

Molecular Dynamics and DFT Study on HIV-1 Nucleocapsid Protein-7 in Complex with Viral Genome

Mattia Mori,[†] Ursula Dietrich,[‡] Fabrizio Manetti,[†] and Maurizio Botta^{*,†}

Dipartimento Farmaco Chimico Tecnologico, Università degli Studi di Siena, Via Alcide de Gasperi 2, I-53100 Siena, Italy and Georg-Speyer-Haus, Institute of Biomedical Research, Paul-Ehrlich-Straße 42-44, 60595 Frankfurt am Main, Germany

Received February 16, 2010

The HIV-1 nucleocapsid protein-7 (NCp7) is a highly basic, small zinc-binding protein involved in both deoxyribonucleic (DNA) and ribonucleic (RNA) acids annealing and in viral particle maturation including genome encapsidation, with an additional chaperoning activity toward reverse transcriptase by promoting the two obligatory strand transfers during reverse transcription. Because of its interaction with highly conserved sequences of the HIV-1 genome, NCp7 is being considered a new potential drug target, resistant to mutation, for antiviral activity. The high flexibility of this protein has, however, limited the identification of structural determinants involved in the interaction with stranded sequences of DNA and RNA. Here, we provide a quantum mechanics (density functional theory) study of the zinc-binding motifs and a molecular dynamics simulation of the protein in complex with RNA and DNA, starting from available nuclear magnetic resonance (NMR) structures. Results show that the interaction between the NCp7 and the viral genome is probably based on electrostatic interactions due to a cluster of basic residues, which is reinforced by the exploitation of nonelectrostatic contacts that further stabilize the complexes. Moreover, a possible mechanism for DNA destabilization that involves amino acids T24 and R26 is also hypothesized. Finally, a network of hydrophobic and hydrogen-bond interactions for the stabilization of complexes with DNA and, especially, with RNA is described here for the first time. The complexes between NCp7 and both DNA and RNA, resulting from computer simulations, showed structural properties that are in agreement with most of the currently available molecular biology evidence and could be considered as reliable models (better than NMR structures currently available) for subsequent structure-based ligand design approaches.

INTRODUCTION

Human immunodeficiency virus-1 (HIV-1) is a lentivirus belonging to the retrovirus family, responsible for the acquired immunodeficiency syndrome (AIDS), one of the most common human infections of the last century. In the past decade, significant steps forward have been made by the scientific community in understanding the molecular bases of the HIV-1 infection and the replication and transmission,¹ allowing for the identification of some key proteins, within the whole virus life cycle, to be selectively targeted for anti-HIV-1 activity. The most common approach was generally to hit the target protein with small molecules able to interfere with its pathological activity. In this field, inhibitors of HIV-1 reverse transcriptase (RT),^{2,3} protease,^{4,5} integrase,^{6–8} and fusion^{9–10} were identified and developed. However, it is well established that lentiviruses, in general, and especially HIV-1, have a high ability to mutate into resistant forms in response to antiviral therapies.¹¹ To overcome virus resistance toward currently used drugs, different targets and strategies need to be speculated.

In the last years, a particular attention was given to atypical highly conserved viral structures that could be resistant to

mutations.¹² In particular, RNA packaging signal ψ (PSI), transactivation-responsive region (TAR), and (\pm)primer binding site (PBS) sequence of DNA, and related protein pathways, are becoming attractive for drug discovery.

In this context, NCp7 is a highly basic, small zinc-binding protein that derives from the cleavage and processing of the HIV structural protein Gag.^{13,14} Literature reports many biological data that describe the mechanism of action of NCp7 and that support its key role in the HIV-1 replication process and life cycle. NCp7 does interact with specific stem loops of viral RNA in the packaging site¹⁵ (PSI, enabling the encapsidation of viral genome into new mature viral particles) as well as with the PBS¹⁶ and TAR sequences of viral DNA¹⁷ (enhancing the reverse transcription and, consequently, the viral maturation and replication process). In addition, NCp7 also possesses nucleic acid chaperone properties that guide RT activity to the synthesis of viral DNA, by operating the obligatory strand transfer necessary for the integration of viral DNA into the cellular genome.^{18–20} In any case, coordination of zinc ions is a crucial step for the exploitation of NCp7 biological activity. In fact, single-point mutations within the zinc fingers^{21,22} as well as the administration of zinc-ejector molecules^{23–25} enhance the *in vitro* and *in vivo* formation of noninfective viral particles.

Zinc fingers are flanked by a cluster of basic residues that play a key role in the recognition of nucleic acids and in the stabilization of their complexes with the protein.²⁶ In fact,

* Corresponding author. Telephone: +39 0577 234306. Fax: +39 0577 234333. E-mail: botta.maurizio@gmail.com.

[†] Università degli Studi di Siena.

[‡] Institute of Biomedical Research.

	*	*****	*	:	*****	*****	*****	*****	*****
1BJ6	-----N-----	VKCFNCGKEGH	TARNCRAPRKKG	CWKC	GKEGH	QMKDCTERQ	--		42
1ESK	-----N-----	VKCFNCGKEGH	TARNCRAPRKKG	CWKC	GKEGH	QMKDCTERQ	--		42
2EXF	-----N-----	VKCFNCGKEGH	TARNCRAPRKKG	CWKC	GKEGH	QMKDCTERQAN			44
2JZW	-----N-----	VKCFNCGKEGH	TARNCRAPRKKG	CWKC	GKEGH	QMKDCTERQAN			44
1A1T	MQKGNFRNQRK	TVKCFNCGKEGHIA	KNCRAPRKKG	CWKC	GKEGH	QMKDCTERQAN			55
1F6U	MQKGNFRNQRK	TVKCFNCGKEGHIA	KNCRAPRKKG	CWKC	GKEGH	QMKDCTERQAN			55
ruler	1.....10.....20.....30.....40.....50.....								



Figure 1. Sequence alignment of NCp7 alone (1ESK), of the complexes with DNA fragments (1BJ6, 1ESK, 2EXF, and 2JZW), and of the complexes with SL3 (1A1T) and SL2 (1F6U) RNA.

nuclear magnetic resonance (NMR) solution structures of the complexes between protein and both DNA and RNA fragments showed that the interaction between the nucleic acids and the protein does involve basic amino acids in close proximity of the zinc fingers, without contacting them directly. In agreement with this evidence, the deletion of these basic residues and the modification of the ionic strength of the medium significantly decrease the affinity of the protein toward nucleic acids.²⁷

Moreover, the proper coordination of zinc ions by the zinc fingers determines the arrangement of the protein in a particular conformation that exposes a hydrophobic core, centered on W37. This is a key residue involved in the stabilization of the bound viral genome, mainly through the formation of a stacking interaction with the central guanosine nucleotides of the single-stranded hairpin loops.²⁸ The importance of W37 (together with F16 that also contributes to the stabilization of the complex with RNA) was also previously demonstrated with a tryptophan-enriched hexapeptide that inhibits the interaction between the NCp7 and the viral genome by mimicking the three-dimensional (3D) structure and the chemical behavior of NCp7.²⁹ Accordingly, single-point mutations at W37 significantly affect the ability of NCp7 to bind viral genomes and, thus, its peculiar activity. Moreover, a biochemical-based high-throughput screening led to the identification of five compounds that interact with NCp7. The binding of these small molecules within the folded zinc fingers of the NCp7 (12–55) prevents the formation of the complex between the protein and the cTAR sequences of DNA, without promoting zinc ejection.³⁰ Knowledge of structural elements as well as biological evidence would allow a structure-based ligand design approach to target this protein. However, high flexibility of the loop region (residues 12–55), involved in the binding of viral genome, represents the major limitation to derive structural information from in vitro experiments.^{31,32} As a matter of fact, NMR solution structures of the complexes between the protein and both DNA and RNA, do not provide elements to account for many of the biological data currently available. In fact, with the only exception of W37, the role of F16 and T24 and of basic residues is not properly accounted for by these structural models. As recently reported,³² this problem is probably related to the intrinsic limitations of high-resolution NMR spectroscopy that, with respect to other techniques such as electron paramagnetic resonance (EPR), is not able to probe the local mobility of a protein, especially when it is lacking in a folded secondary structure, such as NCp7.

