Structural Investigation on the Requirement of CCHH Zinc Finger Type in Nucleocapsid Protein of Human Immunodeficiency Virus 1[†]

Stéphanie Ramboarina, Nelly Moreller, Marie-Claude Fournié-Zaluski, and Bernard P. Roques*

Département de Pharmacochimie Moléculaire et Structurale, INSERM U266−CNRS UMR 8600, UFR des Sciences Pharmaceutiques et Biologiques, 4, avenue de l'Observatoire, 75270 Paris Cedex 06, France, and Département de Pharmacochimie Moléculaire et Structurale, INSERM U266−CNRS UMR 8600, UFR des Sciences Pharmaceutiques et Biologiques, 4, avenue de l'Observatoire, 75270 Paris Cedex 06, France

Received March 5, 1999; Revised Manuscript Received May 11, 1999

ABSTRACT: The nucleocapsid proteins (NCps) of lentiviruses play a key role during the retroviral replication cycle. NCps contain one or two highly conserved domains characterized by a $\mathbf{C}X_2\mathbf{C}X_4\mathbf{H}X_4\mathbf{C}$ sequence which binds zinc with a high affinity. The reasons of the high conservation of zinc fingers of CCHC type in lentiviruses were investigated by a structural study of mutants in which the zinc-coordinated ligands were exchanged. The HCHC form was unable to bind zinc tetrahedrally, whereas in His^{28}(13−30)NCp7 corresponding to the CCHH motif, the zinc was tightly complexed. The mutant peptide exists in two interconverting conformations E and D [ΔG_{DE} (293K) = 0.1 kcal/mol] arising from the zinc coordination of His^{28}, by either its Nϵ2 or its Nδ1, respectively. As compared to the native CCHC zinc finger, the $Cys^{28} \rightarrow$ His mutation induces structural changes in the finger due to a modification in the coordination state of His^{23} bound to zinc by Nϵ2 in the wild-type finger by Nδ1 in both conformers of the mutant. Introduction of these single mutations within the NCp7 proximal zinc finger in the HIV-1 genome was very recently shown to result in a loss of viral infection. This supports the hypothesis that structural changes of the zinc finger domain of NCp7 inhibit the recognition of one or several targets critically involved in the virus life cycle.

A large number of proteins contain zinc finger domains which have structural and/or functional roles in transcription factors. They have been shown to be involved in the direct recognition of DNA (*1*−*4*). In these metalloproteins, the size of the fingers and the sequential nature of the zinc ligands are different. Thus, transcription factors are characterized by successive and independent CCHH domains (*1*), but other proteins contain one or several identical zinc arrays of CCCC, CCHC, CCCH, or HHCC types (*2*−*5*). The $CX_2CX_4HX_4C$ zinc finger present in one or two copies in nucleocapsid proteins (NCps)[1] of lentiviruses, such as the human immunodeficiency virus HIV-1, seems to be the smallest module found yet (*6*−*9*). In these highly conserved finger domains, the zinc atom is bound tetrahedrally and with a great affinity (10^{13} M^{-1}) to the highly folded peptide backbone (*6*−*9*).

The HIV-1 NCp7 protein has two successive CCHC domains (Figure 1A). This protein is crucially involved in several key steps of the viral life cycle (RNA packaging, reverse transcription, and integration) through interactions with single-stranded nucleic acids (RNA and DNA) and viral proteins (*10*−*13*). The main characteristic of the NCp7 structure is a spatial proximity between both zinc fingers which have an almost identical folded conformation around the metal ion (*8*, *14*−*16*). The functional significance of these structural parameters has been clearly evidenced in vivo by the complete loss of infectivity of viruses in which critical residues of NCp7 have been mutated. These mutations were shown to either disrupt the finger proximity ($Pro^{31} \rightarrow$ Leu mutation) (*14*), to produce a loss of tetrahedral zinc complexation (Cys → Ala mutation) (*17*) or to induce a conformational rearrangement of the finger domain around the zinc ion ($His^{23} \rightarrow$ Cys mutation) (*18*). Similar point mutations carried out in the MoMulV NCp10 protein, which contains only one CCHC zinc finger, also resulted in a loss of virus infectivity (*19*, *20*).

It was therefore interesting to investigate the reasons for which the CCHC motif has been selected by the lentiviruses. With this aim, we have synthesized by solid-phase method two new peptide sequences identical to that of the N-terminal (13−30) zinc finger of NCp7, except that the CCHC arrangement was replaced by the HCHC or the CCHH ones. In the first case, no tetrahedral zinc complexation was observed. In contrast, the CCHH motif in His^{28} (13−30)-NCp7 binds zinc tightly and the structure of the peptide has been investigated both by ^1H and heteronuclear ^{15}N NMR. Two folded conformers were found characterized by large structural changes as compared to the native (13−30)NCp7, probably accounting for the high conservation of the CCHC motif in HIV-1 NCp7 and more generally in lentiviruses and explaining the recently reported loss of infectivity of HIV-1 mutants endowed with such single mutations (*35*).

[†] This work was supported by grants of the French research Minister and organizations against AIDS, ANRS.

* To whom correspondence should be addressed. Phone: (33) 1−43.25.50.45. Fax: (33) 1−43.26.69.18. E-mail: roques@pharmacie.univ-paris5.fr.

[1] Abbreviations: NCp, nucleocapsid protein; NOE, nuclear Overhauser enhancement; HSQC, heteronuclear single quantum coherence; HMQC, heteronuclear multiple quantum coherence.

