Journal of Biomolecular NMR, 4 (1994) 153–170
ESCOM

J-Bio NMR 176

Three-dimensional ^{1}H NMR structure of the nucleocapsid protein NCp10 of Moloney murine leukemia virus

Hélène Déméné[a], Nathalie Jullian[a], Nelly Morellet[a], Hugues de Rocquigny[a], Fabrice Cornille[a], Bernard Maigret[b] and Bernard P. Roques[a,*]

[a]*Département de Pharmacochimie Moléculaire et Structurale, U266 INSERM–URA D 1500 CNRS, Faculté de Pharmacie, 4, Avenue de l'Observatoire, 75270 Paris Cedex 06, France*
[b]*Université Nancy I, Laboratoire Chimie Théorique, B.P. 239, 54506 Vandoeuvre-lés-Nancy Cedex, France*

Received 22 June 1993
Accepted 25 October 1993

Keywords: Nucleocapsid protein; Retrovirus Moloney; NMR; Protein structure; Distance geometry

SUMMARY

The nucleocapsid protein of Moloney murine leukemia virus (NCp10) is a 56-amino acid protein which contains one zinc finger of the $CysX_2CysX_4HisX_4Cys$ form, a highly conserved motif present in most retroviruses and retroelements. At pH $\geq$ 5, NCp10 binds one zinc atom and the complexation induces a folding of the $CysX_2CysX_4HisX_4Cys$ box, similar to that observed for the zinc-binding domains of HIV-1 NC protein. The three-dimensional structure of NCp10 has been determined in aqueous solution by 600 MHz ^{1}H NMR spectroscopy. The proton resonances could be almost completely assigned by means of phase-sensitive double-quantum-filtered COSY, TOCSY and NOESY techniques. NOESY spectra yielded 597 relevant structural constraints, which were used as input for distance geometry calculations with DIANA. Further refinement was performed by minimization with the program AMBER, which was modified by introducing a zinc force field. The solution structure is characterized by a well-defined central zinc finger (rmsd of 0.747 ± 0.209 Å for backbone atoms and 1.709 ± 0.187 Å when all atoms are considered), surrounded by flexible N- and C-terminal domains. The Tyr^{28}, Trp^{35}, Lys^{37}, Lys^{41} and Lys^{42} residues, which are essential for activity, lie on the same face of the zinc finger, forming a bulge structure probably involved in viral RNA binding. The significance of these structural characteristics for the various biological functions of the protein is discussed, taking into account the results obtained with various mutants.

INTRODUCTION

The retroviral nucleocapsid (NC) proteins, encoded by the 3' end of the *gag* gene, are small basic single-stranded nucleic-acid-binding proteins (Covey, 1986). In avian, murine and human retroviruses, the NC proteins have been shown to be involved in retroviral RNA dimerization (Coffin, 1984; Prats et al., 1988, 1990; Bieth et al., 1990; Darlix et al., 1990), encapsidation (Méric

*To whom correspondence should be addressed.

and Spahr, 1986; Gorelick et al., 1988; Méric and Goff, 1989) and annealing of the replication primer tRNA to the initiation site for reverse transcription onto the viral genome (Prats et al., 1988). The NC protein of Moloney murine leukemia virus (MoMuLV), NCp10, contains one $CysX_2CysX_4HisX_4Cys$ sequence, a motif which is shared by most retroviral NC proteins and which behaves as a zinc-binding domain (Berg, 1986). The nucleocapsid protein NCp7 of human immunodeficiency virus, HIV-1, which corresponds to the biologically active fragment of NCp15 (De Rocquigny et al., 1991) contains two zinc fingers. The role of these zinc fingers is still ambiguous and the molecular mechanisms accounting for the various NC protein function remain largely unknown.

It was recently shown that, for HIV-1 NCp7 and MoMuLV NCp10, only the peptide domains outside the zinc fingers are critical for RNA annealing activities (De Rocquigny et al., 1992,1993), supporting the proposed minor role played by the CCHC box of NCp10 in inducing an interaction between $tRNA^{Pro}$ and MoMuLV viral genome (Prats et al., 1991). On the other hand, the replacement, by site-directed mutagenesis, of a cysteine or a histidine residue belonging to the CCHC box resulted in a loss of infectivity due to a severe defect in viral RNA packaging (Gorelick et al., 1988; Méric and Goff, 1989). In agreement with these results, both the zinc fingers and the flanking basic residues were shown to be necessary for the production of infectious MoMuLV virions (Housset et al., 1993).

In vitro, the CCHC box of NCp10 was shown to bind one zinc atom with a very high affinity ($K_{app} = 1.2 \times 10^{13}\ M^{-1}$) (Cornille et al., 1990; Green and Berg, 1990; Mély et al., 1991). It has been inferred from UV and ^{113}Cd NMR studies on Rauscher NCp10 that the zinc atom is bound through a tetrahedral coordination mode (Roberts et al., 1989). This has also been confirmed by ^1H NMR studies in the case of HIV-1 NCp7 (Summers et al., 1990; Omichinsky et al., 1991; South et al., 1991; Morellet et al., 1992), and by EXAFS experiments performed with an isolated zinc finger of NCp7 and with equine infectious anemia virus (EIAV) intact particles (Chance et al., 1992; Summers et al., 1992). To date, most of the structural studies on retroviral zinc fingers have been carried out on the HIV-1 NCp7 protein. However, the structure of MoMuLV NC proteins, such as NCp10, could provide a useful murine model to study retroviral gene regulation by nucleocapsid proteins and to investigate the activities of putative antiviral drugs inhibiting the nucleocapsid functions.

With this aim, the complete ^1H NMR sequence-specific assignment of synthetic NCp10 (De Rocquigny et al., 1991) has been carried out and its structure has been derived from distance geometry calculations. The results show similar 3D arrangements for NCp10 and NCp7, including the spatial disposition of the aromatic residues present in the zinc fingers, suggesting that the molecular mechanisms which endow NC proteins with their functions are highly similar, if not identical.

MATERIALS AND METHODS

Material and NMR sample preparation

In order to carry out NMR experiments, large quantities of NCp10 were synthesized using the stepwise solid-phase method and Fmoc amino acids on an automatic reprogrammed Applied Biosystem 431 A synthesizer (De Rocquigny et al., 1991). To preserve the highly oxidizable cysteine residues of the protein (purity $\geq$ 99.9%), all manipulations were carried out under argon.

NMR samples were prepared by dissolving 12.3 mg (3 mM) in 90% H_2O/10% D_2O solution or in 99.96% D_2O in the presence of 1.5 equiv of zinc chloride, with the pH adjusted below 6 in order to slow down the exchange rate of amide protons. The spectrum of NCp10 at pH 7 is not significantly modified, suggesting that the structure of the protein remains almost identical at physiological pH, as already discussed in the case of NCp7 (Morellet et al., 1992).

NMR experiments

All experiments were performed on a Bruker AMX 600 spectrometer, operating at 600 MHz for protons, with the carrier frequency set to the water resonance. Two-dimensional NMR spectra were obtained in the phase-sensitive mode using the time-proportional phase increments method (Marion and Wüthrich, 1983). Water peak suppression was achieved by irradiation of the solvent resonance during the relaxation delay (and during the mixing time for the NOESY spectra). Spectral widths were about 12 ppm in both dimensions. Spectra were recorded with 2K data points and 512 t1 increments with 128 or 64 scans per increment (except for the DQF-COSY, where 4K data points were used and there were 1024 t1 increments). All the 2D experiments described below were performed on both samples (H_2O and D_2O).

Two TOCSY spectra (Braunschweiler and Ernst, 1983; Davis and Bax, 1985) with total mixing times of 55 and 87 ms were recorded. The extended MLEV17 cycle was preceded and followed by a 2.5-ms trim pulse to defocus magnetization not parallel to the spin-lock axis. A DQF-COSY experiment at 20 °C was also performed (Rance et al.,1983). NOESY spectra (Jeener et al., 1979;

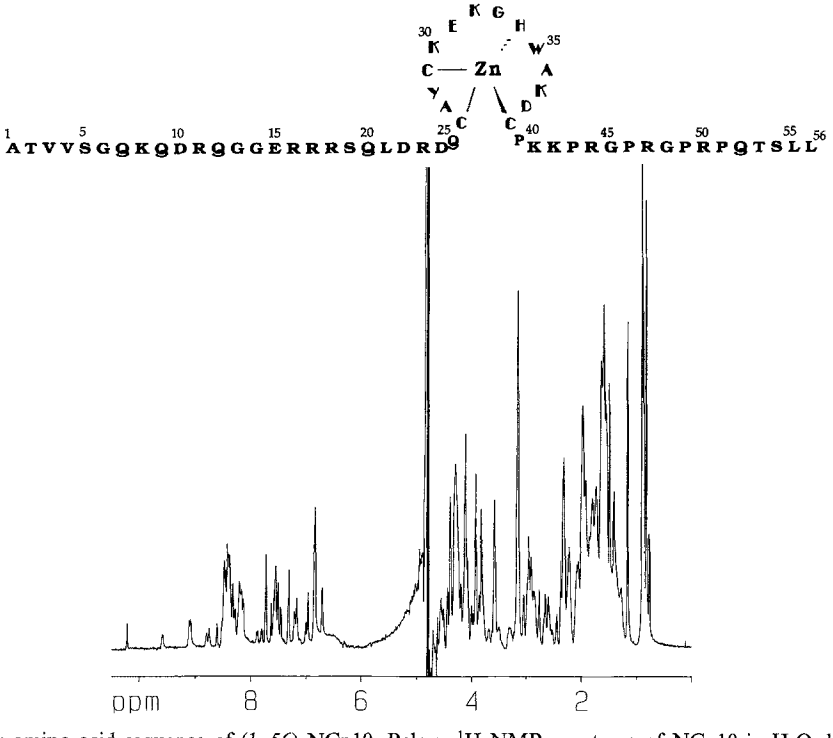

Fig. 1. Above: amino acid sequence of (1–56) NCp10. Below: ^1H NMR spectrum of NCp10 in H_2O, bound to zinc. T = 293 K, pH = 5.55.

