Nucleic Acid Sequence Discrimination by the HIV-1 Nucleocapsid Protein NCp7: A Fluorescence Study†

C. Vuilleumier,‡ E. Bombarda,‡ N. Morellet,§ D. Gérard,‡ B. P. Roques,§ and Y. Mély*,‡

Laboratoire de Pharmacologie et Physico-Chimie des Interactions Cellulaires et Moléculaires, UMR 7034 CNRS, Faculté de Pharmacie, Université Louis Pasteur, Strasbourg 1, 74, Route du Rhin 67401 Illkirch Cedex, France, and Département de Pharmacochimie Moléculaire et Structurale, INSERM U266, CNRS UMR 8600, Faculté de Pharmacie, 4, Avenue de l'Observatoire, 75270 PARIS Cedex 06, France

Received May 18, 1999; Revised Manuscript Received October 13, 1999

ABSTRACT: The critical functions of the HIV-1 nucleocapsid protein NCp7 in genomic RNA packaging and reverse transcription, essentially rely on interactions with nucleic acids. A significant progress in the knowledge of these interactions has been recently achieved with the NMR-derived structures of NCp7 derivatives in complex with two short sequences of the HIV-1 ψ packaging signal, namely ACGCC and the stem-loop 3 (SL3) motif. To further identify the key nucleotides in the formation of both NCp7−d(ACGCC) and NCp7−SL3 complexes, we quantitatively analyzed by steady-state and time-resolved fluorescence, the interaction of NCp7 with d(ACGCC) and SL3 mutants where each nucleotide in interaction with the protein has been systematically substituted. Moreover, by using several NCp7 derivatives, we investigated the contributions of Phe16, Trp37, and Trp61, and the various NCp7 domains, in the binding process. The binding of NCp7 appeared essentially driven by the interaction of the zinc finger domain and notably Trp37 with a G residue, irrespective of its location in the oligonucleotide. The involvement of Trp37 in the binding process depended on its location in the C-terminal finger motif and the proper folding of this motif. Phe16 in the N-terminal finger motif also strongly contributed to the binding energy, while in contrast, Trp61 in the C-terminal domain only marginally interacted with the oligonucleotides. The stem-loop structure of SL3 stabilized the binding of NCp7 by about −7 kJ/mol (at 0.1 M NaCl) by favoring the electrostatic binding of both N- and C-terminal domains. Finally, we found that NCp7 bound to nucleic acid single-stranded regions with the following preference: $X_iTGX_j > X_iGXGX_j \approx X_iTXGX_j > X_iGX_j \gg X_iX_j$, where X corresponds to either A or C. This implies that recognition of nucleic acids by NCp7 may be achieved by a limited number of sites, and hence, no strong affinities are required in order to get a selective binding.

HIV-1[1] viral particles are composed of a capsid surrounded by an outer envelope, and within the capsid lies the nucleocapsid. In mature virions, the nucleocapsid structure contains the genomic 70S RNA extensively coated by about 2000 copies of the nucleocapsid protein NCp7 (*1*, *2*). NCp7 is generated upon protease-directed cleavage of the "*gag*-encoded" Pr55 polyprotein and, depending on HIV-1 isolates considered, is composed of either 55 or 71 amino acids (*3*). Besides its histone-like activity (*1*, *2*, *4*), NCp7 (either as the mature protein or as the Pr55 precursor) plays critical roles in the viral life cycle (for review, see ref *5*). Indeed, NCp7 is thought to specifically interact with the ψ encapsidation sequence of the genomic RNA, enabling its selection from the pool of cellular RNAs and its packaging (*6*−*8*). Interestingly, NCp7 also promotes a variety of nucleic acid annealing and strand transfer processes that are essential during HIV-1 replication, such as RNA dimerization (*1*), hybridization of primer $tRNA^{Lys,3}$ to the genomic RNA (4), DNA strand transfer, and renaturation (*5*, *9*, *10*).

NCp7 is a highly basic protein with a central domain of two zinc fingers or CCHC motifs (*11*), which strongly bind zinc (*12*−*14*). The protein is also characterized by two Trp residues at position 37 and 61, respectively, that constitute sensitive intrinsic fluorescent probes (*14*−*16*). Trp37 is located in the C-terminal CCHC motif and is critical for NCp7 functions since conversion of this residue into a nonaromatic one abolishes viral infectivity and impairs genomic RNA packaging (*17*). Interestingly, a similar phenotype is also obtained by point mutation of Phe16 that is symmetrically located in the N-terminal CCHC motif or by point mutation of any of the zinc ligand in the CCHC motifs (*17*−*21*). On the contrary, no key function of Trp 61 in the C-terminal domain of NCp7 has yet been found.

Recently, a significant progress in the knowledge of the interaction of NCp7 with nucleic acids has been achieved with the NMR-derived structures of (12−53)NCp7 bound

† This work was supported by grants from the Agence Nationale de Recherches sur le SIDA, Centre National de la Recherche Scientifique and Université Louis Pasteur.

* To whom all correspondence should be addressed. Phone: +33 (0)3 88 67 69 28. Fax: +33 (0)3 88 67 40 11. E-mail: mely@pharma.u-strasbg.fr.

‡ Laboratoire de Pharmacologie et Physico-Chimie des Interactions Cellulaires et Moléculaires.

§ Département de Pharmacochimie Moléculaire et Structurale.

[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; NCp7, nucleocapsid protein; SL, stem-loop; MoMuLV, Moloney murine leukaemia virus.

10.1021/bi991145p CCC: $18.00
Published on Web 12/03/1999

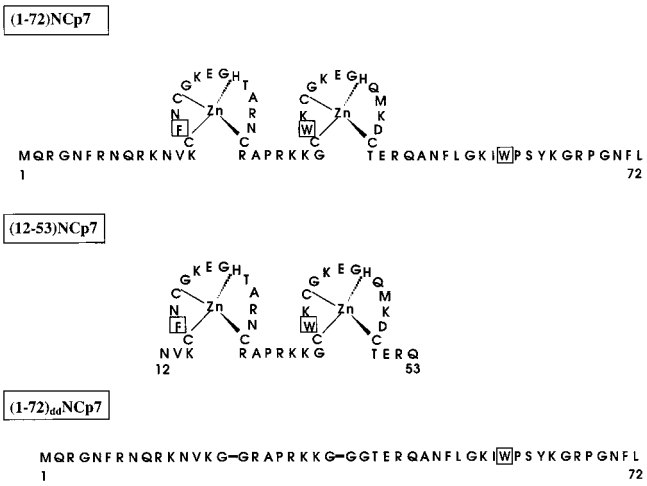

FIGURE 1: Amino acid sequences of NCp7 and its derivatives. The aromatic residues Phe16, Trp37, and Trp61 are framed.

to a d(ACGCC) sequence derived from the SL2 motif of the ψ packaging signal (*22*) and (1−55)NCp7 bound to a 20-nt RNA containing the ψ-SL3 motif (*23*). The two structures exhibit striking similarities, notably, in the position of NCp7 finger domain in respect to the nucleotides and the nature of the bonds (*22*). In both structures, Trp37 stacks with a G residue and Phe16 is involved in hydrophobic interactions with an oligonucleotide base [C^2 in d(ACGCC) and G^9 in SL3].

To identify the nucleotide determinants in both NCp7−d(ACGCC) and NCp7−SL3 complexes, we quantitatively analyzed the interaction of NCp7 with d(ACGCC) and SL3 mutants where each residue in interaction with the protein has been systematically modified. This investigation was performed by steady-state and time-resolved fluorescence using the Trp residues of NCp7 as intrinsic reporters. Moreover, we also investigated the interaction of NCp7 with a 10-nt RNA sequence located in the ψ signal next to ACGCC and critical for efficient genomic RNA packaging (*1*). In addition, several NCp7 derivatives have been used in order to investigate the contributions of Phe16, Trp37, and Trp61 and the various domains of NCp7 in the binding process.

MATERIALS AND METHODS

Materials. Solid-phase synthesis of NCp7 derivatives (Figure 1) was carried out as previously described (*24*). To preserve the highly oxidizable Cys residues, the lyophilized protein was dissolved in freshly degassed 50 mM Hepes (pH 7.5) buffer, poured into anaerobic quartz cells, and saturated with a 2.5 molar ratio of zinc sulfate to protein.