10.1021/bi9905258 CCC: $18.00
Published on Web 07/09/1999

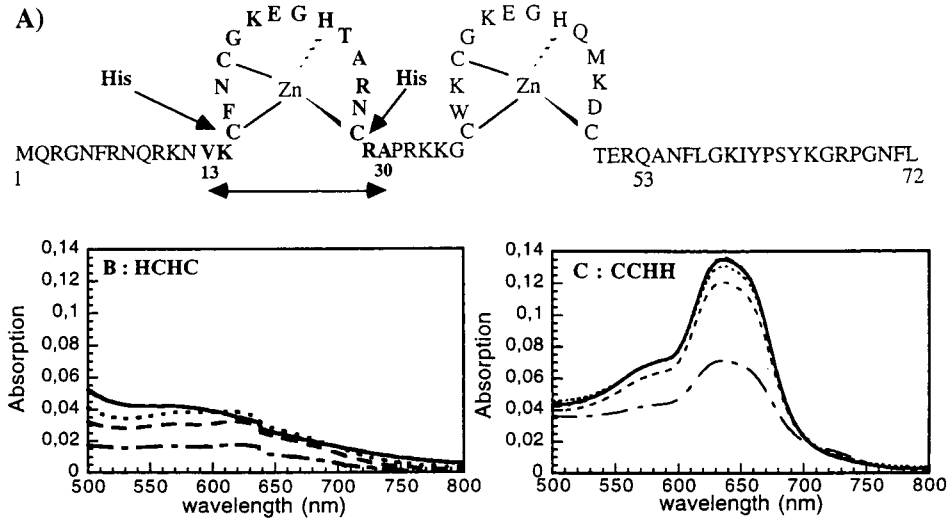

FIGURE 1: (A) Primary sequence of NCp7. The sites of Cys[15] → His and Cys[28] → His mutations in the isolated (13−30)NCp7 zinc finger (in bold) are indicated by arrows. (B) Visible spectrum of the His[15](13−30)NCp7 and (C) His[28](13−30)NCp7 peptides at 0.25 mM, pH 6.5, and 293K in the presence of Co^{2+} ions at increasing concentrations 0.25 (- − -), 0.50 (- - -), 0.75 (···), and 1 (—) molar equivalent. Addition of 2 equiv of Co^{2+} led to curves identical to those observed with one equivalent. The spectrum in panel C is characteristic of a complex formation in which the cobalt atom is tetracoordinated by two cysteine and two histidine residues (*28*).

MATERIALS AND METHODS

Peptide Synthesis. The His[15](13−30)NCp7 (VK**H**FN**C**GKE-G**H**TARN**C**RA) and His[28](13−30)NCp7 (VK**C**FN**C**GKEG-**H**TARN**H**RA) fragments were synthesized on an Automatic Applied Biosystems 433 A peptide synthesizer using the stepwise solid-phase synthesis method and Fmoc amino acids as already described (*21*). The His[28](13−30)NCp7 peptide was also synthesized with the two histidines 95% ^{15}N and 15% ^{13}C enriched. The labeled histidine (Senn Chemicals) has been protected for Fmoc chemistry as reported (*22*). Purification of peptides was carried out by HPLC with C18 VYDAC column using acetonitrile gradient. Electrospray mass spectroscopy measurements: calculated mass = 2035.80 g/mol, experimental mass = 2035.30 g/mol for labeled His[28]-(13−30)NCp7.

UV Measurements. UV−vis absorption spectra between 190 and 800 nm were obtained with a Perkin-Elmer spectrometer (Lambda 3B). The His[15](13−30)NCp7 and His[28](13−30)NCp7 peptides were dissolved in 50 mM Hepes buffer (pH 7), and the complexation of Co^{2+} ions was carried out by adding aliquots of $CoCl_2$ until 2 equiv. In competition experiments, incremented fractions of $ZnCl_2$ solution were added to the Co(II)-peptide (1:1) solution, and the variation of absorption was measured at 685 nm.

Ellman Test. The possible content in free cysteines for both peptides complexed with $CoCl_2$ was determined using the Ellman reagent [5,5′-dithiobis(2-nitrobenzoic acid) (DTNB)]. A 0.5 mM solution of DTNB in 1 M Tris buffer at pH 7.5 was added to a 0.015 mM solution of Co(II)-peptide and the intensity of the absorbance at 412 nm due to the formation of 2-nitro-5-thiobenzoate [$\epsilon_{(412nm)} = 13\,600$ M^{-1} cm^{-1}] in the presence of the free SH group was measured.

NMR Experiments. Two samples of the His[15](13−30)-NCp7 and His[28](13−30)NCp7 peptides in the presence of 1.5 equiv of zinc were prepared at a concentration of 2 mM in 90% H_2O/10% D_2O and D_2O at pH 6.8 ± 0.2. All the NMR spectra were recorded on a Bruker AMX600 spectrometer. 1D NMR experiments in H_2O/D_2O were performed from 278 to 323 K. Addition of 10% CD_3OH in the His[28]-(13−30)NCp7 samples allows a series of 1D spectra from 273 to 263 K to be performed. At these low temperatures, the exchange process is lower resulting in signals sharpening. 2D homonuclear HOHAHA was acquired in the clean mode (75 ms MLEV-17 spinlock duration), using a 3-9-19 pulse sequence (*23*) with gradients for water suppression at 293 K, 278 K, and 268 K. A total of 512 FIDs of 96 scans were collected using the TPPI method (*24*). NOESY spectra were recorded with the same acquisition and water suppression mode at 278 and 268 K with mixing times of 300, 200, and 100 ms. Data were zerofilled to 2K × 1K real points and apodized using 60° phase-shifted cosine-bell function in both dimensions before Fourier transformation. Additional baseline correction was performed using the software supplied by Bruker.

A series of NOESY experiments was carried out at 278 K using several mixing times (25−350 ms) in order to measure the exchange rate between the two forms at this temperature. In these experiments, the intensity of the exchange cross-peaks increased progressively with the enhancement of the mixing time, allowing discrimination between distance-dependent NOEs and exchange process.

A ROESY experiment (with a spin lock of 100 ms) reinforced these results as the exchange peaks were found negative and the ROEs positive (data not shown). Finally, the NOESY (at 100 ms) and ROESY experiments at 268 K gave exactly the same results, excluding spin-diffusion artifacts.

Two ^{15}N-histidine-enriched samples of His[28](13−30)NCp7 were prepared at 2 mM in 90%H_2O/10%D_2O and D_2O. 2D ^1H-^{15}N HSQC (*25*) and HMQC were acquired at pH 6.8 ± 0.2 and 278 K. The HMQC spectrum was recorded with a 22 ms dephasing delay to observe long-range H−N correlations at the histidine ring level (*26*). A total of 256 FIDs of 128 scans were collected using the TPPI method.

Structure Calculations. Structures of His[28](13−30)NCp7 complexed to zinc were generated with a simulated annealing

protocol under NMR distance restraints within the Discover NMRchitect package from Biosym Technologies. A zinc force field obtained from ab initio calculations and parametrized in the AMBER framework was used (*14* and references therein). Distance restraints were obtained from NOESY spectra recorded with mixing times of 100 ms at 268 K (see above). The H−H distances were classified into three categories: 2.0−2.5 Å, 2.0−3.5 Å, and 2.0−5.0 Å, corresponding to strong, medium, and weak NOEs, respectively. Appropriate pseudoatom corrections were added to the upper limits involving methylene, methyl, or aromatic groups given by the default values of the software (*27*).