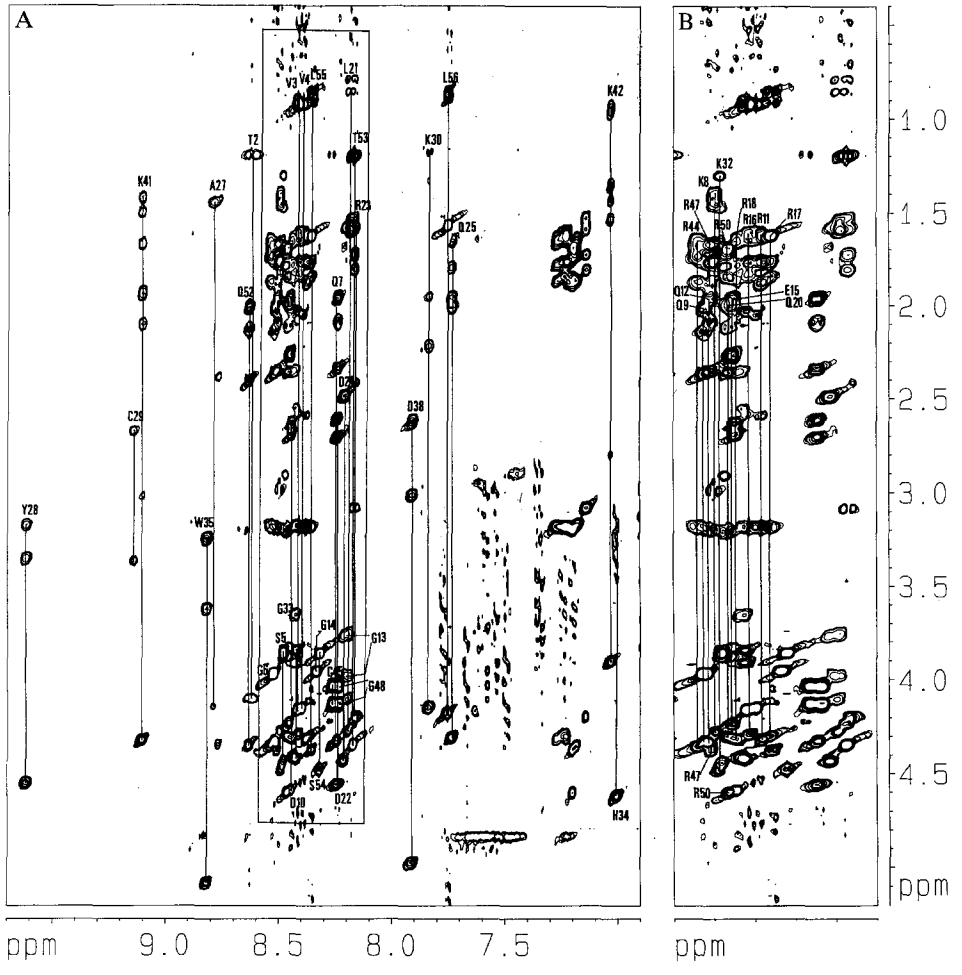

Fig. 2. (A) Part of the TOCSY spectrum of NCp10 in H_2O (T = 293 K, τ_m = 87 ms, pH = 5.65), showing amide proton/aliphatic proton scalar correlations. (B) Expanded plot of the boxed region in (A).

Macura et al., 1981) were recorded with mixing times of 60, 100 and 200 ms at 20 and 30 °C. No zero-quantum suppression techniques were applied. The relaxation delay was 2 s.

In order to observe the slowly exchanging amide resonances, a NOESY experiment was performed at pH 8.7 since, at this pH, the amide protons whose resonances are still observable are considered as hydrogen bonded or at least deeply buried and thus protected from the solvent. All free induction decays were multiplied by $\pi/6$-shifted sine-bell window functions in both dimensions. After zero-filling and double Fourier transformation, the digital resolution was 3.54 Hz/pt in the $\omega 2$ dimension and 7.07 Hz/pt in the $\omega 1$ dimension. Additional baseline corrections were performed using polynomial functions of standard Bruker software, installed on a Bruker X32 computer.

NMR-derived constraints

NOEs observed in NOESY spectra obtained with 60 or 100 ms mixing time were used as

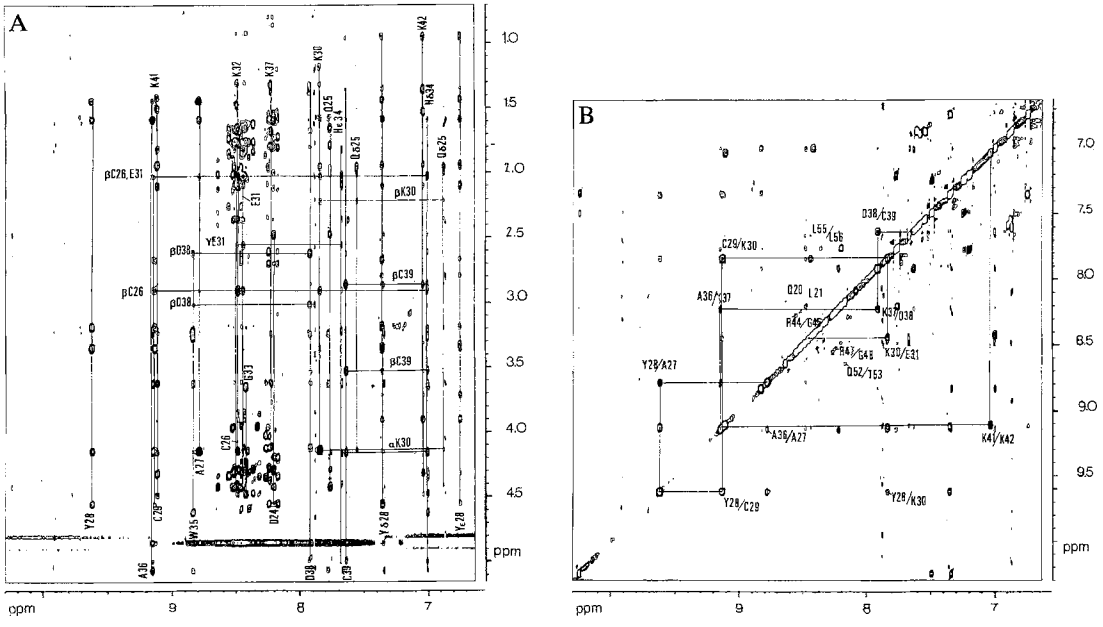

Fig. 3. Parts of the NOESY spectrum (τ_m = 100 ms, pH = 5.5) of NCp10 in H_2O. (A) Through-space interactions between NH and aliphatic protons. Residues labelled vertically are those of the region D^{24}–K^{42}, encompassing the zinc finger. Relevant NOEs for the zinc finger structure are pointed out. (B) NH/NH NOE connectivities. The three stretches of connectivities in the zinc finger region, as well as the NH/NH NOEs observed for the C-terminal part of NCp10 are labelled.

modelling constraints. NOEs obtained in the NOESY experiment performed at 200 ms were taken into account only when spin-diffusion pathways could be discarded (Kochoyan et al., 1991). NOEs were classified as strong, medium or weak on the basis of the relative amplitude of the cysteine geminal protons used as an internal standard.

Structure calculation

Distance constraints were classified into three categories: 2.0–2.5 Å, 2.0–3.5 Å and 2.0–4.5 Å, corresponding to strong, medium and weak NOEs respectively. On the basis of cadmium-com-

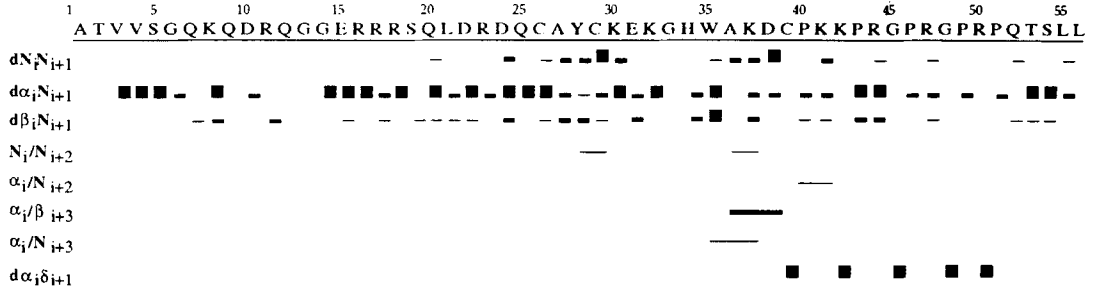

Fig. 4. Summary of short- and medium-range NOEs, involving the NH, H^{α} and H^{β} protons, as well as the H^{δ} protons of proline residues, observed for the Zn^{2+}–NCp10 complex. The NOEs are classified as strong, medium and weak, corresponding to the thickness of the lines.