In the attempt to correlate the dynamic properties of NCp7 structure with the available biological data, unrestricted

molecular dynamics (MD) simulations, based on an implemented version of AMBER force field, were performed on the NMR-derived structures of NCp7–DNA and NCp7–RNA. In detail, the complexes between NCp7 (residues 1–55) and PSI-RNA stem loop 3 (Protein Data Bank³³ (PDB) entry 1A1T),³⁴ and between NCp7 (residues 12–55) and PBS-DNA (PDB entry 2JZW³⁵) were used for this study. However, such structures suffered from several limitations affecting the conformation of zinc fingers, the reciprocal adaptation of the side chains of residues at the binding interface and the electrostatic interactions. Moreover, a steric refinement was also required (in particular for the complex with RNA) to remove steric clashes that affected some key residues strongly involved in the interaction with nucleic acid.

Results of our calculations show that the conformation of NCp7 in complex with DNA is significantly different from that found in the complex with RNA. However, the nature of driving forces for complex stabilization is the same for binding DNA and RNA. In particular, while electrostatic interactions are known to play a key role in the recognition of the viral genome,^{32,36} MD simulations show that non-electrostatic contacts (mainly based on W37 and F16, along with their $\pi-\pi$ and $\pi-$ cation interactions) are also crucial for complex stabilization. Moreover, we are able to hypothesize, for the first time, a possible mechanism for DNA strand destabilization, in which T24 plays a central role, together with R26. In summary, modeled complexes between NCp7 and nucleic acid fragments show structural properties that are consistent with most of the experimental evidence currently available in the literature and could be used as a starting point for subsequent structure-based drug design approaches.

THEORETICAL CALCULATIONS

Quantum Mechanics (QM) Analysis of the Zn(His)₃(Cys)₃ System. Each of the two conserved zinc fingers of NCp7 is characterized by the Cys-x-x-Cys-x-x-x-x-His-x-x-x-x-Cys sequence, and each zinc ion is coordinated by the three Cys and the δ His (HID, a neutral species with a hydrogen at the δ position) residues of such a sequence.¹⁴ A QM-based parametrization of the zinc-binding motif was applied to the Zn(His)(Cys)₃ system isolated in a square-based pyramidal geometry from the best model of the NMR solution structures of NCp7–DNA (PDB entry: 2JZW). To reduce CPU usage and computational resources during QM calculations, carboxy and amino groups of Cys and His were removed, C α of Cys and C β of His were transformed into the carbon atom of a methyl group, and hydrogen atoms were added where necessary (Figure 3).

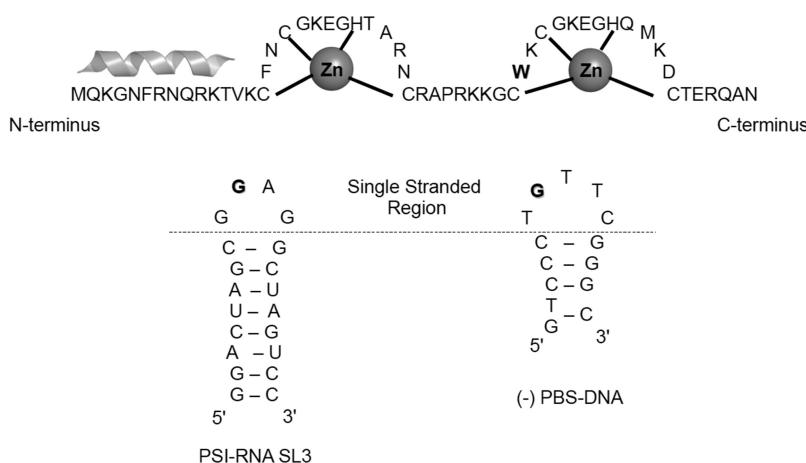


Figure 2. Top: Sequence of NCp7 (obtained from 1A1T by I24T and K26R mutations) showing the N-terminal helix portion and the two zinc fingers with the zinc-binding motifs. W37 is shown in bold text. Bottom: Sequence and schematic representation of SL3 PSI-RNA (left) and (−)PBS-DNA (right). Central guanosines (namely, RNA G10 and DNA G7, following a 5'-3' numbering) are shown in bold text.

The total charge (−1) of the Zn(His)(Cys)₃ system was derived from the protonation state of NMR solution structures of NCp7. The geometry of the Zn-coordinating system was first relaxed and optimized by QM calculations performed with Gaussian03 (G03) software.³⁷ Density functional theory (DFT) methods were applied for geometry optimization at the B3LYP level of theory with the 6-311++G(d,p) basis set. B3LYP/6-311++G(d,p) was also used to compute the restrained electrostatic potential (RESP) charges by a single point calculation^{38–40} and to perform the vibrational analysis of the optimized geometry to derive the force constants. RESP charges and force constants were, in turn, used to build parameter sets for nonstandard residues and to modify the AMBER force field with zinc-binding features, as further described below.

Parameterization of AMBER Force Field. Atomic point charges of the optimized Zn-coordinating geometry were derived by the electrostatic surface potential calculated by QM approaches. A two-step fitting procedure was applied to derive the RESP charges for each atom of the system with ANTECHAMBER⁴¹ in AMBER10. Charges were implemented in the description of new nonstandard residues (the zinc ion and the zinc-binding His and Cys) suitable for AMBER package.⁴² The AMBER force field was also implemented by the addition of the zinc-binding properties, such as equilibrium bond lengths and angles, and their relative force constants derived by means of the vibrational analysis. The van der Waals parameters for the zinc atom were taken from the literature ($\sigma = 1.10 \text{ \AA}$, $\epsilon = 0.0125 \text{ kcal/mol}$).⁴³ The quality of the modified force field was checked by the comparison of frequencies for normal modes calculated with a QM and a molecular mechanics (MM) approach. For this latter, a NMODE⁴⁴ analysis was performed with AMBER10.

Description of Reference Structures. PDB entries 1A1T and 2JZW that describe the structures of NCp7–RNA (5'-GGACUAGCGGAGGUCC-3') and NCp7–DNA (5'-GTCCCTGTTGGGC-3') complexes, respectively, determined by solution NMR spectroscopy, were selected from the PDB. 1A1T was selected because it is the sole solution structure of NCp7 (1–55 sequence) in complex with SL3 of PSI-RNA of the viral genome. SL3 is of particular interest because its primary sequence is highly conserved among

different strains of HIV-1 and because it is the major packaging signal for RNA genome encapsidation.⁴⁵ On the other hand, 2JZW (implemented from a previous structure, 2EXF) was selected because it is one of the complexes with no mutation at the zinc fingers bearing the entire (−)PBS-DNA sequence and because its structure should account for the role of NCp7 in the process of the (−)PBS during reverse transcription.

However, primary sequence of NCp7 slightly differs between these complexes. On the basis of such a structural difference, several residues of 1A1T were computationally replaced by the corresponding amino acids found in the 2JZW sequence and in most of the NCp7 structures deposited in the PDB (IDs 1BJ6, 1ESK, 2EXF, and 1F6U)^{31,35,46,47} (Figure 1), in such a way to perform simulations on identical NCp7 sequences and to obtain results comparable to each other. Accordingly, I24T and K26R mutations were applied at the 13–55 region, leading to the NCp7 sequence reported in Figure 2. Sequence analysis was performed with ClustalX software.⁴⁸

MD Simulations of NCp7–RNA and NCp7–DNA Complexes. Hydrogen atoms were added using the LEAP module of AMBER10. Each complex was solvated in a TIP3PBOX box system of explicit water molecules with a 15 Å buffer.^{49,50} Na⁺ counterions were added at the box limits, to neutralize the whole charge, without forcing initial ionic interactions at the surface of the complex (i.e., Na⁺–phosphate interactions).

The MD protocol consisted of a three-step energy minimization of the solvated complexes. First, water molecules were minimized [1000 steps steepest descendent (SD) and additional 2000 steps conjugate gradient (CG)], while the protein was kept frozen. Then, protein was minimized (2000 steps SD and 4000 CG), keeping the water molecules frozen. Finally, the whole system was minimized by means of 3000 steps SD and 6000 steps CG. A modified version of the AMBER03 force field,⁵¹ implemented with bond and angle terms for the zinc-binding motif, was used in the MD calculations together with forcefield99SB,⁵² that provides a more reliable parametrization for the simulation of nucleic acids.

MD time step was set to 1 fs, and the SHAKE algorithm⁵³ was used to treat hydrogen-containing bonds. Periodic

boundary conditions (PBC) were also applied. Complexes were first heated for 250 ps up to 300 K using the Langevin control of the temperature.^{44,54} Next, the density was equilibrated for further 300 ps before running the equilibration step of MD (10 ns at a constant pressure and 300 K temperature).

