The Nucleocapsid Protein Isolated from HIV-1 Particles Binds Zinc and Forms Retroviral-Type Zinc Fingers†

Terri L. South,‡ Paul R. Blake,‡ Raymond C. Sowder III,§ Larry O. Arthur,§ Louis E. Henderson,*,§ and Michael F. Summers*,‡

Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21228, and Retroviral Protein Biochemistry Laboratory, AIDS Vaccine Program, Program Resources, Inc., NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21701

Received June 4, 1990; Revised Manuscript Received July 2, 1990

ABSTRACT: The role of zinc in retroviral *gag* protein function has been addressed through the application of high-resolution nuclear magnetic resonance spectroscopy to samples of the nucleocapsid protein (NCP, p7) isolated directly from infectious HIV-1 particles. Unlike reports for the NCP from avian myeloblastosis virus (AMV) particles [Jentoft et al. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7094], we find that the HIV-1 NCP binds 2 equiv of zinc tightly and stoichiometrically. Two-dimensional NMR spectroscopic studies reveal that zinc binding induces formation of folded domains that are conformationally similar to (if not identical with) structures observed previously for relevant retroviral-type (RT) zinc finger peptides [formerly called zinc fingerlike peptides; Summers et al. (1990) *Biochemistry 29*, 329]. This finding is consistent with the hypothesis that the inability of mutant proteins (with substituted Cys and His residues) to package viral RNA results from deficient zinc-binding capability, which may have significant consequences in the development of vaccines for the prevention of AIDS.

All retroviruses encode a *gag* gene product that plays a key role in packaging and assembly (Dickson et al., 1985; Bolognesi et al., 1978). Subsequent to viral budding, *gag* proteins are processed by the retroviral protease to give several major structural proteins including the nucleocapsid protein (NCP; MW ca. 7000–12 000). In mature particles, the NCP apparently stabilizes RNA in the virus core through nonspecific NCP–RNA interactions (Karpel et al., 1987).

Without exception, retroviral NCPs (and their *gag* precursors) contain one or two copies of the conserved sequence Cys-X_2-Cys-X_4-His-X_4-Cys (Henderson et al., 1981; Copeland et al., 1984), which may function physiologically as a zinc-binding domain (Berg, 1986). Indeed, in vitro studies involving synthetic peptides provide firm evidence that zinc is capable of binding tightly and stoichiometrically to these sequences, resulting in formation of unique structures stabilized by extensive internal hydrogen bonding (South et al., 1989; Green & Berg, 1989, 1990; Roberts et al., 1989; Summers et al., 1990a,b). Low zinc dissociation constants (10^{-10}–10^{-12} M) measured for synthetic retroviral-type (RT, see Nomenclature) zinc finger peptides and a synthetic protein indicate that, under normal cellular conditions, these domains should be fully populated with zinc (Green & Berg, 1990). In addition, in vivo site-directed mutagenesis experiments introducing substitutions for the conserved Cys and His residues destroy the ability of the *gag* protein to package viral RNA (Gorelick et al., 1988; Meric & Goff, 1989; Aldovini & Young, 1990; Gorelick et al., 1990). This result is consistent with the proposed zinc-binding function and suggests that RT zinc finger formation within the *gag* protein is essential for RNA packaging and infectivity. On the other hand, it has been reported that avian myeloblastosis virus (AMV) and its NCP apparently do not contain significant levels of zinc and that addition of excess zinc to AMV NCP samples leads to weak zinc binding and does not affect the tertiary structure or nucleic acid binding properties of the AMV NCP (Jentoft et al., 1988). From these results, it has been proposed that the conserved sequences in retroviral NCPs do not form zinc-binding fingers (Jentoft et al., 1988; Katz & Jentoft, 1989).

