarcely two turns. This contrasts with the other ribonucleases displaying significant degree of similarity with αS, where a long helical fragment is found and appears to stabilize the molecular structure by forming a hydrophobic core with the β-sheet (Hill et al., 1983; Sevcik et al., 1990). The C-terminal β-sheet of αS is composed of five short strands, β_3 (49–52), β_4 (93–98), β_5 (120–126), β_6 (132–138), and β_7 (144–148), arranged in a +1,+1,+1,+1 topology, as shown in Figure 4. It is remarkable that predictive methods suggested an identical arrangement of β-strands in αS (Mancheño et al., 1995). Taking into account the characterized structure, only one third of the residues are found in repetitive secondary structure elements and the rest constitute the numerous and long loops of the protein structure, where the 13 Pro are located.

Knowledge of the determined secondary structure of αS already sheds some light into the molecular basis of the two main activities of the protein. It has been suggested on the basis of sequence similarity (Sacco et al., 1983; Mancheño et al., 1995) and mutagenesis (Lacadena et al., 1995) studies that His 50, Glu 96, Arg 121, and His 137 would correspond to the active site residues found in the enzymes of the ribonuclease T_1 family (His 40, Glu 61, Arg 121, and His 137, respectively; T_1 numbering). The arrangement of these residues in the C-terminal β-sheet of αS is such that they come close in the folded protein. Therefore, the characterized structure gives additional support to the concept of those residues forming part of the αS recognition/catalytic site. With respect to the lipid interaction capabilities of αS, this study confirms the existence of a hydrophobic

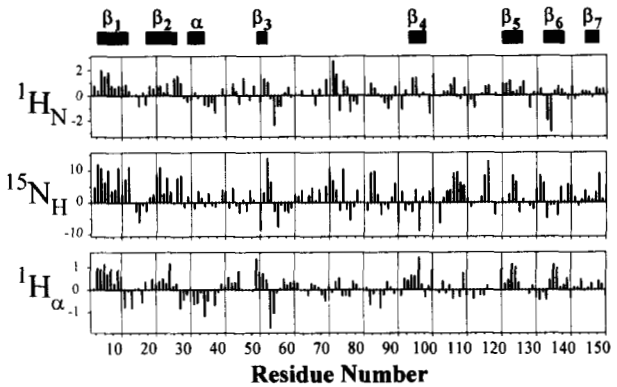

Fig. 3. Analysis of the secondary structure elements of α-Sarcin according to the conformational shifts of α protons ($^1H_\alpha$), backbone amide nitrogens ($^{15}N_H$), and backbone amide protons (^1H_N). Chemical shifts are those obtained from a single 2D NOESY spectrum recorded at 750 MHz with a mixing time of 150 ms on a 2 mM sample of αS at 33 °C and pH 4.0. Conformational shifts are calculated as the difference between the observed value and that expected for the random coil (Wishart et al., 1995). Secondary structure elements identified are indicated on top.

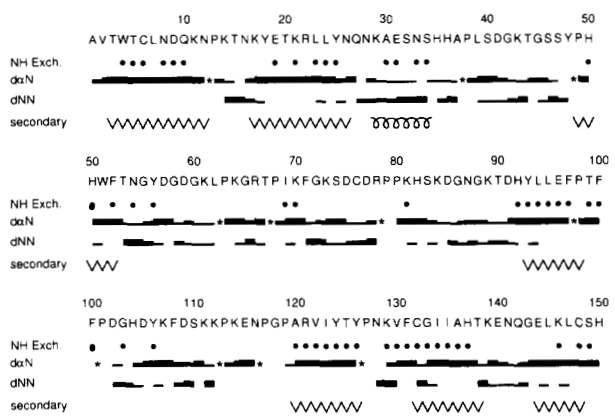

Fig. 2. Summary of the sequential H_N to H_N and H_α to H_N NOE connectivities and the protection of backbone amide protons against exchange found for αS. NOE data were obtained from a single 2D NOESY spectrum recorded at 600 MHz with a mixing time of 120 ms on a 2 mM sample of αS at 38 °C and pH 4.0. The intensity of the corresponding NOE cross peak is encoded in the thickness of the line. The X-Pro connectivities assigned are indicated with the asterisk and are those expected for *cis* (H_α to H_α) or *trans* (H_α to $H_{\delta\delta'}$) conformation, according to the conformation of each X-Pro (see text). In the exchange plot, a dot indicates that the amide proton of that residue remains in D_2O solution at least for 2 h. Secondary structure elements are identified at the bottom.

