J. Biochem. **141**, 269–277 (2007)
doi:10.1093/jb/mvm037

RNA Recognition Mechanism of the Minimal Active Domain of the Human Immunodeficiency Virus Type-2 Nucleocapsid Protein

Takashi Matsui[1], Yoshio Kodera[1,*], Hiroshi Endoh[1], Emi Miyauchi[1], Hiroyoshi Komatsu[2], Kazuki Sato[3], Takeshi Tanaka[4], Toshiyuki Kohno[4] and Tadakazu Maeda[1]

[1]Department of Physics, School of Science, and [2]Department of Immunology, School of Allied Health Sciences, Kitasato University, Sagamihara, Kanagawa 228-8555; [3]Department of Environmental Science, School of Human Environmental Science, Fukuoka Women's University, Higashi-ku, Fukuoka 813-8529; and [4]Mitsubishi Kagaku Institute of Life Sciences (MITILS), Machida, Tokyo 194-8511, Japan

Received October 14, 2006; accepted December 10, 2006; published online January 3, 2007

NCp8 of HIV-2 contains two CCHC-type zinc fingers connected by a linker, and is involved in many critical steps of the virus life cycle. It was previously shown that the first zinc finger flanked by the linker is the minimal active domain for specific binding to viral RNA. In our previous study, we determined the three-dimensional structure of NCp8-f1, including the minimal active domain, and found that a hydrogen bond between Asn^{11} N$^{\delta}$H and Arg^{27} O stabilized the conformation of the linker in the vicinity of the zinc finger [Kodera *et al.* (1998) *Biochemistry* 37, 17704–17713]. In this study, RNA binding activities of NCp8-f1 and three types of its mutant peptides were analysed by native PAGE assay. The activity and three-dimensional structure of NCp8-f1/N11A, in which alanine is substituted for Asn^{11} thereby affecting the conformation of the linker, was analyzed and compared with those of NCp8-f1. We demonstrated that the existence of Arg^{4} and/or Lys^{5} and Arg^{26} and/or Arg^{27} were necessary for binding RNA. Furthermore, the linker's flexible orientation, which is controlled by the hydrogen bond between Asn^{11} N$^{\delta}$H and Arg^{27} O, appears to be a structural basis for NCp8 existing as a multi-functional protein.

Key words: HIV, NMR, nucleocapsid protein, RNA binding protein, zinc finger.

Abbreviations: CCHC, Cys-X_2-Cys-X_4-His-X_4-Cys; DQF-COSY, double quantum filtered-correlation spectroscopy; DSS, 4,4-dimethyl-4-silapentane-1-sulphonic acid; HIV, human immunodeficiency virus; HIV-1, HIV type-1; HIV-2, HIV type-2; HPLC, high performance liquid chromatography; MoMuLV, moloney murine leukemia virus; NC, nucleocapsid; NCp7, HIV-1 NC protein; NCp8, HIV-2 NC protein; NCp10, MoMuLV NC protein; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; NMR, nuclear magnetic resonance; PAGE, polyacrylamide gel electrophoresis; RMSD, root-mean-squared deviation; SL, stem loop; TOCSY, total correlation spectroscopy; ψ site, packaging signal site; WATERGATE, water suppression by gradient-tailored excitation.

Nucleocapsid proteins (NC proteins) are structural proteins of retroviruses that are generated upon processing of the *gag* polyprotein by the viral protease. NC proteins are essential for various steps in retroviral replication. NC proteins specifically discriminate between the viral RNA and cellular RNAs, thereby facilitating the dimerization and packaging of two viral RNA genomes into maturing viruses. In addition, these proteins possess many other functions: they enhance binding of the tRNA primer to the viral RNA, stimulate the synthesis of full-length DNA by reverse transcriptase, improve the efficiency of integrase and protect the viral RNA against nucleases (*1, 2*). Each retroviral NC protein has either one or two conserved zinc finger domain(s), with the amino acid sequence Cys-X_2-Cys-X_4-His-X_4-Cys, which coordinates one zinc ion with very high affinity (X represents amino acids that vary among the different retroviruses) (*3, 4*).

Two types of human immunodeficiency virus (HIV) have been reported, HIV-1 and HIV-2, which contain NC proteins NCp7 and NCp8, respectively (Fig. 1). Both NC proteins, which are 67% identical, contain two zinc fingers connected by a seven amino acid residue linker. In the case of NCp7, several genetic studies have shown that mutations either of the basic amino acid residues in the linker or in the first zinc finger result in the production of non-infectious viral particles (*8, 9*). Moreover, site-directed mutagenesis revealed that substituting a leucine residue for Pro^{31} resulted in the formation of non-infectious and immature viral particles (*10*). *In vitro* investigations demonstrated that the binding of the second zinc finger flanked by the linker with viral RNA was very weak and non-specific, whereas that of the first zinc finger flanked by the linker is sufficient for specific RNA binding (*11–13*). Furthermore, three-dimensional structures of NCp7 and NCp7-related fragments were determined by ^{1}H nuclear magnetic resonance (NMR) spectroscopy (*14–17*). These studies showed that similar folding occurred for the two zinc fingers in the presence of zinc, consistent with the structures determined in previous studies of the NCp7

*To whom correspondence should be addressed. Tel: +81-42-778-9540, Fax: +81-42-778-9541, E-mail: kodera@kitasato-u.ac.jp

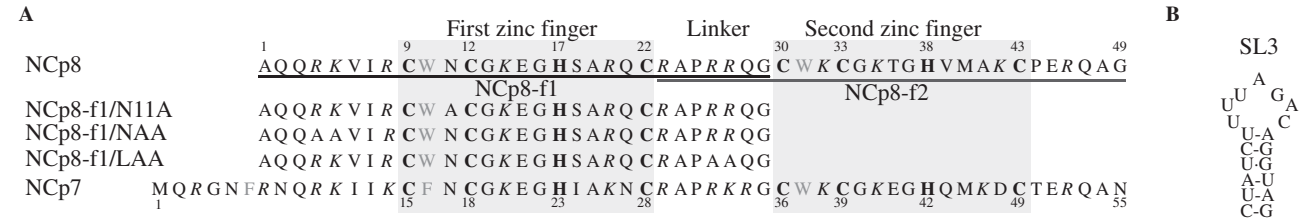

Fig. 1. (A) **Amino acid sequences of the NC protein of HIV-2 (NCp8), HIV-1 (NCp7), NCp8-derived and mutant peptides.** Symbols of basic amino acid residues and zinc-coordinated amino acid residues are indicated by italic and bold characters, respectively. Aromatic amino acid residues are colored grey. (B) Nucleotide sequence and secondary structure of SL3 in HIV-2GH-1 (*5*), which shares homology with SL3 in HIV-2ROD (*29*). This secondary structure was calculated by *m*-fold (*6*, *7*).