The inability of mutant *gag* proteins with defective RT zinc finger arrays to package genomic RNA has important consequences in the area of AIDS vaccine development (Aldovini & Young, 1990; Gorelick et al., 1990), and understanding the mechanism of this effect and the biophysical properties of the HIV-1 *gag* protein is essential for further vaccine development and for rational drug design. To determine if the RT zinc finger domains of the HIV-1 *gag* protein are capable of binding zinc stoichiometrically and with significant affinity, high-resolution nuclear magnetic resonance (NMR) studies have been initiated on NCP samples isolated directly from HIV-1 particles. In contrast to results for the AMV NCP, we find that the HIV-1 NCP binds 2 equiv of zinc tightly and stoichiometrically. Furthermore, two-dimensional (2D) nuclear Overhauser effect (NOESY) spectra provide solid evidence that zinc coordination induces the formation of folded domains that are structurally similar to (if not identical with) the structures observed previously for synthetic peptides with HIV-1 NCP RT zinc finger sequences.

Materials and Methods

Protein Isolation and Purification. HIV-1 (MN strain) was grown from H9 cells and harvested by sucrose density gradient centrifugation. NCP samples for NMR studies were then obtained by disruption of HIV-1 particles with an 8 M guanidine hydrochloride solution (pH 8, 1% 2-mercapto-

† This work was supported by American Cancer Society Institutional Research Grant IN-147F (M.F.S.), NIH Grant GM42561-01 (M.F.S.), and the National Cancer Institute, DHHS, under Contract N01-CO-74102 with Program Resources, Inc. The contents of this publication do not necessarily reflect the views or policies of the DHHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

* To whom correspondence should be addressed.
‡ University of Maryland Baltimore County.
§ Frederick Cancer Research and Development Center.

0006-2960/90/0429-7786$02.50/0

ethanol, T = 50 °C, t = 15 min), followed by cooling and addition of 20% TFA (to pH 2). This solution was then applied to an HPLC column (25 × 150 mm, μBondapak C18; Waters, Milford, MA). Proteins were eluted with a gradient of acetonitrile (0.05% TFA, 5.0 mL/min) as described previously (Henderson et al., 1988a,b). Solvents were then removed by lyophilization.

Protein Characterization. The NCP amino acid sequence was determined with an Applied Biosystems Model 470A protein sequencer. Protein concentrations and composition were determined after acid hydrolysis (6 N HCl) by amino acid analysis with a High Performance System 6300 amino acid analyzer equipped with a Systems Gold Computing Analyzer (Beckman, Palo Alto, CA).

NMR Data Collection and Processing. ^{1}H NMR data were obtained with a GE GN-500 (500 MHz, ^{1}H) NMR spectrometer. Data were transferred via Ethernet to a Silicon Graphics Model 220 GTX-B computer, converted to "readable" files with an in-house program (GENET), and processed with FTNMR (Hare Research, Inc.). 2D NOESY data were obtained with the following parameters and sample conditions: T = 30 °C; no sample spinning; 210-ms mixing period, including a 5-ms homogeneity-spoiling pulse; 2 × 266 × 4096 (D_2O) and 2 × 314 × 4096 (90% H_2O/10% D_2O) data matrix sizes (States et al., 1982), zero-filled to a final matrix size of 2048 × 2048 points; 3-Hz (D_2O) and 6-Hz (90% H_2O/10% D_2O) exponential line-broadening filtering in t_2 and 70°-shifted squared sine-bell filtering in t_1. 2D NOESY data for the sample containing 90% H_2O/10% D_2O were obtained with a 1-1-echo read pulse sequence (Sklenar & Bax, 1987), with τ_1 and τ_2 read delay periods of 80 and 180 μs, respectively, and with weak solvent preirradiation during the 1.8-s recycle delay period.

Results and Discussion

Nomenclature. Since there has been some confusion regarding the nomenclature of zinc finger proteins, a brief discussion and proposal is presented here. The term "zinc finger" has evolved from the original paper by Klug and coworkers where the zinc-binding domains in transcription factor IIIA (TFIIIA) were proposed to form "DNA binding fingers" (Miller et al., 1985). In general, zinc finger has been used to describe virtually any relatively short sequence that (i) contains four (or more) Cys and/or His residues and (ii) is believed to function physiologically as a zinc *and* nucleic acid binding domain. To avoid creating further confusion, we will continue to use this working definition of a zinc finger.

