Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1

MICHAEL F. SUMMERS,¹ LOUIS E. HENDERSON,² MARK R. CHANCE,³
JULIAN W. BESS, JR.,² TERRI L. SOUTH,¹ PAUL R. BLAKE,¹ IRIT SAGI,³
GABRIELA PEREZ-ALVARADO,¹ RAYMOND C. SOWDER III,²
DENNIS R. HARE,⁴ AND LARRY O. ARTHUR²

¹ Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21228

² AIDS Vaccine Program, Program Resources Inc./DynCorp, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21701

³ Department of Chemistry, Georgetown University, 37th and "O" St. N.W., Washington, D.C. 20057

⁴ Hare Research, Inc., 14810 216th Ave. N.E., Woodinville, Washington 98072

(RECEIVED February 5, 1992; ACCEPTED February 11, 1992)

Abstract

All retroviral nucleocapsid (NC) proteins contain one or two copies of an invariant 3Cys-1His array (CCHC = C-X₂-C-X₄-H-X₄-C; C = Cys, H = His, X = variable amino acid) that are essential for RNA genome packaging and infectivity and have been proposed to function as zinc-binding domains. Although the arrays are capable of binding zinc in vitro, the physiological relevance of zinc coordination has not been firmly established. We have obtained zinc-edge extended X-ray absorption fine structure (EXAFS) spectra for intact retroviruses in order to determine if virus-bound zinc, which is present in quantities nearly stoichiometric with the CCHC arrays (Bess, J.W., Jr., Powell, P.J., Issaq, H.J., Schumack, L.J., Grimes, M.K., Henderson, L.E., & Arthur, L.O., 1992, *J. Virol.* 66, 840–847), exists in a unique coordination environment. The viral EXAFS spectra obtained are remarkably similar to the spectrum of a model CCHC zinc finger peptide with known 3Cys-1His zinc coordination structure. This finding, combined with other biochemical results, indicates that the majority of the viral zinc is coordinated to the NC CCHC arrays in mature retroviruses. Based on these findings, we have extended our NMR studies of the HIV-1 NC protein and have determined its three-dimensional solution-state structure. The CCHC arrays of HIV-1 NC exist as independently folded, noninteracting domains on a flexible polypeptide chain, with conservatively substituted aromatic residues forming hydrophobic patches on the zinc finger surfaces. These residues are essential for RNA genome recognition, and fluorescence measurements indicate that at least one residue (Trp³⁷) participates directly in binding to nucleic acids in vitro. The NC is only the third HIV-1 protein to be structurally characterized, and the combined EXAFS, structural, and nucleic acid-binding results provide a basis for the rational design of new NC-targeted antiviral agents and vaccines for the control of AIDS.

Keywords: HIV-1; NMR structure; nucleocapsid protein; zinc finger

During retrovirus assembly and budding, several thousand copies of a gag precursor protein participate in the recognition and packaging of the RNA genome (Bolognesi et al., 1978; Dickson et al., 1985). Subsequent to

budding, the gag precursors of immature virions are cleaved by the viral protease to give several smaller structural proteins, including the nucleocapsid (NC) proteins. Mature particles contain approximately 2,500–3,000 cop-

Reprint requests to Michael F. Summers, Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21228.

Abbreviations: NC, nucleocapsid; HIV, human immunodeficiency virus; SIV, simian immunodeficiency virus; EIAV, equine infectious anemia virus; MMTV, mouse mammary tumor virus; EXAFS, extended X-ray absorption fine structure spectroscopy; NMR, nuclear magnetic

resonance; AIDS, acquired immunodeficiency syndrome; NOESY, nuclear Overhauser effect spectroscopy; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; 2QF-COSY, double quantum-filtered homonuclear correlated spectroscopy; HTLV-I, human T-cell leukemia virus type-I; DG, distance geometry; MoMuLV, Moloney murine leukemia virus; RMSD, root mean square deviation; HPLC, high performance liquid chromatography.

ies of the NC protein that apparently function by stabilizing RNA through formation of a ribonucleoprotein complex (Smith & Bailey, 1979; Henderson et al., 1981; Copeland et al., 1984; Meric et al., 1984; Karpel et al., 1987; Meric & Goff, 1989; Aldovini & Young, 1990; Dupraz et al., 1990; Gorelick et al., 1990). Without exception, retroviral NC proteins (and their gag precursors) contain one or two copies of an invariant CCHC array, Cys-X₂-Cys-X₄-His-X₄-Cys (X = variable or conservatively substituted amino acid) (Henderson et al., 1981), that are essential for RNA genome packaging and infectivity (Gorelick et al., 1988, 1990; Meric & Goff, 1989; Aldovini & Young, 1990) and have been proposed to function as zinc-binding domains (Berg, 1986). Although the arrays are capable of binding zinc in vitro (Green & Berg, 1989, 1990; Roberts et al., 1989; South et al., 1989, 1990a,b, 1991; Summers et al., 1990; Fitzgerald & Coleman, 1991; Omichinski et al., 1991), the physiological relevance of this CCHC zinc finger motif (previously referred to as a retroviral-type or zinc finger-like motif) has not been firmly established. In fact, it has been reported that avian myeloblastosis virus (AMV) particles lack significant amounts of zinc (Jentoft et al., 1988), which has led to the suggestion that the conserved arrays do not function as zinc binding fingers in mature viruses (Jentoft et al., 1988; Katz & Jentoft, 1989).

In contrast to reports for AMV, it was determined recently that zinc copurifies and concentrates with HIV-1 particles, and quantitative analysis of the zinc and NC protein contents revealed tightly bound zinc in a ratio of ca. 1.7 Zn atoms per NC protein, corresponding to approximately one equivalent of zinc per CCHC array (Bess et al., 1992). Similar results were also obtained for a number of other retroviruses, including human T-cell leukemia virus (HTLV-I), mouse mammary tumor virus (MMTV), and equine infectious anemia virus (EIAV, a lentivirus) (Bess et al., 1992). The fact that tightly bound viral zinc is present in quantities nearly stoichiometric with the viral CCHC arrays suggests that the arrays may be populated with zinc in mature virions. We have tested this hypothesis by performing zinc-edge extended X-ray absorption fine structure (EXAFS) experiments on intact retrovirus particles. The viral EXAFS spectra obtained are remarkably similar to the spectrum of a synthetic CCHC zinc finger peptide, providing strong direct evidence that zinc is indeed coordinated to the NC CCHC arrays in mature retroviruses. In light of these findings, we have extended our NMR studies of the HIV-1 NC protein and have determined its three-dimensional solution-state structure by hybrid distance geometry (DG) and simulated annealing computational methods. In addition, nucleic acid interactive studies based on fluorescence quenching experiments have been performed in order to gain insights into the nature of retroviral NC-RNA interactions. The NC protein is only the third HIV-1 protein

to be structurally characterized (X-ray structures of the protease [Navia et al., 1989; Wlodawer et al., 1989] and the RNase domain of reverse transcriptase [Davies et al., 1991] have been reported) and is the only protein for which a structure has been determined using a sample isolated directly from HIV-1 particles.

