

Resonance assignments for Oncostatin M, a 24-kDa α -helical protein

Ross C. Hoffman^a, Franklin J. Moy^b, Virginia Price^a, Jane Richardson^a, Dennis Kaubisch^a,
Eric A. Frieden^a, Jonathan D. Krakover^a, Beverly J. Castner^a, Julie King^a, Carl J. March^a
and Robert Powers^{b,*}

^a*Immunex Corp., 51 University Street, Seattle, WA 98101, U.S.A.*

^b*Department of Structural Biology, Wyeth-Ayerst Research, Pearl River, NY 10965, U.S.A.*

Received 28 November 1995

Accepted 29 February 1996

Keywords: Triple-resonance spectroscopy; Multidimensional NMR; Oncostatin M; Resonance assignments

Summary

Oncostatin M (OM) is a cytokine that shares a structural and functional relationship with interleukin-6, leukemia inhibitory factor, and granulocyte-colony stimulating factor, which regulate the proliferation and differentiation of a variety of cell types. A mutant version of human OM in which two N-linked glycosylation sites and an unpaired cysteine have been mutated to alanine (N76A/C81A/N193A) has been expressed and shown to be active. The triple mutant has been doubly isotope-labeled with ¹³C and ¹⁵N in order to utilize heteronuclear multidimensional NMR techniques for structure determination. Approximately 90% of the backbone resonances were assigned from a combination of triple-resonance data (HNCA, HNCOC, CBCACONH, HBHACONH, HNHA and HCACO), intraregion and sequential NOEs (3D ¹⁵N-NOESY-HMQC and ¹³C-HSQC-NOESY) and side-chain information obtained from the CCONH and HCCONH experiments. Preliminary analysis of the NOE pattern in the ¹⁵N-NOESY-HMQC spectrum and the ¹³C α secondary chemical shifts predicts a secondary structure for OM consisting of four α -helices with three intervening helical regions, consistent with the four-helix-bundle motif found for this cytokine family. As a 203-residue protein with a molecular weight of 24 kDa, Oncostatin M is the largest α -helical protein yet assigned.

Introduction

Oncostatin M (203 residues, 24 kDa) is a member of the cytokine family that includes interleukin-6 (IL-6), leukemia inhibitory factor (LIF), and granulocyte-colony stimulating factor (G-CSF), which are lymphoid cell-derived cytokines that share a number of functional and structural properties (Rose et al., 1991; Bruce et al., 1992a). The close relationship of these cytokines is further demonstrated by the regulation of G-CSF and IL-6 by expression of OM (Brown et al., 1991, 1993). Functionally, they induce differentiation of murine myeloid leukemia cells to macrophages (Bruce et al., 1992b; Tanigawa et al., 1995), act as growth factors for human plasmacytoma cells (Barton et al., 1994; Nishimoto et al., 1994), and induce acute-phase protein synthesis in hepatocytes (Rich-

ards et al., 1992). It has been shown that these shared functions are mediated by the signal-transducing membrane glycoprotein, gp130, that interacts with the extracellular domains of these cytokine receptors (Boulton et al., 1994; Liu et al., 1994; Sporeno et al., 1994). In addition to its familial attributes, OM inhibits the proliferation of several tumor cell lines (Zarling et al., 1986; Horn et al., 1990), inhibits the differentiation of embryonic stem cells (Rose et al., 1994), is a regulator of endothelial cells (Brown et al., 1990) and stimulates the growth of several fibroblast cell lines (Zarling et al., 1986; Horn et al., 1990). OM has also been shown to be a mitogen for AIDS-related Kaposi's sarcoma cells (Miles et al., 1992; Nair et al., 1992), where it functions as an autocrine growth factor (Cai et al., 1994). Structurally the members of this family are believed to be related to growth hor-

*To whom correspondence should be addressed.

Supplementary material available from the authors: tables containing the ¹H, ¹⁵N, ¹³CO, and ¹³C resonance assignments for Oncostatin M.

none, a four-helix bundle with characteristic up-up-down-down helix topology (Zarling et al., 1986; Rose et al., 1991). This has been substantiated by the structural similarity of the LIF X-ray structure with the structure of growth hormone (De Vos et al., 1992; Robinson et al., 1994).

Here, we report the nearly complete backbone resonance assignments of the spectrum of OM from a series of 3D triple-resonance experiments and a preliminary secondary structure analysis from the NOE pattern of the ^{15}N -NOESY-HMQC spectra and the $^{13}\text{C}^\alpha$ secondary chemical shifts. As a 203-residue protein with a molecular weight of 24 kDa, Oncostatin M is the largest α -helical protein yet assigned.

Materials and Methods

Expression

Human recombinant Oncostatin M was expressed in a yeast α -factor vector expression system. The version of OM used included the addition of the octapeptide termed FLAG (DYKDDDDK) to the N-terminus for the purpose of affinity purification (Prickett et al., 1989). In addition, three point mutations were introduced using site-directed mutagenesis: C88A to eliminate an unpaired cysteine and prevent potential oxidation, and N83A/N200A to eliminate the two potential N-glycosylation sites. Note that residues are numbered from the N-terminal Asp as residue 1, i.e. the human sequence starts at Ala⁹ and ends at Arg²⁰³.

The activity of the triple mutant FLAG-OM was measured using a competition assay in which the analyte competes with radiolabeled FLAG-OM for binding to the LIF receptor on a transfected B9/hLIFR cell line (Gearing et al., 1992). The mutant binding activity was identical to that of wild-type FLAG-OM (see discussion below).

The apparent molecular weight of the triple mutant FLAG-OM was determined by injecting 150 μl aliquots of the NMR sample into a DynaPro-801 (Protein Solutions, Charlottesville, VA) dynamic light-scattering instrument.

^{13}C -/ ^{15}N -labeled OM was expressed in yeast *Saccharomyces cerevisiae* XV2181:pIXY755 using a fed-batch fermentation process in which ^{13}C / ^{15}N Celtone medium (Martek Corp., Columbia, MD) was used instead of hydrolyzed casamino acids, essentially as described by Powers and co-workers (Powers et al., 1992a).

Purification

The sterile filtered (0.22 μm) yeast supernatant was buffered by the addition of 10 \times PBS to a final concentration of 1 \times PBS and was supplemented with 1.0 M CaCl_2 to a final Ca^{2+} concentration of 1.0 mM. 100 ml of this treated supernatant was applied to an affinity column equilibrated in 1 \times PBS (no Ca^{2+} /no Mg^{2+}). The affinity

column was prepared by coupling a monoclonal antibody (Ab) raised against the octapeptide FLAG (M1) to 17.5 ml of affigel at 4 mg Ab/ml affigel. The column was then washed with 50 ml of 1 \times PBS + 1.0 mM CaCl_2 . The labeled OM was eluted with 50 mM Na citrate, pH 3.0. The protein was thus purified to homogeneity, as determined by SDS-PAGE, N-terminal sequencing, and amino acid analysis. The yield of pure ^{13}C -/ ^{15}N -labeled OM from the yeast supernatant was 45.6 $\mu\text{g/ml}$.