Energy minimization and MD were performed with SANDER. Statistic analysis for the calculation of the root-mean-square deviation (rmsd), the distances and the other geometrical features during the MD run were performed with PTTRAJ, whereas energy calculations were carried out by means of the MM-PBSA perl script.^{55,56} All these modules belong to the AMBER10 program.

It should be pointed out that in the early 1990s the sole NCp7 protein was submitted to a short MD simulation (10 ps) for the energy refinement of a NMR-based model.⁵⁷ The simulation was based on an AMBER force field implemented with parameters for zinc, derived from semiempirical calculations (MNDO approximation) on the Zn–water system.⁵⁸ A different approach was followed on a single point mutant of NCp7 in complex with DNA,⁵⁹ without reporting parameters for the zinc-coordinating motif. Despite the pioneeristic role of these computational studies in understanding the flexibility of NCp7, they could be considered as obsolete, and their results were not taken into account during our simulations.

Output Structures. Frames of the 10 ns MD simulation were used to generate the average structure of each complex (using the PTTRAJ module). Next, the frame with the lowest rmsd, with respect to the average structure, was taken as the reference output structure used for further calculations. Explicit water molecules and counterions were removed from the structure before performing a short energy minimization of the side chains (200 steps SD and 300 CG). Backbone atoms were restrained by a harmonic force constant of 50 kcal/mol·Å². The resulting complexes were, in turn, analyzed to identify structural determinants involved in the interaction between NCp7 and viral genome (3D coordinates of output structures are available upon request from the corresponding author).

Electrostatic Surface Calculation. Electrostatic surfaces for NCp7, RNA, and DNA and for their complexes were calculated with APBS 1.0.0,⁶⁰ using the default parameters for residues and nucleotides and manually introducing the QM-based parameters for the two zinc ions. Electrostatic surfaces were visualized with PyMol.⁶¹

RESULTS AND DISCUSSION

Considering that NCp7 activity depends from zinc ions properly coordinated within zinc-fingers, QM-based atomic charges and structural parameters were calculated for the Zn-coordination system and implemented in AMBER03 force field. The reliability of the implemented force field was then checked by comparison of the normal-mode frequencies calculated by QM and MM approaches.

Moreover, an analysis of MD results was performed to identify NCp7 structural determinants involved in the interactions with nucleic acids and to check whether these interactions are able to account for the large amount of biological data available in the literature. In particular, the role of electrostatic interactions and basic residues as well

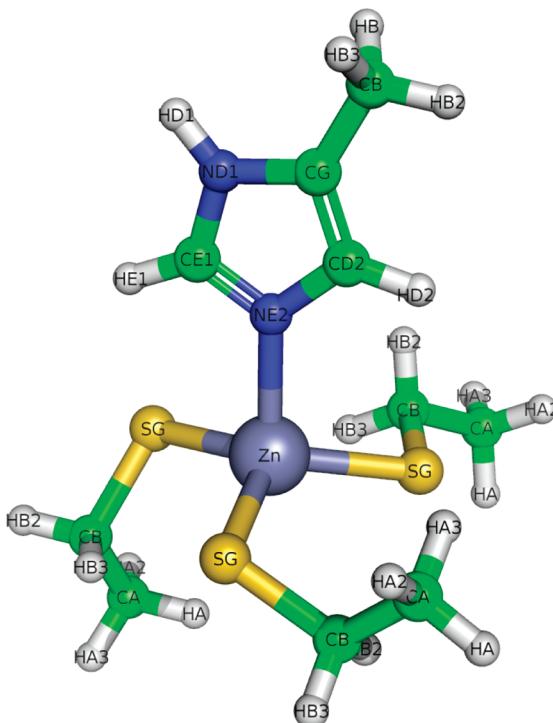


Figure 3. Graphic representation of the optimized zinc-binding motif simplified for QM calculations. Atoms are labeled with atom-type notations.

as key residues W37, F16, and T24 is provided, together with a hypothesis on the possible driving forces involved in the complex stabilization.

Finally, results of MD simulations showed that mutations made on the original NMR-based sequence (1A1T) to obtain the NCp7–RNA complex used in the simulations did not affect significantly the structural properties of the protein and the interactions with RNA. Hence, discussion on the NCp7–RNA complex is hereafter referred to the mutated form, unless declared otherwise.

Modified AMBER Force Field. RESP charges of the zinc-coordination system, calculated by means of the QM approach, were integrated in the parameter set for nonstandard residues, suitable for AMBER program. Although the symmetry option was disabled during the QM calculations, three symmetrical orbital systems, that correspond to the three zinc-binding cysteines, were identified by the software. A comparison between the calculated RESP charges and the point charges available in the AMBER03 force field for HID and Cys with a negative charge (CYM) showed that heavy atoms mostly affected by the positive charge of the zinc ion are the CB of Cys and the imidazole ring of His (Figure 3 and Table 1). The most significant variation was observed for the nitrogen that directly coordinated the zinc ion, whereas sulfur of Cys retained about the same charge as in the zinc-free, negatively charged form. The charge of zinc was reduced to 0.8693.

The projection of normal modes into internal coordinates provided equilibrium values for bonds and angles involved in the zinc-coordination system (Table 2) that were, in turn, implemented in the AMBER03 force field. Quality of the modified force field was then checked by comparison of the normal-mode frequencies calculated with QM and MM approaches on the isolated and simplified zinc-binding

Table 1. Comparison between Atomic RESP Charges for the Zinc Coordination System and the Charges Described in AMBER for the Corresponding Nonzinc-Binding Residues and for the Free Zn(II) Ion^a

atom type	zinc binding motif (RESP)	zinc free system (AMBER)
Zn	zinc ion 0.8693	2.0000
NE2	histidine δ H −0.2139	−0.6015
CD2	−0.2706	0.0437
HD2	0.1673	0.1102
CG	0.3415	−0.0015
CB	−0.4653	−0.1226
HB*	0.1267	0.0863
ND1	−0.4701	−0.2058
HD1	0.3581	0.3183
CE1	0.1963	0.1473
HE1	0.0827	0.1222
SG	cysteine with negative charge −0.7839	−0.8844
CB	0.4103	−0.2413
HB*	−0.0831	0.1122
CA	−0.1278	−0.0351
HA*	0.0030	0.0508

^a Atom types are represented in Figure 3.

Table 2. Force Constants for Angles and Bonds Involving the Two Zinc Ions

bond or angle	equilibrium value ^a	force constant ^b
Zn—NE2	2.163	48.0
Zn—SG	2.394	66.0
Zn—NE2—CD2	128.2	15.2
Zn—NE2—CE1	124.5	16.5
NE2—Zn—SG	101.0	19.2
SG—Zn—SG	116.5	10.8
Zn—SG—CB	102.3	34.7

^a Expressed as r_{eq} (Å) or θ_{eq} (°). ^b Expressed as kcal/mol·Å².

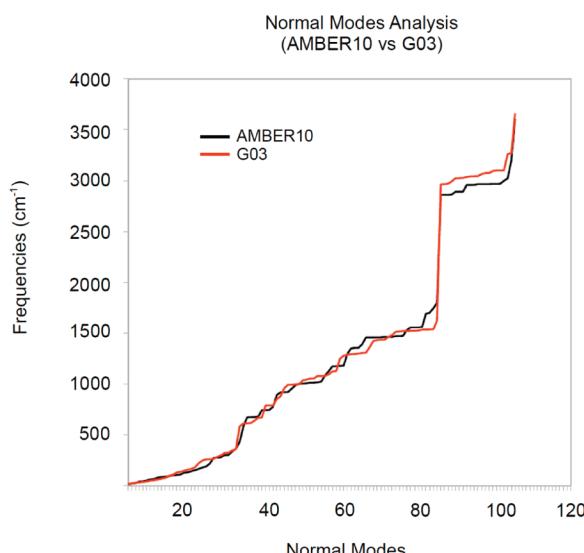


Figure 4. Comparison of the normal-mode frequencies calculated with AMBER (black line) and G03 (red line) for the simplified zinc finger system.

motif,⁶² finding a very good agreement between G03 and AMBER (Figure 4). Quality of the implemented force field was further confirmed by the fine correspondence of equi-

librium distances calculated for Zn—S and Zn—N bonds with those experimentally measured.^{57,63}

Preliminary Analysis of the NCp7 Conformations Bound to Viral Genome. Results of MD simulations showed that the binding conformation of NCp7 within the most stable complex with DNA was significantly different from the protein conformation found in the most stable complex with RNA, in agreement with that already reported in the literature.⁶⁴ This difference was, in general, attributed to the flexibility of NCp7 loops and to their high adaptability in response to small variations of the environmental physicochemical properties.

