

# Heteronuclear ( $^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$ ) NMR Assignments and Solution Structure of the Monocyte Chemoattractant Protein-1 (MCP-1) Dimer<sup>†,‡</sup>

Tracy M. Handel<sup>\*,§,||</sup> and Peter J. Domaille<sup>\*,||</sup>

Department of Molecular and Cell Biology, University of California at Berkeley, Berkeley, California 94720, and Chemical and Physical Sciences, The DuPont Merck Pharmaceutical Company, Wilmington, Delaware 19880-0328

Received January 30, 1996; Revised Manuscript Received March 18, 1996<sup>®</sup>

**ABSTRACT:** A full high-resolution three-dimensional solution structure of the monocyte chemoattractant protein-1 (MCP-1 or MCAF) homodimer has been determined by heteronuclear multidimensional NMR. MCP-1 is a member of a family of small proteins which play a crucial role in immune surveillance by orchestrating the recruitment of specific leukocytes to areas of immune challenge. The protein was uniformly isotopically enriched with  $^{13}\text{C}$  and  $^{15}\text{N}$  by expression in *Escherichia coli*, and complete sequence-specific resonance assignments were obtained by a combination of heteronuclear double- and triple-resonance experiments. The secondary structure was deduced from characteristic patterns of NOEs,  $^{13}\text{C}_{\alpha/\beta}$  chemical shifts, measurements of  $^3J_{\text{HNH}\alpha}$  scalar couplings, and patterns of slowly exchanging amide protons. Because MCP-1 forms symmetrical homodimers, additional experiments were carried out to unambiguously establish the quaternary contacts. NOEs from these novel experiments were merged with more traditional heteronuclear separated NOE measurements in an iterative strategy to partition the restraints between explicit inter/intrasubunit contacts and a class wherein both were retained as ambiguous. With more than 30 restraints per residue, the three-dimensional structure is well-defined with a backbone rmsd of 0.37 Å to the mean over residues 5–69 of the dimer. We compare the structure with those recently reported for the related chemokines MIP-1 $\beta$  and RANTES and highlight the differences in terms of receptor specificity and function as well as interpret the known biological activity data of MCP-1 mutants.

In the last decade several proteins have been identified which possess chemoattractant and activating potential for specific types of leukocytes (Miller & Krangel, 1992; Oppenheim et al., 1991; Schall & Bacon, 1994). Because of their role in cellular recruitment, these proteins (chemokines, chemoattractant and activating cytokines) are intimately involved in coordinating normal inflammatory processes such as tissue remodeling and wound repair as well as those associated with serious pathologic states including rheumatoid arthritis, lung inflammations, and atherosclerosis. The specificity of cytokines for specific cell types contrasts with the more indiscriminate activity of chemoattractants such as C5a,<sup>1</sup> fMLF, and leukatriene B. Consequently, they have superseded these classical molecules as targets for therapeutic intervention and provide a fertile area for more judicious structure-based design of receptor antagonists.

With few exceptions, all chemokines have four completely conserved cysteine residues which form two disulfide bonds within these small (8–10 kDa) proteins. Two major subsets of chemokines have been classified ( $\alpha$  and  $\beta$ ) on the basis of the chromosomal location of their genes and the relative position of the first two cysteine residues in the proteins. The genes for  $\alpha$ -chemokines are tightly clustered on chromosome 4 and encode proteins which have an intervening residue between the first two cysteines (C-X-C), whereas

members of the  $\beta$ -family are derived from chromosome 17, and the first two cysteines are immediately adjacent (C-C). To a first approximation, these sequence differences are paralleled by well-defined functional differences;  $\alpha$ -chemokines such as IL-8 and MGSA/GRO are chemotactic and activating for neutrophils, while most  $\beta$ -chemokines (e.g., MCP-1, RANTES, MCP-3, MIP-1 $\alpha$  and -1 $\beta$ ) are monocyte chemoattractants (Leonard & Yoshimura, 1990; Schall et al., 1990). With the exception of the promiscuous erythrocyte chemokine receptor (Neote et al., 1993a, 1994), proteins within each family but not between families can competitively bind to the same receptor on target cells (Leonard & Yoshimura, 1990; Murphy, 1994; Neote et al., 1993b; Yoshimura & Leonard, 1990). Beyond this, however, the similarity of proteins within each subset is less obvious because all have chemotactic potential for additional diverse cell types. For example, MCP-1, RANTES, and MIP-1 $\beta$  are chemotactic for monocytes but function through two

<sup>1</sup> Abbreviations: MCP-1, monocyte chemoattractant protein-1; chemokine, chemoattractant cytokine; C5a, complement factor 5a; fMLF, formylated methionylleucylphenylalanine; PMSF, phenylmethanesulfonyl fluoride; EDTA, sodium salt of ethylenediaminetetraacetic acid; IPTG, isopropyl  $\beta$ -D-thiogalactopyranoside; TFA, trifluoroacetic acid; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; ROESY, rotating frame nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; CT, constant time; 2D, 3D, and 4D, two, three, and four dimensional; TPPI, time-proportional phase incrementation; SA, simulated annealing;  $T_2$ , transverse relaxation time; rmsd, root mean square deviation; X-PLOR, computer program for molecular dynamics and structure calculations; ANSIG3.0, interactive graphics program for assignment of spectra; NCS, noncrystallographic symmetry restraint term in X-PLOR.

<sup>†</sup> This work was supported by NIH Grant AI 37113 to T.M.H.

<sup>‡</sup> The coordinates have been deposited with the Brookhaven Protein Data Bank. 1DOM is the minimized average structure and 1DON is an ensemble of 20 structures.

<sup>\*</sup> To whom correspondence should be addressed.

<sup>§</sup> University of California at Berkeley.

<sup>||</sup> The DuPont Merck Pharmaceutical Co.

<sup>®</sup> Abstract published in *Advance ACS Abstracts*, May 1, 1996.

separate receptors (Sozzani et al., 1993). MCP-1 and RANTES also both target memory T-cells (Carr et al., 1994; Loetscher et al., 1994; Maghazachi et al., 1994; Schall et al., 1990), while MIP-1 $\beta$  acts preferentially on naive T-cells (Adams et al., 1994). Moreover, of these three chemokines, only RANTES chemotaxes and activates eosinophils (Ebisaawa et al., 1994; Kameyoshi et al., 1994). Understanding the origin of these differences in cellular specificity is a necessary step in developing receptor antagonists that are selective and therapeutically effective. Because structural information can highlight differences between these proteins and provide insight into the basis for their cellular selectivity, we undertook the determination of the solution structure of the MCP-1 dimer using NMR spectroscopy. We chose to focus on this protein because it appears to be the most potent macrophage chemoattractant and is thought to play a major role in the obligatory recruitment of monocytes into arterial lesions during the advancement of atherosclerosis (Li et al., 1993; Nelken et al., 1991; Takeya et al., 1993; Yu et al., 1992).

A complication in determining the three-dimensional solution structure of MCP-1 is that it is dimeric at concentrations above 100  $\mu$ M (Paolini et al., 1994). In most NMR studies of other homodimers, intersubunit contacts have been identified on the basis of the X-ray or NMR structure of homologous proteins or on NOEs which are inconsistent with the monomer structure and lead to violations during structure calculations (Breg et al., 1990; Clore et al., 1990; Fairbrother et al., 1994; Lodi et al., 1994; Skelton et al., 1995). While these are certainly viable strategies, they are not unequivocal, especially for proteins with extensive interfacial contacts which can be nearly satisfied in more than one way. Small errors or incorrect treatment of NOEs at the interface of a dimeric protein can lead to inaccuracies in the overall quaternary structure by incorrect orientation of the two monomers. Recently, X-filtered experiments on 1:1 complexes of asymmetrically labeled proteins have been used to break the magnetic symmetry, allowing direct determination of intersubunit NOEs (Burgering et al., 1994; Clore et al., 1994; Folkers et al., 1994; Lee et al., 1994). The most common approach has been to detect NOEs between  $^1\text{H}$ – $^{12}\text{C}$  and  $^1\text{H}$ – $^{13}\text{C}$  pairs on samples containing mixtures of unlabeled and  $^{13}\text{C}/^{15}\text{N}$ -labeled protein. We have used a unique variation of these experiments to identify the intermonomer contacts in MCP-1. In our method we generated a sample containing a 50:50 mixture of  $^{15}\text{N}$ -labeled and  $^{13}\text{C}$ -labeled MCP-1 and recorded experiments which selectively detect *only* intermonomer NOEs between  $^{15}\text{N}$ – $^1\text{H}$  and  $^{13}\text{C}$ – $^1\text{H}$  pairs. An advantage of this strategy is that the magnetization is filtered through two different magnetically active nuclei, thereby affording excellent suppression of spurious intramonomer NOEs.

While this work was in progress, NMR structural studies of both MIP-1 $\beta$  (Lodi et al., 1994) and RANTES (Chung et al., 1995; Skelton et al., 1995) were published. However, unlike these studies which required extremes of pH to prevent aggregation, our studies of MCP-1 are done near physiological pH and potentially yield different information, particularly if charged interactions are important determinants of receptor specificity. Additionally, although high-resolution X-ray diffraction data from crystals of MCP-1 have been available for some time (A. Wlodawer, private communication), the structure has not been solved, suggesting structural differ-

ences that exceed the conformational search range of molecular replacement. Herein we present the resonance assignments and the high-resolution solution structure of MCP-1 and compare it with those of RANTES and MIP-1 $\beta$ . In addition, we describe our unique mixed isotope experiments used to directly establish the quaternary contacts in MCP-1 and our iterative calculation strategy for determining the structure of homodimers.

## MATERIALS AND METHODS

**Protein Overexpression and Isolation.** Unlabeled protein was generated in a 30° 10 L fermentation growth in M9 minimal media containing [ $^{12}\text{C}_6$ ]glucose and ( $^{14}\text{NH}_4$ ) $_2\text{SO}_4$  as the sole sources of nitrogen and carbon.  $^{15}\text{N}$ -Labeled,  $^{13}\text{C}$ -labeled, and  $^{15}\text{N}/^{13}\text{C}$ -labeled MCP-1 samples were grown in minimal media using ( $^{15}\text{NH}_4$ ) $_2\text{SO}_4$  or [ $^{13}\text{C}_6$ ]glucose/( $^{15}\text{NH}_4$ ) $_2\text{SO}_4$  (ISOTEC, Miamisburg, OH) in a similar manner except on a smaller scale (6–8 L). Typically, 30 g of cell paste was mildly sonicated in 1 L of buffer (10 mM sodium phosphate, pH 7.0, 5 mM EDTA, 1 mM PMSF) and centrifuged at 6000 rpm for 45 min to remove cell debris. The supernatant was then loaded onto a 200 mL Sepharose S Fast Flow column, washed with 2 $\times$  volume of sonication buffer, and eluted with 10 mM sodium phosphate/5 mM EDTA containing a 0–1.0 M gradient of NaCl. The protein elutes at approximately 750 mM NaCl in highly (90%) purified form. Fractions containing the desired protein were identified by injecting aliquots from fractions within the elution range onto an HP1090 HPLC equipped with an analytical C18 column and autosampler. Suitable protein fractions were then pooled, further purified on a semiprep C18 reverse-phase HPLC column with a gradient of 24.9/75/0.1 to 49.9/50/0.1 H $_2\text{O}$ /acetonitrile/TFA, and lyophilized.

Samples of MCP-1, 2 mM (monomer) concentration were prepared by rehydrating the lyophilized powder with either 0.5 mL of 90% H $_2\text{O}$ /10% D $_2\text{O}$  or 100% D $_2\text{O}$  (ISOTEC) and adjusting the pH to 5.42 with NaOD. For the 1 mM sample containing 50%  $^{15}\text{N}$  MCP-1 and 50%  $^{13}\text{C}$  MCP-1, equivalent quantities of protein (measured by OD 280) were dissolved in the HPLC elution buffer, mixed together, and lyophilized. In HPLC buffer, MCP-1 is monomeric and partially unfolded, thus promoting homogeneous mixing of the two isotopically labeled proteins.

**NMR Spectroscopy.** All NMR experiments were carried out under identical conditions using  $\sim$ 2 mM samples of MCP-1 at 35 °C on a Bruker AMX 600 spectrometer equipped with an external fourth channel and class A' amplifiers. Water suppression was achieved using presaturation and/or scrambling pulses (Messerle et al., 1989) with the  $^1\text{H}$  carrier positioned on water. The  $^1\text{H}$  spectral width in the directly detected dimension was 8064.52 Hz for all experiments with the exception of the 3D HCCH-TOCSY in which it was reduced to 6250.0 Hz. The  $^{15}\text{N}$  spectral width in all experiments was 800.0 Hz with aliasing optimized to reduce the number of time increments necessary to achieve the desired digital resolution in 3D and 4D spectra. Quadrature detection in indirect dimensions was achieved by States–TPPI except for certain  $^{13}\text{C}$  dimensions where axials are better placed at the center of the spectrum rather than at the edge. Additional details of the data collection for the following experiments are summarized in Table 1.