The oligonucleotides were purchased by Eurogentec and diluted in water. The r($C_2A_2U_5G$) sequence was a gift from T. Huyn Dynh (Institut Pasteur, Paris).

Spectroscopic Measurements. Absorption spectra were recorded on a Cary 4 spectrophotometer to determine the NCp7 and oligonucleotide concentrations. Extinction coefficients of 12 700, 5700, and 7000 M^{-1} cm^{-1} were used for (1−72)NCp7, (12−53)NCp7, and $(1-72)_{dd}$NCp7, respectively. Oligonucleotide concentrations (expressed in strands) were calculated by using the extinction coefficients provided by the manufacturer. Fluorescence measurements were performed at 20.0 ± 0.5 °C on an SLM 48000 spectrofluorometer. Excitation and emission wavelengths were 295 and 350 nm, respectively. The excitation and emission bandwidths were 4 and 8 nm, respectively.

Fluorescence titrations were performed by adding increasing concentrations of nucleic acid to a fixed amount of peptide in 50 mM Hepes and 0.1 M NaCl, pH 7.5, in the presence of 0.04% PEG 20000 to avoid the adsorption of NCp7 onto quartz cell walls (*25*). Fluorescence intensities were corrected for dilution, buffer fluorescence, and screening effects due to the presence of nucleic acid. The average number, ν, of moles of protein bound per mole of oligonucleotide is calculated from the fluorescence intensities using $\nu = (I_0 - I)/(I_0 - I_t) \times L_t/N_t$, where L_t and N_t designate the total concentration of protein and oligonucleotide, respectively, I_t, the fluorescence at the plateau when all the peptide is bound, whereas I_0 and I correspond to the fluorescence intensities of the peptide in the absence and in the presence of a given concentration of oligonucleotide. Furthermore, the free protein concentration, L, is deduced by $L = L_t - \nu N_t$. As the observed binding constant, K_{obs}, of NCp7 for some oligonucleotides was too low to allow the determination of the plateau of fluorescence, I_t, in our conditions, the Scatchard equation was rewritten to directly fit the fluorescence intensity, I, and calculate the plateau value as a parameter:

$$I = I_0 - \frac{(I_0 - I_t)}{L_t} \times \frac{[1 + (L_t + nN_t)K_{obs}]^2 - \sqrt{[1 + (L_t + nN_t)K_{obs}]^2 - 4L_t nN_t K_{obs}^2}}{2K_{obs}} \quad (1)$$

The parameters were recovered from the fit of eq 1 to experimental data by using the nonlinear procedure of the SAS software. The number, n, of protein-binding sites was determined from titrations under stoichiometric conditions.

Reverse titrations were performed by adding increasing concentrations of NaCl to the previously formed complexes. As the plots of log K_{obs} vs log[NaCl] were linear in a 100−600 mM salt concentration range, this method allows the determination of the nonelectrostatic binding constant K (1 M) and the number, m', of ion pairs between protein and nucleic acid, as previously described (*16*):

$$\log K_{obs} = \log K\,(1\text{ M}) - \Psi_{Na+} m' \log[Na^+] \quad (2)$$

where Ψ_{Na+}, the fraction of cation Na^+ thermodynamically bound per phosphate group, is assumed to be 0.71 as for single-stranded DNA (*26*). As it has been previously shown that no anion uptake or release accompanied the binding of NCp7 to various tRNAs (*16*), we assumed a similar behavior for the oligonucleotides investigated in this study.

Fluorescence lifetime measurements were performed with a time-correlated, single photon counting technique using a pulse-picked frequency-tripled Ti-sapphire laser pumped by a continuous wave argon laser (Spectra Physics). The excitation and emission wavelengths were set at 295 and 350 nm, respectively. The single-photon pulses were detected with a microchannel plate Hamamatsu R3809U photomultiplier and recorded on a multichannel analyzer (Ortec 921) calibrated at 25.5 ps/channel. The instrumental response

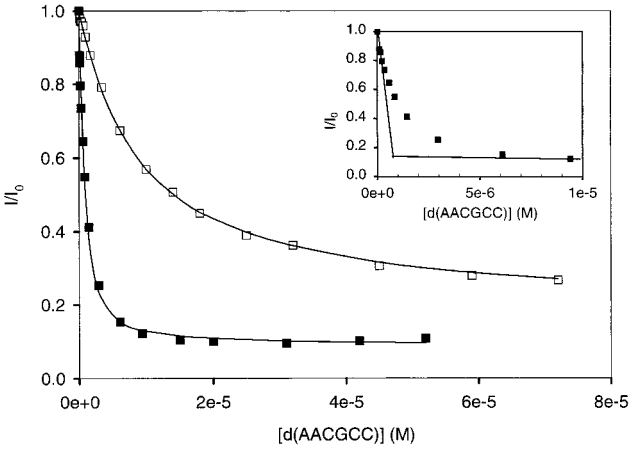

FIGURE 2: Binding curves of (12−53)NCp7 with d(AACGCC). Peptide concentration was 0.8 μM in 50 mM Hepes, pH 7.5, in the absence (closed symbols) or in the presence of 100 mM NaCl (open symbols). The oligonucleotide concentration is expressed in strands. Solid lines correspond to the fit of the experimental points with eq 1 and the parameters of Table 1. Insert, magnification of the titration in the absence of salt. The intercept of the initial slope with the plateau gives the number of protein binding sites on d(AACGCC).

function was recorded with a polished aluminum reflector, and its full-width at half-maximum was 40 ps. The decay data were analyzed with the maximum entropy method (MEM) and the Pulse5 software (*27*). A lifetime domain spanning 200 equally spaced values on a logarithmic scale between 0.01 and 10 ns was routinely used. The mean fluorescence lifetime was calculated by $\langle\tau\rangle = \sum\alpha_i\tau_i/\sum\alpha_i$, where τ_i and α_i designate the fluorescence lifetimes and relative amplitudes.

RESULTS

The use of NCp7 intrinsic fluorescence has proven to be a valuable tool to determine the binding parameters to various oligonucleotides and nucleic acids (*16*, *28*−*31*). To examine the nucleotide dependence of the binding of NCp7 to nucleic acids, we investigated the binding parameters and intrinsic fluorescence properties of NCp7 and NCp mutants in interaction with oligonucleotides representing sequences of HIV-1 ψ packaging signal. The ACGCC sequence was first used since it represents part of the ψ-RNA SL2 motif and since its structure in association with the proximal finger motif, (13−30)NCp7, and the central two-finger domain, (12−53)NCp7, has been solved by NMR (*22*, *32*). The more stable DNA analogue was used instead of the RNA oligonucleotide since both forms gave similar binding curves (*22*). Finally, as the binding site size for (1−72)NCp7 on various nucleic acids is between 6 and 8 (*9*, *16*, *29*, *33*, *34*), an additional A nucleotide was added at the 5′ end of d(ACGCC) to give d(AACGCC).

Binding of NCp7 and NCp Mutants to d(AACGCC)

Titration of the Trp37-containing (12−53)NCp7 by d(AACGCC) in the absence of salt is given in Figure 2. As expected, a single protein-binding site on d(AACGCC) was inferred from the intersection of the initial slope with the plateau (Figure 2, insert). Moreover, both the observed binding constant, K_{obs}, and the maximum extent of quenching, Q_{max}, were close to those reported for d(ACGCC) (*22*), suggesting that the addition of A at the 5′ end did not significantly modify the binding. As in these conditions, the ionic strength is ill-defined and far from the physiological one, the rest of the titrations were performed in the presence of 100 mM NaCl (Figure 2 and Table 1). As expected, Q_{max} was not affected by salt, indicating that the behavior of Trp37 in d(AACGCC)−(12−53)NCp7 did not depend on salt concentration. In contrast, there was a 15-fold decrease in K_{obs}, in keeping with the well-known salt dependence of protein−nucleic acid interactions (*26*). In reasonable agreement with the three electrostatic interactions that have been deduced from NMR data (22), $m' = 2.3(\pm0.2)$ ion pairs were determined in the (12−53)NCp7−d(AACGCC) complex from the slope of log K_{obs} vs log[NaCl] in salt-back-titrations. Since m' was obtained by dividing the slope value by Ψ_{Na+}, the number of Na^+ bound per phosphate group, the slight underestimation of m' in our data may come from the value of Ψ_{Na+}, which is probably lower than 0.71 for short and unstructured oligonucleotides (*26*). Finally, in keeping with the numerous H bonds and hydrophobic interactions evidenced in (12−53)NCp7−d(ACGCC), the nonelectrostatic interactions represented about 70% of the binding energy.

