4314 *Biochemistry* **2009,** *48,* 4314–4323
DOI:10.1021/bi802364b

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

The RNA Recognition Mechanism of Human Immunodeficiency Virus (HIV) Type 2 NCp8 Is Different from That of HIV-1 NCp7[†,‡]

Takashi Matsui,[§,∥] Takeshi Tanaka,[∥] Hiroshi Endoh,[§] Kazuki Sato,[⊥] Hidekazu Tanaka,[§] Emi Miyauchi,[§] Yusuke Kawashima,[§] Misa Nagai-Makabe,[§] Hiroyoshi Komatsu,[@] Toshiyuki Kohno,[∥] Tadakazu Maeda,[§] and Yoshio Kodera*[,§]

[§]*Department of Physics, School of Science, Kitasato University, Sagamihara, Kanagawa 228-8555, Japan,* [∥]*Mitsubishi Kagaku Institute of Life Sciences (MITILS), Machida, Tokyo 194-8511, Japan,* [⊥]*Department of Environmental Science, School of Science, Fukuoka Women's University, Higashi-ku, Fukuoka 813-8529, Japan, and* [@]*Department of Clinical Laboratory Medicine, Faculty of Health Science Technology, Bunkyo Gakuin University, Bunkyo-ku, Tokyo 113-0023, Japan*

Received December 29, 2008; Revised Manuscript Received March 30, 2009

ABSTRACT: The nucleocapsid (NC) protein of HIV, which contains two CCHC-type zinc fingers connected by a linker, is a multifunctional protein involved in many of the critical steps of the HIV life cycle. HIV-1 and HIV-2 contain NC proteins NCp7 and NCp8, respectively. The amino acid sequences of both NC proteins are 67% identical. For NCp7, the important elements for RNA binding were found to be the first zinc finger flanked by the linker, as the minimal active domain, and the 3_{10} helix in the N-terminus, as the secondary active domain. However, for the NCp8 counterpart in HIV-2, the mechanism for binding to viral RNA has not yet been clarified. In this study, we determined NCp8's three-dimensional structure for the first time and examined the dynamic behavior and chemical shift perturbation as a function of the concentration of viral RNA SL3. Moreover, the specific binding activities of NCp8 and the NCp8-derived peptides with SL3 were examined by a native polyacrylamide gel electrophoresis assay. These results indicate that the RNA recognition mechanism for NCp8 is different from that of NCp7 and that the hydrophobic cleft in the second zinc finger acts as a secondary active domain instead of the 3_{10} helix in NCp7. Furthermore, the flexibility of the linker is limited by the hydrogen bond between the first zinc finger (Asn11) and the linker (Arg27), which makes it possible for the sites around Trp10 in the minimal active domain and the secondary active domain to form the binding surface.

The nucleocapsid (NC) protein is a retrovirus structural protein that is generated during the processing of the *gag* polyprotein by viral protease. This protein is multifunctional and is involved in the various steps of retroviral replication. The NC protein specifically discriminates between viral and cellular RNAs, thereby facilitating the dimerization and packaging of the two viral RNA genomes into the maturing viruses. In addition, the protein possesses many other functions (*1*, *2*). Each retroviral NC protein has either one or two conserved CCHC-type zinc finger domains that have the amino acid sequence Cys-X_2-Cys-X_4-His-X_4-Cys. This sequence is able to coordinate one zinc ion with very high affinity (X represents the amino acids that vary among the different retroviruses) (*3*, *4*).

There are two types of human immunodeficiency virus (HIV) that have been reported, HIV-1 and HIV-2, containing NC proteins NCp7 and NCp8, respectively (Figure 1a). These NC proteins are 67% identical and contain two zinc fingers connected by a linker. In the case of NCp7, several *in vitro* investigations of the HIV-1 NCp7-derived peptides have demonstrated that the reactivities of the second zinc finger itself and the zinc finger flanked by the linker are very weak and nonspecific with regard to viral RNA binding. In contrast, this first zinc finger flanked by the

[†]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.

[‡]Atomic coordinates for the 11 converged structures and the chemical shifts have been deposited in the Protein Data Bank as entry 2EC7. The chemical shifts and the ^{15}N relaxation data were deposited in the BioMagResBank as entry 15364.

*To whom correspondence should be addressed: Department of Physics, School of Science, Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan. Telephone: +81-42-778-9541. Fax: +81-42-778-9541. E-mail: kodera@kitasato-u.ac.jp.

[1]Abbreviations: FITC, fluorescein isothiocyanate; HIV, human immunodeficiency virus; HIV-1, HIV type 1; HIV-2, HIV type 2; HSQC, heteronuclear single-quantum correlation spectroscopy; NC, nucleocapsid; NCp7, NC protein of HIV-1; NCp8, NC protein of HIV-2; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; ψ site, packaging signal site in HIV-2 RNA; PAGE, polyacrylamide gel electrophoresis; rmsd, root-mean-square deviation; SD, splice donor site in the ψ site; SL3, stem−loop 3 in the ψ site; TOCSY, total correlation spectroscopy.

 Published on Web 3/31/2009

FIGURE 1: (a) Amino acid sequences for the nucleocapsid proteins of HIV-2 (NCp8) and HIV-1 (NCp7). The italic and bold characters indicate the basic and zinc-coordinated amino acid residues, respectively. Aromatic amino acid residues are highlighted in gray. (b) Nucleotide sequence and secondary structure of HIV-2_{GH-1} (*39*). This secondary structure was calculated with m-fold (*40*, *41*).

linker has been shown to be sufficient for specific binding (*5*−*7*). With respect to the structure of the complexes between the NCp7 and RNA sites, the 3_{10} helix in the N-terminus also has an important role with regard to binding to the RNA sites (*8*, *9*). If there are suitable orientations of the 3_{10} helix and the first zinc finger, NCp7 can bind with similar affinity to both RNA stem−loop 2 (SL2) and stem−loop 3 (SL3) (*2*, *8*, *9*). This conformational behavior, which acts via various subsets of molecular interactions, enables NCp7 to possess the multiple functions that are necessary for recognizing different RNA sites.

On the other hand, the N-terminal segment of NCp8 is too short to form a similar 3_{10} helix. Moreover, a competitive UV cross-linking assay for the NCp8 and NCp8-derived synthetic peptides (*10*) has shown that the peptide corresponding to either the first zinc finger flanked by the linker or the second zinc finger flanked by the linker can interact specifically with viral RNA. RNA footprinting experiments with NCp7 bound to HIV-2 RNA have indicated that SL3 and the splice donor site (SD) are the primary and minor binding sites in HIV-2 RNA for the NC proteins, respectively (*11*). In our previous study, we demonstrated that NCp8-f1, the 29-residue peptide that includes the first zinc finger and the linker in NCp8, played an important part in the process that makes binding to SL3 and SD possible (*12*, *13*). Thus, the minimal active domain, which includes the first zinc finger and the linker in NCp7 and NCp8, has a significant role in specific RNA binding (*5*−*7*, *10*) and acts as the primary active domain. In contrast, we have also found that NCp8-f2, which includes the second zinc finger and the linker, specifically binds to the SD (*13*).