Table 1: NMR Data Acquisition Parameters for All Experiments

experiment	dims <sup>b</sup>	spectral width (Hz)	carrier (ppm)	acq time (ms)	complex points	mix time (ms)	transients	duration (h)	references
2D <sup>15</sup> N- <sup>1</sup> H HSQC	<sup>15</sup> N- <i>t</i> <sub>1</sub>	800.0	117.8	160	128		16		Bax et al., 1990;
	<sup>1</sup> H- <i>t</i> <sub>2</sub>	8064.52	4.72	63.5	512				Norwood et al., 1990
2D CT <sup>13</sup> C- <sup>1</sup> H HSQC	<sup>13</sup> C- <i>t</i> <sub>1</sub>	12500.0	46.0	26.0	325		32	6.3	Vuister & Bax, 1992
	<sup>1</sup> H- <i>t</i> <sub>2</sub>	8064.52	4.72	63.5	512				
2D CT <sup>13</sup> C- <sup>1</sup> H HSQC	<sup>13</sup> C- <i>t</i> <sub>1</sub>	5000.0	126.0	18	90		32	2.0	Vuister & Bax, 1992
	<sup>1</sup> H- <i>t</i> <sub>2</sub>	8064.52	4.72	63.5	512				
2D CγN <sup>13</sup> C- <sup>1</sup> H HSQC	<sup>13</sup> C- <i>t</i> <sub>1</sub>	12500.0	46.0	26.8	670		16	7.5	Grzesiek et al., 1993
2D CγC' <sup>13</sup> C- <sup>1</sup> H HSQC	<sup>1</sup> H- <i>t</i> <sub>2</sub>	8064.52	4.72	63.5	512				
2D CBHD	<sup>13</sup> C- <i>t</i> <sub>1</sub>	4000.0	35.0	9	36		1024	24	Yamazaki et al., 1993
2D CBHE	<sup>1</sup> H- <i>t</i> <sub>2</sub>	8064.52	4.72	63.5	512				
2D (HC)NH NOESY <sup>c</sup>	<sup>15</sup> N- <i>t</i> <sub>1</sub>	800.0	117.8	40.0	32	100	2048	44	Kay et al., 1990
	<sup>1</sup> H- <i>t</i> <sub>2</sub>	8064.52	4.72	63.5	512				
3D <sup>15</sup> N NOESY HMQC	<sup>1</sup> H- <i>t</i> <sub>1</sub>	8064.52	4.72	15.9	128	100	16	3.5	Marion et al., 1989
3D <sup>15</sup> N TOCSY HMQC	<sup>15</sup> N- <i>t</i> <sub>2</sub>	800.0	117.8	40.0	32	52, 30			
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	8064.52	4.72	63.5	512				
3D <sup>15</sup> N ROESY HMQC	<sup>1</sup> H- <i>t</i> <sub>1</sub>	6250.0	4.72	10.2	64	30	32	90.5	Powers et al., 1993
	<sup>5</sup> N- <i>t</i> <sub>2</sub>	800.0	117.8	40.0	32				
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	8064.52	4.72	63.5	512				
3D <sup>15</sup> N HNHB	<sup>1</sup> H- <i>t</i> <sub>1</sub> <sup>a</sup>	8064.52	4.72	12.4	100		32	94.8	Archer et al., 1991
	<sup>15</sup> N- <i>t</i> <sub>2</sub>	800.0	117.8	30.0	24				
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	8064.52	4.72	63.5	512				
3D <sup>15</sup> N HNHA	<sup>15</sup> N- <i>t</i> <sub>2</sub>	800.0	117.8	40.0	32		32	77.6	Vuister & Bax, 1993
	<sup>1</sup> H- <i>t</i> <sub>2</sub>	5813.95	4.72	11.0	64				
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	8064.52	4.72	63.5	512				
3D <sup>13</sup> C TOCSY HMQC	<sup>1</sup> H- <i>t</i> <sub>1</sub>	6250.0	4.72	20.5	128	17	16	82.5	Bax et al., 1990
	<sup>13</sup> C- <i>t</i> <sub>2</sub>	4000.0	45.5	8.0	32				
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	6250.0	4.72	40.0	256				
3D <sup>13</sup> C NOESY HMQC	<sup>1</sup> H- <i>t</i> <sub>1</sub>	8064.52	4.72	15.9	128	80	16	87.8	Zuiderweg et al., 1990
	<sup>13</sup> C- <i>t</i> <sub>2</sub> <sup>a</sup>	4000.0	45.5	8.0	32				
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	8064.52	4.72	63.5	512				
3D HC(N)H NOESY	<sup>1</sup> H- <i>t</i> <sub>1</sub>	6250.0	4.72	12.8	80	100	64	156	Kay et al., 1990
	<sup>13</sup> C- <i>t</i> <sub>2</sub> <sup>a</sup>	4000.0	45.5	6.0	24				
	<sup>1</sup> H- <i>t</i> <sub>3</sub>	8064.52	4.72	126	512				
4D <sup>13</sup> C NOESY HMQC	<sup>1</sup> H- <i>t</i> <sub>1</sub>	8064.52	4.72	7.9	64	100	8	137.0	Clare et al., 1991;
	<sup>13</sup> C- <i>t</i> <sub>2</sub> <sup>a</sup>	4000.0	45.5	2.5	10				Zuiderweg et al., 1991
	<sup>13</sup> C- <i>t</i> <sub>3</sub> <sup>a</sup>	4000.0	45.5	2.5	10				
	<sup>1</sup> H- <i>t</i> <sub>4</sub>	8064.52	4.72	63.5	512				
4D HNCAHA	<sup>1</sup> H- <i>t</i> <sub>1</sub>	1666.7	4.72	14.4	24		32	52.2	Boucher et al., 1992a,b
	<sup>13</sup> C- <i>t</i> <sub>2</sub>	1666.7	58.0	4.8	8				
	<sup>15</sup> N- <i>t</i> <sub>3</sub>	800.0	117.8	6.3	5				
	<sup>1</sup> H- <i>t</i> <sub>4</sub>	8064.52	4.72	63.5	512				
4D HCA(CO)NNH	<sup>1</sup> H- <i>t</i> <sub>1</sub>	1000.0	4.72	18.0	18		16	92.0	Boucher et al., 1992a,b
	<sup>13</sup> C- <i>t</i> <sub>2</sub>	2000.0	58.0	6.0	12				
	<sup>15</sup> N- <i>t</i> <sub>3</sub>	800.0	117.8	12.5	10				
	<sup>1</sup> H- <i>t</i> <sub>4</sub>	8064.52	4.72	63.5	512				
4D HCC(CO)NNH	<sup>1</sup> H- <i>t</i> <sub>1</sub>	5000.0	4.72	3.2	16		16	107.0	Clowes et al., 1993
	<sup>13</sup> C- <i>t</i> <sub>2</sub> <sup>a</sup>	4000.0	45.5	2.0	8				
	<sup>15</sup> N- <i>t</i> <sub>3</sub>	800.0	117.8	25.0	20				
	<sup>1</sup> H- <i>t</i> <sub>4</sub>	8064.52	4.72	63.5	512				
4D HCNH NOESY	<sup>1</sup> H- <i>t</i> <sub>1</sub>	5000.0	4.72	6.4	32	100	8	114.7	Kay et al., 1990
	<sup>13</sup> C- <i>t</i> <sub>2</sub>	4000.0	45.5	5.0	20				
	<sup>15</sup> N- <i>t</i> <sub>3</sub>	800.0	117.8	10.0	8				
	<sup>1</sup> H- <i>t</i> <sub>4</sub>	8064.52	4.72	63.5	512				

<sup>a</sup> Quadrature detection achieved with States rather than States-TPPI. <sup>b</sup> Dimension. <sup>c</sup> <sup>1</sup>H-<sup>1</sup>H and <sup>1</sup>H-<sup>13</sup>C 2D versions of this experiment were also collected as for the 4D HCNH.

Assignment of backbone nuclei (<sup>15</sup>N, H<sub>N</sub>, <sup>13</sup>C<sub>α</sub>, H<sub>α</sub>) was based on the 4D CT HCA(CO)NNH and the 4D CT HNCAHA experiments (Boucher et al., 1992a,b) acquired on a uniformly labeled <sup>15</sup>N/<sup>13</sup>C sample of MCP-1 in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. <sup>13</sup>C and <sup>1</sup>H side chain assignments were obtained via a 4D HCC(CO)NNH experiment (Clowes et al., 1993) on the same sample. Because this experiment suffers from limited digital resolution, a 3D HCCH-TOCSY (<sup>15</sup>N/<sup>13</sup>C MCP-1 in D<sub>2</sub>O) (Bax et al., 1990a) was also recorded to provide higher precision of the indirectly detected <sup>1</sup>H and <sup>13</sup>C chemical shifts and to measure any side chains missing in the HCC(CO)NNH experiment. 3D <sup>15</sup>N-separated TOCSY HMQC, 3D HNHB, and 3D HNHA experiments were collected on a sample of <sup>15</sup>N-labeled MCP-1 (Archer

et al., 1991; Marion et al., 1989; Vuister & Bax, 1993) and provided complementary assignment information and higher precision of <sup>1</sup>H chemical shifts.

NOESY experiments included a 3D <sup>15</sup>N-separated NOESY HMQC (Marion et al., 1989), a 4D <sup>13</sup>C/<sup>15</sup>N-separated NOESY (Kay et al., 1990), 3D and 4D <sup>13</sup>C-separated NOESY HMQC (Clare et al., 1991; Zuiderweg et al., 1990, 1991), a 3D <sup>13</sup>C-separated ROESY, and a 3D <sup>15</sup>N-separated ROESY (Powers et al., 1993) using uniformly labeled <sup>15</sup>N/<sup>13</sup>C MCP-1 or <sup>15</sup>N MCP-1. Mixing times were 100, 100, 80, 100, 30, and 30 ms, respectively. To determine NOEs corresponding to the dimer interface, 2D and 3D versions of the 4D <sup>13</sup>C/<sup>15</sup>N-separated NOESY experiment were recorded on a sample which was 50% <sup>15</sup>N MCP-1 and 50% <sup>13</sup>C MCP-1. Lower

dimensionality experiments were recorded for increased sensitivity because the concentration of the sample was only 1 mM, and only 0.5 mM should have the desired isotope composition. The 2D versions were  $^{15}\text{N}$ -separated/ $^{13}\text{C}$ -filtered and vice versa, or  $^{15}\text{N}$ -filtered/ $^{13}\text{C}$ -filtered/ $^1\text{H}$ -separated, while the 3D version excluded detection of the indirect  $^{15}\text{N}$  dimension.

Slowly exchanging amide protons were identified by recording 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments (Bax et al., 1990b; Norwood et al., 1990) at various intervals following rehydration of lyophilized  $^{15}\text{N}$ -labeled MCP-1 in  $\text{D}_2\text{O}$  (pH 5.42).

A 3D HNHA experiment was used to obtain  $\phi$  angle restraints (Vuister & Bax, 1993). Coupling constants were calculated from the HNHA experiment by measuring the intensity ratio of cross-peaks to diagonal peaks for resolved  $^{15}\text{N}$ - $^1\text{H}$  pairs. Stereospecific assignments and  $\chi_1$  angle restraints of  $\text{H}_\beta$  atoms were determined primarily using a combination of the following experiments: 3D  $^{15}\text{N}$ -separated TOCSY HMQC (Marion et al., 1989), 3D HNHB (Archer et al., 1991), 3D  $^{15}\text{N}$ -separated ROESY, and 3D  $^{13}\text{C}$ -separated ROESY (Powers et al., 1993).  $^3J_{\text{C}_\gamma\text{N}}$  and  $^3J_{\text{C}_\gamma\text{C}'}$  were measured using  $J$ -modulated versions of a 2D CT  $^{13}\text{C}$  HSQC (Grzesiek et al., 1993; Vuister et al., 1993) to provide angular restraints for Val, Thr, and Ile residues.

**Data Processing.** Processing in the directly detected dimension was done with conventional Fourier transforms using the program Felix1.1 (Hare Research) on a Silicon Graphics R4000 Indigo. For 3D experiments, the indirectly detected dimensions were processed with a two-dimensional maximum entropy algorithm (Laue et al., 1986). 4D experiments were processed with 2D maximum entropy for the two most underdigitized dimensions and linearly extended in the third indirect dimension using the linear prediction/extension algorithm in Felix2.1 (Biosym Technologies) or simply windowed and processed with a conventional FT. Some later data were processed with the program Azara (W. Boucher, unpublished results) using related methods.

**Data Reduction.** Spectra were contoured and aligned for optimum overlap of well-resolved cross-peaks in the 3D/4D version of the interactive graphics program ANSIG3.0 (Kraulis, 1989; Kraulis et al., 1994) on a Silicon Graphics R4000 Indigo. All data were analyzed within ANSIG3.0. More than 9500 cross-peaks were assembled in the database with the accumulation of  $\sim 7200$  NOE assignments. The intensity value for each NOE was taken as the average of the corresponding cross-peak in every spectrum where the NOE appeared and classified as distances by box summation integration (very weak, 0–6 Å; weak, 0–5 Å; medium, 0–3.3 Å; and strong, 0–2.7 Å). No attempt was made to account for differences in heteronuclear coherence transfer between different experiments. Corresponding ROESY data were similarly classified as three distance classes (weak, 0–4.2 Å; medium, 0–3.3 Å; and strong 0–2.7 Å). After scaling of methyl intensities by a factor of 3 (Yip, 1990), ignoring cross-peaks reflecting trivial distances by virtue of covalent structure, and addition of pseudoatom distance corrections for nonstereo assigned groups (Wuthrich et al., 1983), the distance restraints were directly output in X-PLOR format. NOESY and ROESY data were treated as separate restraint classes in X-PLOR calculations because of their different distance sensitivities; ROESY data do not suffer from spin diffusion but are limited to a shorter distance range

Table 2: Summary of Input Restraints Used for Structure Calculation of the MCP-1 Dimer

distance restraints	
total inter- and intrasubunit NOE (ROE)	4458 (628)
total intersubunit NOE (ROE)	272 (8)
total intrasubunit NOE (ROE)	4186 (620)
$[i - j] > 4$	1356 (32)
$[i - j] = 4$	146 (4)
$[i - j] = 3$	216 (4)
$[i - j] = 2$	320 (26)
$[i - j] = 1$	1004 (254)
$[i - j] = 0$	1144 (290)
ambiguous	100 (14)
hydrogen bonds, intrasubunit	88
intersubunit	4
angular restraints	
$\chi_1$ ( $\chi_2$ )	66 (4)
$\phi$	96

and therefore are not averaged with NOESY restraints. The restraint count (Table 2) correctly reflects this separate treatment.

**Structure Calculations.** Structures were calculated using the program X-PLOR (Brunger, 1992) on an 8-processor Silicon Graphics Onyx computer using a molecular dynamics/simulated annealing protocol, beginning largely as described by Nilges (1993), and evolving through several rounds of calculations with gradual repartitioning of NOEs into separate classes. This redistribution into different classes (only intermonomer, only intramonomer, and NOEs with both intra- and intermonomer contributions) was accomplished by use of separate Perl scripts to manipulate the X-PLOR restraint files generated in ANSIG. Implementation of this protocol and the iterative calculation approach are detailed further in the Results section. Symmetry was enforced by application of X-PLOR nonstrict, noncrystallographic symmetry (NCS) and intermolecular  $\text{C}_\alpha$  distance symmetry restraints. Force constants were altered during the course of the simulated annealing exactly as described by Nilges (1993).

Structures were visually inspected in INSIGHTII (Biosym Technologies). The program PROCHECK (Laskowski et al., 1993) was used to assess the quality of the covalent geometry and classify regions of secondary structure.

## RESULTS

**Protein Expression and Isolation.** A BL21 strain of *Escherichia coli* was engineered with a thioredoxin reductase knockout mutation designed to modulate the redox potential of the cell and promote proper formation of disulfide bonds (S. Lichter, T. A. Patterson, T. M. Handel, and H. J. George, manuscript in preparation). In this cell line MCP-1 is overexpressed at approximately 5% of the total cell protein on enriched and minimal media (30 °C) in a completely soluble, folded, and homogeneous form. Purification of the properly oxidized protein is then easily achieved by a two-step chromatographic procedure. Because MCP-1 is highly basic ( $pI \sim 10.1$ ), it is strongly adsorbed on a cation-exchange (S-Sepharose) resin at pH 7 and elutes with a salt gradient in a highly purified form (90%). This was followed by reverse-phase HPLC and lyophilization. The protein shows identical NMR spectra after ion-exchange chromatography and after redissolution and pH adjustment of the lyophilized powder.

