Determination of the Secondary Structure and Folding Topology of Human Interleukin-4 Using Three-Dimensional Heteronuclear Magnetic Resonance Spectroscopy†

Daniel S. Garrett,‡ Robert Powers,‡ Carl J. March,§ Eric A. Frieden,§ G. Marius Clore,*,‡ and Angela M. Gronenborn*,‡

Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, and Immunex Corporation, 51 University Street, Seattle, Washington 98101

Received December 23, 1991; Revised Manuscript Received February 28, 1992

ABSTRACT: The secondary structure of human recombinant interleukin-4 (IL-4) has been investigated by three-dimensional (3D) ^{15}N- and ^{13}C-edited nuclear Overhauser (NOE) spectroscopy on the basis of the ^{1}H, ^{15}N, and ^{13}C assignments presented in the preceding paper [Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., & Clore, G. M. (1992) *Biochemistry* (preceding paper in this issue)]. Based on the NOE data involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons, as well as $^{3}J_{HN\alpha}$ coupling constant, amide exchange, and $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ secondary chemical shift data, it is shown that IL-4 consists of four long helices (residues 9–21, 45–64, 74–96, and 113–129), two small helical turns (residues 27–29 and 67–70), and a mini antiparallel β-sheet (residues 32–34 and 110–112). In addition, the topological arrangement of the helices and the global fold could be readily deduced from a number of long-range interhelical NOEs identified in the 3D ^{13}C-edited NOE spectrum in combination with the spatial restrictions imposed by three disulfide bridges. These data indicate that the helices of interleukin-4 are arranged in a left-handed four-helix bundle with two overhand connections.

Human interleukin-4 (IL-4)[1] is a 15.4-kDa protein of 133 residues which plays a key role in the proliferation and differentiation of a wide range of cells within the immune system [see Paul and Ohara (1987), Yokota et al. (1988), and Finkelman et al. (1990) for reviews]. Thus IL-4 stimulates the proliferation of activated B cells, a variety of primary cells, and in vitro cell lines, including factor-dependent T cell and mast cell lines, T lymphocytes, thymocytes, and connective tissue-type mast cells. IL-4 also stimulates or suppresses the in vitro formation of hematopoietic progenitor cells in combination with various colony stimulating factors. In addition, IL-4 is a potent inducer of human cytotoxic $CD8^{+}$ T cells. The latter has recently been shown to be responsible for establishing tumor-specific systemic immunity to an established renal cancer by injection of renal tumor cells genetically engineered to secrete large doses of IL-4 locally (Golumbek et al., 1991). Further, IL-4 induces the expression of class II MHC molecules and the IgE low-affinity receptor on resting B cells and causes immunoglobulin class switching of activated B cells to IgE and IgG1, thereby generating and sustaining in vivo IgE responses and ensuring the dominance of IgG1 in a T cell dependent immune response, respectively (Kühn et al., 1991).

In order to better understand the mode of action of IL-4 and its interaction with its cell surface receptor, we have initiated a structural program aimed at determining its three-dimensional structure in solution by NMR spectroscopy. In the preceding paper (Powers et al., 1992), we presented the ^{1}H, ^{15}N, ^{13}C, and ^{13}CO assignments of IL-4 which comprise the essential foundation for such a study. In this paper, we present the determination of the secondary structure, topology, and global fold of IL-4 derived from an analysis of 3D heteronuclear-edited NOE spectra, ^{13}C chemical shift, amide chemical exchange, and $^{3}J_{HN\alpha}$ coupling constant data.

Experimental Procedures

Sample Preparation. Uniformly (>95%) ^{15}N- and $^{15}N/^{13}C$-labeled human recombinant IL-4 were expressed in yeast and purified as described previously (Powers et al., 1992). The version of recombinant IL-4 used comprises the tetrapeptide sequence Glu-Ala-Glu-Ala added to the N-terminus of the natural sequence. In the present paper, residues are numbered from the N-terminal Glu as residue 1, so that the natural IL-4 sequence starts at residue 5. Samples for NMR contained 2 mM ^{15}N- or $^{15}N/^{13}C$-labeled IL-4, pH 5.7, dissolved in either 90% H_2O/10% D_2O or 99.996% D_2O.

NMR Spectroscopy. All NMR spectra were recorded at 36 °C on a Bruker AM600 spectrometer equipped with a triple-resonance ^{1}H, ^{15}N, ^{13}C probe and modified with additional hardware consisting of synthesizers, amplifiers, and fast switching devices for the X-channel. Quadrature detection in the indirectly detected dimensions was obtained in all 3D experiments using the TPPI-States method (Marion et al., 1989a). The 3D ^{15}N-edited NOESY spectrum was recorded as described by Marion et al. (1989b), the 3D ^{13}C-edited

† This work was supported by the AIDS Directed Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C. and A.M.G.).

‡ NIDDK, NIH.

§ Immunex Corp.

[1] Abbreviations: IL-4, interleukin-4 (the numbering scheme used in the present paper includes the four-residue sequence Glu-Ala-Glu-Ala at the N-terminus of the recombinant protein which is not part of the natural human IL-4; the natural IL-4 sequence therefore starts at residue 5); NMR, nuclear magnetic resonance; CD, circular dichroism; IgE, immunoglobulin E; NOE, nuclear Overhauser effect; NOESY; nuclear Overhauser enhancement spectroscopy; HMQC, heteronuclear multiple-quantum coherence; 3D, three dimensional; 4D, four dimensional; TPPI, time-proportional phase incrementation.

0006-2960/92/0431-4347$03.00/0 © 1992 American Chemical Society

NOESY as described by Ikura et al. (1990a), and the 3D ^1H–^{15}N HMQC-NOESY-HMQC as described by Ikura et al. (1990b). Water suppression in the case of the 3D ^{15}N-edited NOESY spectrum was achieved by means of 3- and 6-ms spin-lock pulses applied along the x axis simultaneously with the first and second ^{15}N 90° pulses of the HMQC part of the sequence, respectively, to randomize magnetization of protons not attached to ^{15}N as described by Messerle et al. (1989). Water suppression in the 3D ^1H–^{15}N HMQC-NOESY-HMQC experiment was achieved by presaturation during the relaxation delay and NOESY mixing time.