Results and discussion

EXAFS spectra of intact retroviruses and the relevance of CCHC zinc fingers

The collection of EXAFS data requires physical handling of highly concentrated viral samples under non-BL3 conditions. As such, only nonhuman pathogenic retroviruses were chosen for EXAFS studies. EIAV was selected as a representative of the lentivirus class and MMTV as a representative of other classes of retroviruses.

EXAFS spectra obtained for MMTV and EIAV particles were evaluated on the basis of their consistence with each other and with EXAFS data obtained for a model CCHC zinc finger peptide, Zn(HIV1-F1) (Chance, M., pers. comm.). This peptide serves as an important model system, as the bound zinc atom is known to exist in a 3Cys-1His tetrahedral environment (South et al., 1989; Summers et al., 1990). Several other peptides and proteins that contain the CCHC motif have been shown by UV-visible and NMR spectroscopic methods to bind zinc via 3Cys and 1His residues, forming a tetrahedral metal coordination site (Green & Berg, 1989, 1990; Roberts et al., 1989; Fitzgerald & Coleman, 1991; South et al., 1991). Although nonlinear least-squares fitting of the Fourier-filtered (first coordination sphere) EXAFS data obtained for HIV1-F1 afforded several solutions that are consistent with the EXAFS data alone (Chance, M., pers. comm.), the best chemically reasonable fit of the peptide EXAFS spectrum was obtained with a 3S-1N (or O; backscattering of N and O atoms cannot be readily differentiated by EXAFS data alone) combination, with the three sulfurs at an average distance of 2.35 Å and the nitrogen at 2.05 Å relative to zinc (Chance, M., pers. comm.). Thus, the EXAFS data for the HIV1-F1 model peptide are fully consistent with the previously established 3Cys-1His tetrahedral zinc coordination environment.

The wave vector-cubed, background-subtracted EXAFS data obtained for EIAV are shown at the top of Figure 1. Similar data were obtained for MMTV (not shown). The Fourier-filtered, back-transformed spectrum of EIAV is compared with the HIV1-F1 model peptide spectrum in the bottom of Figure 1, and the Fourier-transformed EXAFS spectra of EIAV, MMTV, and HIV1-F1 are plotted together in Figure 2. A small phase difference between the spectra of EIAV and HIV1-F1 (Fig. 1) is observed, which reflects a slightly shorter Zn-

Fig. 1. Wave vector-cubed, background-subtracted EXAFS spectrum obtained for intact EIAV particles (top) and Fourier-filtered back-transformed EXAFS spectra obtained for EIAV and for the synthetic model CCHC zinc finger peptide, Zn(HIV1-F1) (bottom). K-space windows for the Fourier transforms were from 0 to 12.5 Å⁻¹, R-space windows for the Fourier filter were from 1.0 to 2.6 Å. The EIAV spectrum is very similar to the Zn(HIV1-F1) spectrum, requiring only slight (0.05 Å) adjustments of the Zn-ligand distances to achieve an optimal fit (see text for details).

ligand distance (ca. 0.05 Å) in the virus sample. A greater apparent disorder for the peptide may account for the lack of a characteristic histidine shoulder in the peptide EXAFS spectrum. Indeed, NMR studies suggest that the charge of the C-terminal residue of the peptide may inhibit the formation of an amide-to-sulfur hydrogen bond that was observed in the intact protein but not in the peptide (see below). The loss of an NH-S hydrogen bond could also explain the slightly longer metal-ligand distances observed in the peptide EXAFS spectrum. Further studies with peptides that contain additional residues at the C-terminus are planned in order to address this point.

Although small differences exist, agreement between the back-transformed EXAFS spectra of the retroviruses and peptide is extremely good, especially when considering the different nature of the samples being compared (i.e., a purified synthetic peptide and preparations of intact retroviruses). As shown in Figure 2, the Fourier-

Fig. 2. Fourier transforms of the EXAFS spectra obtained and processed under identical conditions for an NC zinc finger peptide, Zn(HIV1-F1), and for intact EIAV and MMTV particles. The spectra are remarkably similar and feature a main peak at ca. 1.8 Å (not corrected for the phase shifts of the backscattering atoms) and a shoulder at a radial coordinate of ca. 1.2 Å (due to truncation of a finite data set). The spectra are fully consistent with the 3S-1N Zn coordination environment established by high-resolution structural studies of Zn(HIV1-F1) (Summers et al., 1990).

transformed EXAFS spectra of the retrovirus and peptide samples are also remarkably similar to each other. Furthermore, the integrated intensities of the EXAFS absorption edges are in agreement with the viral zinc concentrations established by atomic absorption analysis. *The striking similarity of the retrovirus and peptide EXAFS spectra provides direct evidence in support of the proposal that the majority of zinc in intact retrovirus particles is coordinated to NC CCHC zinc finger arrays.*

Other combinations of 3Cys and 1His ligands from different retroviral proteins could, in principle, give rise to the observed retrovirus EXAFS spectra. For example, the capsid (CA) and matrix (MA) proteins of EIAV and MMTV contain sufficient numbers of Cys and His residues to comprise a 3Cys/1His combination. In addition, the EXAFS data alone do not rule out the possibility of a heterogeneous mixture of proteins containing, for example, 75% of the zinc in a 4Cys environment and 25% of the zinc in a 4His site. However, these possibilities seem unlikely for several reasons. First, the Cys and His residues of other viral proteins do not occur at conserved positions. As such, it would be difficult to rationalize the

highly similar viral EXAFS spectra and the stoichiometric (with respect to the CCHC arrays) viral zinc contents if the zinc was actually bound to completely different protein structures in different retroviruses. Second, zinc-blotting experiments (Schiff et al., 1988) demonstrate that the NC protein is the only component of the HIV-1 preparations with significant affinity for zinc (Bess et al., 1992). In fact, zinc dissociation constants in the range of 10^{-10} – 10^{-12} M have been measured for synthetic NC zinc finger arrays, indicating that the arrays should be populated with zinc under normal cellular conditions (Green & Berg, 1990). These findings, together with site-directed mutagenesis (Gorelick et al., 1988; Meric & Goff, 1989) and structural results (Summers et al., 1990; South et al., 1991) provide what is now overwhelming evidence for the physiological relevance of the CCHC zinc finger motif in retroviruses. CCHC zinc finger sequences have been observed recently in nucleic acid binding proteins from *Xenopus* (Sato & Sargent, 1991) and humans (Rajavashisth et al., 1989) and may thus provide a generic structural motif for proteins that bind single-stranded nucleic acids (Summers, 1991).