zinc fingers (*18, 19*). The dynamical behaviour of NCp7 indicates that the conformation and solution behaviour of the linker of NCp7 is not stable and might be a consequence of a rapid equilibrium between several different conformations (*20*). On the other hand, the binding of (12–53)NCp7 to a single-strand d(ACGCC) pentanucleotide DNA led to a reduction of the conformational flexibility that characterized (12–53)NCp7. The amplitudes of high-frequency motions were restrained in the complex and the occurrence of conformational exchange was displaced from the second zinc finger to the linker residue Ala^{30} (*21*). Moreover, three-dimensional structures of the NCp7-packaging signal site (ψ site) stem loop 2 ($SL2_{HIV\text{-}1}$) complex (*22*) and the NCp7-stem loop 3 ($SL3_{HIV\text{-}1}$) complex (*23*) were determined. Substantial differences between these structures include the relative orientations of the first and second zinc fingers and the mode of 3^{10} helix binding, highlighting NC's adaptive RNA binding capability (*22*).

A number of studies of the biological function and structure of NCp7 have been published (*2*). Moreover, those of NCp10 from Moloney murine leukaemia virus (MoMuLV) (*24*) and NCp8 from simian immunodeficiency virus (*25, 26*) have been studied, whereas only a few studies have been done for NCp8 of HIV-2. In one previous study of NCp8 that was done, a competitive UV cross-linking assay of NCp8 and NCp8-derived synthetic (*27*) peptides showed that the peptide corresponding to either the first or second zinc finger flanked by the linker interacts specifically with viral RNA corresponding to a candidate sequence for the ψ site (*27, 28*). RNA footprinting experiments with NCp7 bound to HIV-2 RNA found that a high-affinity binding site observed in the HIV-2 RNA leader mapped to stem loop 3 ($SL3_{HIV\text{-}2}$). These results indicated that $SL3_{HIV\text{-}2}$ was the primary binding site in HIV-2 RNA for NC proteins (*29*). In our previous study, the three-dimensional structure of NCp8-f1, a 29 amino acid peptide that includes the minimal active domain of NCp8, revealed that the conformation of the linker was stabilized by a hydrogen bond between Asn^{11} $N^{\delta}H$ in the zinc finger and Arg^{27} O in the linker (*30*).

In this study, the binding activities of two mutant peptides, NCp8-f1/NAA and NCp8-f1/LAA, where two alanine residues were substituted for Arg^4-Lys^5 in the N-terminal of NCp8-f1 and two alanine residues were substituted for Arg^{26}-Arg^{27} in the linker of NCp8-f1, were measured. Moreover, the binding activity and the three-dimensional structure of the mutant peptide NCp8-f1/N11A, in which an alanine residue was substituted for Asn^{11}, thereby affecting the conformation of the linker, were also determined. From these structural and functional experiments, the RNA recognition mechanism of the minimal active domain of NCp8 is discussed, based on comparison with NCp7.

MATERIALS AND METHODS

Preparation of NCp8-derived Peptides—All NCp8-derived peptides (NCp8-f1, NCp8-f1/N11A, NCp8-f1/NAA, NCp8-f1/LAA and NCp8-f2) and NCp8 were obtained commercially from Anygen Inc. (Kwangju, Korea). NCp8-f2 was a 23–49 amino acid residue peptide and included both the linker and the second zinc finger. The purities of these peptides exceeded 95%.

Preparation of ψ Site $SL3_{HIV\text{-}2}$ of Viral RNA—The corresponding region of viral RNA $SL3_{HIV\text{-}2}$ was obtained commercially from Japan Bio Services Co., Ltd. (Saitama, Japan). The RNA was labelled with fluorescein isothiocyanate at its 5′-end and purified by reversed-phase high performance liquid chromatography (HPLC). The purity of RNA exceeded 95%.

NC Titration Experiment—Native PAGE assay was performed using a modification of a previously described procedure (*31*). Native PAGE assays to evaluate NC binding were carried out using samples containing 2 μM $SL3_{HIV\text{-}2}$ and varying NC protein concentrations (0, 2, 4, 6, 8, 10, 14, 20 and 40 μM). $SL3_{HIV\text{-}2}$ samples were heated to 97°C for 3 min and then stored at 37°C for 30 min. Titration experiments were performed in reaction buffer [30 mM Tris–HCl (pH 7.5), 60 mM NaCl, 200 μM $MgCl_2$, 2 mM DTT and 2.5 unit RNase inhibitor (Wako Pure Chemical Industries Ltd., Osaka, Japan)]. All NC samples contained 1.1 equivalents of zinc ions for each zinc finger, and coordination of zinc ion(s) was confirmed by 2D NMR spectra. The titration samples were incubated for 15 min at 37°C prior to electrophoresis. Native 8% (v/v) PAGE gels were pre-run at 210 V for 10 min in 0.5× Tris-borate-EDTA (TBE) buffer (pH 8.0) at 4°C prior to sample loading. The gels were imaged with Typhoon 9400 (GE Healthcare Bio-sciences Corp., NJ, USA). The analysis of imaged data was performed using the ProFINDER 2D version 2004 (PerkinElmer Inc., MA, USA).

NMR Spectroscopy—The samples for NMR experiments each contained 5.0 mM NCp8-f1/N11A and 5.5 mM $ZnCl_2$ in 90% H_2O/10% ^2H_2O at pH 5.8.

All NMR measurements were carried out on a Bruker DMX-500 spectrometer at temperatures of 15, 25 and 35°C.

Nuclear Overhauser effect spectroscopy (NOESY) spectra (*32, 33*) were recorded with mixing times of 80, 100, 200 and 300 ms. Total correlation spectroscopy (TOCSY) spectra were recorded using a MLEV-17 pulse scheme (*34*) with isotropic mixing times of 50 and 80 ms. The suppression of the solvent resonance was achieved using the water suppression by gradient-tailored excitation (WATERGATE) scheme in both the NOESY and TOCSY measurements (*35*). Double-quantum filtered-correlation spectroscopy (DQF-COSY) (*36*) spectra were recorded to obtain the constraints for the torsion angles. The data sizes used for acquisition were 512 (t_1) × 8192 (t_2) for DQF-COSY and 512 × 2048 otherwise, and the spectral widths were 6250 Hz each.