Glycosylation state

It was evident from the apparent molecular weight on the SDS-PAGE gel that despite the mutation of the two N-glycosylation sites, the mutant FLAG-OM was modified by the addition of approximately 1–2 kDa of additional mass. This was confirmed by analysis of the purified protein on a Finnegan triple quadrupole mass spectrometer, which showed a heterogeneous pattern indicating the addition of up to nine hexose units.

Carbohydrate mapping of the O-glycosylation site(s) was carried out as follows. Unlabeled OM (4.0 mg) was reduced in 6 M guanidine-HCl containing 250 mM Tris and 25 mM EDTA, pH 7.0, at a protein concentration of 1 mg/ml by the addition of 100 μl of 1.0 M DTT and incubation at 37 $^\circ\text{C}$ for 1 h under N_2 . The reduced protein was alkylated by the addition of 100 μL of 2.0 M iodoacetic acid, followed by incubation at room temperature for 90 min in the dark. The reaction was quenched by the addition of 100 μl 1.0 M DTT. Amino acid analysis and mass spectrometry revealed that all four cysteine residues had been alkylated.

Two separate digests of the reduced/alkylated OM were carried out using endoproteinase Glu-C and trypsin from Boehringer Mannheim (sequencing grade; Indianapolis, IN). Glu-C conditions were 1 mg/ml protein in 25 mM NH_4HCO_3 , pH 7.8, with 0.01% SDS and 1:20 enzyme:protein (w/w) at room temperature for 24 h. Trypsin conditions were 1 mg/ml protein in 100 mM Tris, pH 8.5, with 1:50 enzyme:protein (w/w) at 37 $^\circ\text{C}$ for 48 h.

50 μg of each digest was run individually on a Vydac C_{18} column using a 0.1% B/min gradient (A: H_2O + 0.1% TFA; B: acetonitrile + 0.1% TFA). The effluent was split with a T-splitter to give a flow rate of 10–20 $\mu\text{l/min}$ into a Finnegan triple quadrupole mass spectrometer for on-line mass spectral analysis of the peptides (LC-MS). The Glu-C digest showed that peptide Met¹³⁶-Asp¹⁶⁶ was a heterogeneous glycopeptide with up to nine attached hexose units (note that Glu-C was cleaved beyond the S-acetyl-Cys¹³⁵). The trypsin digest confirmed this finding, as the tryptic peptide Gly¹⁵⁴-Arg¹⁷⁰ was shown to have up to nine attached hexose units. Overlap of the two digests indicated that the O-glycosylation site(s) was between Gly¹⁵⁴ and Asp¹⁶⁶. MSMS of the tryptic peptide indicated that the hexose residues were attached to Thr¹⁶⁰, Thr¹⁶², and/or Ser¹⁶⁵.

NMR spectroscopy

The sample used in these studies contained 1.3 mM ^{13}C -/ ^{15}N -labeled mutant FLAG-OM in a buffered solution of 50 mM KH_2PO_4 , pH 5.2, containing 2 mM NaN_3 . All NMR spectra were recorded at 35 °C on a Bruker AMX-600 spectrometer equipped with a triple-resonance probe. The water suppression technique was either presaturation during the recycle delay or gradient water suppression via the WATERGATE sequence and homo-spoil pulses (Bax et al., 1992; Sklenář et al., 1993). Quadrature detection in the indirectly detected dimensions was obtained in all cases using the TPPI-States method (Marion et al., 1989a). Spectra were processed using the NMRPipe software package (Delaglio et al., 1995) and analyzed with PIPP (Garrett et al., 1991).

The CBCA(CO)NH (Grzesiek and Bax, 1992a), HBHA(CO)NH (Grzesiek and Bax, 1993), HNCO (Powers et al., 1991), HCACO (Grzesiek and Bax, 1992b), HNHA (Vuister et al., 1993), HNCA (Kay et al., 1990), C(CO)-NH and HC(CO)NH (Grzesiek et al., 1993a) experiments were collected as described previously. In general, the acquisition dimension was collected with a spectral width of 13.44 ppm, using 1024 real points with the carrier set at 4.75 ppm. Spectral widths in the indirectly detected ^{15}N and ^{13}C dimensions were 27.13 and 12.13 ppm, respectively, with the carrier positions at 114.5 and 177.7 ppm, respectively. The numbers of points acquired were 32 complex points for ^{15}N and 64 complex points for ^{13}C . For the CBCACONH and C(CO)NH experiments the spectral width for the indirectly detected ^{13}C dimension was 55.97 ppm, using 52 complex points with the carrier set at 46.5 ppm. For the HBHACONH and HC(CO)NH experiments the spectral width for the indirectly detected ^1H dimension was 7.08 ppm, using 62 complex points with the carrier set at 4.75 ppm. For the HNCA and HCACO experiments the spectral widths for the indirectly detected ^{13}C dimensions were 31.86 and 33.13 ppm, respectively, with the carrier set at 56.0 and 56.44 ppm, respectively. The numbers of points acquired were 50 and 32 data points, respectively. For the HNHA experiments the spectral width for the indirectly detected ^1H dimension was 9.50 ppm, using 48 complex points with the carrier set at 4.75 ppm.

The ^{15}N -edited NOESY-HMQC, ^{13}C -edited HSQC-NOESY and ^{15}N -edited HOHAHA experiments were collected as described previously with 150, 150 and 36.3 ms mixing times, respectively (Marion et al., 1989b,c; Clore et al., 1991). Since a ^{13}C -/ ^{15}N -labeled NMR sample was used for all experiments, ^{15}N -WALTZ decoupling was done during acquisition for the ^{13}C -edited HSQC-NOESY and ^{13}C decoupling was achieved in the ^{15}N -edited NOESY-HMQC and ^{15}N -edited HOHAHA experiments by applying a 180° pulse during indirect ^1H evolution. Spectral widths in the indirectly detected ^{15}N and ^{13}C dimensions were 27.13 and 26.51 ppm, respectively,

with the carrier positions at 114.5 and 46.0 ppm, respectively. For the direct and indirect ^1H dimensions the spectral width was 13.44 ppm with the carrier at the water frequency of 4.75 ppm. The numbers of points acquired in the various dimensions were 32 complex points in F1 (^{15}N or ^{13}C), 128 complex points in F2 (^1H) and 1024 real points in F3 (^1H).

