Journal of Biomolecular NMR, 8 (1996) 123–135
ESCOM

J-Bio NMR 366

Complete ^{1}H, ^{15}N and ^{13}C assignments, secondary structure, and topology of recombinant human interleukin-6

Guang-Yi Xu[a,*], Jin Hong[a], Tom McDonagh[a], Mark Stahl[a], Lewis E. Kay[b], Jasbir Seehra[a] and Dale A. Cumming[a]

[a]*Small Molecule Drug Discovery Group, Genetics Institute, 87 CambridgePark Drive, Cambridge, MA 02140, U.S.A.*
[b]*Department of Medical Genetics, Biochemistry and Chemistry, University of Toronto, Toronto, ON, Canada M5S 1A8*

Received 25 March 1996
Accepted 10 May 1996

Keywords: Interleukin-6; Secondary structure; Topology

Summary

Essentially complete backbone and side-chain ^{1}H, ^{15}N and ^{13}C resonance assignments for the 185-amino-acid cytokine interleukin-6 (IL-6) are presented. NMR experiments were performed on uniformly [^{15}N]- and [^{15}N,^{13}C]-labeled recombinant human IL-6 (rIL-6) using a variety of heteronuclear NMR experiments. A combination of ^{13}C-chemical shift, amide hydrogen-bond exchange, and ^{15}N-edited NOESY data allowed for analysis of the secondary structure of IL-6. The observed secondary structure of IL-6 is composed of loop regions connecting five α-helices, four of which are consistent in their length and disposition with the four-helix bundle motif present in other related cytokines and previously postulated for IL-6. In addition, the topology of the overall fold was found to be consistent with a left-handed up-up-down-down four-helix bundle based on a number of long-range interhelical NOEs. The results presented here provide deeper insight into structure–function relationships among members of the four-helix bundle family of proteins.

Introduction

Interleukin-6 (IL-6) is a 185-amino-acid cytokine which exerts multiple biological effects in vivo. Originally identified as a T-cell-derived factor regulating B-cell growth and differentiation (Hirano et al., 1986), IL-6 has also been shown to induce the production of acute-phase proteins in liver (Kishimoto, 1989), to serve as an autocrine and paracrine growth factor in the pathogenesis of human multiple myeloma (Kawano et al., 1988), to stimulate the recruitment and formation of osteoclasts (Ishimi et al., 1990), and to enhance, in conjunction with IL-3, the formation of multilineage blast cells and the proliferation of stem cells in the hematopoietic system (Bernad et al., 1994). IL-6 is also an inflammatory mediator. IL-6 production is markedly stimulated by the inflammatory cytokines IL-1 and TNF-α (Van Damme et al., 1987; Mawatari et al., 1989) and neutralizing antibodies to IL-6 protect mice from lethal challenges of *E. coli* or TNF (Starnes et al., 1990). Dysregulation of IL-6 production has been implicated in a variety of inflammatory/autoimmune disease states, including rheumatoid arthritis, cardiac myxoma, Castleman's disease, and mesangial proliferative glomerulonephritis (reviewed in Hirano et al., 1990).

The biological effects of IL-6 on target cells are mediated through a heterotrimeric receptor complex formed by the sequential engagement of IL-6 with an IL-6-specific receptor subunit (IL-6Rα) followed by recruitment of gp130 (Kishimoto et al., 1992). While both IL-6Rα and gp130 are integral membrane proteins with a single transmembrane domain, activation of intracellular signalling pathways is dependent on association of gp130 (Taga et al., 1989). Specifically, IL-6/IL-6Rα induces disulfide-linked homodimerization of gp130, which alone leads to activation of the JAK/STAT signalling pathway (Murakami et al., 1993; Ivashkiv, 1995). Thus, the final receptor complex which can minimally mediate signal transduction is believed to be hexameric and composed of two mol-

*To whom correspondence should be addressed.
Supplementary Material is available upon request.

ecules each of IL-6, IL-6Rα, and gp130, with intermolecular contacts between all components. Indeed, residues in IL-6 critical for specific binding to IL-6Rα or gp130 have been identified by site-directed mutagenesis (Savino et al., 1993,1994; Paonessa et al., 1995). Several other cytokines, including LIF, Oncastatin M, and CNTF, also employ receptors that are members of the hematopoietin receptor superfamily and that also associate with gp130 (or gp130-like molecules) for subsequent signal transduction (Kishimoto et al., 1992; Davis et al., 1993).

It is therefore of interest that, despite a modest level of amino acid sequence homology, all four of these cytokines have been suggested to share a common three-dimensional structural motif, the 'up-up-down-down' left-handed four-helix bundle (Sprang and Bazan, 1993). A family of cytokines and their receptors with shared biochemical and three-dimensional structural features, yet vastly different biological activities, provides an exciting opportunity for delineating structure–function relationships in cytokine biology. Thus, we have undertaken the three-dimensional structure determination of rIL-6 by multidimensional heteronuclear NMR spectroscopy. As a first step in this process, we report here the sequence-specific assignments, secondary structure, and the folding topology for the isotopically labeled, recombinant IL-6 produced in *E. coli*.

Materials and Methods

Sample preparation

The polypeptide for rIL-6 was overexpressed in *Escherichia coli*. [U-^{15}N]rIL-6 and [U-^{13}C,U-^{15}N]rIL-6 were purified from cells grown at 37 °C on minimal medium containing 2 g/l [^{15}N]ammonium sulfate alone or in combination with 3 g/l [U-^{13}C]glucose. The cells were lysed in 50 mM Tris-HCl, pH 8.5, at 4 °C, 1 mM EDTA, 5 mM DTT, 5 mM PABA. The lysate was centrifuged at 9000 × g at 4 °C for 10 min. The pellet containing insoluble, unfolded rIL-6 was resuspended in 50 mM Tris-HCl, 8M Urea, pH 8.5, and subjected to a refolding protocol essentially as previously described for IL-4 (Van Kimmenade et al., 1988) Purified rIL-6 was concentrated and dialyzed using vacuum dialysis into 10 mM MES, pH 6.1, 200 mM $MgSO_4$, 10% D_2O, 0.05% NaN_3.

NMR spectroscopy

All NMR experiments were performed on Varian Unity$^+$ spectrometers operating at either 500 or 600 MHz. Both instruments are equipped with triple-resonance (^{1}H, ^{13}C, ^{15}N) probes with an actively shielded z-gradient coil and pulsed field gradient accessories. Both ^{15}N-labeled and ^{15}N,^{13}C double-labeled samples of rIL-6 were dissolved in 90% H_2O/10% D_2O containing 200 mM $MgSO_4$, pH 6.1 (25 °C) to a final protein concentration of 1.3 mM. Slowly exchanging amide protons were identified by recording a series of two-dimensional ^{1}H-^{15}N HSQC experiments over a time span of 5 min to 26 h, beginning immediately after dissolution of lyophilized ^{15}N-labeled rIL-6 in D_2O.

^{1}H-^{15}N HSQC spectra were recorded with the enhanced-sensitivity pulsed field gradient approach (Bax and Pochapsky, 1992; Kay et al., 1992a; Schleucher et al., 1993; Muhandiram and Kay, 1994). All triple-resonance HSQC-based experiments (e.g. CBCA(CO)NNH, C(CC)TOCSY_NNH, H(CC)TOCSY_NNH, HNCACB) for correlating backbone or/and aliphatic side-chain chemical shifts (Kay et al., 1990; Grzesiek and Bax, 1992, 1993; Wittekind and Mueller, 1993) and which detect amide proton magnetization also employed enhanced-sensitivity pulsed field methods. This approach provides coherence transfer selection both to improve sensitivity and eliminate artefacts as well as for solvent suppression. No presaturation or spin-lock purging pulses (Messerle et al., 1989) were used in order to avoid saturation of amide protons by spin diffusion or chemical exchange with water. The (HB)CBCACO(CA)HA experiment (Kay, 1993) provided an independent means to complement assignments obtained by the HNCO experiment (Ikura et al., 1990). Constant-time ^{1}H-^{13}C 2D HSQC experiments (Vuister and Bax, 1992), optimized for selection of aromatic side chains or aliphatic terminal methyl groups, were performed to confirm aromatic carbon and methyl carbon assignments, respectively. A constant-time 2D methyl-relay experiment (Kay et al., unpublished results) was used to obtain correlations between δ-carbons of leucine residues, as well as γ-carbons of valine. The 3D long-range-^{13}C-^{13}C (LRCC) experiment (Bax et al., 1992, 1994) was employed to correlate ε-methyls to the C^β and C^γ of methionine residues, as well as to correlate δ-methyls to C^α in leucine/isoleucine. ^{15}N-edited NOESY (Marion et al., 1989a; Zuiderweg and Fesik, 1989), ^{15}N-edited TOCSY (Marion et al., 1989b) and ^{1}H-^{15}N HMQC-NOESY-HMQC (Ikura et al., 1990) experiments were carried out on a sample of singly labeled rIL-6, using the enhanced-sensitivity pulsed field gradient method (Kay et al., 1992a; Schleucher et al., 1993; Muhandiram and Kay, 1994). For ^{15}N-edited NOESY experiments, mixing times of 50 or 100 ms were used, whereas in HMQC-NOESY-HMQC experiments a 100-ms mixing time was employed. ^{15}N-edited TOCSY experiments used a spinlock time of 40 ms. Pulsed field gradient HCCH-TOCSY (Kay et al., 1993a) experiments were recorded on a H_2O sample with mixing times of 8 and 16 ms. Simultaneous ^{15}N,^{13}C-edited NOESY (Pascal et al., 1994) experiments were recorded with mixing times of 80 and 100 ms. HMQC-J experiments (Kay and Bax, 1990) were used for determination of $^3J_{H\alpha NH}$ values. In all experiments, the ^{1}H carrier was placed on the H_2O resonance at 4.76 ppm, the ^{15}N carrier was placed at 115.0 ppm, and the ^{13}C carrier was set differentially, depending on the experiment, as shown in Table 1. Quadrature detection in all indirectly detected