All NMR spectra were processed using the program package Azara (*37*). Phase-shifted sine-squared window functions were applied prior to Fourier transformation, with shifts from $\pi/3$ to $\pi/2$ in both dimensions. Final matrix sizes were 1024 × 8192 real points for DQF-COSY and 2048 × 2048 otherwise, in the F_1 and F_2 dimensions, respectively.

To detect the slowly exchanging backbone amide protons, we recorded a NOESY spectrum with a mixing time of 200 ms at pH 8.8. The intensity of each intra-residue NH-C$^{\alpha}$H cross-peak was then compared with that obtained at pH 5.8. The chemical shifts of C$^{\alpha}$H were not significantly changed when the pH was changed from 5.8 to 8.8.

Chemical shifts were referenced to the methyl resonance of 4,4-dimethyl-4-silapentane-1-sulphonic acid (DSS) used as an internal standard. Complete sets of 2D spectra were recorded at 15°C and pH 5.8.

Experimental Constraints and Structure Calculations—All of the spectra were analysed with a combination of customized macroprograms from the Ansig Version 3.3 software (*38, 39*). Sequence-specific resonance assignments were achieved according to the standard method established by Wüthrich and co-workers (*40*). All resonance assignments are indicated in supplementary data 1. Inter-proton distance restraints were obtained from the NOESY spectra with a mixing time of 200 ms. All nuclear Overhauser effect (NOE) data were divided into four classes, strong, medium, weak and very weak, corresponding to distance upper limits of 2.5, 3.5, 5.0 and 6.0 Å, respectively, in the inter-proton distance restraints. Pseudoatoms were used for non-stereospecifically assigned protons, and intra-residue and long-range correcting factors were added to the distance restraints (*41*). In addition, 0.5 Å was added to the upper limits for distance restraints involving methyl protons (*42*).

The backbone NH-C$^{\alpha}$H coupling constants were estimated from the DQF-COSY spectrum, and were converted to the backbone torsion angle ϕ constraints according to the following rules (*43*): For $^3J_{\mathrm{NH\text{-}c}\alpha_H}$ smaller than 5.5 Hz and larger than 8.0 Hz, the ϕ angles were constrained in the ranges of $-65^\circ \pm 25^\circ$ and $-120^\circ \pm 40^\circ$, respectively. A total of nine backbone torsion ϕ angles for NCp8-f1/N11A were constrained: one residue (Cys^{22}) with $^3J_{\mathrm{NH\text{-}C}\alpha_H}$ smaller than 5.5 Hz and eight residues (Ile^7, Arg^8, Cys^{12}, Lys^{14}, Glu^{15}, Ser^{18}, Gln^{21} and Arg^{23}) with $^3J_{\mathrm{NH\text{-}C}\alpha_H}$ larger than 8.0 Hz. Backbone dihedral constraints were not applied for residues with $^3J_{\mathrm{NH\text{-}C}\alpha_H}$ values between 5.5 and 8.0 Hz.

All calculations were carried out on a Linux workstation with the XPLOR-NIH version 2.9.9 (*44*). The three-dimensional structure of NCp8-f1/N11A in solution was calculated on the basis of nine dihedral angle constraints and 371 distance constraints. The 371 distance constraints included 146 intra-residue and 225 inter-residue NOE distance constraints, and an additional four constraints between the zinc ion and the sulphur atoms of Cys^9, Cys^{12} and Cys^{22} and between the zinc ion and the N$^{\varepsilon}$ of His^{17} (*18, 30*). We carried out simulated annealing calculations starting with 100 initial random structures, and selected 13 final structures that were in good agreement with NMR experimental constraints, for which the NOE distance and torsion angle violations were smaller than 0.5 Å and 5°, respectively.

Evaluation Methods—The convergence of the calculated structures was evaluated in terms of the structural parameters, that is, the root-mean-squared deviation (RMSD) from the experimental distance and dihedral constraints, the values of the energy statistics (F_{NOE}, F_{tor}, F_{repel} and $F_{\mathrm{L\text{-}J}}$), and the RMSD from idealized geometry. The structures were analysed using PROCHECK_NMR (*45*) and MOLMOL (*46*). The distributions of backbone dihedral angles of the final converged structures were evaluated using the representation of the Ramachandran dihedral pattern, which indicates the deviations from the sterically allowed (ϕ, ψ) angle limits (*47*). The degrees of angular variation between the converged structures were further assessed by using an angular order parameter (*48*).

RESULTS AND DISCUSSION

NC Titration Experiments of the Minimal Active Domain and its Mutant Peptides—Fig. 2, A–F shows the binding activities of NC peptides inferred by native PAGE data obtained for $SL3_{\mathrm{HIV\text{-}2}}$ as a function of added NCp8-f1, NCp8-f1/N11A, NCp8-f1/NAA, NCp8-f1/LAA, NCp8 and NCp8-f2, respectively. Addition of NC proteins resulted in the appearance of a new band that was predicted for an NC-$SL3_{\mathrm{HIV\text{-}2}}$ complex. Fig. 2G summarizes the analysis of the gel shift data for the six types of peptides shown in Fig. 2, A–F. The data points are based on the amount of bound RNA, which was calculated as the difference between the fluorescence intensity of the total $SL3_{\mathrm{HIV\text{-}2}}$ in the absence of NC and that of free $SL3_{\mathrm{HIV\text{-}2}}$ after addition of NC protein, and standard deviations versus the NC protein:$SL3_{\mathrm{HIV\text{-}2}}$ ratios for eight independent experiments. These experiments showed that the binding activities of NCp8-f1 and NCp8-f1/N11A are the same, and are greater than those of NCp8-f2 but weaker than that of NCp8. For both NCp8-f1/NAA and NCp8-f1/LAA, the band intensity of the free $SL3_{\mathrm{HIV\text{-}2}}$ did not change even if the peptide concentration was increased up to a peptide:$SL3_{\mathrm{HIV\text{-}2}}$ ratio of 20:1. These results indicate that the basic amino acid residues Arg^4 and/or Lys^5 and Arg^{26} and/or Arg^{27} are necessary for binding to $SL3_{\mathrm{HIV\text{-}2}}$.