In the case of the 3D ^{15}N-edited NOESY experiment, recorded with a mixing time of 120 ms, the spectral widths in the indirectly detected $^1H(F_1)$ and $^{15}N(F_2)$ dimensions were 11.4 and 30.01 ppm, respectively, with the carrier positions at 4.67 and 115 ppm, respectively, and the spectral width in the ^1H acquisition dimension (F_3) was 13.44 ppm with the carrier at 4.67 ppm. For the 3D ^{13}C-edited NOESY experiment recorded with a mixing time of 110 ms, the spectral widths in the indirectly detected $^1H(F_1)$ and $^{13}C(F_2)$ dimensions were 9.26 and 20.71 ppm, respectively, with the carrier positions at 4.3 and 63.71 ppm, respectively, and the spectral width in the ^1H acquisition dimension (F_3) was 10.04 pm with the carrier at 4.3 ppm. For the 3D ^1H–^{15}N HMQC-NOESY-HMQC experiment, recorded with a mixing time of 150 ms, the spectral width in the two ^{15}N dimensions (F_1 and F_2) was 30.01 ppm with the carrier at 115 ppm, and the spectral width in the ^1H acquisition dimension (F_3) was 13.44 ppm with the carrier at 4.67 ppm.

For the 3D ^{15}N-edited NOESY experiment, there were 128 complex points in F_1 (^1H), 32 complex points in F_2 (^{15}N), and 1024 real points in F_3 (^1H); for the 3D ^{13}C-edited NOESY experiment there were 128 complex points in F_1 (^1H), 32 complex points in F_2 (^{15}N), and 512 real points in F_3 (^1H); and for the 3D ^1H–^{15}N HMQC-NOESY-HMQC experiment there were 64 complex points in F_1 (^{15}N), 32 complex points in F_2 (^{15}N), and 1024 real points in F_3 (^1H).

$^3J_{HN\alpha}$ coupling constants were obtained from a 2D HMQC-J spectrum (Kay & Bax, 1990) as described by Forman-Kay et al. (1990). The HMQC-J spectrum was recorded using TPPI (Marion & Wüthrich, 1983) for quadrature detection in the t_1 dimension with 800 t_1 increments of 2K data points to give an acquisition time in t_1 of 219.2 ms. Zero-filling in F_1 gave a final digital resolution of 0.45 Hz/point in the ^{15}N dimension.

Slowly exchanging NH protons were identified by recording a series of ^{15}N–^1H Overbodenhausen correlation spectra (Bodenhausen & Ruben, 1980; Bax et al., 1990; Norwood et al., 1990) over a period of ~24 h starting within 5 min of dissolving an unexchanged sample of lyophilized protein in D_2O (Driscoll et al., 1990a).

All spectra were processed on a Sun Sparc Workstation using in-house routines for Fourier transformation (Kay et al., 1989), together with the commercially available software package NMR2 (New Methods Research, Ind., Syracuse, NY). Analysis of the 3D spectra and peak picking was carried out using the in-house programs CAPP and PIPP (Garrett et al., 1991).

Results and Discussion

Secondary Structure Determination. Elements of regular secondary structure can be reliably delineated from a qualitative analysis of NOE data involving the backbone protons, together with information on NH exchange rates and $^3J_{HN\alpha}$ coupling constants (Wüthrich, 1986; Clore & Gronenborn, 1987). Specifically, a stretch of consecutive sequential NH(i)–NH(i+1) NOEs, in conjunction with C$^\alpha$H(i)–NH(i+2,3,4), NH(i)–NH(i+2), and C$^\alpha$H(i)–C$^\beta$H(i+3) NOEs and $^3J_{HN\alpha}$ coupling constants less than 6 Hz, is characteristic of an α-helix. A stretch of strong sequential C$^\alpha$H(i)–NH(i+1) NOEs in combination with interstrand NOEs involving the NH and C$^\alpha$H protons, as well as $^3J_{HN\alpha}$ coupling constants of >8 Hz, is characteristic of a β-sheet. Slowly exchanging amide protons are indicative of hydrogen bonding and thus provide further evidence of the presence and nature of the secondary structure.

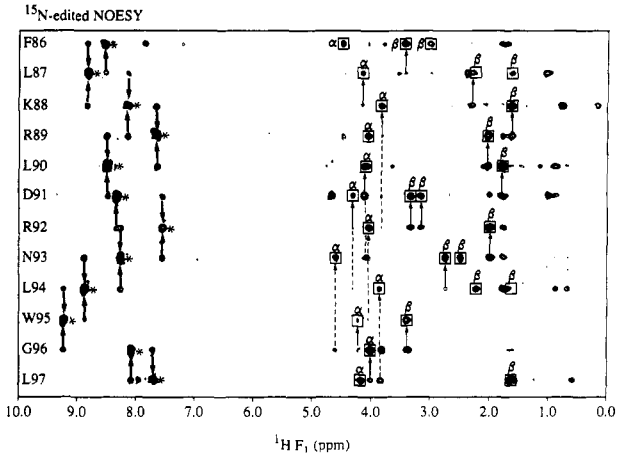

FIGURE 1: Composite of amide strips taken from the 120-ms mixing time 3D ^{15}N-edited NOESY spectrum of IL-4 for the stretch of sequence from Phe-86 to Leu-97 comprising helix C which illustrates typical NOEs found in a helical region. The diagonal peaks are indicated by an asterisk, NH(i)–NH($i\pm1$) NOEs are indicated by the thick arrows, NH(i+1)–C$^\alpha$H(i) and NH(i+1)–C$^\beta$H(i) NOEs by the thin arrows, and NH(i+3,4)–C$^\alpha$H(i) NOEs by the arrows with the dashed lines. Intraresidue C$^\alpha$H–NH and C$^\beta$H–NH NOEs are boxed. The $^1H(F_3)/^{15}N(F_2)$ chemical shifts for Phe-86, Leu-87, Lys-88, Arg-89, Leu-90, Asp-91, Arg-92, Asn-93, Leu-94, Trp-95, Gly-96, and Leu-97 are 8.50/118.8, 8.79/120.1, 8.18/118.8, 7.69/120.3, 8.51/120.2, 8.35/117.3, 7.57/114.9, 8.24/116.8, 8.83/121.8, 9.15/119.2, 8.06/105.7, and 7.69/121.2 ppm, respectively. The strips were generated from the 3D ^{15}N-edited NOESY spectrum as described by Driscoll et al. (1990b). It should be noted that within this stretch of sequence there are five NH(i)–NH(i+2) NOEs present which cannot be seen at the contour level plotted: these are between residues 86 and 88, 87 and 89, 88 and 90, 92 and 94, and 93 and 95.