Solution-state NMR studies of the HIV-1 NC protein

Because there appears to be some confusion in the literature regarding the exact composition of the HIV-1 NC protein, a brief discussion is presented here. Although initial reports indicated that this protein is similar in molecular weight to the HTLV-1 NC protein (ca. 15 kDa), more recent studies have shown that the protein is actually 55 residues in length and has a molecular weight of 6,430 Da (Henderson et al., 1987a, 1992). Highly efficient proteolysis of the gag precursor affords six smaller proteins that are (from N- to C-terminus): p17(MA)-p24(CA)-p2-p7(NC)-p1-p6 (Henderson et al., 1992). Homologous cleavage patterns have also been observed for the gag proteins of the lentiviruses HIV-2 (Henderson et al., 1990), SIV (Henderson et al., 1988, 1990), and EIAV (Henderson et al., 1987b). A 15-kDa HIV-1 gag-product protein such as the one engineered for ^{113}Cd NMR investigations (Fitzgerald & Coleman, 1991) could exist as a transient intermediate in the maturation pathway, although no such protein has been detected thus far in mature particles.

The fingerprint region of the 2D NOESY spectrum obtained for the HIV-1 NC protein is shown in Figure 3. Other diagnostic regions of the 2D NOESY data have been published elsewhere (South et al., 1990a). In general, cross peak patterns and chemical shifts for protons associated with residues Lys¹⁴–Cys²⁸ and Gly³⁵–Cys⁴⁹ of the N- and C-terminal CCHC zinc fingers are similar to data obtained for synthetic CCHC zinc finger peptides with similar amino acid sequences (Summers et al., 1990; South et al., 1991). The ^1H NMR chemical shift assign-

ments and representative spectra are included in the supplementary materials. Important NOE data observed for HIV-1 NC but *not* for the synthetic peptides include (1) strong cross peaks involving the Cys²⁸-H α -Ala³⁰NH and Cys⁴⁹-H α -Thr⁵⁰NH proton pairs, (2) cross peaks from the Ala³⁰H α proton to Pro³¹H δ , δ' , and (3) cross peaks from Ala³⁰CH₃ to the side chain amide of Asn¹⁷, to Cys¹⁸H α , to the β -protons of Cys¹⁸, and to the δ -protons of Pro³¹. In addition, narrow, well-resolved ^1H NMR signals were observed in spectra collected at 30 °C for residues Gly³⁵–Cys⁴⁹, indicating the presence of a dynamically nonlabile structure. This contrasts with the corresponding synthetic peptide, Zn(HIV1-F2), which gave rise to broad signals indicative of conformational lability at temperatures above ca. 5 °C (South et al., 1991). No NOE cross peaks were observed between residues of the N- and C-terminal zinc fingers. Selected sequential and long-range NOE connectivities are summarized in Figure 4.

In contrast to the structured CCHC domains, NMR data indicate that residues Met¹-Ile¹³, Arg³²-Arg³⁴, and Arg⁵²-Asn⁵⁵ do not exist in a unique solution conformation. Thus, for data collected at 30 °C, pH 7.2, and with solvent presaturation, no NH signals were detected for these residues due to rapid chemical exchange with the saturated H₂O protons. This rapid chemical exchange indicates that the backbone amide protons are not involved in significant hydrogen bonding. Backbone NH proton signals were observed for these residues in NOESY data collected at 0.0 °C, where the proton exchange rate is retarded. However, the additional NH signals exhibited poor chemical shift dispersion ($\Delta\delta$ ca. 1.2 ppm), which is also indicative of a random coil conformation (Wuthrich, 1976, 1986). In addition, whereas the aliphatic regions of the two NMR spectra (Supplementary materials) exhibited good chemical shift dispersion typical of a globular protein for signals associated with the CCHC domains, the ^1H NMR signals associated with residues Met¹-Ile¹³, Arg³²-Arg³⁴, and Arg⁵²-Asn⁵⁵ exhibited poor chemical shift dispersion, with chemical shift values typical of a random coil conformation.

Three-dimensional structure of the HIV-1 NC protein

A total of 191 distance restraints derived from the NOE data and 15 NH-S and NH-O hydrogen bond restraints identified on the basis of reduced NH exchange rates (see Materials and methods) and characteristic NOE cross peak patterns were employed for structural modeling (Table 1). The complete set of distance restraints used for structural modeling are listed in the Supplementary materials and have been submitted to the Brookhaven Protein Data Bank along with the structural coordinates. Good convergence was achieved for residues Cys¹⁵–Ala³⁰ and Cys³⁶–Glu⁵¹, with pairwise root mean square

Fig. 3. Fingerprint region of the 500-MHz two-dimensional NOESY spectrum obtained for a 1.7-mg sample of the HIV-1 NC protein isolated from disrupted virus particles under denaturing conditions and reconstituted with zinc ($\tau_m = 210$ ms). The NOESY spectrum required 5 days of continuous data collection. Signals for the NH protons of the CCHC residues only were observed; NH signals for residues outside the zinc finger domains were not observed due to rapid exchange with presaturated solvent protons.

deviation (RMSD) values in the ranges 0.16–0.87 and 0.38–1.6 Å, respectively (see Fig. 5; Kinemages 2, 3; Table 1). The degree of convergence is lower than that obtained in the high-resolution structural studies of the synthetic peptides due to (1) the reduced amount of NOE

Table 1. Structural restraints and refinement data for the HIV-1 NC protein

Number of amino acid residues: 55
Molecular weight: 6,430
Sequence: M-Q-R-G-N-F-R-N-Q-R-K-I-I-K-C-F-N-C-G-K-E-G-H-I-A-K-N-C-R-A-P-R-K-R-G-C-W-K-C-G-K-E-G-H-Q-M-K-D-C-T-E-R-Q-A-N
Number of distance geometry structures: 20
Total number of distance restraints: 224
Intraresidual NOEs: 94
Sequential NOEs ($i, i + 1$): 55
Medium range NOEs ($1 < i - j \leq 5$): 27
Long range NOEs ($ i - j > 5$): 15
Restraints to zinc: 18
Hydrogen bond restraints: 15
Dihedral angles: None
Penalties for DG structures 1–20: 0.075–0.095 Å ²
Largest single bounds violation: 0.063 Å
Convergence for backbone atoms of HIV1–NCP (RMSD, Å)
C(15)–H(23): 0.079–0.541
C(15)–C(28): 0.084–0.623
C(15)–A(30): 0.157–0.865
C(36)–H(44): 0.159–0.976
C(36)–C(49): 0.256–1.43
C(36)–E(51): 0.375–1.61
Comparison of the folding observed for synthetic peptides:
C(3)–C(16) of Zn(HIV1–F1) onto C(15)–C(28) of Zn ₂ (HIV1–NCP): 0.413–0.635 Å
C(3)–C(16) of Zn(HIV1–F2) onto C(36)–C(49) of Zn ₂ (HIV1–NCP): 0.371–1.11 Å