In this study, the three-dimensional structure of NCp8 was determined for the first time. We also examined the structural dynamics and the chemical shift perturbation as a function of the concentration of viral RNA SL3. Moreover, the specific binding activities of NCp8 and the NCp8-derived peptides with SL3 were examined by a native polyacrylamide gel electrophoresis (PAGE) assay. These results indicate that the function of the second zinc finger in NCp8 is different from that of NCp7. We discuss the structural elements required for the viral RNA binding process.

MATERIALS AND METHODS

Protein Expression and Purification. NCp8, which was uniformly enriched with ^{15}N and/or ^{13}C, was expressed in *Escherichia coli* cells and subsequently purified using the ubiquitin fusion protein system, as described elsewhere (*14*). NCp8 was overexpressed in *E. coli* strain BL21 (DE3) pLysE as a recombinant fusion protein with an N-terminal decahistidine tag and in *Saccharomyces cerevisiae* ubiquitin, which was under the control of the T7 promoter. The *E. coli* cells were grown in a ^{15}N- and/or ^{13}C-labeled CHL medium (Chlorella Inc., Tokyo, Japan) with 0.1 mM kanamycin. The cells reached an OD_{600} of 0.7 at 37 °C and were induced via addition of 0.4 mM isopropyl β-D-thiogalactopyranoside, 0.1 mM $ZnCl_2$, and various organic compounds, in addition to the ^{15}N- and/or ^{13}C-labeled amino acids (Chlorella Inc.). Growth was continued for 9 h at 42 °C, followed by centrifugation of the cells at 6800*g*, with resuspension of the cells in buffer A [50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM 2-mercaptoethanol, and 6 M guanidine hydrochloride]. Cell lysis was carried out on ice, with the lysate sonicated to reduce the viscosity, and then stirred for 12 h. After being stirred, the lysate was centrifuged at 8500*g*, and the supernatant was then loaded onto Ni^{2+}-NTA agarose (Qiagen GmbH, Hilden, Germany). Refolding of the ubiquitin fusion protein was performed using a linear gradient of buffer A to buffer B [50 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 1 mM 2-mercaptoethanol] and eluted from the column by use of a linear gradient of buffer B to buffer C [50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM 2-mercaptoethanol, and 600 mM imidazole]. Fractions containing the fusion protein were incubated with 10 mM dithiothreitol and 1 mM $ZnCl_2$ for 1 h at 57 °C. Yeast ubiquitin hydrolase 1 (YUH1), which specifically cleaves the ubiquitin fusion protein at the C-terminus of ubiquitin, was separately overexpressed in *E. coli* strain BL21 (DE3) with a C-terminal hexahistidine tag using the T7 promoter (*15*). After incubation, the fusion protein was cleaved by YUH1 for 1 h at 37 °C, and the mixtures were then loaded onto Ni^{2+}-NTA agarose. This procedure ensured that the resultant NCp8 was not trapped on the Ni^{2+}-NTA agarose. The flow was purified by being passed through a Sep-Pak Plus C_{18} cartridge (Waters, Milford, MA) and then pooled and lyophilized. Finally, the NCp8 was purified by reverse phase HPLC using a TOSOH 8020 system (Tosoh, Tokyo, Japan) with an Imtakt Cadenza CD-C18 reverse phase C_{18} column [250 mm × 4.6 mm (inside diameter), Imtakt, Kyoto, Japan] and then lyophilized. The samples for the NMR experiments contained 1.5 mM NCp8 and 3.2 mM $ZnCl_2$ in either a 90% H_2O/10% D_2O mixture or 99.96% D_2O at 15 °C and pH 5.8. For all NC samples, coordination of the zinc ions was confirmed by ^1H NMR spectra as described previously (*12*, *13*).

NMR Spectroscopy. NMR data were collected with a Bruker DMX-500 spectrometer and processed using the program package Azara (W. Boucher, unpublished work, available at http://www.bio.cam.ac.uk/azara/) followed by analysis with Ansig version 3.3 (*16*, *17*). Sequence-specific assignments of the polypeptide backbone and side chain resonances were obtained by standard triple-resonance techniques (*18*) using the following three-dimensional pulse sequences: HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, (H)CC(CO)NH, H(CCCO)NH, ^{15}N TOCSY-HSQC, and HCCH-TOCSY spectra. NOEs were assigned and collected on the basis of ^{15}N NOESY-HSQC with a mixing time of 200 ms and ^{13}C NOESY-HSQC spectra with a mixing time of 200 ms. Backbone torsion ϕ angles were obtained from HNHA spectra (*19*). Side chain torsion χ_1 angles were obtained from HN(CO)HB (*20*) and HNHB spectra (*21*). The ^{1}H assignment of the aromatic side chain was obtained from ^{1}H TOCSY with a mixing time of 80 ms and from ^{1}H NOESY with mixing times of 50, 100, and 200 ms using a nonlabeled sample with either a 90% H_2O/10% D_2O mixture or 99.96% D_2O at 15 °C and pH 5.8.

Structure Calculation. NOE cross-peaks were divided into strong, medium, weak, and very weak classes, which corresponded to upper distance limits of 2.5, 3.5, 5.0, and 6.0 Å, respectively. Pseudoatoms were used for nonstereospecifically assigned protons, and intraresidue correcting factors were added to the distance restraints (*22*). In addition, for distance restraints involving methyl protons, 0.5 Å was added to the upper limits. The dihedral angle ϕ restraints were converted according to the following rules (*23*, *24*). The $^3J_{NH\text{-}C^\alpha H}$ was less than 5.5 Hz and greater than 8.0 Hz; the ϕ angles were constrained in the ranges of $-65 \pm 25°$ and $-120 \pm 40°$, respectively, and the χ_1 angles were constrained in the range of $60 \pm 40°$ for the g^2g^3 conformation and $180 \pm 40°$ for the g^2t^3 conformation (*25*).

All calculations were carried out using XPLOR-NIH version 2.9.9 (*26*). Starting with 100 initial random structures, we carried out simulated annealing calculations that led to the selection of 11 final structures. Structural statistics for the 11 final structures are summarized in Table 1. Structures were analyzed and displayed using PROCHECK_NMR (*27*), MOLMOL (*28*), and PyMOL (W. L. DeLano, unpublished work, available at http://www.pymol.org/).

NMR Relaxation Measurements. ^{15}N NMR relaxation data were collected using the Bruker instrument. The ^{15}N T_1 data were collected using relaxation recovery delays of 20, 100, 300, 600, 1000, 1500, and 2500 ms. The ^{15}N T_2 data were collected with 17, 51, 85, 119, 153, 187, and 221 ms relaxation recovery delays. The relaxation rate constants were obtained by fitting the peak heights to a single-exponential function using the nonlinear least-squares method. The ^{15}N$-^1$H NOE data were obtained by recording interleaving pulse sequences with and without proton saturation. One spectrum was collected with a 3 s predelay and 3 s saturation, while another spectrum was collected with 6 s predelay and no saturation. The ^{15}N$-^1$H NOE was determined from the ratio of the peak heights for the experiments with and without ^{1}H saturation pulses.