Results and Discussion

The continuous advancements in NMR technology have allowed for the assignments of ever larger proteins. The recent assignments of serine protease PB92 (27 kDa) and subtilisin 309 (28 kDa), both containing 269 residues, are prime examples (Fogh et al., 1994,1995; Remerowski et al., 1994). The application of deuterium enrichment (Grzesiek et al., 1993b; Yamazaki et al., 1994a,b) suggests that the upper limit in protein size may reach the 30–40 kDa range. It is also well known that the number of residues and mass of a protein are not the only limiting factors in assigning NMR spectra. Mobility and the nature of the protein's secondary structure play a role in determining whether the spectra of a protein may be assigned. Additionally, the presence of any complicating factors such as low solubility, aggregation and glycosylation tends to significantly reduce the upper size limit of acceptable proteins for NMR analysis. The backbone resonance assignments for Oncostatin M (24 kDa) probably represent the current size limit in the assignment capability of the triple-resonance protocol without the use of deuterium enrichment for a protein with less than ideal properties.

Oncostatin M is predicted to have a predominately α -helical structure similar to human growth hormone, consisting of a four-helix bundle with characteristic up-up-down-down helix topology (Zarling et al., 1986; Rose et al., 1991). This results in the chemical shift dispersion of OM being very narrow relative to β -sheet proteins. For OM, the amide proton resonances are confined to a spectral region of only 2.2 ppm, extending from 7.1 to 9.3 ppm. Additionally, the lowest field C^αH proton resonates at 4.96 ppm, while most of the C^αH protons are centered about 4.33 ppm. In addition to the poor spectral dispersion, OM also has significantly broad line widths with an estimated ^{15}N T_2 value of ~42 ms at 35 °C. Dynamic light-scattering analysis yielded an apparent molecular weight of 46 kDa, which is evidence for aggregation of the protein.

The triple-mutant FLAG-OM protein was shown to be as active as wild-type OM ($K_d = 2.2 \times 10^{-10}$ M; Gearing et al., 1992) by its ability to bind to the LIF receptor with a K_d of 1.7×10^{-10} M, indicating that the presence of the FLAG (DYKDDDDK) peptide and the three point mutations did not drastically affect the integrity of the OM structure. While it is plausible that the presence of the

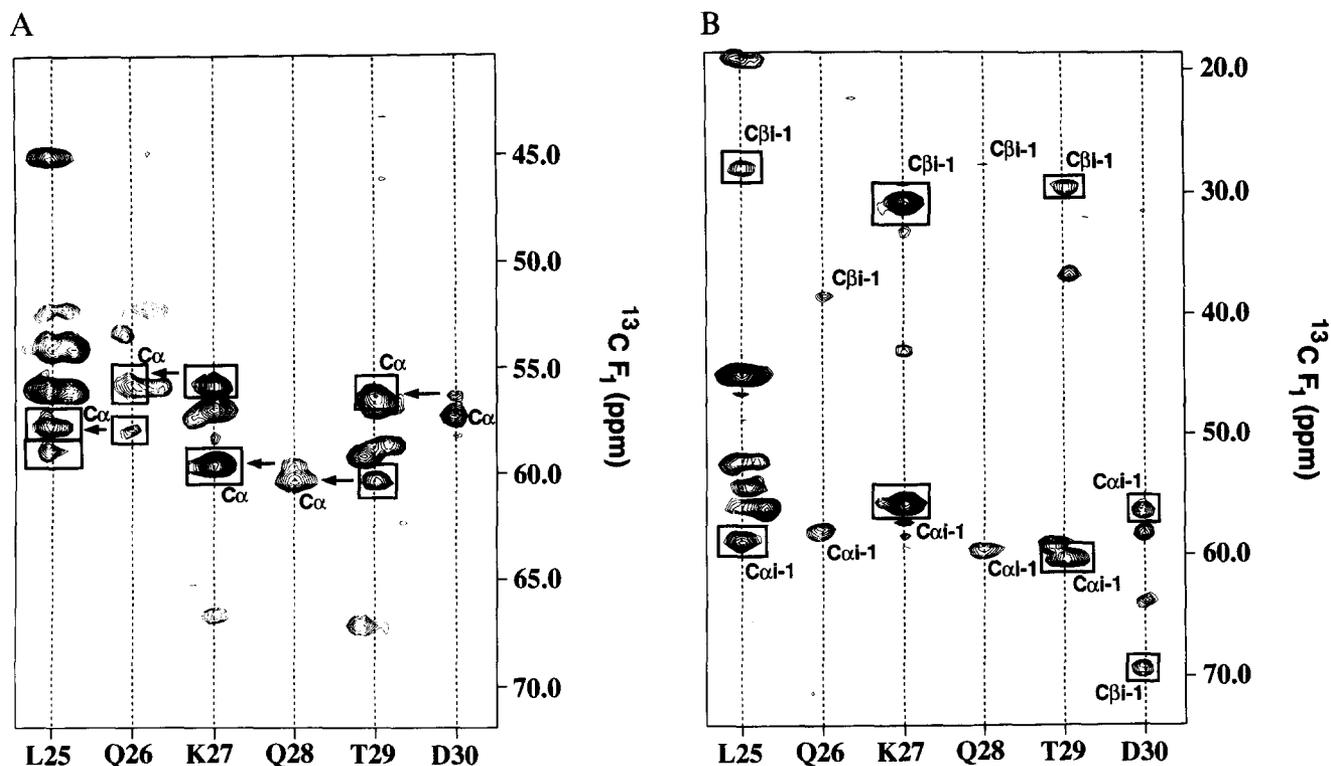


Fig. 1. Composite of amide strips taken from the (A) HNCA; and (B) CBCACONH spectrum of OM for residues Leu²⁵ to Asp³⁰. The C^α(i)-N(i+1),NH(i+1) to C^α(i)-N(i),NH(i) correlations in the HNCA spectra are indicated by an arrow. For strips with multiple spin systems the correct assignments are boxed.

FLAG peptide may have an adverse effect on the quality of the NMR spectra, this effect could not be examined, since the FLAG construct was prepared to facilitate the purification of OM, and an OM sample without the FLAG peptide was not available. Despite the broad linewidth problem, initial 2D ¹⁵N-¹H HSQC NMR spectra collected at 35 °C suggested that it was still possible to assign the OM spectra without the addition of detergents or other factors. Since we were limited to a single ¹³C-/¹⁵N-labeled OM NMR sample, extensive manipulation of the solvent conditions was not practical.