Moreover, backbone conformation of NCp7 in complex with DNA was very close to that found in the original NMR structure ($rms = 2.3$), whereas more pronounced changes were found for side-chain conformations, especially for several basic residues. The major difference affected the highly flexible C-terminal loop (D48–N55) that was not involved in binding nucleotides. A remarkable conformational change, with respect to the original NMR structure, was also observed in the single-stranded region of DNA because of the interactions between DT8 and the T24–R26 pair.

On the other hand, deviation of NCp7 backbone atoms of the complex with RNA in comparison to the starting NMR structure was significantly higher than that found in the NCp7–DNA complex ($rms = 3.9$). The C-terminal loop, the two zinc fingers, and the R29–K34 basic loop underwent the most significant conformational changes. The sole variation of the RNA conformation involved the two nucleotides at the 3' side of RG10 that are interfaced with the basic loop R29–K34. Nucleotides of the viral genome will be hereafter identified by R to specify the RNA nucleotides (i.e., RG10 stands for the guanosine nucleotide number 10 of the PSI-RNA sequence), while D will be used to list the DNA nucleotides (i.e., DG7 stands for the guanosine nucleotide number 7 of the (−)PBS-DNA sequence).

Energetic Aspects of the Interactions. Free energies of binding of NCp7–RNA and NCp7–DNA complexes were calculated by solving the Poisson–Boltzmann equation (PBE)⁶⁵ for trajectories of the MD run. A total of 200 snapshots were extracted around the frame with the lowest rmsd with respect to the average structure, for a total of 400 ps. Interaction energy between receptor (NCp7) and ligand (nucleic acids) was calculated by combining the MM energies with the continuum solvent approaches (MM-PBSA). Due to the different composition in nucleotides between PSI-RNA and (−)PBS-DNA, the free energy of interaction was then decomposed in pairwise residue contributions, to allow for a direct comparison of interaction energies of the two complexes. The interaction energy of the 15 most strongly interacting pairwise residues of both complexes, together with the description of the part of residues involved in the interaction (backbone or side chains) as well as the nature of the driving forces (electrostatic, hydrophobic-stacking, and hydrogen bond) are reported in Tables 3 and 4. Electrostatic interactions mainly contributed to the contacts that occurred at the binding interface. Moreover, also hydrophobic interactions between aromatic rings seemed to play a key role in both complexes, together with hydrogen bonds.

Protein residues involved in the interactions with nucleic acids were different in the two complexes, with the exception

Table 3. Free δ Energy of Binding Calculated for the 15 Residue Pairwise with the Strongest Interactions in the NCp7–RNA Complex

residues (protein– RNA)	part of residues involved (S = side chain, B = backbone)	energy (kcal·mol ⁻¹)	main contribution
K47–RG9	S–B	-10.18 ± 1.51	electrostatic
R26–RG12	S–B	-9.68 ± 1.48	electrostatic
M1–RA3	B–B	-9.16 ± 1.46	electrostatic
M1–RC4	B–B	-8.94 ± 1.04	electrostatic
Q45–RG10	S–S	-8.11 ± 0.92	hydrophobic (stacking)
N5–RG13	S–B	-8.03 ± 0.98	H-bonding
R10–RA6	S–B	-7.24 ± 1.47	electrostatic
K47–RG10	S–B	-7.09 ± 1.41	electrostatic
F16–RG12	S–S	-6.85 ± 0.56	hydrophobic (stacking)
W37–RG10	S–S	-6.80 ± 1.00	hydrophobic (stacking)
F16–RG10	B–S	-6.33 ± 0.57	H-bonding
M46–RG10	B–S	-6.02 ± 0.59	H-bonding
K33–RA11	B–S	-5.86 ± 1.39	H-bonding
G35–RG10	B–S	-5.38 ± 0.83	H-bonding
F16–RA11	B–S	-5.17 ± 0.50	H-bonding

Table 4. Free δ Energy of Binding Calculated for the 15 Residue Pairwise with the Strongest Interactions in the NCp7–DNA Complex

residues (protein– DNA)	part of residues involved (S = side chain, B = backbone)	energy (kcal·mol ⁻¹)	main contribution
R26–DT8	S–B/S	-12.05 ± 1.59	elec/H-bonding
N12–DC4	B–B	-11.54 ± 1.45	electrostatic
R26–DT9	S–B	-10.63 ± 1.50	electrostatic
W37–DG7	S–S	-8.95 ± 0.91	hydrophobic (stacking)
R32–DT8	S–B	-8.25 ± 1.36	electrostatic
N12–DC5	B–B	-7.61 ± 1.03	electrostatic
R29–DC10	S–B	-7.28 ± 1.54	electrostatic
R32–DG7	S–B	-6.79 ± 1.24	electrostatic
Q45–DG7	S–S	-6.78 ± 0.63	hydrophobic (stacking)
V13–DC4	B–B	-5.89 ± 1.03	H-bonding
G35–DG7	B–S	-5.44 ± 0.74	H-bonding
M46–DG7	B–S	-4.71 ± 0.45	H-bonding
A25–DT8	B–S	-4.65 ± 0.46	H-bonding
T24–DT8	S–S	-4.05 ± 0.73	H-bonding
C28–DT9	B–S	-3.91 ± 0.54	H-bonding

of W37, Q45, and M46 that constituted the binding site for a highly conserved central guanosine (namely, RG10 for the complex with RNA and DG7 for the complex with DNA), and R26.

Moreover, nucleotides at the 3' side of the central guanosine (namely, RA11 and RG12 of RNA and DT8, DT9, and DC10 of DNA, respectively; nucleotides were enumerated in the 5'-3' direction) gave a strong contribution to the complex stabilization (in particular, for their interaction with W37), as already reported in previous literature.^{15,34,63} Analysis of interactions between residues and nucleic acid bases showed a network of stacking contacts and hydrogen bonds in proximity of the binding interface with the stranded region, flanked by a set of electrostatic interactions that involved many nucleotides and basic residues (Figure 5). Moreover, it is worth noting that only a few of these

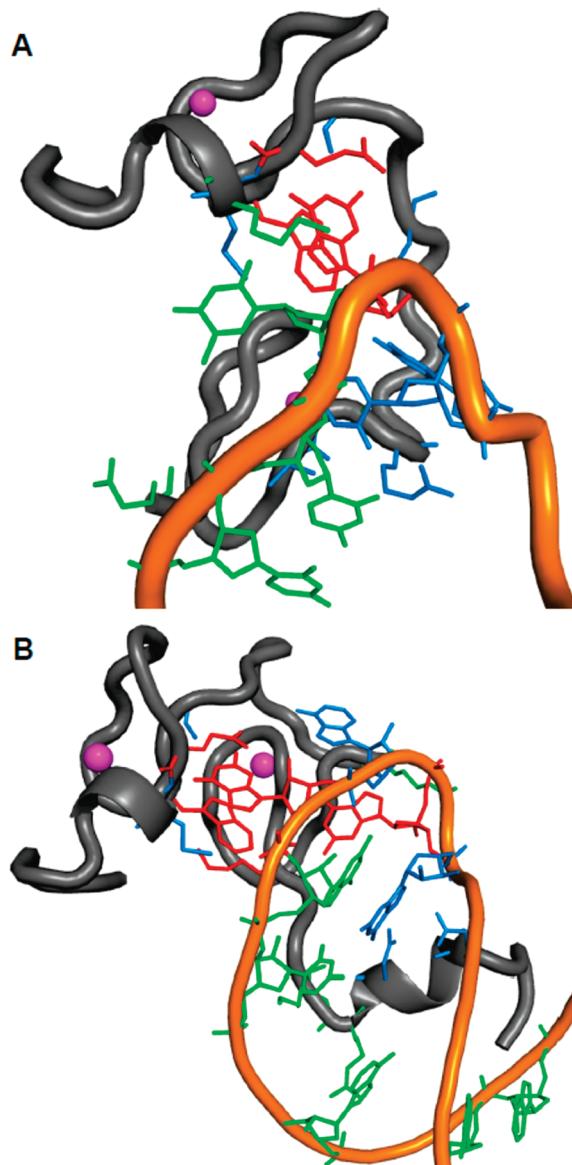


Figure 5. Graphic representation of amino acids and nucleic acid bases most involved in the energetic of (A) NCp7–DNA and (B) NCp7–RNA interactions. NCp7 is displayed as a dark-gray cartoon, while the 5'-3' nucleic acid chain is the orange cartoon. Sticks represent residues (amino acids or bases) involved in electrostatic, hydrophobic (stacking), and hydrogen-bond interactions (green, red, and blue, respectively).

interactions were described in the original NMR models of the NCp7–RNA complex, whereas contacts between the NMR conformation of NCp7 and DNA appeared very similar to the MD output.