Further subclassifications based on sequence homology are proposed. Sequences similar to those observed in TFIIIA are referred to as "classical-type" zinc fingers, and all other sequences are labeled "nonclassical" zinc fingers. Nonclassical zinc finger sequences are further divided according to sequence homology. We refer to the C-X_2-C-X_4-H-X_4-C sequence as a retroviral-type zinc finger since this conserved sequence was identified originally in retroviral *gag* proteins (Henderson et al., 1981; Copeland et al., 1984). Note, however, that several copies of this conserved sequence have also been observed in a *human* cellular nucleic acid binding protein (Rajavashisth et al., 1989). Other nonclassical zinc finger sequences have been identified (Berg, 1990; South & Summers, 1990).

Amino Acid Sequence of HIV-1 NCP. After HPLC purification, a sample of HIV-1 NCP was subjected to amino acid composition and N-terminal amino acid sequence analyses. Edman degradation (53 cycles) enabled direct assignment of all but the last two amino acids in the sequence. The sequence for the last two residues was determined by amino acid composition and gene sequence analysis (Gurgo et al., 1988). The amino acid sequence, size, and cleavage sites for HIV-1 NCP (Figure 1) are consistent with previously published data for HIV-1 and SIV (simian immunodeficiency virus) (Henderson et al., 1988a,b). The results are fully consistent with the sequence predicted via nucleotide sequence analysis.

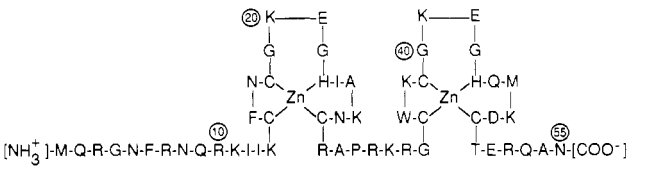

FIGURE 1: Amino acid sequence and zinc coordination mode for the zinc adduct with the HIV-1 nucleocapsid protein (MN strain).

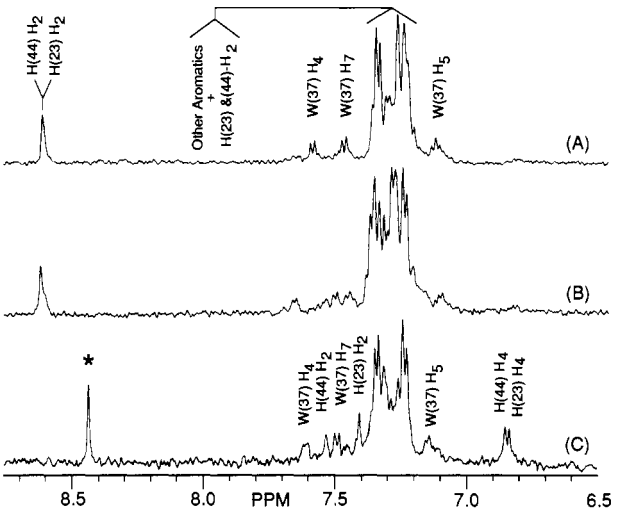

FIGURE 2: ^{1}H NMR spectra obtained for HIV-1 NCP. (Top) A pH 4.3, 600-μg sample; spectrum obtained immediately after dissolution into D_2O (180 μM). (Middle) Spectrum obtained 24 h after the top spectrum was collected (same sample). (Bottom) A 300-μg sample of HIV-1 NCP (pH 7.2, 90 μM) prepared anaerobically in the presence of zinc (1.5 equiv of Zn per finger domain; * = impurity in our $ZnSO_4$ solution). Chemical shifts in the top and bottom spectra are consistent with shifts observed for synthetic zinc finger peptides under similar sample conditions.