data available for the intact protein, which was the result of substantially reduced sensitivity (protein concentration is 10-fold lower than that of the peptides) and (2) the reduced amount of NOE data for the C-terminal CCHC domain due to the increased NH-proton exchange rate at $T = 30^\circ\text{C}$. However, the folding of these residues is essentially identical to that observed for the synthetic peptides, with pairwise RMSD values for residues Cys¹⁵–Cys²⁸ and Cys³⁶–Cys⁴⁹ of the NC protein superpositioned with relevant residues of the peptides in the ranges 0.41–0.64 and 0.37–1.1 Å, respectively (Table 1). In addition, residues Cys²⁸–Ala³⁰ and Cys⁴⁹–Glu⁵¹ are folded to allow type-I NH-S tight turns. Similar folding was not observed in the peptides, possibly due to the negative charges of the carboxy-terminal Ala and Glu residues in the peptides. The lack of additional NH-S hydrogen bonding may explain the increased stability of the C-terminal zinc finger of the intact NC protein compared to that of the synthetic peptide, Zn(HIV1–F2).

Residues Met¹–Ile¹³, Arg³²–Arg³⁴, and Arg⁵²–Asn⁵⁵ did not converge to unique conformations due to the absence of NOE-derived distance information; superposition of these residues afforded pairwise RMSD values in the ranges 1.1–7.1, 0.29–1.7, and 0.42–2.0 Å, respectively. As described above, two-dimensional NMR data obtained at 30 °C and at 0 °C provide solid evidence that these segments of the protein exist in a random coil conformation.

The structural and dynamic properties of the HIV-1 NC protein in solution are summarized in the C α drawing of Figure 6 and Kinemage 1, where blue and red colors denote the conformationally structured zinc finger domains and the labile N- and C-terminal “tails” and “linker” segments, respectively. A space-filling representation is shown in Figure 7 to allow visualization of conservatively substituted aromatic and hydrophobic res-

Fig. 4. Scheme showing the amino acid sequence and selected sequential and long-range NOE connectivities observed for the HIV-1 NC protein. Filled and open circles reflect the relative sensitivity of observable amide proton signals to solvent presaturation (low and moderate-to-high, respectively; determined by qualitative evaluation of relative signal intensities). The symbols d_{NN} , $d_{\alpha N}$, and $d_{\beta N}$ represent NOESY connectivities involving amide protons of residue i and the αH , NH , and βH protons of residue $i + 1$; connectivities involving residues separated by one or more amino acids are denoted with horizontal lines. For sequential connectivities, NOE intensities (classified as strong, medium, or weak) are indicated by the thickness of the horizontal lines. Secondary structural elements are summarized at the bottom of the figure.

idues (see below). The protein may be best described as a flexible polypeptide that contains two structurally independent, conformationally nonlabile CCHC zinc finger domains.

As observed previously in studies with synthetic peptides, the CCHC zinc finger domains are stabilized by extensive internal NH-O and NH-S hydrogen bonding and also by the conservative placement of sterically nonde-

Fig. 5. Superposition of residues Cys¹⁵-Cys²⁸ (bottom) and Cys³⁶-Cys⁴⁹ (top) of the 30 HIV-1 NC DG structures. For clarity, only the backbone atoms of residues Lys¹¹-Gln⁵³ are shown in the figure. This figure illustrates the finding that only the CCHC arrays exist in a unique solution-state conformation.

Fig. 6. Structure of the HIV-1 NC protein displayed as a $C\alpha$ tube tracing of diameter 1 Å. The red and blue colors denote the conformationally labile N- and C-terminal tails and linker segments and the structured CCHC zinc finger domains, respectively. The figure was rendered on a Silicon Graphics computer with the Ray-T software package (T. Palmer, Cray Research Inc.).

manding glycine residues at positions immediately following and preceding the second cysteine and histidine residues, respectively (Summers et al., 1990). The backbone folding is unique compared to the folding observed for classical CCHH zinc fingers from transcription factor proteins (Parraga et al., 1988; Lee et al., 1989; Omichinski et al., 1990; Kochoyan et al., 1991; Pavletich & Pabo, 1991) and for other nonclassical zinc fingers such as those found in steroid hormone receptor proteins (Hard et al., 1990; Schwabe et al., 1990; Luisi et al., 1991) and in GAL-4 (Gadhavi et al., 1990, 1991; Gardner et al., 1991). The CCHC zinc finger must also be structurally different from the zinc binding domain of

gene 32 protein, which also binds zinc via a 3Cys-1His combination (Giedroc et al., 1992).

Structural implications

NC CCHC zinc finger arrays contain conservatively substituted aromatic and hydrophobic residues that are essential for proper RNA genome recognition and packaging. Mutagenesis of the aromatic Tyr and Trp residues in the single 3Cys-1His array of MoMuLV (posi-

Fig. 7. Space-filling model of a representative HIV-1 NC protein DG structure; note that residues of the N- and C-terminal tails and the linker region are conformationally labile (see Figs. 5, 6). The structure has been rotated ca. 180° about the long axis compared to the view in Figure 6. The Shapely model coloring scheme was employed to denote basic (blue), acidic (red), and hydrophobic (green) side chains. For emphasis, the conservatively substituted aromatic residues that participate in RNA genome recognition and packaging are shaded light green. These residues are located on the "knuckles" of the zinc fingers and protrude into the solvent (to the left in the current view). This figure was generated with a Silicon Graphics computer using the Raster-3D software package (D. Bacon, Maryland-CARB).

tionally homologous to Phe¹⁶ and Ile²⁴ in HIV-1 NC protein) generally results in reduced levels of RNA packaging and loss of infectivity (Gorelick et al., 1988; Meric & Goff, 1989). Interestingly, replacement of the Tyr residue by Gly resulted in a 100-fold reduction of MoMuLV RNA packaging but did not affect the packaging of an endogenous VL30 RNA, whereas replacement of this Tyr by Ser decreased MoMuLV RNA packaging by a factor of 5 but actually enhanced the packaging of VL30 RNA (Meric & Goff, 1989). Based on these findings, it was suggested that the conservatively substituted aromatic and hydrophobic residues might participate directly in specific RNA binding (Meric & Goff, 1989).

