C-Terminal Retroviral-Type Zinc Finger Domain from the HIV-1 Nucleocapsid Protein Is Structurally Similar to the N-Terminal Zinc Finger Domain†

Terri L. South,‡ Paul R. Blake,‡ Dennis R. Hare,§ and Michael F. Summers*,‡

Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, Baltimore, Maryland 21228, and Hare Research, Incorporated, 14810 216th Avenue North East, Woodinville, Washington 98072

Received February 4, 1991; Revised Manuscript Received March 15, 1991

ABSTRACT: Two-dimensional NMR spectroscopic and computational methods were employed for the structure determination of an 18-residue peptide with the amino acid sequence of the C-terminal retroviral-type (r.t.) zinc finger domain from the nucleocapsid protein (NCP) of HIV-1 [Zn(HIV1-F2)]. Unlike results obtained for the first retroviral-type zinc finger peptide, Zn(HIV1-F1), [Summers et al. (1990) *Biochemistry 29*, 329], broad signals indicative of conformational lability were observed in the ^{1}H NMR spectrum of Zn(HIV1-F2) at 25 °C. The NMR signals narrowed upon cooling to −2 °C, enabling complete ^{1}H NMR signal assignment via standard two-dimensional (2D) NMR methods. Distance restraints obtained from qualitative analysis of 2D nuclear Overhauser effect (NOESY) data were used to generate 30 distance geometry (DG) structures with penalties (penalty = sum of the squared differences between interatomic distances defined in the restraints file and in the DG structures) in the range 0.02–0.03 Å^2. All structures were qualitatively consistent with the experimental NOESY spectrum based on comparisons with 2D NOESY back-calculated spectra. Superposition of the backbone atoms (C, Cα, N) for residues C(1)–C(14) gave pairwise RMSD values in the range 0.16–0.75 Å. The folding of Zn(HIV1-F2) is very similar to that observed for Zn(HIV1-F1). Small differences observed between the two finger domains are localized to residues between His(9) and Cys(14), with residues M(11)–C(14) forming a 3_{10} helical corner. Superposition of Zn(HIV1-F2) structures onto Zn(HIV1-F1) structures gave pairwise RMSD values in the ranges 0.4–0.7 Å [backbone atoms of residues C(1)–H(9) superpositioned] and 0.7–1.2 Å [backbone atoms of residues C(1)–C(14) superpositioned]. These results indicate that the r.t. zinc finger sequences observed in retroviral NCPs, simple plant virus coat proteins, and in a human single-stranded nucleic acid binding protein share a common structural motif.

Without exception, retroviral nucleocapsid proteins (NCPs) contain one or two copies of a conserved "retroviral-type" (r.t.) zinc finger sequence, Cys-X_2-Cys-X_4-His-X_4-Cys (Henderson et al., 1981; Copeland et al., 1984), which has been proposed to function physiologically as a zinc-binding domain (Berg, 1986). This motif is not limited to retroviral proteins and has been observed in proteins from yeast transposable element copia, cauliflower mosaic virus, and more recently from humans in a cellular nucleic acid binding protein (CNBP) that contains seven sequential copies of the motif (South & Summers, 1990; Berg, 1990). A common feature of proteins containing the r.t. zinc finger sequence is that they appear to be involved at some stage in sequence-specific single-stranded nucleic acid interactions. Thus the conserved sequence may represent a generic zinc finger motif for single-stranded nucleic acid binding, analogous to the classical-type zinc finger motif found widely in duplex-DNA-binding proteins (Summers, 1991).

Studies from several laboratories involving synthetic r.t. zinc finger peptides and synthetic proteins have shown that these sequences are capable of binding zinc stoichiometrically and with high affinity (South et al., 1989, 1990a; Green & Berg, 1989; Roberts et al., 1989; Summers et al., 1990). Low zinc dissociation constants (10^{-10}–10^{-12} M) measured for synthetic r.t. zinc finger peptides and a synthetic protein indicate that, under normal cellular conditions, these domains should be fully populated with zinc (Green & Berg, 1990). Indeed, HIV-1 particles appear to contain sufficient quantities of zinc to populate the NCP r.t. zinc finger domains (South et al., 1990b). Detailed NMR-based structural studies of a peptide with sequence of the first HIV-1 NCP zinc finger domain, Zn(HIV1–F1), revealed a novel structural motif containing extensive internal hydrogen bonding (Summers et al., 1990). The folding of residues C(1)–K(6) [a numbering scheme is used where the first Cys of the domain is labled C(1)] is virtually identical with the folding exhibited by related iron-binding residues of rubredoxin (Adman et al., 1975).

More recently, we established that the intact HIV-1 nucleocapsid protein, isolated from virus particles and reconstituted with zinc, gives NMR spectra with features that are remarkably similar to features observed in the NMR spectra of Zn(HIV1-F1) (South et al., 1990b) and with a peptide containing the sequence of the C-terminal r.t. zinc finger domain, Zn(HIV1-F2) [the amino acid sequence of Zn(HIV1-F2) is K-G-C-W-K-C-G-K-E-G-H-Q-M-K-D-C-T-E]. To determine if the C-terminal r.t. zinc finger domain adopts folding similar to that of the N-terminal domain, 2D NMR and structural studies have been carried out on the synthetic metallopeptide Zn(HIV1-F2). This represents the second structural study of an r.t. zinc finger peptide and is an important step toward the NMR signal assignment and structure determination of the intact HIV-1 nucleocapsid protein.

† This work was supported by ACS Institutional Research Grant IN-147F, administered by the MD Cancer Program/University of Maryland (M.F.S.), and by NIH Grant GM42561 (M.F.S.).

* Author to whom correspondence should be addressed.

‡ University of Maryland, Baltimore County.

§ Hare Research, Inc.

0006-2960/91/0430-6342$02.50/0

Experimental Procedures

Computations. NMR data processing and structure calculations were carried out with Silicon Graphics 4D-20 and 4D-220 computers. NMR data were transferred via ethernet and converted to "readable" files with an in-house program (GENET). NMR data processing and analysis, distance geometry calculations, and NOESY back-calculations were performed with FTNMR, DSPACE, and BKCALC software packages (Hare Research, Inc).