Given the limited ^1H chemical shift dispersion exhibited by IL-4 (Powers et al., 1991), as well as its relatively large size by NMR standards (133 residues), it is no longer feasible to rely solely on 2D spectra, and recourse has to be made to 3D heteronuclear-edited spectroscopy to identify the appropriate NOEs. NOEs involving the NH protons were identified from 3D ^{15}N-edited NOESY (Marion et al., 1989b) and ^1H–^{15}N HMQC-NOESY-HMQC (Ikura et al., 1990b) spectra. The latter is particularly useful for the identification of NOEs between NH protons with degenerate or near degenerate chemical shifts. An example of the quality of the 3D ^{15}N-edited NOESY spectrum is provided in Figure 1, which illustrates a series of amide strips (Driscoll et al., 1990b) for the stretch of sequence extending from Phe-86 to Leu-97. This stretch of sequence illustrates typical NOEs found in a helical region including a stretch of consecutive sequential NH(i)–NH(i+1) NOEs together with some C$^\alpha$H(i)–NH(i+3) NOEs (from residue 86 to 89, 88 to 91, 90 to 93, 91 to 94, 92 to 95, 93 to 96, and 94 to 97) and a single C$^\alpha$H(i)–NH(i+4) NOE from residue 88 to 92. Figure 2 illustrates a $^{15}N(F_1)$–$^1H(F_3)$ plane of the ^1H–^{15}N HMQC-NOESY-HMQC spectrum at $\delta^{15}N(F_2)$ = 121.57 ppm. The diagonal peaks appear along a line parallel to F_3 at the $^{15}N(F_1)$ frequency equal to the $^{15}N(F_2)$ frequency. Only NOEs between amide protons are

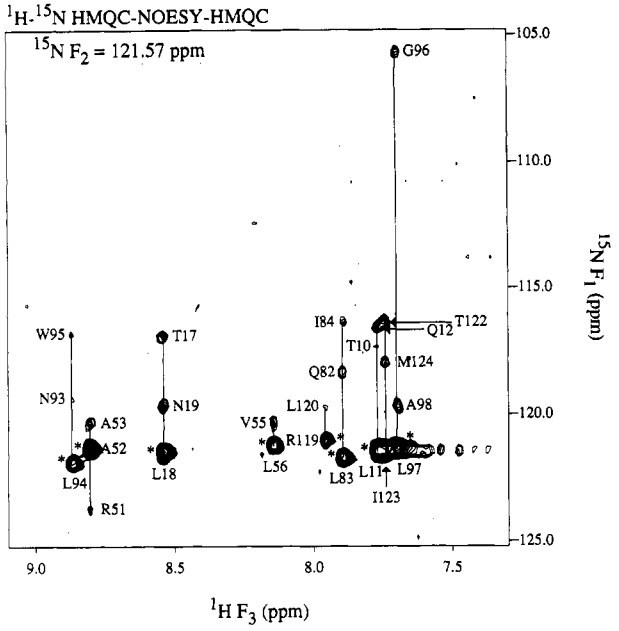

FIGURE 2: Typical $^{15}N(F_1)$–$^1H(F_3)$ plane of the 150-ms mixing time 3D ^{1}H–^{15}N HMQC-NOESY-HMQC spectrum of IL-4 at $\delta^{15}N(F_2)$ = 121.57 ppm, illustrating a number of NH(i)–NH($i\pm1$) NOEs. The "diagonal" peaks which occur at $\delta^{15}N(F_2) = \delta^{15}N(F_1)$ are indicated by asterisks.

detected in this spectrum, and 16 sequential NH(i)–NH($i\pm1$) NOEs are clearly observed. Of these 16 NOEs, four involve pairs of NH resonances separated by less than 0.1 ppm (Gln-12–Leu-11, Asn-19–Leu-18, Val-55–Leu-56, and Gln-82–Leu-83), of which the resonances of one pair (Gln-82–Leu-83) are separated by only 0.01 ppm.

Also important with regard to secondary structure determination are NOEs involving the C$^\alpha$H and C$^\beta$H protons which can be readily identified in a 3D ^{13}C-edited NOESY spectrum (Ikura et al., 1990a). A portion of the $^1H(F_1)$–$^1H(F_3)$ plane at $\delta^{13}C(F_2)$ = 57.56 ppm of the 3D ^{13}C-edited NOESY spectrum is illustrated in Figure 3. Despite the use of 3D spectroscopy, it is evident that there is still a significant degree of chemical shift overlap within this plane. Thus, for example, the C$^\alpha$H and ^{13}C$^\alpha$ shifts of Glu-23 and Leu-27 are degenerate, and the C$^\alpha$H shifts of Leu-11 and Leu-117 are near degenerate. Nevertheless, the resolution is still sufficient to identify a large number of NOE interactions, and six C$^\alpha$H(i)–C$^\beta$H(i+3) NOEs can be readily discerned (from Leu-11 to Ile-14, Leu-27 to Glu-30, Asp-91 to Leu-94, Glu-107 to Gln-110, Leu-117 to Leu-120, and Leu-120 to Ile-123).