NMR and Structural Findings. 1D ^{1}H NMR studies were performed on a small sample of HIV-1 NCP (600 μg) to determine the influence of pH, oxygen, and zinc on the ^{1}H NMR spectral properties (line widths, chemical shifts, etc.). The downfield region of the 1D ^{1}H NMR spectrum obtained for HIV-1 NCP immediately after dissolution in D_2O (0.18 mM, pH* 4.3) is shown in Figure 2A. No efforts were made to exclude oxygen from this sample. Sharp ^{1}H NMR signals are observed at chemical shifts expected for the aromatic residues and protonated histidines. This spectrum is consistent with data obtained under similar conditions for the synthetic RT zinc finger peptides HIV1-F1 [sequence of HIV-1 NCP residues 13–30 (South et al., 1989; Summers et al., 1990)] and HIV1-F2 [sequence of HIV-1 NCP residues 34–51 (South and Summers, in preparation)]. Major changes were observed in the ^{1}H NMR spectrum obtained 24 h later for the same sample, indicating that >50% of the sample had undergone a structural change (Figure 2B). Similar spectral changes have been observed previously in our laboratory for the synthetic finger peptides and are consistent with formation of cysteine disulfide bonds and polymerization.

A second sample of the HIV-1 NCP (300 μg) was obtained subsequently, and this sample was stored initially under inert atmosphere. The 1D ^{1}H NMR spectrum obtained on disso-

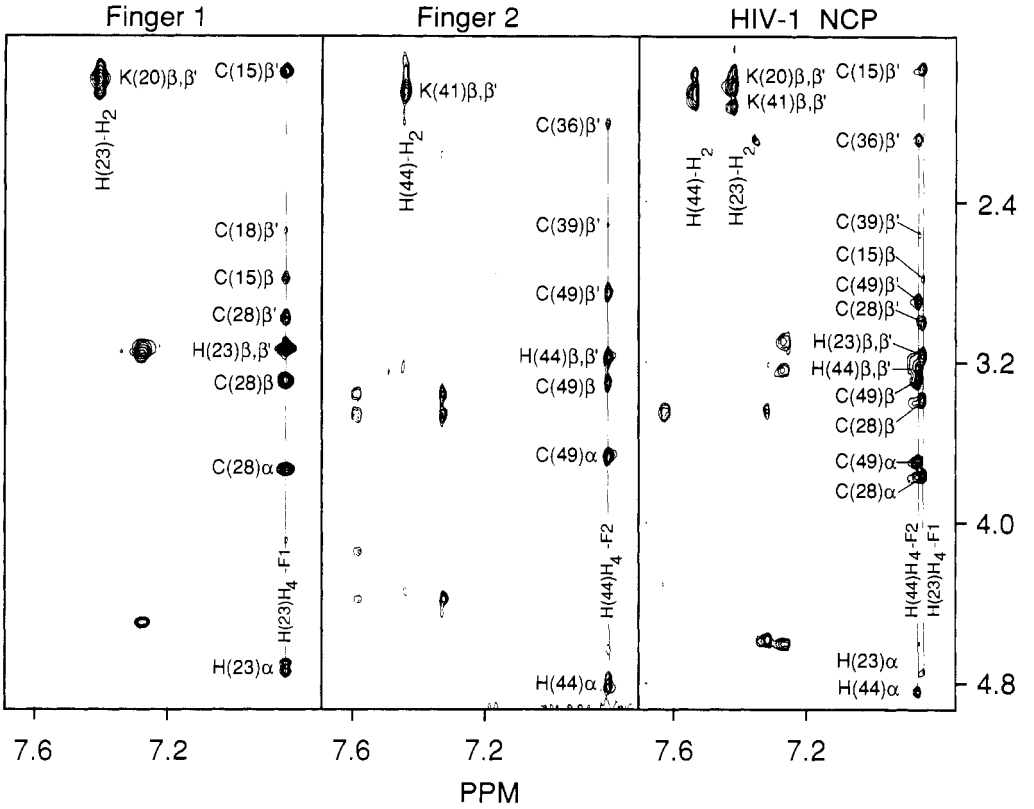

FIGURE 3: Portions of the 2D NOESY spectra (D_2O solutions) obtained for Zn(HIV1-F1) (left: T = 30 °C; τ_m = 300 ms), Zn(HIV1-F2) (middle: T = −2 °C; τ_m = 300 ms), and Zn_2(HIV1-NCP) (right: T = 30 °C; τ_m = 210 ms). Cross peaks were assigned via standard procedures (Wuthrich, 1986).