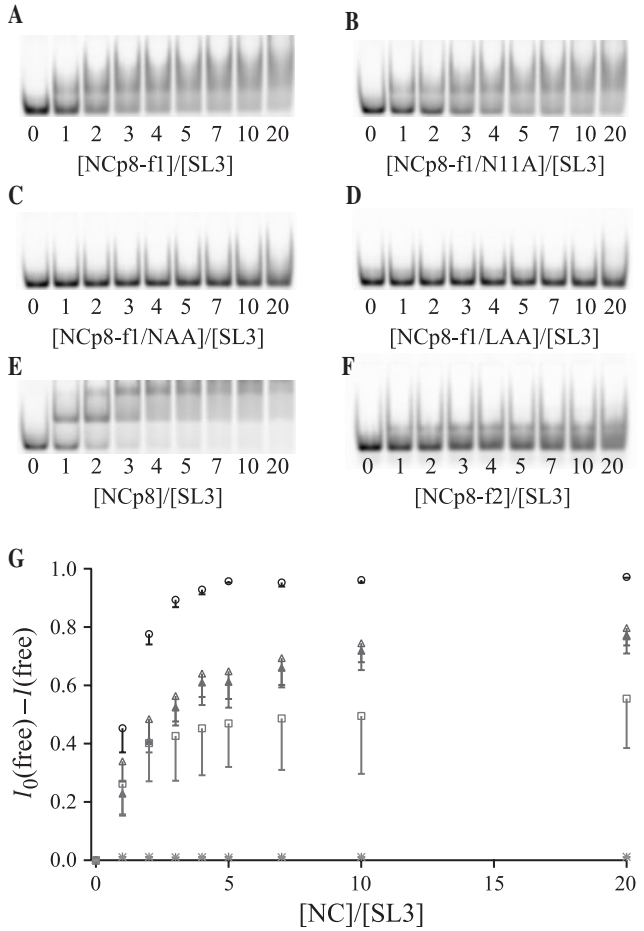

Fig. 2. **Native PAGE NC titration results obtained for (A) NCp8-f1:$SL3_{HIV-2}$, (B) NCp8-f1/N11A:$SL3_{HIV-2}$, (C) NCp8-f1/NAA:$SL3_{HIV-2}$, (D) NCp8-f1/LAA:$SL3_{HIV-2}$, (E) NCp8:$SL3_{HIV-2}$ and (F) NCp8-f2:$SL3_{HIV-2}$.** The gel images were excited at 488 nm and the resulting fluorescent emission was detected using Emission Filter 520 BP 40 (GE Healthcare Bio-sciences Corp., NJ, USA). The amount of bound RNA was calculated as the difference between the total $SL3_{HIV-2}$ intensity in the absence of NC minus the total free $SL3_{HIV-2}$ after addition of NC. (G) The amount of bound $SL3_{HIV-2}$ is represented by the mean and standard deviation of several independent experiments (independent experiments for NCp8 was only three times, otherwise it was eight times.): open circle, NCp8; open triangle, NCp8-f1; filled triangle, NCp8-f1/N11A; open rectangle, NCp8-f2; plus, NCp8-f1/NAA; cross, NCp8-f1/LAA; error bars represent the relevant standard deviations.

The binding and folding of the N-terminal domain in NCp7 is likely driven by the optimization of electrostatic interaction between the 3^{10} helix and the phosphate strand of the RNA duplex (*22, 23*). Moreover, computational methods revealed that a number of Lys and Arg residues in the 3^{10} helix and zinc finger of NCp7 were critical for binding to $SL3_{HIV-1}$ (*49*). For NCp8, this study demonstrated for the first time that the N-terminal region in NCp8, which would be too short to form a similar helix, was also important for binding to $SL3_{HIV-2}$. Arg^{26} and Arg^{27} in the linker, which are conserved in all strains of NCp7 and NCp8, were also necessary to bind $SL3_{HIV-1}$. It has been shown for NCp7 that the aromatic residue in the first zinc finger, which forms a hydrophobic cleft, and the basic residue in the helical conformation of the first zinc finger interact with viral RNA (*22, 23*). These facts and results of this study indicate that Arg^4 and/or Lys^5, Arg^{26} and/or Arg^{27}, Trp^{10} and Arg^{20} may collaborate in the binding to $SL3_{HIV-2}$.

Other studies also showed that point mutations in the linker in which a leucine was substituted for Pro^{31} and D-Pro for L-Pro^{31} led to a non-infectious virus and poor RNA dimerization, respectively (*10, 14*). These studies suggested that the conformation of the linker was important for both RNA binding and biological function. On the other hand, in our previous study, the conformation of the linker was stabilized by a hydrogen bond between Asn^{11} $N^{\delta}H$ in the zinc finger and Arg^{27} O in the linker (*30*). However, the RNA binding activity of NCp8-f1/N11A, in which alanine is substituted for Asn^{11} thereby affecting the hydrogen bond between the zinc finger and the linker, was the same as that of NCp8-f1. To help clarify this linker conformation issue, in the present study we determined the structure of NCp8-f1/N11A and discuss the structure–activity relationships of the linker of NCp8-f1.

Structure of NCp8-f1/N11A and Structural Comparison with NCp8-f1—Fig. 3A shows a stereopair representation of the best-fit superposition of the backbone atoms (N, C^{α} and C) of all of the residues and the side-chains of Cys^9, Ala^{11}, Cys^{12}, His^{17} and Cys^{22} for the 13 convergent structures of NCp8-f1/N11A, which were obtained by best-fitting superposition of the backbone atom (N, C^{α}, C and O) coordinates for residues 9–22. Statistics for the converged structures were evaluated in terms of the structural parameters (Table 1). The deviations from idealized covalent geometry were very small, and the Lennard-Jones van der Waals energies were negative, indicating that there were no distortions and no non-bonded bad contacts in the converged structures.

The N-terminal region of NCp8-f1/N11A, Ala^1-Lys^5, did not converge well. The conformation of the N-terminal region of the zinc finger appears to be stable, as suggested by the observation of HN-$C^{\alpha}H$ NOESY cross-peaks for Ala^{11}-Lys^{14} even at pH 8.8 (Supplementary Data 2). It is possible that this region is stabilized by the same hydrogen bond network as that previously found in NCp7 (*18*). The Ala^{19}-Gln^{21} region adopts a helical conformation. Such a conformation is also found in the corresponding regions of NCp8-f1 (*30*) and NCp7 (*19*) (Fig. 1). The Ala^{24}-Pro^{25} bond was considered to be in the *trans* conformation, because some NOEs, Ala^{24} $C^{\alpha}H$ to Pro^{25} $C^{\delta}H$ and Ala^{24} NH to Pro^{25} $C^{\delta}H$, were observed in the NOESY spectrum. Moreover, those intensities did not change when the temperature was increased from 5°C to 25°C (data not shown). However, the C-terminal segment (Arg^{26}-Gly^{29}) in the linker was not restricted by NMR constraints, reflecting the randomness of the conformationally disordered backbone. This may be the result of a lack of medium- and long-range NOE constraints due to the inherent flexibility of the segment.