The three-bond $^3J_{HN\alpha}$ coupling constants are related to the ϕ backbone torsion angles by an empirical Karplus relationship and therefore provide useful additional information for the assessment of secondary structure elements (Pardi et al., 1984). For small proteins, these couplings are generally obtained from the separation of the antiphase components of the NH–C$^\alpha$H cross-peaks in a 2D ^{1}H–^{1}H COSY type spectrum. In the case of IL-4, this was not possible due to two factors. First, the resolution afforded in such a spectrum was insufficient to permit reliable extraction of coupling constants. Second, as the ^{1}H line widths for IL-4 are broad and the minimum separation of the antiphase components is equal to ~0.58 times the linewidth at half-height (Neuhaus et al., 1986), the measured couplings no longer reflect the true couplings but are mainly a reflection of the linewidths. We therefore resorted to measuring the couplings from the splittings in the $^{15}N(F_1)$ dimension of a ^{1}H–^{15}N HMQC-J spectrum (Kay & Bax, 1990; Forman-Kay et al., 1990). A portion of the spectrum, illustrating how different size couplings can be easily discerned, is shown in Figure 4. For example, Ser-133 shows a large splitting reflecting a coupling of 8.3 Hz, the splitting of A38 is somewhat smaller (6.1 Hz), while the splittings of the other cross-peaks are all less than 5 Hz.

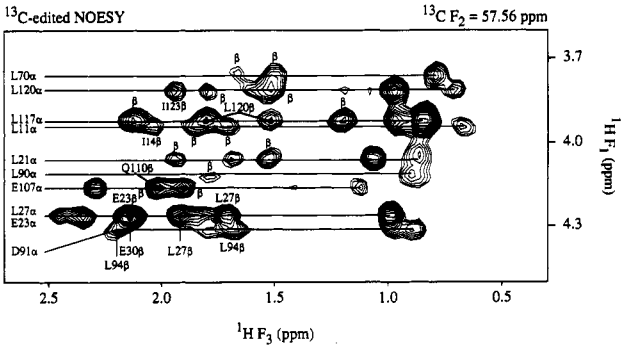

FIGURE 3: Portion of a typical $^1H(F_1)$–$^1H(F_3)$ plane of the 110-ms mixing time 3D ^{13}C-edited NOESY spectrum of IL-4 at $\delta^{13}C(F_2)$ = 57.56 ppm. This portion of the spectrum illustrates NOEs originating on aliphatic protons and going to C$^\alpha$H protons attached to ^{13}C$^\alpha$ atoms resonating at 57.56 ppm. The destination C$^\alpha$H protons are indicated on the left-hand side of the figure, and a number of C$^\alpha$H(i)–C$^\beta$H(i+3) NOEs are clearly discerned.

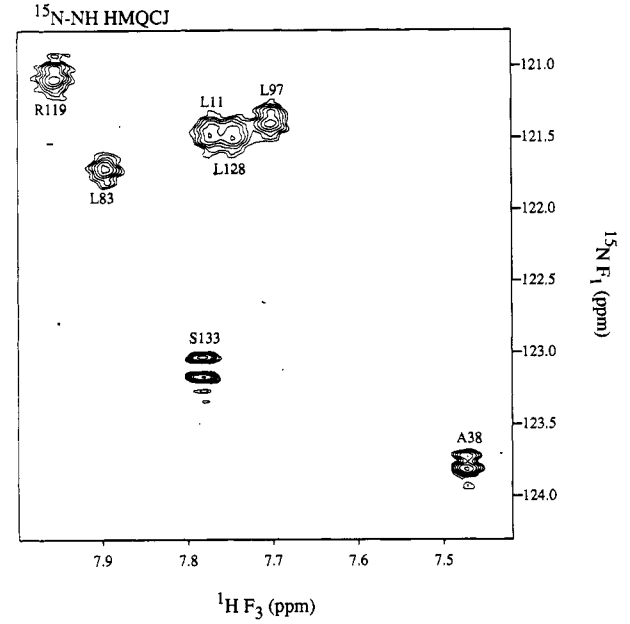

FIGURE 4: Portion of the ^{1}H–^{15}N HMQC-J spectrum of IL-4. The splittings in the $^{15}N(F_1)$ dimension are used to determine the $^3J_{HN\alpha}$ coupling constants. The spectrum has been processed with a −2 Hz negative exponential multiplication factor in the ^{15}N dimension and by a Lorenz-to-Gaussian apodization function in the $^1H(F_2)$ dimension. The digital resolution in the ^{15}N dimension (after zero-filling to 8K points) is 0.45 Hz per point.

A summary of the sequential and medium-range ($1 < |i - j| \leq 4$) NOE data involving the NH, C$^\alpha$H and C$^\beta$H protons, the $^3J_{HN\alpha}$ couplings, the NH exchange data, and the ^{13}C$^\alpha$ and ^{13}C$^\beta$ secondary chemical shifts is shown in Figure 5, together with the secondary structure elements deduced from this data. In addition to the NOE, coupling constant, and amide exchange data, the ^{13}C$^\alpha$ and ^{13}C$^\beta$ secondary chemical shifts provide a reliable indicator for the existence of β-sheet and α-helical structure (Spera & Bax, 1991). Specifically, for α-helices, the average ^{13}C$^\alpha$ and ^{13}C$^\beta$ secondary shifts are 3.09 ± 1.00 and −0.38 ± 0.85 ppm, respectively; while for β-sheets they are −1.48 ± 1.23 and +2.16 ± 1.91 ppm, respectively. Further, a histogram of C$^\alpha$ and C$^\beta$ secondary chemical shift