RESULTS

Folding of the HCHC and CCHH Fingers of His[15](13−30)NCp7 and His[28](13−30)NCp7, Respectively, around the Zinc Atom. Absorption spectra between 500 and 800 nm following the addition of $CoCl_2$ to a solution of a His[15](13−30)NCp7 at pH 7, 20 °C (Figure 1B), did not present the typical transition bands observed in a cobalt tetracoordinated complex. Moreover, for a metal−peptide ratio of 1.5, the Ellman test showed the presence of one free SH group in solution indicating that the HCHC motif is unable to bind the zinc cation under tetracoordinated form.

In contrast, the addition of CoCl2 to a solution of His28-(13−30)NCp7 (Figure 1C) led to a shape of the absorption spectra in the range of 500−800 nm which represented the *d*−*d* transition bands at 580 and 640 nm, consistent with the formation of a stable tetrahedral complex similar to that observed in other CCHH zinc fingers (*28*). From this experiment, the affinity of the peptide for Co(II) ions was estimated to be 5×10^7 M^{-1} comparable to that of the wild-type zinc finger ($K_d \approx 10^7$ M^{-1}) (*29*). The immediate and complete displacement of the cobalt from the complex by 1 equiv of $ZnCl_2$ (not shown) reflected the greater affinity of His[28](13−30)NCp7 for zinc as previously observed for wild-type zinc motifs of NCps (*28*−*30*).

Structural analysis by ^{1}H NMR of the His[15](13−30)NCp7 peptide with 1.5 molar equivalent of zinc confirms that the HCHC arrangement did not form a stable tetracoordinated zinc array, since the 1D NMR spectrum exhibited broadened resonances very likely due to an exchange between several unstable zinc coordinated forms. Temperature variations from 5 to 50 °C did not result in an improvement of the 1D ^{1}H NMR spectra (data not shown).

In contrast, the 1D ^{1}H NMR spectrum of the His[28](13−30)NCp7 peptide in the presence of zinc displayed a wide chemical shift dispersion characteristic of a folded conformation. However, the 2D HOHAHA spectrum at 293 K shows the presence of numerous NH−NH cross-peaks for this 18 amino acid peptide suggesting an exchange between two folded forms (Figure 1, Supporting Information).

Conformational Exchange between Two Zinc Complexed CCHH Forms in His[28](13−30)NCp7. Sequence-specific assignments of both structures were carried out by standard methods (*28*) at 268 K. The signals from both species were easily assigned since two well-distinguished sets of cross-peaks along the entire peptide are observed in the 2D HOHAHA and NOESY spectra (Figure 2).

The two folded peptides present a similar pattern of NH−NH NOEs from Val[13] to His[28], with NH−NH connectivities

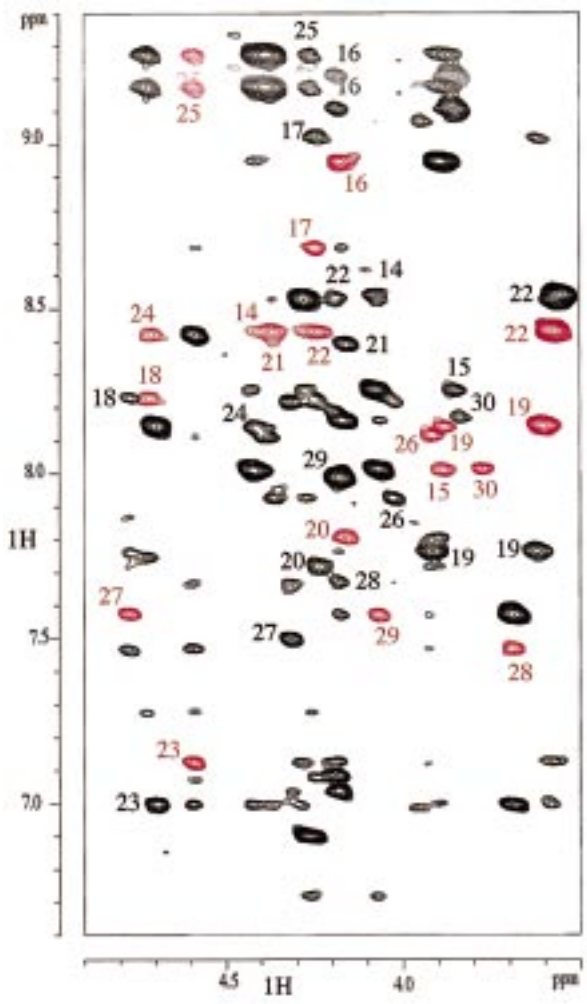

FIGURE 2: Part of the NOESY spectrum of (13−30)CCHH recorded at pH 6.8 and 268K with a mixing time of 100 ms showing connectivities between NH and αH resonances which are distinct for each conformation. The black and red assignments refer to the first and second forms, respectively.

of medium and strong intensity observed particulary from Phe[16] to Lys[20], and from Gly[22] to Thr[24] (Figure 3). In the first form, some dαN(*i*,*i*+2) of medium intensity are found between Lys[14]−Phe[16], Cys[15]−Asn[17], and Lys[20]−Gly[22], whereas weak dαN(*i*,*i*+2) NOEs are observed between Glu[21]−His[23] and Arg[26]−His[28]. Only one medium intensity dαN(*i*,*i*+3) and one dα*β*(*i*,*i*+3) connectivity is observed between Ala[25] and His[28] (Figure 3). A strong dαN(*i*,*i*+4) NOE between Glu[21] and Ala[25] and long-range d*β*N(*i*,*i*+4) and d*β*N(*i*,*i*+5) cross-peaks of weak intensity between Cys[15]−Gly[19] and Cys[15]−Lys[20] are observed in the N-terminal part of this E form (Figure 3). In the second form, dαN(*i*,*i*+2) NOEs are found, two of weak intensity between Cys[18]−Lys[20] and Arg[26]−His[28] and one medium dαN(*i*,*i*+2) involving Glu[21] and His[23]. A weak dαN(*i*,*i*+3) correlation between Phe[16] and Gly[19] was observed. Two weak d*β*N(*i*,*i*+2) connectivities and one medium are evidenced between Asp[17]−Gly[19], Glu[21]−His[23], and Cys[18]−Lys[20], respectively (Figure 3). The N-terminal fragment of this form is characterized by one d*β*N(*i*,*i*+4) and d*β*N(*i*,*i*+5) of medium and weak intensity between Cys[15] and Gly[19] and Cys[15] and Lys[20], respectively. In the C-terminal part, only one medium dα*β*(*i*,*i*+3) connectivity has been assigned between His[23] and Arg[26].