dimensions was obtained through States-TPPI phase cycling (Marion et al., 1989c). Table 1 summarizes the acquisition parameters for each experiment employed.

Data processing

Data sets were typically processed and displayed on SGI workstations using the program packages NMRDraw and NMRPipe (Delaglio et al., 1995). Post-acquisition water suppression was employed in all spectra where proton magnetization was detected in the acquisition time-domain (Marion et al., 1989d). In the case of simultaneous ^{15}N,^{13}C-edited NOESY experiments, a time domain deconvolution procedure was used to minimize the signal from residual water. A 60° phase-shifted sine-bell function and a single zero-filling was used in each of the t_3 and t_1 dimensions prior to Fourier transformation. For most HSQC-based experiments, the time domain was extended by a factor of two using forward–backward linear prediction in the ^{15}N (t_2) dimension. In all constant-time ^1H-^{13}C correlation experiments, mirror image linear prediction (Zhu and Bax, 1990) was used prior to zero-filling to the double time-domain data points. For HMQC-J data sets, VNMR software was employed for processing and spectral apodization. The programs PSC and PIPP (Garrett et al., 1991) were used for spectral plotting, peak picking and analysis.

Results

Complete $^1H/^{15}N/^{13}C$ assignments

Figure 1 shows the ^1H-^{15}N HSQC spectrum of uniformly ^{15}N-labeled rIL-6. Of the 225 cross peaks which theoretically should be observed in this spectrum (comprising 176 amide NHs, 24 pairs of side-chain amide NH_2's, and an indole NH of tryptophan), roughly 200 are observed. As shown in Fig. 1, there are spectral regions which exhibit a high degree of resonance overlap and degeneracy, as expected for proteins of this size and predicted helical content (Powers et al., 1992). It can also be noted that a large number of peaks are broadened (short apparent T_2 values) which is caused, in part, by slow relative motion in some α-helical regions. Taken together, these observations indicate that triple-resonance experiments are essential for NMR-based structural analysis of rIL-6.

Of the predicted 24 Gln/Asn side-chain amide NH_2 groups, 22 can be observed in Fig. 1, along with their corresponding NHD signals. These latter peaks result from chemical exchange with the 10% D_2O present in the solution for field-frequency locking. These NHD groups appear about 0.5 ppm upfield of their NH_2 counterparts, and can be used in amide proton/nitrogen-based triple-resonance experiments to assist the assignment process (Benjamin et al., 1995). In these HSQC and related experiments 13 very strong signals with narrow line widths were observed, all which lie within the N-terminal 19 residues of IL-6 and indicate that this segment is much more mobile than the rest of the molecule. Preliminary ^{15}N-relaxation studies further substantiate these observations, yielding typical T_2 values of about 250 ms in the N-terminal region versus T_2 values of about 40 ms in the rest of molecule (data not shown).

The sequential assignment procedure used for rIL-6 is similar to those reported previously for NMR studies of proteins such as interleukin-4 (Powers et al., 1992), Profi-

TABLE 1
ACQUISITION PARAMETERS FOR NMR EXPERIMENTS ON rIL-6

NMR experiments	Nucleus			No. of complex points			Spectral width (ppm)			Reference (ppm)		
	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
HSQC	^{15}N	^1H		256	512		32.6	13.3		115	4.76	
HSQC_J	^{15}N	^1H		400	512		32.6	13.3		115	4.76	
HSQC_CT (arom)	^{13}C	^1H		100	512		59.7	13.3		125	4.76	
HSQC_CT (CH_2)	^{13}C	^1H		100	512		29.8	13.3		20	4.76	
Methyl-relay	^{13}C	^1H		064	512		30.0	13.3		20	4.76	
HNCO	^{13}CO	^{15}N	^1H	052	040	512	12.0	32.6	13.3	177	115	4.76
CBCA(CO)NNH	^{13}C	^{15}N	^1H	050	040	512	60.9	32.6	13.3	43	115	4.76
C(CC)TOCSY_NNH	^{13}C	^{15}N	^1H	48	32	512	60.9	32.6	13.3	43	115	4.76
H(CC)TOCSY_NNH	^1H	^{15}N	^1H	64	32	512	4.16	32.6	13.3	4.76	115	4.76
HNCACB	^{13}C	^{15}N	^1H	50	40	512	60.9	32.6	13.3	43	115	4.76
HN(COCA)HA	^1H	^{15}N	^1H	50	32	512	3.10	32.6	13.3	4.76	115	4.76
HN(CA)HA	^1H	^{15}N	^1H	50	32	512	3.10	32.6	13.3	4.76	115	4.76
(HB)CBCACO(CA)HA	^{13}C	^{13}CO	^1H	50	52	384	60.9	10.0	13.3	43	177	4.76
^{15}N-edited TOCSY	^1H	^{15}N	^1H	128	32	512	11.7	32.6	13.3	4.76	115	4.76
^{15}N-edited NOESY	^1H	^{15}N	^1H	128	32	512	11.7	32.6	13.3	4.76	115	4.76
HMQC-NOESY-HMQC	^{15}N	^{15}N	^1H	48	64	512	32.6	32.6	13.3	115	115	4.76
HCCH-TOCSY	^1H	^{13}C	^1H	128	32	416	7.10	23.9	13.3	4.76	43	4.76
Simultaneous ^{15}N/^{13}C-edited NOESY	^1H	^{13}C	^1H	128	32	416	10.0	23.9	13.3	4.76	43	4.76
		^{15}N								4.76	115	4.76
Long-range-CCJ	^{13}C	^{13}C	^1H	64	32	416	58.0	20.0	13.3	43	23	4.76

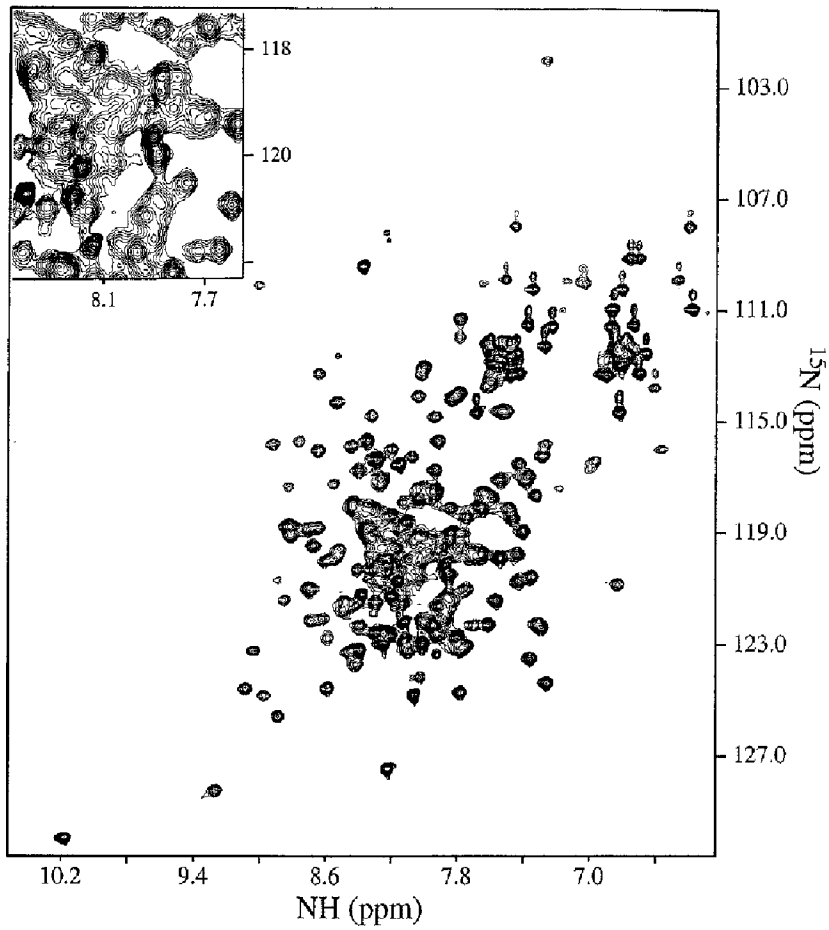

Fig. 1. 2D ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled rIL-6. In the upper left corner is an enlargement of a spectral region exhibiting severe overlap.

lin I (Archer et al., 1993), Recoverin (Ames et al., 1994), Protein S (Bagby et al., 1994) and CBDcex (Xu et al., 1995). However, additional experiments were needed to determine assignments, since peak broadening and overlap for rIL-6 were more extensive than for the above proteins.