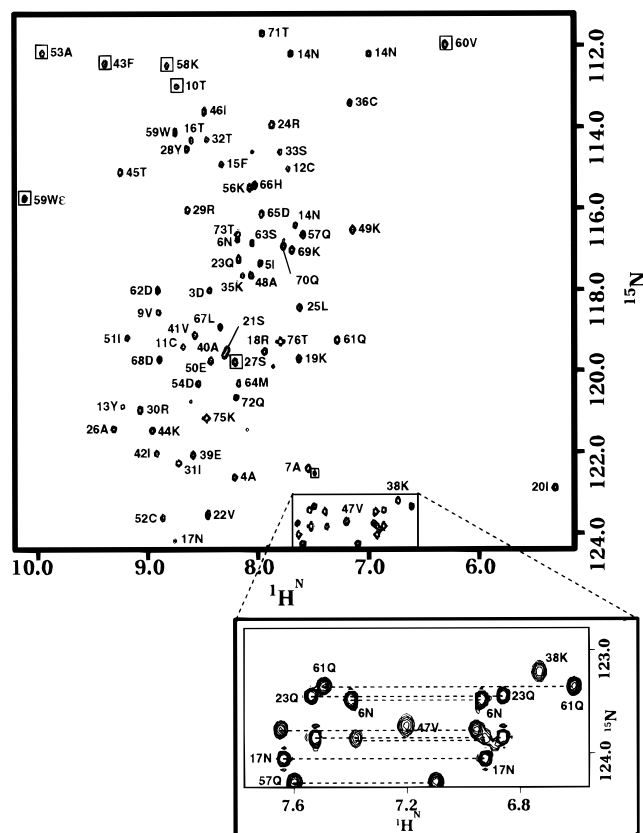


FIGURE 1: 2D  $^1\text{H}$ – $^{15}\text{N}$  HSQC of MCP-1 at 35  $^\circ\text{C}$ , pH 5.4. The spectral width in the  $^{15}\text{N}$  dimension was 800.0 Hz with aliased cross-peaks (indicated by boxes) displayed without sign discrimination. The region containing most of the  $\text{NH}_2$  resonances is enlarged.

Analytical ultracentrifugation indicates that the protein has an effective molecular mass of 17.2 kDa, consistent with a dimeric species in solution at concentrations above 100  $\mu\text{M}$  (data not shown).  $^1\text{H}_\text{N}$  line widths of 15 Hz and  $^{15}\text{N}$   $T_2$  values of ca. 65–70 ms in structured regions of the protein further support a dimer, and the number of cross-peaks in the  $^{15}\text{N}$ – $^1\text{H}$  HSQC spectrum (*vide infra*) implies that it is a symmetrical homodimer.

**Sequence-Specific NMR Assignments.** In general, the NMR spectra of MCP-1 were well dispersed and the sequential assignments straightforward with the exceptions described below. Figure 1 shows the 2D  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectrum of MCP-1 to demonstrate the extent of dispersion. Several cross-peaks, highlighted with boxes, are aliased because the spectral width was reduced to 800.0 Hz in the  $^{15}\text{N}$  dimension. All backbone amide cross-peaks were visible with the exception of Met0, Gln1, and Ser34. Recently, we identified Ser34 from a pulsed field gradient version of the HSQC and verified its identity from sequential and intraresidue cross-peaks in the  $^{15}\text{N}$ -separated ROESY which was collected without presaturation. On the basis of  $^{15}\text{N}$   $T_1$ ,  $T_2$ , and heteronuclear NOE experiments (data not shown) the C-terminal residues from Thr71 to Thr76 are significantly more mobile than the first observed N-terminal residue, Asp3; nonetheless, these residues are visible in the HSQC and are punctuated by truncation artifacts in the indirect  $^{15}\text{N}$  dimension. We also identified a second set of cross-peaks for residues Thr76, Lys75, Thr73, and possibly Gln72. These have approximately 10% of the intensity of the major cross-peaks and are presumably due to *cis*–*trans* isomerization of Pro74 in the disordered C-terminus of the protein.

The 4D HNNCAHA and HCA(CO)NNH experiments were used in combination with the  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  HSQC to determine assignments of backbone nuclei ( $^{13}\text{C}_\alpha$ ,  $^{15}\text{N}$ ,  $^1\text{H}_\alpha$ ,  $^1\text{H}_\text{N}$ ) using previously established methods (Boucher et al., 1992a; Campbell-Burk et al., 1992). The HNNCAHA provides intraresidue correlations between the ( $^1\text{H}_\alpha$ ,  $^{13}\text{C}_\alpha$ ) and ( $^{15}\text{N}$ ,  $^1\text{H}_\text{N}$ ) resonances of a given residue, as well as the interresidue correlation to the ( $^1\text{H}_\alpha$ ,  $^{13}\text{C}_\alpha$ ) chemical shifts of the preceding residue. For MCP-1, we found this experiment to be significantly more sensitive than the related net transfer HCANNH experiment, consistent with arguments given earlier (Boucher et al., 1992b) that proteins with short  $^{13}\text{C}$   $T_2$  values should have a higher sensitivity in the HNNCAHA experiment because the  $\text{C}_\alpha$  magnetization is transverse for a shorter duration than in the HCANNH experiment (6 ms versus 24.8 ms). Parenthetically, similar comparisons of these two experiments on a protein with higher molecular mass (ras p21 1–166, 19.2 kDa) gave the opposite results, confirming other indications that at NMR concentrations MCP-1 behaves like a protein with an effective hydrodynamic radius significantly larger than a 76-residue dimer. This is not surprising, however, given the elongated nature of the dimer (*vide infra*).

Due to an untimely arc in the NMR probe, only five complex points ( $t_3(\text{max}) = 6.25$  ms) were collected in the  $^{15}\text{N}$  dimension of the 4D HNNCAHA experiment, but the data were entirely usable without deleterious truncation artifacts by using maximum entropy reconstruction. Since the sequential correlations are often weak or missing in this experiment, we also recorded the 4D HCA(CO)NNH experiment which detects only the interresidue connectivities with higher sensitivity than the HNNCAHA experiment. These data were collected with a different  $^{13}\text{C}$  spectral width than the HNNCAHA so that cross-peaks aliased in the  $^{13}\text{C}$  dimension were easily identified by their different apparent  $^{13}\text{C}$  chemical shifts in the two experiments.

In concert with backbone assignments, we also measured side chain resonances using a 4D HCC(CO)NNH experiment (Clowes et al., 1993) and a 3D HCCH-TOCSY (Bax et al., 1990a). By analogy to the HCA(CO)NNH experiment, the HCC(CO)NNH experiment correlates  $^{15}\text{N}$ – $^1\text{H}_\text{N}$  resonances of a given residue with  $^1\text{H}$ – $^{13}\text{C}$  chemical shifts of the entire preceding side chain usually providing a unique amino acid spin type. Generally, we first assigned side chain resonances using the HCC(CO)NNH experiment and subsequently also assigned them in the more highly digitized 3D HCCH-TOCSY in order to increase the precision of the chemical shifts and to provide reliable statistics of chemical shift differences between different samples in  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$ .

**Assignment Strategy.** Making concurrent backbone and side chain assignments for each residue was greatly facilitated by the multiwindow graphics environment and the powerful macro language of ANSIG3.0 (Kraulis, 1989; Kraulis et al., 1994). Within each window, one can view, overlay, and correlate 2D planes or 2D slices from multiple 2D, 3D, and 4D data sets. The windows are designed with two adjustable scroll bars which control the extent of the chemical shift ranges in the unviewed dimensions and allow for real-time scrolling through all 2D planes of 3D and 4D data sets. Macros can be used to read and set cross-peak and scroll bar coordinates in specific windows and thereby automate viewing of specific planes in 3D and 4D data sets. These and many other features of the program, such as the ability

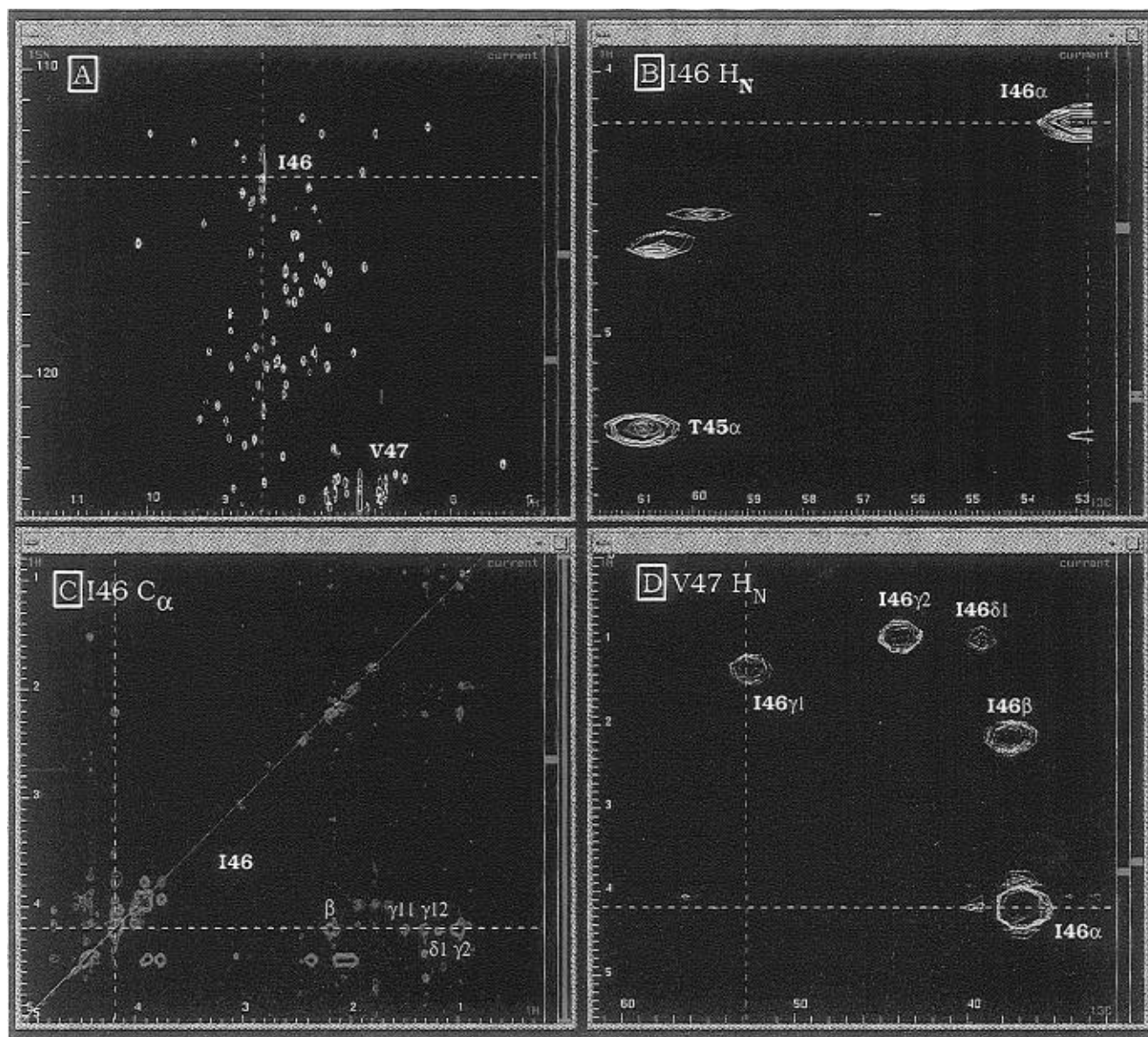


FIGURE 2: ANSIG sequential assignment strategy. The 2D  $^{15}\text{N}$  HSQC (white contours) and a plane from the 4D HNCAHA experiment (yellow contours) at the  $^{13}\text{C}_\alpha$ – $^1\text{H}_\alpha$  chemical shifts of Ile46 are displayed in panel A. The  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  coordinates of Ile46 are used to automatically set the scroll bars of panel B for display of the  $^1\text{H}$ – $^{13}\text{C}$  plane from the HNCAHA experiment which contains the intrasidue cross-peak of Ile46  $\{^1\text{H}_\text{N}(\text{I46}), ^{15}\text{N}(\text{I46}), ^1\text{H}_\alpha(\text{I46}), ^{13}\text{C}_\alpha(\text{I46})\}$  and the correlation to the preceding residue  $\{^1\text{H}_\text{N}(\text{I46}), ^{15}\text{N}(\text{I46}), ^1\text{H}_\alpha(\text{T45}), ^{13}\text{C}_\alpha(\text{T45})\}$ . Also shown in this window is the HCA(CO)NNH experiment (red contours) which allows definitive identification of the interresidue cross-peak. The  $^{13}\text{C}_\alpha$  chemical shift of Ile46 is then used to set the scroll bar of the 3D  $^{13}\text{C}$ -separated TOCSY in panel C to begin assignment of Ile46 side chain  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts. This is done in parallel with assignment of the same side chain via the 4D HCC(CO)NNH experiment in panel D which is displayed by automatically setting the scroll bars to the  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  chemical shift of Val47 from panel A.

to readily account for spectral aliasing, were used to make virtually complete residue assignments of MCP-1 and construct a multinuclear chemical shift database, starting at the C-terminus and working toward the N-terminus.

Operationally (Figure 2), we started with a window (A) containing a  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  plane from the 4D HNCAHA experiment containing a cross-peak for a given residue ( $i$ ) and overlaid it with the 2D  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  HSQC to visually check the peak center. Under macro control, the  $^1\text{H}$ – $^{15}\text{N}$  coordinates of the cross-peak were then used to automatically restrict the scroll bars of a second window (B) and display the orthogonal 2D  $^1\text{H}_\alpha$ – $^{13}\text{C}_\alpha$  plane of the HNCAHA experiment containing  $\text{H}_\alpha(i)$ – $\text{C}_\alpha(i)$  and  $\text{H}_\alpha(i-1)$ – $\text{C}_\alpha(i-1)$  cross-peaks. Generally the intrasidue  $^1\text{H}_\alpha$ – $^{13}\text{C}_\alpha$  cross-peaks could be distinguished from the interresidue cross-peak on the basis of their higher intensity. However, the

interresidue cross-peak was confirmed by displaying the same 2D  $^1\text{H}$ – $^{13}\text{C}_\alpha$  plane from the HCA(CO)NNH experiment (i.e., a plane at the same  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  coordinates) in the same window.