To get additional information, a time-resolved fluorescence investigation was performed. In the absence of d(AACGCC), the decay of (12−53)NCp7 was characterized by three lifetimes with a very narrow distribution (Table 2). These decay parameters should be considered as refinements of those previously published (*15*) since the device used to measure the decays was significantly improved. In the presence of an oligonucleotide concentration that saturates more than 98% of the peptide, the three lifetimes were not or only poorly affected, but their relative proportions strongly decreased to the benefit of an ultrashort lifetime (60 ps) whose appearance is in full agreement with a stacking of Trp37 with d(AACGCC) (*35*). Noticeably, the ratio of the mean fluorescence lifetime in the presence of d(AACGCC) to that in its absence was in excellent agreement with the corresponding quantum yield ratio (Table 2), suggesting that no additional static quenching occurred. Stacking was clearly the major mode of interaction of Trp37 in the complex but the 25% relative proportion not associated to the ultrashort lifetime suggested that, in a significant population of (12−53)NCp7−d(AACGCC) complexes, Trp37 may not be involved in stacking interactions.

To check the dependence of the binding and fluorescence parameters of d(AACGCC)−(12−53)NCp7 on the location of Trp37 in the C-terminal finger, a W16F37(12−53)NCp7 derivative was synthesized, where Phe16 and Trp37 were inverted. This double mutation only moderately decreased the binding constant for d(AACGCC) but strongly modified the ratio of electrostatic versus nonelectrostatic interactions (Table 1) in keeping with significant changes in the binding mode. Moreover, the Q_{max} value of Trp16 in W16F37(12−53)NCp7−d(AACGCC) was only half of that of Trp37 in (12−53)NCp7−d(AACGCC). In fact, time-resolved fluorescence measurements indicated that a stacking of Trp16 with d(AACGCC) bases occurred only for 33% of the W16F37(12−53)NCp7−d(AACGCC) complexes (data not shown).

The replacement of (12−53)NCp7 by (1−72)NCp7 only led to a 2.5-fold K_{obs} increase (Table 1), suggesting that the N- and C-terminal domains flanking the finger domain of

Table 1: Binding Parameters of NCp7 and NCp-Mutants to d(AACGCC)[a]

peptide	[NaCl] (mM)	K_{obs}[b] × 10^{-5} (M^{-1})	Q_{max}[b] (%)	K (1 M)[b] × 10^{-4} (M^{-1})	m' [b]
(12−53)NCp7	0	17(±2)	87(±2)	0.5(±0.1)	2.3(±0.3)
	100	1.1(±0.2)	85(±3)		
(1−72)NCp7	100	2.6(±0.1)	65(±1)	0.7(±0.1)	2.3(±0.2)
$(1-72)_{dd}$NCp7	100	<0.05	<10	ND[c]	ND[c]
W16F37(12−53)NCp7	100	0.70(±0.05)	48(±1)	0.03(±0.01)	3.3(±0.2)

[a] The experiments were performed in 50 mM Hepes, pH 7.5. The primary sructures of the peptides used are given in Figure 1. [b] The observed affinity constant, K_{obs}, the maximum extent of fluorescence quenching, Q_{max}, the number of ion pairs, m', and the nonelectrostatic binding constant, K (1 M), were expressed as mean (±standard error of the mean) for three experiments. [c] Non determined.

Table 2: Time-Resolved Fluorescence Parameters of NCp7 and NCp-Mutants in the Presence of d(AACGCC)[a]

	[AACGCC] (μM)	τ_i[b] (ns)	α_i[b] (%)	f_i[b] (%)	$\langle\tau\rangle/\langle\tau_0\rangle$[c]	I/I_0[c]
(12−53)NCp7	0	7.3 (±0.1)	35 (±1)	65 (±5)	1 (4.1)	1
		3.4 (±0.1)	40 (±1)	33 (±2)		
		0.64 (±0.02)	25 (±1)	4 (±1)		
	20	7.4 (±0.4)	3 (±1)	37 (±9)	0.15 (0.6)	0.14
		2.7 (±0.2)	10 (±1)	44 (±9)		
		0.63 (±0.03)	11 (±3)	11 (±5)		
		0.06 (±0.01)	76 (±2)	8 (±3)		
$(1-72)_{dd}$NCp7	0 or 20	4.8 (±0.1)	19 (±3)	38 (±5)	1 (2.5)	1
		3.0 (±0.1)	43 (±3)	51 (±5)		
		0.74 (±0.03)	38 (±1)	11 (±1)		
NCp7	0	6.5 (±0.1)	24 (±2)	50 (±4)	1 (3.1)	1
		3.3 (±0.1)	39 (±3)	41 (±4)		
		0.72 (±0.03)	36 (±1)	8 (±1)		
	20	5.6 (±0.1)	8 (±1)	35 (±4)	0.42 (1.3)	0.38
		2.9 (±0.1)	22 (±2)	50 (±5)		
		0.64 (±0.02)	24 (±1)	12 (±2)		
		0.07 (±0.01)	46 (±2)	3 (±1)		
linear combination[d]		5.3	10	37	0.44 (1.4)	
model		2.9	23	50		
		0.72	21	11		
		0.06	46	2		

[a] Experiments were performed in the absence of salt to get a full complexation of (12−53)NCp7 and (1−72)NCp7 with d(AACGCC). Peptide concentration was 1 μM. [b] The barycenter value, τ_i, the relative proportion, α_i and the fractional intensity, f_i of each lifetime are expressed as mean (± standard error of the mean) for three experiments. [c] $\langle\tau\rangle$ and I designate the mean fluorescence lifetime and steady-state fluorescence intensity in the presence of d(AACGCC). $\langle\tau\rangle_0$ and I_0 are the corresponding values in the absence of d(AACGCC). The mean fluorescence lifetimes (expressed in nanoseconds) are given in parentheses. [d] In the linear combination model, the values are calculated as described in the text.

NCp7 were not critically involved in the binding to d(AACGCC). Moreover, the absence of changes in m' further suggested that the numerous positively charged amino acids of the N- and C-terminal domains might not interact with the oligonucleotide phosphate groups. The absence of strong interaction of the N- and C-terminal domains of NCp7 with the oligonucleotide was confirmed with the fingerless $(1-72)_{dd}$NCp7 peptide, since competition with (12−53)NCp7 for d(AACGCC) indicated that K_{obs} of $(1-72)_{dd}$NCp7 was very low and did not exceed 5×10^3 M^{-1} (Table 1). Interestingly, the presence of zinc in NCp7 was critical for the binding to d(AACGCC), since its removal by EDTA decreased K_{obs} by about 2 orders of magnitude (data not shown).

The Q_{max} value of (1−72)NCp7 was significantly lower than the (12−53)NCp7 one. To get further information, time-resolved fluorescence measurements were performed. In the absence of d(AACGCC), the fluorescence decay of (1−72)-NCp7 was characterized by a trimodal distribution (Table 2) that has been shown to correspond to a linear combination, in respect to both lifetimes and fractional intensities, of the individual emitting Trp residues (namely Trp 37 and 61) in the two single-Trp containing derivatives, (12−53)NCp7 and $(1-72)_{dd}$NCp7 (*36*). Similarly, the fluorescence decay parameters of (1−72)NCp7−d(AACGCC) were in excellent agreement with the theoretical parameters calculated from Trp37 in (12−53)NCp7−d(AACGCC) and Trp61 in free $(1-72)_{dd}$NCp7, assuming that a linear combination applies (Table 2). This suggests an identical behavior of Trp37 in (1−72)NCp7−d(AACGCC) and (12−53)NCp7−d(AACGCC) and marginal interaction of Trp61 with the oligonucleotide.

Finally, the substitution of d(AACGCC) with its RNA analogue, r(AACGCC), led only to a limited change in the binding and fluorescence parameters (Table 3), confirming that the complexes of NCp7 with both the sequences were similar (*22*).