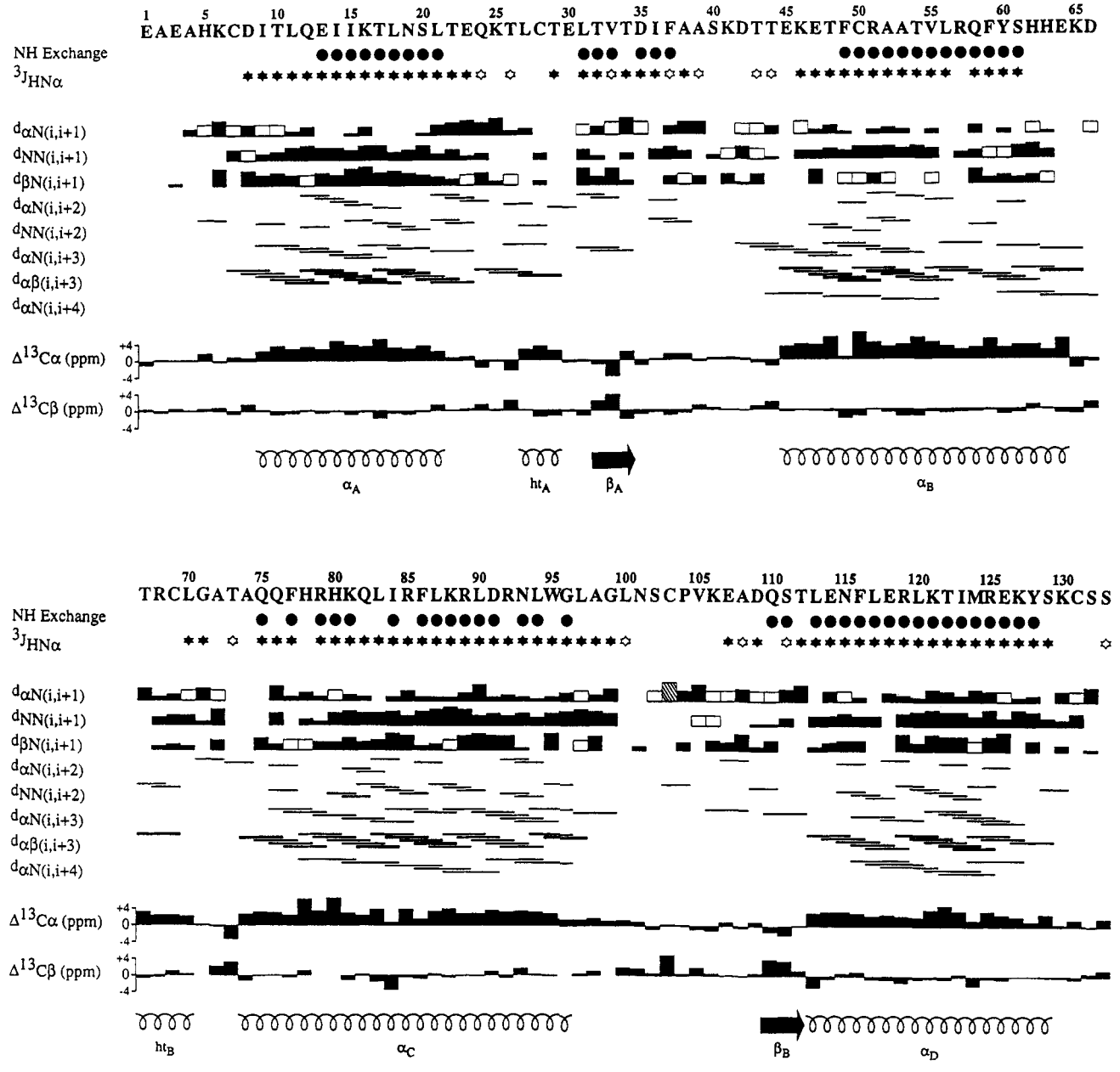

FIGURE 5: Summary of the sequential and medium range NOEs involving the NH, $C^\alpha H$, and $C^\beta H$ protons, the amide exchange and $^3J_{HN\alpha}$ coupling constant data, and the $^{13}C^\alpha$ and $^{13}C^\beta$ secondary chemical shifts observed for IL-4 together with the secondary structure deduced from these data. The thickness of the lines reflects the strength of the NOEs. Amide protons which are still present 20 min after taking up a lyophilized sample in D_2O are indicated by closed circles. The filled-in and open stars indicate $^3J_{HN\alpha}$ values <6 and >8 Hz, respectively. The open boxes represent potential sequential (i,i+1) NOEs which were obscured by resonance overlap and could therefore not be assigned unambiguously. The hatched box on the same line as the $C^\alpha H(i)$–$NH(i+1)$ NOEs represents the sequential NOE between the $C^\alpha H$ proton of Cys-103 and the $C^\delta H$ proton of Pro-104 and is indicative of a *trans*–proline.

values for α-helices and β-sheets indicates that in the case of the secondary $^{13}C^\alpha$ shifts there is no overlap between helix and β-sheet above +2 ppm (α-helix) and below 0 ppm (β-sheet); similarly, for the secondary $^{13}C^\beta$ shifts there is no overlap above +1.2 ppm (β-sheet). Inspection of Figure 5 indicates that this correlates perfectly with the information available from the NOE and coupling constant data and permits a more reliable definition of the beginnings and ends of secondary structure elements than would have been possible otherwise. Particularly telling in the case of helices is the general tendency for the secondary $^{13}C\alpha$ shifts to be small at the beginning of helix, increase to a maximum toward the middle of the helix, and then decrease again as one approaches the end of the helix.

Analysis of the data in Figure 5 clearly reveals the following secondary structure elements. There are four long helices: helix A extends from residue 9 to 21, helix B from residue 45 to 64, helix C from residue 74 to 96, and helix D from residue 113 to 129. Interestingly, in the case of helix B, there is a break in the general $^{13}C^\alpha$ secondary shift trend at Phe-49 such that Phe-49 exhibits a very small $^{13}C^\alpha$ secondary shift of only 0.5 ppm, while the two residues on either side, Thr-48 and Cys-50, display large secondary $^{13}C^\alpha$ shifts of 5.31 and 6.37 ppm. Both the NOEs in this region and the secondary $^{13}C^\beta$ chemical shift of Phe-49 (−2.1 ppm) are indicative of a helical conformation. The discontinuity in the secondary $^{13}C^\alpha$ chemical shift at Phe-49 may therefore suggest the presence of a distortion in the helix, most likely associated with the presence of the disulfide bridge involving the neighboring Cys-50. In addition to the four long helices, there are two small helical turns (residues 27–29 and 67–70), as well as a mini-antiparallel β-sheet comprising residues 32–34 and 110–112. The alignment of the β-sheet is shown in Figure 6

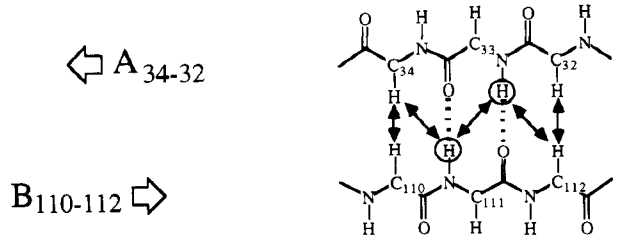

FIGURE 6: Mini-antiparallel β-sheet structure of IL-4 as determined from a qualitative analysis of the NOE and amide exchange data. The β-strands are indicated on the left by letters and the residue number range. Interstrand NOEs derived from the 3D ^{15}N- and ^{13}C-edited NOESY spectra are indicated by arrows, and slowly exchanging amide protons are encircled. The hydrogen bonds deduced from these data are shown as broken lines.

and illustrates classical interstrand NH–NH, C$^{\alpha}$H–C$^{\alpha}$H, and C$^{\alpha}$H–NH NOEs. This together with the presence of slowly exchanging NH protons for Val-33 and Ser-111 are characteristic of interstrand hydrogen bonding between the backbone amide groups of Val-33 and Ser-111 and the backbone carbonyls of Ser-111 and Val-33, respectively.