Consistent with the above proposal, our structural findings demonstrate that the conservatively substituted aromatic and hydrophobic residues in the HIV-1 NC protein (including residues Phe¹⁶, Ile²⁴, and Ala²⁵ of the N-terminal finger and Trp³⁷, Gln⁴⁵, and Met⁴⁶ of the C-terminal finger) are directed away from the body of the protein and into the solvent, comprising hydrophobic surface patches that could serve as nucleic acid interactive sites (see Fig. 7 and Kinemage 1). Indeed, preliminary nucleic acid binding studies by NMR revealed that synthetic peptide Zn(HIV1-F1) is capable of binding oligodeoxyribonucleic acids, and that the hydrophobic side chains of residues Phe¹⁶ and Ile²⁴ participate in the binding (South et al., 1990b; Delahunty et al., 1992). In the extended form of the HIV-1 NC protein, the aromatic residues of the N- and C-terminal fingers are separated by ca. 28 Å. This distance bears potential significance, as it corresponds to the repeat length of single-stranded A-helical RNA. This form of RNA has been observed previously by X-ray crystallography in structural studies of an intact plant virus capsid particle (Chen et al., 1989). It is thus possible that, in mature particles, the NC protein binds to RNA via hydrophobic interactions between the zinc fingers and the RNA bases, with electrostatic stabilization involving the positively charged linker region and the phosphate backbone of the RNA.

Nucleic acid interactive properties of the HIV-1 NC protein

To determine if the conservatively substituted zinc finger residues are capable of participating in nucleic acid binding, we have measured the zinc and nucleic acid dependence of the intrinsic fluorescence associated with Trp³⁷. The initial slope of the titration curve obtained upon addition of Zn to apo-NC is near linear and requires two equivalents of zinc for maximum fluorescence (Fig. 8A), indicating that the fingers have similar affinities for zinc. A similar effect has been observed for the MoMuLV NC protein (Cornille et al., 1990). The intrinsic fluorescence decreases dramatically upon addition of RNA (Fig. 8B) or DNA polynucleotides (Fig. 8C) at low ionic strength, reflecting a less polar Trp environment in the NC-nucleic

acid complex and supporting previous NMR evidence for an intercalative-type of interaction (South et al., 1990b). Stabilization of RNA secondary structure by addition of Mg²⁺ (Abouhaidar et al., 1979) decreases the initial slope and extent of quenching (Fig. 8B). These data suggest that the HIV-1 NC protein may be capable of unwinding folded RNA structures by binding selectively to single-stranded RNA.

As part of the gag precursor protein, the NC zinc fingers participate in RNA genome recognition, possibly via sequence-specific interactions with an essential "Psi-packaging signal" of the viral RNA (Lever et al., 1989; Aldovini & Young, 1990; Reitz, M., pers. comm.). As an initial probe for sequence-specific binding, fluorescence quenching experiments were performed with a 21-residue oligonucleotide with sequence of the HIV-1 (MN-strain) Psi-packaging signal (dUACGCCAAAUCUUGACUAG) and with a non-Psi sequence (dCCCAGAAATTC AACACTTAAT) (Fig. 8C). At low ionic strength, both oligonucleotides bind the HIV-1 NC protein with similar affinities, although the Psi-sequence oligomer induces slightly greater fluorescence quenching. At higher ionic strength (150 mM NaCl) where ionic contributions to binding are minimized, the NC protein binds the Psi-sequence oligonucleotide with significantly greater affinity (Fig. 8C). In addition, titration studies performed with 7- and 8-residue oligonucleotides representing overlapping regions of the Psi sequence (dAAUUCUUG, dUGACUAG, dUACGCCA, dCCAAAAT) indicate that the NC protein has greatest affinity for the dAAUUCUUG site (Supplementary materials). Sequence-dependent affinities were also observed for studies with octanucleotide homopolymers (Supplementary materials). Although further studies with short oligo-RNA sequences are certainly in order, the present data indicate that the HIV-1 NC protein possesses inherent sequence-specific nucleotide binding capabilities.

Conclusions

The high similarity of the EXAFS data obtained for preparations of intact retroviruses and for a model CCHC zinc finger peptide provides strong evidence that the NC CCHC arrays in mature retroviral particles are populated with zinc. This adds to what is now overwhelming support for the physiological relevance of the CCHC zinc finger motif. The implications of this finding extend to areas outside of the retrovirus field, as the array has recently been discovered in proteins from *Xenopus* (Sato & Sargent, 1991) and humans (Rajavashisth et al., 1989). A common feature of proteins that contain this motif is that they appear to be involved in binding to single-stranded nucleic acids. Thus, the conserved sequence may represent a generic zinc finger motif for single-stranded nucleic acid-binding proteins, analogous to the classical-type

Fig. 8. **A:** The intrinsic fluorescence of the HIV-1 NC protein (123 nM in 5 mM phosphate buffer, pH 7) increases about 2.2-fold and reaches a maximum upon titration with two equivalents of zinc. The linearity of the titration curve indicates that the N- and C-terminal fingers bind zinc with comparable affinities. **B:** The intrinsic tryptophan fluorescence of the protein (182 nM, in 5 mM phosphate buffer at pH 7) is readily quenched by addition of yeast tRNA^{Phe}. The quenching is inhibited by salts that destabilize ionic interactions (NaCl) or stabilize RNA secondary structure (MgCl₂). The fluorescence quenching is significantly reduced in the presence of 37 mM MgCl₂. Thus, the HIV-1 NC protein has a substantially reduced affinity for folded tRNA^{Phe} and appears to be capable of unwinding tRNA^{Phe} at low Mg²⁺ concentrations by binding preferentially to single-stranded RNA. **C:** Preliminary studies with oligodeoxyribonucleotide probes dUACGCCAAAAUUCUUGACUAG (Psi) (sequence analogous to the Psi-packaging signal; see text) and dCCCAGAATTCCAACACTTAAT (nonPsi) demonstrate that the fluorescence quenching is sensitive to nucleic acid composition. The preference is most noticeable under conditions of 0.15 M NaCl, where electrostatic contributions to binding are minimized.

CCHH zinc finger motif found widely in proteins that bind duplex DNA (Summers, 1991).

The solution-state structure of the HIV-1 NC may be best described as a flexible polypeptide chain containing independently folded, noninteracting CCHC-type zinc finger domains. This finding is consistent with results obtained recently for a synthetic peptide with sequence corresponding to residues 13–51 of the HIV-1 3B strain NC protein (Omichinski et al., 1991). Conservatively substituted hydrophobic residues that are essential for viral genome recognition and packaging form hydrophobic patches on the zinc finger surfaces. In the extended form of the protein, these surface patches are separated by ca. 28 Å, a distance of potential importance, as it corresponds to the repeat length of single-stranded A-helical RNA. Thus, nonspecific NC–RNA binding in the mature particles could occur via (1) hydrophobic interactions involving the CCHC hydrophobic patches and the nucleic

acid bases and (2) electrostatic interactions involving the positively charged linker region of the NC protein and the negatively charged phosphate backbone of the RNA. Studies designed to test this hypothesis are underway. In addition, fluorescence measurements indicate that one of the conservatively substituted residues (Trp³⁷) participates directly in binding to single-stranded DNA and RNA. This finding is consistent with the above proposal for nonspecific binding and also supports previous proposals that NC–RNA interactions involving the conservatively substituted hydrophobic residues help mediate viral genome recognition (Meric & Goff, 1989). The NC protein also appears to be capable of readily unwinding folded RNA structures by binding selectively to single-stranded RNA.