Table 1: Structural Statistics for the 11 Converged Structures of NCp8[a]

rmsd from experimental distance constraints (Å)	
all (787)	0.028 ± 0.008
intraresidue (332)	0.028 ± 0.006
sequential (236)	0.030 ± 0.009
medium-range ($\|i-j\| < 5$) (102)	0.023 ± 0.006
long-range ($\|i-j\| \geq 5$) (117)	0.024 ± 0.009
rmsd from experimental dihedral constraints (deg) (24)	0.652 ± 0.211
energetic statistics (kcal/mol)[b]	
F_{NOE}	7.88 ± 1.64
F_{tor}	89.75 ± 11.72
F_{repel}	18.78 ± 4.36
$F_{L\text{-}J}$	−52.0 ± 17.4
rmsd from idealized geometry	
bonds (Å)	0.003 ± 0.0003
angles (deg)	0.650 ± 0.042
impropers (deg)	0.466 ± 0.081
Ramachandran analysis (%) (residues 9−22)	
most favored regions	54.5
additionally allowed regions	37.3
generously allowed regions	6.6
disallowed regions	1.6
average rmsd (Å) (residues 9−43)	
backbone (N, C^α, and C) atoms	1.90 ± 0.58
all heavy atoms	2.55 ± 0.60
average rmsd (Å) (residues 9−22)	
backbone (N, C^α, and C) atoms	0.42 ± 0.09
all heavy atoms	1.19 ± 0.20
average rmsd (Å) (residues 30−43)	
backbone (N, C^α, and C) atoms	0.64 ± 0.17
all heavy atoms	1.43 ± 0.27

[a] The values are means ± the standard deviation. [b] The values of the force constants used for these terms are the standard values, as depicted in XPLOR-NIH (*26*).

Specific Binding Activity between NCp8 and SL3. NCp8-f1 and NCp8-f2 (Figure 1a) were obtained commercially from Anygen Co., Ltd. (Kwangju, Korea). Viral RNA SL3 (Figure 1b) and competitive 19-mer adenine RNA were obtained commercially from Japan Bio Services Co., Ltd. (Saitama, Japan), and labeled with or without fluorescein isothiocyanate (FITC) at their 5′-ends. The purity of peptides and RNAs exceeded 95%. The native PAGE assay was performed using a previous procedure (*13*). Specific binding activities between NCp8 and FITC-labeled SL3 (FITC-SL3) were examined by varying the concentration of the competitor RNAs, which included 19-mer adenine RNA and nonlabeled SL3. These experiments were conducted using samples that contained 2 μM FITC-SL3, 8 μM NCp8, or 16 μM NCp8-derived peptide and the varying competitor RNA concentrations. The ratio of the concentration between the competitor RNA and FITC-SL3 ranged from 0 to 50. FITC-SL3 and competitor RNA were heated to 97 °C for 3 min and stored at 37 °C for 30 min. FITC-SL3 was preincubated with NC protein in reaction buffer [30 mM Tris-HCl (pH 7.5), 60 mM NaCl, 200 μM $MgCl_2$, 2 mM DTT, and 2.5 units RNase inhibitor (Wako Pure Chemical Industries Ltd., Osaka, Japan)] for 10 min at 37 °C. After addition of competitor RNA, the mixtures were incubated for 2 min at 37 °C prior to electrophoresis. Native 8% (v/v) PAGE gels were run at 210 V for 15 min in 0.5× Tris-borate-EDTA (TBE) buffer (pH 8.0) at 4 °C. The gels were imaged using a Typhoon 9400 instrument (GE Healthcare Bio-Sciences Corp.). The gel images were

excited at 488 nm with the resulting fluorescent emission detected using a 520 BP 40 emission filter (GE Healthcare Bio-Sciences Corp.). The analysis of imaged data was performed using ProFINDER 2D version 2004 (PerkinElmer Inc.).

NMR Analysis of the SL3 Binding Site of NCp8. The two-dimensional (2D) $^{15}N-^{1}H$ HSQC spectra of 2 mM ^{15}N-labeled NCp8 and 4.4 mM $ZnCl_2$ with 0, 0.25, 0.5, and 1 equiv of SL3 were recorded at 35 °C and pH 5.8. The combined chemical shift perturbation was calculated with the equation

$$\Delta\delta_{ppm} = \sqrt{(\Delta\delta_{HN})^2 + (\Delta\delta_N\alpha_N)^2} \quad (1)$$

with a scaling factor (α_N) of 0.14. The backbone assignments of NCp8 and SL3-bound states in the presence of a 0.5-fold molar excess of SL3 were determined with $^{1}H-^{15}N$ HSQC, ^{15}N TOCSY-HSQC, and ^{15}N NO ESY-HSQC experiments performed at 35 °C and pH 5.8.

RESULTS AND DISCUSSION

Structure Determination. Analysis of the $^{1}H-^{15}N$ and $^{1}H-^{15}N-^{13}C$ correlated NMR spectra enabled the unambiguous assignment of all backbones and most side chain atoms of NCp8, with the exception of Ala1 and Gln2 residues. The three-dimensional (3D) structures of NCp8 were calculated on the basis of 19 backbone torsion ϕ angle constraints, five side chain torsion χ_1 angle constraints, and 787 distance constraints. The side chains of Cys9, Cys12, and Ser18 assumed a g^2g^3 conformation, while the side chains of Cys33 and Lys42 assumed a g^2t^3 conformation. The 787 distance constraints included 332 intraresidue and 455 inter-residue NOE distance constraints, and four hydrogen bond constraints between Ile7 O and Gly16 NH and between Cys9 O and Lys14 NH [obtained using HNCO and HN(CA)CO spectra (*29*)], along with eight additional constraints between the zinc ion and sulfur atoms of Cys9, Cys12, Cys22, Cys30, Cys33, and Cys43, and between the zinc ion and the N^{ε} atoms of His17 and His38 (*13*, *30*, *31*). Statistics for the converged structures were evaluated in terms of the structural parameters (Table 1). Figure 2 shows the superposition of the resulting 11 lowest-energy structures of the 100 total calculations. The structure (Protein Data Bank entry 2EC7) shows the stereo pairs of the backbone heavy atoms (N, C^{α}, and C) for the converged structures of NCp8 that were obtained by the best-fitting superposition of the backbone atom (N, C^{α}, C, and O) coordinates for residues 9−43 (Figure 2a), residues 9−22 (Figure 2b), and residues 30−43 (Figure 2c). There was no convergence noted for the N-terminal segment (Ala1−Lys5) and the C-terminal segment (Gln47−Gly49). This may be the result of a lack of medium- and long-range NOE constraints due to the inherent flexibility of both segments. The two zinc fingers adopt almost the same secondary and tertiary structure with two turns involving the Cys9−Lys14 and Cys30−Lys35 domains on one hand and Glu15−His17 and Thr36−His38 fragments on the other hand. The C-terminal parts of each zinc finger, Ala19−Cys22 and Met40−Cys43, adopt helical conformations (Figure S1 of the Supporting Information). These structural features are the same as those reported for the NCp8-derived peptides (*13*, *31*) and the other NC proteins (*32*, *33*). While there was good convergence for the first zinc finger, there was less for the second zinc finger (Table 1). Also, the conformations of the first and second zinc fingers were not converged on a single conformation. Figure 2 shows the convergence of the N-terminal region of the linker (Arg23−Arg26) and the first zinc finger. However, the C-terminal region of the linker (Gln28 and Gly29) did not converge with the first zinc finger. For the linker in NCp8, several long-range NOEs were observed between the residues in the first zinc finger and the residues in the N-terminal region of the linker. However, there were only a few very weak long-range NOEs between the second zinc finger and the C-terminal region of the linker (Gln28 and Gly29). Moreover, the NOEs between the Asn11 $N^{\delta}H$ group in the first zinc finger and the Arg27 $C^{\alpha}H$ group in the linker were still present even at pH 8.8 (Figure 3). This NOE was also found in NCp8-f1 and proved to be the characteristic hydrogen bond between the Asn11 $N^{\delta}H$ group and Arg27 O in NCp8-f1 (*31*). The NOE indicated that the first zinc finger and the linker were in the proximity of each other (*12*). Thus, the conformation of the first zinc finger and the linker converged together, which is a requirement for this hydrogen bond (*12*, *31*).