In summary, while the 1–55 sequence of the full length NCp7 (72 residues)^{31,66} is fundamental for protein activity, analysis of pairwise interactions found with MD simulations suggested that, in addition to Q45 and M46, several residues of the 10–37 sequence are directly involved in the interactions with the single-stranded region of the viral genome, as also reported in the literature.^{15,67,68} Electrostatic interactions involved phosphate groups of double-stranded portions of nucleic acids, whereas nucleotides of the single-stranded region exploited principally hydrophobic interactions with aromatic rings of the protein, having the geometry and the distances of the stacking interactions.⁶⁹ A network of hydrogen bonds was located within the binding interface.

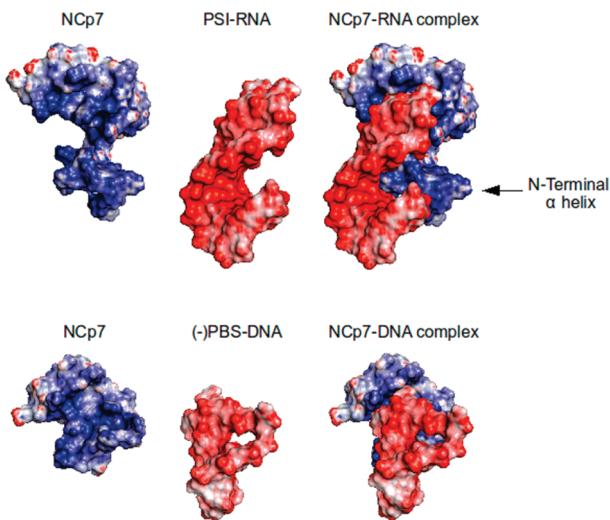


Figure 6. Top: (1–55) NCp7 (left) in its binding conformation with PSI-RNA (middle), and with the relative complex (right). Bottom: (12–55) NCp7 (left) in its binding conformation with (−)PBS-DNA (middle), and the respective complex (right). Electrostatic properties are color coded: positive charges are in blue; negative charges are in red; and neutral residues are in white. Color intensity is proportional to the charge value. All surfaces are calculated at the same salt concentration of the aqueous medium.

Relevance of Electrostatic Interactions. Electrostatic properties of the NCp7–RNA and NCp7–DNA complexes were widely discussed in the recent literature. The charge complementarity between the protein and the nucleic acids is generally considered the main driving force involved in the recognition of these macromolecules. Here, we evaluated the contribution of electrostatic interactions in the stabilization of the already formed complex.^{26,32,36,70} The PBE of the MD output structures was solved by means of APBS, and electrostatic surfaces were separately calculated for the protein and the nucleic acids in the binding conformation extracted from MD complexes. Despite the different conformation of protein residues and the different composition in nucleotides of the binding interface of complexes with DNA and RNA, no significant difference was found between the overall electrostatic surface of the protein in the two complexes (Figure 6). This result suggested that the strong electrostatic contribution to the binding was independent from the nature of interacting nucleic acids. In fact, interaction between the positively charged core of the protein (constituted by the residues between the two zinc fingers) and the viral genome involved the phosphate groups of nucleic acids (common to both DNA and RNA). As a result, electrostatic complementarity appeared as the driving force for the interactions between the NCp7 and the viral genome, especially at the binding interface, in addition to their role in the recognition step.²⁶ Moreover, electrostatic contribution was also independent from the type of amino acids. As an example, a comparison of electrostatic surfaces calculated for the original and the mutated NCp7 model structure led to no significant differences (data not shown).

MD results, showing the relevance of NCp7 electrostatic interactions, are in agreement with previous literature describing a fine analysis of the role of electrostatic contribution to the strength of NCp7 complexes^{71,72} and hypothesizing a possible correlation between the ionic strength and the tendency to complex aggregation. Similar conclusions emerged

from an experimental correlation between the dynamics of the protein and the ionic strength of the medium.³² In particular, ionic strength was proportionally correlated with the aggregation state of the complexes between NCp7 and nucleic acids, suggesting that the electrostatic-based aggregation of NCp7–RNA complexes with an additional RNA strand is possible at the 1–12 N-terminal helix (that is, the terminal helix of a complex interacts with the RNA strand of an additional complex). This hypothesis could be supported by our MD models showing that the positively charged N-terminal α -helix of the protein in the complex with RNA remained partially exposed to solvent, thus, providing a surface for complex aggregation (Figure 6). In fact, N-terminus was characterized by a highly basic surface that could represent a dimerization site for the binding, via electrostatic interactions, of the nucleic acid portion of another complex.

Relevance of Basic Residues. Basic residues can play a key role in biological activity of NCp7. As an example, they are fundamental especially for the recognition of nucleic acids, through interactions involving the positively charged side chains of Lys and Arg residues with the negatively charged phosphate groups of nucleic acids.¹⁵ Analysis of MD results revealed that they could also be involved in the stabilization of the complexes already formed through nonelectrostatic interactions. Here, we report evidence for both electrostatic and nonelectrostatic interactions made by basic residues of modeled complexes with DNA and RNA.

R26, R29, and R32 are the basic residues of NCp7 most strongly involved in the interaction with DNA (Table 4). Analysis of distances between these residues and the neighboring phosphate groups of DNA during the time of MD simulation showed that the Arg side chains and the phosphate groups perform electrostatic interactions stable in time. Differently, most of the original NMR structure showed the sole electrostatic interaction between R32 and phosphate groups of both DG7 and DT8 nucleotides. These interactions, that become stable during the first half of simulation (Figure 7), are crucial for the recognition and the stabilization of complexes between NCp7 and DNA.⁷³ In addition to electrostatic interactions, basic residues are also able to make π –cation interactions. As an example, the charged terminal group of R32 made a contact with the aromatic ring of DT9 that became stable after about 1 ns MD simulation and was not found in the original NMR structure.

Moreover, R26 is considered one of the key residues contributing to the destabilizing activity of NCp7 toward the stranded region of DNA.⁷⁴ In particular, the R26 guanidine moiety, upon anchoring to the DNA strand through the stable electrostatic interaction with neighboring DT9, exploited a hydrogen bond with the carbonyl oxygen at position 2 (O2) of the DT8 pyrimidine ring. This interaction guided the formation of the additional hydrogen bond between DT8 (NH at position 3 of the pyrimidine ring) and T24 side chain, thus, creating a network of hydrogen bonds resembling what is usually found in the interactions of a nucleic acid purine and a pyrimidine rings (Figure 9).