Fig. 3, B and C shows the best-fit superposition of these structures in NCp8-f1/N11A (black) and in

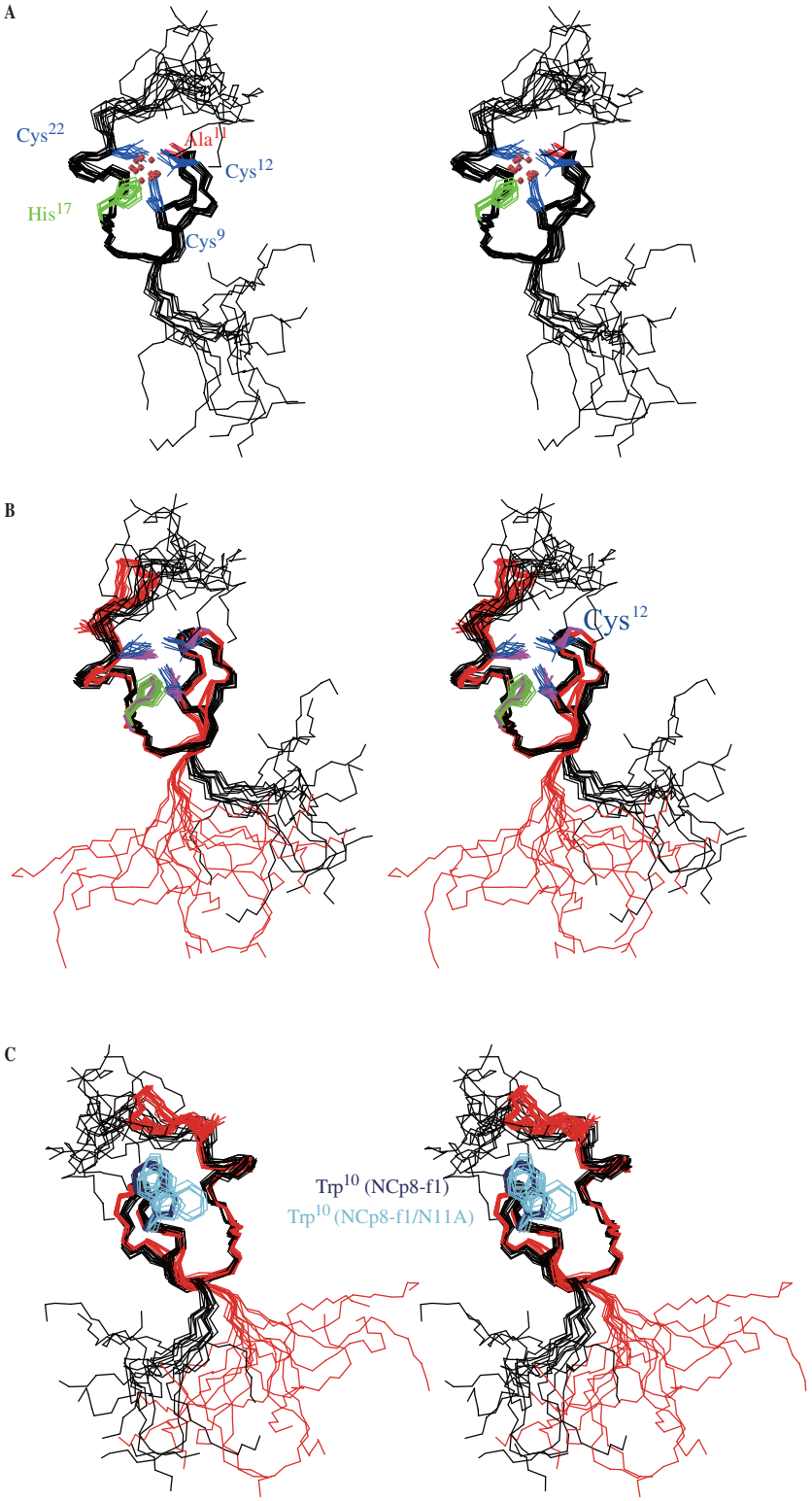

Fig. 3. **(A) Stereopairs of backbone heavy atoms (N, C^{α} and C) for the 13 converged structures of NCp8-f1/N11A.** These are the results of the best-fit superposition of the backbone heavy atoms for all of the residues of the molecule. Zinc-coordinated side-chains are coloured blue (Cys^{9}, Cys^{12} and Cys^{22}) and green (His^{17}). The side-chain of Ala^{11}, which was substituted for Asn^{11} in NCp8-f1, was coloured red. Zinc ions are indicated as red balls. (B) and (C) Conformational comparison of NCp8-f1 (PDB ID: 1NC8) (*30*) and NCp8-f1/N11A: (B) Stereopairs of backbone heavy atoms (N, C^{α} and C) and the side-chains bound with zinc ion for the 15 and 13 converged structures of NCp8-f1 (red) and NCp8-f1/N11A (black), respectively. The location of Cys^{12} is indicated. (C) Obtained by a 90 degree rotation about the vertical axis of stereopair (A) and (B). The side-chains of Trp^{10} in NCp8-f1 and in NCp8-f1/N11A are coloured blue and sky blue, respectively.

Table 1. **Structural statistics for the 13 converged structures of NCp8-f1/N11A[a].**

RMSD from experimental distance constraints (Å)[a]	
All (371)	0.020 ± 0.003
Intraresidue (146)	0.017 ± 0.007
Sequential (98)	0.023 ± 0.008
Medium range ($\|i-j\|<5$) (60)	0.019 ± 0.006
Long range ($\|i-j\| \geq 5$) (67)	0.020 ± 0.006
RMSD from experimental dihedral constraints (deg)[a] (9)	0.922 ± 0.298
Energetic statistics (kcal/mol)[b]	
F_{NOE}	2.53 ± 0.45
F_{tor}	39.16 ± 2.94
F_{repel}	6.39 ± 2.28
$F_{L\text{-}J}$	-28.6 ± 15.8
RMSD from idealized geometry	
Bonds (Å)	0.002 ± 0.0002
Angle (deg)	0.553 ± 0.021
Impropers (deg)	0.384 ± 0.028
Ramachandran analysis (%) (residues 9–22)	
Most favoured regions	68.6
Additionally allowed regions	25.6
Generously allowed regions	5.8
Disallowed regions	0
Average RMSD (Å) (residues 9–22)	
Backbone (N, C^{α} and C)	0.34 ± 0.05
All heavy atoms	1.14 ± 0.13
Average RMSD (Å) (residues 23–29)	
Backbone (N, C^{α} and C)	1.35 ± 0.24
All heavy atoms	3.11 ± 0.58

[a]The values in the table are given as mean ± standard deviation. The number of each experimental constraint used in the calculations is given in parentheses.
[b]F_{NOE}, F_{tor}, F_{repel} and $F_{L\text{-}J}$ are the energies related to the NOE violations, the torsion angle violations, the van der Waals repulsion term and Lennard-Jones van der Waals energy, respectively. The values of the force constants used for these terms are the standard values, as depicted in the XPLOR-NIH (*44*).