Dependence of the Binding Process on the Hexanucleotide Sequence

Critical Involvement of G Residues. To further examine the binding of NCp7 to oligonucleotides, we analyzed the effects of systematic substitutions in d(AACGCC) on the binding process. Since G has been shown to stack with Trp37 in (12−53)NCp7−d(ACGCC) (*22*), G was changed to either A, C, or T (Table 3). In each case, a dramatic K_{obs} decrease, associated with a corresponding decrease in the nonelectrostatic binding constant, K (1 M), was observed. The loss of binding energy was between 5 and 7 kJ/mol, suggesting that the interactions of the guanine base with NCp7 residues contributed to about 20% of the total binding energy of (1−72)NCp7 to d(AACGCC). In contrast, Q_{max} was only poorly

Table 3: Sequence Dependence of the Binding of (1−72)NCp7 to Hexanucleotides[a]

oligo	K_{obs}[b] $\times 10^{-5}$ (M^{-1})	Q_{max}[b] (%)	K (1 M)[b] $\times 10^{-4}$ (M^{-1})	m'[b]	$\Delta\Delta G$[c] (kJ/mol)
d(AACGCC)	2.6(±0.1)	65(±3)	0.7(±0.1)	2.3(±0.2)	0.0
r(AACGCC)	3.9(±0.3)	68(±3)	0.8(±0.1)	2.3(±0.2)	−1.0
d(AAC**T**CC)	0.35(±0.02)	70(±1)	0.14(±0.05)	2.0(±0.2)	4.9
d(AAC**C**CC)	0.16(±0.05)	71(±2)	ND	ND	6.8
d(AAC**A**CC)	0.13(±0.03)	66(±3)	ND	ND	7.3
d(**G**CC**C**CC)	1.5(±0.2)	70(±3)	0.4(±0.1)	2.3(±0.2)	1.3
d(**C**CC**C**CC)	0.18(±0.03)	67(±3)	0.16(±0.04)	2.0(±0.3)	6.5
d(AA**AAAA**)	0.07(±0.02)	47(±4)	ND	ND	8.8
d(AA**A**GCC)	1.3(±0.1)	53(±1)	0.14(±0.01)	2.9(±0.1)	1.7
d(AA**G**GCC)	1.6(±0.2)	65(±1)	0.15(±0.03)	2.7(±0.1)	1.2
d(AA**T**GCC)	38[d] (±3)	65(±1)	7.0(±0.5)	2.4(±0.1)	−6.5
r(AA**U**GCC)	35[d] (±2)	66(±2)	7.2(±0.4)	2.3(±0.2)	−6.3
d(**CCGTAA**)	4.3(±0.4)	68(±3)	0.8(±0.2)	2.3(±0.2)	−1.2
d(AACG**A**C)	1.9(±0.4)	49(±2)	0.29(±0.02)	2.1(±0.1)	0.8
d(AACG**G**C)	4.8(±0.3)	71(±2)	0.5(±0.1)	2.8(±0.2)	−1.5
d(AACG**T**C)	2.8(±0.3)	64(±3)	0.49(±0.01)	2.4(±0.1)	−0.2
d(A**G**CGCC)	8(±3)	66(±2)	1.4(±0.2)	2.4(±0.1)	−2.7
d(A**T**CGCC)	9(±2)	69(±2)	1.2(±0.2)	2.5(±0.2)	−3.0
d(A**C**CGCC)	4(±1)	68(±2)	0.6(±0.1)	2.4(±0.2)	−1.0
d(**TGT**GCC)	71[d] (±6)	66(±2)	30(±1)	2.0(±0.1)	−8.1

[a] The experiments were performed in 50 mM Hepes and 100 mM NaCl, pH 7.5. [b] Parameter significance and expression are as in Table 1. [c] $\Delta\Delta G$ corresponded to the difference between the free energy measured with a given oligonucleotide and the free energy measured with d(AACGCC) at 0.1 M NaCl. [d] For the mutants with the highest affinities, K_{obs} was deduced from salt-back-titrations.

affected by G^4 substitution, suggesting that Trp37 also stacks with A, C, or T bases but probably with a lower stability. It may also be inferred that the 5′ AAC and 3′ CC residues in d(AACGCC) were unable to strongly interact with Trp37 or other NCp7 residues. This inability of A's or C's to strongly interact with NCp7 was confirmed by the low K_{obs} obtained with either d(**C_6**) or d(**A_6**).

The critical role of G in hexanucleotide-NCp7 interaction was further investigated by using d(**GC_5**). In contrast to the marginal binding of NCp7 to d(**C_6**), the K_{obs} value for d(**GC_5**) was only 2-fold reduced as compared to d(AACGCC). Moreover, the time-resolved fluorescence decays of NCp7−d(**GC_5**) and NCp7−d(AACGCC) were largely similar (data not shown), suggesting that the two Trp residues of NCp7 have the same behavior in the two complexes. Thus, we conclude that the interaction of NCp7 and probably the stacking of Trp37 with a G drives the binding of NCp7 to hexanucleotides.

Involvement of the Nucleotides Flanking the G Residue. The dependence of NCp7 binding on the oligonucleotide sequence was further investigated by substituting C^3 in d(AACGCC) by either A, G, or T (Table 3). Only a limited K_{obs} decrease was observed when C^3 was replaced by G or A. In contrast, a 15-fold increase in K_{obs} accompanied the replacement of C^3 by T. This large increase in K_{obs} was associated with a corresponding increase in K (1 M) that stabilizes the NCp7−oligonucleotide complex by about −6.5 kJ/mol. Interestingly, Q_{max} and the fluorescence decay parameters were not affected by T^3 substitution, suggesting that the behavior of the two Trp residues was not modified. A similar sharp K_{obs} increase was observed for (12−53)-NCp7 (data not shown), suggesting that the stabilization induced by T^3 did not involve the N- and C-terminal domains of NCp7. Interestingly, the binding and fluorescence parameters of the interaction of NCp7 with r(AA**U**GCC) and d(AA**T**GCC) were strikingly similar (Table 3), suggesting that neither the ribonucleotidic nature nor the replacement of T by U significantly affected the binding process. Finally, the binding of NCp7 to d(AA**T**GCC) was found to be polarized since the K_{obs} for the palindromic sequence, d(**CCGTAA**), showed only a limited increase as compared to K_{obs} of d(AACGCC).

Next, we substituted C^5 in d(AACGCC) by either A, G, or T (Table 3). Neither of these substitutions significantly modified K_{obs} or Q_{max}, except for the A substitution that induced a limited decrease in Q_{max}. A similar conclusion held true for (12−53)NCp7, suggesting that the substitution of C^5 by A modified to some extent the interaction of Trp37 with the hexanucleotide.

Involvement of the 5′ End Nucleotides. By substituting A^2, we found that either G or T (but not C) at this position induced a significant 3-fold increase in K_{obs} due to additional nonelectrostatic interactions (Table 3). To further investigate the stabilization effect of G residues, we analyzed the interaction of NCp7 with d(**G_6**). Unfortunately, the appearance of light scattering indicated an aggregation of d(**G_6**)−NCp7 complexes as in the case of polyA (*37*). This prevented an accurate determination of K_{obs}, but a rough comparison of the binding curves suggested that the binding parameters of NCp7 for d(**G_6**) and d(A**G**CGCC) were similar (data not shown), and thus, no strong interactions of G residues at positions 1, 3, 5, or 6 with NCp7 amino acids may occur.

Since Fisher et al. (*42*) have reported that sequences with TG repeats present very high affinities for NCp7, we investigated the binding parameters of NCp7 with d(**TGT**-GCC). This sequence was found to show the highest affinity among all the hexanucleotides we tested, although its K_{obs} was only twice that of d(AC**TG**CC). The binding energy associated to this oligonucleotide exceeded that of d(AACGCC) by −8.1 kJ/mol, a stabilization that is entirely related to nonelectrostatic interactions.