For (12−53)NCp7, a few NOEs between the zinc fingers were observed. In particular, (12−53)NCp7 exhibited interfinger NOEs between the Phe16 side chain (corresponding to Trp10 in NCp8) and the side chain of Trp37 (corresponding to Trp31 in NCp8). However, for (12−53)NCp7 with D-Pro31 (corresponding to Pro25 in NCp8), all spatial cross-correlations between the zinc fingers disappeared (*34*) and the point mutations in the linker, for which a leucine was substituted for Pro31 in NCp7 and D-Pro for L-Pro31, led to a noninfectious virus and poor RNA dimerization, respectively (*32*, *34*). NCp8 also exhibited weak interfinger NOEs between the aromatic protons of Trp31 and the Hβ protons of Trp10 (Figure 3b). Moreover, because a strong NOE was observed, i.e., Ala24 $C^{\alpha}H$ to Pro25 $C^{\delta}H$, the Ala24−Pro25 bond was considered to be in the *trans* conformation. Therefore, Pro25 and the hydrogen bond between Asn11 and Arg27 induced a bend in the linker for the spatial proximity of the two zinc fingers (Figure 2).

Conformational Dynamics of NCp8. To examine the dynamic behavior of the two zinc fingers in more detail, T_1 and T_2 relaxation times and $^{15}N-^{1}H$ NOE data were independently measured at two different times. The results are plotted in Figure 4, with mean values (±standard deviation) listed in Table S2 of the Supporting Information. In Figure 4d, the $^{15}N-^{1}H$ NOE values are mapped in continuous-scale colors on the average structure of NCp8.

Qualitative analysis of the relaxation data revealed some significant points. As shown in Figures 4c and 5, the $^{15}N-^{1}H$ NOE values for the residues of the second zinc finger (average of 0.696) were ~10% lower than those of the first zinc finger (0.772). In general, the T_1 and T_2 values of the residues in the second zinc finger were lower than those of the first zinc finger. T_1/T_2 ratios of the residues in the first zinc finger were also lower than those

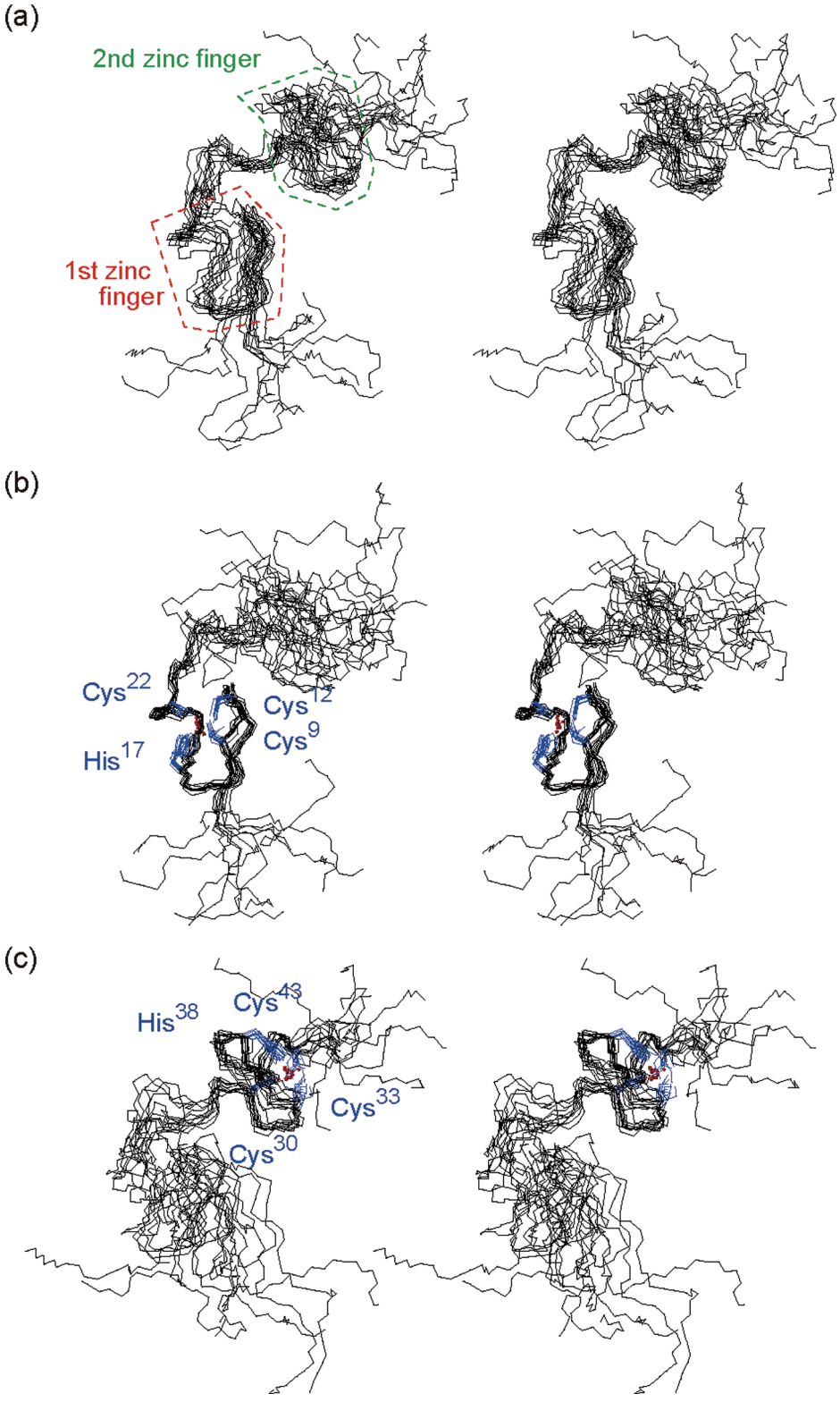

FIGURE 2: Stereo pairs of the backbone heavy atoms (N, C^{α}, and C) for the 11 converged structures of NCp8. Panels a−c are the best-fitting superpositions of the backbone atom (N, C^{α}, C, and O) coordinates for (a) residues 9−43, (b) residues 9−22 (first zinc finger), and (c) residues 30−43 (second zinc finger), respectively. The side chains of the zinc-bound residues are colored blue. Zinc ions are represented as red spheres.