For side chain assignments, a plane from the 4D HCC(CO)NNH experiment containing  $^1\text{H}$ – $^{13}\text{C}$  cross-peaks of side chain ( $i$ ) was displayed in a fourth window (D) by automatically setting the scroll bars on the basis of  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  coordinates of residue ( $i+1$ ) from the HNCAHA experiment in the first window. The 3D HCCH-TOCSY was also analyzed in a third window (C), starting at the  $^{13}\text{C}_\alpha$  plane defined by the intrasidue  $\text{H}_\alpha(i)$ – $\text{C}_\alpha(i)$  cross-peak from the HNCAHA experiment in the second window. To proceed with assignments of the next N-terminal residue, the coordinates of the interresidue  $\text{H}_\alpha(i-1)$ – $\text{C}_\alpha(i-1)$  cross-peak in the second window were used to limit the display of the 4D

HNNCAHA experiment in the first window to the 2D  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  plane containing the intraresidue  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  cross-peak of the N-terminal residue. The process is then repeated.

**Assignment Results.** General exceptions to our strategy for assignment of side chains included side chains N-terminal to prolines which are predictably missing in the HCC(CO)-NNH experiment. These side chains were assigned exclusively with the 3D HCCH-TOCSY starting with  $^1\text{H}_\alpha$ – $^{13}\text{C}_\alpha$  chemical shifts derived from the HNNCAHA experiment. Assignment of the proline residues was also problematic. Prolines at the unstructured N- and C-termini (Pro2 and Pro74) were visible but highly overlapped in the HCCH-TOCSY. By contrast, the three remaining prolines which are in ordered regions of MCP-1 were not apparent in the HCCH-TOCSY. We therefore relied entirely on the HCC(CO)NNH experiment which has the advantage over the 3D HCCH-TOCSY that even if the side chains have indistinguishable chemical shifts, they can be assigned on the basis of the dispersion of the  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  resonances of the C-terminal residue. Finally, aromatic side chains are not detected in either the HCCH-TOCSY or HCC(CO)NNH experiments. We initially made connections from  $\text{C}_\beta\text{H}_\beta$  to the  $\text{C}_\delta\text{H}_\delta$  resonances in the 3D  $^{13}\text{C}$ -separated NOESY, and in some cases from the  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  to the  $\text{C}_\delta\text{H}_\delta$  in the 4D  $^{13}\text{C}/^{15}\text{N}$ -separated NOESY, assuming that these intraresidue NOEs would be the strongest. Sequential connections were then made from  $\text{C}_\delta\text{H}_\delta$  resonances to other ring nuclei with the 3D  $^{13}\text{C}$ -separated NOESY and a 2D TOCSY in  $\text{D}_2\text{O}$ . Subsequently, we recorded 2D  $\text{C}_\beta\text{H}_\delta$ ,  $\text{C}_\beta\text{H}_\epsilon$  experiments (Yamazaki et al., 1993) and an aromatic version of the CT  $^{13}\text{C}$  HSQC (Vuister & Bax, 1992) to validate these assignments. One interesting observation is that the  $\text{C}_\beta\text{H}_\delta$  correlation of His66 shows two distinct cross-peaks reflecting the conformational disorder of two different  $\text{H}_\delta$  orientations after rotation about the  $\text{C}_\beta$ – $\text{C}_\gamma$  bond; only the major conformer had been detected using the NOE method.

The only unexpected breaks in the sequential assignments appeared between Cys11 and Thr10 (most likely due to the strong  $^{13}\text{C}_\alpha$ – $^{13}\text{C}_\beta$  coupling within Thr10), between Ser34 and Ser33 because the  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  cross-peak of Ser34 is missing in both 4D experiments, and between Ser63 and Met64, again because of non-first-order  $J_{\text{C}\alpha\text{C}\beta}$  coupling in Ser63. Other regions of the protein requiring special attention included (i) assignment of the side chains of the disulfide-bonded pair, Cys11 and Cys36, which were weak in the  $^{13}\text{C}$ -separated TOCSY but visible, in the HCC(CO)NNH experiment, (ii) the C-terminal helix residues Asp68, Asp65, and Asp62, which are highly overlapped, and (iii) the side chain of Met64 which has severely degenerate  $\text{C}_\beta\text{H}_\beta$  and  $\text{C}_\gamma\text{H}_\gamma$  cross-peaks.

Chemical shift assignments for all  $^{15}\text{N}$ ,  $^{13}\text{C}$ , and  $^1\text{H}$  nuclei are provided in the Supporting Information. Most chemical shifts appear in standard regions with a few interesting exceptions. Lys19 has highly upfield-shifted  $^1\text{H}$  chemical shifts ( $\text{H}_\alpha = 1.75$  ppm;  $\text{H}_\beta$ 's  $\sim 0.8$  ppm) induced by susceptibility anisotropy from the indole side chain of the proximal Trp59. Leu20 also has the most upfield-shifted amide proton (5.3 ppm) for the same reason. Val60 has highly shifted methyls ( $\text{H}_{\gamma 1,2} = 0.54, 0.45$  ppm) due to the nearby aromatic ring of Phe43, and the Cys  $\text{C}_\beta$  residues display a large spread of  $^{13}\text{C}$  chemical shifts ranging from 35.2 ppm (Cys36) to 46.8 ppm (Cys52).

**NOE Assignment and Distance Restraints.** NOE distance restraints were determined from 3D  $^{15}\text{N}$ -, 3D  $^{13}\text{C}$ -, 4D  $^{13}\text{C}/$

$^{13}\text{C}$ -, and 4D  $^{13}\text{C}/^{15}\text{N}$ -separated NOESY spectra. Although less sensitive than the 3D spectra, both 4D experiments contained a large number of useful NOEs and resolved many ambiguities present in the 3D experiments. The most difficult residues to assign in the NOESY data involved several Ile side chains (5, 31, 42, and 51) which were heavily overlapped even in the 4D experiments; these were particularly critical to assign correctly since they contribute to the dimer interface. Similarly, the Pro55  $\text{C}_\delta\text{H}_\delta$  is very close to Asp54  $\text{C}_\alpha\text{H}_\alpha$ , and although clearly resolved in 2D  $J$ -correlated data, they are seriously overlapped in higher dimensionality NOESY data due to the lower resolution in these spectra.

The 3D  $^{15}\text{N}$ - and 4D  $^{13}\text{C}/^{15}\text{N}$ -separated NOESY data were analyzed together in ANSIG by viewing contoured strips along the indirectly detected proton dimension at the  $^{15}\text{N}$  chemical shift of a particular residue in the 3D, while simultaneously viewing 2D  $^1\text{H}$ – $^{13}\text{C}$  planes at the same  $^1\text{H}_\text{N}$ – $^{15}\text{N}$  coordinate in the 4D. The 3D and 4D  $^{13}\text{C}$ -separated NOESY data were examined in a similar fashion.  $^1\text{H}$  contoured strips at a given  $^{13}\text{C}$  chemical shift in the 3D were analyzed along with 2D  $^1\text{H}$ – $^{13}\text{C}$  planes in the 4D at the same  $^1\text{H}/^{13}\text{C}$  position. Intraresidue NOEs were verified, integrated, and assigned using automated interactive macro procedures which display vectors and 2D planes on the basis of the chemical shifts of user-specified nuclei and residues from the assignment database. Other NOESY cross-peaks were assigned by reference to the list of possible assignments within a user-defined chemical shift tolerance output from ANSIG. The NOE intensity was converted into one of four different restraint classes (0.0–2.7, 0.0–3.3, 0.0–5.0, and 0.0–6.0 Å) (Hommel et al., 1992; Kraulis et al., 1994) on the basis of the average of the corresponding cross-peak in every spectrum where it was assigned. With this strategy we were able to accumulate more than 1865 unambiguous distance restraints per monomer prior to carrying out any structure calculations. Following initial structure calculations, we increased this number to 2250 (based on 7200 NOE-derived cross-peaks) by importing calculated structures into ANSIG and using distance limits to assign previously ambiguous NOE assignments.

**Other Restraints.**  $\phi$  angle restraints were determined from  $^3J_{\text{HNH}\alpha}$  measurements;  $J$ -values less than 5 Hz were assigned  $\phi$ -values of  $-65 \pm 30^\circ$ , and  $J$ -values greater than 8 Hz were restrained to  $-120 \pm 30^\circ$ . Residues with indeterminate  $\phi$  angles because of overlap in the  $^3J_{\text{HNH}\alpha}$  experiment were assigned angular restraints of  $-90 \pm 90^\circ$  if the intraresidue  $\text{H}_\text{N}\text{H}_\alpha$  NOE was weaker than the sequential  $\text{H}_\text{N}\text{H}_\alpha$  NOE (Clubb et al., 1994).

Stereospecific assignments of  $\beta$ -methylene protons and estimates of  $\chi_1$  torsion angles were made from qualitative judgment of  $^3J_{\text{NH}\beta}$  and  $^3J_{\alpha\beta}$  from the 3D HNHB and 3D  $^{15}\text{N}$ -separated TOCSY HMQC experiments, in concert with the relative strengths of  $\text{H}_\text{N}$ – $\text{H}_{\beta 1}$  versus  $\text{H}_\text{N}$ – $\text{H}_{\beta 2}$  NOEs from a 3D  $^{15}\text{N}$ -separated ROESY experiment (Powers et al., 1993). These distinctions based on intensity differences were not always possible even for well-resolved methylene pairs. The  $\text{H}_{\delta 21}$  and  $\text{H}_{\delta 22}$  resonances of the three asparagine residues were distinguished by the dominance of the intraresidue  $\text{H}_\beta$ – $\text{H}_{\delta 21}$  cross-peak in the  $^{15}\text{N}$ -separated ROESY. Proline  $\text{H}_\beta$ 's were stereospecifically assigned on the basis of the intensity of  $\text{H}_\alpha\text{H}_\beta$  NOEs in  $^{13}\text{C}$ -separated ROESY and NOESY (Kline et al., 1988). The  $\chi_1$  angles of Ile, Val, and Thr were



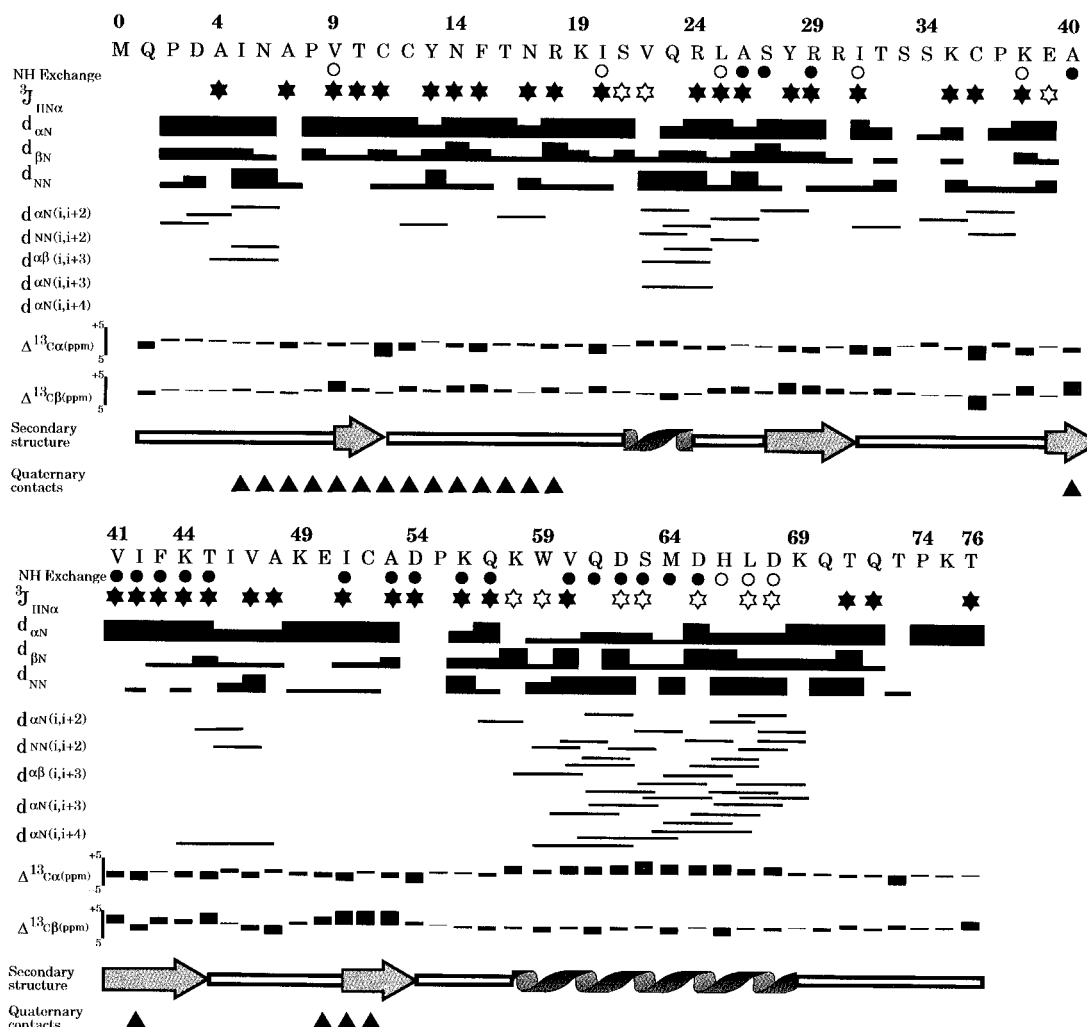


FIGURE 3: Summary of short- and medium-range NOE distances,  $^3J_{\text{HNH}\alpha}$  coupling constants, deviations of  $^{13}\text{C}_\alpha/^{13}\text{C}_\beta$  chemical shifts from random coil values (Spera & Bax, 1991; Wishart & Sykes, 1994), and slowly exchanging amide protons. Relative NOE intensity values are indicated by the bar thickness. Open and filled stars represent  $^3J_{\text{HNH}\alpha}$  less than 5 Hz and greater than 8 Hz, respectively. Open and filled circles designate amide protons which persist up to and beyond 30 min at 35 °C, except for Val9 where slow amide exchange was observed only at 20 °C. Residues which have intermonomer contacts in the dimer are designated with filled triangles.

measured using the  $^3J_{\text{C}_\gamma\text{N}}$  and  $^3J_{\text{C}_\gamma\text{C}'} J$ -modulated CT HSQC experiments. In the case of valine residues, these experiments also allowed stereospecific assignment of the  $\text{C}_{\gamma 1}/\text{C}_{\gamma 2}$  pairs. Side chain orientations of these residues were verified in an ensemble of structures which were computed without  $\chi_1$  restraints. Stereospecific assignments of Leu  $\text{H}_\delta$  chemical shifts and  $\chi_2$  torsion angles were determined by analysis of the structures.

A tightly bound water molecule and hydrogen bond acceptors of slowly exchanging  $\text{H}_\text{N}$  protons (carbonyl oxygens as well as several side chain hydroxyl protons) were located at later stages of the refinement by inspection of the structure. Particularly noteworthy is the hydroxyl of Thr45, identified by strong NOEs to the backbone amides of Val47, Ala48, and Lys49. A water molecule was identified in the  $^{15}\text{N}$ -separated ROESY experiment and verified in a doubly filtered 2D ROE difference experiment (Grzesiek & Bax, 1993) in close proximity to Cys12  $\text{H}_\text{N}$ ; this presumably bridges to the carbonyl of Lys38 or possibly the  $\text{H}_\text{N}$ . A summary of restraints used in deriving the structure is collected in Table 2.