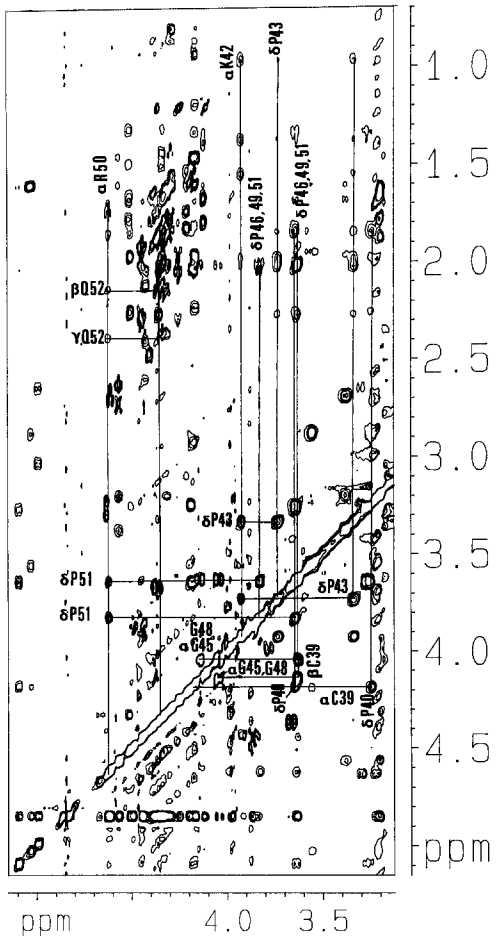

Fig. 5. Part of the NOESY spectrum (τ_m = 200 ms, pH = 5.5) of NCp10 in D_2O, showing $d\alpha_i\delta_{i+1}$ connectivities between prolines and the preceding residues in the sequence. The unique nonsequential NOEs observed in the C-terminal part of NCp10 between the α-proton of Arg^{50} and the β- and γ-protons of Gln^{52} are indicated.

plexed metallothionein (Arseniev et al., 1988), additional constraints were added in order to define the interactions between the cysteine and histidine residues and the zinc atom. The zinc was introduced as a pseudoatom QZn, linked to the S^{γ} of each cysteine residue with a bond length of 2.3 Å. The distance between QZn and the $N^{\varepsilon 2}$ atom of His^{34} was constrained to 2.0 Å. The distance between two QZn atoms was imposed to be less than 0.1. Å, the three QZn pseudoatoms mimicking the unique zinc ion.

Structures were calculated using the distance geometry program DIANA (Güntert et al., 1991a). Preliminary calculations generated a group of structures which were used to identify ambiguities in the assignment of NOESY spectra, and thus to determine a refined set of NOE distance constraints. Finally, DIANA was run with standard parameters and methodology, as described by Güntert et al. (1991b). The 20 best structures, according to the target function value, were then subjected to energy minimization with the AMBER software (Singh et al., 1986), using the conjugate-gradient method. A zinc force field derived from ab-initio calculations was parame-

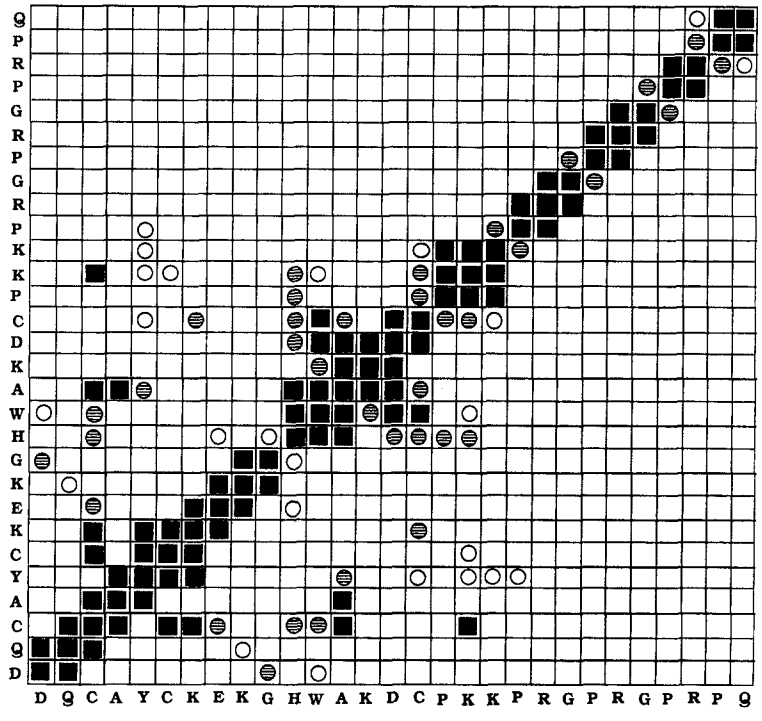

Fig. 6. Diagonal plot of NOEs for the region D^{24}–Q^{52}. Black boxes: backbone/backbone NOE; hatched circles: backbone/side chain NOE; open circles: side chain/side chain NOE.

terized in the AMBER framework in order to take into account metal–protein interactions (Jacob, 1990).

The resulting structures were graphically analyzed with the INSIGHT molecular modelling software (BIOSYM Technologies, San Diego, CA) on a Silicon Graphics IRIS 4D35 workstation. All structure calculations were performed on an IBM RS6000/550 workstation.

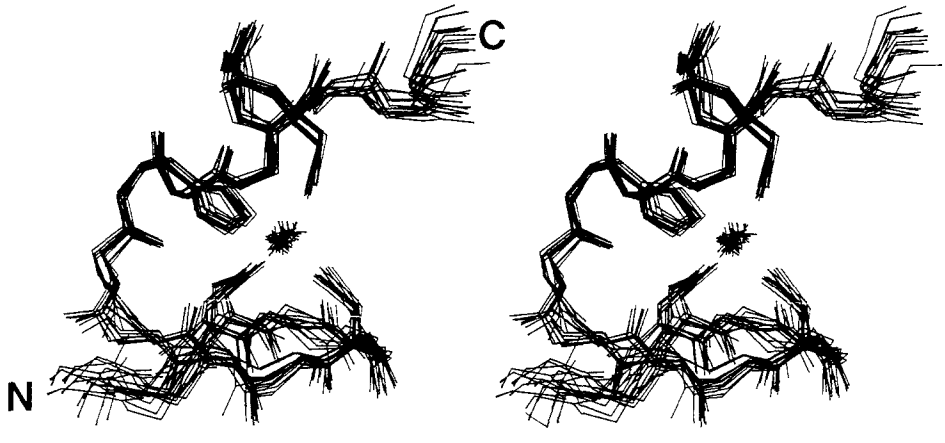

Fig. 7. Stereo representation of the DIANA-derived 20 superimposed zinc finger domains.

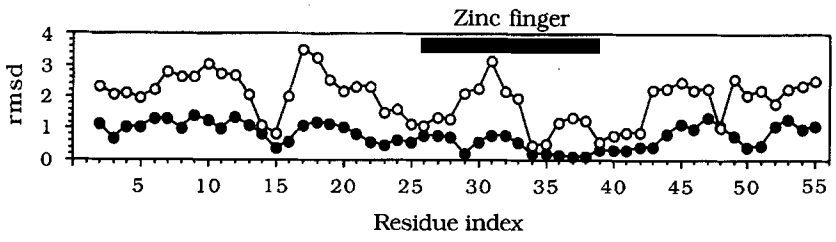

Fig. 8. Plot of the mean rmsd of the 20 calculated structures for tripeptide backbone (closed circles) and all-atom (open circles) segments against the sequence number of the central residue.

RESULTS

NMR experiments

The 1D ^{1}H NMR spectrum of NCp10 in the presence of one equivalent of Zn^{2+} (Fig. 1) exhibited marked dispersion of chemical shifts: downfield shifted $C^{\alpha}H$ resonances and upfield shifted methyl resonances, reflecting the presence of a folded metal-complexed structure. Furthermore, the spectrum of Zn-bound NCp10 contains only a single set of resonances, indicating the presence of a unique species in solution.