in the second zinc finger (Figure 5c). It has been reported that large amplitude motions with time scales greater than a few hundred picoseconds affect the T_1/T_2 ratios (*35*). For NCp7, the ^{15}N relaxation parameters of the residues in the second zinc finger were also lower than those of the first zinc finger, which was considered to indicate that the two zinc fingers possess different intrinsic rotational correlation times (*2*, *36*). These results suggest that NCp8's second zinc finger has a different correlation time as compared to that seen for the first zinc finger. For the NCp8 linker, residues 23−27 had higher $^{15}N-^{1}H$ NOE values (Figure 4c) as compared to that seen for the same NCp7 region (*36*). For NCp8, the values for Arg23−Arg27 were almost the same as those observed in the first

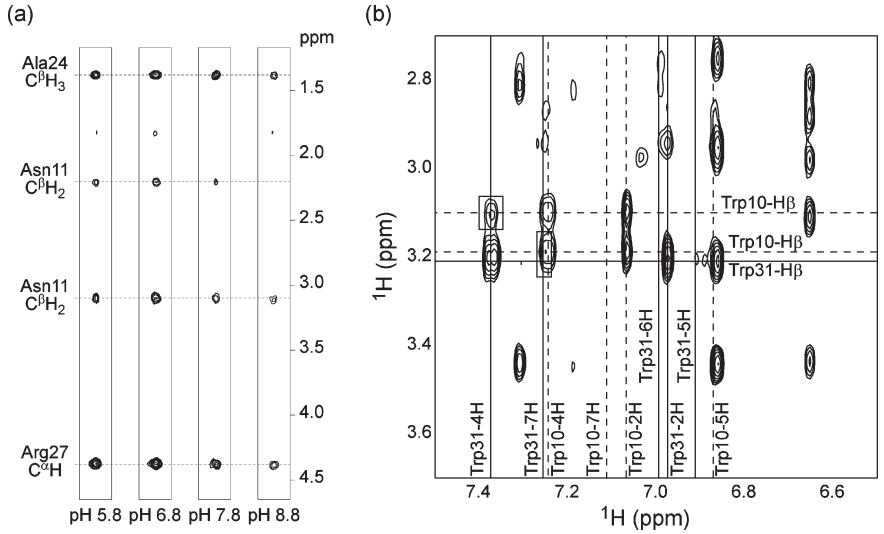

FIGURE 3: (a) NOE cross-peaks of the Asn11 $N^{\delta}H_2$ group ($N^{\delta}H_2$, 6.73 ppm) were recorded at pH 5.8, 6.8, 7.8, and 8.8. The NOEs of the Asn11 $N^{\delta}H_2$ group were still present even at pH 8.8. (b) Portion of the NOESY data obtained for nonlabeled NCp8 in a D_2O solution showing cross-peaks associated with the nonexchangeable aromatic protons. The Trp10 and Trp31 aromatic and $C^{\beta}H_2$ proton signals (solid line, Trp10 signals; dashed line, Trp31 signals) are labeled. Several interdomain cross-peaks are denoted in the boxes.

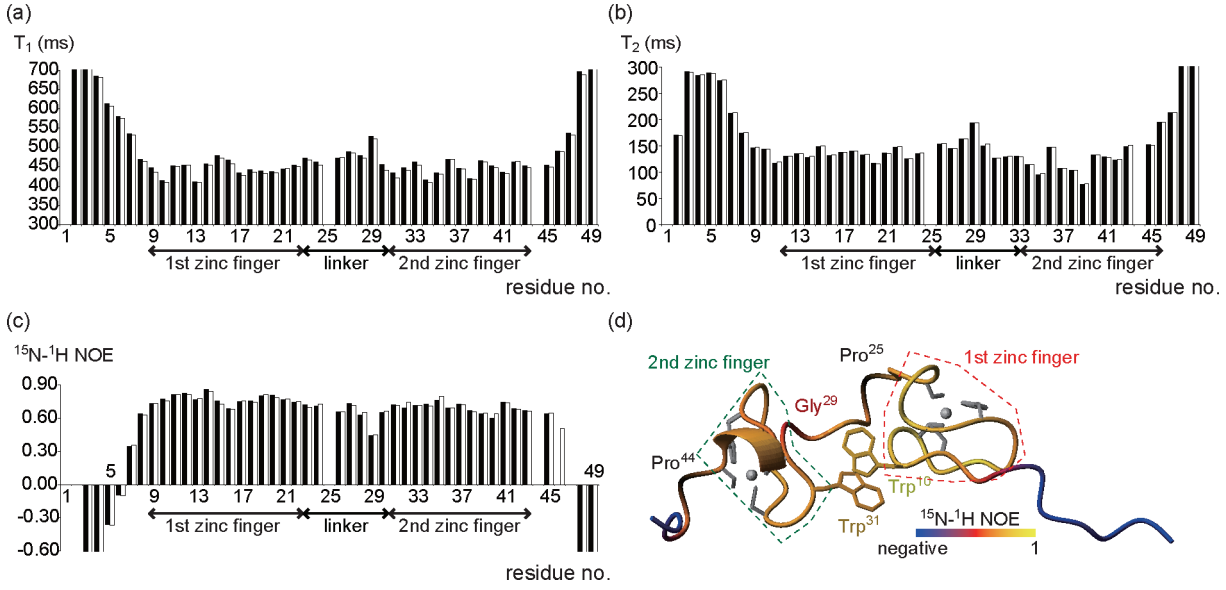

FIGURE 4: Backbone (a) ^{15}N T_1 and (b) T_2 relaxation data and (c) ^{15}N−^{1}H NOE values obtained for NCp8 at 15 °C and pH 5.8. Results for two separate experiments are plotted as shaded and open rectangles and illustrate the degree of experimental reproducibility. Residues 25 and 44 are prolines. Residues Cys9−Cys22, Arg23−Gly29, and Cys30−Cys43 comprise the first zinc finger, the linker, and the second zinc finger, respectively. (d) ^{15}N−^{1}H NOE values are mapped in continuous-scale colors on the average structure of NCp8. Prolines are colored black. Two zinc ions and the zinc-coordinated side chains are represented as gray spheres and colored gray, respectively.

zinc finger. The structural features (Figure 2) also suggest that the motion in the N-terminal region of the linker (Arg23−Arg26) conforms to that in the first zinc finger. On the other hand, for NCp7, the convergence conformations between the N-terminal region in the linker and the first zinc finger have been shown to occur only in the RNA complex states (*8*, *9*, *36*). Moreover, for NCp7, the T_1 and T_2 values of the linker were higher than those of the two zinc fingers, while the NOE values of the linker were lower than those of the two zinc fingers (*36*). Thus, the residues of the linker in NCp7 are more labile than those of the zinc fingers (*36*). On the other hand, for NCp8, only the Gln28−Gly29 region in the linker were observed to have higher T_1 and T_2 values and lower NOE values. These results indicate that the flexibility of the linker was only induced by the Gln28−Gly29 region, which suggests that the conformational limitations were caused by the hydrogen bond between Asn11 in the first zinc finger and Arg27 in the linker. Therefore, the conformational flexibility between the two zinc fingers in NCp8 was lower than that for NCp7.