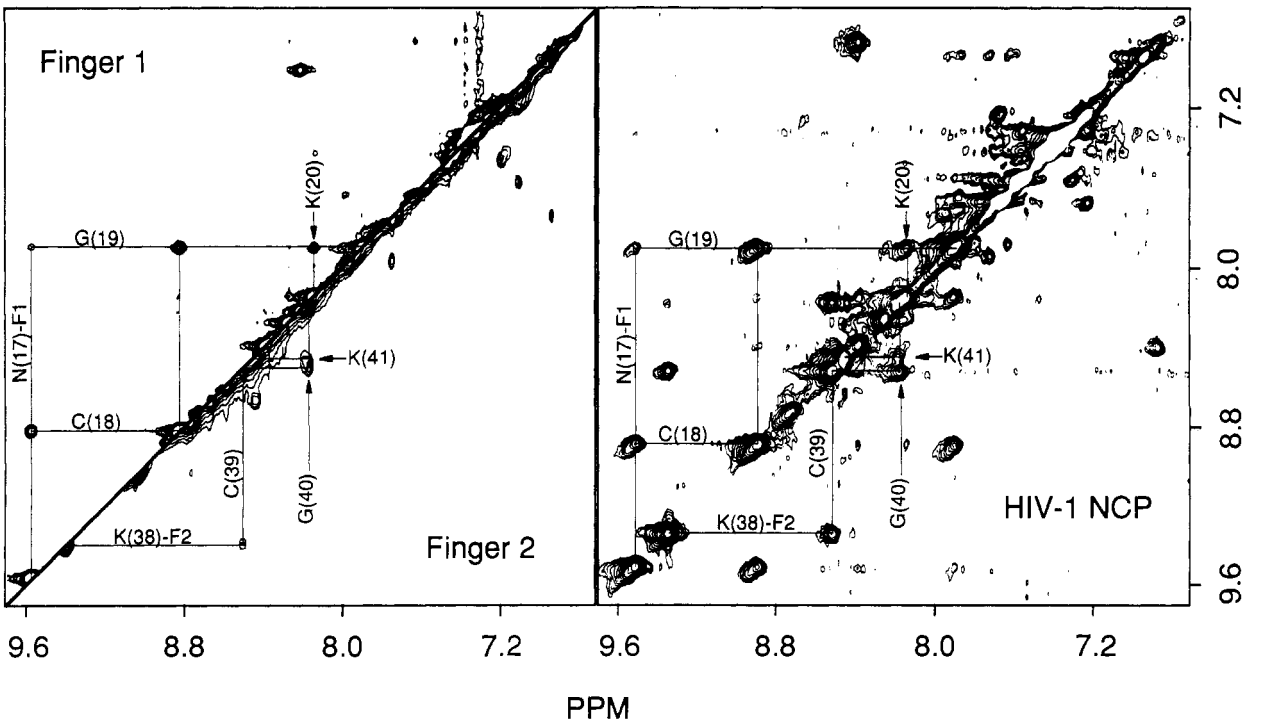

FIGURE 4: Portions of the 2D NOESY spectra obtained for Zn(HIV1-F1) (15 mM; left spectrum above the diagonal), Zn(HIV1-F2) (9.5 mM; left spectrum below the diagonal), and Zn_2(HIV1-NCP) (0.6 mM; right spectrum) in 90% H_2O/10% D_2O solution, pH 7. Data were obtained with presaturation of the water signal during the recovery period, resulting in loss of signals due to moderately and rapidly exchanging amide protons.