The methodology employed in the assignment of 2D NMR spectra is well established

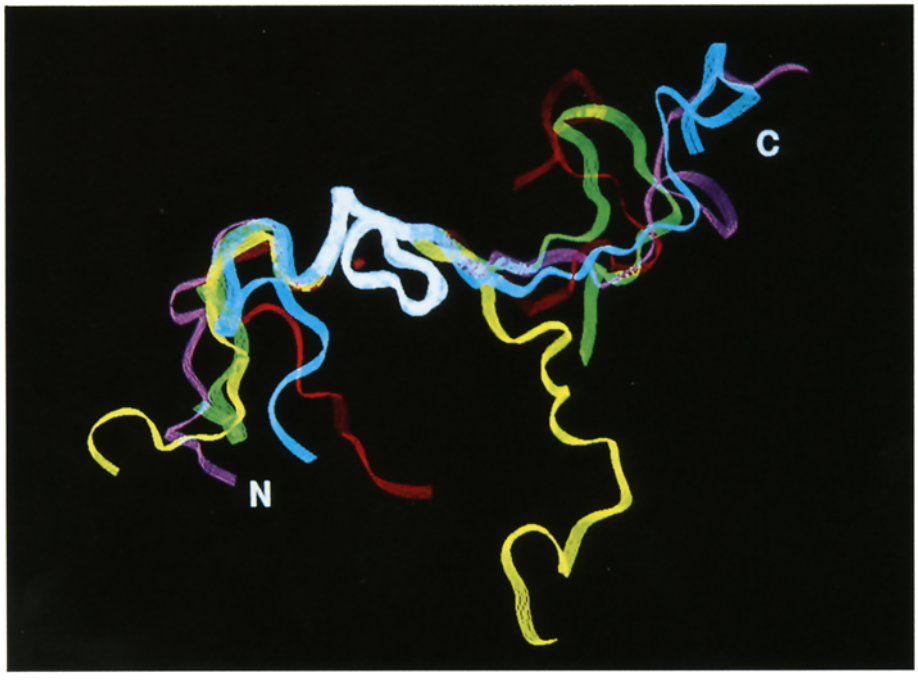

Fig. 9. Ribbon representation of five out of the 20 calculated structures: the superimposition was performed with the backbone atoms of the (26–39) zinc finger residues, as indicated by white color. The zinc atom is represented as a star. Each structure is displayed with a different color for the extra zinc finger domain.

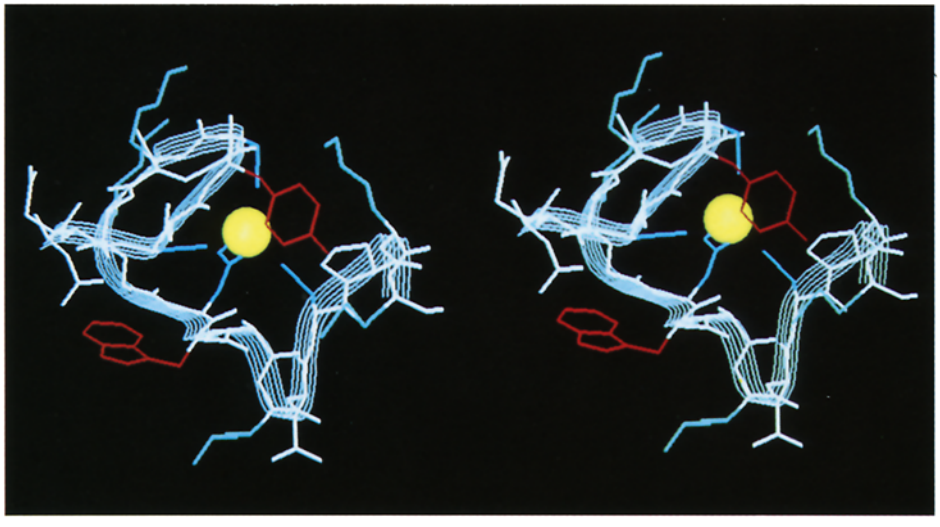

Fig. 10. Stereoview of the NCp10 zinc finger domain.

(Wüthrich, 1986). Spin systems for amino acid side chains were subsequently assigned, based on intraresidue connectivities identified in DQF-COSY and TOCSY spectra recorded in H_2O and ^2H_2O at 293 K.

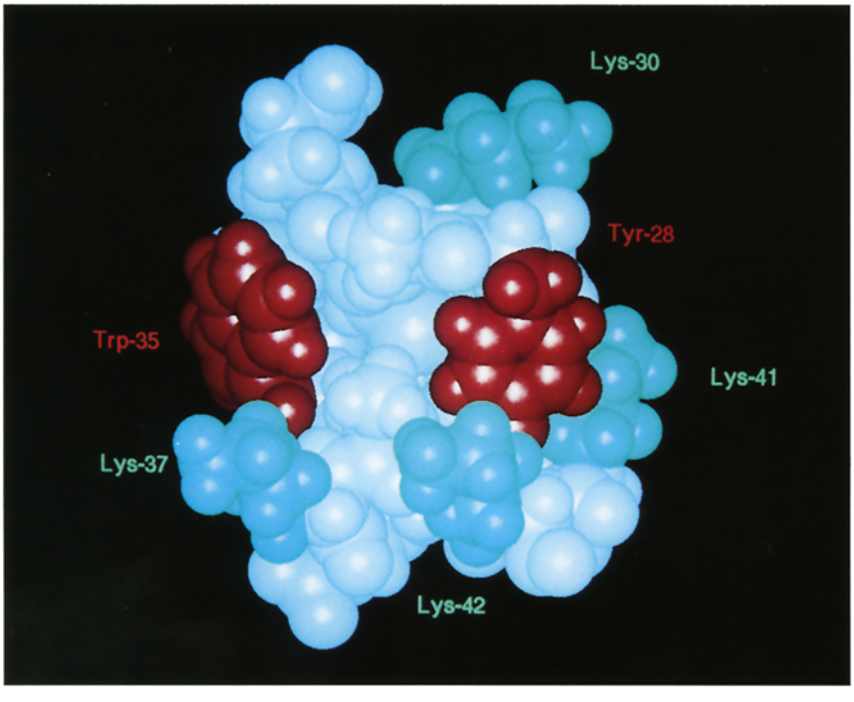

Fig. 11. CPK representation of the bulge-like structure of NCp10, evidencing the relative spatial disposition of the critical residues. Trp^{35} and Tyr^{28}: red; Lys^{37}, Lys^{42} and Lys^{41}: green.

In the TOCSY spectrum (Fig. 2), a total of 45 NH-CH^{α} and NH-CH^{β} cross peaks were obtained. The few missing NH-CH^{α} and NH-side chain cross peaks may have been lost due to incomplete magnetization transfer, overlap of the resonances, or H_2O signal irradiation. Due to the large number of overlapping resonances, identification of spin systems was done mainly using TOCSY and NOESY spectra and confirmed by DQF experiments. Total coherence transfer was more efficient for the residues belonging to the N- and C-terminal parts than for residues of the zinc finger domain. In contrast, dipolar relaxation was more efficient for the residues within the zinc finger sequence due to a lower degree of freedom of this sequence. The almost complete NMR assignment of NCp10 was facilitated by the analysis of the (24–42) NCp10 zinc finger alone (unpublished results).

The TOCSY spectra analysis readily allowed identification of seven NH resonances of the eight arginines (R^{11}, R^{16}, R^{17}, R^{18}, R^{23}, R^{44}, R^{47} and R^{50}) through NH/CH^{δ} correlation peaks (Fig. 2B). At pH 5.5, the NH and aliphatic resonances of R^{18} were found to overlap with those of R^{50}, except for the H^{α} proton. The NH^{ε} protons of these arginines were observed in H_2O TOCSY spectra and could be unambiguously assigned for R^{23}, R^{44} and R^{50} only. The three alanines (A^1, A^{27} and A^{36}), two threonines (T^2 and T^{53}), two valines (V^3 and V^4) and three leucines (L^{21}, L^{55} and L^{56}) are the 10 methyl-containing spin systems of the NCp10 protein. Identification of backbone resonances of A^{27} and A^{36} could be achieved only with the help of NOESY spectra (Fig. 3A), whereas the two threonines T^2 and T^{53} were readily identified (see Fig. 2A) through the downfield resonances of their H^{β} protons, the valines V^3 and V^4 through the upfield resonances of their H^{β} protons (Fig. 2A) and the leucines L^{21}, L^{55} and L^{56} through their nonequivalent upfield shifted H^{δ} proton (Fig. 2A). Of the six lysines (K^8, K^{30}, K^{32}, K^{37}, K^{41} and K^{42}), except for K^{37} all NH resonance lines are directly visible in the TOCSY spectrum (Figs. 2A and B). For some of them (K^{30}, K^{32} and K^8), a lack of total-coherence transfer occurs. The NOESY spectra (Fig. 3A) allowed identification of the missing resonances of K^{30}, K^{32}, K^{37}, K^{41} and K^{42}. The six glutamines Q^7, Q^9, Q^{12}, Q^{25} and Q^{52} were characterized by their side-chain connectivity patterns. Amide/aliphatic cross peaks were observable for all of them in TOCSY spectra (Figs. 2A and B). An intraresidue NOE connectivity between $NH^{\varepsilon 2}$ and H^{γ} could be detected only for Q^{25}. The identification of E^{15} was straightforward (Fig. 2B), but E^{31} could only be completely assigned through NOESY spectra (Fig. 3A). The resonances of the three aromatic residues of NCp10, i.e., H^{34}, W^{35} and Y^{28}, were easily identified, all the more because their position within the zinc finger sequence causes shifting of their proton resonances to less crowded regions of the spectrum (Fig. 2A). Of the four aspartic acids (D^{10}, D^{22}, D^{24} and D^{38}), D^{10} and D^{22} were recognizable through their downfield shifted (close to random coil values) H^{β} resonances (Fig. 2A). Because the H^{β} protons of D^{24} are chemically equivalent, this residue was identified during the assignment step. Likewise, the resonances of D^{38} (Figs. 2A and 3A) could be assigned because the localization of this residue within the zinc finger sequence leads to unusual proton chemical shifts. The three cysteines (C^{26}, C^{29} and C^{39}) were also identified at the sequential-assignment step (Fig. 3A). The three serines (S^5, S^{19} and S^{54}) were easily identified in TOCSY and DQF spectra through the particular downfield resonance of their H^{β} proton (Fig. 2A). Of the six glycines, G^{13}, G^{14}, G^{33}, $G^{45}$and G^{48} could be identified from the large resonance splittings (resulting from large coupling constants) in the DQF-COSY spectra. G^6, whose H^{α} protons are chemically equivalent, was assigned later, using NOESY interresidual connectivities. Finally, analysis of the aliphatic part of TOCSY and DQF spectra allowed the characterization of the five proline residues through their long side-chain connectivity patterns.