Figures 4c and 5 show that the Val39−Met40 region in the second zinc finger exhibited NOE values lower than those of the other residues in the second zinc finger. In contrast, the values for the same region in NCp7 were almost the same as the average NOE values that were noted for the second zinc finger (*36*). This suggests that only the second zinc finger in NCp8 has a more flexible region.

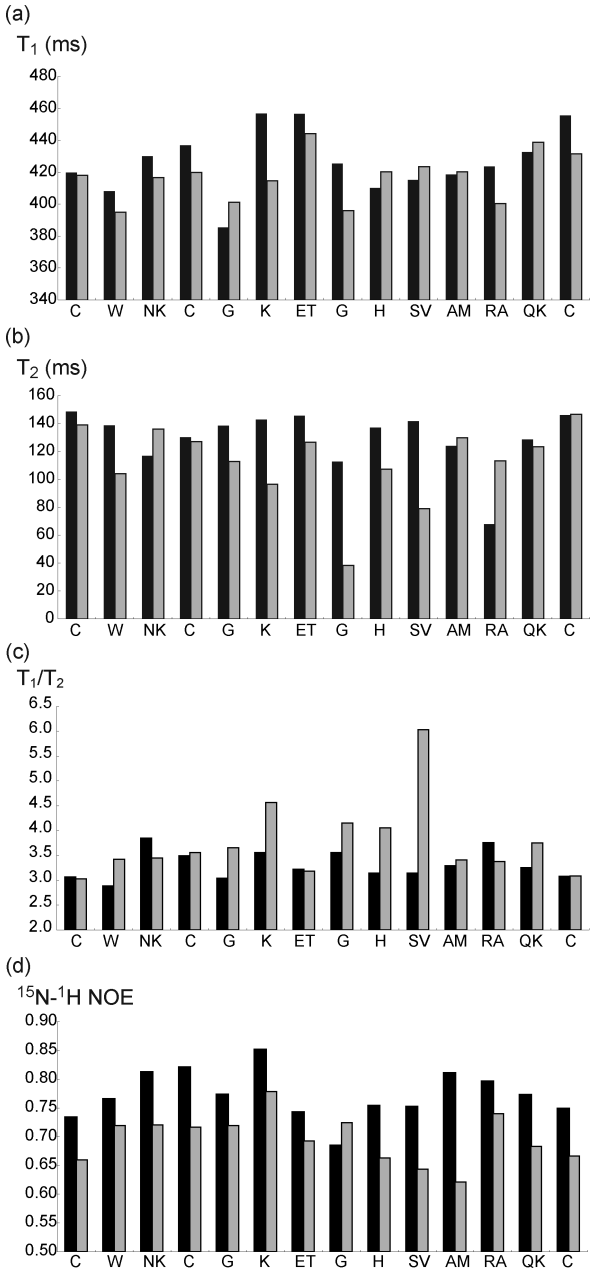

FIGURE 5: Comparison of the ^{15}N backbone relaxation NMR data (average values from two independent experiments) for the first (black rectangles) and second (gray rectangles) zinc fingers of NCp8. The arrays for the first and second zinc fingers begin at Cys9 and Cys30, respectively: (a) ^{15}N T_1, (b) ^{15}N T_2, (c) T_1/T_2 ratios, and (d) $^{15}N-^{1}H$ NOE values obtained for the NCp8 at 15 °C and pH 5.8.

RNA Recognition Mechanism of HIV-2 NCp8. To determine the structural element responsible for binding with SL3, we examined the NMR chemical shift perturbations as a function of the added SL3. Figure 6a shows the ^{15}N HSQC spectra for 2 mM ^{15}N-labeled NCp8 in the absence of SL3 (black) and after addition of 1 mM SL3 (red). The resonances for Trp10, Asn11, and Ala19, which are located in the first zinc finger, exhibited a particularly severe broadening. For the complex structures between NCp7 and the RNAs, the hydrophobic cleft in the first zinc finger, which was formed by Phe16 and Ala25 (corresponding to Trp10 and Ala19 in NCp8, respectively), bound specifically to an exposed guanine residue within the nucleic acid sequences (*2*, *8*, *9*). Thus, these facts suggest that it is the broadened residues that form the binding surface in the first zinc finger of NCp8 to allow binding with SL3. Figure 6b shows the chemical shift perturbations as a function of the amount of SL3 added. Arg20, Arg27, Trp31, Lys32, Cys33, Val39, Met40, and Ala41, which were observed as large chemical shift perturbations ($\Delta\delta > 0.4$ ppm), were mainly found in the second zinc finger. In the NCp8-derived peptide, NCp8-f2, it has been suggested that Trp31 and the Val39−Ala41 region form a hydrophobic cleft (*13*). In addition, for full-length NCp8, Trp31 and the Val39−Ala41 region formed the same hydrophobic cleft, as the conformations between the second zinc finger in NCp8 and that in NCp8-f2 were well-converged (Figure S3 of the Supporting Information). Moreover, Arg27 and Lys32 are located near the cleft (Figure 6c). These perturbation results suggest that the hydrophobic cleft and the basic residues around the cleft form the binding surface in the second zinc finger that allows binding with SL3.

It has been previously reported that ligands that bind with a K_d of $< 10^{-5}$ M will be in the fast-exchange limit for both chemical shifts, and those in the $10^{-6}-10^{-7}$ M range often fall into the intermediate (*37*). Therefore, it is possible that these broadened residues (within the first zinc finger) and the perturbed residues (within the second zinc finger) undergo intermediate and fast-exchange processes, respectively. Thus, the broadened residues may have a stronger interaction with SL3 as compared to the perturbed residues. This finding is supported by our previous results, which showed that the binding activity of NCp8-f1 was stronger than that of NCp8-f2 (*12*, *13*). However, the binding activity of NCp8 with SL3 was much higher than that seen for NCp8-f1. Therefore, both the perturbed and broadened residues are important for the binding to SL3, and this is considered to be the surface that forms the nucleotide interface (Figure 6c).

The NMR signals corresponding to the hydrophobic cleft of the second zinc finger were observed to have large chemical shift perturbations. However, for NCp7 during the RNA complex states, only a few residues in the second zinc finger were found to interact with the viral RNA (*8*, *9*). The hydrophobic cleft of the second zinc finger in NCp8 is formed by Trp31, Val39, Met40, and Ala41, whereas Trp37, Gln45, and Met46 (corresponding to Trp31, Val39, and Met40, respectively, in NCp8) form the cleft in NCp7 (*8*, *9*). These data suggest that the hydrophobicity of the hydrophobic cleft in the second zinc finger of NCp8 is stronger than that found for NCp7. Also, the flexibility of the Val39−Met40 region is only found in NCp8 (Figures 4c,d and 5). For NCp7, the minimal active domain and the 3_{10} helix in the N-terminus are the important elements that are required to be able to specifically bind to SL2 and SL3, and it is the 3_{10} helix that acts as the secondary active domain. When binding to each of the RNA sites, the relative orientations between the minimal active domain and the secondary active domain change in an adaptive manner (*2*, *8*, *9*). This conformational behavior, which acts through different subsets of molecular interactions, enables NCp7 to possess the multiple functions that are necessary for the