OM was over-expressed in a yeast α -factor vector expression system. To eliminate the possibility of glycosylation, two potential N-glycosylation sites corresponding to residues Asn⁸³ and Asn²⁰⁰ were mutated to alanines. Unfortunately, OM was post-translationally modified with O-linked carbohydrates, as determined by mass spectral analysis (data not shown). Overlapping enzymatic digests and LC-MS confirmed that up to nine hexose units were attached to Thr¹⁶⁰, Thr¹⁶², and/or Ser¹⁶⁵. In addition to increasing the protein mass by ~1–2 kDa, the presence of the O-linked glycosylation site probably results in some local heterogeneity and may possibly affect the local structure.

Over the course of data collection, a new set of peaks were observed in the ¹⁵N-¹H HSQC spectra of OM, which were consistent with chemical shifts of a random coil peptide. A 0.5 mg OM sample treated at pH 2.0 and 70

°C over a 72-h period yielded a degradation product with an apparent molecular weight of 18–19 kDa by gel electrophoresis (data not shown). The OM sequence contains a single Asp-Pro site that can be susceptible to acid autolysis (Landon, 1977) and would yield two fragments with molecular weights of 4.5 and 18.3 kDa. Apparently, these degradation products are formed slowly under the NMR conditions of pH 5.2 and 35 °C. Verification by NMR was not possible, since significant quantities of the 4.5 kDa and 18.3 kDa fragments were not available.

It was readily apparent that the Oncostatin M sample would be a major challenge for NMR and determining a high-resolution NMR solution structure would be a monumental task. It is possible that a number of the observed difficulties with the OM protein may eventually be solved by molecular biology, i.e. by removal of the Asp-Pro cleavage site, the O-linked glycosylation site(s) and the FLAG peptide, but since a new OM construct was not readily available and an X-ray structure does not exist, it seemed reasonable to proceed with the NMR assignments of OM in an effort to learn as much about its structure as possible.

The sequential assignments of OM essentially followed the semiautomated protocol described previously (Powers et al., 1992b; Friedrichs et al., 1994). The primary sequential connectivities used to assign the backbone resonances were the C^α(i) and C^α(i–1) correlations observed in the HNCA and CBCA(CO)NH experiments, respectively, and

the CO(i) and CO(i-1) correlations in the HCACO and HNCO experiments, respectively. Examples of amide strips taken from the HNCA and CBCACONH spectra of OM for residues Leu²⁵ to Asp³⁰ used for the sequential assignments are shown in Fig. 1, and a summary of the sequential connectivities is given in Fig. 2. Despite the difficulties with the OM NMR sample it is quite apparent that a reasonable number of sequential connectivities was observed and that the quality of the NMR data was acceptable for obtaining reliable assignments for approximately 90% of the sequence. The ¹H, ¹⁵N, ¹³C and ¹³CO sequential resonance assignments are provided in Table S1 (supplementary material) and are deposited in the BioMagRes Databank (Seavey et al., 1991).

The major obstacle in completing the OM backbone assignments was the difficulty in observing the H^α(i) res-

onances. Obtaining a significant number of NH(i)-H^α(i) correlations was crucial for the assignments, since it established a third sequential connectivity with the H^α(i-1) resonances in the HBHA(CO)NH experiment and it provided the means to correlate the HCACO data with the HNCA data to make use of the CO(i) and CO(i-1) correlations. Since OM is predominately α-helical and the line widths were significantly broad, only a minimal number of NH(i)-H^α(i) resonances were observed in either the HNHA or ¹⁵N-edited HOHAHA experiments. Therefore it was necessary to collect a ¹³C-edited HSQC-NOESY experiment on an H₂O OM sample to resolve any ambiguities and to increase the number of NH(i)-H^α(i) correlations. Figure 3 shows the expanded NH-H^α region of a selected slice from the ¹³C-edited HSQC-NOESY experiment.