Sequential analysis of NMR spectra was performed by the usual assignment strategy, resulting in the nearly complete interpretation of the ^{1}H NMR spectra (Table 1). The sequential connectivities, which were determined by using through-space NOE interactions, are depicted in Fig. 4. In this 56-residue protein, only 50 sequential dαN peaks were expected in the H_2O NOESY spectrum, because of the presence of the P^{40}, P^{43}, P^{46}, P^{49} and P^{51} proline residues. Of these 50 expected peaks, only 37 were observed. The remaining 13 dαN cross peaks could not be unambiguously determined due to overlap between intraresidue cross peaks or as a result of radio-frequency irradiation used for suppression of the water resonance. In six cases, this problem was overcome by using other sequential contacts, such as dNN (K^{30}/E^{31}, A^{36}/K^{37}, Q^{52}/T^{53}, dβN (R^{11}/Q^{12}, S^{19}/Q^{20}, E^{31}/K^{32}, Q^{52}/T^{53}) and/or dγN connectivities. No sequential connectivities were observed for the following residue pairs: A^1/T^2, T^2/V^3, Q^7/K^8, Q^9/D^{10}, Q^{12}/G^{13}, G^{13}/G^{14}, R^{18}/S^{19}.

Sequential connectivities between the five prolines and their preceding amino acids were determined from the $H_i^\alpha/H_{i+1}^\delta$ (Fig. 5) and NH_i/H_{i+1}^δ NOE peaks. The resonances of the H^β, H^γ and H^δ protons of P^{46}, P^{49} and P^{51} were somewhat overlapping, thus assignment was tentatively done on the basis of the H_i^α/NH_{i+1} and then of the $H_i^\alpha/H_i^{\beta,\gamma,\delta}$ cross-peak identification. These NOEs are characteristic of proline residues in a trans conformation. No evidence of an equilibrium between cis and trans forms of proline could be detected. After sequential assignment, medium- and long-range connectivities supplying secondary and tertiary structure information were searched in the NOESY spectra. The corresponding cross-correlation signals are depicted in Figs. 4 and 6.

A^1–Q^{25}: N-terminal sequence

The rather intense sequential dαN connectivities are characteristic of an extended structure (Wüthrich, 1986) for this part of the molecule (Fig. 4), except for the D^{22}–Q^{25} segment for which medium- and long-range NOEs were found (Figs. 4 and 6). The chemical shifts listed in Table 1 provide additional evidence for the absence of regular secondary structure elements in this region, since most of the chemical shifts are very similar to those of random coil sequences. Interestingly, the $NH^{\varepsilon 2}$ chains of Q^{25} also exhibited interresidual connectivities with the aliphatic chain of K^{30}.

C^{26}–C^{39}: zinc finger domain

This sequence corresponds to the unique zinc finger of NCp10 and a comparison of the NOE pattern with those observed in the first and second zinc fingers of NCp7 from HIV-1 (MN strain) (South et al., 1990; Summers et al., 1990) and HIV-1 (LAV strain) (Omichinski et al., 1991; Morellet et al., 1992) gives insight into the similarities and differences between these zinc-complexed domains. As for the two zinc fingers of NCp7, two segments of intense NH/NH connectivities were observed in NCp10: A^{27}–D^{31} and A^{36}–C^{39}. The major differences between the NMR data of the NCp10 and NCp7 zinc fingers appeared in the middle-range NOE patterns, whereas the long-range NOEs were well conserved. Only three medium-range NOEs (N_i/N_{i+2} (28/30) and (36/38), α_i/β_{i+3} (36/39)), characteristic of regular secondary structure, were found for the NCp10 zinc finger instead of the 10 reported for each zinc finger of NCp7 (Summers et al., 1990; Omichinski et al., 1991; South et al., 1991; Morellet et al., 1992). Not enough medium-range NOES were present in any turn in the N-terminal part of the zinc finger, although all the slowly exchanging amide protons present at pH 8.0, and probably involved in hydrogen bonds in the NCp7 zinc fingers, are also present at pH 8.7 in the NCp10 zinc finger (NH 26, 28, 29, 30, 31, 32 and 39). Interestingly, the NH resonance of A^{27} is also still observable at pH 8.7, which was not

TABLE 1
^{1}H NMR CHEMICAL SHIFTS OF NCp10 AT 20 °C AND pH 5.5

Residue	Chemical shifts				Residue	Chemical shifts			
	NH	$C^{\alpha}H$	$C^{\beta}H$	Others		NH	$C^{\alpha}H$	$C^{\beta}H$	Others
Ala[1]		4.14	1.51		Lys[32]	8.49	4.49	1.75, 1.64	$C^{\gamma}H_2$ 1.30; $C^{\delta}H_2$ 1.47
Thr[2]	8.63	4.36	3.99	$C^{\gamma}H_3$ 1.18					$C^{\varepsilon}H_2$ 2.94
Val[3]	8.44	4.15	2.02	$C^{\gamma}H_3$ 0.90	Gly[33]	8.42	4.35, 3.65		
Val[4]	8.37	4.14	2.04	$C^{\gamma}H_3$ 0.95	His[34]	6.98	4.62	3.27, 3.21	$C^{\delta}H$ 7.00; $C^{\varepsilon}H$ 7.67
Ser[5]	8.49	4.43	3.85		Trp[35]	8.83	5.06	3.62, 3.23	$C^{\delta}H$ 7.34; $C^{\varepsilon}H$ 7.77
Gly[6]	8.53	3.96							$C^{\zeta}H$ 7.20, 7.49; $C^{\eta}H$ 7.24
Gln[7]	8.26	4.31	2.07, 1.95	$C^{\gamma}H_2$ 2.33					$N^{\varepsilon}H$ 10.23
Lys[8]	8.49	4.26	1.82, 1.75	$C^{\gamma}H_2$ 1.44, 1.38; $C^{\delta}H_2$ 1.66	Ala[36]	9.14	5.01	1.58	
				$C^{\varepsilon}H_2$ 2.98	Lys[37]	8.22	4.11	1.80, 1.67	$C^{\gamma}H_2$ 1.31; $C^{\delta}H_2$ 1.38
Glu[9]	8.51	4.33	2.10, 1.98	$C^{\gamma}H_2$ 2.34					$C^{\varepsilon}H_2$ 2.93, 2.85
Asp[10]	8.46	4.58	2.67, 2.02		Asp[38]	7.92	4.99	3.01, 2.62	
Arg[11]	8.40	4.31	1.82, 1.75	$C^{\gamma}H_2$ 1.64, 1.62; $C^{\delta}H_2$ 3.17	Cys[39]	7.63	4.16	3.53, 2.86	
				$N^{\varepsilon}H$ 7.21	Pro[40]		4.48	2.22, 1.95	$C^{\gamma}H_2$ 1.82; $C^{\delta}H_2$ 3.58, 3.28
Gln[12]	8.52	4.29	2.14, 2.01	$C^{\gamma}H_2$ 2.35	Lys[41]	9.11	4.31	2.09, 1.94	$C^{\delta}H_2$ 1.64; $C^{\gamma}H_2$ 1.50, 1.40
Gly[13]	8.19	3.97, 3.74							$C^{\varepsilon}H_2$ 3.02
Gly[14]	8.33	3.95			Lys[42]	7.04	3.90	1.53, 1.42	$C^{\gamma}H_2$ 0.93; $C^{\delta}H_2$ 1.35
Glu[15]	8.46	4.22	2.01, 1.94	$C^{\gamma}H_2$ 2.25					$C^{\varepsilon}H_2$ 2.78
Arg[16]	8.42	4.28	1.82, 1.75	$C^{\gamma}H_2$ 1.61; $C^{\delta}H_2$ 3.17	Pro[43]		4.34	2.24, 1.99	$C^{\gamma}H_2$ 1.95, 1.82
Arg[17]	8.36	4.29	1.83, 1.76	$C^{\gamma}H_2$ 1.62; $C^{\delta}H_2$ 3.17					$C^{\delta}H_2$ 3.71, 3.31
Arg[18]	8.42	4.32	1.84, 1.72	$C^{\gamma}H_2$ 1.66; $C^{\delta}H_2$ 3.19	Arg[44]	8.55	4.35	1.86, 1.73	$C^{\gamma}H_2$ 1.65; $C^{\delta}H_2$ 3.17
Ser[19]	8.43	4.44	3.90, 3.84						$N^{\varepsilon}H$ 7.16
Gln[20]	8.48	4.29	2.11, 1.98	2.35	Gly[45]	8.27	4.12, 4.02		
Leu[21]	8.21	4.27	1.58	$C^{\gamma}H$ 1.54; $C^{\delta}H$ 0.85, 0.79	Pro[46]		4.42	2.26, 1.98	$C^{\gamma}H_2$ 1.90; $C^{\delta}H_2$ 3.80, 3.61
Asp[22]	8.25	4.54	2.69, 2.61		Arg[47]	8.51	4.37	1.88, 1.75	$C^{\gamma}H_2$ 1.66; $C^{\delta}H_2$ 3.18
Arg[23]	8.18	4.19	1.80, 1.70	$C^{\gamma}H_2$ 1.50, 1.53; $C^{\delta}H_2$ 3.08					$N^{\varepsilon}H$ 7.18
				$N^{\varepsilon}H$ 7.12	Gly[48]	8.24	4.12, 4.03		
Asp[24]	8.22	4.42	2.47		Pro[49]		4.41	2.25, 1.98	$C^{\gamma}H_2$ 1.90; $C^{\delta}H_2$ 3.80, 3.60
Gln[25]	7.76	4.30	1.78, 1.65	$C^{\gamma}H_2$ 1.99, 2.01	Arg[50]	8.46	4.59	1.84, 1.72	$C^{\gamma}H_2$ 1.67; $C^{\delta}H_2$ 3.19
				$N^{\varepsilon}H$ 7.54, 6.87					$N^{\varepsilon}H$ 7.18
Cys[26]	8.48	4.14	2.90, 2.02		Pro[51]		4.42	2.29, 2.00	$C^{\gamma}H_2$ 1.89; $C^{\delta}H_2$ 3.79, 3.61
Ala[27]	8.78	4.14	1.44		Gln[52]	8.64	4.33	2.12, 2.01	$C^{\gamma}H_2$ 2.40
Tyr[28]	9.62	4.54	3.34, 3.18	$C^{\delta}H$ 7.35; $C^{\varepsilon}H$ 6.74	Thr[53]	8.08	4.34	4.22	$C^{\gamma}H_3$ 1.18
Cys[29]	9.13	4.15	3.35, 2.66		Ser[54]	8.32	4.46	3.85	
Lys[30]	7.84	4.14	2.20, 1.94	$C^{\gamma}H_2$ 1.31, 1.17	Leu[55]	8.37	4.36	1.62	$C^{\gamma}H_2$ 1.62; $C^{\delta}H_3$ 0.91, 0.84
				$C^{\delta}H_2$ 1.58, 1.49; $C^{\varepsilon}H_2$ 2.89	Leu[56]	7.76	4.16	1.55	$C^{\gamma}H_2$ 1.55; $C^{\delta}H_3$ 0.88, 0.84
Glu[31]	8.45	4.39	2.08, 2.02	$C^{\gamma}H_2$ 2.55, 2.35					