**Secondary Structure.** Data reflecting the secondary structure of MCP-1 are summarized in Figure 3. The presence of strong  $\text{H}_\alpha\text{--H}_\text{N}$  ( $i, i+1$ ) and weak  $\text{H}_\text{N}\text{--H}_\text{N}$  ( $i, i+1$ ) NOEs

coupled with values of  $^3J_{\text{HNH}\alpha}$  greater than 8 Hz indicates four regions of  $\beta$ -sheet. These include residues 9–11 ( $\beta_0$ ), 27–31 ( $\beta_1$ ), 40–45 ( $\beta_2$ ), and 51–54 ( $\beta_3$ ). Interestingly, in  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$ , many amide protons are significantly protected from exchange but the first strand shows only protection of Val9. In addition to the four strands of sheet there are two helical regions. A long helix extends from approximately residue 58 to residue 69 and is well-defined by  $^3J_{\text{HNH}\alpha}$  less than 5 Hz, several slowly exchanging amide protons, and very strong  $\text{H}_\text{N}\text{--H}_\text{N}$  ( $i, i+1$ ) NOEs as well as  $\text{H}_\alpha\text{--H}_\text{N}$  ( $i, i+3$ ) and ( $i, i+4$ ) NOEs. In addition, the persistent positive deviation of  $^{13}\text{C}_\alpha$  chemical shifts from random coil values is diagnostic of helical structure (Spera & Bax, 1991; Wishart & Sykes, 1994). Likewise, patterns of data for residues 21–24 suggest that they are involved in a single turn of  $3_{10}$  helix.

**Determination of Interfacial Contacts.** MCP-1 is analogous to other chemokines in that it forms homodimers at concentrations significantly below the concentrations used for NMR studies. Early in the process of accumulating NOEs and prior to the publication of both the MIP-1 $\beta$  (Lodi et al., 1994) and RANTES (Skelton et al., 1995) structures, we became convinced that the NMR dimerization interface of MCP-1 was not the same as that of IL-8 on which it was originally modeled (Gronenborn & Clore, 1991) since we



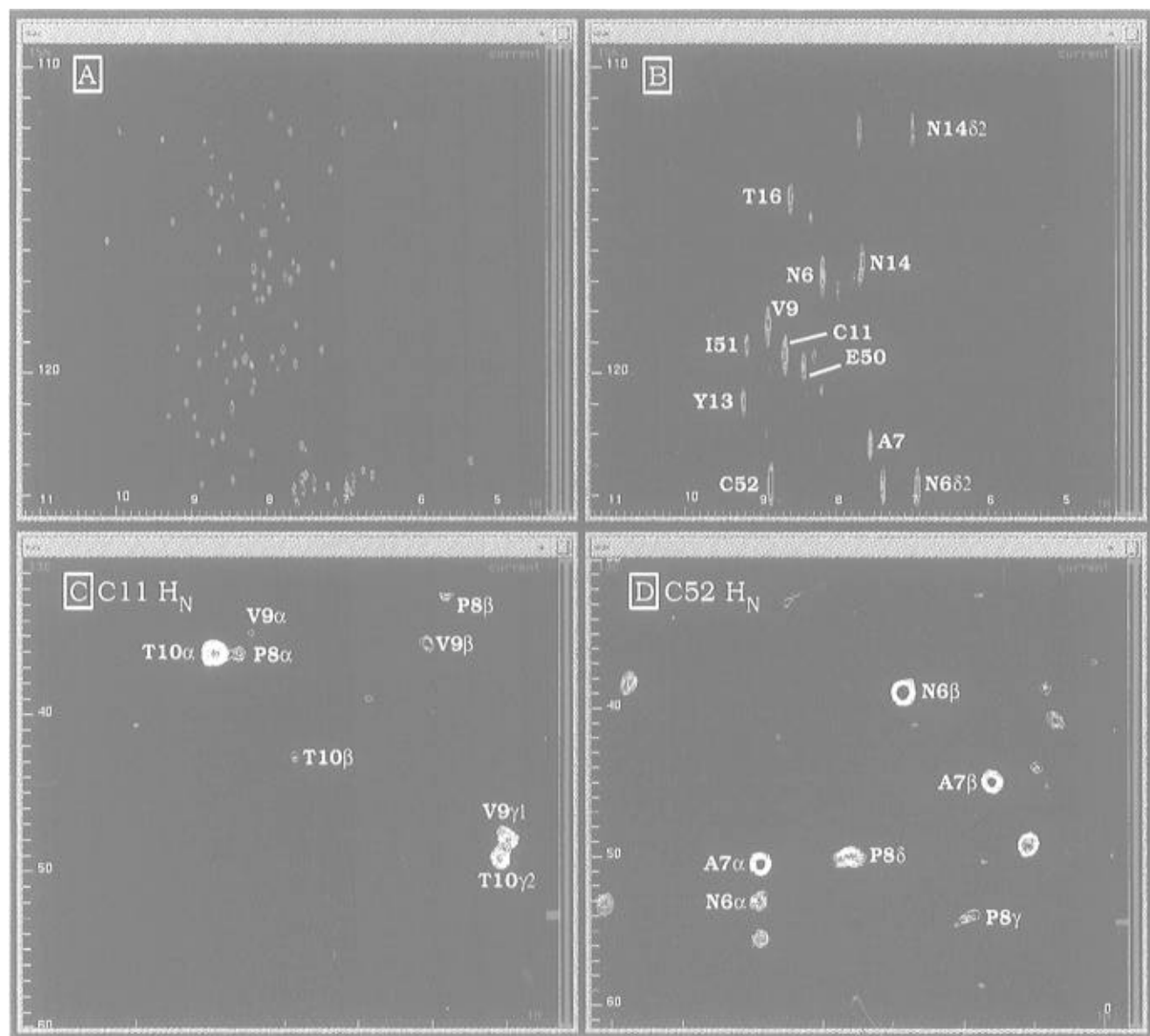


FIGURE 4: Representative spectra from the uniformly ( $^{15}\text{N}$ - $^{13}\text{C}$ ) and asymmetrically ( $^{15}\text{N} + ^{13}\text{C}$ ) labeled MCP-1 dimer. Panel A: 2D  $^{15}\text{N}$  HSQC of uniformly labeled MCP-1. Panel B: 2D  $^1\text{H}$ - $^{15}\text{N}$  version of the (HC)NH NOE experiment on a 50:50 mixture of ( $^{15}\text{N} + ^{13}\text{C}$ ) labeled MCP-1. Only a fraction of the cross-peaks observed in the 2D  $^{15}\text{N}$  HSQC are observed and unambiguously establish the backbone and side chain amide residues involved in the dimerization interface. Panel C: 2D  $^1\text{H}$ - $^{13}\text{C}$  plane from the mixed isotope 3D HC(N)H at the  $^1\text{H}_\text{N}$  chemical shift of Cys11 showing the side chains in contact across the dimer interface. Panel D: 2D  $^1\text{H}$ - $^{13}\text{C}$  plane from the same experiment at the  $^1\text{H}_\text{N}$  chemical shift of Cys52.

could not find NOEs consistent with  $\beta_1$ - $\beta_1$  interfacial contacts. In MCP-1 it was clear that the interface involved the N-terminal residues 6–14 of each monomer; these formed either an additional  $\beta$ -strand ( $\beta_0$ ) with an antiparallel orientation between two monomers or a parallel orientation with the extra strand occurring within one subunit. However, there were also a significant number of NOEs between the N-terminus and other regions of the protein which were difficult to reconcile at this stage.

To determine exactly which residues made contact in the dimerization interface, we carried out  $^{13}\text{C}/^{15}\text{N}$ -separated and filtered NOESY experiments on a sample containing 50%  $^{13}\text{C}$  MCP-1 and 50%  $^{15}\text{N}$  MCP-1. These experiments exclusively detect intermonomer NOEs by removing the magnetic symmetry inherent in a uniformly labeled sample. The 3D HC(N)H experiment yields NOEs between  $^1\text{H}_\text{N}$  on one monomer with proximal  $^1\text{H}$ - $^{13}\text{C}$  pairs on the other monomer. This experiment is identical to the 4D  $^{13}\text{C}/^{15}\text{N}$ -

separated NOESY experiment, but to optimize sensitivity and resolution needs, we only carried out 2D and 3D versions. In the 3D experiment, NOEs were measured using evolution of both  $^{13}\text{C}$  and  $^1\text{H}$  to identify the  $^{13}\text{C}$ - $^1\text{H}$  pairs, since the dispersion of the  $\text{H}_\text{N}$  is sufficient to distinguish them without  $^{15}\text{N}$ -separation once the amide assignments are known. All 2D versions of the experiments were also collected.

Figure 4B illustrates the 2D  $^1\text{H}_\text{N}$ - $^{15}\text{N}$  version of the experiment. Although highly underdigitized, substantially fewer cross-peaks are visible compared to the  $^1\text{H}_\text{N}$ - $^{15}\text{N}$  HSQC of Figure 4A. With these data we could immediately identify the  $^1\text{H}_\text{N}$ - $^{15}\text{N}$  pairs of residues involved in the interface; these include Asn6, Ala7, Val9, Cys11, Tyr13, Asn14, Phe15, and Thr16 near the N-terminus, and Glu50, Ile51, and Cys52. The symmetry axis of the dimer occurs at Thr10 but its  $^1\text{H}_\text{N}$  is directed away from intersubunit side chains and it is not observed. Panels C and D of Figure 4

show 2D  $^{13}\text{C}$ – $^1\text{H}$  planes at the  $^1\text{H}_\text{N}$  chemical shifts of Cys11 and Cys52, respectively. Figure 4C shows NOEs between Cys11  $\text{H}_\text{N}$  and Thr10  $\text{H}_\beta$ ,  $\text{H}_{\gamma/2\#}$ , Val9  $\text{H}_{\gamma1\#}$ , and Pro8  $\text{H}_{\beta\#}$ , all of which could have been mistaken for intramonomer contacts had we not used the mixed isotope sample. At the  $^1\text{H}_\text{N}$  chemical shift of Cys52 (Figure 4D) longer range intersubunit contacts with residues Asn6, Ala7, and Pro8 are observed. In total, we obtained assignments for  $\sim 65$  intersubunit NOEs which unequivocally establish the orientation of the dimer interface as that of an antiparallel  $\beta$ -sheet between the subunits. Initial rounds of structure calculations used the distances from this experiment as a distinct set of intersubunit NOEs, some of which were treated as ambiguous, but in later rounds these cross-peaks were incorporated with the other distance restraints.

**Structure Calculations.** Prior to determining the interface contacts via the experiment described above, we began the first round of structure calculations ambitiously by treating all the distance restraints as 2-fold ambiguous (with no distinction between intrasubunit or intersubunit) using the protocol detailed by Nilges (1993). Starting with completely overlapped, extended monomer templates, our enthusiasm for this impartial approach was tempered by low rates of convergence and unacceptably high NOE energies. While the major elements of secondary and tertiary structure of the protein were well reproduced, the N-terminal dimerization domain and the mobile loop between residues 31 and 35 (*vide infra*) often became severely entangled. Upon analysis of the mixed isotope experiment, however, we observed a clear correspondence between violations in the structure calculations and the contacts independently determined from the experimental data. Two monomers were then crudely docked to satisfy the interfacial NOEs from the mixed isotope experiment and the structures recomputed with two explicit classes of NOEs; all those not comprising the interface were treated with  $(1/r^6)^{-1/6}$  sum averaging (Nilges, 1993) while the explicit interface restraints were maintained in a separate class. The calculated ensemble showed a much higher frequency of successful convergence and substantially reduced, NOE energies. To ensure that docking the starting structure had not induced bias, we also randomly separated preformed monomers and achieved similar high rates of convergence.

Several rounds of structure calculations were then used to gradually repartition the distance restraints between the ambiguous class and a specific inter- and intrasubunit class as well as to correct assignment inconsistencies within ANSIG. At each round of structure calculation approximately 20 converged structures were selected at random, and all interproton distances less than 8 Å were back-calculated for each structure. If both inter- and intrasubunit distances for symmetrically related protons were less than 3 Å apart for any of the structures, the restraint was maintained in the ambiguous class. Of course, it was possible to quickly allocate the majority of distances into the specific inter- or intrasubunit class. Near completion of the structure, we repeated the calculations, starting with exactly overlapped extended monomers but with the NOEs correctly partitioned into their appropriate classes; the convergence rate was identical to that obtained starting with the docked structures.

It is important to note that throughout the calculations the distances were rigorously maintained in the categories automatically generated by ANSIG; if NOE violations

Table 3: Restraint Violations and Energy Statistics for the Ensemble of 20 Structures and Atomic rms Differences to the Unminimized Mean Structure (Residues 5–69)

	energy (kcal/mol)	rms deviation to target
total	747 $\pm$ 39	
NOE (all)	288 $\pm$ 20	0.033 $\pm$ 0.001 Å
dihedral angles	4.5 $\pm$ 0.9	0.66 $\pm$ 0.07°
van der Waals	28 $\pm$ 4	
bond lengths	109 $\pm$ 4	0.0066 $\pm$ 0.0001 Å
bond angles	241 $\pm$ 9	0.60 $\pm$ 0.01°
improper angles	77 $\pm$ 9	0.64 $\pm$ 0.04°
NCS	0.0092 $\pm$ 0.0002	
atomic rms differences		
heavy atoms	0.70 $\pm$ 0.08 (monomer)	0.78 $\pm$ 0.12 (dimer)
backbone	0.22 $\pm$ 0.07 (monomer)	0.37 $\pm$ 0.17 (dimer)

occurred, the data were examined for the cause. Usually the source of the problem was overlap with another cross-peak (particularly in the 3D spectra) resulting in distance restraints being assigned to a distance class that was too short. When this occurred, the overlapped cross-peak was removed from the list of cross-peaks included in the integration and the restraint maintained only if present as an isolated cross-peak in the 4D experiment. During structure calculations the restraint files automatically generated from ANSIG and Perl scripts were used without modification except that the ambiguous class was manually edited to remove pseudoatoms because of an X-PLOR atom selection restriction in  $(1/r^6)^{-1/6}$  sum averaging (Nilges, 1993). Typically, from a total of 100 random starting structures, 40–80 converged to produce no distance violations greater than 0.4 Å or angle violations greater than 4°; violations beyond these ranges usually occurred in the mobile N- and C-termini. The stereochemical quality of the ensemble was checked with PROCHECK and showed no residues in disallowed regions. The energy statistics and rmsd values for the final ensemble are summarized in Table 3. The distribution of the NOEs across the primary sequence is shown in Figure 5A and specific contacts are depicted in Figure 5B. The reproducibility of the structures in the ensemble and the well-ordered regions of the protein can be judged by the variation in the backbone  $\text{C}_\alpha$  rmsd and the backbone  $\phi$ ,  $\varphi$ , and side chain  $\chi_1$  angular order parameters (Hyberts et al., 1992) shown in panels C, D, E, and F of Figure 5, respectively.