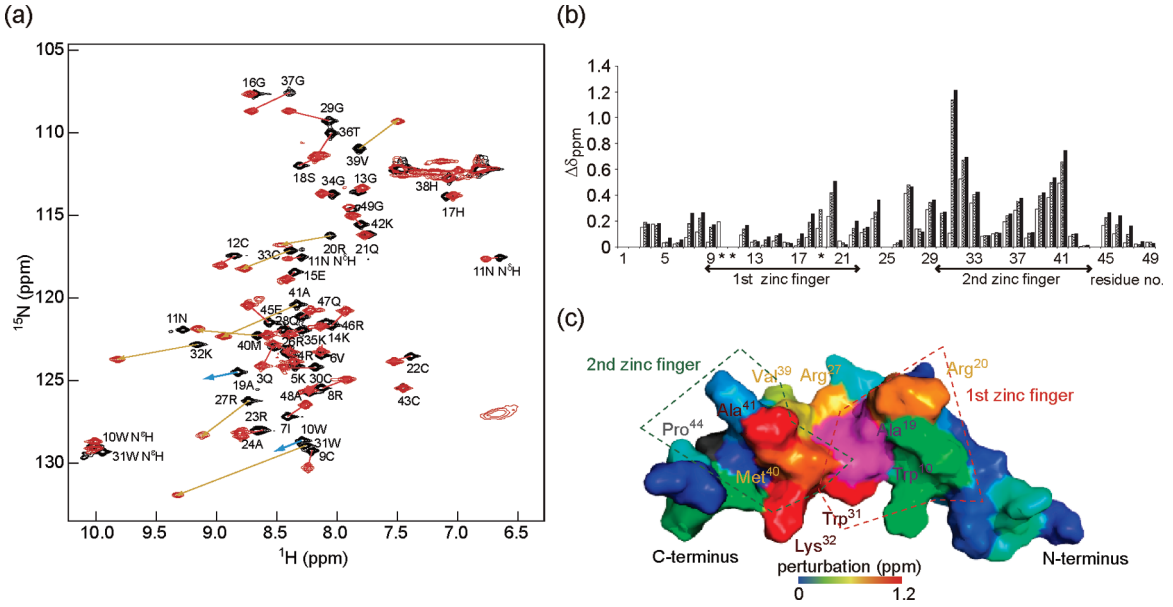

FIGURE 6: (a) NMR spectra of NCp8 free and bound to SL3. ^{15}N HSQC spectra of 2 mM NCp8 in the absence (black) or presence of a half-equivalent of SL3 (red) at 35 °C and pH 5.8 are shown. The orange and blue arrows indicate the large perturbation and the severe broadening cross-peaks, respectively. The sequence-specific assignment for the cross-peak at 6.6/127 ppm could not be determined by 3D ^{15}N NOESY and ^{15}N TOCSY spectra. (b) Chemical shift perturbations obtained for NCp8 as a function of added SL3 at 35 °C and pH 5.8. White rectangles indicate 2 mM NCp8 with 0.5 mM SL3, striped rectangles 2 mM NCp8 with 1 mM SL3, and black rectangles 2 mM NCp8 with 2 mM SL3. Residues 25 and 44 are prolines. Asterisks indicate the resonances of the residues that exhibited a particularly severe broadening and then disappeared. (c) Chemical shift perturbation values mapped in continuous-scale colors on the average structure of NCp8. The large perturbation residues, the broadened residues, and the prolines are colored with continuous-scale colors, purple, and gray, respectively.

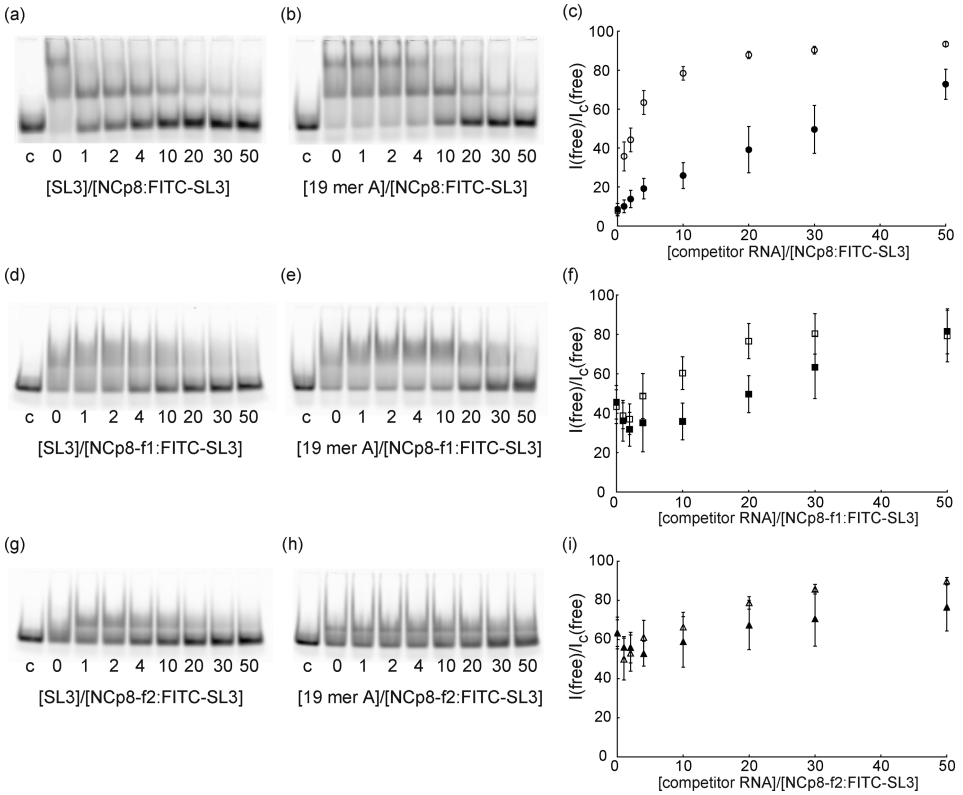

FIGURE 7: Specific binding activity of (a–c) NCp8, (d–f) NCp8-f1, and (g–i) NCp8-f2 to HIV-2 RNA. (a, d, and g) Nonlabeled SL3 and (b, e, and h) 19-mer adenines were used as competitors. Lane c in the gel images was loaded with only 2 μM FITC-RNA as the control. The concentration ratios of the competitor RNA to 2 μM FITC-SL3 ranged from 0 to 50. The graphs summarize the specific binding activities of (c) NCp8, (f) NCp8-f1, and (i) NCp8-f2. The data points are based on the amount of bound SL3 and were calculated as the ratio of the fluorescence intensity of the total FITC-SL3 in the control to that of the free FITC-SL3 after addition of the competitor: (○) NCp8–FITC-SL3 with nonlabeled SL3, (●) NCp8–FITC-SL3 with 19-mer adenines, (□) NCp8-f1–FITC-SL3 with nonlabeled SL3, (■) NCp8-f1–FITC-SL3 with 19-mer adenines, (△) NCp8-f2–FITC-SL3 with nonlabeled SL3, and (▲) NCp8-f2–FITC-SL3 with 19-mer adenines. The error bars represent the relevant standard deviations from five independent experiments.

recognition of the different RNA sites (*2*, *8*, *9*). On the other hand, for NCp8, the dynamics and hydrophobicity in the hydrophobic cleft of NCp8's second zinc finger may induce the region that forms the nucleotide-binding surface. Thus, these results suggest that the second zinc finger of NCp8 has a much more significant role in the RNA binding process than that seen for NCp7, which only serves in a restrictive accessory role (*38*). Additionally, the second zinc finger of NCp8 acts as a secondary active domain, similar to the role of the 3_{10} helix in NCp7. Therefore, the mechanisms that lead to the recognition of RNA between NCp8 and NCp7 are indeed different.