reported for the corresponding residue in HIV-1 zinc fingers. On the other hand, the C-terminal part of the NCp10 zinc finger exhibited $d(N_i, N_{i+2})$ and $d(\alpha_i, \beta_{i+3})$ NOE patterns characteristic of helical segments. Although not all the possible NOE patterns which describe this type of turn are present, the A^{36}–C^{39} segment could adopt the same 3_{10} helical corner already found in one of the two zinc fingers of NCp7. This was confirmed by the presence of the same exchanging amide protons at pH 8.7 in the NCp10 (NH 38, 39) and NCp7 zinc fingers and consequently, the same hydrogen bonds are probably present in this part of NC proteins. Most of the long-range NOEs of the two zinc fingers of NCp7 are also found in the case of the unique fingers of NCp10, as, for example: $H^{\beta}(C^{26})/NH(K^{30})$, $H^{\beta}(C^{26})/NH(E^{31})$, $H^{\beta}(C^{26})/H^{\varepsilon 1}(H^{34})$, $H^{\beta}(G^{31})/H^{\varepsilon 1}(H^{34})$. We also observed the NOEs between the side chains of Y^{28}-K^{41} and Y^{28}-P^{43}, corresponding to the NOEs reported for NCp7 between the side chains of N^{17}-A^{30} and N^{37}-R^{32}, which confirms the conservation of the mutual orientation of N- and C-terminal parts in both zinc fingers. Interestingly, a few NOEs were found in NCp10 that had no equivalent in NCp7 zinc fingers, namely NOEs involving the imidazole ring protons of H^{34} and the backbone protons of P^{40} and K^{41}, the side-chain protons of C^{39} and the side-chain protons of K^{42}, the aromatic protons of Y^{28} and the backbone protons of K^{42}.

P^{40}–L^{56}: C-terminal sequence

The sequential attribution obtained by using NH resonances is interrupted in the C-terminal region at the level of the proline residues (P^{40}, P^{43}, P^{46}, P^{49} and P^{51}). The medium-range NOE patterns and the slowly exchanging amide protons present at pH 8.7 (NH 41 and 42) support the presence of a conformationally restricted short domain (C^{39}–K^{42}), whereas medium and strong $d\alpha N$ connectivities, associated with small dNN cross peaks (Figs. 3B and 4) between successive amino acids, are characteristic of random coil structure in the P^{43}–L^{56} segment (Wright et al., 1988). Nevertheless, the possibility of a turn involving residues R^{50}–Q^{52} was suggested by the NOE $H^{\alpha}(R^{50})/H^{\beta}(Q^{52})$ (Fig. 5).

Structure calculations

A total of 597 NOE-derived distance constraints were used for DIANA calculations: 237 were intraresidual and 360 were interresidual. The latter set contained 179 sequential (i,i+1), 66 medium-range ($i,i+2 \geq j \geq i+4$) and 115 long-range ($i,j \geq i+5$) distance constraints. The long-range interactions were only observed in the (24–42) sequence, which constitutes the zinc finger domain (Fig. 6).

A set of 20 structures was generated using the distance geometry program DIANA and optimized using the program AMBER. The final total average energy was 85 ± 14 kcal mol^{-1} (Fig. 7).

The three-dimensional structure of (1–56)NCp10

Figure 8 shows the local variability and the related conformational space available for the backbone and side-chain atoms. A ribbon representation of five calculated structures is shown in Fig. 9. The zinc finger core (residues 26–39) displays a high degree of structural definition, with a backbone rmsd of 0.747 ± 0.209 Å and an rmsd of 1.709 ± 0.187 Å when all atoms are considered. The rmsd are higher for residues 1–23 and 43–56. It is interesting to note that low rmsd

values are observed for the side chain of the zinc ligand residues and for the two aromatic amino acids. Following the zinc finger domain is the ^{40}PKK sequence, which is also well defined with all atom rmsd values lower than 1.0 Å. Lower values of 0.946 ± 0.186 Å and 0.517 ± 0.211 Å are found for the backbone atoms of residues 13–17 and 49–52, respectively. The latter tetrapeptide is involved in a kink, defined by the dihedral angle values of R^{50} ($\phi = -70 \pm 18$; $\psi = 125 \pm 12$) and P^{51} ($\phi = -80 \pm 3$; $\psi = -5 \pm 23$). When the backbone atoms of the NCp10 zinc finger were superimposed onto the N- and C-terminal HIV-1 zinc fingers, the rmsd were found to be 1.427 ± 0.058 Å and 1.473 ± 0.059 Å, respectively.

Description of the (24–42) zinc finger domain

The 3D arrangement of the zinc finger is stabilized by numerous hydrogen bonds involving the backbone amide protons of residues 22, 23, 24, 26, 27, 28, 29, 35, 38 and 42, in agreement with the NMR data (see below). For residues 22, 23 and 24, the amide protons are involved in H-bonds with the oxygen atoms (OD1-22, OD2-22, O-20), (O-21, O-31) and (O31, O-32), respectively. The amide protons of the C^{26} and C^{29} residues form hydrogen bonds with the carbonyl oxygens of D^{24} and A^{27}. The NH^{27} is hydrogen bonded with the sulfur atom of the C^{26} residue. Residues C^{26}–Y^{28} form a γ-type local conformation, as can be concluded from a hydrogen bond between O^{26} and NH^{28} and the dihedral angle values of A^{27} ($\phi = -80 \pm 39$; $\psi = 78 \pm 82$).

The G^{33}–W^{35} segment forms a γ-turn, defined by the dihedral angle values of H^{34} ($\phi = 65 \pm 12$; $\psi = -37 \pm 27$) and a tight hydrogen bond between the O^{33} and NH^{35} atoms. It is of interest to note that the corresponding G^{22} and G^{43} residues in both NCp7 zinc fingers were reported to form a 3_{10} corner (Summers et al., 1990; South et al., 1991). The W^{35}–D^{38} residues are involved in a type-III turn, characterized by a hydrogen bond between the O^{35} and NH^{38} atoms, and the dihedral angle values of A^{36} ($\phi = -53 \pm 4$; $\psi = -37 \pm 7$) and K^{37} ($\phi = -57 \pm 3$; $\psi = 37 \pm 7$). A similar turn has not been reported for NCp7 zinc fingers (South et al., 1990; Summers et al., 1991) and it was not observed when the NCp10 (24–42) peptide was studied alone (unpublished results). The residues C^{39}–K^{42} form a type-I turn, with the dihedral angle values of P^{40} being ($\phi = -64 \pm 10$; $\psi = -25 \pm 33$) and K^{41} ($\phi = -111 \pm 26$; $\psi = 6 \pm 35$), which is stabilized by a hydrogen bond between O^{39} and NH^{42}.