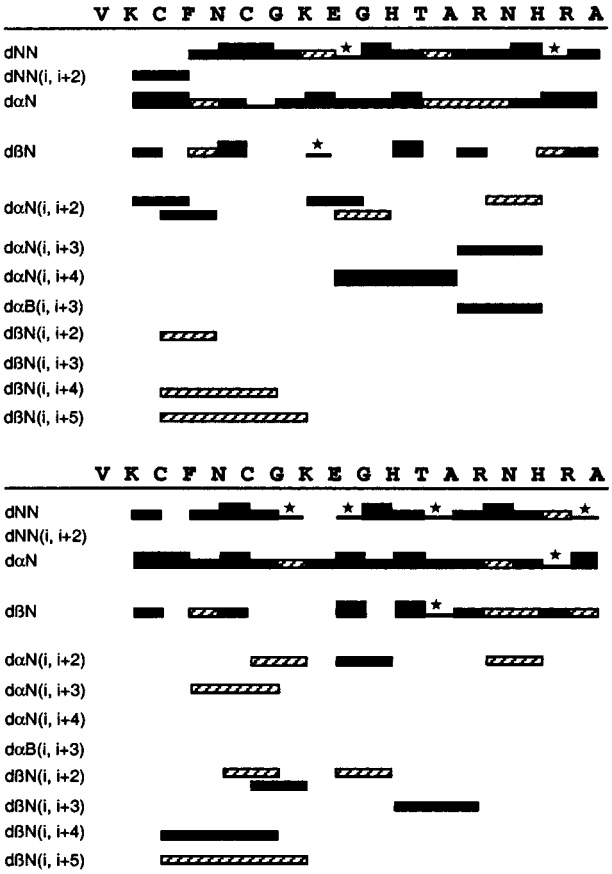

FIGURE 3: Summary of sequential (i,i+1), medium and long range (i+2 ≤ i,j ≤ i+5) NH−NH, αH−NH, αH-βH, and βH−NH NOE correlations identified in (13−30)CCHH from the 2D NOESY spectra at pH 6.8 and 268K in the first form (above) and in the second form (below). The thickness of the bars is proportional to the observed NOE cross-peaks intensity. The streak bars correspond to NOEs of weak intensities. Asterisks refer to NOES that could not be unambiguously assigned on account of signal overlap.

The ^{1}H-^{15}N HSQC spectra at 278 K display clearly two ^{1}H-^{15}N cross-peaks for both His[23] and His[28] in the His[28]-(13−30)NCp7 peptide, confirming the existence of two forms (data not shown). From the ^{1}H-^{15}N HMQC spectra of the Zn(II)-His[28](13−30)NCp7 complex, which recorded with a 22 ms dephasing delay to observe only long-range H−N correlations at the histidine ring level (*26*), two cross-peaks patterns for each histidine His[23] and His[28] were observed which could be assigned from their chemical shifts and coupling constants (Figure 4). It has been established that the chemical shifts of the protonated and unprotonated histidine nitrogen are around 170 and 200 ppm, respectively (*32*), and that the values of the imidazole coupling constants $^2J_{N\delta1H\epsilon1}$, $^2J_{N\epsilon2H\epsilon1}$, and $^2J_{N\epsilon2H\delta2}$ are of 10, 8, and 5.6 Hz, respectively, whereas the three bonds $^3J_{N\delta1H\delta2}$ is around 2 Hz (*32*). From this combined information, obtained from measurements of cross-peaks intensities, we can deduce that His[23] coordinates zinc atom by its Nδ1 atom in both forms [weak Nδ1/Hδ2 cross-peaks of the unprotonated nitrogen around 196 ppm and two strong Nϵ2/Hδ2 and Nϵ2/Hϵ1 correlations of protonated nitrogen at 172 ppm (Figure 4)]. Concerning His[28], a pattern with two strong Nϵ2/Hϵ1 and Nϵ2/Hδ2 cross-peaks at the unprotonated nitrogen chemical shift indicates that His[28] coordinates zinc by its Nϵ2 atom in one form and that, in the second form, His[28] coordinates zinc by its Nδ1 as a weaker Nδ1/Hδ2 correlation was

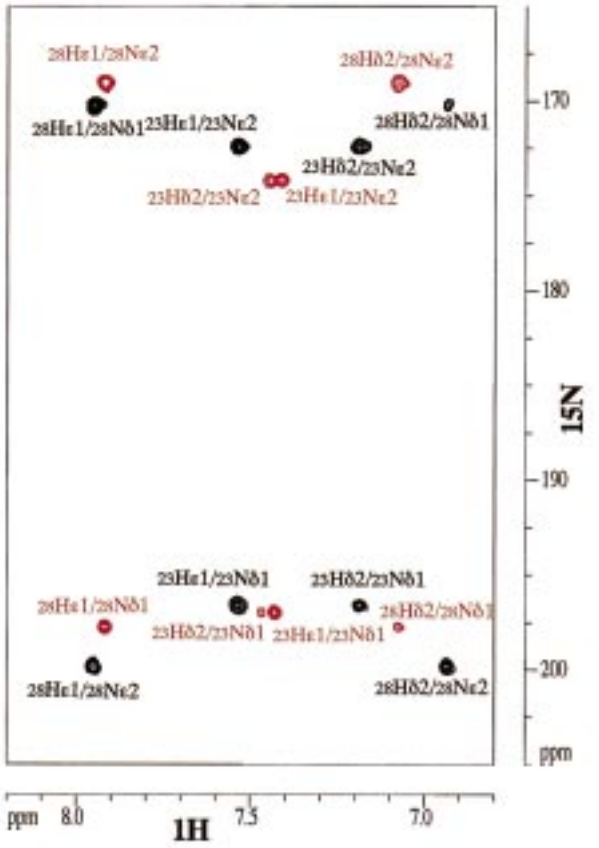

FIGURE 4: Long range ^{1}H-^{15}N correlation spectrum observed in the ^{1}H-^{15}N HMQC at pH 6.8 and 293K showing the protonation states of each histidine ring. The black and red cross-peak patterns refer to the zinc coordination of His[23] and His[28] in E and D forms, respectively.

observed around 198 ppm for this nitrogen atom (Figure 4). From these experiments, the two forms are designated E and D on the basis of zinc coordination by the Nϵ2 atom and Nδ1 atom of His[28], respectively (Figure 6A).