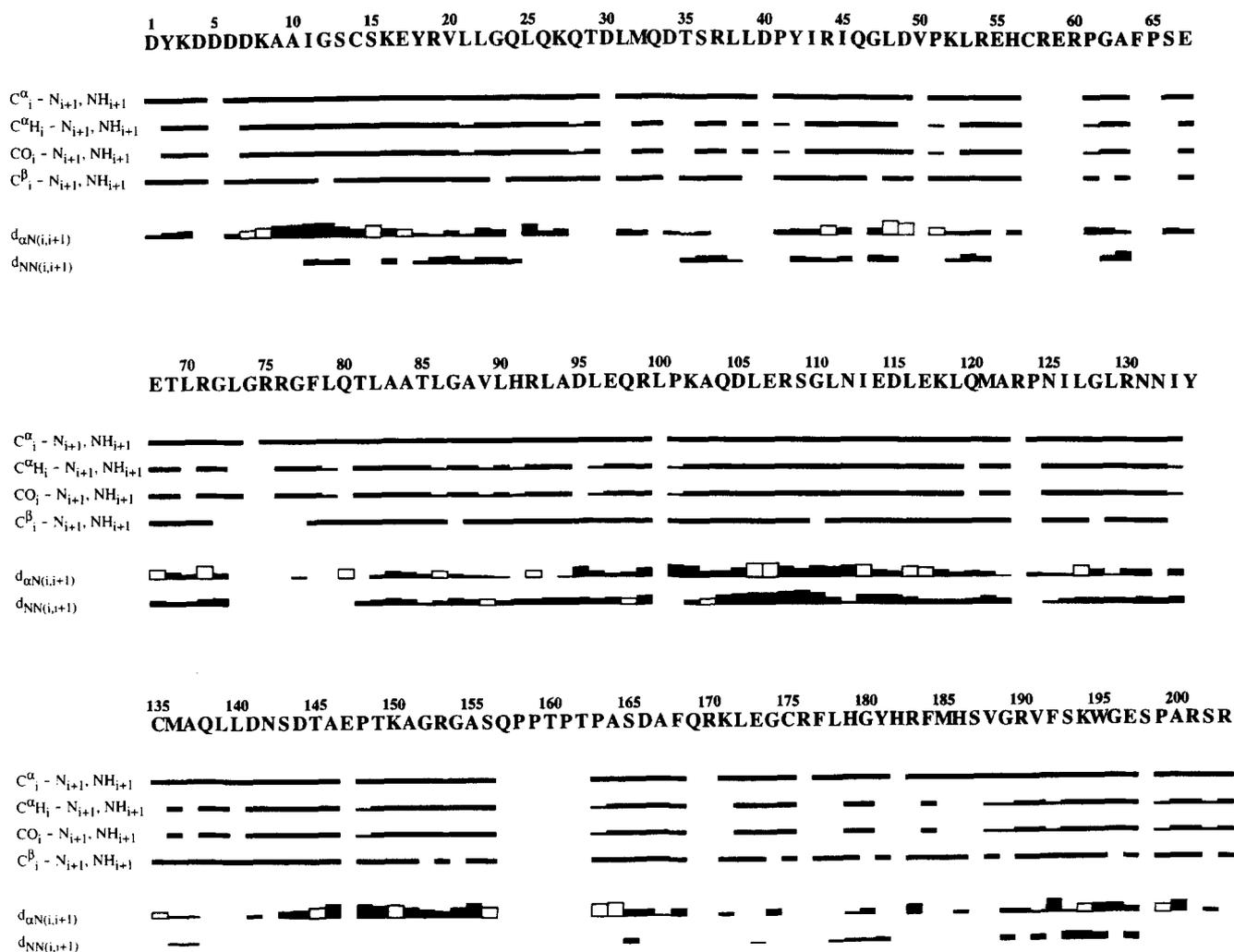


Fig. 2. Summary of the sequential scalar connectivities observed in the 3D triple-resonance experiments of OM. The $C^{\alpha}(i)-N(i+1), NH(i+1)$ correlations in the HNCA and CBCA(CO)NH spectra, the $C^{\alpha}H(i)-N(i+1), NH(i+1)$ correlations in the HBHA(CO)NH spectrum, the $CO(i)-N(i+1), NH(i+1)$ correlations in the HNCO spectrum and the $C^{\beta}(i)-N(i+1), NH(i+1)$ correlations in the CBCA(CO)NH spectra are presented. A half-height bar is shown in cases where a $C^{\alpha}H(i)-N(i), NH(i)$ correlation was not observed, but an HCACO cross peak, consistent with the $C^{\alpha}(i)-N(i), NH(i)$ correlation from the HNCA spectrum, the $CO(i)-N(i+1), NH(i+1)$ correlation in the HNCO spectrum and the $C^{\alpha}H(i)-N(i+1), NH(i+1)$ correlation in the HBHA(CO)NH spectrum, was observed instead. The open boxes represent potential sequential assignment NOEs that are obscured by resonance overlap and could therefore not be assigned unambiguously.

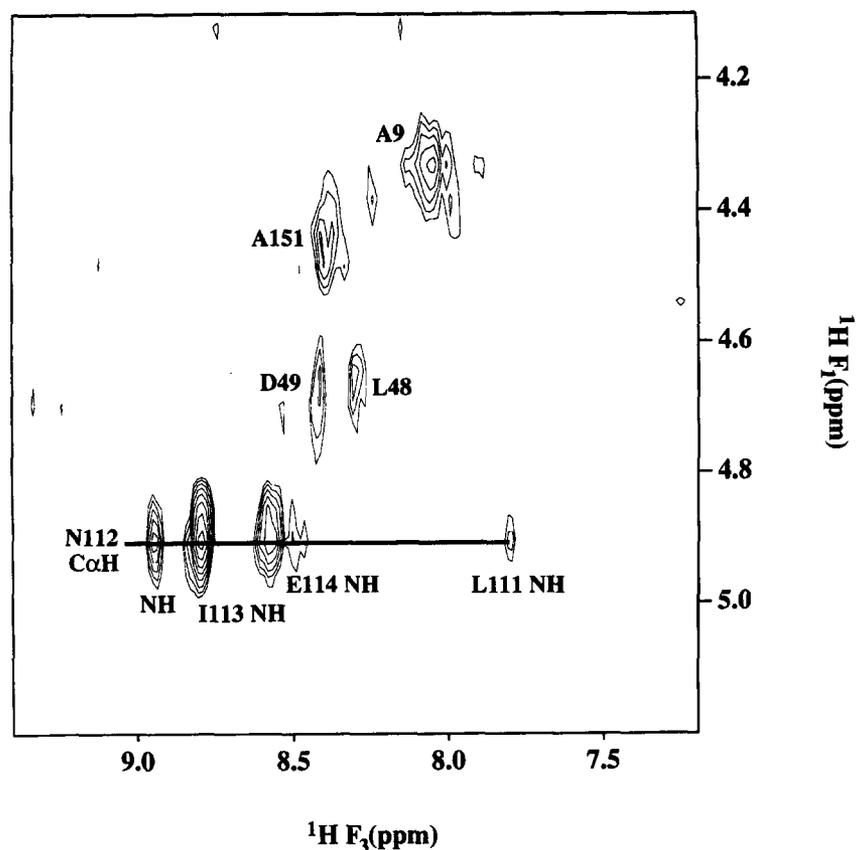


Fig. 3. Selected slice at the ^{13}C chemical shift of 52.83 ppm from the ^{13}C -edited HSQC-NOESY spectrum of Oncostatin M in 90% $\text{H}_2\text{O}/10\%$ D_2O . The expanded NH- C^αH region is shown. Because of the weak $^3J_{\text{HN}^\alpha}$ coupling constants and large line widths, the ^{13}C -edited HSQC-NOESY spectra were crucial for identifying NH(i)- $\text{C}^\alpha\text{H}(i)$ correlations. The sequential assignments are indicated, where the solid line connects the NOEs from the Asn 112 C^αH proton.

Residue typing was based initially on C^α and C^β chemical shifts (Grzesiek and Bax, 1993). Further verification of the residue type was made, when possible, from the analysis of the C(CO)NH and HC(CO)NH experiments. Because of the nature of the OM sample, complete spin systems were not typically observed in the C(CO)NH and HC(CO)NH experiments. Therefore corrections to the assignments were based only on the presence of clearly contradictory data, and not on a lack of information. This was done by simply comparing the number of correlations observed and the corresponding chemical shifts to what was expected for a particular residue. The side-chain aliphatic ^{13}C resonance assignments obtained from the C(CO)NH experiment are provided in Table S1 (supplementary material).

As a final step in verifying the assignments, a composite of amide strips was taken from the 3D ^{15}N -edited NOESY-HMQC experiment and aligned according to the current assignments. Analysis of the NOESY strips provided confirmation of assignments by the presence of sequential NH(i)-NH(i-1) and NH(i)- $\text{H}^\alpha(i-1)$ NOEs. An example of the ^{15}N -edited NOESY-HMQC amide strips for residues Ile 113 to Arg 123 is shown in Fig. 4 and a summary of the sequential NOEs is given in Fig. 2.