It is interesting to note that no standard NH··S turns (Adman et al., 1975) are observed in this 3D zinc finger structure. We have previously noted that the NH··S turns observed in the NCp7 zinc finger peptides were not found in the intact protein (unpublished data). We observed the same phenomenon when we studied the (24–42) peptide alone and compared its local folding to the segment in the entire protein (not shown). This shows that the N- and C-terminal domains induce some local structural modifications, even if they do not modify the global folding of the zinc finger. This is in agreement with previously reported fluorescence data (Mély et al., 1993).

In the sequences of other retroviral NC proteins which have a unique zinc finger, the aromatic residues are well conserved: a tyrosine residue is found just before the second cysteine zinc ligand for the NC proteins of Feline leukemia virus type-2 (FeLV-2) (Copeland et al., 1984), Gibbon ape leukemia virus (GALV) (Delassus et al., 1989) and Spleen necrosis virus (SNV) (Weaver et al., 1990). Moreover, the location of a tryptophan residue just after the histidine zinc ligand is also conserved in these related viruses. It is interesting to note that, in our 3D model, the distance between the aromatic rings of the corresponding Y^{28} and W^{35} residues is about 10.847 ± 0.640 Å.

When the 24–42 peptide was studied alone, this distance increased to 14.908 ± 0.487 Å (unpublished results), further suggesting, as previously discussed, some local changes induced by the N-and C-terminal domains. These data are in agreement with fluorescence studies, which indicated a distance around 13 Å between Y^{28} and W^{35} (Mély et al., 1993). Both Y^{28} and W^{35} are on the surface of the zinc finger and situated on the same side (Fig. 10). The structural changes concerning the side-chain positions probably arise due to additional NMR-derived constraints, involving the side-chain protons of Y^{28} and P^{43} (Fig. 6). The aromatic rings of Y^{28} and P^{43} are 9.705 ± 0.160 Å apart, accounting for the experimental NOEs.

As shown in Fig. 11, the 3D folding of the CCHC box induced a proximity between Y^{28} and the K^{41} and K^{42} residues, while the distance between the basic K^{37} residue and W^{35} is about 4 Å. Moreover, all these residues located at the surface of the finger are situated on the same side (Figs. 10 and 11). The N- and C-terminal domains tend to turn away from this core. Nevertheless, residues R^{16}, R^{17} and R^{18} are found in the direction of Y^{28} for about 17 structures out of 20. The smallest distance between the ^{16}RRR segment and the zinc finger core is around 7 Å.

DISCUSSION AND CONCLUSIONS

The ^{1}H NMR-derived 3D structure of MoMuLV NCp10 in aqueous solution is characterized by a well-defined zinc finger domain, surrounded by flexible N- and C-terminal regions. The zinc complexation leads to a large scattering of proton chemical shifts, especially those corresponding to the amide, α- and aromatic protons, and to a pattern of medium- and long-range NOEs similar to that reported for the two zinc finger domains of HIV-1 NCp7 (Summers et al., 1990; South et al., 1991; Morellet et al., 1992; Omichinski et al., 1992). This particular zinc-induced folding structure of the CCHC box very likely plays an important role, since mutations of any zinc ligand residue of MoMuLV NCp10 induced a strong impairment of genomic RNA packaging (Gorelick et al., 1988). Likewise, mutations affecting the cysteine or histidine residues were reported to be lethal in the case of Rous Sarcoma virus (RSV) NC protein (Dupraz et al., 1990).

In addition, point mutations affecting various conserved amino acids other than the zinc ligands in CCHC boxes of NC proteins, such as replacement of Tyr^{28} or Trp^{35} in NCp10 by a glycine residue, led to an inhibition of virus replication through alteration of viral genomic RNA packaging (Méric and Goff, 1989). The distance between the Tyr^{28} and Trp^{35} aromatic rings, located on the solvent-accessible surface of the quasispherical zinc finger core of the protein, is about 10.8 Å (Fig. 10). The same distance was also found between two aromatic rings in drugs able to intercalate or bis-intercalate into DNA, following the usual excluded-site model (Delbarre et al., 1987). Interestingly, a recent ^{1}H NMR study of the interaction between the N-terminal zinc finger of NCp7 and the oligonucleotide d(ACGCC), present at the 3′ end of an identified segment of the Psi packaging signal, showed that the phenylalanine and isoleucine residues, which correspond to Tyr^{28} and Trp^{35} in NCp10, form a cleft, making the Phe^{16} side chain a possible candidate for intercalation between C-G or A-C base steps (South and Summers, 1993).

The replacement by valine of the glycine residue positioned just before the histidine zinc ligand (Gly^{33} in NCp10) and highly conserved in NC proteins was shown to induce a decrease in the genomic RNA encapsidation (Méric and Goff, 1989). In our 3D model, Gly^{33} is involved in a γ-turn conformation, while in both the N- and C-terminal zinc fingers of HIV-1 NCp7, the corresponding Gly^{22} and Gly^{43} were shown to form a 3_{10} corner in the polypeptide chain (unpub-

lished results). In addition to its tendency to be involved in a turn-like conformation, the inherent flexibility of glycine may be important to allow correct spatial orientations of critical residues, such as Tyr^{28} and Trp^{35}. Changes of this particular glycine in a sterically larger residue in RSV NC led to virus particles containing RNA material, but unable to replicate (Dupraz et al., 1990). These modifications probably induce conformational constraints incompatible with the biologically active structure of the CCHC box.

In addition to zinc finger domains, most of the retroviral NC proteins possess a large number of basic residues. The Lys^{41} and Lys^{42} residues which follow the zinc finger sequence in MoMuLV NCp10 and which are part of the minimal peptide sequences retaining full NCp10 activity in vitro (De Rocquigny et al., 1993) have been reported to be critical for virus infectivity (Housset et al., 1993). Thus, mutants in which both Lys^{41} and Lys^{42} were replaced by neutral residues were replication defective (Housset et al., 1993). It is interesting to observe that, in the 3D structure of NCp10, these biologically relevant basic amino acids are located on the surface of the zinc finger and in the proximity of the Tyr^{28} residue (Fig. 9).

Thus, the Tyr^{28}, Trp^{35}, Lys^{37}, Lys^{41} and Lys^{42} residues constitute a sort of bulge on which a nucleic acid sequence could coil, the complex being stabilized by salt bridges between the charged lysines and the phosphate groups of the nucleotides and possible intercalation of the aromatic residues into the single-stranded genomic RNA (Karpel et al., 1987). Moreover, Van der Waals interactions between the hydrophobic side chains of the residues belonging to the bulge and RNA bases could stabilize the protein–RNA complex (Summers et al., 1992). The size of the bulged structure is compatible with fluorescence studies, indicating that NCp10 could cover about six residues along the nucleic acid chain when stoichiometrically bound to single-stranded RNA (Karpel et al., 1987).

In the solvated structure of NCp10, the strongly constrained zinc finger domain appears flanked by rather flexible N- and C-terminal domains. Nevertheless, the NOEs between the α-proton of R^{50} and the β- and γ-protons of Q^{52} suggest that the C-terminal pentapeptide Pro^{49}-Arg-Pro-Gln-Thr^{53}, which has been shown to be necessary for complete NCp10 function in vitro (De Rocquigny et al., 1993), is probably involved in a bent structure.

It has been reported that the addition of the N- and C-terminal sequences to the finger domain of NCp10 increased the zinc affinity by about one order of magnitude (Mély et al., 1993). This is probably due to the increased structural constraints induced in the zinc finger by the surrounding sequences (unpublished results). However, fluorescence studies indicated that the C-terminal chain was not in direct or indirect interaction with the Trp^{35} environment (Mély et al., 1993), a result which is confirmed by our data showing that the C-terminal chain did not contact the biologically relevant domain defined by Trp^{35}, Tyr^{28} and the triplet of lysine residues.

Given these findings, and because of their flexibility, the domains surrounding the zinc finger may function as arms, anchoring the viral RNA onto the NCp10 bulged structure. In NCp7, this spatial arrangement appears to be a necessary condition for full in vitro activity, since synthetic peptides corresponding to the N- and C-terminal domains no longer have in vitro activity if they are not linked at least by a flexible Gly-Gly spacer (De Rocquigny et al., 1993). This C-terminal domain of HIV-1 NCp7 was also found to contain small sequences, endowed with some elements of secondary structure (Morellet et al., 1994), supporting similar structure–function relationships between both proteins. These domains, rich in proline and basic amino acids, have been identified as structural motifs, able to interact with the minor groove of A · T-rich sequences of DNA

(Reeves and Nissen, 1990). The sequence of NCp10 following the CCHC box, i.e. K^{41}KPR-GPRGPRP^{51}, has a number of analogies with the nonlinear sequence V^{13}K- R^{30}APRKKG^{35}-T^{50} of NCp7, which forms a spatially continuous domain induced by the folding of the two successive zinc fingers.