An increase of the temperature from 278 to 313 K induced a transition from a slow exchange toward an intermediate exchange between the two forms (Figure 5). The respective amounts of exchange forms were easily measured at the level of the two histidine protons such as Hδ2 of His[28] or Hϵ1 of His[23] for which a coalescence was observed at a temperature around 315 K in D_2O (Figure 5). From this experiment, a free enthalpy of activation $\Delta G^{\neq}$ of 13.4 kcal/mol has been calculated at 293 K.

From a series of 2D ^{1}H-^{1}H NOESY spectra recorded at 278 K with mixing times from 25 to 300 ms (Figure 2a, Supporting Information), a rate constant of 3.6 ± 1.2 s^{-1} for the overall exchange process has been determinated. The difference of enthalpy between the two forms ΔG_{DE} at 293 K estimated to 0.092 kcal/mol, ($\Delta G_{DE} = RT \ln(C_D^\circ/C_E^\circ)$ does not vary significantly from 278 to 313 K (Figure 2b, Supporting Information).

Differences between E and D forms are reinforced by significant changes in the HN and αH proton chemical shifts from the Cys[18] residue to the C-terminus of the peptide (Figure 3, Supporting Information). These NMR results suggest that the structural changes are not localized only around the site of point mutation, but affect all the Cys[18]−His[28] sequence.

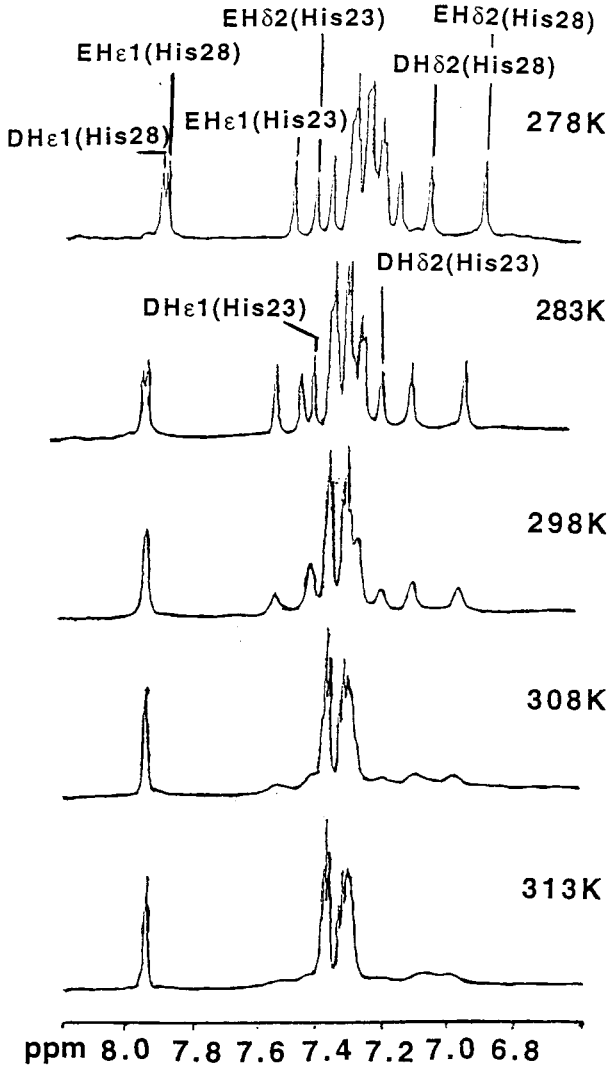

FIGURE 5: Aromatic proton region of the 1D NMR spectra of His[28]-(13−30)NCp7 in the presence of Zn^{2+} in D_2O, pH 6.8, as a function of temperature illustrating the exchange between two conformations. ΔG_{DE} and $\Delta G^{\neq}$ were calculated using different pairs of histidine proton signals.

Besides, the features of the NOEs provide strong evidence that both E and D species fold in conformations different from that of the native peptide through all the (13−30) peptide sequence (Figure 4, Supporting Information).

E and D His[28](13−30)NCp7 Structure Determination. D Conformation of the CCHH Zinc Finger. Simulated annealing calculations were performed with 145 relevant distance restraints distributed into 65 intraresidual and 80 interresidual upper limits. The distribution of the interresidual constraints was as follows: 45 were sequential ($i,i+1$), 18 were medium range ($2 \leq |i-j| \leq 4$), and 17 were long-range interactions ($i,j \geq i+5$). A total of 50 structures characterized by the lowest restraint energy (<10 kcal mol^{-1}) and no distance violation greater than 0.3 Å were selected. They were generated (see Materials and Methods), and the 16 best structures were further minimized with restraints (Table 1, Supporting Information) and in a final step without restraints. We currently used a zinc force field derived from ab initio calculations and parametrized in AMBER framework in order to account for metal protein interactions (*14*). The root-mean-square deviation (rmsd) on the backbone from Cys[15] to His[28] residues of the 16 best structures is 0.25 ± 0.06 Å (Figure 5a, Supporting Information). The final total energy is 65 ± 7 kcal/mol. All the structures are characterized by a γ-turn involving Cys[15], Phe[16], and Asn[17] followed by the [19]Gly−Gly[22] sequence which folds in a type III β-turn characterized by the dihedral angle values of Lys[20] ($\Phi \approx -95°$, $\psi \approx 45°$) and Glu[21] ($\Phi \approx -75°$, $\psi \approx 62°$). Finally the sequence Thr[24], Ala[25], and Arg[26] is involved in a γ-turn with the dihedral angle values ($\Phi \approx -59°$, $\psi \approx 98°$) for Ala[25].

In contrast, the wild-type (13−30)NCp7 is formed of three successive turns with the Cys15−Cys18, Gly22−Ala25 sequences generating a type VII β-turn, a type X β-turn, and a type I β-turn, respectively.

Different combinations of zinc coordination by the nitrogen atoms of the two histidines have been tried (without taking into account the unambiguous assignment of zinc coordination of the two histidines deduced from ^{15}N NMR data). Obviously, the zinc finger conformation which did not correspond to the experimentally observed structures led to important NOEs violations and structure distortions.