NMR Spectroscopy. NMR data were collected with a GE GN500 (500.11-MHz, ^1H) NMR spectrometer. Sample conditions were as follows: 10 mg (9.5 mM) of Zn(HIV1-F2) (Peptide Technologies Inc., Washington, D. C.) in D_2O (99.9%, MSD) or 90% H_2O/10% D_2O solutions, pH 7.0, −2 °C. ^1H NMR chemical shifts were referenced to internal H_2O (5.03 ppm at −2 °C). 2D NMR data were collected without sample spinning. All homonuclear 2D NMR data were processed with zero filling to a final spectrum size of 2048 × 2048 noncomplex data points. Pulse sequences and parameters used to collect and process the 2D NMR data are described below.

2D NOESY data (τ_m = 50, 150, and 300 ms) for the sample containing 90% H_2O/10% D_2O were obtained with a recycle delay of 2.5 s and were processed with 6-Hz exponential line broadening in t_2 and 90°-shifted squared sine bell filtering in t_1, with third-order polynomial baseline correction in the F_1 domain subsequent to the final Fourier transform. Solvent suppression was achieved either with a 1−$\bar{1}$−echo read pulse sequence (Sklenar & Bax, 1987; read delay periods τ_1 and τ_2 of 90 and 210 μs, respectively) or with a combination of DANTE presaturation (+90° phase shifted relative to the preparation pulse) (Morris & Freeman, 1978; Zuiderweg et al., 1986), followed by a SCUBA pulse train (Brown et al., 1988) for partial recovery of saturated α-protons. Other parameters were as follows: 2 × 256 × 2048 data matrix sizes; 64 scans per t_1 increment. For D_2O solutions: 2 × 256 × 2048 raw data matrix sizes; 48 scans per t_1 increment.

Phase-sensitive HOHAHA data (Davis & Bax, 1985) were obtained for aqueous samples with a 1−$\bar{1}$−echo read pulse (Bax et al., 1987) with read delay periods of 90 (τ_1) and 210 μs (τ_2). Other parameters were as follows: 2 × 256 × 4096 data matrix size [two separate sets of States−Haberkorn-type data (States et al., 1982), with 256 data points in t_1 and 4096 data points in t_2]; 64 scans per t_1 increment; recycle delay = 2 s; MLEV-17 mixing period = 45 ms, preceded and followed by 2.0-ms trim pulses; 8.1 kHz spin-lock field strength; 90° squared sine bell filtering in the t_2 and t_1 dimensions.

2QF phase-sensitive COSY (Pinantini et al., 1982) data were obtained for the Zn(HIV1-F2) sample (D_2O solution) with solvent preirradiation to eliminate the residual HDO signal. The 2D data matrix (2 × 512 × 2048 complex points) was processed with 4-Hz exponential line broadening in the t_2 dimension and 8-Hz exponential plus trapezoidal filtering in the t_1 dimension.

Results

Effect of Temperature on 1D Spectra. Broad signals observed in the 1D ^1H NMR spectrum obtained for Zn(HIV1-F2) at temperatures above approximately 10 °C indicate that the molecule is conformationally labile. Variable temperature ^1H NMR data obtained for Zn(HIV1-F2) are shown in Figure 1. At 40 °C all aromatic protons gave rise to broad ^1H NMR signals. When the temperature was lowered to −2 °C, the aromatic proton NMR signals narrowed considerably, and several of the Trp aromatic proton signals exhibited expected splitting due to scalar coupling (Figure 1). These data reflect a decrease in conformational lability at lower temperatures.

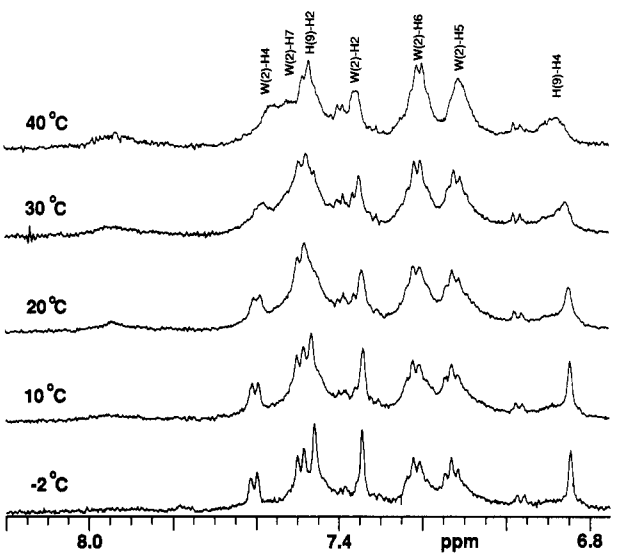

FIGURE 1: Downfield region of the ^1H NMR spectra obtained for Zn(HIV1-F2) over the temperature range −2–40 °C. The signal narrowing at lower temperatures reflects reduced conformational lability.

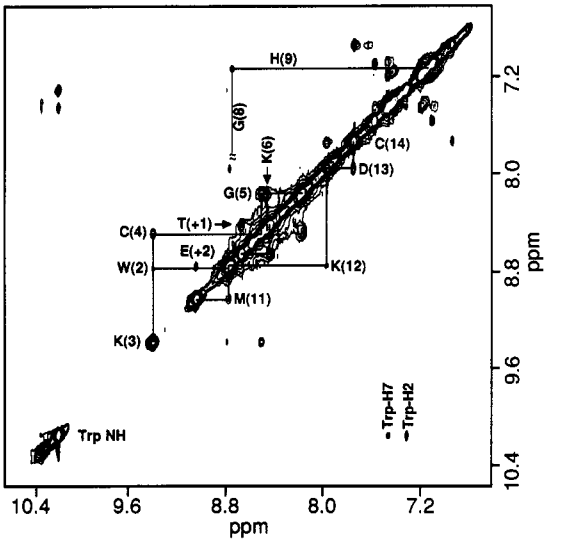

FIGURE 2: Downfield region of the 2D NOESY spectrum obtained for Zn(HIV1-F2) (9.5 mM; 90% H_2O/10% D_2O; pH 7) showing connectivities involving amide and aromatic protons. The spectrum is not symmetric with respect to the diagonal due to the nonlinear 1−$\bar{1}$−echo excitation pulse employed.