The present results are in excellent agreement with the secondary structure content determined from CD measurements (Curtis et al., 1991). Specifically, the percentages of α-helix (excluding the two small helical turns of three and four residues) and antiparallel β-sheet derived from the NMR data are 54% and 4.5%, respectively, compared to 55% and 5% from CD measurements.

Helical wheel projections of the four helices indicate that they are amphiphatic in character and contain hydrophobic residues along their entire lengths separated by three and four residues. Thus, these helices possess large hydrophobic faces providing ideal interaction surfaces for helix–helix packing. In addition, negatively charged capping residues are found at the N-terminal ends of helices A (Asp-8), B (Glu-45), and D (Glu-114) (Presta & Rose, 1988).

Comparison of the Secondary Structure with That Determined by Redfield et al. (1991). In a recent paper, Redfield et al. (1991) presented the ^{15}N and ^{1}H backbone assignments of IL-4 together with a secondary structure determination based on 3D ^{15}N-edited NOESY spectroscopy. A comparison of their secondary structure and ours is given in Table I. There is good agreement in the delineation of the helices and β-strands. Redfield et al. (1991) did not, however, provide any evidence for the existence of the two small helical turns extending from residues 27–29 and 67–70. With the exception of helix A, there are some differences in the exact positioning of the beginnings and ends of helices. Thus helix B stops three residues earlier, helix C starts two residues later and stops two residues earlier, and helix D starts one residue later in the secondary structure determined by Redfield et al. (1991). These minor differences can be attributed to the fact that Redfield et al. (1991) did not have either the C$^{\alpha}$H(i)–C$^{\beta}$H(i+3) NOEs or the ^{13}C$^{\alpha}$ and ^{13}C$^{\beta}$ secondary chemical shifts available to them, which permit the start and end points of the helices to be determined more reliably.

Comparison of the Secondary Structure with That Predicted by Curtis et al. (1991). Recently, Curtis et al. (1991) presented a prediction of a consensus secondary structure of human and mouse IL-4 based on sequence alignment of IL-4 from the two species, secondary structure prediction from a combined use of the Chou and Fasman (1978), Garnier et al. (1978), and Cohen et al. (1986) methods, helical wheel projections to assess helix length based on preserving amphiphatic character, and the location of exon boundaries to suggest corresponding helix boundaries. A comparison of their prediction with our experimental findings is also presented in Table I. There is excellent agreement between the experiment and prediction for helices A, B, and D, with only minor differences of 1–3 residues in either the start or end of these helices. In the case of helix C, on the other hand, a much smaller helix (10 residues) located approximately in the middle of the experimentally determined helix (23 residues) was predicted. In addition, another helix extending from residues 26–39 was initially predicted but excluded on packing grounds. It is interesting to note that the start of this helix coincides with the small helical turn observed experimentally from residues 27 to 29. No prediction of the existence of the mini-antiparallel β-sheet was made. This is hardly surprising given the intrinsic limitations of the prediction methods employed. Nevertheless, the excellent agreement for three out of the four helices, and the partial agreement for the fourth helix, suggests that for highly helical proteins the prediction strategy employed can yield surprisingly good results.

Table I: Comparison of the Secondary Structure of IL-4 Determined in the Present Study with the Secondary Structure Derived by Redfield et al. (1991) from 3D ^{15}N-Edited NOESY Spectroscopy and That Predicted by Curtis et al. (1991)[a]

secondary structure	present work	Redfield et al. (1991)	Curtis et al. (1991)
helix			
α_A	9–21	9–21	8–21
α_B	45–64	45–61	47–62
α_C	74–96	76–94	81–90
α_D	113–129	114–129	113–126
helical turns			
ht_A	27–29		26–39 (helix)
ht_B	67–70		
β-strands			
β_A	32–34	32–34	
β_B	110–112	110–112	

[a] The numbering scheme is that of the present paper which includes the four-reside sequence Glu-Ala-Glu-Ala at the N-terminus which is not part of the natural human IL-4. The natural IL-4 sequence therefore starts at residue 5.

Topology and Global Fold of IL-4. Given the presence of four long helices, as well as the spatial and topological restrictions imposed by the three disulfide bridges between Cys-7 and Cys-131, Cys-28 and Cys-69, and Cys-50 and Cys-103 (Carr et al., 1990; Curtis et al., 1991), only a relatively small number of long-range NOEs are required to determine the relative orientation of the helices with respect to one another and to define the global fold of IL-4. Due to the extensive overlap and chemical shift degeneracy present in the aliphatic region of the ^{1}H spectrum, only three long-range NOEs involving NH protons could be identified unambiguously at this stage in the 3D ^{15}N-edited NOESY spectrum. These NOEs occurred between residues of the two strands, β_A and β_B (see Figure 6). The 3D ^{13}C-edited NOESY spectrum, on the other hand, is easier to interpret owing to its symmetric properties. Namely, an NOE between protons x and y will appear in two planes at the ^{13}C chemical shifts of the directly attached carbon atoms. By only using peaks which have a symmetry-related partner, each NOE can therefore potentially be specified by four chemical shift coordinates. By this means, we were able to assign about 150 long-range NOEs between aliphatic protons and between aromatic and aliphatic protons. Further, we were able to identify NOEs between all pairs of helices, which was sufficient to unambiguously specify their relative directionality.