The first step in the assignment process is to identify the residue type for all contributing spin systems. This was done using a combination of triple-resonance ($^{1}H/^{15}N/^{13}C$) experiments: CBCA(CO)NNH, C(CC)TOCSY_NNH and H(CC)TOCSY_NNH. All three experiments correlate backbone amide proton and nitrogen chemical shifts of one residue with the aliphatic side-chain carbon or proton chemical shifts of the preceding residue. Together, these correlations form diagnostic patterns for each residue type. The CBCA(CO)NNH experiment is the most useful of the three, providing information on not only the C^{α} chemical shift, but also the C^{β} chemical shift. However, distinguishing between certain residue pairs (e.g. proline and valine, leucine and aspartic acid, as well as lysine and arginine) is difficult with this experiment alone. This is due principally to degeneracy in C^{α} and C^{β} chemical shift values among these amino acid pairs. In these cases, the C(CC)TOCSY_NNH experiment is useful since it correlates backbone amide proton and nitrogen chemical shifts of one residue with all aliphatic side-chain carbon chemical shifts of the preceding residue. For example, the C^{α} and C^{β} chemical shifts of valine and proline residues can be essentially indistinguishable, making it difficult to assign a residue type based on these data alone. On the other hand, the C(CC)TOCSY_NNH-derived ^{13}C chemical shifts for these two residues allow unambiguous assignment due to the characteristic C^{γ} and C^{δ} chemical shifts for proline versus C^{γ} chemical shifts for valine. Figure 2 shows a representative 2D slice from the C(CC)TOCSY_NNH experiment, illustrating the correlation of backbone NH/^{15}N chemical shifts to the carbon chemical shifts of the previous residue side chains. In those cases where the C(CC)TOCSY_NNH experiment failed to provide sufficient side-chain information, the H(CC)TOCSY_NNH experiment was employed. This latter experiment correlates NH/^{15}N chemical shifts for a given residue with aliphatic side-chain proton chemical shifts of the preceding residue and thus helps assign the residue type. Distinguishing phenylalanine and tyrosine residues is more difficult, since the side chains of both contain AMX spin systems with similar ^{13}C and proton chemical shifts. In this case, analysis of aromatic side chains in a constant-time ^{1}H-^{13}C HSQC experiment provided proton chemical

shifts which were, in turn, correlated by homonuclear-NOESY experiments to H^{β}'s, thus allowing the assignment of the residue type. Glutamine and aspartate residues were distinguished from their amidated counterparts by virtue of the fact that the latter had C^{γ} and C^{β} chemical shifts which correlated to side-chain NHD groups. In principle, analysis of the data from the above three experiments should yield assignments for all residue types in the protein, except from residues which precede proline. However, in practice, magnetization transfer was inadequate for more than 25 residues, as evidenced by missing correlation peaks.

The sequential connectivity of residues was derived by comparison of spectra from HNCACB and from CBCA(CO)NNH experiments. Ideally, HNCACB spectra show correlations between two pairs of C^{α} and C^{β} chemical shifts and the NH/^{15}N chemical shift of a given residue. One pair arises from intraresidue correlations while the other pair derives from the preceding residue. As noted above, CBCA(CO)NNH spectra exhibit correlations only to the C^{α} and C^{β} chemical shifts of the preceding residue. Thus, comparison of spectra from these two experiments can give unambiguous information about the sequence of the residues. Figure 3 illustrates such a spectral comparison for a stretch of residues from Glu^{56} to Asn^{62} of rIL-6 and highlights the importance of C^{β} chemical shifts in resolving ambiguities arising from degenerate (or nearly degenerate) C^{α} chemical shifts.

Although the combined use of these two experiments comprises a powerful sequential assignment strategy for many proteins, they were inadequate in the case of rIL-6. One reason is that degeneracy of C^{α} and C^{β} chemical shifts precludes unambiguous assignment. For example, the C^{β} chemical shift for 12 of the 23 leucine residues in rIL-6 were within ~0.5 ppm, which is approximately the resolution in the carbon dimension. To resolve such ambiguity, four experimental approaches were pursued: (i) application of HN(CA)HA (Clubb et al., 1992; Kay et al., 1992b) and HN(COCA)HA (Kay et al., 1992b) experiments to provide correlations between an NH and the H^{α}'s of the same and preceding residues; (ii) application of the (HB)CBCACO(CA)HA experiment (Kay, 1993b)

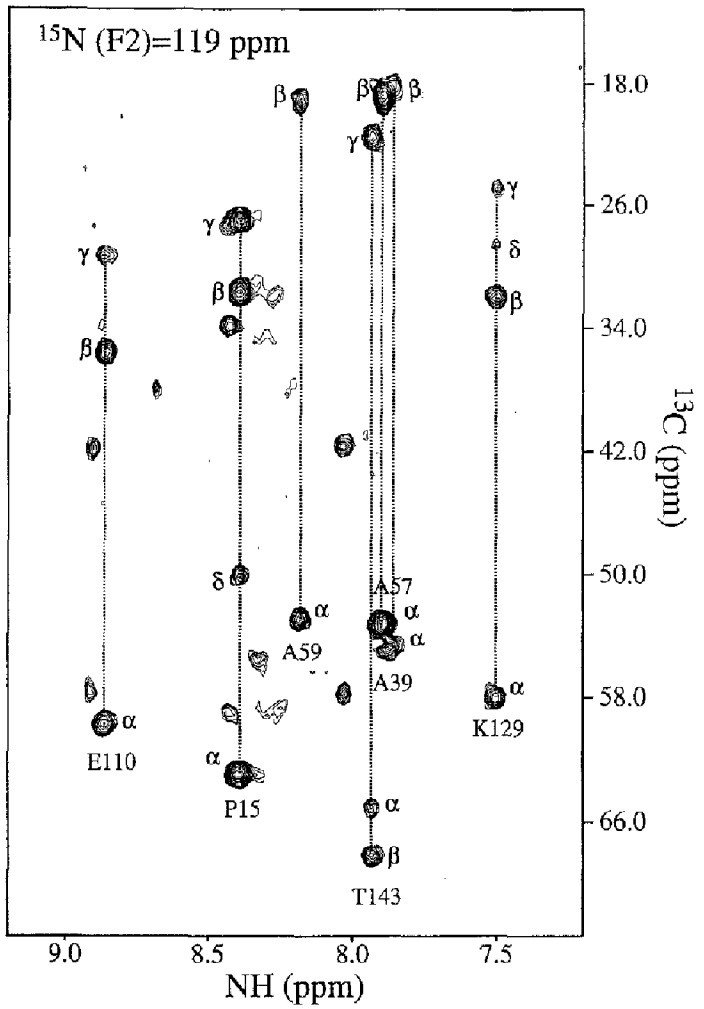

Fig. 2. A selected F1-F3 plane at ^{15}N (F2) = 119 ppm in the C(CC)TOCSY_NNH spectrum of rIL-6. The aliphatic ^{13}C chemical shifts of each residue are correlated with the amide NH/^{15}N chemical shifts of their succeeding residue.