NMR Signal Assignments. All 2D NMR data for signal assignment and structure determination were obtained for 9.5 mM Zn(HIV1-F2) samples at −2 °C. ^1H NMR signal assignments were made by determining scalar connectivities within amino acid residues from COSY and HOHAHA spectra and then correlating the signals of adjacent residues on the basis of dipolar connectivities obtained from 2D NOESY data (Table I). This approach has been described in detail (Wuthrich, 1986). Three stretches of NH−NH connectivities were observed: W(2)−K(6), M(11)−C(14), and T(+1)−E(+2) (Figure 2). The most intense NH−NH connectivities were observed for the K(3)−C(4), C(4)−G(5), and D(13)−C(14) proton pairs. Dipolar connectivities from amide protons to α- and side-chain protons were also used for sequential signal assignments, and the "fingerprint" region of the NOESY spectrum that contains these connectivities is

Table I: ^{1}H NMR Chemical Shift Assignments for Zn(HIV1-F2)[a]

residue	NH	CαH	CβH	others
K(−2)	*b*	3.93	1.84	γCH$_2$, 1.42[c]
				δCH$_2$, 1.64[c]
				εCH$_2$, 2.89
G(−1)	8.74	3.74		
		3.37		
C(1)	8.19	3.96	2.84	
			2.00	
W(2)	8.79	4.36	3.46	H2, 7.31
			3.35	H4, 7.57
				H5, 7.10
				H6, 7.19
				H7, 7.46
				NH, 10.16
K(3)	9.39	4.16	2.26	γCH$_2$, 1.64[c]
				δCH$_2$, 1.25[c]
				εCH$_2$, 2.88
C(4)	8.50	4.89	3.17	
			2.49	
G(5)	8.16	4.07		
		3.80		
K(6)	8.46	4.34	1.85	γCH$_2$, 1.39
				δCH$_2$, 1.56
				εCH$_2$, 2.94
E(7)	8.58	4.16	1.79	γCH$_2$, 2.26[d]
			1.98[d]	2.14[d]
G(8)	8.75	4.36		
		3.66		
H(9)	7.16	4.78	3.16	H2 7.44
				H4 6.80
Q(10)	9.03	4.63	1.89	γCH$_2$, 2.03[d]
			2.03[d]	
M(11)	9.03	4.76	2.16[d]	γCH$_2$, 2.49[d]
				CH$_3$, 2.10
K(12)	8.77	4.12	1.71[d]	γCH$_2$, 1.85[c,d]
				δCH$_2$, 1.21[c,d]
				εCH$_2$, 2.87
D(13)	7.96	4.88	2.94	
			2.51	
C(14)	7.75	3.65	3.28	
			2.85	
T(+1)	8.44	4.43	4.43	CH$_3$, 1.06
E(+2)	8.67	4.04	1.82	γCH$_2$, 2.34[d]
				2.16[d]

[a] Chemical shift in ppm relative to internal H_2O (5.03 ppm) at −2 °C. [b] Signal for terminal NH_3^+ not observed. [c] Overlap precluded unambiguous differentiation of the γ and δ methylene proton signals for a given amino acid residue. [d] Overlap precluded unambiguous differentiation of the β and γ methylene proton signals for a given amino acid residue.

shown in Figure 3. Sequential NOESY connectivity data and relative amide proton labilities are summarized in Figure 4.

Structure Determination

Modification of our Original Approach. In our earlier studies of Zn(HIV1-F1), 2D NOESY back calculations were used extensively as part of the structure refinement approach (Summers et al., 1990). In that approach, randomly embedded initial atom positions were treated with minimal (low-velocity) simulated annealing, affording DG structures with modest penalty values (ca. 3–5 Å^2). These DG structures were then subjected to conjugate gradient minimization (CGM), affording new structures with penalties in the range of ca. 0.5–2 Å^2. When additional CGM was unable to further reduce the penalty for a particular structure, 2D NOESY back calculations were performed, and new distance restraints dictated by discrepancies between the experimental and back-calculated spectra were added to the experimental restraints list. Freshly embedded DG structures minimized with the modified restraints list generally exhibited penalty values lower than those of the previously refined structures, and the new DG structures

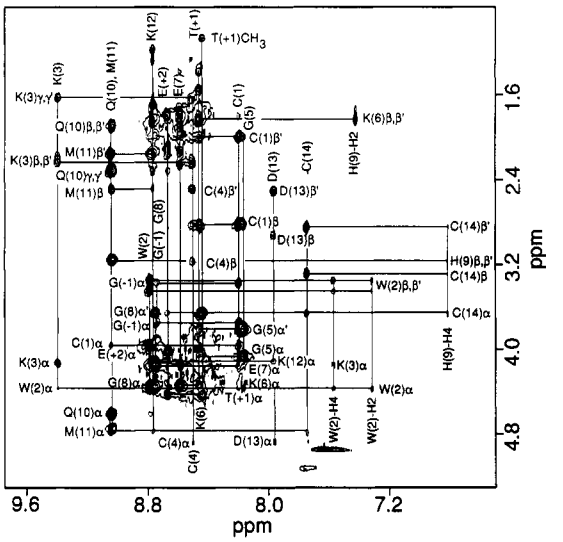

FIGURE 3: "Fingerprint" region of the 2D NOESY spectrum obtained for Zn(HIV1-F2) showing connectivities from amide and aromatic protons to α- and side-chain protons.

FIGURE 4: Sequential NOESY connectivities used in the assignment of the ^{1}H NMR spectrum of Zn(HIV1-F2) (T = −2 °C). For comparison, relevant results for Zn(HIV1-F1) (T = 30 °C) are included. Filled, shaded, and open circles reflect estimated low-, moderate-, and high-sensitivity, respectively, of the amide proton signal intensity to solvent preirradiation. d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ represent NOESY connectivities from backbone amide protons to neighboring amide protons, α-protons, and β-protons, respectively. The relative intensities of the NOEs, classified as strong, medium, and weak, are indicated by the thickness of the horizontal lines. The star (*) represents potential connectivities that could not be unambiguously assigned due to signal overlap.

generally gave back-calculated NOESY spectra that were more consistent with experimental data. This cycle of (1) random embedding, followed by (2) minimal simulated annealing/CGM, (3) back calculation, and (4) restraints modification was repeated iteratively until structures consistent with the experimental data could be obtained. Thus, in our earlier approach, addition of restraints based on iterative 2D NOESY back calculations led to modifications of the ψ,ϕ penalty surface function, and enough restraints were eventually added to allow generation of low-penalty structures with only minimal (low-velocity) simulated annealing and conjugate gradient minimization steps. Final DG structures generated in this manner exhibited penalties of less than 0.21 Å^2. Most important, low penalties (<0.21 Å^2) were also obtained when these final structures were reanalyzed with the initial set of experimental restraints (i.e., restraints generated prior to any back calculating). Clearly, our earlier application of minimal simulated annealing and GCM was insufficient for adequate sampling of conformational space.