## DISCUSSION

A stereoview of an ensemble of 20 structures of the MCP-1 dimer is shown in two orientations in Figure 6 along with MOLSCRIPT (Kraulis, 1991) ribbon diagrams. A closer view of some well-ordered side chains in the hydrophobic core where the C-terminal helix packs onto the three-stranded  $\beta$ -sheet is depicted in Figure 7. The large number of experimental restraints define a high quality structure with the backbone rmsd to the mean structure of 0.22  $\pm$  0.07 and 0.37  $\pm$  0.17 Å over residues 5–69 of the monomer and dimer, respectively, and 0.70  $\pm$  0.08 and 0.78  $\pm$  0.12 Å for all heavy atoms (Table 3).

The bulk of the experimental restraints are derived from heteronuclear separated NOESY spectra collected with mixing times of 80–100 ms. Spin diffusion is rampant as evidenced by the nearly equal intensities of NOEs involving geminal methylene pairs. ROESY spectra ( $\tau_\text{m}$  = 30 ms) which do not suffer from this limitation were useful in stereo

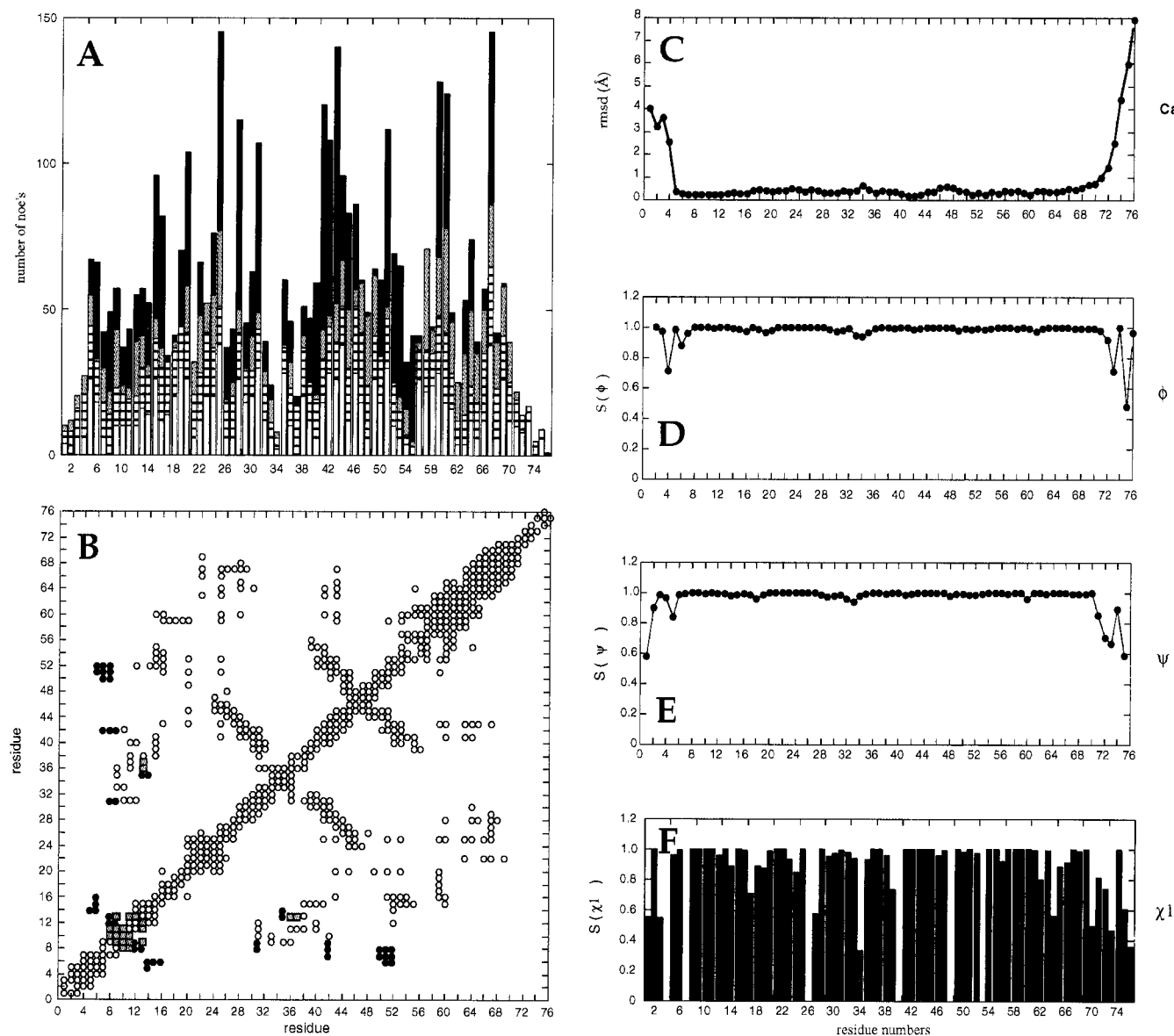


FIGURE 5: (A) Distribution of NOEs along the primary sequence:  $|i-j| > 4$  (black),  $2 < |i-j| < 4$  (shaded), sequential (horizontal bars), and intraresidue (unshaded). (B) Contact map illustrating the intramolecular NOEs (open circles), intermolecular NOEs (filled circles), and ambiguous NOEs (shaded squares). (C) Deviation of  $C_{\alpha}$  rmsd from the mean of the ensemble from residues 1–76. The structures are ordered from residues 5–69. (D and E) Backbone angular order parameters  $\phi$  and  $\psi$  (Hyberts et al., 1992) showing the ordered region from residues 5–69. (F) Side chain  $\chi_1$  angular order parameters for all residues.

assigning many pairs and providing  $\chi_1$  angular restraints. These restraints are particularly important for small amino acids such as aspartic acid and serine where there are few side chain NOEs to define a particular rotamer conformation; inclusion of the  $^{15}\text{N}$ -separated ROESY as a separate class of distance restraints improves the local geometry. The greatest difficulty we found in making stereo assignments was in using a short mixing time (30 ms)  $z$ -filtered  $^{15}\text{N}$ -separated TOCSY as an indirect measure of the  $^3J_{\alpha\beta}$  coupling constant because we frequently did not see substantial intensity differences between  $H_{\beta}$ 's. Our approach was conservative in that unless the intensity differences were greater than approximately 2, and the data were self-consistent with the  $^{15}\text{N}$  ROESY and HNHB data, we did not use the stereo assignment. In spite of these limitations, the  $\chi_1$  angular order parameters (Figure 5F) show that most side chains are well determined; deviations occur for small residues without well-defined angular restraints and for polar side chains on the surface.

It is important to reemphasize the conservative separation we used in distinguishing discrete intrasubunit and intersubunit NOEs of the dimer from potentially ambiguous contacts. A good example is between Tyr13 and the side chains of Lys35, Cys36, Pro37, and Lys38. Our intermediate back-calculation of all distances less than 8 Å which show differences between inter- and intrasubunit contacts of less than 3 Å indicates that many restraints between these residues must be retained as ambiguous. Since all cross-peaks in a symmetrical dimer can have contributions from both intra- and intersubunit NOEs, this conservative criterion and  $(1/r^6)^{-1/6}$  sum averaging of ambiguous restraints do not impose any assumptions about the NOE assignments.

The finer details of the secondary structure emerge more fully when the tertiary and quaternary structures are complete. Residues on the N-terminal side of Ile5 are not well-defined (Figure 5C) although there is a tendency in some structures for a distorted helical turn to allow the interaction of the hydrophobic side chains of Ile5 and Pro2. A bend occurs

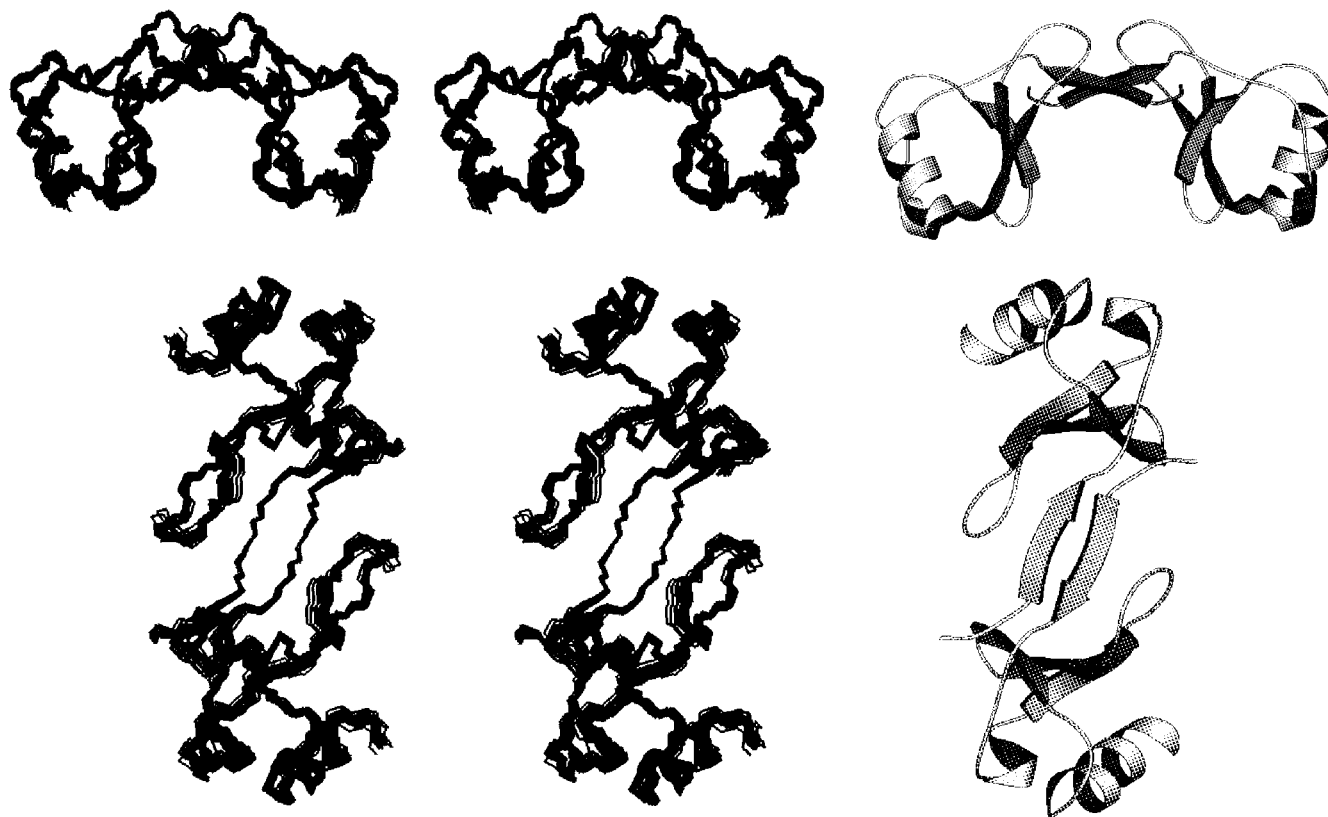


FIGURE 6: Stereoviews and MOLSCRIPT (Kraulis, 1991) representations of an ensemble of 20 structures of MCP-1 from two perspectives (top and bottom). Residues from 1–4 and 70–76 are excluded.

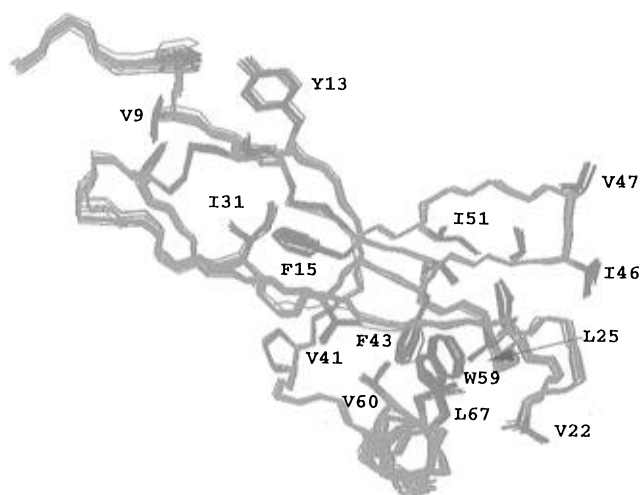


FIGURE 7: View of some ordered side chains in the structure, including those which stabilize the packing of the C-terminal helix onto the three-stranded Greek key.

at Pro8 which initiates the interfacial strand  $\beta_0$  from Pro8 to Cys12. The central three residues Val9, Thr10, and Cys11 form a well-ordered sheet with an intersubunit hydrogen bond between Val9  $H_N$  and Cys11' carbonyl. An extended series of bends from Tyr13 to Ile20 are followed by a single helical turn from Ser21 to Arg24 which immediately leads into the first  $\beta$ -strand of a Greek key motif (residues Ser27 to Ile31). A less structured hydrogen-bonded turn forms from Ser33 through Cys36 characterized by relatively few NOE restraints, lower angular order parameters, rapid solvent exchange of Ser34  $H_N$ , and possible intersubunit contacts with N-terminal residues. The central strand of the antiparallel  $\beta$ -sheet from Ala40 to Thr45 shows an extensive hydrogen-bonded network. A type 3:5 turn highlighted by

strong NOEs between Thr45  $O_H$  and Val47, Ala48, and Lys49  $H_N$  reverses the final  $\beta_3$  strand from Glu50 to Ala53 to orient it with respect to  $\beta_2$ . Another loop from Asp54 to Lys58 with a central Pro55–Lys56 reverse turn orients the regular  $\alpha$ -helix from Trp59 to Lys69, allowing it to pack onto the antiparallel sheet. Beyond Gln70 to the C-terminus the structure is highly mobile with  $^{15}N$ – $H$  heteronuclear NOE values (data not shown) of +0.5 to –2.0 and evidence of two distinct Pro74 isomers.

The orientation of the helix relative to the Greek key is stabilized by extensive interactions between a number of hydrophobic side chains including Val60 of the helix with Phe43 and Val41 in  $\beta_2$  and Leu67 of the helix with Leu25 and Phe43 (Figure 7). Hydrophobic interactions between Val22, Leu25, Phe43, and Trp59 also stabilize the helical turn between residues 21 and 24. Hydrophobic groups which are not part of the core and do not have an obvious role in stabilizing the protein include Ile46 and Val47. It is interesting to note that both corresponding residues in MIP-1 $\beta$  and RANTES are polar (Arg and Lys). A number of electrostatic and H-bonding interactions between side chains also contribute to the overall structure and stability. For example, a salt bridge network involving Lys58–Asp62 and Asp65–Lys69 lies along the exposed hydrophilic face of the helix, Lys44 and Glu50 interact across the 40's 3:5 turn (45–49), and the Tyr28 OH of strand  $\beta_1$  potentially forms a hydrogen bond to the side chain carboxylate of Asp68.