For helices A and B, NOEs were observed between the residue pairs L18–L56; for helices A and C, NOEs were observed between the residue pairs T10–N93, I14–I90, I14–L94,

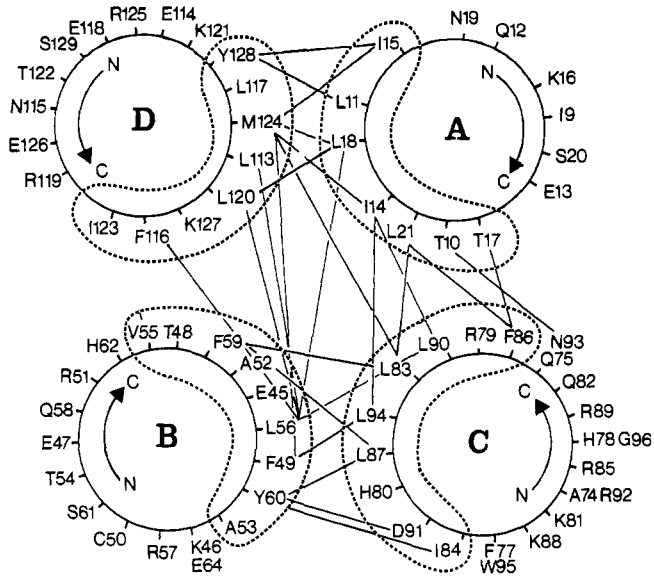

FIGURE 7: Summary of the interhelical NOE contacts derived from a preliminary analysis of the 3D ^{13}C-edited NOESY spectrum of IL-4 illustrating the left-handed four-helix bundle in a helical wheel representation. The hydrophobic faces are enclosed by the dotted lines, and the pairs of residues for which NOEs could be assigned are connected by straight lines. Helices A and B are viewed from their N-termini, while helices C and D are viewed from their C-termini. Adjacent helices are antiparallel, while diagonally related helices (i.e., A and B, C and D) run parallel to each other.

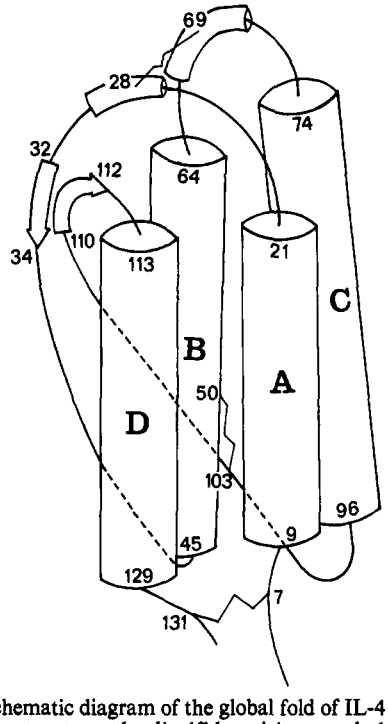

FIGURE 8: Schematic diagram of the global fold of IL-4 deduced from the secondary structure, the disulfide pairings, and about 150 long-range NOEs observed in the 3D ^{13}C-edited NOESY spectrum. There are two overhand connections between helices A and B and between helices C and D. The disulfide pairings are also indicated.

T17–F86, L21–F86, and L21–L83; and for helices A and D, NOEs were observed between the residue pairs L11–Y128, I14–M124, I15–Y128, I15–M124, L18–L120, and L18–M124. Thus, helix A must be antiparallel to helices C and D and parallel to helix B. For helices B and C, NOEs were observed between the residue pairs F49–L94, L56–L90, F59–L83, F59–L87, Y60–I84, Y60–L87, and Y60–D91; and for helices B and D, NOEs were observed between the residue pairs F49–M124, L56–L113, L56–F116, and L56–L120. Hence, helix B must be antiparallel to helices C and D. Finally, for helices C and D, NOEs were observed between the residue pairs L83–M124, indicating that these two helices are parallel.

The NOE data therefore clearly indicate that the helices are arranged in an antiparallel four-helix bundle. The handedness of the bundle is readily ascertained by visualizing the interhelical NOE interactions diagrammatically in the form of a helical wheel representation, as depicted in Figure 7. In this view, helices A and B are viewed from their N-termini, while helices C and D are viewed from their C-termini. The amphiphatic nature of the helices is clearly evident with the hydrophobic residues directed toward the interior of the bundle and the hydrophilic residues on the exterior. The handedness is unambiguously defined by the NOEs observed between pairs of residues at the outer edges of the hydrophobic surfaces. Specifically, the NOEs from Leu-21, Thr-17, and Thr-10 of helix A to Phe-86, Phe-86, and Asn-93 of helix C, respectively, from Ile-84, Asp-91, and Leu-87 of helix C to Tyr-60 of helix B, from Leu-56 of helix B to Phe-116 of helix D, and finally from Tyr-128 and Met-124 of helix D to Ile-15 of helix A are only consistent with a left-handed bundle. If the bundle were right-handed, on the other hand, a quite different set of NOEs would be observed between helices A and C and between helices D and B, as can be easily ascertained by simply placing the helical wheels of helices A and D below those of helices C and B, instead of above as shown in Figure 7.

The handedness of the helical bundle and the relative arrangement of the four helices deduced from the data in Figure 7 readily permit the determination of the global fold, which is illustrated schematically in Figure 8. The four long helices are arranged as a left-handed four-helix bundle with two overhand connections. Helix A is adjacent to helix D, which is located next to helix B, followed by helix C. The first long loop (residues 22–44) connects the diagonally located helices A and B and represents an overhand connection. This loop contains both a helical turn (ht_A) centered around Cys-28 and the three-residue strand β_A (residues 32–34). The second loop (residues 65–73) connects the adjacent helices B and C and contains the second helical turn ht_B (residues 67–70). The two helical turns are linked together in an antiparallel fashion by a disulfide bridge between Cys-28 and Cys-69. Helix C is connected to helix D in an overhand fashion by the third long loop (residues 97–112). The third loop is anchored to helix B by the disulfide bridge between Cys-103 and Cys-50 and to the first loop by the mini-antiparallel β-sheet between strands β_B (residues 110–112) and β_A (residues 30–32). From the helical wheels shown in Figure 7, one can deduce that Cys-50 is located at the back of helix B in the view depicted in Figure 8. Consequently, the path of both overhand connections must go behind helices B and C. Finally, the N- and C-termini of the protein are linked by a disulfide bridge between Cys-7 and Cys-131.