Table II: Covalent and Experimental Restraints Used To Generate Zn(HIV1-F2) Structures[a]

Covalent Restraints		
Atoms with Assigned Prochiral Protons		
cb[3]	ca[7]	cb[16]
cb[6]	cb[11]	
Restraints to Zinc		
cb[3] 3.3581	sg[3] 2.3	ne2[11] 2.0
cb[6] 3.3581	sg[6] 2.3	nd1[11] 4.13 ∞
cb[16] 3.3581	sg[16] 2.3	cg[11] 4.13 ∞
Experimental Restraints Based on Direct NOE Cross Peaks		
Alpha to Alpha		
ha[3] ha[12] 2.0 3.0	has[2] ha[9] 2.0 3.0	
har[2] ha[9] 2.0 3.0		
Amide to Amide		
hn[4] hn[5] 2.0 3.0	hn[13] hn[14] 2.0 3.0	
hn[5] hn[6] 2.0 2.5	hn[14] hn[15] 2.0 3.0	
hn[6] hn[7] 2.0 2.5	hn[15] hn[16] 2.0 2.5	
hn[7] hn[8] 2.0 2.5	hn[17] hn[18] 2.0 3.0	
hn[10] hn[11] 2.0 2.5		
Amide to Alpha		
hn[2] har[2] 2.0 4.0	hn[11] ha[9] 2.0 4.0	
hn[2] has[2] 2.0 4.0	ha[12] ha[12] 2.0 3.0	
hn[3] har[2] 2.0 3.0	hn[13] ha[3] 2.0 4.0	
hn[3] has[2] 2.0 3.0	hn[13] ha[13] 2.0 3.0	
hn[3] ha[3] 2.0 3.0	hn[14] ha[13] 2.0 4.0	
hn[4] ha[3] 2.0 2.5	hn[14] ha[14] 2.0 3.0	
hn[4] ha[4] 2.0 3.0	hn[15] ha[14] 2.0 4.0	
hn[5] ha[4] 2.0 4.0	hn[15] ha[15] 2.0 4.0	
hn[6] ha[6] 2.0 4.5	hn[16] ha[13] 2.0 4.0	
hn[7] har[7] 2.0 2.5	hn[16] ha[16] 2.0 3.0	
hn[8] has[7] 2.0 3.0	hn[17] ha[16] 2.0 3.0	
hn[8] ha[8] 2.0 3.0	hn[17] ha[17] 2.0 3.0	
hn[9] ha[8] 2.0 2.5	hn[18] ha[16] 2.0 3.0	
hn[9] ha[9] 2.0 4.0	hn[18] ha[17] 2.0 3.0	
hn[10] ha[9] 2.0 4.0	hn[18] ha[18] 2.0 3.0	
Amide to Beta		
hn[3] hbr[3] 2.0 3.0	hn[12] hbr[12] 2.0 3.0	
hn[3] hbs[3] 2.0 3.0	hn[12] hbr[15] 2.0 4.0	
hn[4] hbr[4] 2.0 3.0	hn[13] hbs[13] 2.0 3.0	
hn[4] hbs[4] 2.0 4.0	hn[14] hbs[13] 2.0 3.0	
hn[6] hbr[3] 2.0 4.0	hn[15] hbs[11] 2.0 4.0	
hn[6] hbr[6] 2.0 3.0	hn[15] hbs[15] 2.0 3.0	
hn[7] hbr[3] 2.0 3.0	hn[16] hbs[16] 2.0 2.5	
hn[8] hbr[3] 2.0 3.0	hn[16] hbr[16] 2.0 3.0	
hn[12] hbs[11] 2.0 3.0		
alpha to side chain		
ha[3] hbs[3] 2.0 3.0	ha[13] hbr[13] 2.0 4.0	
ha[4] hbr[4] 2.0 3.0	ha[13] hbs[13] 2.0 3.0	
ha[4] hbs[4] 2.0 3.0	ha[13] hbs[16] 2.0 3.0	
ha[6] hbs[6] 2.0 2.5	ha[15] hbs[15] 2.0 3.0	
ha[6] hbr[6] 2.0 2.5	ha[16] hbr[16] 2.0 4.0	
ha[11] hbr[11] 2.0 3.0	ha[16] hbs[16] 2.0 3.0	
Histidine Aromatics		
hd2[11] hbs[3] 2.0 4.0	hd2[11] hbs[16] 2.0 3.0	
hd2[11] ha[16] 2.0 3.0	he1[11] hbr[8] 2.0 4.0	
hd2[11] hbs[11] 2.0 3.0	he1[11] hbs[8] 2.0 4.0	
hd2[11] hn[12] 2.0 4.0	he1[11] hn[8] 2.0 4.0	
hd2[11] hn[16] 2.0 4.0		
Methyls		
mg2[17] hn[17] 2.0 4.0	mg2[17] ha[17] 2.0 4.0	
Betas		
hbs[11] hbr[15] 2.0 3.0		
Tryptophan Aromatics		
hd1[4] ha[4] 2.0 4.0	he3[4] ha[4] 2.0 4.0	
hd1[4] hn[4] 2.0 4.0	he3[4] ha[5] 2.0 4.0	

[a] In this table only, residues (in brackets) are numbered 1–18, beginning with the N-terminal Lys. Numbers for pairs of protons represent lower and upper (∞ = infinite) distance restraints. Atom definitions: h = single proton; m = methyl; a, b, c, d, and e = alpha, beta, gamma, delta, and epsilon, respectively; r = pro-R and s = pro-S stereochemistry.