In our previous studies, we examined the structure−activity relationships of the NCp8-derived peptides. We demonstrated that the interaction of NCp8-f2 with SL3 was weaker than that observed for NCp8-f1 with SL3 and that the binding activity of NCp8 with SL3 was much higher than that seen for NCp8-f1 (*12*, *13*). The results of the specific binding assay (Figure 7) show that the specific binding activities of NCp8-f1 and NCp8-f2 with SL3 were weaker than those of NCp8 with SL3. These results provided an interesting insight into the fact that while each zinc finger itself is nonspecifically bound to SL3, both zinc fingers in NCp8 are able to strongly and specifically bind to SL3. Because the chemical shifts of the Asn11 $N^{\delta}H_2$ group in the NCp8−SL3 complex were the deviations from that of the random coil as seen for the NCp8 free state (Figure 6a), Asn11 may be able to form a hydrogen bond with Arg27 in the linker, similar to that seen for the NCp8 free state. Therefore, it is most likely that the limited flexibility of the linker, which is caused by the hydrogen bond between the first zinc finger (Asn11) and the linker (Arg27), makes it possible for the hydrophobic cleft of the second zinc finger to locate the spatial proximity in the first zinc finger. Furthermore, the flexibility of the second zinc finger makes it possible for the conformation to change in an adaptive manner. Hence, the cooperative effect of the primary active domain and the flexible secondary active domain of NCp8, which is different from that of NCp7, possesses the multiple functions required to be able to recognize several different RNA sites.

ACKNOWLEDGMENT

We are grateful to Ms. Kuniko Kobayashi of the Mitsubishi Kagaku Institute of Life Sciences for her technical assistance.

SUPPORTING INFORMATION AVAILABLE

Summary of the sequential NOE connectivities and $^3J_{NH\text{-}C^{\alpha}H}$ coupling constants observed in NCp8 (Figure S1), ^{15}N relaxation times (milliseconds) and $^{15}N-^{1}H$ NOE ratios observed for ^{15}N-labeled NCp8 at 15 °C and pH 5.8 (Table S2), stereo pairs of the backbone heavy atoms (N, C^{α}, and C) for the 11 converged structures of residues 23−49 of NCp8 and 12 converged structures of NCp8-f2 (Protein Data Bank entry 2E1X) (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

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. 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. U.S.A. 89*, 6472–6476.
6. 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.
7. 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.
8. 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 ψ-RNA recognition element. *Science 279*, 384–388.
9. 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.
10. 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.
11. 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.
12. Matsui, T., Kodera, Y., Endoh, H., Miyauchi, E., Komatsu, H., Sato, K., Tanaka, T., Kohno, T., and Maeda, T. (2007) RNA recognition mechanism of the minimal active domain of the human immunodeficiency virus type-2 nucleocapsid protein. *J. Biochem. 141*, 269–277.
13. Matsui, T., Kodera, Y., Miyauchi, E., Tanaka, H., Endoh, H., Komatsu, H., Tanaka, T., Kohno, T., and Maeda, T. (2007) Structural role of the secondary active domain of HIV-2 NCp8 in multi-functionality. *Biochem. Biophys. Res. Commun. 358*, 673–678.
14. Kohno, T., Kusunoki, H., Sato, K., and Wakamatsu, K. (1998) A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: An application to the production of mastoparan-X uniformly enriched with ^{15}N and $^{15}N/^{13}C$. *J. Biomol. NMR 12*, 109–121.
15. Sakamoto, T., Tanaka, T., Ito, Y., Rajesh, S., Iwamoto-Sugai, M., Kodera, Y., Tsuchida, N., Shibata, T., and Kohno, T. (1999) An NMR analysis of ubiquitin recognition by yeast ubiquitin hydrolase: Evidence for novel substrate recognition by a cysteine protease. *Biochemistry 38*, 11634–11642.
16. Kraulis, P. J. (1989) ANSIG: A program for the assignment of protein ^{1}H 2D NMR spectra by interactive computer graphics. *J. Magn. Reson. 84*, 627–633.
17. 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.
18. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) in *Protein NMR Spectroscopy, Principles and Practice*, Academic Press, Inc., San Diego.
19. Vuister, G. W., and Bax, A. (1993) Quantitative *J* correlation: A new approach for measuring homonuclear three-bond $J(H^{N}H^{\alpha})$ coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc. 115*, 7772–7777.

20. Grzesiek, S., Ikura, M., Marius Clore, G., Gronenborn, A. M., and Bax, A. (1992) A 3D triple-resonance NMR technique for qualitative measurement of carbonyl-Hβ J couplings in isotopically enriched proteins. *J. Magn. Reson. 96*, 215–221.
21. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) An alternative 3D NMR technique for correlating backbone ^{15}N with side chain Hβ resonances in larger proteins. *J. Magn. Reson. 95*, 636–641.
22. 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.
23. Pardi, A., Billeter, M., and Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C$^{\alpha}$ proton coupling constants, $^{3}J_{\mathrm{HN}\alpha}$, in a globular protein. Use of $^{3}J_{\mathrm{HN}\alpha}$ for identification of helical secondary structure. *J. Mol. Biol. 180*, 741–751.
24. Kline, A. D., Braun, W., and Wüthrich, K. (1988) Determination of the complete three-dimensional structure of the α-amylase inhibitor Tendamistat in aqueous solution by nuclear magnetic resonance and distance geometry. *J. Mol. Biol. 204*, 675–724.
25. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., and Wüthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol. 196*, 611–639.
26. 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.
27. 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.
28. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51–55.
29. Cordier, F., and Grzesiek, S. (1999) Direct observation of hydrogen bonds in proteins by interresidue $^{3H}J_{\mathrm{NC'}}$ scalar couplings. *J. Am. Chem. Soc. 121*, 1601–1602.
30. 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.
31. 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.
32. 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 ^{1}H NMR. *J. Mol. Biol. 235*, 287–301.
33. 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.
34. 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.
35. Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) Rotational diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation. *J. Am. Chem. Soc. 117*, 12562–12566.
36. 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.
37. Pellecchia, M., Sem, D. S., and Wüthrich, K. (2002) NMR in drug discovery. *Nat. Rev. Drug Discovery 1*, 211–219.
38. Heath, M. J., Derebail, S. S., Gorelick, R. J., and DeStefano, J. J. (2003) Differing roles of the N- and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-enhanced nucleic acid annealing. *J. Biol. Chem. 278*, 30755–30763.
39. 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.
40. 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.
41. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res. 31*, 3406–3415.