On the other hand, R26, R10, K11, and K47 were responsible of the most relevant electrostatic interactions of the NCp7–RNA complex, in agreement with previous experimental data.¹⁵ R26 and R10 showed their terminal basic group close enough to the phosphate of RG12 and RA6,

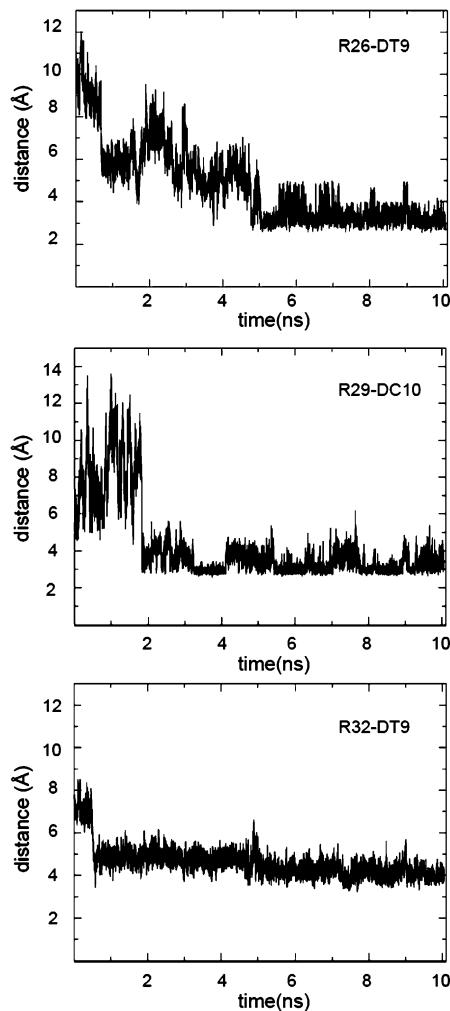


Figure 7. Time evolution of the distance between R26 and DT9 (top), R29 and DC10 (middle), and R32 and DT9 (bottom). Distances are calculated between the mass centers of the Arg—guanidine moiety and the phosphate group (first and second case) or between the guanidine moiety of Arg and the pyrimidine aromatic ring of DT9 (third case).

respectively, to make a stable electrostatic interaction. The original NMR structure of the complex, where R26 was replaced by K26, showed a similar interaction involving the ammonium group, thus, suggesting that a basic residue at that position could have an important role for the protein in recruiting RNA.

K47 and the phosphate groups of RG9 and RG10 exploited strong and stable electrostatic interactions that were not described in the starting NMR structures. The time profile of the distance between the protonated K47 Ne and the mass center of RG9 and RG10 phosphate groups showed that this electrostatic interaction became stable after about 3 ns MD simulations (Figure 8). Moreover, as observed for the complex with DNA, basic residues can perform nonelectrostatic intermolecular and intramolecular interactions that participate to complex stabilization. As an example, K14 played a key role in the interaction with RNA because of the intramolecular π —cation contact with the aromatic side chain of W37 (stable after 3 ns MD) that acted as a constraint avoiding W37 side-chain rotations around CA—CB or CB—CG bonds and promoting the formation of a stable stacking interaction between the aromatic rings of W37 and RG10. This result was in agreement with literature reporting

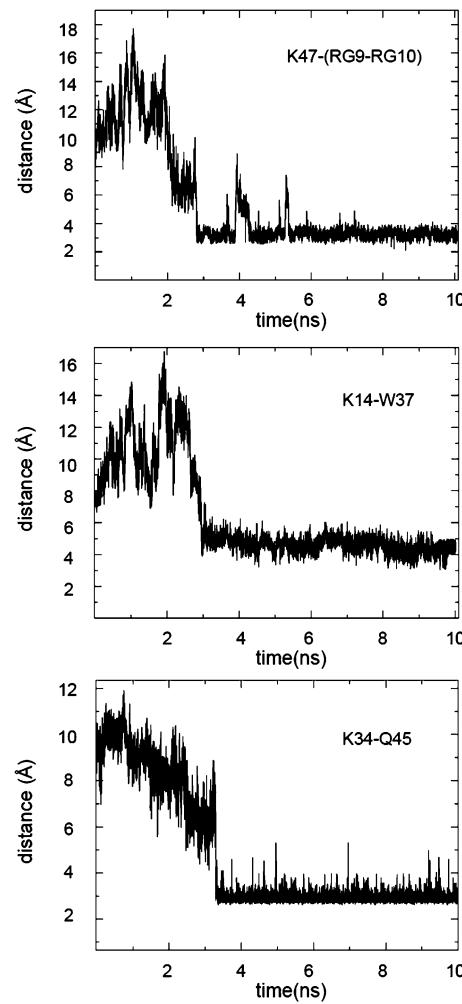


Figure 8. Time evolution of the distance between the intermolecular pairs K47—RG9 and K47—RG10 (top) and between the intramolecular pairs K14—W37 (middle) and K34—Q45 (bottom). Distances are calculated between the mass centers of the Arg—guanidine moiety and the phosphate groups of G9 and G10 (first case), between the K14 Ne and the aromatic ring of W37 (second case), and between the Q45 NE2 and K34 carbonyl group.

W37 as the most important NCp7 residue for the interaction with viral genome. On the contrary, the K14 side chain of all the original NMR structures showed a completely random disposition without any favorable interaction with the surrounding environment, suggesting a low resolution of structural details.

The side chain of Q45 is involved in a hydrogen bond with the carbonyl group of K34 and in the network of stacking interactions, also involving RG10, W37 and K14.

Finally, the role of K33 should be also mentioned for its intermolecular interaction with RA11 (together with R32) and for intramolecular hydrogen bonds with the N17 side chain crucial to stabilize the stranded region of bound RNA and to maintain the geometry of the flanked N-terminal zinc finger. Literature reported that these residues should not be mutated to avoid a decrease of the viral activity of NCp7.⁷⁵ Accordingly, single-point mutations of specific basic residues led to a decrease of the NCp7 binding affinity for RNA.⁷⁶ These observations were supported by MD simulations showing that the basic residues of NCp7, and particularly those of flanking zinc fingers, were crucial for the recognition of PSI-RNA and for the stabilization of its complex.

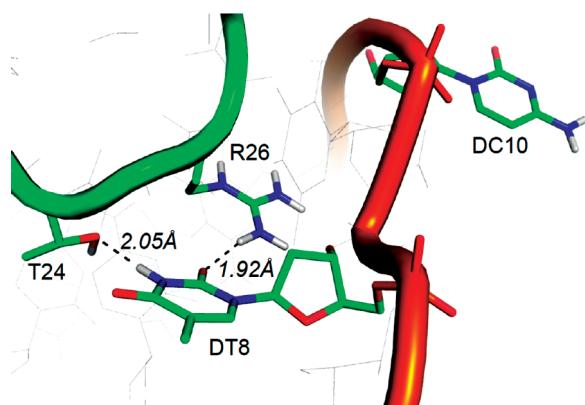


Figure 9. Stick representation of T24 and R26 in contact with the pyrimidine ring of DT8. DC10 is projected toward the solvent, thus, favoring the NCp7 viral activity of DNA destabilization. Hydrogen bonds are showed as dotted lines.

Moreover, the MD simulations also suggested that the most important factor for NCp7 activity could be the high number of basic residues rather than the amino acid sequence. This hypothesis was confirmed by the fact that incorporation of basic residues in NCp7-mimic hexapeptides increased the affinity toward RNA,^{77,78} independently from the position of this amino acid in the sequence.

In summary, in addition to the contribution of basic residues to the electrostatic interactions, MD simulations also provided evidence for nonelectrostatic contacts of the same basic residues involved in inter- and intramolecular hydrogen bonds and in π -cation interactions that strongly contributed to the stabilization of the already formed complexes.

Role of W37. W37 is reported to play a central role in the interaction between the NCp7 and the viral genome by interacting with a guanosine nucleotide of the stranded region, by hydrogen bond, and by hydrophobic and stacking interactions. W37 is directly involved in the annealing and dimerization of viral genomes, whereas its Ala mutant negatively affects the maturation of a dimeric partial transcript, resulting in a less efficient viral replication into mature particles.⁷⁹ Moreover, additional works describe the central role of a Trp residue of NCp7 in the interaction with RNA and DNA. As an example, a library of Trp-enriched hexapeptides was reported to inhibit the NCp7-induced strand destabilization by interacting, in a nonelectrostatic manner, with stranded sequences of the viral genome.²⁹ The driving force for this interaction was attributed to the π -stacking activity of a Trp of the peptide that mimics the NCp7 W37.⁶³ All these considerations are consistent with the NMR structures of complexes between the NCp7 and the viral genome. In fact, the hydrophobic interaction between aromatic rings, probably of stacking nature, is also found in NMR models and is conserved during the MD simulation. The sole exception was the complex with RNA, where a mutual adaptation of W37 and RG10 side chains that occurs during a MD run to avoid steric clashes observed in all NMR-based models. Finally, W37 also binds the viral genome through a network of hydrogen bonds and through π -cation and additional stacking interactions, which are better described below.