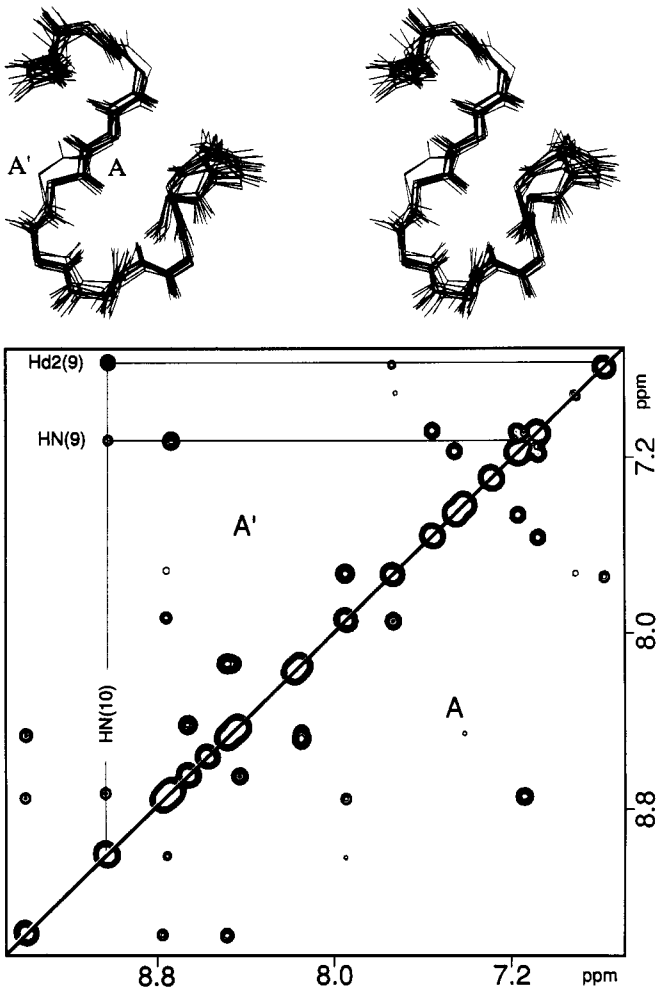

FIGURE 5: (Top) Superposition of the group A′ structures onto the group A structures showing backbone atoms only of residues C(1)–C(14). Note that the H(9)–Q(10) amide–carbonyl peptide bond is reversed in the group A′ structures. (Bottom) Portion of the back-calculated spectrum of a representative group A′ DG structure (Left) showing the HN(9)–HN(10) and Hd2(9)–HN(10) cross peaks that are more intense than those observed in the experimental data (Figure 2). (Right) Portion of the back-calculated spectrum obtained after correction of the H(9)–Q(10) peptide linkage.

We recently described a modified structure determination approach that utilizes several variable-velocity simulated annealing and conjugate gradient minimization steps in the refinement scheme (Blake et al., 1990). With this approach, low-penalty conformational barriers are overcome without the application of additional restraints based on NOE back calculations. This approach was applied previously to the structure determination of Zn(HIV1-F1) (Blake et al., 1990). Low-penalty (ca. 0.05 Å^2) structures were obtained by using loose distance restraints of 2.0–2.5, 2.0–3.5, and 2.0–4.5 Å for qualitatively observed strong, medium, and weak direct dipolar cross-peak intensities, respectively. Structures generated by the modified approach exhibited very high convergence, with pairwise RMSD values for backbone atoms (C, Cα, N) of residues C(1)–C(14) in the range 0.06–0.18 Å (Blake et al., 1990). However, despite the high degree of convergence, the "goodness" of the structures (defined as the ability of the DG structures to reproduce the experimental NOESY spectrum upon back calculation) was not ideal. Addition of restraints to account for minor differences between experimental and back-calculated spectra enabled the generation of new DG structures with substantially reduced penalties (ca. 0.02 Å^2). This additional reduction in penalty is attributed again to

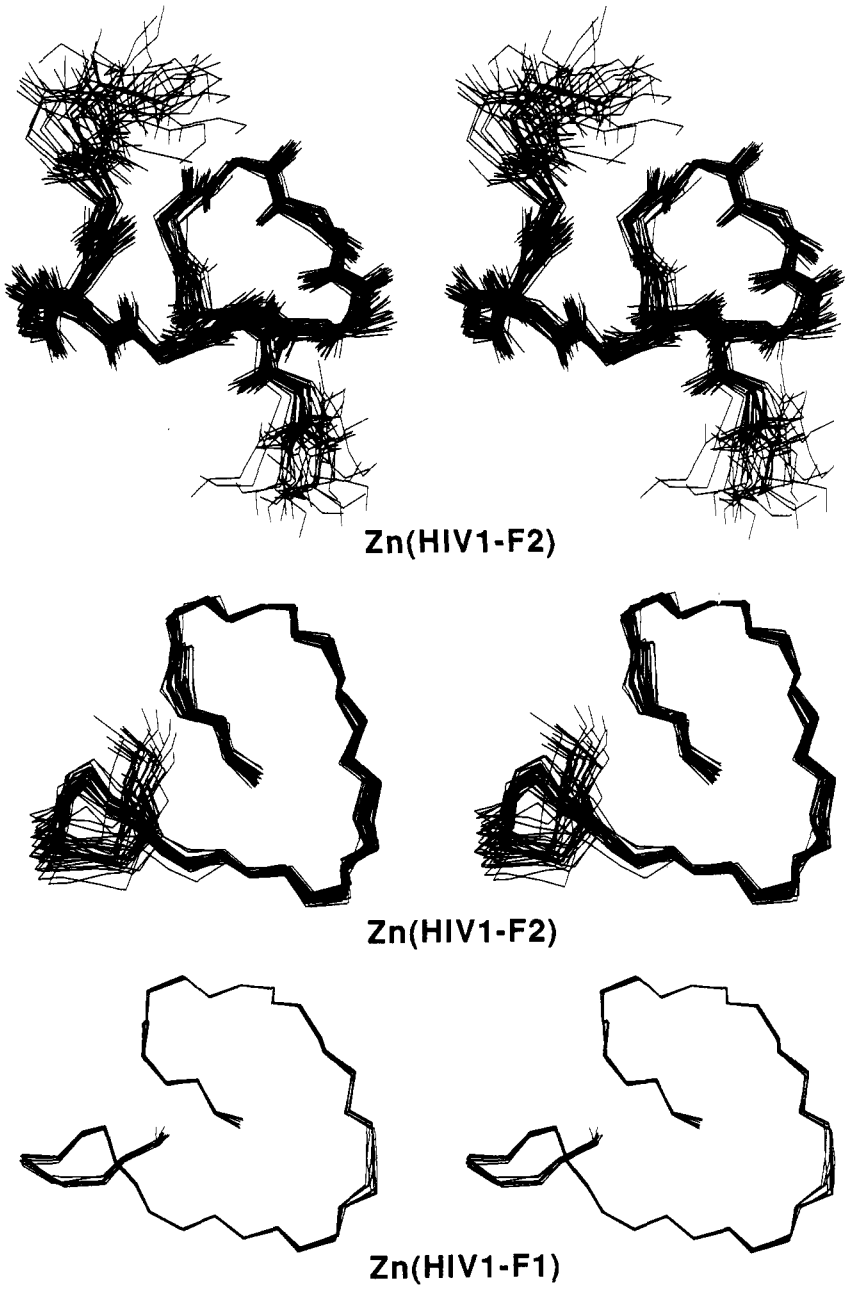

FIGURE 6: (Top) Superposition of residues C(1)–C(14) for DG structures 1–30 of Zn(HIV1-F2) showing backbone atoms including amide hydrogens and carbonyl oxygen atoms for all residues. (Middle) Superposition of residues C(1)–H(9) for DG structures 1–30 of Zn(HIV1-F2) showing backbone atoms only for residues C(1)–C(14). From this orientation, the helical nature of the M(11)–C(14) 3_{10} turn is visible. (Bottom) Superposition of residues C(1)–H(9) for DG structures 1–8 of Zn(HIV1-F1) showing backbone atoms only for residues C(1)–C(14). For comparison, the orientation of Zn(HIV1-F1) is similar to the orientation used to generate the middle figure of Zn(HIV1-F2).