A preliminary analysis of the OM secondary structure was based on a qualitative analysis of sequential NOEs and the $^{13}\text{C}^\alpha$ secondary chemical shifts (Clare et al., 1989; Spera et al., 1991). Regions of α -helices were identified by stretches of sequential NH-NH NOEs and by positive $^{13}\text{C}^\alpha$ secondary chemical shifts (>3 ppm). Figure 5 shows the $^{13}\text{C}^\alpha$ secondary chemical shifts plotted versus the residue number for OM. It is clearly evident from this plot and from the sequential NH correlations observed in the ^{15}N -edited NOESY spectra (Figs. 2 and 4) that OM is predominately helical in nature. The structure consists of four major α -helices (residues 18–45, 78–98, 113–140, and 168–194) with three short intervening helical regions (residues 51–57, 66–71, and 103–110), which may be short helices or helical turns. It was not possible to distinguish between a true helix or a helical turn based on the limited information available in the ^{15}N -edited NOESY spectra. The stretches of negative $^{13}\text{C}^\alpha$ secondary chemical shifts for residues 145–157 and 163–167 are consistent with a β -strand conformation (note that Pro 158 -Thr 162 is unassigned and also is the region that is glycosylated).

There are two residues with apparently inconsistent $^{13}\text{C}^\alpha$ secondary chemical shifts in α -helices 1 and 3, corresponding to residues Thr 29 and Pro 124 , respectively. Amide

strips taken from the HNCA spectrum of OM for residues Leu²⁵ to Asp³⁰ in Fig. 1 and the summary of the sequential connectivities in Fig. 2 support the Thr²⁹ C^α assignment. The C^β(i-1) of 69.3 ppm in the CBCA(CO)NH experiment for the spin system assigned to Asp³⁰ is clearly consistent with a threonine side chain with an unusually upfield shifted C^α of 56.4 ppm. There is a total of nine threonines in the sequence of OM and only three do not correspond to an α -helical region or precede a proline: Thr⁶⁹, Thr¹⁴⁵ and Thr¹⁴⁹. All attempts at exchanging the assignment of Thr²⁹ with either Thr⁶⁹, Thr¹⁴⁵ or Thr¹⁴⁹ to eliminate the inconsistent ¹³C^α secondary chemical shifts were not successful. Therefore, the current assignment for Thr²⁹ appears to be the best interpretation of the data. A possible explanation for the unusual ¹³C^α secondary chemical shift might be a distortion in the structure of the α -helix. This is suggested by the lack of sequential NH-NH NOEs and the relatively small ¹³C^α secondary chemical shifts for residues near Thr²⁹.

Unlike Thr²⁹, the only source for the C^α assignment of Pro¹²⁴ is the C^α(i-1) correlation at 59.6 ppm observed in the HNCA for Asn¹²⁵. This residue is nearly overlapped with Arg¹⁷⁶ and Phe¹⁸⁴ and only correlations from these

two residues are observed in the CBCA(CO)NH experiment, but the C(CO)NH experiment does indicate that the (i-1) residue from Asn¹²⁵ is probably a proline because of the presence of a C^δ peak at 50.2 ppm. While there is reason to question the reliability of the Pro¹²⁴ C^α assignment because of the overlap, the lack of corroboration and the inconsistent secondary chemical shift data, it is also difficult to ignore an apparently real peak in the HNCA spectra that might indicate a local perturbation in structure. A schematic representation of the proposed secondary structure of OM along with the known disulfide bonds is shown in Fig. 6 (Kallestad et al., 1991).

A model structure for OM had been proposed based on the X-ray structure of porcine growth hormone (Abdel-Meguid et al., 1987; Bazan, 1990; Rose et al., 1991). This structure predicts four α -helices corresponding to residues 19-40, 86-107, 113-139 and 165-192 (the residue numbering was corrected for the additional eight amino acids in the NMR sequence). The NMR structure is consistent with the model structure, except for the eight-residue shift in helix 2 and the missing short helical regions. Since the model structure was based on porcine growth hormone (PGH) instead of human growth hormone (HGH), this

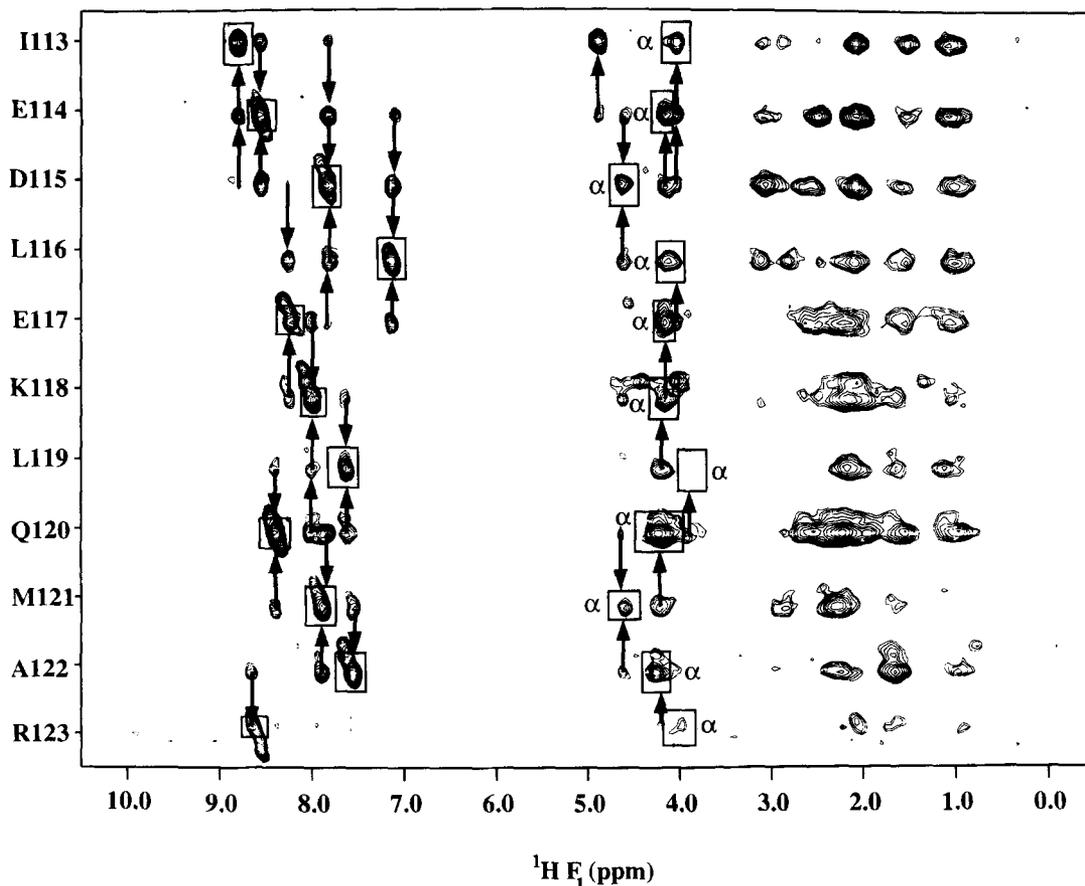


Fig. 4. Composite of amide strips taken from the 150-ms mixing time 3D ¹⁵N-edited NOESY-HMQC spectrum of OM for the stretch Ile¹¹³ to Arg¹²³ comprising part of helix 3. The diagonal peaks and NH(i)-C^αH(i) NOEs are boxed. The NH(i+1)-C^αH(i) and NH(i)-NH(i+1,2) NOEs are indicated by arrows.