lution into oxygen-free D_2O (0.4 mL; 90 μM NCP) was identical with the spectrum shown in Figure 2A. Zinc was then added to this sample (in slight excess of 1:1 equiv of zinc per finger domain or 2:1 equiv of zinc per NCP) and the solution pH adjusted under inert atmosphere to pH 7. Signals in the downfield region of the spectrum obtained for this sample (Figure 2C) are at chemical shifts similar to the shifts observed for synthetic peptides Zn(HIV1-F1) and Zn(HIV1-F2). In fact, superposition of the relevant regions of the ^1H NMR spectra of Zn(HIV1-F1) and Zn(HIV1-F2) (not shown) nearly reproduces the spectrum obtained for HIV-1 NCP in the presence of zinc (Figure 2C). Unlike the apo-

NCP, the Zn_2(HIV1-NCP) is extremely stable in the presence of oxygen (as also observed for the RT zinc finger peptides). Thus, after a 3-month period, the above sample (aqueous, pH 7, T = 22–30 °C, no effort to exclude oxygen) had not undergone significant oxidation or decomposition.

To determine if the zinc adduct of HIV-1 NCP contains appreciable secondary structure, 2D nuclear Overhauser effect (NOESY) spectra were obtained for samples in both D_2O (0.3 mM HIV-1 NCP, 0.9 mM $ZnSO_4$, pH* 7.0, 30 °C) and 90% H_2O/10% D_2O (0.6 mM HIV-1 NCP, 1.8 mM $ZnSO_4$, pH 7.0, 30 °C) solutions. Signal assignments were made for all but a few side-chain protons located within the RT zinc finger domains (residues 13–30 and 35–49) of the HIV-1 NCP, using a combination of double-quantum-filtered homonuclear correlated (2QF-COSY), ^1H–^{13}C heteronuclear multiple-quantum coherence (HMQC), and homonuclear Hartmann–Hahn (HOHAHA) spectroscopies (Wuthrich, 1989) (signal assignments will be published elsewhere).

The "fingerprint region" of the 2D NOESY spectrum obtained for HIV-1 NCP in D_2O solution is shown in Figure 3. For comparison, relevant portions of the 2D NOESY spectra obtained for zinc adducts with 18 amino acid RT zinc finger peptides are included in Figure 3. In this region of the NOESY spectrum, cross peaks diagnostic for retroviral-type zinc fingers are observed; specifically, all RT zinc finger peptides examined in our laboratory exhibit cross peaks between the His-H^4 aromatic proton and the side-chain protons of Cys(1) (NOE to β_R; spin diffusion to β_S), Cys(3) (NOE to β_R), and Cys(14) (NOE to α, β_R; spin diffusion to β_S) (Figure 3). Similar cross peaks are clearly visible in the NOESY spectrum obtained for Zn_2(HIV1-NCP) (Figure 3). In addition, most of the cross peaks observed for Zn_2(HIV1-NCP) have chemical shifts and intensities similar to the shifts and intensities observed for the isolated finger peptides (Figure 3).

Amide-to-amide NOE connectivities observed for Zn_2-(HIV1-NCP) are also consistent with data obtained for Zn-(HIV1-F1) and Zn(HIV1-F2). Thus, both synthetic peptides exhibit sequential amide-to-amide cross peaks for amino acid residues within the "rubredoxin-like" portions of the finger units [residues N(17)–K(20) and K(38)–K(41); Summers et al., 1990] (Figure 4). As shown in Figure 4, the zinc adduct with HIV-1 NCP exhibits amide-to-amide cross peaks and intensities that are remarkably similar to the data obtained for the relevant RT zinc finger peptides. *The presence of diagnostic zinc finger cross-peak patterns and the striking similarity of chemical shifts to those observed for the synthetic peptides provide hard evidence that the HIV-1 NCP binds zinc and forms RT zinc finger structures that are conformationally similar to (if not identical with) structures observed for the synthetic peptides.*

For Zn(HIV1-F2), sharp ^1H NMR signals were observed only after cooling the sample below ca. 5 °C, apparently due to increased structural lability of Zn(HIV1-F2) compared to Zn(HIV1-F1) (South and Summers, in preparation). In contrast, the corresponding RT zinc finger domain within the HIV-1 NCP exhibits sharp, well-resolved signals at temperatures as high as 30 °C (Figures 2–4). Additional experiments will be necessary to explain the apparent increase in structural stability associated with the second RT zinc finger when located within the NCP.