deficiencies in sampling conformational space, where low-velocity simulated annealing enabled low-penalty barriers to be surmounted but led to circumvention of high-penalty barriers. Thus, back calculations do not appear to be necessary to obtain high convergence but are important for maximizing consistence between the DG structures and the experimental NOESY data. In addition, even with more sophisticated refinement schemes, back calculations may be necessary since conformational space can not realistically be sampled entirely for proteins (Skolnick & Kolinski, 1990). This technique approaches the ideal of refining against the actual experimental data.

Application of the Modified Approach to Zn(HIV1-F2). Our modified structure determination approach (Blake et al., 1990) was applied to Zn(HIV1-F2) as follows. Loose distance restraints of 2.0–2.5, 2.0–3.0, and 2.0–4.0 Å representing strong, medium, and weak cross-peak intensities, respectively, were employed. NOE-derived interproton distance restraints are summarized in Table II. For NOEs involving geminal protons where spin diffusion is significant, restraints were included only for the most intense of the pair of cross peaks. Distance restraints were included to enforce sp^3 and sp^2 hybridization for the Cys sulfurs and the His Nϵ coordinating atoms, respectively. Bond length (but not bond angle) restraints for Zn–S(Cys) and Zn–Nϵ(His) bonding were employed as described (Summers et al., 1990).

Covalent and experimental distances were randomly embedded into 3-space and subjected to simulated annealing and conjugate gradient minimization (CGM). Additional structures were generated subsequently by performing two sequential 10-Å randomizations of atom positions (Metzler et al., 1989), followed by simulated annealing and CGM. Thirty

structures generated automatically with this approach were treated with an additional refinement protocol that included variable velocity simulated annealing, SHAKE (Ryckaert, et al. 1977), and CGM algorithms. The resulting structures exhibited penalty values in the range 0.02–0.04 Å^2.

The backbone folding in the 30 initial models was self-consistent except for variations at the Q(10)–M(11) and H(9)–Q(10) amide linkages. In 75% of the structures (group A), dihedral ψ and ϕ angles for the Q(10)–M(11) amide linkage of ca. −170° and −61°, respectively, were observed. The remaining structures (group B) exhibited ψ and ϕ values of −27° and −178°, respectively. In addition, two of the structures exhibited alternate folding for the H(9)–Q(10) linkage (group A′), with ψ and ϕ values of ca. −67° and 79°, respectively, compared to values of 170° and −105°, respectively, observed for the group A structures. Figures showing relative orientations of backbone atoms for the group A and A′ structures are included in Figure 5, and the group B structures are shown in the supplementary material.

To determine which of the three groups of DG structures most accurately reflect the experimental NOESY data, 2D NOESY back calculations were carried out. Comparisons were made only for well-behaved proton signals. NOE cross peaks involving protons of the K(−2), T(+1), E(+2) residues and the side-chain protons of residues K(3), K(6), E(7), Q(10), M(11), and K(12) were not evaluated for consistence with experimental data since these protons are most likely to be associated with conformationally mobile groups. Unfortunately, rapid $t_1\rho$ relaxation precluded the collection of rotating frame Overhauser effect (ROESY) (Bothner-by et al., 1984) data, which generally enables unambiguous differentiation of cross peaks resulting from spin diffusion and direct-dipolar cross relaxation. Here, cross peaks resulting from direct cross relaxation were assigned on the basis of relative cross-peak intensities observed in NOESY data obtained with 50-, 150-, and 300-ms mixing intervals. Unlike in our earlier studies of Zn(HIV1-F1), quantitative analysis of the relaxation matrix based on NOE buildup curves was not performed. Cross relaxation and z-leakage parameters of −50 and 1 s^{-1}, respectively, were employed on the basis of comparisons of the relative intensities of the auto- and cross peaks for the Gly(−1) α-protons.

Back-calculated spectra of the group A structures were generally consistent with the experimental NOESY data. In contrast, back-calculated NOESY spectra for the group A′ and B DG structures contained features inconsistent with the experimental NOESY data. Consider the group A′ structures first. For all group A′ DG structures, back-calculated NOESY spectra exhibited strong cross peaks for the HN(9)–HN(10) and Hd2(9)–HN(10) proton pairs (see Figure 5). In the experimental spectrum, a cross peak was not observed for the HN(9)–HN(10) protons, and a very weak cross peak was observed for the Hd2(9)–HN(10) proton pair (Figure 5). For each group A′ structure, a 180° rotation of the HN(9)–HN(10) amide linkage was performed, followed by low-velocity simulated annealing and CGM minimization. This manipulation resulted in a decrease in the penalty and led to folding consistent with the folding of the group A structures. Thus, low-velocity simulated annealing employed at the latter stages of the refinement procedure was sometimes insufficient to overcome the large penalty barriers associated with a major reorientation of the peptide linkages.

Back-calculated NOESY spectra of the group B structures exhibited cross peaks of weak intensity for the HN(11)–HA(10), HN(11)–HBR(11), and HN(11)–HBS(11) proton pairs. In the experimental spectrum, these cross peaks are very intense (Figure 3 and supplementary figures). These discrepancies were corrected by manual 180° rotation of the Q(10)–M(11) amide linkage, followed by simulated annealing and CGM minimization. In every case, this manipulation led to a reduction in the total penalty and to backbone conformations consistent with the group A structures.