Retroviral NC proteins provide potential new targets for the design of both antiviral drugs and vaccines. Mutations of the metal-binding residues of the NC zinc fingers have

enabled production of noninfectious genome-deficient viruses (Aldovini & Young, 1990; Gorelick et al., 1990) that are currently being evaluated as possible vaccine candidates. In addition, the rational design of agents that interfere with the molecular events necessary for genome packaging, such as metal chelating agents or compounds that competitively bind to the hydrophobic surface patches, could provide entirely new classes of compounds for the treatment of AIDS.

Materials and methods

EXAFS data collection and processing

EIAV and MMTV were produced from infected cells and purified by ultracentrifugation in a sucrose density gradient containing 1 mM EDTA as described (Bess et al., 1992). Exogenous or weakly associated zinc was removed by successive (threefold) ultracentrifugation pelleting of particles resuspended in 1.0 mM EDTA solutions (Bess et al., 1992). After final centrifugation, the supernatant was removed, and the pelleted viral paste (0.5 mL) was stored at -70°C in an EXAFS sample holder. The total zinc concentrations of the final pastes (determined by atomic absorption) were 1.46 mM and 0.81 mM for EIAV and MMTV, respectively.

X-ray measurements were carried out in fluorescence mode using an energy resolving 13-element Germanium detector. Beamline X-9A at the Brookhaven National Laboratories was used for all studies in unfocused mode with Si(111) crystals. Harmonics were rejected by a mirror positioned in front of the monochromator. The Ge detector was run at internal count rates of 40,000 counts $\text{channel}^{-1} \text{ s}^{-1}$ to minimize dead time losses. Data for EIAV and MMTV samples were collected at 170 K using a liquid nitrogen cryostat to limit sample damage and to minimize the thermal component of Debye-Waller factors (which leads to maximum differences in backscattering for chemically distinct ligands) (Sagi et al., 1990).

NMR data collection and processing

The HIV-1 NC protein was isolated and purified as described (South et al., 1990a). ^1H NMR data obtained for a 1.7-mg sample of HIV-1 NC protein in D_2O and 90% $\text{H}_2\text{O}/10\% \text{D}_2\text{O}$ solutions, pH 7.2, $T = 30^{\circ}\text{C}$. The protein was isolated from HIV-1 particles (MN-strain), purified under denaturing conditions by HPLC, completely sequenced (Henderson et al., 1992), and reconstituted with zinc. Data obtained with GE GN-500 and Omega PSG-600 NMR spectrometers were transferred to Silicon Graphics computers for processing with FTNMR (Hare Research, Inc.).

NMR data analysis and signal assignments

^1H NMR signal assignments were made by determining scalar connectivities within amino acid residues from 2QF-COSY and HOHAHA spectra and then correlating the signals of adjacent residues on the basis of dipolar connectivities obtained from two-dimensional NOESY data. This general approach has been described in detail (Wuthrich, 1986). Relative amide proton chemical exchange rates were determined qualitatively by measuring the relative sensitivity of NH signals to solvent presaturation.

Structure calculations

Structures were calculated via a standard DG-based simulated annealing/conjugate gradient minimization (SA/CGM) approach (Blake et al., 1991; Gronenborn et al., 1991; South et al., 1991). Distance restraints of 2.0–2.5 Å, 2.0–3.5 Å, and 2.0–4.5 Å were used to represent qualitatively observed strong, medium, and weak cross peak intensities, respectively. The Zn-S(Cys) and Zn-N ϵ (His) bond distances and the Zn-S-C β bond angles were constrained at 2.3 Å and 2.0° (consistent with the EXAFS data) and 109.5° , respectively. Atoms embedded into random positions in 3-space were subjected to variable velocity-simulated annealing and conjugate gradient minimization with DSPACE™ (Hare Research, Inc.). Additional structures were generated by performing two sequential 10-Å randomizations and by subjecting atoms to variable velocity-simulated annealing and CGM refinement (Blake et al., 1991).

Fluorescence measurements

Fluorescence measurements were made with a SHIM-ADZU model RF5000 spectrofluorophotometer using a 1-cm light path and excitation wave length set to 285 nm and emission wave length set to 348 nm. Concentrations of stock protein were determined by amino acid analysis after hydrolysis in 6 N HCl and by UV absorption at 280 nm using an experimentally determined molar extinction coefficient of 6,979 for HIV-1 NC. Concentrations of stock oligonucleotides were determined by UV absorption at 260 nm using calculated molar extinction coefficients obtained by summation of the monomer molar extinction coefficients: A, 15,400; G, 11,700; C, 7,300; T, 8,800; U, 8,900. All titrations were performed in 5 mM sodium phosphate buffer at pH 7 by adding increments of oligonucleotide to a protein solution. All reported values were corrected for dilution to a final volume of 1.0 mL. Oligonucleotides were synthesized as deoxyribonucleotides and purified by reversed-phase HPLC, and yeast tRNA^{Phe} was obtained from Boehringer Mannheim, Germany.

Supplementary materials available

Tables of ^1H NMR signal assignments and distance restraints for structural modeling, the upfield portion of the two-dimensional HOHAHA spectrum of $\text{Zn}_2(\text{HIV1-NC})$, and fluorescence titration curves obtained for synthetic oligodeoxyribonucleic acids with sequences that comprise portions of the Psi-packaging signal will be provided on request. A complete listing of restraints employed for structural modeling and coordinates for 30 HIV-1 NC protein DG structures has been submitted to the Brookhaven Protein Data Bank with the atomic coordinates. The signal assignments and distance restraints appear on the May 1992 *Protein Science* Diskette Appendix in files \SUPLEMNT\Summers.cms and \SUPLEMNT\Summers.no.

Acknowledgments

Support from the National Institutes of Health (NIH grants GM42561 and AI30917 to M.F.S.), the Petroleum Research Fund (grant 23257-G3 to M.R.C.), and the National Cancer Institute, Division of Health and Human Services (DHHS), under contract number N01-C0-74012 with Program Resources Inc./DynCorp, is gratefully acknowledged. The construction and operation of beamline X-9A is supported by NIH grant RR-01633 and by NSF grant DMR-85190959. The National Synchrotron Light Source at Brookhaven National Laboratories is supported by the Department of Energy, Division of Materials Sciences and Division of Chemical Sciences. 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 endorsements of the U.S. Government. We gratefully acknowledge Drs. George McGregor, John Erickson, and Erick Baldwin (FCRDC) for helpful discussions and for the generous gift of tRNA^{Phe}. The Psi-packaging sequence of the HIV-1 MN isolate was kindly provided by Dr. Marvin Reitz (NCI).