Preliminary atomic absorption experiments in our laboratory indicate that zinc copurifies with mature HIV-1 particles. Although further studies are necessary, this finding raises the possibility that zinc may be associated with the NCP in virus particles, in addition to functioning in retroviral gene recognition at the *gag* protein level. Experiments aimed at quantitatively assessing the zinc content of the NCP within mature virus particles are in progress.

References

Aldovini, A., & Young, R. A. (1990) *J. Virol. 64*, 1920.

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

Berg, J. M. (1990) *J. Biol. Chem. 265*, 6513.

Bolognesi, D. P., Montelaro, R. C., Frank, H., & Schafer, W. (1978) *Science 199*, 183.

Copeland, T. D., Morgan, M. A., & Oroszlan, S. (1984) *Virology 133*, 137.

Dickson, C., Eisenman, R., Fan, J., Hunter, E., & Teich, N. (1985) in *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H., & Coffin, J., Eds.) 2nd ed., Vols. 1 and 2, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Gorelick, R. J., Henderson, L. E., Hanser, J. P., & Rein, A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8420.

Gorelick, R. J., Nigida, S. M., Jr., Bess, J. W., Jr., Arthur, L. O., Henderson, L. E., & Rein, A. (1990) *J. Virol. 64*, 3207.

Green, L. M., & Berg, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4047.

Green, L., & Berg, J. M. (1990) *Proc. Natl. Acad. Sci. U.S.A.* (in press).

Gurgo, C., Guo, H.-G., Franchini, G., Aldovini, A., Collalti, E., Farrell, K., Wong-Staal, F., Gallo, R. C., & Reitz, M. S., Jr. (1988) *J. Virol. 164*, 531.

Henderson, L. E., Copeland, T. D., Sowder, R. C., Smythers, G. W., & Oroszlan, S. (1981) *J. Biol. Chem. 256*, 8400.

Henderson, L. E., Copeland, P. D., Sowder, R. C., Schultz, A. M., & Oroszlan, S. (1988a) *UCLA Symp. Mol. Cell. Biol. 71*, 135.

Henderson, L. E., Benveniste, R. E., Sowder, R., Copeland, T. D., Schultz, A. M., & Oroszlan, S. (1988b) *J. Virol. 62*, 2587.

Jentoft, J. E., Smith, L. M., Fu, X., Johnson, M., & Leis, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7094.

Karpel, R. L., Henderson, L. E., & Oroszlan, S. (1987) *J. Biol. Chem. 262*, 4961.

Katz, R. A., & Jentoft, J. E. (1989) *BioEssays 11*, 176.

Meric, C., & Goff, S. P. (1989) *J. Virol. 63*, 1558.

Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609.

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.

Sklenar, V., & Bax, A. (1987) *J. Magn. Reson. 74*, 469.

South, T. L., & Summers, M. F. (1990) in *Advances in Inorganic Biochemistry* (Marzilli, L. G., & Eichhorn, G., Eds.) Vol. 8, Elsevier Science Publishing Co., New York, NY.

South, T. L., Kim, B., & Summers, M. F. (1989) *J. Am. Chem. Soc. 111*, 395.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.

Summers, M. F., South, T. L., Kim, B., & Hare, D. R. (1990a) *Biochemistry 29*, 329.

Summers, M. F., South, T. L., Kim, B., & Hare, D. R. (1990b) *Biochem. Pharmacol. 40*, 123.

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

Wuthrich, K. (1989) *Science 243*, 45.