The global fold and helix packing shown in Figure 8 are the mirror image of the right-handed four-helix bundle predicted by Curtis et al. (1991). In addition, as the experimentally determined helix C is considerably longer than the predicted one (cf. Table I), it is involved in many more interactions with helix B. The topology and global fold shown in Figure 8 appears to be similar to that of the preliminary structures presented by Redfield et al. (1991), although this is very difficult to ascertain from the monoviews of the C^{α} traces shown by these authors. Further, the packing of the helices is different. In particular, the preliminary Redfield et al. (1991) structures have a more open conformation than that

depicted in Figure 8 and implied by the NOE data in Figure 7. The reason for this difference can be attributed to the fact that all interhelical NOEs involve side chains. As Redfield et al. (1991) did not have ^{13}C-labeled IL-4 at their disposal, they were only able to obtain a very limited number of ^{1}H side-chain assignments and had to rely on conventional 2D ^{1}H–^{1}H NOESY spectroscopy to identify these NOEs. Consequently, they could not observe interhelical NOEs between as many residue pairs as in the current work and were unable to observe any NOEs between helices A and B and between helices C and D. In contrast, we were able to identify a much larger network of interhelical NOEs involving all pairs of helices from the 3D ^{13}C-edited NOESY spectrum which place much more stringent restrictions on the helix packing (cf. Figure 7).

Concluding Remarks. In this paper we have presented the determination of the secondary structure and folding topology of IL-4 based on a qualitative analysis of NOE data derived from 3D ^{15}N- and ^{13}C-edited NOESY spectra, secondary $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts, $^{3}J_{HN\alpha}$ coupling constants, and NH exchange data. The combined use of NOE data and secondary $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts permits a more precise delineation of the helices than would otherwise have been possible. In addition, we were able to determine the arrangement of the helices, and hence the topology and global fold of IL-4, by simple model building considerations on the basis of interhelical NOEs involving aliphatic and aromatic side chains which could be readily identified in the 3D ^{13}C-edited NOESY spectrum. Because of extensive chemical shift overlap, it is not warranted to go beyond a schematic representation of the fold using the current data, and a high-resolution structure determination will require a detailed analysis of 4D $^{15}N/^{13}C$ and ^{13}C-edited NOESY spectra (Kay et al., 1990, Clore et al., 1991a,b; Clore & Gronenborn, 1991). These studies are currently in progress.

References

Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990) *J. Magn. Reson. 86*, 304–318.

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

Carr, C., Aykent, S., Kimack, N. M., & Levine, A. D. (1990) *Biochemistry 30*, 1515–1523.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275–288.

Clore, G. M., & Gronenborn, A. M. (1991) *Science 252*, 1390–1399.

Clore, G. M., Kay, L. E., Bax, A., & Gronenborn, A. M. (1991a) *Biochemistry 30*, 12–18.

Clore, G. M., Wingfield, P. T., & Gronenborn, A. M. (1991b) *Biochemistry 30*, 2315–2323.

Cohen, F. E., Abarbanel, R. M., Kuntz, I. D., & Fletterick, R. J. (1986) *Biochemistry 25*, 266–275.

Curtis, B. M., Presnell, S. R., Srinivasan, S., Sassenfeld, H., Klinke, R., Jeffery, E., Cosman, D., March, C. J., & Cohen, F. E. (1991) *Proteins: Struct., Funct., Genet. 11*, 111–119.

Garnier, J., Osguthorp, D. J., & Robson, B. (1978) *J. Mol. Biol.* 97–120.

Golumbek, P. T., Lazenby, A. J., Levitzky, H. I., Jaffee, L. M., Karasuyama, H., Baker, M., & Pardoll, D. M. (1991) *Science 254*, 713–716.

Driscoll, P. C., Gronenborn, A. M., Wingfield, P. T., & Clore, G. M. (1990a) *Biochemistry 29*, 4668–4682.

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

Finkelman, F. D. Holmes, J., Katonn, I. M., Urban, J. F., Beckmann, M. P., Peik, L. S., Hooley, K. A., Coffman, R. L., Momman, T. R., & Paul, W. E. (1990) *Annu. Rev. Immunol. 8*, 303–333.

Forman-Kay, J. D., Gronenborn, A. M., Kay, L. E., Wingfield, P. T., & Clore, G. M. (1990) *Biochemistry 29*, 1566–1572.

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

Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990a) *J. Magn. Reson. 86*, 204–209.

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

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

Kay, L. E., Marion, D., & Bax, A. (1989) *J. Magn. Reson. 84*, 72–84.

Kay, L. E., Clore, G. M., Bax, A., & Gronenborn, A. M. (1990) *Science 249*, 411–414.

Kühn, R., Rajewsky, K., & Müller, W. (1991) *Science 254*, 707–710.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

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

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

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

Neuhaus, D., Wagner, G., Vasak, M., Kägi, J. R. H., & Wüthrich, K. (1986) *Eur. J. Biochem. 151*, 257–273.

Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N., & Campbell, I. D. (1990) *J. Magn. Reson. 87*, 488–510.

Pardi, A., Billetter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741–751.

Paul, W. E., & Ohara, J. (1987) *Annu. Rev. Immunol. 5*, 429–460.

Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., & Clore, G. M. (1992) *Biochemistry* (preceding paper in this issue).

Presta, L. G., & Rose, G. D. (1988) *Science 240*, 1632–1641.

Redfield, C., Smith, L. J., Boyd, J., Lawrence, G. M. P., Edwards, R. G., Smith, R. A. G., & Dobson, C. M. (1991) *Biochemistry 30*, 11029–11033.

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

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

Wüthrich, K. (1989) *Acc. Chem. Res. 22*, 36–44.

Yokota, T., Arai, N., deVries, J., Spits, H., Banchereau, J., Zlotnik, A., Rennick, D., Howard, M., Takebe, Y., Miyatake, S., Lee, F., & Arai, K. (1988) *Immunol. Rev. 103*, 137–187.