References

- Abouhaidar, M., Erickson, J.W., & Bancroft, J.B. (1979). The inhibition of papaya mosaic virus assembly related to the effect of cations on its RNA. *Virology* 98, 116–120.
- Aldovini, A. & Young, R.A. (1990). Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J. Virol.* 64, 1920–1926.
- Berg, J.M. (1986). Potential metal binding domains in nucleic acid binding proteins. *Science* 232, 485–487.
- Bess, J.W., Jr., Powell, P.J., Issaq, H.J., Schumack, L.J., Grimes, M.K., Henderson, L.E., & Arthur, L.O. (1992). Tightly bound zinc in HIV-1, HTLV-1 and other retroviruses. *J. Virol.* 66, 840–847.
- Blake, P.R., Hare, D.R., & Summers, M.F. (1991). Is 2D-NOESY back-calculation necessary for NMR-based structure determination? A systematic study involving a retroviral-type zinc finger peptide. In *Techniques in Protein Chemistry II*, Vol. II (Villafranca, J.J., Ed.), p. 357. Academic Press, New York.
- Bolognesi, D.P., Montelaro, R.C., Frank, H., & Schafer, W. (1978). Assembly of type C oncornaviruses: A model. *Science* 199, 183–186.
- Chen, Z., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kamer, G., Shanks, M., Lomonosoff, G., & Johnson, J.E. (1989). Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science* 245, 154–159.
- Copeland, T.D., Morgan, M.A., & Oroszlan, S. (1984). Complete amino acid sequence of the basic nucleic acid binding protein of feline leukemia virus. *Virology* 133, 137–145.
- Cornille, F., Mely, Y., Ficheux, D., Savignol, I., Gerard, D., Darlix, J.-L., Fournie-Zaluski, M.-C., & Roques, B.P. (1990). Solid phase synthesis of the retroviral nucleocapsid protein NCp10 of Moloney murine leukaemia virus and related "zinc-fingers" in free SH forms. Influence of zinc chelation on structural and biochemical properties. *Int. J. Pept. Protein Res.* 36, 551–558.
- Davies, J.F., II, Hostomska, Z., Hostomosky, Z., Jordan, S.R., & Matthews, D.A. (1991). Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* 252, 88–95.
- Delahunty, M.D., South, T.L., Summers, M.F., & Karpel, R.L. (1992). Nucleic acid interactive properties of a peptide corresponding to the N-terminal zinc finger domain of HIV-1 nucleocapsid protein. *Biochemistry* (in press).
- Dickson, C., Eisenman, R., Fan, J., Hunter, E., & Teich, N. (1985). Protein biosynthesis and assembly. In *RNA Tumor Viruses*, 2nd Ed., Vol. 1 (Weiss, R., Teich, N., Varmus, H., & Coffin, J., Eds.), pp. 513–648. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Dupraz, P., Oertle, S., Meric, C., Damay, P., & Spahr, P.-F. (1990). Point mutations in the proximal Cys-His box of Rous sarcoma virus nucleocapsid protein. *J. Virol.* 64, 4978–4987.
- Fitzgerald, D.W. & Coleman, J.E. (1991). Physicochemical properties of cloned nucleocapsid protein from HIV. Interactions with metal ions. *Biochemistry* 30, 5195–5201.
- Gadhavi, P.L., Davis, A.L., Povey, J.F., Keeler, J., & Laue, E.D. (1991). Polypeptide-metal cluster connectivities in Cd(II)GAL4. *FEBS Lett.* 281, 223–226.
- Gadhavi, P.L., Raine, A.R.C., Alefounder, P.R., & Laue, E.D. (1990). Complete assignment of the ^1H NMR spectrum and secondary structure of the DNA binding domain of GAL4. *FEBS Lett.* 276, 49–53.
- Gardner, K.H., Pan, T., Narula, S., Rivera, E., & Coleman, J.E. (1991). Structure of the binuclear metal-binding site in the GAL4 transcription factor. *Biochemistry* 30, 11292–11302.
- Giedroc, D.P., Qiu, H., Khan, R., King, G.C., & Chen, K. (1992). Zn(II) coordination domain mutants of T4 gene 32 protein. *Biochemistry* 31, 765–774.
- Gorelick, R.J., Henderson, L.E., Hanser, J.P., & Rein, A. (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: Evidence for specific RNA recognition by a "zinc finger-like" protein sequence. *Proc. Natl. Acad. Sci. USA* 85, 8420–8424.
- Gorelick, R.J., Nigida, S.M., Bess, J.W., Arthur, L.O., Henderson, L.E., & Rein, A. (1990). Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* 64, 3207–3211.
- Green, L.M. & Berg, J.M. (1989). A retroviral Cys-Xaa2-Cys-Xaa4-His-Xaa4-Cys peptide binds metal ions: Spectroscopic studies and a proposed three-dimensional structure. *Proc. Natl. Acad. Sci. USA* 86, 4047–4051.
- Green, L.M. & Berg, J.M. (1990). Retroviral nucleocapsid protein-metal ion interactions: Folding and sequence variants. *Proc. Natl. Acad. Sci. USA* 87, 6403–6407.
- Gronenborn, A.M., Filpula, D.R., Essig, N.Z., Achari, A., Whitlow, M., Wingfield, P.T., & Clore, G.M. (1991). A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 253, 657–661.
- Hard, T., Kellenbach, E., Boelens, R., Maler, B.A., Dahlman, K., Freedman, L.P., Carlstedt-Duke, J., Yamamoto, K.R., Gustafsson, J.-A., & Kaptein, R. (1990). Solution structure of the glucocorticoid receptor DNA-binding domain. *Science* 249, 157–160.
- Henderson, L.E., Benveniste, R.E., Sowder, R., Copeland, T.D., Schultz, A.M., & Oroszlan, S. (1988). Molecular characterization of gag proteins from simian immunodeficiency virus (SIVMne). *J. Virol.* 62, 2587–2595.
- Henderson, L.E., Bowers, M.A., Sowder, R.C., III, Sarabyn, S.A., Johnson, D.G., Bess, J.W., Jr., Arthur, L.O., Bryant, D.K., & Fenselau, C. (1992). Gag proteins of highly replicative HIV-1_{MN}: Post-translational modifications, proteolytic processings, and complete amino acid sequences. *J. Virol.* 66, 1856–1865.