Binding of NCp7 and NCp-Mutants to d($C_2A_2T_5G$) and SL3

Binding to d($C_2A_2T_5G$). To confirm the conclusions drawn from the hexanucleotides, we investigated the interaction of NCp7 with d($C_2A_2T_5G$), the DNA analogue of r($C_2A_2U_5G$),

Table 4: Binding Parameters of NCp7 to d($C_2A_2T_5G$) Derivatives[a]

oligonucleotide	n[b]	K_{obs}[c] $\times 10^{-5}$ (M^{-1})	Q_{max}[c] (%)	K (1 M)[c] $\times 10^{-4}$ (M^{-1})	m'[c]
d($C_2A_2T_5G$)	1.1(±0.2)	27(±3)	68(±1)	4.0(±0.1)	2.5(±0.1)
d($C_2A_2T_5$)	1.0(±0.2)	2.5(±0.2)	66(±2)	0.6(±0.1)	2.3(±0.2)
r($C_2A_2U_5G$)	1.2(±0.2)	21(±3)	70(±2)	3.1(±0.1)	2.8(±0.3)

[a] Experiments were performed in 50 mM Hepes and 100 mM NaCl, pH 7.5. [b] The number n of protein molecules bound per oligonucleotide was obtained by fitting the titration curves with eq 1. [c] Parameter significance and expression are as in Table 1.

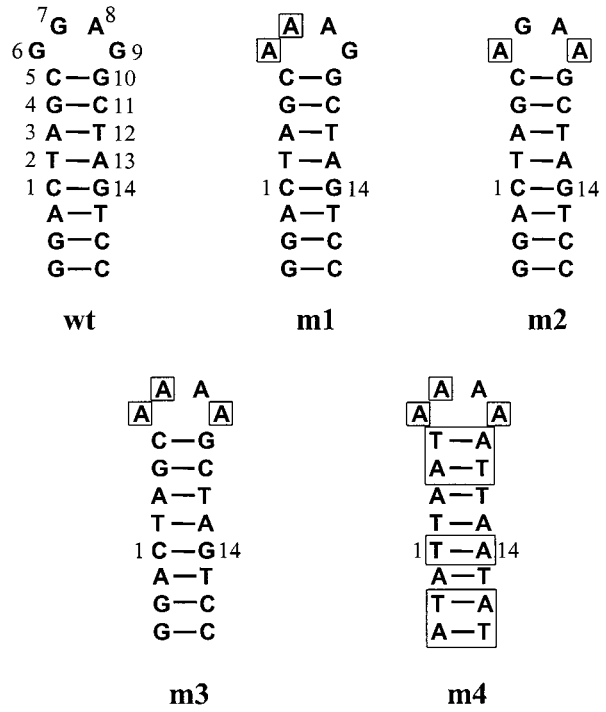

FIGURE 3: Nucleotide sequence of the 20 nucleotide SL3 ψ packaging signal and its mutants. Mutated nucleotides are in boxes. The numbering of nucleotides is as in De Guzman et al. (*23*).

a putative NCp7-binding site (*1*), located next to ACGCC in the ψ packaging signal from HIV-1 strain MAL. Stoichiometric titrations performed in the absence of salt and fits of titrations performed at 0.1 M NaCl (Table 4) revealed a single protein-binding site on d($C_2A_2T_5G$). This was not unexpected since the hexanucleotide data indicate that the 5′ end CCAA sequence may not promote a tight binding of NCp7. Moreover, Q_{max} and K_{obs} were close to those obtained with d(AA**T**GCC), suggesting that NCp7 may bind to the 3′ end of d($C_2A_2T_5G$) in order to achieve a strong interaction with the TG pair. This hypothesis was confirmed by the analysis of the interaction of NCp7 with a d($C_2A_2T_5$) sequence where the 3′ end G residue has been removed. Indeed, a 10-fold decrease in K_{obs} associated to a corresponding K (1 M) decrease and a 5.7 kJ/mol loss in binding energy resulted from 3′G removal. This energy loss was close to that accompanying the substitution of G[4] in d(AACGCC) (Table 3). These data confirmed that the interaction of NCp7 with G is a driving event in the binding process. Moreover, since no large differences in the binding parameters were observed when NCp7 was substituted by (12−53)NCp7, it is suggested that the N- and C-terminal domains do not significantly contribute to the binding of NCp7 to d($C_2A_2T_5G$) (data not shown). Finally, the binding parameters of NCp7 with d($C_2A_2T_5G$) and r($C_2A_2U_5G$) were highly similar, again suggesting that neither the ribonucleotidic nature nor the replacement of T by U significantly affected the binding process.

Table 5: Time-Resolved Fluorescence Parameters of NCp7 and NCp-Mutants Complexed to SL3[a]

	τ_i[b] (ns)	α_i[b] (%)	f_i[b] (%)	$\langle\tau\rangle/\langle\tau_0\rangle$[b]	I/I_0[b]
(12−53)NCp7	7.4 (±0.2)	3 (±1)	39 (±9)	0.14 (0.6)	0.15
	2.5 (±0.1)	7 (±1)	31 (±9)		
	0.60 (±0.02)	18 (±2)	19 (±2)		
	0.09 (±0.01)	72 (±3)	11 (±2)		
$(1-72)_{dd}$NCp7	4.7 (±0.1)	7 (±2)	23 (±5)	0.60 (1.45)	0.64
	2.7 (±0.1)	32 (±1)	60 (±7)		
	0.60 (±0.02)	37 (±2)	15 (±2)		
	0.10 (±0.01)	24 (±1)	2 (±1)		
NCp7	5.9 (±0.3)	5 (±1)	28 (±5)	0.32 (1.0)	0.33
	2.6 (±0.1)	22 (±1)	55 (±7)		
	0.58 (±0.02)	22 (±1)	12 (±1)		
	0.09 (±0.01)	51 (±1)	5 (±1)		
linear combination	5.65	4.5	27	0.30 (0.95)	
model	2.67	18.5	52		
	0.60	25	16		
	0.09	52	5		

[a] The SL3 concentration in nucleotides was 20 μM. [b] Parameter significance and expression are as in Table 2.

Binding of NCp7 to SL3 ψ-RNA. We next investigated the interaction of NCp7 with the DNA analogue of the SL3 ψ-RNA recognition element (Figure 3) whose structure in complex with (1−55)NCp7 has been solved (*23*). The Q_{max} value of NCp7 was somewhat higher with SL3 than with the shorter oligonucleotides. This was confirmed by the significant increase of the relative proportion associated to the ultrashort lifetime in NCp7−SL3 (Table 5) as compared to NCp7−d(AACGCC) (Table 2). In contrast, binding of SL3 to (12−53)NCp7 induced the same Q_{max} and fluorescence decay parameters (Table 5) as did d(AACGCC) (Table 2), suggesting that the stacking of Trp37 was similar in both complexes. Consequently, the high Q_{max} of NCp7−SL3 is probably linked to a partial quenching of Trp61. This was confirmed by the significant fluorescence decrease that accompanied the interaction of $(1-72)_{dd}$NCp7 with SL3 (Table 5). The appearance of an ultrashort lifetime in $(1-72)_{dd}$NCp7−SL3 complexes suggested that Trp61 may also be involved in stacking interactions, but the rather low relative proportion associated to this lifetime indicated that stacking only occurred for a small fraction of these complexes. Moreover, the time-resolved fluorescence parameters of NCp7−SL3 corresponded to a linear combination of the (12−53)NCp7−SL3 and $(1-72)_{dd}$NCp7−SL3 parameters, suggesting that the behavior of Trp37 and Trp61 in NCp7 may be similar to that in the mutants. These results were strikingly similar to those obtained with $tRNA^{Phe}$ (*36*) and suggested that, in contrast to the complexes with the hexanucleotides, the C-terminal domain of NCp7 probably stabilize the binding to SL3.