NCp8-f1 (red), which were fitted using the backbone atom (N, C^{α} and C) coordinates in the zinc finger regions. For the zinc finger, the conformations of the N-terminal regions are similar to each other. However, the backbone orientation of Cys^{12} and the convergence of side-chain conformations of Trp^{10} are rather different from each other. These differences were almost the same as that between NCp8-f1 and NCp7 in which there was no hydrogen bond between Asn^{17} (Asn^{11} in NCp8) and the linker. These results suggested that the conformational change of the linker induced by the loss of hydrogen bonding caused a slight alteration of the zinc finger structure.

The most significant difference between the two structures lies in the linker: the linker of NCp8-f1 adopts a well-defined conformation, whereas that of NCp8/N11A does not. NOESY cross-peaks that were observed between the atoms of Arg^{27}-Arg^{29} and those in the zinc finger for NCp8-f1 (Trp^{10}-Arg^{27}, Trp^{10}-Gly^{29}, Asn^{11}-Arg^{27} and Arg^{20}-Gln^{28}) disappeared in NCp8-f1/N11A, presumably due to the substitution of alanine for Asn^{11}. Most probably, this substitution resulted in a loss of hydrogen bonding, which allowed the linker to move far away from the N-terminal region of the zinc finger. Fig. 4 shows the finger print region of the NOESY spectra of NCp8-f1 (red) and NCp8-f1/N11A (black). The residues indicated by black and red characters showed clear differences in chemical shifts between NCp8-f1 and NCp8-f1/N11A. In particular, for the Trp^{10}, Ala^{11} and Arg^{27} residues, large chemical shift differences between NCp8-f1/N11A and NCp8-f1 were observed. These changes in chemical shift are consistent with the above-mentioned structural difference between NCp8-f1 and NCp8-f1/N11A.

The conformation of the linker in NCp7 was discussed in a published ^{15}N NMR relaxation study, in which it was concluded that the conformation of the linker was not completely random, but rather a rapid equilibrium between several conformational substates (*20*). For NCp7, the corresponding asparagine residue (Asn^{17}) forms a hydrogen bond with Lys^{33} in NCp7 (Arg^{27} in NCp8) that appears in the $SL3_{HIV\text{-}1}$ complex (*23*), but not in the $SL2_{HIV\text{-}1}$ complex (*22*). The linker's conformation in NCp7, under the $SL2_{HIV\text{-}1}$ or $SL3_{HIV\text{-}1}$ complex state, corresponds to one of the several substable conformations with rapid equilibrium that characterize NCp7. For NCp8-f1 (*30*) and NCp8-f1/N11A, Pro^{25} of both peptides forms a *trans* conformation, like that found in NCp7, in which Pro^{31} induces a bend (*14*). For residues 23–26 in the linker of NCp8-f1/N11A, rather intense medium-range NOEs were observed, except for Pro^{25} (Supplementary Data 2). The conformation of Cys^{22}-Pro^{25} of NCp8-f1/N11A was not completely dispersed (Fig. 3). From these results, it is possible that the linker's conformation in NCp8-f1/N11A also is not completely random, but rather in rapid equilibrium, which is the same kind of feature as in NCp7.

In our previous study, the conformation of the linker of NCp8-f1 was stabilized by a hydrogen bond between Asn^{11} and Arg^{27}. Since the interaction of a hydrogen bond is weak, this bond in NCp8-f1 would be broken due to the interaction between $SL3_{HIV\text{-}2}$ and the linker of NCp8-f1, which is thought to behave like that of NCp8-f1/N11A, the latter having several substable conformations. It is conceivable that both NCp8-f1 and NCp8-f1/N11A are able to form complexes with $SL3_{HIV\text{-}2}$, in which the conformation of the linkers are stabilized in one of the conformational substates and Arg^{26} and/or Arg^{27} are placed in a suitable position to stabilize the complex. With regard to the necessity of the hydrogen bond, that issue still remains unresolved in this study. However, it is possible that the hydrogen bond is an essential conformational factor for complex formation between NCp8-f1 and the other RNA sites. The asparagine residue located at the third amino acid residue in the first zinc finger (Asn^{11} of NCp8-f1) is well-conserved in the NC proteins that include two zinc fingers connected with a short linker (*1*). From this fact and our results, the conformational variety of the linker, one of which is induced by the hydrogen bond, may be an essential factor for NC proteins being multi-functional proteins.

On the other hand, the binding activity of NCp8-f2 to $SL3_{HIV\text{-}2}$ was lower than that of both NCp8-f1 and NCp8-f1/N11A (Fig. 2, F and G). This result showed that NCp8-f2 was unable to bind to $SL3_{HIV\text{-}2}$ either in the