Table III: Structure Restraints and Refinement Data

number of amino acid residues: 18	
molecular weight: 2055	
sequence: K-G-C-W-K-C-G-K-E-G-H-Q-M-K-D-C-T-E	
total number of distance restraints: 96	
intraresidual NOEs: 42	
sequential NOEs ($i,i+1$): 24	
medium range NOEs ($1 < \|i-j\| \leq 5$): 16	
long range NOEs ($\|i-j\| > 5$): 5	
restraints to zinc: 9	
dihedral angles: none	
penalties for DG structures 1–30: 0.022–0.026 Å^2	
largest single bounds violation: 0.032 Å	
Atomic Root Mean Square Differences	
superposition of all atoms: 1.39–2.99 Å	
superposition of backbones of residues C(1)–H(9): 0.06–0.63 Å	
superposition of backbones of residues C(1)–C(14): 0.16–0.75 Å	
superposition of backbones of residues K(−2)–E(+2): 0.33–1.60 Å	

Thirty final superpositioned DG structures are shown in Figure 6. Superposition of backbone atoms (C, Cα, N) of residues C(1)–H(9), C(1)–C(14), and K(−2)–E(+2) afforded pairwise RMSDs in the ranges 0.06–0.63, 0.16–0.75, and 0.33–1.60 Å, respectively. Pairwise RMSDs obtained upon superposition of all atoms were in the range 1.39–2.99 Å. Structure refinement data are summarized in Table III.

Description of Structure and Comparison with Zn(HIV1-F1). Residues C(1)–K(6) fold in a manner virtually identical with the folding observed for the first zinc finger domain [Zn(HIV1-F1)] with RMSDs for backbone atoms (C, Cα, N) of these residues in the range 0.24–0.59 Å. As described earlier (Summers et al., 1990), the folding of residues C(1)–K(6) is also identical to the folding observed via X-ray crystallography for related residues in the iron-binding domain of rubredoxin (Adman et al., 1975; Watenpaugh et al., 1980). Superposition of all backbone atoms of residues C(1)–C(14) of Zn(HIV1-F2) onto the corresponding residues of Zn(HIV1-F1) gives RMSDs in the range 0.73–1.2 Å (Figure 6, bottom). The K(3) and C(4) backbone amide protons of Zn(HIV1-F2) are oriented in a manner consistent with hydrogen bonding to the C(1) sulfur in what has been called a type I NH–S tight turn (Adman et al., 1975). The C(4) α-carbon serves as a corner between the type I NH–S tight turn and a short orthogonally directed type II NH–S tight turn with hydrogen bonding between the amide of K(6) and the C(4) sulfur. The carbonyl oxygen of K(6) is pointing in a direction consistent with hydrogen bonding to the amide hydrogen of C(1), and the amide hydrogen of G(5) exhibits hydrogen bonding to the C(1) carbonyl oxygen. With this folding, the β-protons of K(6) are in the vicinity of the H(9) imidazole H^2 proton, and the G(−1) α-protons are in the vicinity of the E(7) α-proton. The folding implicitly directs the hydrophobic side chain of W(2) away from the body of the peptide and into the solvent. This folding is identical with that observed in Zn(HIV1-F1), in which the phenyl group of F(2) is also directed away from the body of the peptide and toward solvent.

The G(8) α-carbon serves as a corner that leads to β-like stretch. As observed for Zn(HIV1-F1), the G(8) residue exhibits the greatest degree of scatter (i.e., poorest convergence) relative to other residues from C(1) to C(14). The β-like stretch extends from the carbonyl of G(8) through the amide of Q(10). None of the backbone atoms within this

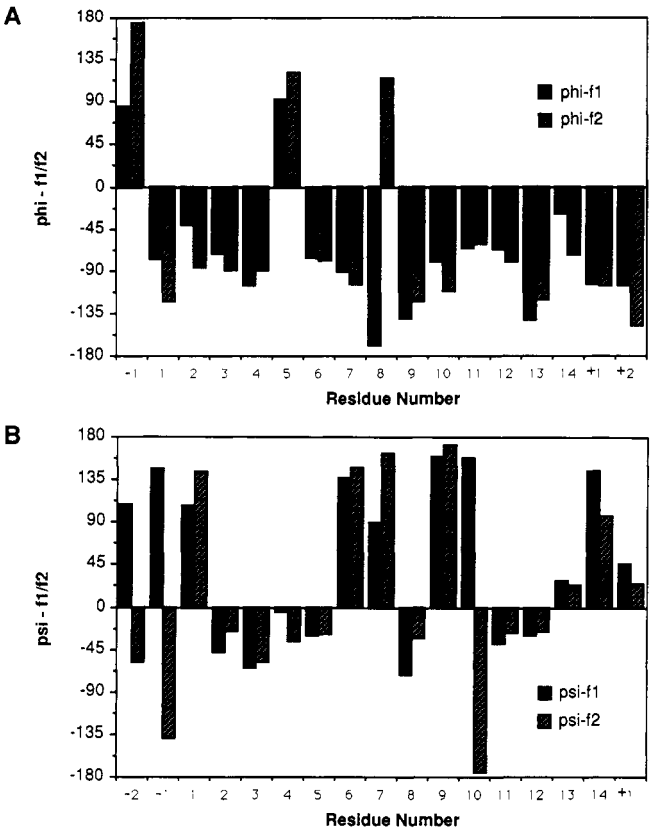

FIGURE 7: Graph of the relative ϕ (A) and ψ (B) values for representative Zn(HIV1-F1) and Zn(HIV1-F2) structures.

stretch appear to be involved in hydrogen bonding.

Backbone atoms of residues M(11)–C(14) form what may be best described as a 3_{10} helical corner (Watenpaugh et al., 1979), which includes apparent hydrogen bonding between the Q(10) carbonyl and the D(13) amide and between the M(11) carbonyl and the C(14) amide. This helical segment is clearly distinguishable in the central drawing of Figure 6. Two features about this region are worth noting: First, this region of the model exhibits somewhat poorer convergence compared to residues C(1)–H(9) (see Figure 6, middle drawing). Structural variations could be due to actual thermal motion or to the limited availability of structural data. In fact, overlap of the amide signals of residues Q(10) and M(11) precluded unambiguous assignment of several NOE cross peaks for these residues. This region was well-defined in Zn(HIV1-F1) (see Figure 6, bottom drawing), even for structures generated with loose distance restraints (Blake et al., 1990). This is probably due to the presence of significantly more NOE data obtained from the NOESY spectrum of Zn(HIV1-F1); in particular, several long-range NOESY cross peaks involving the A(11) methyl group were observed, whereas no similar cross peaks involving the M(11) residue of Zn(HIV1-F2) were observed.

Second, it appears that minor structural differences exist between Zn(HIV1-F1) and Zn(HIV1-F2) for residues at positions 11–14. The most notable differences involve the orientations of backbone atoms of residues D(13) and C(14). Comparison of dihedral ψ and ϕ values reveal that the largest differences occur at residues 13 and 14 (see Figure 7). Note, however, that the structural variations are actually rather minor as indicated in the middle [Zn(HIV1-F2)] and bottom [Zn(HIV1-F1)] drawings of Figure 6.