Role of F16. Together with W37, F16 is one of the most important residues described for the interaction of the NCp7 to the viral genome. However, in contrast to that observed

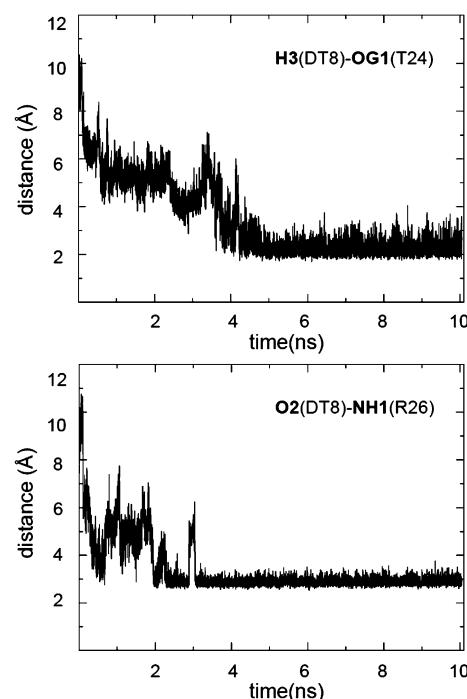


Figure 10. Time evolution of hydrogen bonds made by T24 and R26 with DT8.

for W37, the mutation of F16 to Ala does not affect the binding affinity of the NCp7 to the viral genome. MD simulations show that F16 adopts a different orientation, depending from the type of nucleic acid involved in the interaction, with the protein. In particular, the major difference observed concerns the reciprocal orientation of F16 and W37 in the complexes with DNA and RNA.

F16 is not directly involved in interactions with nucleotides of the (-)PBS-DNA. In fact, only a partial π -stacking interaction (T-shaped) with the aromatic ring of DT6 could be observed in the range between 3.0 and 6.5 ns of MD simulation. However, according to a NMR-based model,¹⁴ the main role of F16 suggested by the MD results is to create, together with V13, T24, A25, W37, and Q45, the high-affinity hydrophobic core for the binding of DNA at (-)PBS.³⁵

On the contrary, F16 is involved in the direct binding of nucleotides of the stranded region of Psi-RNA. MD simulations show a well-defined binding site for F16 within the nucleotides flanking the central guanosine RG10 in the stranded region (namely, RG9 and RG12, Figure 12A). Moreover, F16 does make a series of hydrogen-bond interactions (Figure 12B) that strongly contribute to the complex stabilization. These networks of favorable interactions, that also involve additional NCp7 residues, are not present in the starting NMR structure and in the complex with DNA. Interestingly, the distance between centroids of aromatic side chains of F16 and W37 was measured for NCp7 in the mean binding conformation with RNA. The value of 5.9 Å was in very good agreement with the value of about 6 Å experimentally measured by a previous NMR spectroscopy study⁵⁷ and well accounted for the NOE effect observed between Phe16 and Trp37 side-chain protons. Also in this case, agreement between the NMR-based observations and the MD model further supported the reliability of the model itself.

In conclusion, F16 seems to be able to induce the preferred protein conformation suitable for the binding to a particular type of nucleic acid. In particular, when NCp7 binds (−)PBS-DNA, F16 does make mainly aspecific hydrophobic interaction toward bases of the stranded region. Differently, in the complex between NCp7 and Psi-RNA, F16 exploits a series of π -stacking and hydrogen-bond contacts with bases of the stranded region that largely contribute to the complex stabilization and that renders F16 as a key residue for protein–nucleic acid interactions.

Role of T24. In the current literature, T24 is often described, together with W37 and F16, as a key residue for the interaction between the NCp7 and the viral genome. However, its role seems to be more relevant for protein–nucleic acid recognition rather than for direct interaction and stabilization of the complex.^{15,31,80} In fact, MD simulations for the complex with RNA show that T24 does not exploit a direct role in stabilizing nucleotides. Only an unstable hydrogen bond with RG12 is observed between 1.2 and 3.0 ns. As a consequence, a specific role of T24 in binding RNA cannot be identified by MD simulations, suggesting the hypothesis that this residue probably plays a role in the recognition phase rather than in the stabilization of the already formed complex. In fact, if T24 has a crucial role for the stabilization of the complex, then NMR structures (where T24 is mutated into Ile) should be characterized by a low definition of the region close to I24 because of the lack of T24 and its interactions. On the contrary, all the original NMR models show a well-defined folding of that region, without any specific binding interaction between I24 and RNA. This evidence suggests that I24 does not contribute to the stabilization of complexes, in agreement with that already found for T24 in MD models.

On the contrary, results of MD simulations carried out for NCp7–DNA complex show a significantly different situation. In fact, T24 participates to direct interactions with nucleotides. In particular, together with R26, it contacts DT8 with a network of hydrogen bonds similar to that found in purine–pyrimidine interactions (Figure 9). Orientation of R26 also forces DC10 to point toward the solvent (the stability of interactions between DT8 and both T24 and R26 during MD simulations is reported in Figure 10). As a consequence of these interactions, DNA probably undergoes a strand destabilization, thus, inducing the NCp7 viral activity of strand transfer.

This possible mechanism of action is proposed for the first time in this work and could account for the high relevance given to T24 in NCp7 annealing and in DNA strand transfer.^{15,31,80} On the contrary, interactions involving T24 are not found in the original NMR structure because T24, R26, DT8, and DC10 show a high variability of their reciprocal orientation. As a consequence, these residues do not make stable interactions that could be common to more than one NMR model.

Networks of Nonelectrostatic Interactions. It is well-known that electrostatic contributions represent the most important driving force for the formation of NCp7–viral genome complexes and take place principally during the recognition phase. Moreover, the major interactions for the stabilization of complexes already formed are represented by nonelectrostatic contributions, mainly involving residues at the binding interface. These different networks of interac-

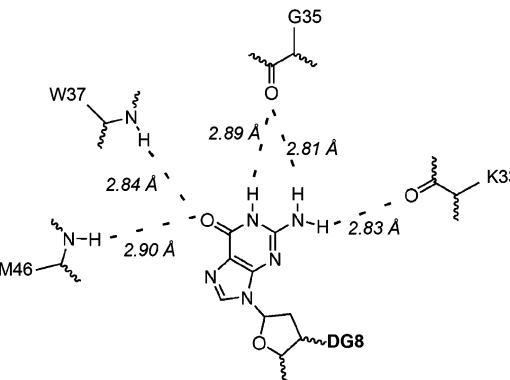


Figure 11. Schematic representation of the binding site for the central guanosine DG7 of the (−)PBS-DNA. Hydrogen bonds are displayed as dashed blue lines.

tions were studied during MD simulations. The most relevant nonelectrostatic interactions of the NCp7–DNA complex are located around the central guanosine DG7, where W37 and Q45 are involved in a small network of hydrophobic and stacking interactions. In addition, a network of hydrogen bonds further constraints DG7, its carbonyl oxygen at position 6 (O6) being the hydrogen-bond acceptor for the backbone NH of both M46 and W37 (Figure 11). Analysis of MD simulations reveals that the contact with M46 NH is present during 83% of simulation, while the interaction with W37 NH is found for about 24% of simulation time. This result is probably due to a slight rotation of the W37 aromatic ring around CB–CG bond to optimize the interaction with the aromatic ring of DG7, probably of a stacking nature, negatively affecting the orientation and the distance between O6 and W37 for a proper hydrogen-bond interaction. Most of the NMR models are only in partial agreement with that because the interaction between O6 and W37 NH seems to be preferred to that with M46 NH. Moreover, NH at position 1 and 2 of the purine ring system of DG7 acts as a hydrogen-bond donor toward the backbone carbonyl of G35 (interactions are found for 42 and 89% of MD simulation time, respectively). On the other hand, NH at position 2 is also involved in a stable hydrogen bond with the backbone carbonyl of K33 (62% of MD simulation time).

The side chain of Q45 is also involved in this network of interactions by two intramolecular hydrogen bonds with the backbone atoms of K34 and with the zinc-binding histidine H44 of the C-terminal zinc finger. The hydrogen bond with K34 is found for 62% of the whole MD simulation, whereas interaction with H44 shows a 92% of permanence during the whole MD simulation. Both of these interactions are not present in the starting NMR structure.

Analysis of these interactions clearly shows that hydrogen bonds and hydrophobic interactions between aromatic rings represent the driving forces for the stabilization of the complex between NCp7 and stranded DNA. A unique and well-defined binding pocket for the crucial guanosine DG7 can be described in the region within the two zinc fingers of NCp7 (Figure 11).