According to the data of de Guzman et al. (*23*), the existence of a single (1−55)NCp7−SL3 complex for equimo-

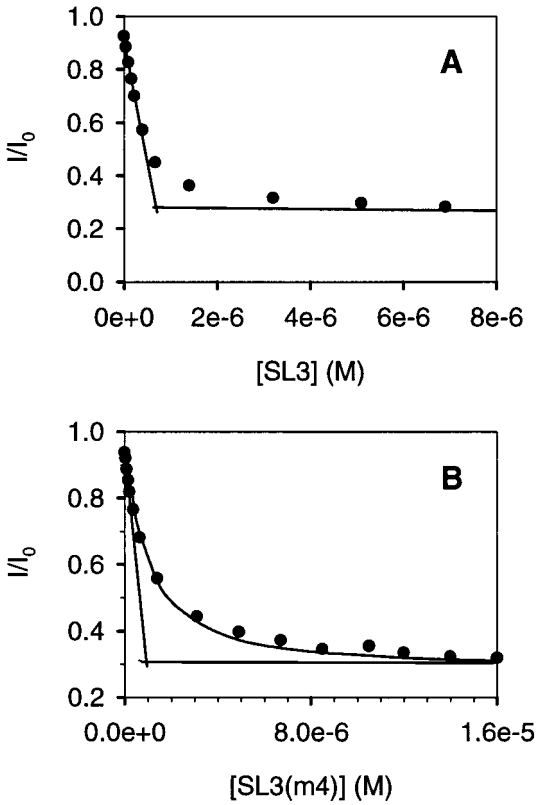

FIGURE 4: Binding curves of NCp7 with SL3 derivatives. NCp7 concentration was 0.8 μM in 50 mM Hepes, 0.1 M NaCl, pH 7.5. The SL3 (A) and SL3(m4) (B) concentrations are expressed in strands. The intercept of the initial slope with the plateau gives the number of protein binding sites on SL3 mutants. Experimental points in panel B are fitted with eq 1 and the parameters of Table 6.

lar concentrations of its components strongly suggests that the affinity of the protein is much higher for the SL3 loop than for the other putative binding sites on SL3. In line with this conclusion, we found that, in the presence of 0.1 M salt, NCp7 binds to SL3 with a 1:1 stoichiometry (Figure 4, Table 6). This suggested that no significant contribution of additional SL3-binding sites could be observed in our conditions. As K_{obs} was too high to be accurately determined from a titration at 0.1 M NaCl, it was deduced from a salt-back-titration. K_{obs} was about 20-fold larger for SL3 (Table 6) than d(A**G**CGCC) (Table 3), which is comparable to the SL3 loop sequence. In contrast, the binding parameters of (12−53)NCp7 with both oligonucleotides were similar, strongly suggesting that the additional −6.9 kJ/mol energy stabilization for NCp7−SL3 was related to the N- and C-terminal domains of NCp7. This additional stabilization was essentially due to the presence of two additional ion pairs and thus strongly depended on the salt concentration. The contribution of both NCp7 terminal domains was further assessed by the significant interaction of the fingerless peptide, $(1-72)_{dd}$NCp7 with SL3 (Table 6).

In a second step, we investigated the binding of NCp7 and (12−53)NCp7 to SL3 mutants (Figure 3). As Trp37 has been shown to stack with G^7 (*22*, *23*), we investigate the energetic contribution of this stacking by replacing G^7 by A in the SL3(m1) mutant. Moreover, to avoid the occurrence of a compensatory stacking of Trp37 with G^6, this residue was also substituted by an A residue. As expected, a decrease in K_{obs} and K (1 M), associated to a 1.2−1.7 kJ/mol energy loss as compared to SL3, was observed for both NCp7 and (12−53)NCp7 (Table 6).

The energetic contribution of the interaction of G^9 with NCp7 residues and notably Phe16 was investigated by examining the interaction of both peptides with a SL3 mutant, SL3(m2), where G^9 was replaced by A, and to facilitate the comparison with SL3(m1), G^6 residue was also replaced by A. For both NCp7 and (12−53)NCp7, the binding parameters were similar to those obtained with SL3(m1), suggesting that Phe16−G^9 and Trp37−G^7 interactions equally contributed to NCp7−SL3 stability.

A further K_{obs} and K (1 M) decrease as compared to SL3 was evidenced for the interaction of both peptides with the SL3(m3) mutant where the three G's of the SL3 loop were simultaneously replaced by A. The binding energy loss for SL3(m3) as compared to SL3 roughly corresponded to the sum of the energy losses observed for SL3(m1) and SL3-(m2) but was much lower than expected from the hexanucleotide results. Indeed, while a 2 orders of magnitude decrease in K_{obs} was observed for d(**A_6**) as compared to d(A**G**CGCC), an only 4−5-fold decrease was evidenced for SL3(m3) as compared to SL3. Since the residues of the (12−53) sequence in the (1−55)NCp7−SL3 complex (*23*) have been shown to exclusively bind to the four residues of the SL3 loop, this suggested that the unexpected high affinity of (12−53)NCp7−SL3(m3) may be due either to (i) a strengthening, induced by the stem-loop structure, of the interaction of the central finger domain with the A_4 loop or (ii) some differences in the location of (12−53) sequence on SL3(m3) as compared to the wild-type SL3. In the last hypothesis, some interactions with residues of the stem may be involved. Since the G residues may be good candidates for these additional interactions, the SL3(m4) mutant where all C−G pairs in the stem were replaced by A−T pairs was tested. The affinity of (12−53)NCp7 for SL3(m4) was further decreased by a factor of 7 as compared to SL3(m3). Moreover, the significant Q_{max} decrease suggested a less efficient stacking of Trp37. We checked that the substitution of C−G by A−T pairs has no deleterious effect by itself since the same mutation in SL3 did not induce any K_{obs} decrease as compared to the native SL3 (data not shown). Accordingly, our data suggested that the central finger domain of NCp7 is stabilized by the binding to at least one residue (probably a G) of the SL3(m3) stem, in keeping with a shift of the protein on SL3(m3) as compared to the native SL3.

DISCUSSION

Oligonucleotide and Protein Determinants in the Binding of NCp7 to d(AACGCC) and d($C_2A_2T_5G$). A comparison of the binding properties of (12−53)NCp7 and NCp7 to d(ACGCC) or d($C_2A_2T_5G$) indicated that the central finger domain drives the interaction of NCp7 with both oligonucleotides. In contrast, the lateral domains flanking the finger domain did only marginally interact with these oligonucleotides in line with the perpendicular orientation of d(ACGCC) with the short interfinger sequence (*22*), an orientation that is not favorable for an interaction with NCp7 terminal domains.

Substitution of G^4 in d(AACGCC) dramatically weakened the interaction with NCp7, indicating a critical role of G in

Table 6: Binding Parameters of NCp7 and NCp-Mutants to SL3 and Its Mutants

peptide	SL3	K_{obs}[a] × 10^{-5} (M^{-1})	Q_{max}[a] (%)	m'[a]	K (1 M)[a] × 10^{-4} (M^{-1})	$\Delta\Delta G$ (kJ/mol)
(1−72)NCp7	wt	140(±20)	75(±2)	4.6(±0.3)	0.74(±0.05)	0.0[b]
	m1	85(±9)	75(±3)	4.5(±0.2)	0.50(±0.05)	1.2[b]
	m2	66(±5)	75(±2)	4.5(±0.3)	0.37(±0.04)	1.8[b]
	m3	30(±4)	76(±2)	4.2(±0.2)	0.35(±0.02)	3.8[b]
	m4	11(±3)	70(±2)	4.5(±0.3)	0.07(±0.01)	6.2[b]
(12−53)NCp7	wt	10(±2)	87(±3)	2.9(±0.2)	0.89(±0.05)	0.0[c]
	m1	4.9(±0.3)	90(±2)	2.7(±0.3)	0.55(±0.03)	1.7[c]
	m2	3.7(±0.2)	87(±2)	2.8(±0.2)	0.34(±0.03)	2.4[c]
	m3	2.7(±0.3)	90(±3)	2.8(±0.2)	0.28(±0.04)	3.2[c]
	m4	0.37(±0.08)	83(±2)	2.8(±0.2)	0.033(±0.008)	8.0[c]
$(1-72)_{dd}$NCp7	wt	15(±2)	42(±3)	4.3(±0.4)	0.12(±0.02)	

[a] Parameter significance and expression are as in Table 1. The stoichiometry of binding, determined as described in Figure 4, was about one for each complex. [b] $\Delta\Delta G$ corresponded to the difference between the free energy measured with a given SL3 mutant and that measured with the native SL3 at 0.1 M NaCl, for (1−72)NCp7. [c] $\Delta\Delta G$ corresponded to the difference between the free energy measured with a given SL3 mutant and that measured with the native SL3 at 0.1 M NaCl for (12−53)NCp7.

the interaction. Moreover, comparison of the interaction of NCp7 derivatives with d(GC_5) and d(C_6) on one hand and with d($C_2A_2T_5G$) and d($C_2A_2T_5$) on the other hand strongly suggested that the interaction of the central finger domain with G drives the positioning of NCp7, irrespective of the position of G in the oligonucleotide sequence. According to the theoretical studies of Kumar and Govil (*38*), we infer that the stacking of Trp37 with G^4 base may largely contribute to the −5 to −7 kJ/mol stabilization energy associated to the interaction of G^4 with NCp7. The strong involvement of Trp37 is in line with the strong binding decrease observed with either a single-stranded DNA (*39*) or poly(ϵA) (*34*) when Trp37 has been replaced by a nonaromatic residue. Moreover, the biological relevance for a preference of a Trp37 residue in NCp7 is supported by the inability of Leu37(12−53)NCp7 to initiate the annealing of $tRNA^{Lys,3}$ to the primer-binding site PBS in vitro, in contrast to its native counterpart (*40*). However, the presence of Trp37 is not sufficient for strong binding since the zinc-free protein only weakly bound to d(AACGCC). As the solvent accessibility of Trp37 does not depend on zinc (*28*), this suggests that the proper folding of the finger motifs is also critical for binding. This requirement of a bona fide conformation of NCp7 has also been demonstrated by the effects of His23 → Cys mutation since the structural changes induced by this mutation modify the interaction of the finger domain with d(ACGCC) (*22*) and leads to noninfectious viruses (*41*).