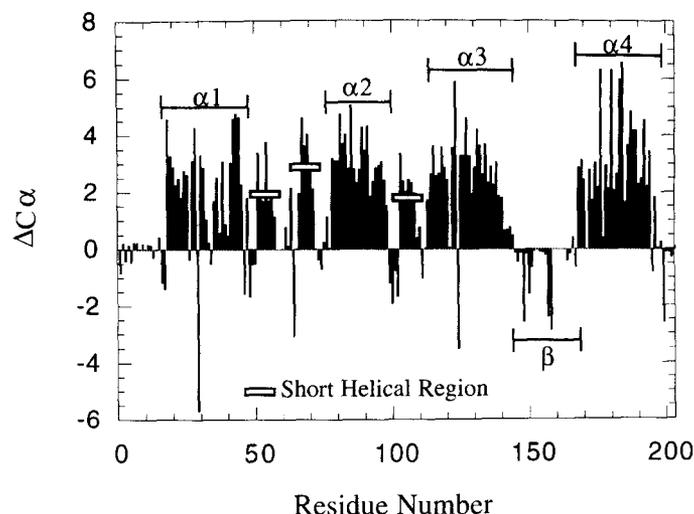


Fig. 5. $^{13}\text{C}^{\alpha}$ secondary chemical shifts, plotted versus the residue number for Oncostatin M, based on the assignments reported in Table S1 (supplementary material). Positive deviations are characteristic of an α -helix and negative deviations are characteristic of a β -sheet. Regions with a proposed secondary structure are indicated on the graph.

difference may not be unexpected. While the X-ray structures of PGH and HGH are very similar, there is a distinct difference between the two structures, particularly around helix 2 (Abdel-Meguid et al., 1987; De Vos et al., 1992). The X-ray structure of HGH contains three additional short helices, which are not seen in the PGH X-ray structure. Two of these short helices directly precede and follow helix 2 in the HGH structure. As a result, helix 2 in HGH is shifted 4–5 residues towards the N-terminus relative to PGH, which allows for the short helix to fit between helices 2 and 3. The third short helix observed in the HGH structure directly follows helix 1. Thus, the

secondary structure determined for OM by NMR (i.e. a four-helix bundle with a short helical region directly after helix 1 and two short helical regions directly before and after helix 2) appears to be more consistent with the HGH X-ray structure instead of the model based on the PGH X-ray structure. This suggests that the current NMR results may be used to refine the existing homology model for OM based on the HGH X-ray structure. Based on the inherent difficulties observed with the OM NMR sample, a combined use of NMR data with homology modeling may prove to be the best approach for obtaining a structure for OM.

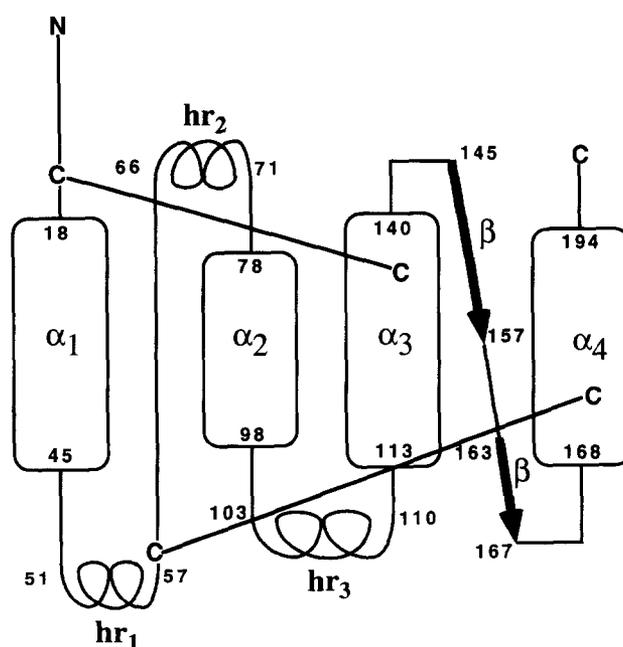


Fig. 6. Schematic representation of the secondary structure predicted for Oncostatin M based on the $^{13}\text{C}^{\alpha}$ secondary chemical shifts and the NH-NH correlations in the ^{15}N -edited NOESY-HMQC spectrum.

Conclusions

We have shown that it is possible to assign a 24 kDa predominately α -helical protein with less than ideal properties using standard 3D triple-resonance experiments. A significant problem in obtaining the assignments for OM was the difficulty in obtaining complete NH(i)-H $^{\alpha}$ (i) assignments, which was solved by using ^{13}C -edited HSQC-NOESY spectra collected on an H $_2$ O sample. The secondary structure of OM determined by stretches of sequential NH-NH NOEs and by positive $^{13}\text{C}^{\alpha}$ secondary chemical shifts consists of four major α -helices (residues 18–45, 78–98, 113–140, and 168–194) with three short intervening helical regions (residues 51–57, 66–71, and 103–110). This is consistent with the X-ray structure of HGH, which provides further validation for the accuracy of the NMR assignments. Despite the inherent difficulties in the OM sample it appears that current NMR protocols are robust enough to provide useful structural data for large (> 20 kDa) and ill-behaved protein samples.

Acknowledgements

We would like to thank Dr. Stephan Grzesiek for helpful discussions and sharing his pulse programs, and Dr. Dan Garrett and Frank Delaglio for the use of their programs for NMR data analysis and processing.