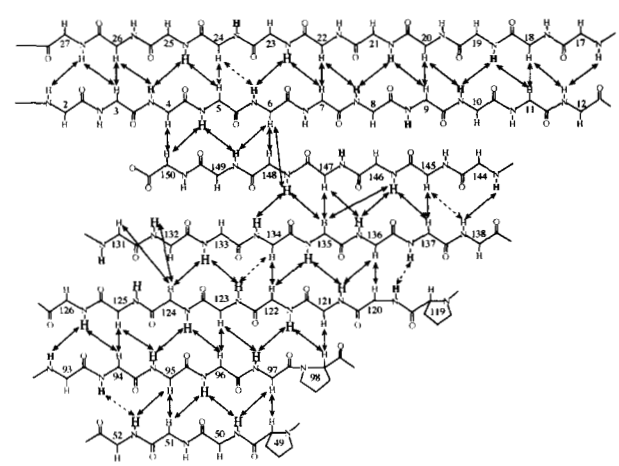

Fig. 4. Schematic diagram showing the NOE connectivities between backbone protons that allow definition of the extension of the β-strands and their connection topology, as present in αS. An arrow connecting two protons is drawn when an NOE between them is observed in the NOESY spectrum at 120-ms mixing time. Dashed arrows represent NOE connectivities that cannot be assigned unambiguously or observed due to overlap with sequential signals or closeness to the diagonal, respectively. Slow exchanging amide protons are represented in bold, and those remaining after 140 h in D_2O are indicated with the big H.

domain defined by a nucleus of β-sheet that would allow the hydrophobic interaction with bilayers (Mancheño et al., 1995). In addition, the presence of many potentially flexible loops could be envisioned as essential parts of a conformational switch mechanism, which on contacting the bilayer, would change their relative position and allow the β-sheet hydrophobic core to be displayed to interact with the acyl chains of phospholipids.

In summary, the almost complete assignment of the proton and nitrogen spectra of αS has been achieved, thus paving the way for the determination of the 3D structure of αS in solution. Structure calculation is in progress. Detailed knowledge of this structure will provide important advances in our understanding of this relevant group of enzymes.

Acknowledgments: This work was supported by the Dirección General de Investigación Científica y Técnica (Spain) (PB93/0090) and the Comunidad de Madrid (AE00328/95). R.C.O. was recipient of a predoctoral grant (FP92-02882770) from the Ministerio de Educación y Ciencia (Spain). We are grateful to Dr. B.J. Bachman of Yale University for kindly providing the auxotrophic *E. coli* strains. *A. giganteus* MDH 18894 was a generous gift from Dr. Y. Endo from Ehime University. Dr. R. Boelens, Dr. K. Hård, and Dr. K. Tamminga are acknowledged for their help with NMR spectroscopy at the SON-NMR Facility at the University of Utrecht, and the EC for financial support to that visit. We thank Mr. A. Gómez, Mrs. C. López, and Mr. L. de la Vega for excellent technical assistance.

References

Aue WP, Bartholdi E, Ernst RR. 1976. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J Chem Phys 64*:2229–2246.

Bax A, Davis DG. 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson 65*:355–360.

Bodenhausen G, Ruben DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett 69*:185–189.

Endo Y, Glück A, Chan YL, Tsurugi K, Wool IG. 1990. RNA-protein interaction. An analysis with RNA oligonucleotides of the recognition by α-Sarcin of a ribosomal domain critical for function. *J Biol Chem 265*: 2216–2222.

Endo Y, Huber PW, Wool IG. 1983. The ribonuclease activity of the cytotoxin α-Sarcin. The characteristics of the enzymatic activity of α-Sarcin with ribosomes and ribonucleic acids as substrates. *J Biol Chem 258*: 2662–2667.

Englander SW, Kallenback NR. 1984. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q Rev Biophys 16*:521–655.