In addition to the pivotal role of G, we evidenced that the nature of the nucleotide at position 3 was important, too, since only a T at this position gave a strong increase in K_{obs} and K (1 M). The −6.5 kJ/mol stabilization energy associated to the presence of T^3 may be related to the creation of specific hydrophobic interactions between T and several residues of NCp7 finger domain or, alternatively, the strengthening of the interactions described with C at the same position in the (12−53)NCp7−d(ACGCC) complex (*22*). In contrast, the nucleotide at position 5 seemed not critical. This conclusion held notably true for T, confirming that the binding of NCp7 to oligonucleotides may be polarized. Moreover, in agreement with the limited interactions of the 3′ end residue of d(ACGCC) with (12−53)NCp7 (*22*), this residue does not seem to strongly contribute to the stability of the complex.

The stability of the NCp7−hexanucleotide complex was further increased when A^2 was substituted by G or T, and an even higher affinity was reached with d(TGTGCC), in full keeping with the strong binding constants reported for sequences with TG repeats (*31*, *42*). The potency of both TG pairs to optimally interact with the two finger motifs of NCp7 may suggest some symmetry in the interaction of Phe16 and Trp37 with the two TG pairs. However, the nature of the aromatic amino acid at positions 16 and 37 was clearly not indifferent, since the inversion of the two residues in W16F37(12−53)NCp7 led to profound changes in the interaction with d(AACGCC). The critical role of the aromatic amino acids and their proper location in the finger motif may be a general rule for NC proteins since high-affinity binding to the NC of murine leukemia virus also requires that the Tyr and Trp residues be in their native positions in the zinc finger motif (*43*).

Taken together, our data on the hexanucleotides suggested a strong dependence of K_{obs} on the nucleotide sequence; the difference being more than 4 orders of magnitude between the strongest and the weakest binding sequences. According to the well-known preference of NCp7 for single-stranded sequences (*42*) and to the overall equivalence of A and C in the binding process, we propose that NCp7 may bind to the single-stranded regions of DNA with the following binding preference: $X_iTGX_j > X_iGXGX_j \approx X_iTXGX_j > X_iGX_j \gg X_iX_j$, where X correspond to either A or C residues. This rule is in full keeping with the negligible binding of (1−55)NCp7 to d[$A_4(TC)_2TA$] and the strong binding to d($GACT_2GTG_2$) and d(GCAGTGCAT) (*42*) or to sequences selected by the SELEX method (*44*, *45*). Moreover, in keeping with previous investigations on the interaction of NCp7 with the DNA and RNA forms of various oligonucleotide sequences (*22*, *42*), our data on selected sequences suggested that the conclusions drawn from DNA experiments were applicable to the interaction with RNA too, and thus, our preference rule probably applied for RNA as well. This observed sequence dependence was somewhat unexpected according to the histone-like function of NCp7. This suggests that either some structural features in the genomic RNA allow a sufficient binding constant to be reached even when the sequence is not favorable for binding or alternatively that protein−protein interactions stabilize to some extent the binding of NCp7 to low-affinity binding sites.

Oligonucleotide and Protein Determinants in the Binding of NCp7 to SL3. The analysis of the interaction of NCp7 with the DNA analogue of the SL3 ψ-RNA recognition

element confirmed the critical role of Trp37-G stacking for the proper positioning of NCp7 on its target nucleic acids. Similarly, Phe16−G interaction seems critical, too, and contributes to the same extent than Trp37 in the stability of NCp7−SL3. Accordingly, the deleterious effects of Phe16 and Trp37 mutations (substituted by the hydrophobic but nonaromatic Ala residues) on virus infectivity (*17*) may be partly related to the binding energy decrease that prevented the recognition of the ψ-RNA-specific binding sites.

The preference rule established for the binding of NCp7 to hexanucleotides was found to fully apply when the binding of NCp7 to various SL3 mutants was investigated. Moreover, the binding parameters of (12−53)NCp7 to SL3, SL3(m1), and SL3(m2) were strikingly similar to the binding parameters of (12−53)NCp7 or (1−72)NCp7 to hexanucleotides with sequences that are analogous to the SL3 tetraloop sequence. This strongly suggested that the central finger domain of NCp7 essentially interacts with the single-stranded tetraloop, in full agreement with NMR data (*23*). Moreover, the stem-loop structure of SL3 mutants does not seem to stabilize the binding of the central finger motif itself but favors the binding of the lateral N- and C-terminal domains, providing a high stabilization energy of $-6.9(\pm 0.7)$ kJ/mol for the five SL3 derivatives at 0.1 M NaCl. In keeping with the NMR data of (1−55)NCp7−SL3 (*23*) and our previous results on $tRNA^{Lys,3}$-(1−72)NCp7 (*29*), the stability brought about by the N- and C-terminal domains of NCp7 is essentially due to electrostatic interactions. A stacking interaction was also evidenced for Trp61 in a limited population of NCp7−SL3 complexes, but the marginal K (1 M) increase for NCp7 as compared to (12−53)NCp7 suggested that this stacking is not critical. Finally, the binding and fluorescence parameters did not evidence any strong affinity binding sites for NCp7 on the double-stranded stem of SL3 mutants in our conditions.

One of the consequences of the strong sequence dependence of NCp7 binding to nucleic acids is the large decrease of the number of potential binding sites as compared to the Mc Ghee and von Hippel binding model (*46*), where the nucleic acid is considered as a lattice of overlapping sites of identical affinities. This notably implies that no huge binding constant is required to get a selective binding to the limited number of potential binding sites in DNA. If this conclusion also applies to NCp7−RNA interactions, this suggests that the recognition of genomic RNA by NCp7 may be achieved by a limited number of sites with moderate affinity. This is in keeping with the modest binding constants that have been reported for the specific binding sites in the HIV-1 ψ encapsidation sequence (*8*, *45*). In this context, both the SL3 loop and r($C_2A_2U_5G$) sequences are likely candidates for preferential binding in the encapsidation process since both sequences are single stranded in the packaging signal of HIV-1 genomic RNA (*44*), contain either a favorable UG or GAG sequence, and bind to NCp7 with affinities that are close to that of the overall ψ packaging signal (*8*). The same conclusion held true for the loops of the three other stem-loops of the HIV-1 ψ sequence since all of them contain a GXG sequence. In contrast, due to its location in the double-stranded stem of SL2 (*8*), the ACGCC sequence is probably not a preferential binding site for NCp7. In fact, a recent model indicated that the secondary structure of the HIV-1 ψ sequence (*47*) is largely double-stranded, and thus, the preferential binding sites for NCp7 may be limited to the four SL loops, the UG-containing single-stranded sequence between SL2 and SL3 and the GXGXG-containing single-stranded sequence between SL3 and SL4. Furthermore, examination of the secondary structure model of the 350 nt packaging signal of MoMuLV (*48*) reveals that the majority of the loop sequences of this packaging signal do not contain either a UG, GXG, or TXG sequence. This may partly explain the absence of selective recognition of this packaging signal by NCp7 (*49*, *50*).

Taken together, our data provide new clues for the selective binding of NCp7 to its nucleic acid targets despite the absence of very high-affinity binding sites. However, to straightforwardly quantify the binding of NCp7 to the HIV-1 ψ sequence and, consequently, to give a valid model for the encapsidation process, efforts should be made to elaborate a new formalism that should take into account both sequence-dependent and overlapping binding sites. Similarly, efforts should also be made to explain how, despite the low-binding affinities of numerous nucleotide sequences, NCp7 could efficiently cover the genomic RNA. Both aspects are under current investigation.