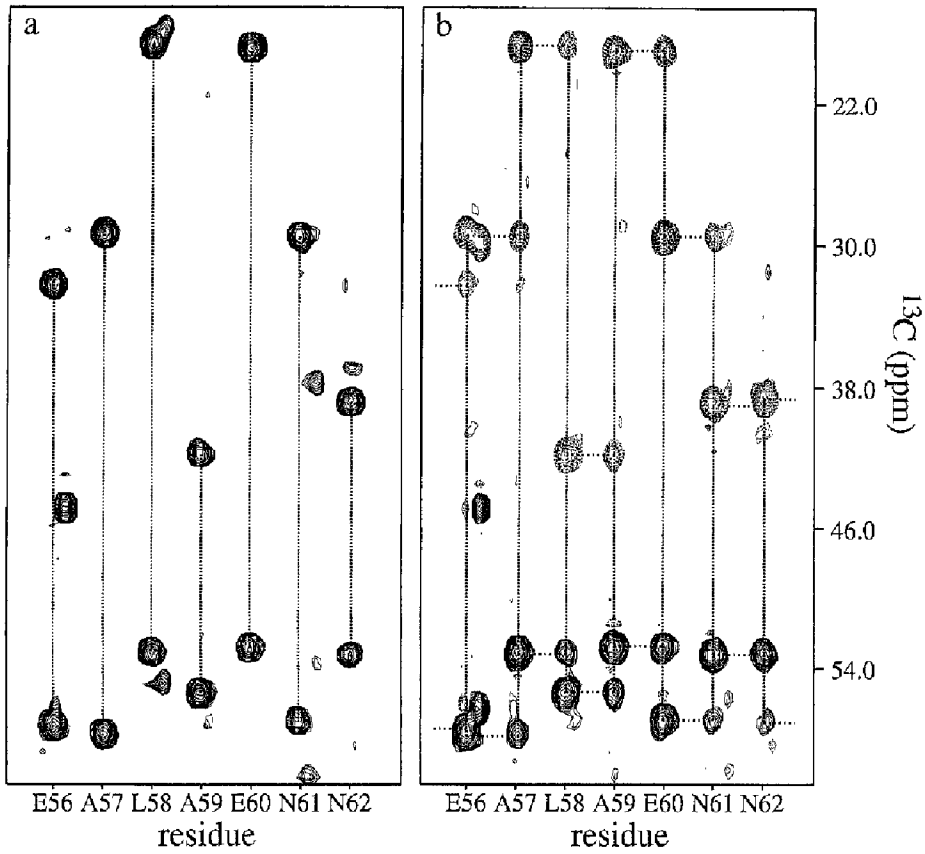

Fig. 3. F1 ($C^{\alpha/\beta}$) strip plots for residues Glu^{56} to Asn^{62} from (a) CBCA(CO)NNH and (b) HNCACB spectra of rIL-6. In (b), dashed contours (negative peaks) are C^{β}s and horizontal dashed lines indicate sequential C^{α}, C^{β} connectivities.

to establish correlations between adjacent residues independent of amide protons; (iii) application of ^{15}N-edited NOESY experiments (Marion et al., 1989a), including the 3D ^{1}H-^{15}N HMQC-NOESY-HMQC experiment (Ikura et al., 1990), to take advantage of the high helical content of rIL-6 (see below). The ^{15}N-edited NOESY yields a characteristic set of interresidue NOEs (Fig. 4) and HMQC-NOESY-HMQC is particularly useful for the recognition of NOEs between sequential, but overlapping, amide protons based on their different amide ^{15}N chemical shifts. Figure 5a is a strip plot from the HMQC-NOESY-HMQC experiment for a helical region in rIL-6 (Thr^{143} to Leu^{152}), showing quite distinct ^{15}N chemical shifts in the F1 dimension, whereas the corresponding amide proton chemical shifts were degenerate or near degenerate (Fig. 5b); (iv) application of three constant-time ^{1}H-^{13}C correlation experiments (a 'methyl-group-optimized' ^{1}H-^{13}C HSQC experiment, a ^{1}H-^{13}C methyl-relay experiment, and a long-range ^{13}C-^{13}C J-coupling ($^{2}J_{C-C}$ and $^{3}J_{C-C}$) correlation experiment (Bax et al., 1992; Vuister et al., 1994)) to exploit the methyl groups of residues such as leucine, isoleucine, or methionine for probing the shifts of other side-chain atoms. The latter set of experiments were most useful in conjunction with HCCH-TOCSY or NOESY data and are illustrated in Fig. 6. By this combined approach, essentially all 185 sequential assignments were obtained for rIL-6. A listing of the tabulated ^{1}H, ^{15}N, and ^{13}C chemical shifts for rIL-6 is available as Supplementary Material.

Secondary structure

The secondary structure of rIL-6 was determined through the use of four sets of experimental data: NOEs (Wüthrich, 1986), three-bond J-coupling constants between amide protons and their corresponding H^{α}'s ($^{3}J_{HN-H\alpha}$), chemical exchange rates of amide hydrogens with D_2O, and ^{13}C chemical shift perturbations (Spera and Bax, 1991; Wishart et al., 1991; Wishart and Sykes, 1994). These data are summarized for rIL-6 in Fig. 7.

Distinct secondary structure elements can be defined by unique sets of internuclear distances and hence yield characteristic NOE patterns. rIL-6 has abundant α-helical content, as evidenced by contiguous amide proton NOEs [$d_{NN(i,i+1)}$], NOEs between H^{α} of residue i and NH of residues i + 3 and i + 4 [$d_{\alpha N(i,i+3)}$, $d_{\alpha N(i,i+4)}$], as well as medium to weak NOEs between H^{α} of residue i and NH of residue i + 1 [$d_{\alpha N(i,i+1)}$] (Fig. 7). Moreover, residues within α-helices usually have $^{3}J_{\alpha N}$ values of 5 Hz or less and hydrogen bonds between carbonyl oxygens and amide hydrogens four residues away ($CO_i - NH_{i+4}$). Examination of amide hydrogen exchange rates and $^{3}J_{\alpha N}$ values in rIL-6 further support the position and length of α-helices sug-

gested by NOE patterns. The C^{α} and C^{β} secondary chemical shifts indices measure deviations from random-coil ^{13}C chemical shifts for corresponding amino acids when followed by alanine in the polypeptide (Spera and Bax, 1991; Wishart et al., 1994,1995) and have been shown to correlate with the protein secondary structure. In general, the minimal criteria for defining a helical segment are positive deviations from random-coil C^{α} chemical shifts and the presence of $d_{\alpha N(i,i+3)}/d_{\alpha N(i,i+4)}$ NOEs. The presence of other indicators was taken as supportive of the minimal criteria. Analysis of the data in Fig. 7 clearly delineates the following secondary structure elements. IL-6 has four long helices: helix A extends from Thr^{21} to Lys^{47}, helix B from Glu^{81} to Asn^{104}, helix C from Glu^{110} to Lys^{130}, and helix D from Gln^{157} to Met^{185}. The length and relative position of these secondary structure elements fit the four-helix bundle motif previously postulated as the topological fold for rIL-6 (Sprang and Bazan, 1993). In addition, Fig. 7 shows that all four experimental indicators are consistent with a fifth helix located at residues Pro^{142}–Gln^{153}. The presence of a short six-residue turn between helices B and C yields strong $d_{\alpha N(i,i-1)}$ NOEs between Ser^{108}–Ser^{109} and Ser^{109}–Glu^{110}, and several other weak $d_{\alpha N(i,i+2)}$, $d_{NN(i,i+2)}$ NOEs between Phe^{106}–Ser^{108}, Glu^{107}–Ser^{109}, Ser^{108}–Glu^{110} and Ser^{109}–Glu^{111}, as well as a large (>8 Hz) $^{3}J_{H\alpha\text{-}NH}$ for Ser^{109} and a slow-exchange amide NH for Glu^{110}. This latter observation is presumably due to a hydrogen bond between the NH of Glu^{110} and the carbonyl of Glu^{107} (Wüthrich, 1986).

In addition to these features, there is a long loop between helices A and B, extending from Ser^{48} to Asn^{80}. Within this loop, Cys^{51} forms a disulfide bridge to Cys^{45} (Clogsten et al., 1989). The continuous sets of $d_{NN(i,i-1)}$ NOEs from Ser^{48} to Glu^{52} as well as an NOE between the H^{α} of Cys^{45} and the NH of Met^{50}, plus NOEs from the H^{α}'s of Lys^{47} and Ser^{48} to the NHs of Asn^{49} and Met^{50}, respectively, indicate that this loop has a tight turn because of the disulfide bridge. After Glu^{52}, the loop appears to be more disordered, as indicated by the lack of NOEs until near the end of this loop. In this region, Cys^{74} forms a disulfide bond to Cys^{84} (Clogston et al., 1989), located within helix B.