**Quaternary Structure.** To a first approximation, the overall quaternary structure of the MCP-1 dimer resembles RANTES (Chung et al., 1995; Skelton et al., 1995) and MIP-1 $\beta$  (Lodi et al., 1994). The shape of the protein is elongated with the two monomers oriented to give a fairly large pocket (Figure 8, center). The dimerization interface in MCP1

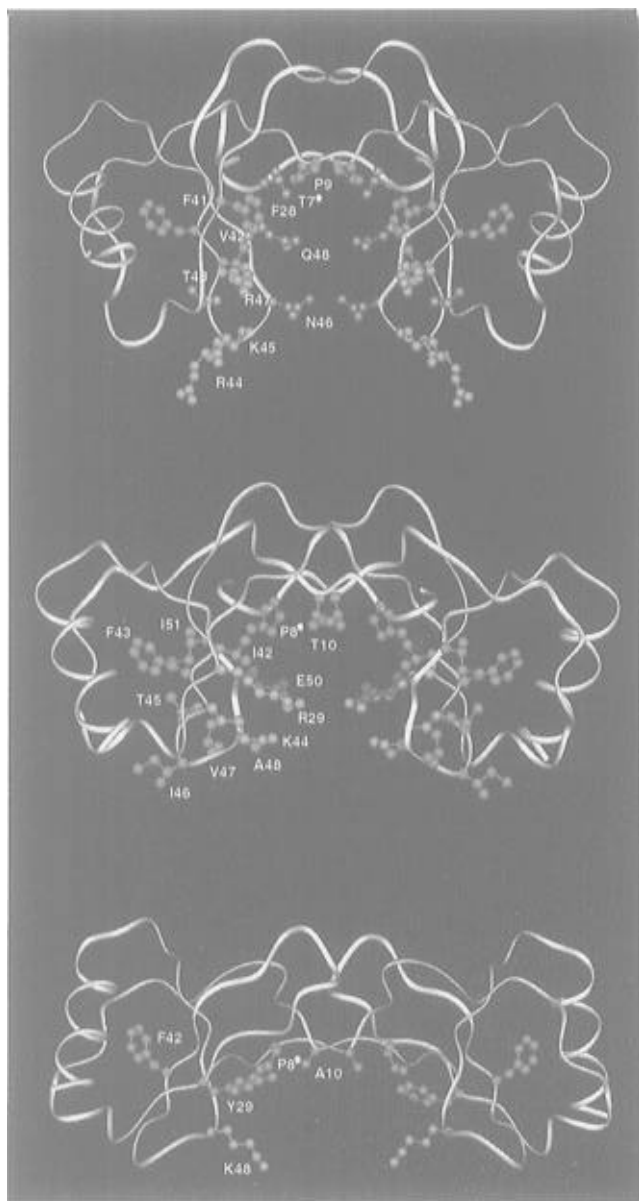


FIGURE 8: Comparison of the backbone topology and side chains which project into the cavity formed by association of the two monomers: RANTES (top), MCP1 (center), and MIP-1 $\beta$  (bottom).

involves N-terminal residues from Ile5 to Thr16 which are extended and oriented in an antiparallel fashion. The core of the interface is the  $\beta_0$  strand centered around Thr10, Thr10' whose methyl groups interact on the face of the strand pointing into the pocket. The side chain of Pro8 is also oriented into the pocket and makes contacts with the side chains of Ile42 and Glu50 from the opposing monomer. Additional contacts are made between the N-terminus and more distant regions of the protein including the 30's loop (31–35), Ile42, Glu50, Ile51, and Cys52. On the convex side of the  $\beta_0$  strand, the aromatic side chains of Tyr13 from each monomer point toward each other with the potential of forming hydrogen bonds between the two hydroxyl groups (Figure 9, center). This orientation is further stabilized by packing onto the side chain of Val9 from the other monomer and between Pro37 from both monomers. N-terminal to the  $\beta_0$  strand Asn6 shows multiple intersubunit NOEs to Asn14, Phe15, Thr16, Ile51, and Cys52. Likewise Ala7 contacts residues Glu50, Ile51, and Cys52. Although only in a partially ordered region of the N-terminus, Asp3 is within

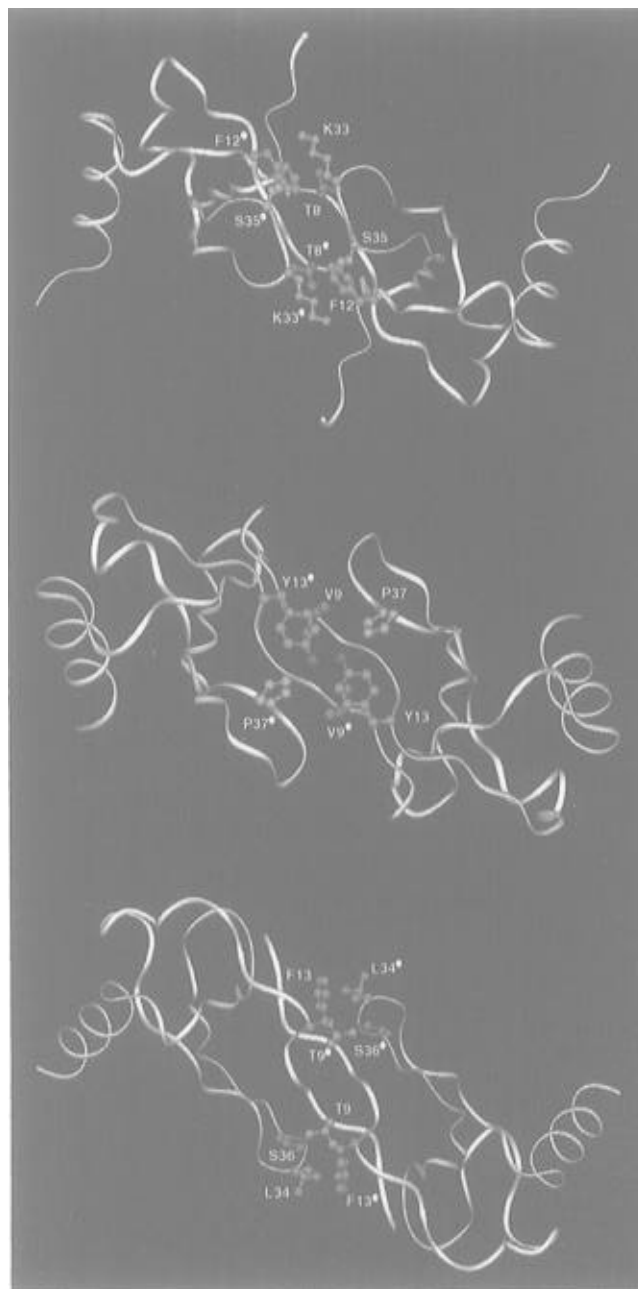


FIGURE 9: Comparison of the side chain orientations of aromatic residues in the dimer interface and the hydrophobic residues on the opposing monomer with which they interact. Separate monomers are indicated by the presence or absence of a filled circle attached to the residue label. RANTES (top), MCP1 (center), and MIP-1 $\beta$  (bottom).

range to form an intersubunit salt bridge with Lys49. Between residues Tyr13 and Asn14 the backbone of one monomer loops out and over the N-terminal residues (Ala7 and Asn6) of the other subunit. This occurs to allow packing of the bulky side chain from Phe15 into a cavity created by Cys12, Cys52, Lys38, and Ala40.

**Comparison to RANTES and MIP-1 $\beta$ .** The secondary structure of the MCP-1 monomer strongly resembles that observed in MIP-1 $\beta$  and RANTES. There are, however, substantial differences at the tertiary level. Because both MIP-1 $\beta$  and MCP-1 have an additional residue at the N-terminus forming the dimer interface, and MCP-1 has an extra residue between  $\beta_1$  and the following turn, we restrict the overlap to the  $\beta_2$ ,  $\beta_3$ , and  $\alpha$ -helical regions of secondary structure (40–45, 50–53, and 59–69) to compute the rmsd

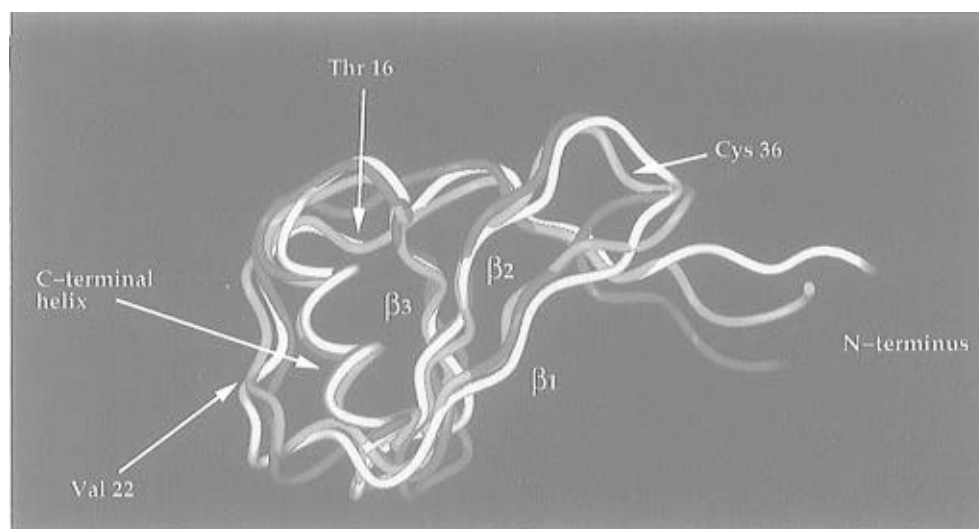


FIGURE 10: Optimal superposition of monomers from MIP-1 $\beta$  (yellow), RANTES (red), and MCP-1 (blue) showing divergent N-termini and other regions of difference and similarity.

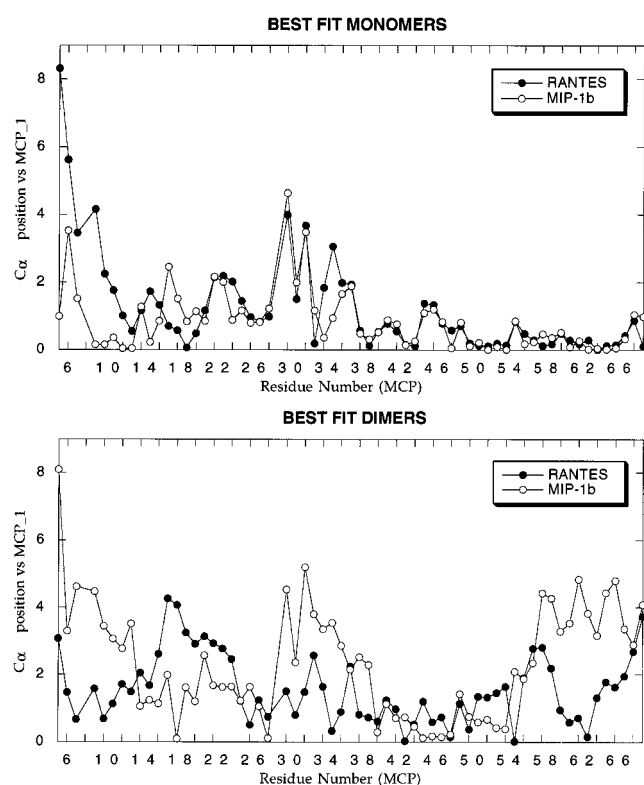


FIGURE 11: Absolute value of the difference in  $C_{\alpha}$  coordinates after optimal superposition of monomers (A, top) and dimers (B, bottom) of MIP-1 $\beta$  and RANTES onto MCP-1.

values. Between monomers, MIP-1 $\beta$  and RANTES show an rmsd of 0.7 Å, while the fit of MCP-1 to both RANTES and MIP-1 $\beta$  is 1.1 Å. The regions of greatest structural variability between the monomers are shown by the superposition of ribbon structures for the three proteins in Figure 10 and the differences in  $C_{\alpha}$  coordinates relative to MCP-1 in Figure 11A. The N-terminal residues orient at distinctly different angles with respect to the much better fits over the remainder of the structures. The angles between MIP-1 $\beta$  and RANTES are the most disparate with MCP-1 intermediate between the other two. Other regions where there are significant differences include the 30's loop and the region of the 20's turn (residues 21–24).

At the quaternary level the differences are even greater with an rmsd of 2.8–4.4 Å between MIP-1 $\beta$ , MCP-1, and RANTES dimers. The regions of difference are shown in the plots of Figure 11B. The dissimilar angle of the N-terminal residues cause a significant change in the relative orientations of the two monomers and in the overall topology of the proteins. RANTES has the most compact structure, MIP-1 $\beta$  is the most elongated, and MCP-1 falls somewhere in the middle (Figure 8). Several additional factors may also account for these effects. One reason for the more open structure of MCP-1 relative to RANTES may be due to differences in the chemical nature of the side chains pointing into the pocket. In RANTES (Figure 8, top), these are either hydrophobic residues (Val42, Phe28, Pro9, and Thr7) or uncharged polar residues (Gln45 and Asn46). Asn46 from each monomer is also within hydrogen-bonding distance (2.9 Å) which may favor the more compact structure. In the case of MCP-1, the hydrophobic residues are Ile42, Pro8, and Thr10 (Figure 8, center). There are also several charged residues (Arg29, Glu50, and Lys44) which may disfavor close approach of the subunits due to electrostatic repulsion between the charged basic residues, although Arg29 and Glu50 have the potential of forming a favorable attraction. Nevertheless, the net result is that the distance between residues in the 40's loop at the base of the pocket of MCP-1 is significantly further apart compared to the 40's loop in RANTES. The same distance in MIP-1 $\beta$  is even longer as described by Skelton et al. (1995).

Another region of the protein which shows a high degree of variability is the C-terminal helix which is significantly rotated away from the sheet in RANTES compared to MCP-1. This is caused primarily by the different sizes of residues which stabilize the packing of the helices onto the sheet and against the 20's loop. In RANTES, these interactions include Ile24 and Leu19 which bracket the 20's loop and interact with Tyr61 from the helix. In MCP-1, the corresponding residues are Leu25, Ile20, and the much smaller side chain of Ser63.