E Conformation of the CCHH Zinc Finger. The 15 best structures were generated by simulated annealing calculations with 144 relevant NOEs including 68 intraresidual and 75 interresidual NOEs distributed into 39 sequential ($i,i+1$), 18 medium range ($i,i+2 \leq j \leq i+4$), and 19 long-range distances ($i,j = i+5$) (Table 1, Supporting Information). The rmsd on the (15−28) backbone atoms of the 15 best structures is 0.48 ± 0.15 Å with a final total energy of 58 ± 8 kcal/mol (Figure 5b, Supporting Information). The main characteristics of this structure are a γ-turn observed between Cys[15] and Asn[17] followed by a second γ-turn including Gly[19], Lys[20], and Glu[21]. The end of the peptide is constituted by a type V′ β-turn from Thr[24] to Asn[27] residues, with dihedral angle values ($\Phi \approx 59°$, $\psi \approx -80°$) for Ala[25] and ($\Phi \approx -85°$, $\psi \approx 61°$) for Arg[26].

Comparison between the E and D forms of (13−30)CCHH zinc finger shows a 3.05 Å pairwise rmsd on the (15−28) backbone atoms. In addition to the modification of the (23−28) sequence which presents a pairwise backbone rmsd of 1.45 Å, an important difference in the two folded structures is observed at the level of Lys[20] and Glu[21] increasing the rmsd value to 1.93 Å when the (18−23) backbone atoms are superimposed (Figure 6B). Although the difference between the two CCHH zinc finger concerns a change in nitrogen coordination of zinc occurring at the level of His[28] (Figure 6A), this change also induces important conformational modifications from Lys[20] to His[28] residues.

Comparison of Native and Mutant Zinc Finger Structures. The wild-type (13−30)CCHC zinc finger is composed of three turns: the Cys[15]−Cys[18], Gly[22]−Ala[25], and Ala[25]−Cys[28]. Fragments form a type VII β-turn, a type X β-turn, and a type I β-turn, respectively. In the case of the CCHH mutant, the turns found in both conformers are slightly different as illustrated in Figure 4 (Supporting Information) when a NOE's diagonal plot shows the differences between the E and D mutated forms and the wild-type zinc finger. The contribution of the Cys[28] → His mutation on the entire zinc finger is estimated in a first step from the difference evidenced in the NOEs diagonal plot between each mutated form and the (13−30)wild-type zinc finger (Figure 4, Supporting Information). Moreover, the superposition of the (15−28) backbone atoms of the CCHH averaged D and E

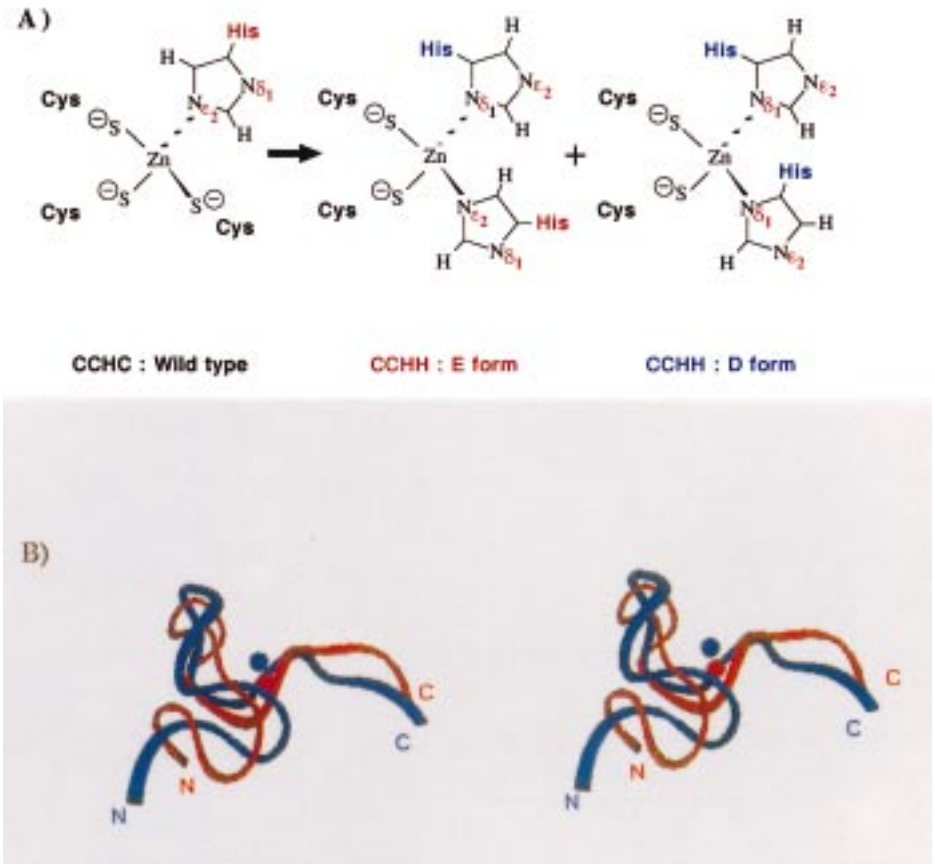

FIGURE 6: (A) Substitution of the CCHC wild-type motif onto a CCHH one, leads to two interconverting E and D forms characterized by Nϵ2 and Nδ1 of His[28] as zinc ligands, respectively. In the CCHH motif, the first His[23] ligand coordinates zinc by Nδ1 contrary to Nϵ2 in the CCHC wild-type motif (*7*, *9*, *14*). (B) Stereoview of the superimposition of the backbone of the average 3D structure of E (in red) and D (in blue) forms, with a pairwise rmsd of 3.05 Å on the (15−28) backbone atoms.