- Henderson, L.E., Copeland, T.D., Sowder, R.C., Schultz, A.M., & Oroszlan, S. (1987a). Analysis of proteins and peptides purified from sucrose gradient banded at HTLV-III. In *Human Retroviruses, Cancer and AIDS: Approaches to Prevention and Therapy* (Bolognesi, D., Ed.), pp. 135-147, Liss Publishing Co., New York.
- Henderson, L.E., Copeland, T.D., Sowder, R.C., Smythers, G.W., & Oroszlan, S. (1987b). Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses. *J. Biol. Chem.* 256, 8400-8406.
- Henderson, L.E., Sowder, R.C., Copeland, T.D., Oroszlan, S., & Benveniste, R.E. (1990). Gag precursors of HIV and SIV are cleaved into six proteins found in the mature virions. *J. Med. Primatol.* 19, 411-419.
- Henderson, L.E., Sowder, R.C., Smythers, G.W., & Oroszlan, S. (1987b). Chemical and immunological characterizations of equine infectious anemia virus gag-encoded proteins. *J. Virol.* 61, 1116-1124.
- Jentoft, J.C., Smith, L.M., Fu, X., Johnson, M., & Leis, J. (1988). Conserved cysteine and histidine residues of the avian myeloblastosis virus nucleocapsid protein are essential for viral replication but are not "zinc-binding fingers." *Proc. Natl. Acad. Sci. USA* 85, 7094-7098.
- Karpel, R.L., Henderson, L.E., & Oroszlan, S. (1987). Interactions of retroviral structural proteins with single-stranded nucleic acids. *J. Biol. Chem.* 262, 4961-4967.
- Katz, R.A. & Jentoft, J.E. (1989). What is the role of the Cys-His motif in retroviral nucleocapsid (NC) proteins? *BioEssays* 11, 176-181.
- Kochoyan, M., Havel, T.F., Nguyen, D.T., Dahl, C.E., Keutemann, H.T., & Weiss, M.A. (1991). Alternating zinc fingers in the human male associated protein (ZFY): 2D NMR structure of an even finger and implications for "jumping-linker" DNA recognition. *Biochemistry* 30, 3371-3386.
- Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., & Wright, P.W. (1989). Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* 245, 635-637.
- Lever, A., Gottlinger, H., Haseltine, W., & Sodroski, J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol.* 63, 4085-4087.
- Luisi, E.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R., & Sigler, P.B. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* 352, 497-505.
- Meric, C., Darlix, J.-L., & Spahr, P.-F. (1984). It is Rous sarcoma virus protein p12 and not p19 that binds tightly to Rous sarcoma virus RNA. *J. Mol. Biol.* 173, 531-538.
- Meric, C. & Goff, S.P. (1989). Characterization of Moloney murine leukemia virus mutants with single-amino-acid substitutions in the Cys-His box of the nucleocapsid protein. *J. Virol.* 63, 1558-1568.
- Navia, M.A., Fitzgerald, P.M.D., McKeever, B.M., Leu, C.-T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L., & Springer, J.P. (1989). Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* 337, 615-620.
- Omichinski, J.G., Clore, G.M., Appella, E., Sakaguchi, K., & Gronenborn, A.M. (1990). High-resolution three-dimensional structure of a single zinc finger from a human enhancer binding protein in solution. *Biochemistry* 29, 9324-9334.
- Omichinski, J.G., Clore, G.M., Sakaguchi, K., Appella, E., & Gronenborn, A.M. (1991). Structural characterization of a 39-residue synthetic peptide containing the two zinc binding domains from the HIV-1 p7 nucleocapsid protein by CD and NMR spectroscopy. *FEBS Lett.* 292, 25-30.
- Parraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T., & Kleit, R.E. (1988). Zinc-dependent structure of a single-finger domain of yeast ADR1. *Science* 241, 1489-1492.
- Pavletich, N.P. & Pabo, C.O. (1991). Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809-817.
- Rajavashisth, T.B., Taylor, A.K., Andalibi, A., Svenson, K.L., & Luisi, A.J. (1989). Identification of a zinc finger protein that binds to the sterol regulatory element. *Science* 245, 640-643.
- Roberts, W.J., Pan, T., Elliott, J.I., Coleman, J.E., & Williams, K.R. (1989). p10 single-stranded nucleic acid binding protein from murine leukemia virus binds metal ions via the peptide sequence Cys26-X2-Cys29-X4-His34-X4-Cys39. *Biochemistry* 28, 10043-10047.
- Sagi, I., Wirt, M.D., Chen, E., Frisbie, S., & Chance, M.R. (1990). Structure of an intermediate of coenzyme B12 catalysis by EXAFS: Cobalt(II)B12. *J. Am. Chem. Soc.* 112, 8639-8644.
- Sato, S.M. & Sargent, T.D. (1991). Localized and inducible expression of *Xenopus*-posterior (Xpo), a novel gene active in early frog embryos encoding a protein with a 'CCHC' finger domain. *Development* 112, 747-753.
- Schiff, L.A., Nibert, M.L., & Fields, B.N. (1988). Characterization of a zinc blotting technique: Evidence that a retroviral gag protein binds zinc. *Proc. Natl. Acad. Sci. USA* 85, 4195.
- Schwabe, J.W.R., Neuhaus, D., & Rhodes, D. (1990). Solution structure of the DNA-binding domain of the estrogen receptor. *Nature* 348, 458-461.
- Smith, B.J. & Bailey, J.M. (1979). The binding of an avian myeloblastosis virus basic 12,000 dalton protein to nucleic acids. *Nucleic Acids Res.* 7, 2055-2072.
- South, T.L., Blake, P.R., Hare, D.R., & Summers, M.F. (1991). C-terminal retroviral-type zinc finger domain from the HIV-1 nucleocapsid protein is structurally similar to the N-terminal zinc finger domain. *Biochemistry* 30, 6342-6349.
- South, T.L., Blake, P.R., Sowder, R.C., Arthur, L.O., Henderson, L.E., & Summers, M.F. (1990a). The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers. *Biochemistry* 29, 7786-7789.
- South, T.L., Kim, B., Hare, D.R., & Summers, M.F. (1990b). Zinc fingers and molecular recognition. Structure and nucleic acid binding studies of an HIV zinc finger-like domain. *Biochem. Pharmacol.* 40, 123-129.
- South, T.L., Kim, B., & Summers, M.F. (1989). ¹¹³Cd NMR studies of a 1:1 Cd adduct with an 18-residue finger peptide from HIV-1 nucleic acid binding protein, p7. *J. Am. Chem. Soc.* 111, 395-396.
- Summers, M.F. (1991). Zinc finger motif for single-stranded nucleic acids? Investigations by nuclear magnetic resonance. *J. Cell. Biochem.* 45, 41-48.
- Summers, M.F., South, T.L., Kim, B., & Hare, D.R. (1990). Structure of an HIV zinc finger via a new NMR-based distance geometry approach. *Biochemistry* 29, 329-340.
- Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J., & Kent, S.B. (1989). Conserved folding in retroviral proteases: Crystal structure of a synthetic HIV-1 protease. *Science* 245, 616-621.
- Wuthrich, K. (1976) *NMR in Biological Research: Peptides and Proteins*. Elsevier, New York.
- Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley, New York.