References

- Abdel-Meguid, S.S., Shieh, H.S., Smith, W.W., Dayringer, H.E., Violand, B.N. and Bentle, L.A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6434–6437.
- Barton, B.E., Jackson, J.V., Lee, F. and Wagner, J. (1994) *Cytokine*, **6**, 147–153.
- Bax, A. and Pochapsky, S.S. (1992) *J. Magn. Reson.*, **99**, 638–643.
- Bazan, J.F. (1990) *Immunol. Today*, **11**, 350–354.
- Boulton, T.G., Stahl, N. and Yancopoulos, G.D. (1994) *J. Biol. Chem.*, **269**, 11648–11655.
- Brown, T.J., Rowe, J.M., Shoyab, M. and Gladstone, P. (1990) *UCLA Symp. Mol. Cell. Biol., New Ser.*, **131**, 195–206.
- Brown, T.J., Rowe, J.M., Liu, J. and Shoyab, M. (1991) *J. Immunol.*, **147**, 2175–2180.
- Brown, T.J., Liu, J., Brashem-Stein, C. and Shoyab, M. (1993) *Blood*, **82**, 33–37.
- Bruce, A.G., Hoggatt, I.H. and Rose, T.M. (1992a) *J. Immunol.*, **149**, 1271–1275.
- Bruce, A.G., Linsley, P.S. and Rose, T.M. (1992b) *Prog. Growth Factor Res.*, **4**, 157–170.
- Cai, J., Gill, P.S., Masood, R., Chandrasoma, P., Jung, B., Law, R.E. and Radka, S.F. (1994) *Am. J. Pathol.*, **145**, 74–79.
- Clore, G.M. and Gronenborn, A.M. (1989) *Crit. Rev. Biochem. Mol. Biol.*, **24**, 479–564.
- Clore, G.M. and Gronenborn, A.M. (1991) *Annu. Rev. Biophys. Biophys. Chem.*, **20**, 29–63.
- De Vos, A.M., Ultsch, M. and Kossiakoff, A.A. (1992) *Science*, **255**, 306–312.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.
- Fogh, R.H., Schipper, D., Boelens, R. and Kaptein, R. (1994) *J. Biomol. NMR*, **4**, 123–128.
- Fogh, R.H., Schipper, D., Boelens, R. and Kaptein, R. (1995) *J. Biomol. NMR*, **5**, 259–270.
- Friedrichs, M.S., Mueller, L. and Wittekind, M. (1994) *J. Biomol. NMR*, **4**, 703–726.
- Garrett, D.S., Powers, R., Gronenborn, A.M. and Clore, G.M. (1991) *J. Magn. Reson.*, **95**, 214–220.
- Gearing, D.P. and Bruce, A.G. (1992) *New Biologist*, **4**, 61–65.
- Grzesiek, S. and Bax, A. (1992a) *J. Am. Chem. Soc.*, **114**, 6291–6293.
- Grzesiek, S. and Bax, A. (1992b) *J. Magn. Reson.*, **96**, 432–440.
- Grzesiek, S. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 185–204.
- Grzesiek, S., Anglister, J. and Bax, A. (1993a) *J. Magn. Reson. Ser. B*, **101**, 114–119.
- Grzesiek, S., Anglister, J., Ren, H. and Bax, A. (1993b) *J. Am. Chem. Soc.*, **115**, 4369–4370.
- Horn, D., Fitzpatrick, W.C., Gompper, P.T., Ochs, V., Bolton-Hanson, M., Zarlring, J., Malik, N., Todaro, G.J. and Linsley, P.S. (1990) *Growth Factors*, **22**, 157–165.
- Kallestad, J.C., Shoyab, M. and Linsley, P.S. (1991) *J. Biol. Chem.*, **266**, 8940–8945.
- Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.
- Landon, M. (1977) *Methods Enzymol.*, **47**, 145–149.
- Liu, J., Modrell, B., Aruffo, A., Scharnowske, S. and Shoyab, M. (1994) *Cytokine*, **6**, 272–278.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989a) *Biochemistry*, **28**, 6150–6156.
- Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989b) *J. Magn. Reson.*, **85**, 393–399.
- Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989c) *J. Am. Chem. Soc.*, **111**, 1515–1517.
- Miles, S.A., Martinez-Maza, O., Rezai, A., Magpantay, L., Kishimoto, T., Nakamura, S., Radka, S.F. and Linsley, P.S. (1992) *Science*, **255**, 1432–1434.
- Nair, B.C., DeVico, A.L., Nakamura, S., Copeland, T.D., Chen, Y., Patel, A., O'Neil, T., Oroszlan, S., Gallo, R.C. and Sarngadharan, M.G. (1992) *Science*, **255**, 1430–1432.
- Nishimoto, N., Ogata, A., Shima, Y., Tani, Y., Ogawa, H., Nakagawa, M., Sugiyama, H., Yoshizaki, K. and Kishimoto, T. (1994) *J. Exp. Med.*, **179**, 1343–1347.
- Powers, R., Gronenborn, A.M., Clore, G.M. and Bax, A. (1991) *J. Magn. Reson.*, **94**, 209–213.
- Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1992a) *Biochemistry*, **31**, 4334–4346.
- Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1992b) *Science*, **256**, 1673–1677.
- Prickett, K.S., Amberg, D.C. and Hopp, T.P. (1989) *Biotechniques*, **7**, 580–589.
- Remerowski, M.L., Domke, T., Groenewegen, A., Pepermans, H.A.M., Hilbers, C.W. and Van de Ven, F.J.M. (1994) *J. Biomol. NMR*, **4**, 257–278.
- Richards, C.D., Brown, T.J., Shoyab, M., Baumann, N. and Gauldie, J. (1992) *J. Immunol.*, **148**, 1731–1736.
- Robinson, R.C., Grey, L.M., Staunton, D., Vankelcom, H., Vernallis, A.B., Moreau, J.F., Stuart, D.I., Heath, J.K. and Jones, E.Y. (1994) *Cell*, **77**, 1101–1116.
- Rose, T.M. and Bruce, A.G. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 8641–8645.

- Rose, T.M., Weiford, D.M., Gunderson, N.L. and Bruce, A.G. (1994) *Cytokine*, **6**, 48–54.
- Seavey, B.R., Farr, E.A., Westler, W.M. and Markley, J.L. (1991) *J. Biomol. NMR*, **1**, 217–230.
- Sklenář, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J. Magn. Reson. Ser. A*, **102**, 241–245.
- Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.
- Sporeno, E., Paonessa, G., Salvati, A.L., Graziani, R., Delmastro, P., Ciliberto, G. and Toniatti, C. (1994) *J. Biol. Chem.*, **269**, 10991–10995.
- Tanigawa, T., Nicola, N., McArthur, G.A., Strasser, A. and Begley, C.G. (1995) *Blood*, **85**, 379–390.
- Vuister, G.W. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 7772–7777.
- Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994a) *J. Am. Chem. Soc.*, **116**, 11655–11666.
- Yamazaki, T., Lee, W., Revington, M., Mattiello, D.L., Dahlquist, F.W., Arrowsmith, C.H. and Kay, L.E. (1994b) *J. Am. Chem. Soc.*, **116**, 6464–6465.
- Zarling, J.M., Shoyab, M., Marquardt, H., Hanson, M.B., Lioubin, M.N. and Todaro, G.J. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9739–9743.