Topology and protein folding

The overall topological fold of the five helices of rIL-6 was probed using data from simultaneous ^{15}N,^{13}C-edited

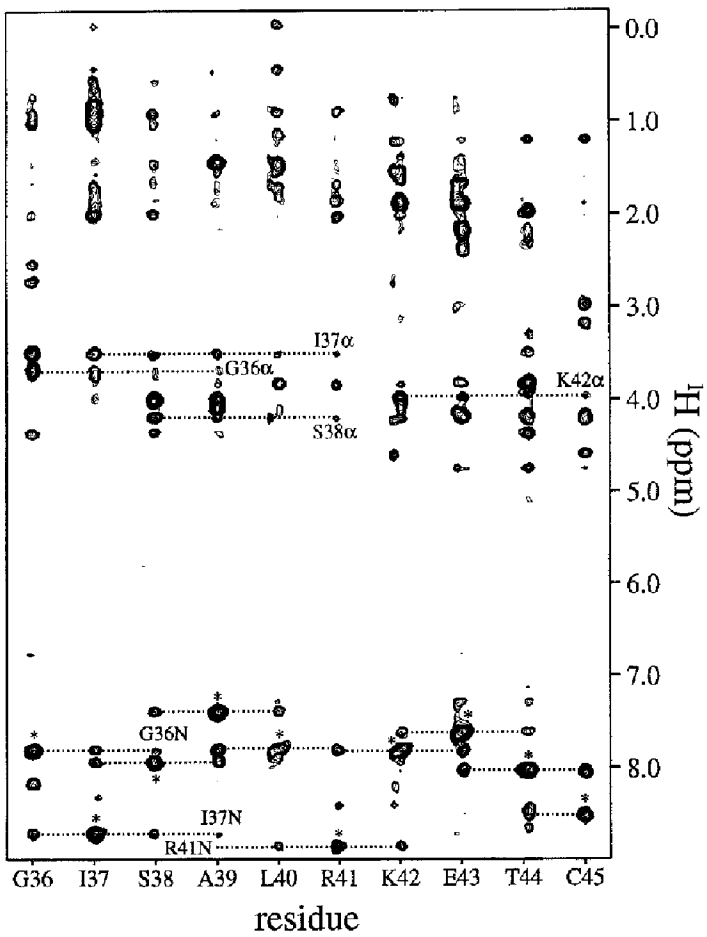

Fig. 4. A strip plot from the 3D ^{15}N-edited NOESY spectra (100 ms mixing time) of ^{15}N-labeled rIL-6 for residues Gly^{36} to Cys^{45} in the first α-helical (αA) region. The dashed lines (····) illustrate some of the expected $d_{NN(i,i+1)}$, $d_{NN(i,i+2)}$, $d_{\alpha N(i,i+3)}$ and $d_{\alpha N(i,i+4)}$ NOEs observed. Diagonal peaks are labeled with an asterisk (*).

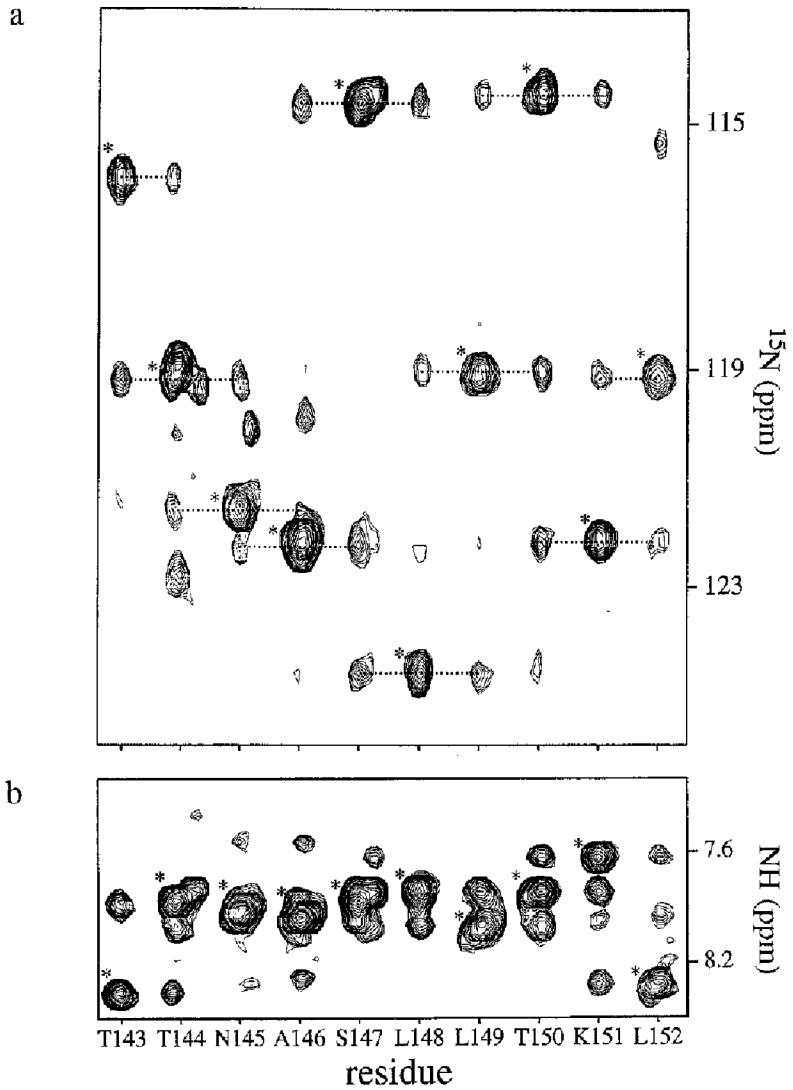

Fig. 5. Representative strip plots from (a) a 3D ^{1}H-^{15}N HMQC-NOESY-HMQC and (b) the NH region of a 3D ^{15}N-edited NOESY spectrum of ^{15}N-labeled rIL-6, corresponding to residues Thr^{143} to Leu^{152} in a α-helical (αE) segment. Diagonal peaks are labeled with an asterisk (*). Dashed lines (····) in (a) indicate sequential NH-NH connectivities.

NOESY and 2D homonuclear-NOESY experiments. All together, more than 100 long-range, interhelical NOEs were obtained, which defined spatially proximal interhelical residue pairs. For example, NOEs were found between the residue pairs Leu^{40}–Leu^{99} and Leu^{40}–Phe^{106}, defining in part the relative disposition of helices A and B. Similarly, NOEs were found between the residue pairs Ile^{26}–Leu^{127}, Gln^{29}–Leu^{123}, Ile^{33}–Thr^{120}, Ile^{37}–Val^{116}, Leu^{40}–Val^{116} and Leu^{40}–Ala^{113}, corresponding to helix A–C interactions, between the residue pairs Asp^{27}–Leu^{179}, Ile^{30}–Leu^{179}, Ile^{30}–Leu^{175}, Ile^{34}–Leu^{175}, Ile^{37}–Leu^{175}, Ile^{37}–Phe^{171} and Arg^{41}–Leu^{168}, corresponding to helix A–D interactions, between the residue pairs Glu^{96}–Thr^{120}, Leu^{99}–Thr^{120}, Gln^{103}–Ala^{113} and Phe^{106}–Ala^{113}, corresponding to helix B–C interactions, and between the residue pair Cys^{84}–Met^{185}, Cys^{84}–Ala^{181}, Ile^{88}–Ser^{178}, Ile^{88}–Phe^{174}, Gly^{91}–Phe^{174}, Phe^{95}–Phe^{174}, Phe^{95}–Phe^{171}, corresponding to helix B D interactions.

Figure 8 summarizes, in a helical wheel format, the interhelical NOE contacts derived from these NOESY experiments. It can be readily appreciated from Fig. 8 that helix A is parallel to helix B, helix C is parallel to helix D, and that the A–B helices are antiparallel to helices C and D. In addition, several NOEs were found between residue pairs Val^{97}–Asn^{145}, Val^{97}–Leu^{148}, Glu^{100}–Leu^{149} and Tyr^{101}–Leu^{152}, thus demonstrating that helices B and E are parallel to one another. Therefore, the four long α-helices are connected by cross-over loops to form an up-up-down-down antiparallel bundle. Moreover, the interhelical contacts summarized in Fig. 8 show very clearly the left-handed nature of the four-helix bundle and thus define the overall fold of the molecule. Figure 8 also illustrates the amphiphilic nature of the helices, with hydrophobic residues directed towards the interior of the bundle and hydrophilic residues oriented towards the exterior (solvent).