Conclusions

Except for minor structural differences involving residues 11–13, the backbone folding exhibited by Zn(HIV1-F2) is essentially identical with the folding exhibited by Zn(HIV1-F1). This finding is significant since is demonstrates that, for at least two peptides, the conserved r.t. zinc finger sequence gives rise to a common structural motif. 2D NOESY spectra obtained recently for samples of the HIV-1 nucleocapsid protein, isolated from viral particles and reconstituted with zinc, contain cross-peak patterns remarkably similar to those observed in the NOESY spectra of Zn(HIV1-F1) and Zn(HIV1-F2) (South et al., 1990b). The similarity of ^{1}H NMR chemical shifts and NOESY cross-peak intensities provide solid evidence that conformationally similar structures are formed by the synthetic peptides and the intact protein.

We have yet to determine the reason for the differences in conformational lability of Zn(HIV1-F2) compared to Zn(HIV1-F1). The increased lability of Zn(HIV1-F2) could be due to structural differences involving residues 11–14. Position 11 contains a sterically demanding Met in Zn(HIV1-F2), compared to the sterically smaller Ala in Zn(HIV1-F1). In this regard, it is interesting to note that all but one of the seven r.t. finger sequences in the human CNBP contain an Ala at position 11 (Rajavashisth et al., 1989). It is also significant that narrow ^{1}H NMR signals are observed for residues in the C-terminal r.t. zinc finger domain of the intact HIV-1 NCP. Thus, although the synthetic peptide is conformationally labile at 35 °C, the corresponding C-terminal r.t. zinc finger domain of the intact HIV-1 NCP is conformationally nonlabile under similar sample conditions (South et al., 1990b). This difference cannot be attributed to amino acid differences since the HIV1-F2 sequence matches the relevant sequence in the HIV1-NCP. Preliminary analysis of the 2D NOESY data obtained for Zn_2(HIV1-NCP) suggest that residues immediately following the r.t. zinc finger domains are folded in a manner to facilitate hydrogen bonding between the backbone amide of residue +2 and the Cys(14) sulfur. This interaction was not observed in our peptide structures, possibly due to the presence of the terminal carboxylate at the C-terminal (+2-position) residue.

Finally, the Gly at position 8 is conservatively substituted in retroviral NCPs and is completely conserved in the human CNBP. Recent site-directed mutagenesis experiments reveal that this residue can be replaced by Ala in Rous sarcoma virus without affecting RNA packaging, whereas replacement by a bulkier Val leads to a reduction in packaging by 90% and a 30-fold decrease in infectivity (Dupraz et al., 1990). These findings are consistent with our structural findings, where some flexibility at position 8 is apparent. Structural studies of mutant peptides for correlation with mutagenesis (Meric & Goff, 1989; Dupraz et al., 1990) and metal-binding (Green & Berg, 1989, 1990) studies are in order.

Supplementary Material Available

Figures S1 and S2 showing the relative orientations of the backbone atoms for the group B structures and a portion of the back-calculated spectrum for a representative group B DG structure (3 pages). Ordering information is given on any current masthead page.

Registry No. Zn(HIV1-F2), 133833-03-1.

References

Adman, E., Watenpaugh, K. D., & Jensen, L. H. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4854.

Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511.

Berg, J. M. (1986) *Science 232*, 485.

Berg, J. M. (1990) *J. Biol. Chem. 265*, 6513.
Blake, P. R., Hare, D. R., & Summers, M. F. (1991) *Techniques in Protein Chemistry* (Villafranca, J. J., Ed.) Vol. II, p 357, Academic Press, New York.
Bothner-by, A., Stephens, R. L., Lee, J. T. Warren, C. D., & Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811.
Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 71*, 166.
Copeland, T. D., Morgan, M. A., & Oroszlan, S. (1984) *Virology 133*, 137.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820.
Dupraz, P., Oertle, S., Meric, C., Damay, P., & Spahr, P.-F. (1990) *J. Virol. 64*, 4978.
Green, L. M., & Berg, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4047.
Green, L. M., & Berg, J. M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6403.
Henderson, L. E., Copeland, T. D., Sowder, R. C., Smythers, G. W., & Oroszlan, S. (1981) *J. Biol. Chem. 256*, 8400.
Meric, C., & Goff, S. P. (1989) *J. Virol. 63*, 1558.
Metzler, W. J., Hare, D. R., & Pardi, A. (1989) *Biochemistry 28*, 7045.
Morris, G. A., & Freeman, R. (1978) *J. Magn. Reson. 29*, 433.
Pinantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800.
Rajavashisth, T. B., Taylor, A. K., Andalibi, A., Svenson, K. L., & Lusis, A. J. (1989) *Science 245*, 640.
Roberts, W. J., Pan, T., Elliott, J. I., Coleman, J. E., & Williams, K. R. (1989) *Biochemistry 28*, 10043.
Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327.
Sklenar, V., & Bax, A. (1987) *J. Magn. Reson. 74*, 469.
Skolnick, J. & Kolinski, A. (1990) *Science 250*, 1121.
South, T. L., & Summers, M. F. (1990) in *Advances in Inorganic Biochemistry* (Marzilli, L. G., & Eichhorn, G. L., Eds.) Vol. 8, Elsevier Science Publishing Co., New York.
South, T. L., Kim, B., & Summers, M. F. (1989) *J. Am. Chem. Soc. 111*, 395.
South, T. L., Kim, B., Hare, D. R., & Summers, M. F. (1990a) *Biochem. Pharmacol. 40*, 123.
South, T. L., Blake, P. B., Sowder, R. C., III, Arthur, L. O., Henderson, L. E., & Summers, M. F. (1990b) *Biochemistry 29*, 7786.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.
Summers, M. F. (1991) *J. Cell. Biochem. 45*, 41.
Summers, M. F., South, T. L., Kim, B., & Hare D. R. (1990) *Biochemistry 29*, 329.
Watenpaugh, K. D., Sieker, L. C., & Jensen, L. H. (1979) *J. Mol. Biol. 131*, 509.
Watenpaugh, K. D., Sieker, L. C., & Jensen, L. H. (1980) *J. Mol. Biol. 204*, 483.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zuiderweg, E. R. P., Hallenga, K., & Olejniczak, E. T. (1986) *J. Magn. Reson. 70*, 336.