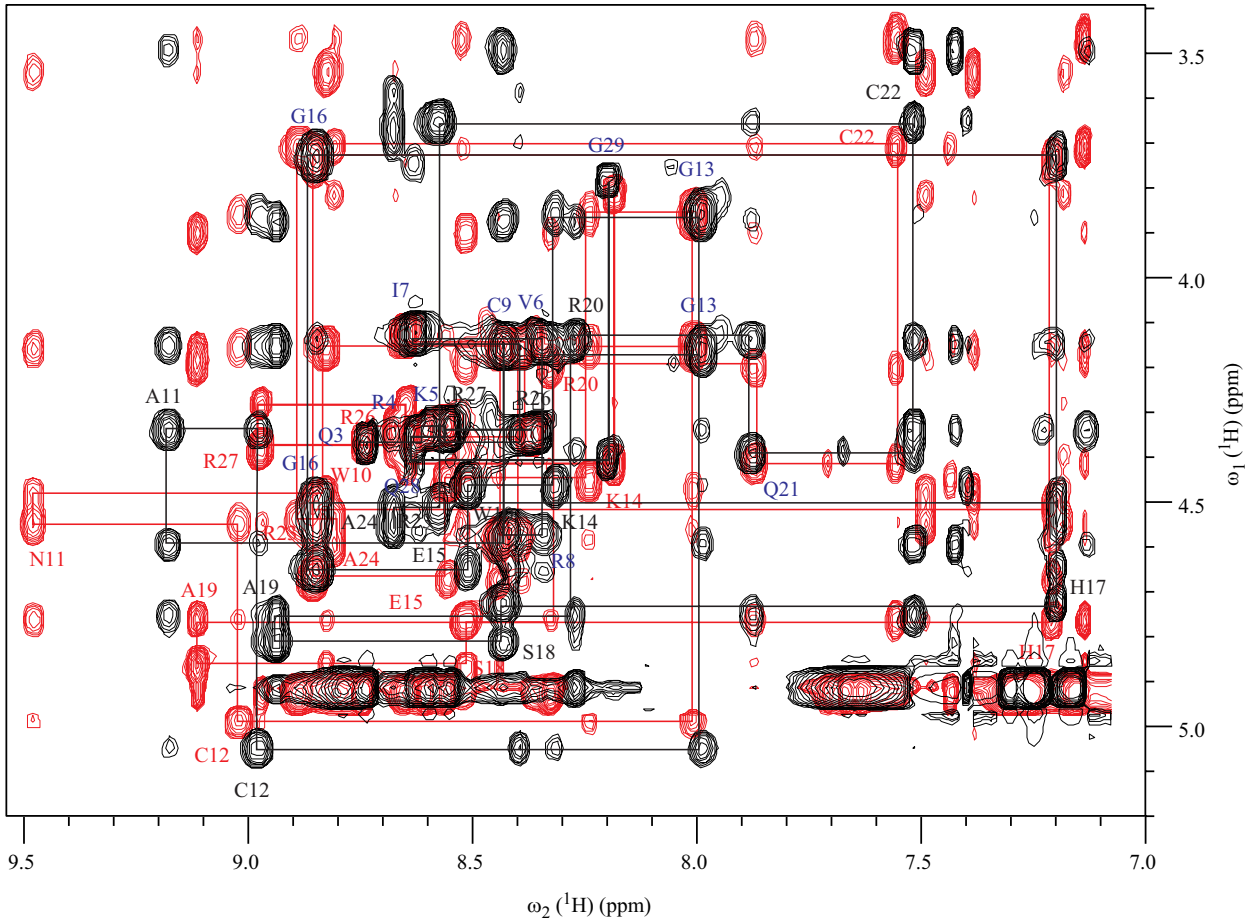

Fig. 4. **Portions of the 500 MHz 2D NMR spectra of NCp8-f1 (red) (30) and NCp8-f1/N11A (black) at pH 5.8 and 15°C.** Sequential $d_{\alpha N}(i, i+1)$ NOE connectivities for residues 1–29 in the NOESY spectra observed with a mixing time of 200 ms (red line: NCp8-f1, black line: NCp8-f1/N11A). Intra-residue NH-C$^{\alpha}$H crosspeaks are labelled with the residue number by standard single-letter amino acid abbreviations (red: NCp8-f1, black: NCp8-f1/N11A). The overlapped cross-peaks between NCp8-f1 and NCp8-f1/N11A are indicated by blue characters.

same manner as NCp8-f1 and NCp8-f1/N11A or intrinsically, even though it includes a zinc finger and basic amino acid residues in its N-terminal site (Arg^{26} and Arg^{27} in the linker) and in its C-terminal site (Arg^{46}). However, the binding activity of full-length NCp8 was higher than that of either NCp8-f1 or NCp8-f1/N11A. These results suggest that the linker also plays an important role in determining the orientation of the second zinc finger, which enables NCp8 to bind specifically to $SL3_{HIV-2}$ with high affinity. It is expected that further experiments using full-length NCp8 and other binding sites of RNA will give us additional important insights into the functions of NC proteins.

The structures of NCp8-f1/N11A have been registered in the Protein Data Bank under accession code 2DI2. This study was supported in part by the Ministry of Education, Culture, Sports, Science and Technology of Japan, a grant for the National Project on Protein Structural and Function Analyses and Kitasato University Research Grants for Young Researchers. We thank Professor Jae Il Kim at Kwangju Institute of Science and Technology and Ms Kuniko Kobayashi at Mitsubishi Kagaku Institute of Life Sciences for useful discussions.

REFERENCES

1. Darlix, J.L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B.P. (1995) First glimpses at structure–function relationships of the nucleocapsid protein of retroviruses. *J. Mol. Biol.* **254**, 523–537
2. Levin, J.G., Guo, J., Rouzina, I., and Musier-Forsyth, K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* **80**, 217–286
3. Henderson, L.E., Copeland, T.D., Sowder, R.C., Smythers, G.W., and Oroszlan, S. (1981) Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses. *J. Biol. Chem.* **256**, 8400–8406
4. Covey, S.N. (1986) Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. *Nucleic Acids Res.* **14**, 623–633
5. Hasegawa, A., Tsujimoto, H., Maki, N., Ishikawa, K., Miura, T., Fukasawa, M., Miki, K., and Hayami, M. (1989)