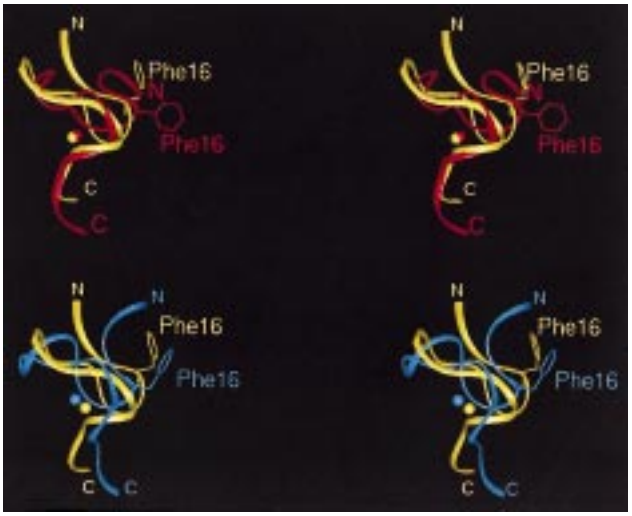

FIGURE 7: Stereoview of the comparaison between (top) the E(13−30)CCHH (red) and wild-type (13−30)CCHC zinc finger (yellow) and (bottom) the D(13−30)CCHH (blue) and wild-type (13−30)-CCHC zinc finger (yellow).

conformations onto the CCHC native one showed rmsd values of 2.34 and 3.15 Å, respectively (Figure 7). The change of zinc coordination of His[23] from Nϵ2 into the native CCHC motif (*7*, *14*) to Nδ1 into the mutant CCHH motif (Figure 6A), strongly modifies the N-terminal (*18*−*22*) segment in addition to the (*23*−*28*) domain (Figure 7) (Table 2, Supporting Information).

DISCUSSION

The aim of this study was to investigate the structural and functional specificity of the retroviral CCHC zinc finger which is highly conserved in lentiviruses. This CCHC domain represents the smallest sequence able to bind zinc with a high affinity, and its solution structure determined by ^{1}H NMR spectroscopy is characterized by four successive well-defined turns (*7*, *9*, *14*). The two repetitive fingers of NCp7 have a similar structure, and their relative orientation was demonstrated to be crucial for the virus infectivity. The exchange of zinc ligand, which consists of the replacement of His[23] by Cys in NCp7 leading to a CCCC domain was found to preserve the affinity for the zinc but was shown by ^{1}H NMR to induce a different peptide folding around the metal atom (*32*), leading to a loss of proximity between both fingers (*18*). This single mutation introduced in the retroviral genome led to a noninfectious virus (*18*, *33*).

It was therefore interesting to investigate if other exchanges in zinc ligands could induce modifications in zinc finger structure and subsequent loss of biological activity. The first modification performed in this study was the replacement of Cys[15] in the N-terminal finger of NCp7 by His leading to a HCHC motif. This motif has never been described as a potential finger domain in a protein, and we show here that the sequence $HX_2CX_4HX_4C$ is unable to provide a stable tetracoordinated zinc or cobalt complex as shown by UV spectroscopy and by ^{1}H NMR study of the His[15](13−30)-NCp7 peptide.

The second modification concerns the last Cys of the finger Cys^{28}, which was replaced by His leading to a CCHH structure which is a classical zinc binding motif found in a great number of nucleic acid binding proteins (*1*, *2*). These motifs, containing in most cases a $CX_{2-4}CX_{12}HX_{3-5}H$ type sequence, have a highly conserved 3D structure consisting of a β-hairpin which encompasses the two cysteine ligands, followed by an α-helix including the two His (*1*−*4*, *34*).

The structure of the CCHH motif in the finger domain of NCp7 is very different to that of the CCHC motif found in the native protein. First, the CCHH zinc array exists in two interconverting folded conformations (E and D forms) that differ by the coordination of the last His side chain to zinc. Indeed, whereas in both forms, His^{23} coordinates zinc by its Nδ1 atom, the metal is coordinated by the Nϵ2 atom and by the Nδ1 atom of His^{28} in the E and D forms, respectively (Figure 6A). An interesting observation is that both complexes are energetically almost equivalent with a ΔG_{DE} around 0.1 kcal/mol at 293 K.

It can be noticed that mutation of the last Cys ligand into a histidine does not induce helical folding at the level of the $H-X_4-H$ sequence as observed in the classical CCHH zinc finger. It has been generally observed that the helical folding induces a zinc coordination to the imidazole Nϵ2 of the histidine ligands (*34*). Our heteronuclear NMR results show that the zinc binding of the first histidine ligand His^{23} did not occur by coordination of Nϵ2 as found in the native zinc finger but by Nδ1 in the mutated zinc array (Figure 6A). This is probably due to restraints for zinc binding in the **C-X_4-H**-X_4-H sequence.

The presence of two interconverting structures was previously observed in the N-terminal $^{12}HX_3HX_{23}CX_2C^{43}$ zinc binding domain of HIV-1 integrase (*5*). NMR studies of this HHCC domain have shown that His^{16} is coordinated to the zinc by its Nδ1 atom in both forms, while the metal is coordinated by the Nϵ2 and the Nδ1 atoms of His^{12} in the E and D forms, respectively. In the E form, the two histidines are included in a well-defined helix, while in the D form, the helical arrangement is broken by a turn at the His^{12} level (*5*).

In addition to the discussed differences in zinc coordination, a comparison of the structure of the wild-type (13−30) zinc finger with that of the His^{28}(13−30) mutant shows important structural modifications. Thus, the native zinc finger is characterized by type I and II NH−S tight turns with numerous internal hydrogen bonds between the sulfur atom of the cysteine amino acids and the amide proton of the (i+1) and (i+2) residues (*7*, *9*). The replacement of Cys^{28} by His in the mutant results in a large rearrangement of the peptide backbone from the Cys^{18} residue arising in part from the zinc coordination of His^{23} by its Nδ1 atom. Only one NH−S hydrogen bond is observed in the His^{28}(13−30)NCp7, between the sulfur atom of Cys^{15} and the amide proton of His^{23} in the E form and between the sulfur atom of Cys^{18} and the amide proton of His^{23} in the D form. Therefore, the $Cys^{28} \rightarrow$ His point mutation provides large structural changes in all the peptide backbone in the two forms of the mutant.

Thus, it can be noticed that the side chains of amino acids (Val^{13}, Phe^{16}, Thr^{24}, Ala^{25}, Arg^{26}) critically involved in nucleic acid binding (*37*, *38*) or RT recognition (*13*) are not found in the same spatial orientation in the WT zinc finger and in the CCHH mutant especially in the case of the E conformer. These structural changes are very likely responsible for the loss of infectivity induced by the corresponding mutations within the zinc finger sequence in the genomes of HIV-1 (*35*) and MoMuLV viruses (*19*, *20*). These mutants are essentially defective in reverse transcription, leading to a loss of full-length viral DNA synthesis (*19*, *20*, *35*).