Discussion

Essentially complete sequence-specific ^{1}H, ^{15}N, and ^{13}C assignments have been made for uniformly [^{15}N]-/[^{13}C,^{15}N]-labeled rIL-6. Previous studies (Kruttgen et al., 1990; Rock et al., 1992), as well as inspection of the 1D ^{1}H NMR and ^{15}N-edited NOESY spectra obtained during the course of these studies, indicated a high helical content for the protein, potentially complicating NMR structural studies due to substantial chemical shift degeneracy and resulting spectral overlap. Indeed, these complications have been observed in NMR studies of other highly helical proteins such as IL-4 (Garrett et al., 1992) and the Gail domain of the heterotrimeric G proteins (Benjamin et al., 1995). In the latter case, application of an assignment strategy, largely predicated upon two heteronuclear experiments (CBCA(CO)NNH and HNCACB), nonetheless provided the experimental means to assign virtually the entire molecule. rIL-6 proved to be more difficult, presumably due to its greater size (185 versus 145 residues in Gail) and this basic assignment strategy required significant augmentation. However, application of a battery of heteronuclear experiments (CBCA(CO)NNH, HNCACB, HN(CA)HA, HN(COCA)HA, and HMQC-NOESY-HMQC) did allow sequence-specific assignment of the vast majority of atoms in rIL-6.

rIL-6 appears to be largely monomeric and structured in solution under the conditions employed here, as evidenced by the observed line widths in 2D-HSQC spectra and gel filtration studies (data not shown). This is in contrast to previous reports (May et al., 1991), as well as our own data (not shown), where multimers of rIL-6 could be observed under certain conditions. The biological significance of multimers of rIL-6 is not clear, since other closely related cytokines (e.g. CNTF, M-CSF) clearly are dimeric under physiological conditions (Pandit et al., 1992; McDonald et al., 1995). Residues at the N-terminus of rIL-6 show narrower line widths, enhanced signal intensity in 2D-HSQC experiments, very rapid amide exchange rates and few NOEs, indicating that this segment is unstructured and/or highly mobile in solution. In contrast, residues which comprise the hydrophobic core of rIL-6 exhibit notably broader line widths, consistent with a relatively immobile, tightly packed core.

Secondary structure analysis of rIL-6 was performed utilizing several experimental indicators: ^{13}C chemical shift, amide hydrogen-bond exchange, and ^{15}N-edited NOESY data. The predominant secondary structural element is the α-helix, as demonstrated by all indicators.

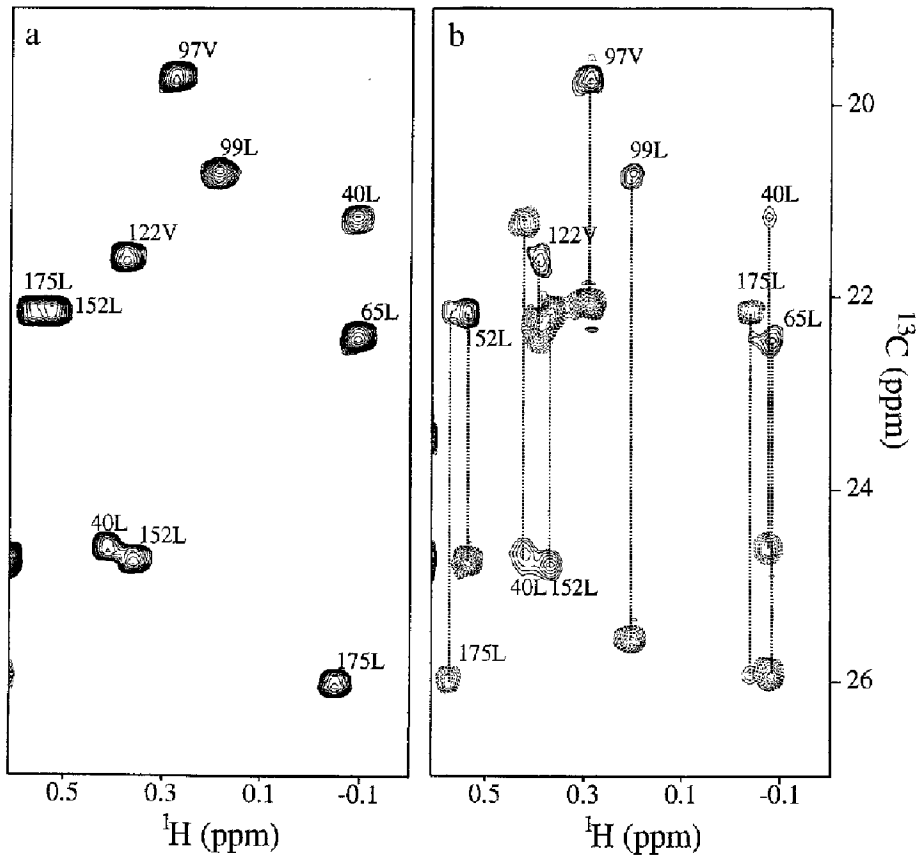

Fig. 6. (a) A representative spectral region from a 2D constant-time ^{1}H-^{13}C HSQC spectrum optimized for methyls and (b) a 2D methyl-relay spectrum. The region shown encompasses a portion of the high-field contact shift region in the proton spectrum of rIL-6. Together, these experiments facilitate intraresidue side-chain assignments for residues such as valine and leucine.

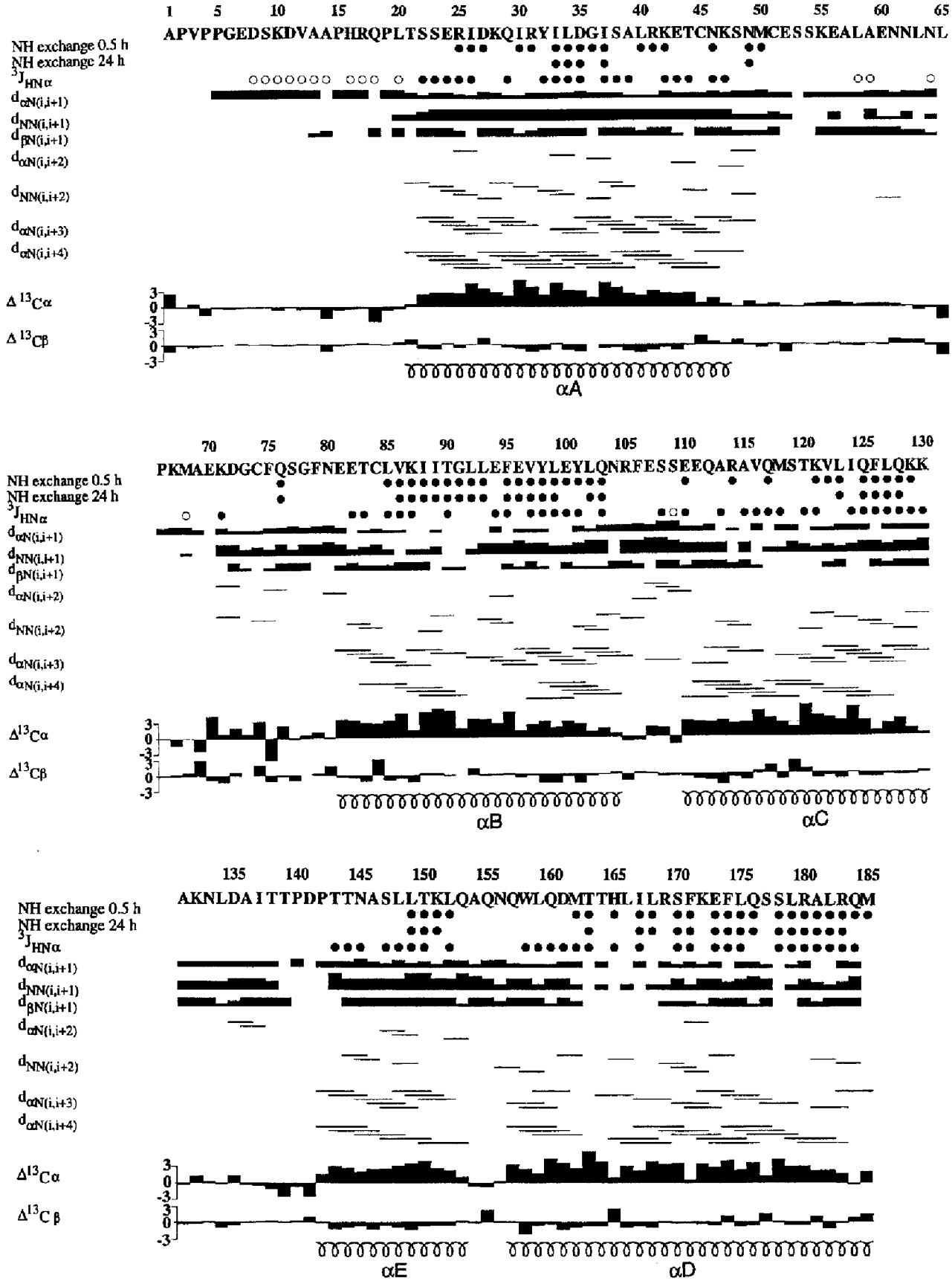

Fig. 7. Summary of NMR-derived secondary structure indicators measured for rIL-6 together with the deduced secondary structure. For NOE data, the thickness of the lines reflects the strength of sequential NOEs. Filled (●) and open (○) circles indicate $^3J_{HN\alpha}$ values of less than 5 Hz and greater than 8 Hz, respectively.