Gasset M, Mancheño JM, Lacadena J, Turnay J, Olmo N, Lizarbe MA, Martínez del Pozo A, Oñaderra M, Gavilanes JG. 1994. α-Sarcin, a ribosome-inactivating protein that translocates across the membrane of phospholipid vesicles. *Curr Topics Pept Protein Res 1*:99–104.

Hill C, Dodson G, Heinemann U, Saenger W, Mitsui Y, Nakamura K, Borisov S, Tischenko G, Polyakov K, Pavlosky S. 1983. The structural and sequence homology of a family of microbial ribonucleases. *Trends Biochem Sci 8*:364–369.

Jiménez MA, Nieto JL, Herranz J, Rico M, Santoro J. 1987. ^{1}H NMR and CD evidence of the folding of the isolated ribonuclease 50–61 fragment. *FEBS Lett 221*:320–324.

Kumar A, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross relaxation networks in biological macromolecules. *Biochim Biophys Res Commun 95*:1–6.

Lacadena J, Mancheño JM, Martínez-Ruiz A, Martínez del Pozo A, Gasset M, Oñaderra M, Gavilanes JG. 1995. Substitution of histidine-137 by glutamine abolishes the catalytic activity of the ribosome-inactivating protein α-sarcin. *Biochem J 309*:581–586.

Lacadena J, Martínez del Pozo A, Barbero JL, Mancheño JM, Gasset M, Oñaderra M, López-Otín C, Ortega S, García J, Gavilanes JG. 1994. Overproduction and purification of biologically active native fungal α-Sarcin in *Escherichia coli*. *Gene 142*:147–151.

Mancheño JM, Gasset M, Lacadena J, Martínez del Pozo A, Oñaderra M, Gavilanes JG. 1995. Predictive study of the conformation of the cytotoxic protein α-Sarcin: A structural model to explain α-Sarcin-membrane interaction. *J Theor Biol 172*:259–267.

Martínez del Pozo A, Gasset M, Oñaderra M, Gavilanes JG. 1988. Conformational study of the antitumor protein α-Sarcin. *Biochim Biophys Acta 953*:280–288.

Martínez SE, Smith JL. 1991. Crystallization and preliminary characterization of Mitogillin, a ribosomal ribonuclease from *Aspergillus restrictus*. *J Mol Biol 218*:489–492.

McIntosh LP, Dahlquist FW. 1990. Biosynthetic incorporation of ^{15}N and ^{13}C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q Rev Biophys 23*:1–38.

Otting G, Senn H, Wagner G, Wüthrich K. 1986. Editing of 2D NMR spectra using X half-filters. Combined use with residue-selective ^{15}N labelling of proteins. *J Magn Reson 70*:500–505.

Palmer AG, Cavanagh J, Byrd RA, Rance M. 1992. Sensitivity improvement in three-dimensional heteronuclear correlation NMR spectroscopy. *J Magn Reson 96*:416–424.

Sacco G, Drickamer K, Wool IG. 1983. The primary structure of the cytotoxin α-Sarcin. *J Biol Chem 258*:5811–5818.

Sevcik J, Sanishvili RG, Pavlosky AG, Polyakov KM. 1990. Comparison of active sites of some microbial ribonucleases: Structural basis for guanylic specificity. *Trends Biochem Sci 15*:158–162.

Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. 1995. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigation of nearest-neighbour effects. *J Biomol NMR 5*:67–81.

Wishart DS, Sykes BD, Richards FM. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol 222*:311–333.

Wool IG. 1984. The mechanism of action of the cytotoxic nuclease α-Sarcin and its use to analyse ribosome structure. *Trends Biochem Sci 9*:14–17.

Wool IG, Glück A, Endo Y. 1992. Ribotoxin recognition of ribosomal RNA and a proposal for the mechanism of translocation. *Trends Biochem Sci 17*:266–269.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: J. Wiley & Sons.

Yamakasi K, Muto Y, Ito Y, Wälchli M, Miyazawa T, Nishimura S, Yokoyama S. 1992. A ^{1}H-^{15}N NMR study of human c-Ha-*ras* protein: Biosynthetic incorporation of ^{15}N-labelled amino acids. *J Biomol NMR 2*: 71–82.