Genomic divergence of HIV-2 from ghana. *AIDS Res. Hum. Retroviruses* **5**, 593–604
6. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415
7. Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940
8. Dupraz, P., Oertle, S., Meric, C., Damay, P., and Spahr, P.F. (1990) Point mutations in the proximal cys-his box of rous sarcoma virus nucleocapsid protein. *J. Virol.* **64**, 4978–4987
9. Fu, X.D., Katz, R.A., Skalka, A.M., and Leis, J. (1988) Site-directed mutagenesis of the avian retrovirus nucleocapsid protein, pp 12. mutation which affects RNA binding in vitro blocks viral replication. *J. Biol. Chem.* **263**, 2140–2145
10. Ottmann, M., Gabus, C., and Darlix, J.L. (1995) The central globular domain of the nucleocapsid protein of human immunodeficiency virus type 1 is critical for virion structure and infectivity. *J. Virol.* **69**, 1778–1784
11. De Rocquigny, H., Gabus, C., Vincent, A., Fournie-Zaluski, M.C., Roques, B., and Darlix, J.L. (1992) Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc. Natl. Acad. Sci. USA* **89**, 6472–6476
12. Lapadat-Tapolsky, M., De Rocquigny, H., Van Gent, D., Roques, B., Plasterk, R., and Darlix, J.L. (1993) Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. *Nucleic Acids Res.* **21**, 831–839
13. Dannull, J., Surovoy, A., Jung, G., and Moelling, K. (1994) Specific binding of HIV-1 nucleocapsid protein to PSI RNA in vitro requires N-terminal zinc finger and flanking basic amino acid residues. *EMBO J.* **13**, 1525–1533
14. Morellet, N., de Rocquigny, H., Mély, Y., Jullian, N., Déméné, H., Ottmann, M., Gerard, D., Darlix, J.L., Fournie-Zaluski, M.C., and Roques, B.P. (1994) Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. *J. Mol. Biol.* **235**, 287–301
15. Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J.L., and Roques, B.P. (1992) Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR. *EMBO J.* **11**, 3059–3065
16. Omichinski, J.G., Clore, G.M., Sakaguchi, K., Appella, E., and 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
17. Summers, M.F., Henderson, L.E., Chance, M.R., Bess, J.W. Jr., South, T.L., Blake, P.R., Sagi, I., Perez-Alvarado, G., Sowder, R.C. 3rd., and Hare, D.R. (1992) Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. *Protein Sci.* **1**, 563–574
18. Summers, M.F., South, T.L., Kim, B., and Hare, D.R. (1990) High-resolution structure of an HIV zinc fingerlike domain via a new NMR-based distance geometry approach. *Biochemistry* **29**, 329–340
19. South, T.L., Blake, P.R., Hare, D.R., and 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
20. Lee, B.M., De Guzman, R.N., Turner, B.G., Tjandra, N., and Summers, M.F. (1998) Dynamical behavior of the HIV-1 nucleocapsid protein. *J. Mol. Biol.* **279**, 633–649
21. Ramboarina, S., Srividya, N., Atkinson, R.A., Morellet, N., Roques, B.P., Lefevre, J.F., Mély, Y., and Kieffer, B. (2002) Effects of temperature on the dynamic behaviour of the HIV-1 nucleocapsid NCp7 and its DNA complex. *J. Mol. Biol.* **316**, 611–627
22. Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., Chancellor, K.J., Wu, Z.R., and Summers, M.F. (2000) NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* **301**, 491–511
23. De Guzman, R.N., Wu, Z.R., Stalling, C.C., Pappalardo, L., Borer, P.N., and Summers, M.F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **279**, 384–388
24. D'Souza, V. and Summers, M.F. (2004) Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* **431**, 586–590
25. Urbaneja, M.A., McGrath, C.F., Kane, B.P., Henderson, L.E., and Casas-Finet, J.R. (2000) Nucleic acid binding properties of the simian immunodeficiency virus nucleocapsid protein NCp8. *J. Biol. Chem.* **275**, 10394–10404
26. Morellet, N., Meudal, H., Bouaziz, S., and Roques, B.P. (2006) Structure of the zinc finger domain encompassing residues 13–51 of the nucleocapsid protein from simian immunodeficiency virus. *Biochem. J.* **393**, 725–732
27. Komatsu, H., Tsukahara, T., and Tozawa, H. (1996) Viral RNA binding properties of human immunodeficiency virus type-2 (HIV-2) nucleocapsid protein-derived synthetic peptides. *Biochem. Mol. Biol. Int.* **38**, 1143–1154
28. Tsukahara, T., Komatsu, H., Kubo, M., Obata, F., and Tozawa, H. (1996) Binding properties of human immunodeficiency virus type-2 (HIV-2) RNA corresponding to the packaging signal to its nucleocapsid protein. *Biochem. Mol. Biol. Int.* **40**, 33–42
29. Damgaard, C.K., Dyhr-Mikkelsen, H., and Kjems, J. (1998) Mapping the RNA binding sites for human immunodeficiency virus type-1 gag and NC proteins within the complete HIV-1 and -2 untranslated leader regions. *Nucleic Acids Res.* **26**, 3667–3676
30. Kodera, Y., Sato, K., Tsukahara, T., Komatsu, H., Maeda, T., and Kohno, T. (1998) High-resolution solution NMR structure of the minimal active domain of the human immunodeficiency virus type-2 nucleocapsid protein. *Biochemistry* **37**, 17704–17713
31. Amarasinghe, G.K., Zhou, J., Miskimon, M., Chancellor, K.J., McDonald, J.A., Matthews, A.G., Miller, R.R., Rouse, M.D., and Summers, M.F. (2001) Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein. structural studies and implications for genome recognition. *J. Mol. Biol.* **314**, 961–970
32. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553
33. Macura, S., Huang, Y., Suter, D., and dErnst, R.R. (1981) Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.* **43**, 259–281
34. Bax, A. and Davis, D.G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **65**, 355–360
35. Piotto, M., Saudek, V., and Sklenár, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665
36. Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R., and Wüthrich, K. (1983) Improved spectral resolution in COSY 1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479–485

37. Boucher, W. Azara 2.7 software package. http://www.bio.cam.ac.uk/azara/
38. Kraulis, P.J. (1989) ANSIG: a program for the assignment of protein 1H 2D NMR spectra by interactive computer graphics, *J. Magn. Reson.* **84**, 627–633
39. Kraulis, P.J., Domaille, P.J., Campbell-Burk, S.L., Van Aken, T., and Laue, E.D. (1994) Solution structure and dynamics of ras p21.GDP determined by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry* **33**, 3515–3531
40. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids.* Eiley, New York
41. Wüthrich, K., Billeter, M., and Braun, W. (1983) Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton–proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949–961
42. Hyberts, S.G., Marki, W., and Wagner, G. (1987) Stereospecific assignments of side-chain protons and characterization of torsion angles in eglin c. *Eur. J. Biochem.* **164**, 625–635
43. Pardi, A., Billeter, M., and Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C alpha proton coupling constants, 3JHN alpha, in a globular protein. Use of 3JHN alpha for identification of helical secondary structure. *J. Mol. Biol.* **180**, 741–751
44. Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. (2003) The xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73
45. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486
46. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–5, 29–32
47. Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99
48. Hyberts, S.G., Goldberg, M.S., Havel, T.F., and Wagner, G. (1992) The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures. *Protein Sci.* **1**, 736–751
49. Khandogin, J., Musier-Forsyth, K., and York, D.M. (2003) Insights into the regioselectivity and RNA-binding affinity of HIV-1 nucleocapsid protein from linear-scaling quantum methods. *J. Mol. Biol.* **330**, 993–1004