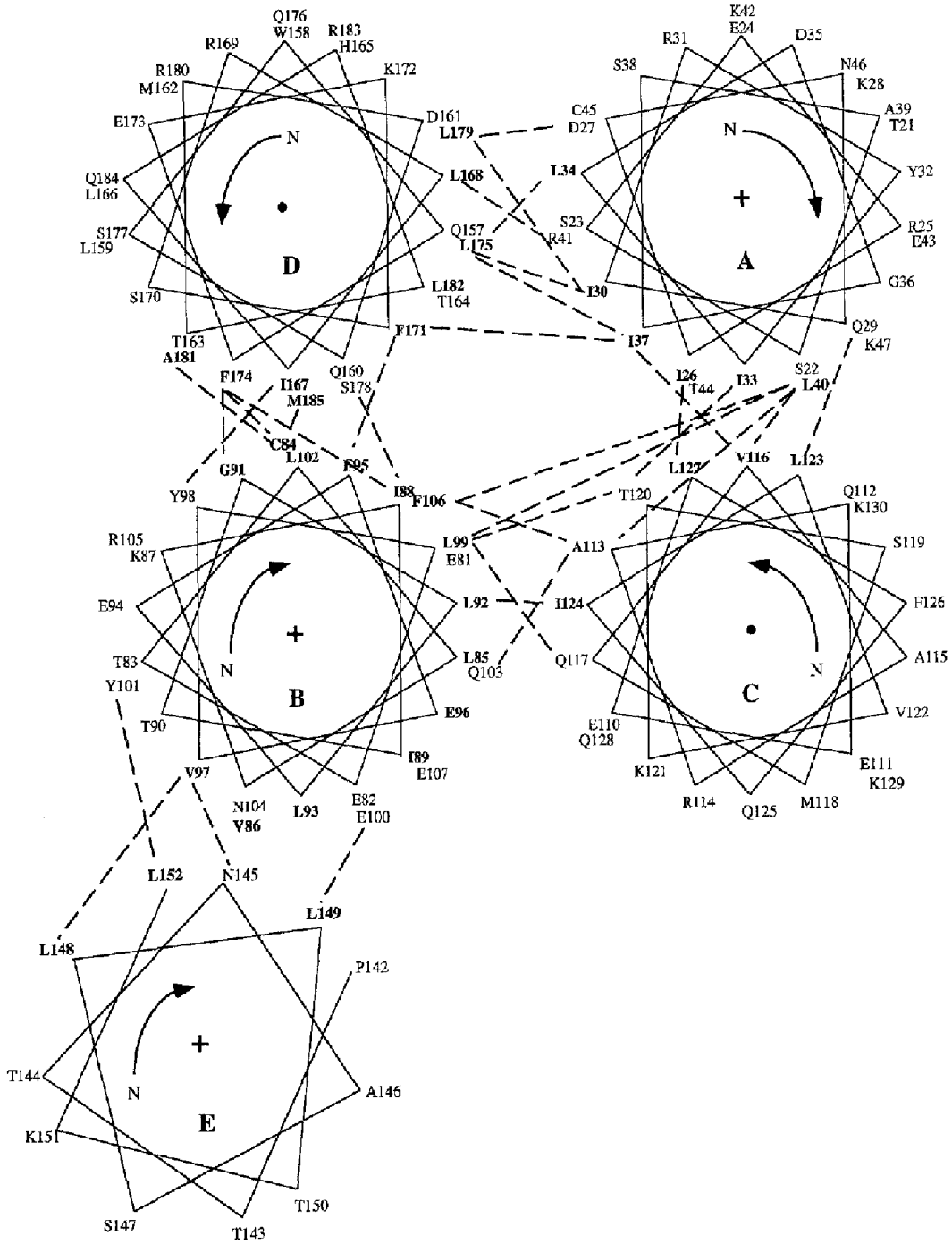

Fig. 8. Summary of the interhelical NOE contacts. The helices A, B, and E are viewed from their N-termini, while helices C and D are viewed from their C-termini. Hence, helices A, B and E are parallel to each other and antiparallel to helices C and D. Interhelical NOE contacts are denoted by dashed lines (– – –), while hydrophobic residues in the interior of the bundle are marked in bold.

Overall, the secondary structure is characterized by four long helical stretches, one short one, and interconnecting loops (Fig. 7). The secondary structure analysis of rIL-6 reported here is consistent with previous predictions, based on sequence homology to cytokines of known structure (Bazan, 1991). The topology for rIL-6 is a left-handed, antiparallel four-helix bundle. Thus, the four long helices form the bundle with the long extended loops after the first and third helices presumably corresponding to the prototypical AB and CD cross-over loops. Additional topological constraints derive from the two disulfide bonds in rIL-6 between Cys^{45}–Cys^{51} and Cys^{74}–Cys^{84} (Kruttgen et al., 1990; Rock et al., 1992,1994).

Indeed, other cytokine members of this family have been divided into two subgroups based upon, among other things, overall polypeptide chain length and average helix length (Sprang and Bazan, 1993). One of these subgroups is the 'short-chain' four-helix bundle, as typified

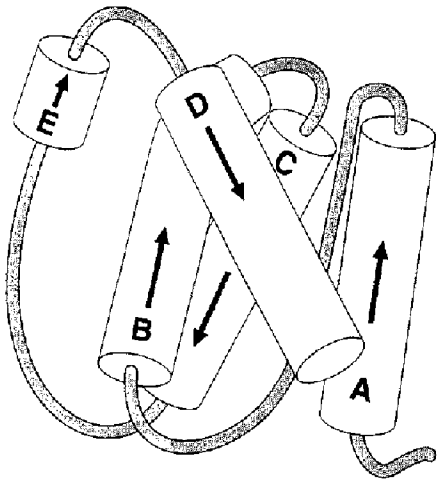

Fig. 9. Schematic diagram of the topological fold of rIL-6, as derived in this study.

by Il-2, IL-4, or IL-5. Members of this subgroup typically possess polypeptide chain lengths of 105–145 amino acids, with typical helices of about 15 residues in length and cross-over loops containing short, twisted antiparallel β-strands. The other subgroup is the 'long-chain' four-helix bundle and is typified by G-CSF, IL-10, and interferon-β. The polypeptide chain lengths for members of this subgroup are longer (160–200 amino acids), as are the typical helix lengths (about 25 residues). The cross-over loops in members of this subgroup lack the antiparallel β-strands present in short-chain cytokines, but can contain additional, 'extra-core' helices, such as in the CD cross-over loop of INF-β. Thus, NMR-based secondary structure analysis is consistent with the prediction that rIL-6 is also a member of the long-chain subgroup of four-helix bundles with an overall chain length of 185 amino acids, long helices, no β-strands in the cross-over loops, and an additional, 'extra-core' short helix in the CD loop.

A set of over 100 interhelical NOEs, obtained from homo- and heteronuclear NOESY experiments, provides further experimental support for a four-helix bundle topology for rIL-6. The topology of a four-helix bundle can be completely described in terms of three characteristics: the polypeptide connectivity between helices, the unit direction vectors of each helix, and the handedness of the bundle (Presnell and Cohen, 1989). While experimental evidence for the first characteristic is derived from the secondary structure analysis, the interhelical NOEs provide direct evidence for the latter two. As summarized in Fig. 8, the set of interhelical NOEs clearly show the antiparallel disposition of the helices and the left-handed nature of the bundle. The overall topology of rIL-6 is summarized as a schematic diagram in Fig. 9.

The topology of rIL-6 would be anticipated to include other specific structural features, as observed in the X-ray crystal structures of other subgroup members, such as LIF and CNTF (Robinson et al., 1994; McDonald et al., 1995). These include characteristic packing between the A–D and B–C helix pairs, distinct packing of the AB loop over the CD loop, and characteristic 'kinks' in either or both of the A or D helices. This latter feature is especially interesting, since such 'kinks' in the A and D helices of LIF and CNTF have been postulated to be important in forming appropriate receptor contacts (McDonald et al., 1995). The specifics of these and other structural details await three-dimensional structure determination.

Acknowledgements

We thank Elizabeth DeBlasio and Dr. John McCoy for the expression vector for rIL-6 and Dr. Neil Schauer for microbial fermentations. We also thank Dr. Ranjith Muhandiram (University of Toronto) for invaluable technical advice and assistance, Drs. Dan Garrett and Frank DeLaglio (NIH) for providing their software packages for data processing and display, and Dr. Hsiang-Ai Yu (Genetics Institute) for valuable discussions and advice.