In account of these biological data and according to our NMR results on the isolated His^{28}(13−30)NCp7 peptide, we can assume that the substitution of the N-terminal CCHC zinc finger to a CCHH one in the entire NCp7 would strongly affect the structural integrity of the protein. This could result in a drastic alteration in the recognition by the modified NCp7 of its biological partners such as nucleic acids as well as viral or cellular proteins (Vpr, RT, integrase) (*11*−*13*, *39*) (work in progress in our laboratory).

In conclusion, this structural study of the isolated proximal CCHH zinc finger mutant of NCp7, associated to that performed on the CCCC mutant (*18*, *32*), allows the demonstration of the critical requirement of a CCHC type structure for the zinc fingers of NCps in lentiviruses.

ACKNOWLEDGMENT

We thank P. Petitjean for peptide synthesis. We acknowledge C. Dupuis for expert assistance in drafting this manuscript.

SUPPORTING INFORMATION AVAILABLE

Tables of NMR restraints and statistical parameters and figures of fingerprint region, build-up curves, chemical shift differences, distance lattices, and stereoviews. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Klug, A., and Rhodes, D. (1987) *Trends Biochem. Sci. 12*, 464−469.
2. Klug, A., and Schwabe, W. R. J. (1995) *FASEB 9*, 597−604.
3. Berg, J. M., and Shi, Y. (1996) *Science 271*, 1081−1085.
4. Mackay, J. P., and Crossley, M. (1998) *Trends Biochem. Sci. 23*, 1−4.
5. Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1997) *Nat. Struct. Biol. 4*, 567−577.
6. Cornille, F., Mely, Y., Ficheux, D., Savignol, I., Gerard, D., Darlix, J. L., Fournié-Zaluski, M. C., and Roques, B. P. (1990)- *Int. J. Pept. Res. 36*, 551−558.
7. Morellet, N., Jullian, N., de Rocquigny, H., Maigret, B., Darlix, J. L., and Roques, B. P. (1992) *EMBO J. 11*, 3059−3065.
8. Mély, Y., Cornille, F., Fournié-Zaluski, M. C., Roques, B. P., and Gerard, D. (1991) *Biopolymers 31*, 899−906.
9. Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., South, T. L., Blake, P. R., Sagi, I., Perez-Alavarado, G., Sowder, R. C. I., Hare, D. R., and Arthur, L. O. (1992) *Protein Sci. 1*, 563−574.
10. Darlix, J. L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) *J. Mol. Biol. 254*, 523−537.
11. Roques, B. P., Morellet, N., de Rocquigny, H., Déméné, H., Schueler, W., and Jullian, N. (1997) *Biochimie 79*, 673−680.
12. De Rocquigny, H., Petitjean, P., Tanchou, V., Décimo, D., Drouot, L., Delaunay, T., Darlix, J. L., and Roques, B. P. (1997) *J. Biol. Chem. 272*, 30753−30759.
13. Druillennec, S., Caneparo, A., and Roques, B. P. (1999) *J. Biol. Chem. 274*, 11283−11288.
14. 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.

15. Mély, Y., de Rocquigny, H., Morellet, N., Roques, B. P., and Gérard, D. (1996) *Biochemistry 35*, 5175−5182.
16. Lee, B. M., De Guzman, R. N., Turner, B. G., Tjandra, N., and Summers, M. F. (1998) *J. Mol. Biol. 279*, 633−649.
17. Aldovini, A., and Young, R. A. (1990) *J. Virol. 64*, 1920−1926.
18. Déméné, H., Dong, C. Z., Ottmann, M., Rouyez, M. C., Jullian, N., Morellet, N., Mély, Y., Darlix, J. L., Fournié-Zaluski, M. C., Saragosti, S., and Roques, B. P. (1994) *Biochemistry 33*, 11707−11716.
19. Gorelick, R. J., Chabot, D. J., Ott, D. E., Gagliardi, T. D., Rein, A., Henderson, L. E., and Arthur, L. O. (1996) *J. Virol. 70*, 2593−2597.
20. Yu Q., and Darlix, J. L. (1996) *J. Virol. 70*, 5791−5798.
21. De Rocquigny, H., Gabus, C., Vincent, A., Fournié-Zaluski, M. C., Roques, B. P., and Darlix, J. L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6472−6476.
22. Barlos, K., Papaioannou, D., and Theodoropoulos, D. (1982) *J. Org. Chem. 47*, 1324−1326.
23. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biol. NMR 2*, 661−665.
24. Marion, D., and Wütrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
25. Bodenhausen, G., and Ruben. (1980) *Chem. Phys. Lett. 69*, 185−188.
26. Pelton, J. G., Torchia, D. A., Meadow, N. D., Roseman, S. (1993) *Protein Sci. 2*, 543−558.
27. Wüthrich, K. (1986) in *NMR of proteins and nucleic acids* (Wiley, A., Ed.) pp 177−179, Interscience publication.
28. Krizek, B A, Amann, T A, Kilfoil, J. V., Merkle, D. L., and Berg, J. M. (1991) *J. Am. Chem. Soc. 113*, 4518−4523.
29. 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.
30. Green, L. M., and Berg, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4047−4051.
31. Bloomberg, F., Murer, W., and Rüterjans, H. (1977) *J. Am. Chem. Soc. 99*, 8149−8159.
32. Jullian, N., Déméné, H., Morellet, N., Maigret, B., and Roques, B. P. (1993) *Febs Lett. 331*, 43−48.
33. Tanchou, V., Decimo, D., Péchoux, C., Lener, D., Rogemond, V., Berthoux, L., Ottman, M., and Darlix, J. L. (1998) *J. Virol. 72*, 4442−4447.
34. Lee, S. M., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) *Science 245*, 635−637.
35. Gorelick, R. J., Gagliardi, T. D., Bosche, W. J., Wiltrout, T. A., Coren, L. V., Chabot, D. J., Lifson, J. D., Henderson, L. E., and Arthur, L. O. (1999) *Virology 256*, 92−104.
36. Rémy, E., de Rocquigny, H., Petitjean, P., Muriaux, D., Theilleux, V., Paoletti, J., and Roques, B. P. (1998) *J. Biol. Chem. 27*, 4819−4822.
37. Morellet, N., Déméné, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournié-Zaluski, M.-C., and Roques, B. P. (1998) *J. Mol. Biol. 283*, 419−434.
38. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) *Science 279*, 384−388
39. Carteau, S., Batson, S. C., Poljak, L., Mouscadet, J. F., de Rocquigny, H., Darlix, J. L., Roques, B. P., Kas, E., and Auclair, C. (1997) *J. Virol. 71*, 6225−6229.

BI9905258