ACKNOWLEDGMENT

We are grateful to J. L. Darlix for a critical reading of the manuscript, Patrice Petitjean for peptide synthesis, and T. Hyunh Dynh for RNA decanucleotide synthesis.

REFERENCES

1. Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., and Barré-Sinoussi, F. (1990) *J. Mol. Biol. 216*, 689−99.
2. Tanchou, V., Gabus, C., Rogemond, V., and Darlix, J. L. (1995) *J. Mol. Biol. 252*, 563−71.
3. Sheng, N., and Erickson-Viitanen, S. (1994) *J. Virol. 68*, 6207−14.
4. Barat, C., Lullien, V., Schatz, O., Keith, G., Nugeyre, M. T., Gruninger-Leitch, F., Barré-Sinoussi, F., LeGrice, S. F., and Darlix, J. L. (1989) *EMBO J. 8*, 3279−85.
5. Darlix, J. L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) *J. Mol. Biol. 254*, 523−37.
6. Berkowitz, R. D., Luban, J., and Goff, S. P. (1993) *J. Virol. 67*, 7190−200.
7. Berkowitz, R. D., and Goff, S. P. (1994) *Virology 202*, 233−46
8. Clever, J., Sassetti, C., and Parslow, T. G. (1995) *J. Virol. 69*, 2101−9.
9. Dib-Hajj, F., Khan, R., and Giedroc, D. P. (1993) *Protein Sci. 2*, 231−43.
10. Tsuchihashi, Z., and Brown, P. O. (1994) *J. Virol. 68*, 5863−70.
11. Berg, J. M. (1986) *Science 232*, 485−7.
12. Bess, J. W., Jr. Powell, P. J., Issaq, H. J., Schumack, L. J., Grimes, M. K., Henderson, L. E., and Arthur, L. O. (1992) *J. Virol. 66*, 840−7.
13. 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., II, Hare, D. R., and Arthur, L. O. (1992) *Protein Sci. 1*, 563−74.
14. Mély, Y., De Rocquigny, H., Morellet, N., Roques, B. P., and Gérard, D. (1996) *Biochemistry 35*, 5175−82.
15. Mély, Y., Jullian, N., Morellet, N., De Rocquigny, H., Dong, C. Z., Piemont, E., Roques, B. P., and Gérard, D. (1994) *Biochemistry 33*, 12085−91.
16. Mély, Y., de Rocquigny, H., Sorinas-Jimeno, M., Keith, G., Roques, B. P., Marquet, R., and Gérard, D. (1995) *J. Biol. Chem. 270*, 1650−6.
17. Dorfman, T., Luban, J., Goff, S. P., Haseltine, W. A., and Gottlinger, H. G. (1993) *J. Virol. 67*, 6159−69.

18. Aldovini, A., and Young, R. A. (1990) *J. Virol. 64*, 1920−6.
19. Gorelick, R. J., Nigida, S. M., Jr., Bess, J. W., Jr., Arthur, L. O., Henderson, L. E., and Rein, A. (1990) *J. Virol. 64*, 3207−11.
20. 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*, 11707−16.
21. Tanchou, V., Decimo, D., Pechoux, C., Lener, D., Rogemond, V., Berthoux, L., Ottmann, M., and Darlix, J. L. (1998) *J. Virol. 72*, 4442−7.
22. 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*, 419−34.
23. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) *Science 279*, 384−8.
24. De Rocquigny, H., Ficheux, D., Gabus, C., Fournié-Zaluski, M. C., Darlix, J. L., and Roques, B. P. (1991) *Biochem. Biophys. Res. Commun. 180*, 1010−8.
25. Vuilleumier, C., Maechling-Strasser, C., Gérard, D., and Mély, Y. (1997) *Anal. Biochem. 244*, 183−5.
26. Record, M. T., Jr., Lohman, M. L., and De Haseth, P. (1976) *J. Mol. Biol. 107*, 145−58.
27. Livesey, A. K., and Brochon, J. C. (1987) *Biophys. J. 52*, 693−706.
28. Mély, Y., Piemont, E., Sorinas-Jimeno, M., de Rocquigny, H., Jullian, N., Morellet, N., Roques, B. P., and Gérard, D. (1993) *Biophys. J. 65*, 1513−22.
29. Mély, Y., Bombarda, E., Vuilleumier, C., De Rocquigny, H., Roques, B. P., and Gérard, D. (1998) In *Fluorescent Microscopy and Fluorescent Probes* (Slavik, J., Ed.) Vol. 2, pp 181−6, Plenum Press, New York and London.
30. Lam, W. C., Maki, A. H., Casas-Finet, J. R., Erickson, J. W., Sowder, R. C., II, and Henderson, L. E. (1993) *FEBS Lett. 328*, 45−8.
31. Wu, J. Q., Ozarowski, A., Maki, A. H., Urbaneja, M. A., Henderson, L. E., and Casas-Finet, J. R. (1997) *Biochemistry 36*, 12506−18.
32. South, T. L., and Summers, M. F. (1993) *Protein Sci. 2*, 3−19.
33. Khan, R., and Giedroc, D. P. (1994) *J. Biol. Chem. 269* (36), 22538−46.
34. Urbaneja, M. A., Kane, B. P., Johnson D. G., Gorelick, R. J., Henderson, L. E., and Casas-Finet, J. R. (1999) *J. Mol. Biol. 287*, 59−75.
35. Montenay-Garestier, T., Toulmé, F., Fidy, J., Toulmé, J. J., Le Doan, T., and Hélène, C. (1983) In *Structure, dynamics interactions and evolution of biological macromolecules* (Hélène, C., Ed.) pp 113−28, Reidel, D. Publishing Company.
36. Bombarda, E., Ababou, A., Vuilleumier, C., Gérard, D., Roques, B. P., Piémont, E., and Mély, Y. (1999) *Biophys. J. 76* (3), 1561−70.
37. Stoylov, S. P., Vuilleumier, C., Stoylova, E., De Rocquigny, H., Roques, B. P., Gérard, D., and Mély, Y. (1997) *Biopolymers 41*, 301−12.
38. Kumar, N. V., and Govil, G. (1984) *Biopolymers 23*, 2009−24.
39. De Rocquigny, H., Delaunay, T., Petitjean, P., Fournié-Zaluski, M. C., and Roques, B. P. (1998) *Regard sur la Biochim. 2*, 44−50.
40. Remy, E., de Rocquigny, H., Petitjean, P., Muriaux, D., Theilleux, V., Paoletti, J., and Roques, B. P. (1998) *J. Biol. Chem. 273*, 4819−22.
41. Morellet, N., de Rocquigny, H., Mély, Y., Jullian, N., Déméné, H., Ottmann, M., Gérard, D., Darlix, J. L., Fournie-Zaluski, M. C., and Roques, B. P. (1994) *J. Mol. Biol. 235*, 287−301.
42. Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E. (1998) *J. Virol. 72*, 1902−9.
43. Wu, J. Q., Maki, A. H., Ozarowski, A., Urbaneja, M. A., Henderson, L. E., and Casas-Finet, J. R. (1997) *Biochemistry 36*, 6115−23.
44. Allen, P., Collins, B., Brown, D., Hostomsky, Z., and Gold, L. (1996) *Virology 225*, 306−15.
45. Berglund, J. A., Charpentier, B., and Rosbach, M. (1997) *Nucleic Acids Res. 25*, 1042−49.
46. McGhee, J. D., and Von Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469−89.
47. Pappalardo, L., Kerwood, D. J., Pelczer, I., and Borer, P. N. (1998) *J. Mol. Biol. 282*, 801−18.
48. Tounetki, N., Mougel, M., Roy, C., Marquet, R., Darlix, J. L., Paoletti, J., Ehresmann, B., and Ehresmann, C. (1992) *J. Mol. Biol. 223*, 205−20.
49. Berkowitz, R. D., Ohagen, A., Hoglund, S., and Goff, S. P. (1995) *J. Virol. 69*, 6445−56.
50. Zhang, Y., and Barklis, E. (1995) *J. Virol. 69*, 5716−22 [published erratum appears in *J. Virol.* (1997) *71*, 5712].

BI991145P