References

Ames, J.B., Tanaka, T., Stryer, L. and Ikura, M. (1994) *Biochemistry*, **33**, 10734–10753.

Archer, S.J., Bax, A., Roberts, A.B., Sporn, M.B., Ogawa, Y., Piez, K.A., Weatherbee, J.A., Tsang, M.L.-S., Lucas, R., Zheng, B.-L., Wenker, J. and Torchia, D.A. (1993) *Biochemistry*, **32**, 1152–1163.

Bagby, S., Harvey, T.S., Kay, L.E., Eagle, S.G., Inouye, S. and Ikura, M. (1994) *Biochemistry*, **33**, 2409–2421.

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

Bax, A., Max, D. and Zax, D. (1992) *J. Am. Chem. Soc.*, **114**, 6924–6925.

Bax, A., Delaglio, F., Grzesiek, S. and Vuister, G.W. (1994) *J. Biomol. NMR*, **4**, 787–797.

Bazan, J.F. (1991) *Neuron*, **7**, 197–208.

Benjamin, D.R., Markby, D.W., Bourne, H.R. and Kuntz, I.D. (1995) *Biochemistry*, **34**, 155–162.

Bernad, A., Kopf, M., Kulbaki, R., Weich, N., Koehler, G. and Gutierrez-Ramos, J.C. (1994) *Immunity*, **1**, 725–731.

Clogston, C.L., Boone, T.C., Crandall, B.C., Mendiaz, E.A. and Lu, H.S. (1989) *Arch. Biochem. Biophys.*, **272**, 144–151.

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

Davis, S., Aldrich, T.H., Stahl, N., Pan, L., Taga, T., Kishimoto, T., Ip, N.Y. and Yancopoulos, G.D. (1993) *Science*, **260**, 1805–1808.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Garrett, D.S., Powers, R., Gronenborn, A.M. and Clore, G.M. (1991) *J. Magn. Reson.*, **95**, 214–220.

Garrett, D.S., Powers, R., March, C.J., Frieden, E.A., Clore, G.M. and Gronenborn, A.M. (1992) *Biochemistry*, **31**, 4347–4353.

Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **96**, 432–440.

Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Hirano, T., Yasukawa, K., Harada, H., Taga, T., Watanabe, Y., Matsuda, T., Kashiwamura, S., Nakajima, K., Koyama, K., Iwamatsu, A., Tsunasawa, S., Sakiyama, F., Matsui, H., Takahara, Y., Taniguichi, T. and Kishimoto, T. (1986) *Nature*, **324**, 73–76.

Hirano, T., Akiro, S., Taga, T. and Kishimoto, T. (1990) *Immunol. Today*, **11**, 443–449.

Ikura, M., Bax, A., Clore, G.M. and Gronenborn, A.M. (1990) *J. Am. Chem. Soc.*, **112**, 9020–9022.

Ishimi, Y., Miyaura, C., Jin, C.H., Akatsu, T., Abe, E., Nakamura, Y., Yamaguchi, A., Yoshiki, S., Matsuda, T., Hirano, T., Kishimoto, T. and Suda, T. (1990) *J. Immunol.*, **145**, 3297 3303.

Ivashkiv, L.B. (1995) *Immunity*, **3**, 1–14.

Kawano, M., Hirano, T., Matsuda, T., Taga, T., Horii, Y., Iwato, K., Asaoku, H., Tang, B., Tanabe, O., Tanaka, H. and Kishimoto, T. (1989) *Nature*, **332**, 83–85.

Kay, L.E. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 110–126.

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

Kay, L.E., Keifer, P. and Saarinen, T. (1992a) *J. Am. Chem. Soc.*, **114**, 1063–1065.

Kay, L.E., Wittekind, M., McCoy, M.A., Friedrichs, M.S. and Mueller, L. (1992b) *J. Magn. Reson.*, **98**, 443–450.

Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R. and Forman-Kay, J.D. (1993a) *J. Magn. Reson.*, **B101**, 333–337.

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

Kishimoto, T. (1989) *Blood*, **74**, 1–10.

Kishimoto, T., Akira, S. and Taga, T. (1992) *Science*, **258**, 593–597.

Kruttgen, A., Rose-John, S., Moller, C., Wroblowski, B., Wollmer, A., Mullberg, J., Hirano, T., Kishimoto, T. and Heinrich, P.C. (1990) *FEBS Lett.*, **262**, 323 326.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989a) *J. Am. Chem. Soc.*, **111**, 1515–1517.

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

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

Marion, D., Ikura, M. and Bax, A. (1989d) *J. Magn. Reson.*, **84**, 425 430.

Mawatari, M., Kohno, K., Mizoguchi, H., Matsuda, T., Asoh, K., Van Damme, J., Welgus, H.G. and Kuwano, M. (1989) *J. Immunol.*, **143**, 1619–1627.

May, L.T., Santhanam, U. and Sehgal, P.B. (1991) *J. Biol. Chem.*, **226**, 9950–9955.

McDonald, N.Q., Panayotatos, N. and Hendrickson, W.A. (1995) *EMBO J.*, **14**, 2689–2699.

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

Muhandiram, D.R. and Kay, L.E. (1994) *J. Magn. Reson.*, **B103**, 203–216.

Murakami, M., Hibi, M., Nakagawa, N., Nakagawa, T., Yasukawa, K., Yamanishi, K., Taga, T. and Kishimoto, T. (1993) *Science*, **260**, 1808–1810.

Pandit, J., Bohm, A., Jancarik, J., Halenbeck, R., Koths, K. and Kim, S.-H. (1992) *Science*, **258**, 1358–1362.

Paonessa, G., Graziani, R., De Serio, A., Savino, R., Ciapponi, L., Lahm, A., Salvati, A.L., Toniatti, C. and Ciliberto, G. (1995) *EMBO J.*, **14**, 1942 1951.

Pascal, S.M., Muhandiram, D.R., Yamazaki, T., Forman-Kay, J.D. and Kay, L.E. (1994) *J. Magn. Reson.*, **B103**, 197–201.

Powers, R., Garrett, D.S., March, C.J., Frieden, E.A., Gronenborn, A.M. and Clore, G.M. (1992) *Biochemistry*, **31**, 4334–4346.

Presnell, S.R. and Cohen, F.E. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6592–6596.

Robinson, R.C., Grey, L.M., Staunton, D., Vankelecom, H., Vernallis, A.B., Moreau, J.F., Stuart, D.I., Heath, J.K. and Jones, E.Y. (1994) *Cell*, **77**, 1101–1116.

Rock, F., Everett, M. and Klein, M. (1992) *Protein Eng.*, **5**, 583 591.

Rock, F.L., Li, X., Chong, P., Ida, N. and Klein, M. (1994) *Biochemistry*, **33**, 5146–5154.

Savino, R., Lahm, A., Giorgio, M., Cabibbo, A., Tramontano, A. and Cilberto, G. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4067–4071.

Savino, R., Lahm, A., Salvati, A.L., Ciapponi, L., Sporeno, E., Altamura, S., Paonessa, G., Tonoatti, C. and Ciliberto, G. (1994) *EMBO J.*, **13**, 1357–1367.

Schleucher, J., Sattler, M. and Griesinger, C. (1993) *Angew. Chem. Int. Ed. Engl.*, **32**, 1489–1491.

Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.

Sprang, S.R. and Bazan, J.F. (1993) *Curr. Opin. Struct. Biol.*, **3**, 815–827.

Starnes, H.F., Pearce, M.K., Towari, A., Yim, J.H., Zou, J.C. and Abrams, J.S. (1990) *J. Immunol.*, **145**, 4185–4191.

Taga, T., Hirata, Y., Yamasali, K., Yasukawa, K., Matsuda, T., Hirano, T. and Kishimoto, T. (1989) *Cell*, **58**, 573–581.

Van Damme, J., Opdenakker, G., Simpson, R.J., Rubira, M.R., Cayphas, S., Vink, A., Billiau, A. and Van Snick, J. (1987) *J. Exp. Med.*, **165**, 914–919.

Van Kimmenade, A., Bond, M.A., Schumacher, J.H., Laquoi, C. and Kastelein, R.A. (1988) *Eur. J. Biochem.*, **173**, 109–114.

Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.

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

Wishart, D.S., Richards, F.M. and Sykes, B.D. (1991) *J. Mol. Biol.*, **222**, 311–333.

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

Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) *J. Biomol. NMR*, **5**, 67–81.

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

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

Xu, G.Y., Ong, E., Gilkes, N.R., Kilburn, D.G., Muhandiram, D.R., Harris-Brandts, M., Carver, J.P., Kay, L.E. and Harvey, T.S. (1995) *Biochemistry*, **34**, 6993–7009.

Zhu, G. and Bax, A. (1990) *J. Magn. Reson.*, **90**, 405–410.

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