The 3D structure of the NCp10 zinc finger is very similar (global folding), if not identical (local modifications), to those reported for the two zinc finger domains of HIV-1 NCp7. Structure–activity relationships tend to indicate that both the zinc finger and the basic residues are involved synergistically in NCp10 activities (De Rocquigny et al., 1993), as was also proposed for HIV-1 NCp7 (De Rocquigny et al., 1991; Morellet et al., 1994). Thus, it has been recently suggested that an additional function of NC protein is to protect, in a histone-like manner, RNA from degradation by nucleases in the virions and that this function requires the presence of the CCHC box (Aronoff et al., 1993; Bowles et al., 1993). Our bulge-like model induced by the zinc finger folding may explain the ability of NCp10 to bind and protect nucleic acids. Since similar structural and biological properties have been reported for HIV-1 NCp7 and MoMuLV NCp10, it is of interest to use the NCp10 murine model to better understand the HIV-1 NCp7 functions. Investigations are currently in progress to understand more fully NCp10–RNA interactions and particularly to define the NCp10 zinc finger role and the possible participation of Tyr^{28} and Trp^{35} residues in nucleic acid interaction.

ACKNOWLEDGEMENTS

The authors would like to thank C. Dupuis for her assistance in preparing the manuscript and N. Goudreau and A. Beaumont for critical reading and stylistic revision. Particular acknowledgements are due to both referees for their pertinent suggestions. This work was supported by the Association Nationale de Recherche sur le SIDA and the Institut de Formation Supérieure Biomédicale.

REFERENCES

Adman, E., Watenpaugh, K.D. and Jensen, L.H. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 4854–4858.

Aronoff, R., Hajjar, A.M. and Linial, M.L. (1993) *J. Virol.*, **67**, 178–188.

Arseniev, A., Schultze, P., Wörgötter, T., Braun, W., Wagner, G., Vasak, M., Kägi, J.H.R. and Wüthrich, K. (1988) *J. Mol. Biol.*, **201**, 637–657.

Berg, J.M. (1986) *Science*, **232**, 485–487.

Bieth, E., Gabus, C. and Darlix, J.L. (1990) *Nucleic Acids Res.*, **18**, 119–127.

Bowles, N.E., Damay, P. and Spahr, P.F. (1993) *J. Virol.*, **67**, 623–631.

Braunschweiler, L. and Ernst, R.R. (1983) *J. Magn. Reson.*, **43**, 521–528.

Chance, M.R., Sagi, I., Wirt, M.D., Frisbie, S.M., Scheuring, E., Bess, J.W., Henderson, L.E., Arthur, L.O., South, T.L., Perez-Alvarado, G. and Summers, M.F. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 10041–10045.

Coffin, J.M. (1984) In *RNA Tumor Viruses*, Vol. I (Eds, Weiss, R., Teich, N., Varmus, H. and Coffin, J.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., pp. 261–368.

Copeland, T.D., Morgan, M.A. and Oroszlan, S. (1984) *Virology*, **133**, 137–145.

Cornille, F., Mély, Y., Ficheux, D., Salvignol, I., Gérard, D., Darlix, J.L., Fournié-Zaluski, M.C. and Roques, B.P. (1990) *Int. J. Pept. Protein Res.*, **36**, 551–558.

Covey, S.N. (1986) *Nucleic Acids Res.*, **14**, 623–633.

Darlix, J.L., Gabus, C., Nugeyre, H.T., Clavel, F. and Barré-Sinoussi, F. (1990) *J. Mol. Biol.*, **216**, 689–699.

Davis, D.G. and Bax, A. (1985) *J. Am. Chem. Soc.*, **107**, 2821–2822.

Delassus, S., Sonigo, P. and Wain-Hobson, S. (1989) *Virology*, **173**, 205–213.
Delbarre, A., Delepierre, M., Garbay, C., Igolen, J., Le Pecq, J.B. and Roques, B.P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2155–2159.
De Rocquigny, H., Ficheux, D., Gabus, C., Fournié-Zaluski, M.C., Darlix, J.L. and Roques, B.P. (1991) *Biochem. Biophys. Res. Commun.*, **180**, 1010–1018.
De Rocquigny, H., Gabus, C., Vincent, A., Fournié-Zaluski, M.C., Roques, B.P. and Darlix, J.L. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6472–6476.
De Rocquigny, H., Ficheux, D., Gabus, C., Allain, B., Fournié-Zaluski, M.C., Darlix, J.L. and Roques, B.P. (1993) *Nucleic Acids Res.*, **21**, 823–829.
Dupraz, P., Oertle, S., Méric, C., Damay, P. and Spahr, P.F. (1990) *J. Virol.*, **64**, 4978–4987.
Gorelick, R.J. Henderson, L.E., Hanser, J.P. and Rein, A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 8420–8424.
Green, L.M. and Berg, J.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 6403–6407.
Güntert, P., Braun, W. and Wüthrich, K. (1991a) *J. Mol. Biol.*, **217**, 517–530.
Güntert, P., Qian, Y.Q., Otting, O., Müller, M., Gehring, W. and Wüthrich, K. (1991b) *J. Mol. Biol.*, **217**, 531–540.
Housset, V., De Rocquigny, H., Roques, B.P. and Darlix, J.L. (1993) *J. Virol.*, **67**, 2537–2545.
Jacob, O. (1990) Ph.D. Thesis, Université Louis Pasteur, Strasbourg.
Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.
Karpel, R.L., Henderson, L.E. and Groszlan, S. (1987) *J. Biol. Chem.*, **262**, 4961–4997.
Kochoyan, M., Havel, T.F., Nguyen, D., Dahl, C.E., Keutmann, H.T. and Weiss, M.A. (1991) *Biochemistry*, **30**, 3371–3386.
Macura, S., Huang, Y., Suter, D. and Ernst, R.R. (1981) *J. Magn. Reson.*, **43**, 259–281.
Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
Mély, Y., Cornille, F., Fournié-Zaluski, M.C., Darlix, J.L., Roques, B.P. and Gérard, D. (1991) *Biopolymers*, **31**, 899–906.
Mély, Y., De Rocquigny, H., Piémont, E., Déméné, H., Jullian, N., Fournié-Zaluski, M.C., Roques, B.P. and Gérard, D. (1993) *Biochim. Biophys. Acta*, **1161**, 6–18.
Méric, C. and Spahr, P.F. (1986) *J. Virol.*, **60**, 450–459.
Méric, C. and Goff, S.J. (1989) *J. Virol.*, **63**, 1558–1568.
Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J.L. and Roques, B.P. (1992) *EMBO J.*, **11**, 3059–3065.
Morellet, N., De Rocquigny, H., Mély, Y., Jullian, N., Déméné, H., Ottmann, M., Gérard, D., Darlix, J.L., Fournié-Zaluski, M.C. and Roques, B.P. (1994) *J. Mol. Biol.*, **235**, 287–301.
Omichinski, J.C., Clore, G.M., Sakaguchi, K., Appella, E. and Gronenborn, A.M. (1991) *FEBS Lett.*, **292**, 25–30.
Prats, A.C., Sarih, L., Gabus, C., Litvak, S., Keith, G. and Darlix, J.L. (1988) *EMBO J.*, **7**, 1777–1783.
Prats, A.C., Roy, C., Wang, P., Erard, M., Houset, V., Gabus, C., Paoletti, C. and Darlix, J.L. (1990) *J. Virol.*, **64**, 774–783.
Prats, A.C., Housset, V., De Billy, G., Cornille, F., Prats, H., Roques B.P. and Darlix, J.L. (1991) *Nucleic Acids Res.*, **13**, 3533–3541.
Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 479–485.
Reeves, R. and Nissen, M.S. (1990) *J. Biol. Chem.*, **265**, 8573–8582.
Roberts, W.J., Pan, T., Eliott, J.I., Coleman, J.E. and Williams, K.R. (1989) *Biochemistry*, **28**, 10043–10047.
Singh, U.C., Weiner, P.K., Caldwell, J. and Kollman, P.A. (1986) *AMBER* (version 4.0), School of Pharmacy, University of California, San Francisco, CA.
South, T.L., Blake, P.R., Hare, D.R. and Summers, M.F. (1991) *Biochemistry*, **30**, 6342–6349.
South, T.L. and Summers, M.F. (1993) *Protein Sci.*, **2**, 3–19.
Summers, M.F., South, T.L., Kim, B. and Hare, D.R. (1990) *Biochemistry*, **29**, 329–340.
Summers, M.F., Henderson, L.E., Chance, M.R., Bess Jr., J.W., South, T.L., Blake, P.R. Sagi, I., Perez-Alvarado, G., Sowder III, R.C., Hare, D.R. and Arthur, L.O. (1992) *Protein Sci.*, **1**, 563–574.
Weaver, T.A., Talbot, K.J. and Panganiban, A.T. (1990) *J. Virol.*, **64**, 2642–2652.
Wright, P.E., Dyson, H.J. and Lerner, R.A. (1988) *Biochemistry*, **27**, 7167–7175.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.