Many of the conserved core residues have similar side chain orientations relative to the backbone in the three  $\beta$ -chemokines. However, there are distinct differences in conformations for a number of interfacial residues which promote favorable contacts with other side chains or accom-

moderate changes in side chain size. A dramatic example of the former involves aromatic residues at the dimer interface. In MCP-1, the hydroxyl group of Tyr13 from one monomer points toward the hydroxyl group of Tyr13' in the other monomer and makes hydrophobic intermonomer contacts with Val9 as well as with Pro37 from both monomers (Figure 9, center). In RANTES the corresponding residue is Phe12. This side chain has a similar  $\chi_1$  rotamer compared to MCP-1 (gauche<sup>+</sup>), but the ring is oriented at a very different angle that optimizes packing against the methylene group of Lys33 from the opposing monomer (Figure 9, top). The Phe12–Tyr13' separation is also significantly larger than Tyr13–Tyr13' distance in MCP-1 and allows closer approach of the two 30's loops in RANTES compared to that in MCP-1. The corresponding side chain of Phe13 in MIP-1 $\beta$  is oriented entirely in the opposite direction (gauche<sup>-</sup>) (Figure 9, bottom) and interacts with the hydrophobic groups of Thr9 and Leu34 from the opposing subunit. It is important to emphasize that these structural features in MCP-1 are the result of a very conservative calculational approach in which NOEs involving Tyr13 were treated as ambiguous and no  $\chi_1/\chi_2$  restraints were imposed in order to avoid bias.

**Correlation with Biological Data.** On the basis of the solution structures of several chemokines from the C-C and C-X-C families, it appears likely that all  $\beta$ -chemokines will have a quaternary topology similar to MCP-1, MIP-1 $\beta$ , and RANTES, whereas all  $\alpha$ -chemokines will resemble IL-8 and MGSA/GRO. This has prompted speculation that the functional distinction between the two families is due to their different quaternary folds (Lodi et al., 1994). The dimer interface of IL-8 (Clower et al., 1990) and MGSA/GRO (Fairbrother et al., 1994) involve hydrophobic interactions between residues in the  $\beta_1$  strands from each monomer; this arrangement is disfavored in MCP-1, RANTES, and MIP-1 $\beta$  because the corresponding residues are polar. Conversely, the N-terminal residues in the  $\beta$ -chemokines are more hydrophobic than those in the  $\alpha$ -chemokines and provide a driving force for dimerization mediated by the  $\beta_0$  strand.

If functional segregation of the two families is indeed encoded by the quaternary topology, it requires these proteins bind to their receptors as homodimers. Although this is an appealing mechanism, the biology is still emerging with data for and against homodimers as the functionally competent form of the protein (Covell et al., 1994). On the basis of mutagenesis and cross-linking studies, Zhang and Rollins (1995) have argued that the active form of MCP-1 is a dimer. However, their data conflict with studies of an IL-8 derivative that does not dimerize but is fully active (Rajaraman et al., 1994). If, on the other hand, monomers are the functionally relevant form, is there a purpose to the distinct quaternary structure classes adopted by C-C and C-X-C chemokines? Clarke-Lewis (Gong & Clarke-Lewis, 1995) has done extensive mutational analysis of the N-terminal residues of MCP-1 and has shown that residues 1–6 are required for functional activity, while residues 7–10 are important for receptor desensitization. Several of these residues are hydrophobic, and as pointed out by Skelton et al. (1995), dimerization may be a simple consequence of the exposure of hydrophobic surface area from residues required for function.

Although the N-terminal residues are required for functional activity of the protein, other regions are necessary since a peptide spanning residues 1–10 was inactive by itself

(Gong & Clarke-Lewis, 1995). Some of the most instructive mutational studies have involved deletion mutants (Gong & Clarke-Lewis, 1995; Zhang & Rollins, 1995), particularly a fragment retaining only residues 9–76 of MCP-1. Although functionally inactive, this mutant has a  $K_d$  only 3-fold higher than the intact protein (8.3 vs 2.8 nM). More importantly, it is a potent antagonist of wild-type MCP-1 and MCP-3, suggesting that important determinants of binding reside within this domain. It is not clear whether this protein functions as an antagonist by forming inactive heterodimers with the wild-type protein which would imply that MCP-1 is active as a homodimer or if it simply competes with MCP-1 for receptor binding sites. Our current structure of the MCP-1 homodimer can guide mutational studies which will ultimately resolve these issues and lead to the identification of additional regions of the protein which are important for binding and signal transduction. In turn, the combination of structural and mutagenesis data should contribute to the development of effective therapies against chronic macrophage-mediated inflammatory diseases.

## ACKNOWLEDGMENT

We acknowledge and thank the Cambridge "supergroup", Ernest Laue, Per Kraulis, and Wayne Boucher for providing the ANSIG software and the maximum entropy code interface and the continuing helpful advice of Ernest Laue on all aspects of the NMR, including the idea for the mixed isotope experiment. We also thank Henry George for construction of the MCP-1 vector, Sue Lichter, Tom Patterson, and Henry George for construction of the TAP302 cells, Ann Yetter and Daniel Camac for purifying the protein, and Richard Yates and Stuart Rosenfeld for the fermentation growths. Finally, we thank Jim Krywko for computer and graphics support and Sharon Archer for critical reading of the manuscript.

## SUPPORTING INFORMATION AVAILABLE

A table containing <sup>1</sup>H, <sup>15</sup>N, and <sup>13</sup>C chemical shift assignments for MCP-1 at pH 5.42, 35 °C (3 pages). Ordering information is given on any current masthead page.

## REFERENCES

- Adams, D. H., Harvath, L., Bottaro, D. P., Interrante, R., Catalano, G., Tanaka, Y., Strain, A., Hubscher, S. G., & Shaw, S. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 7144–7148.
- Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson.* 95, 636–641.
- Bax, A., Clore, G. M., & Gronenborn, A. M. (1990a) *J. Magn. Reson.* 88, 425–431.
- Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990b) *J. Magn. Reson.* 86, 304–318.
- Boucher, W., Laue, E. D., Campbell-Burk, S., & Domaille, P. J. (1992a) *J. Am. Chem. Soc.* 114, 2262–2264.
- Boucher, W., Laue, E. D., Campbell-Burk, S. L., & Domaille, P. J. (1992b) *J. Biomol. NMR* 2, 631–637.
- Breg, J. N., van Opheusden, J. H., Burgering, M. J., Boelens, R., & Kaptein, R. (1990) *Nature* 346, 586–589.
- Brunger, A. T. (1992) *X-PLOR Manual, Version 3.0*, Yale University, New Haven, CT.
- Burgering, M. J., Boelens, R., Gilbert, D. E., Breg, J. N., Knight, K. L., Sauer, R. T., & Kaptein, R. (1994) *Biochemistry* 33, 15036–15045.
- Campbell-Burk, S. L., Domaille, P. J., Starovasnik, M. A., Boucher, W., & Laue, E. D. (1992) *J. Biomol. NMR* 2, 639–646.
- Carr, M. W., Roth, S. J., Luther, E., Rose, S. S., & Springer, T. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 3652–3656.



- Chung, C. W., Cooke, R. M., Proudfoot, A. E. I., & Wells, T. N. C. (1995) *Biochemistry* 34, 9307–9314.
- Clore, G. M., Appella, E., Yamada, M., Matsushima, K., & Gronenborn, A. M. (1990) *Biochemistry* 29, 1689–1696.
- Clore, G. M., Kay, L. E., Bax, A., & Gronenborn, A. M. (1991) *Biochemistry* 30, 12–18.
- Clore, G. M., Omichinski, J. G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E., & Gronenborn, A. M. (1994) *Science* 265, 386–391.
- Clowes, R. T., Boucher, W., Hardman, C. H., Domaille, P. J., & Laue, E. D. (1993) *J. Biomol. NMR* 3, 349–354.
- Clubb, R. T., Ferguson, S. B., Walsh, C. T., & Wagner, G. (1994) *Biochemistry* 33, 2761–2772.
- Covell, D. G., Smythers, G. W., Gronenborn, A. M., & Clore, G. M. (1994) *Protein Sci.* 11, 2064–2072.
- Ebisawa, M., Yamada, T., Bickel, C., Klunk, D., & Schleimer, R. P. (1994) *J. Immunol.* 153, 2153–2160.
- Fairbrother, W. J., Reilly, D., Colby, T. J., Hesselgesser, J., & Horuk, R. (1994) *J. Mol. Biol.* 242, 252–270.
- Folkers, P. J., Nilges, M., Folmer, R. H., Konings, R. N., & Hilbers, C. W. (1994) *J. Mol. Biol.* 236, 229–246.
- Gong, J.-H., & Clarke-Lewis, I. (1995) *J. Exp. Med.* 181, 631–640.
- Gronenborn, A. M., & Clore, G. M. (1991) *Protein Eng.* 4, 263–269.
- Grzesiek, S., & Bax, A. (1993) *J. Biomol. NMR* 3, 627–638.
- Grzesiek, S., Vuister, G. W., & Bax, A. (1993) *J. Biomol. NMR* 3, 487–493.
- Hommel, U., Harvey, T. S., Driscoll, P. C., & Campbell, I. D. (1992) *J. Mol. Biol.* 227, 271–282.
- Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci.* 1, 736–751.
- Kameyoshi, Y., Schroder, J. M., Christophers, E., & Yamamoto, S. (1994) *Int. Arch. Allergy Immunol.* 104, 49–51.
- Kay, L. E., Clore, G. M., Bax, A., & Gronenborn, A. M. (1990) *Science* 249, 411–414.
- Kline, A. D., Braun, W., & Wuthrich, K. (1988) *J. Mol. Biol.* 204, 675–724.
- Kraulis, P. J. (1989) *J. Magn. Reson.* 84, 627–633.
- Kraulis, P. J. (1991) *J. Appl. Crystallogr.* 24, 946–950.
- Kraulis, P. J., Domaille, P. J., Campbell-Burk, S. L., Vanaken, T., & Laue, E. D. (1994) *Biochemistry* 33, 3515–3531.
- Laskowski, R. A., Macarthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
- Laue, E. D., Mayger, R., Skilling, J., & Staunton, J. (1986) *J. Magn. Reson.* 68, 14–29.
- Lee, W., Harvey, T. S., Yin, Y., Yau, P., Litchfield, D., & Arrowsmith, C. H. (1994) *Nat. Struct. Biol.* 1, 877–890.
- Leonard, E. J., & Yoshimura, T. (1990) *Immunol. Today* 11, 97–101.
- Li, Y. S., Shyy, Y. J., Wright, J. G., Valente, A. J., Cornhill, J. F., & Kolattukudy, P. E. (1993) *Mol. Cell. Biochem.* 126, 61–68.
- Lodi, P. J., Garrett, D. S., Kuszewski, J., Tsang, M. L., Weatherbee, J. A., Leonard, W. J., Gronenborn, A. M., & Clore, G. M. (1994a) *Science* 263, 1762–1767.
- Loetscher, P., Seitz, M., Clark-Lewis, I., Baggiolini, M., & Moser, B. (1994) *FASEB J.* 8, 1055–1060.
- Maghazachi, A. A., al-Aoukaty, A., & Schall, T. J. (1994) *J. Immunol.* 153, 4969–4977.
- Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry* 28, 6150–6156.
- Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wuthrich, K. (1989) *J. Magn. Reson.* 85, 608–613.
- Miller, M. D., & Krangel, M. S. (1992) *Crit. Rev. Immunol.* 12, 17–46.
- Murphy, P. M. (1994) *Annu. Rev. Immunol.* 12, 593–633.
- Nelken, N. A., Coughlin, S. R., Gordon, D., & Wilcox, J. N. (1991) *J. Clin. Invest.* 88, 1121–1127.
- Neote, K., Darbonne, W., Ogez, J., Horuk, R., & Schall, T. J. (1993a) *J. Biol. Chem.* 268, 12247–12249.
- Neote, K., DiGregorio, D., Mak, J. Y., Horuk, R., & Schall, T. J. (1993b) *Cell* 72, 415–425.
- Neote, K., Mak, J. Y., Kolakowski, L. F., Jr., & Schall, T. J. (1994) *Blood* 84, 44–52.
- Nilges, M. (1993) *Proteins: Struct., Funct., Genet.* 17, 297–309.
- Norwood, T. J., Boyd, J., Heritage, J., Soffe, N., & Campbell, I. D. (1990) *J. Magn. Reson.* 488–501.
- Oppenheim, J. J., Zachariae, C. O., Mukaida, N., & Matsushima, K. (1991) *Annu. Rev. Immunol.* 9, 617–648.
- Paolini, J. F., Willard, D., Consler, T., Luther, M., & Krangel, M. S. (1994) *J. Immunol.* 153, 2704–2717.
- Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., & Clore, G. M. (1993) *Biochemistry* 32, 6744–6762.
- Rajarathnam, K., Sykes, B. D., Kay, C. M., Dewald, B., Geiser, T., Baggiolini, M., & Clarke-Lewis, I. (1994) *Science* 264, 90–92.
- Schall, T. J., & Bacon, K. B. (1994) *Curr. Opin. Immunol.* 6, 865–873.
- Schall, T. J., Bacon, K., Toy, K. J., & Goeddel, D. V. (1990) *Nature* 347, 669–671.
- Skelton, N. J., Aspiras, F., Ogez, J., & Schall, T. J. (1995) *Biochemistry* 34, 5329–5342.
- Sozzani, S., Molino, M., Locati, M., Luini, W., Cerletti, C., Vecchi, A., & Mantovani, A. (1993) *J. Immunol.* 150, 1544–1553.
- Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc.* 113, 5490–5492.
- Takeya, M., Yoshimura, T., Leonard, E. J., & Takahashi, K. (1993) *Hum. Pathol.* 24, 534–539.
- Vuister, G. W., & Bax, A. (1992) *J. Magn. Reson.* 98, 428–435.
- Vuister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc.* 115, 7772–7777.
- Vuister, G. W., Wang, A. C., & Bax, A. (1993) *J. Am. Chem. Soc.* 115, 5334–5335.
- Wishart, D. S., & Sykes, B. D. (1994) *J. Biomol. NMR* 4, 171–180.
- Wuthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol.* 169, 949–961.
- Yamazaki, T., Forman-Kay, J. D., & Kay, L. E. (1993) *J. Am. Chem. Soc.* 115, 11054–11055.
- Yip, P. (1990) *J. Magn. Reson.* 90, 382–383.
- Yoshimura, T., & Leonard, E. J. (1990) *J. Immunol.* 145, 292–297.
- Yu, X., Dluzy, S., Graves, D. T., Zhang, L., Antoniadis, H. N., Hollander, W., Prusty, S., Valente, A. J., Schwartz, C. J., & Sonenshein, G. E. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 6953–6957.
- Zhang, Y., & Rollins, B. J. (1995) *Mol. Cell Biol.* 15, 4851–4855.
- Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., & Fesik, S. W. (1990) *J. Magn. Reson.* 86, 210–216.
- Zuiderweg, E. R. P., Petros, A. M., Fesik, S. W., & Olejniczak, E. T. (1991) *J. Am. Chem. Soc.* 113, 370–372.

BI9602270