In the NCp7–RNA complex (Figure 12), RG10 interacts with W37 in a π – π stacking conformation also extended to Q45 (as in the protein–DNA complex). Moreover, K14 contacts the aromatic ring of W37 in a probable π –cation interaction, further stabilizing the K14–W37–RG10–Q45 system. A key role also emerges for F16, that was not

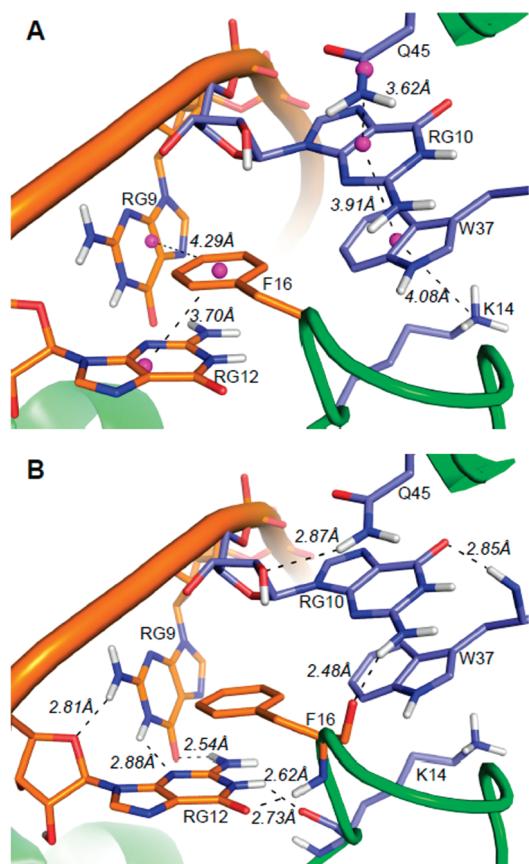


Figure 12. (A) Residues at the binding interface involved in hydrophobic interactions of a stacking nature. Carbon atoms of the system centered on F16 are displayed in orange, and those of the system focused on W37 are displayed in blue. Magenta spheres represent the centroids of each π system. Distance between the centroids and the parallelism between the π systems are in agreement with the description of the π -stacking interaction. (B) Hydrogen-bond interactions that involve the residues of the stacking systems in the NCp7–RNA complex.

displayed by the starting NMR structure. In fact, F16 constitutes the core of an additional stacking system involving RG9 and RG12, with which F16 interacts in a T-shaped and sandwich-like conformation, respectively.

These two interaction patterns, proposed for the first time in this work, are well folded and connected by a network of hydrogen bonds (Figure 12B). The carbonyl group of F16 makes a hydrogen bond with the NH at position 2 of the purine ring system of RG10, whereas the carbonyl of K14 interacts with the NH at position 1 of the RG12 purine ring. Such interactions are detectable starting from about 2.5 ns of MD simulation and are represented for 59 and 52% of the total time, respectively. In addition, NH of F16 performs also a highly stable hydrogen bond with the carbonyl oxygen at position 6 (O6) of RG12. Thus, K14 and F16 can be considered the bridge residues between these two systems, contributing to the extension of the nonelectrostatic network of interactions along the binding interface.

Despite the significant conformational differences observed for NCp7 in complex with DNA and RNA, the binding sites for the central guanosine nucleotides (namely, DG7 for DNA and RG10 for RNA) are very similar. In fact, several residues were observed to be equally relevant for both RNA and DNA complex stabilization during the energetic analysis of the interactions (see Tables 3 and 4). As in the case of the

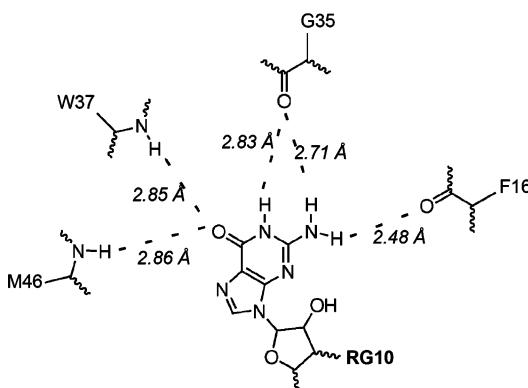


Figure 13. Schematic representation of the binding site for the central guanosine RG10 of the PSI-RNA hydrogen bonds are displayed as dashed blue lines.

NCp7–DNA complex, hydrogen bonds between RG10 O6 and NH of both W37 and M46 (Figure 13) seem to be equivalent in all NMR models, whereas MD simulations show the interaction with M46 for 67% of simulation time and the interaction with W37 for 39% of time. Moreover, RG10 does also make interactions with backbone carbonyl groups of G35 (present for 62 and 70% of the simulation time) and with the carbonyl group of F16 (54% of simulation time).

Several of the interactions involving F16 were already identified by experimental works,⁶³ whereas the remaining contacts are described for the first time in this work, providing a significant contribution to understand the binding mechanism of the NCp7 to the viral genome and to identify differences between the binding of RNA and DNA.

CONCLUSIONS

Unrestrained MD simulations with an implemented version of AMBER force field were performed for the first time on two available NMR structures of NCp7 in complex with viral DNA and RNA. Results provided a strong contribution to the understanding of structural determinants involved in the interactions between this highly flexible protein and nucleic acids. Significant conformational differences were found for the protein in complex with RNA and DNA, although the same amino acids strongly contributed to the interaction with the central guanosine (RG10 of RNA and DG7 of DNA, respectively). MD simulations showed that electrostatic contributions, known to guide NCp7 and nucleic acids during the recognition step, are also fundamental for complex stabilization, mainly involving basic residues. Moreover, nonelectrostatic interactions, such as hydrogen bonds, π – π , and π –cation contacts, were also found to strongly stabilize complexes. Finally, agreement between the structural properties of complexes (in terms of residues and interactions between them) and the experimental data available in the literature supported the reliability of the computational models. Such models, significantly better than those of NMR-based structures, could be considered as useful tools to be implemented in computational protocol for structure-based ligand design and for virtual screening, aimed at facilitating the drug discovery process.

ACKNOWLEDGMENT

We acknowledge the financial support provided by the European Union for the Collaborative Project HEALTH-

2007-2.3.2-1 “Targeting HIV integration co-factors, targeting cellular proteins during nuclear import or integration of HIV (THINC”).

REFERENCES AND NOTES

- (1) Gallo, R. C.; Montagnier, L. The discovery of HIV as the cause of AIDS. *N. Engl. J. Med.* **2003**, *349*, 2283–2285.
- (2) Radi, M.; Maga, G.; Alongi, M.; Angeli, L.; Samuele, A.; Zanol, S.; Bellucci, L.; Tafi, A.; Casaluce, G.; Giorgi, G.; Armand-Ugon, M.; Gonzalez, E.; Este, J. A.; Baltzinger, M.; Bec, G.; Dumas, P.; Ennifar, E.; Botta, M. Discovery of chiral cyclopropyl dihydro-alkylthio-benzyl-oxopyrimidine (S-DABO) derivatives as potent HIV-1 reverse transcriptase inhibitors with high activity against clinically relevant mutants. *J. Med. Chem.* **2009**, *52*, 840–851.
- (3) Srivastava, S.; Sluis-Cremer, N.; Tachedjian, G. Dimerization of human immunodeficiency virus type 1 reverse transcriptase as an antiviral target. *Curr. Pharm. Des.* **2006**, *12*, 1879–1894.
- (4) Anderson, J.; Schiffer, C.; Lee, S. K.; Swanstrom, R. Viral protease inhibitors. *Handb. Exp. Pharmacol.* **2009**, *189*, 85–110.
- (5) McKeage, K.; Perry, C. M.; Keam, S. J. Darunavir: a review of its use in the management of HIV infection in adults. *Drugs* **2009**, *69*, 477–503.
- (6) Emery, S.; Winston, A. Raltegravir: a new choice in HIV and new chances for research. *Lancet* **2009**, *374*, 764–766.
- (7) Mugnaini, C.; Rajamaki, S.; Tintori, C.; Corelli, F.; Massa, S.; Witvrouw, M.; Debysier, Z.; Veljkovic, V.; Botta, M. Toward novel HIV-1 integrase binding inhibitors: molecular modeling, synthesis, and biological studies. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5370–5373.
- (8) Rajamaki, S.; Innitzer, A.; Falciani, C.; Tintori, C.; Christ, F.; Witvrouw, M.; Debysier, Z.; Massa, S.; Botta, M. Exploration of novel thiobarbituric acid-, rhodanine- and thiohydantoin-based HIV-1 integrase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3615–3618.
- (9) Caporuscio, F.; Tafi, A.; González, E.; Manetti, F.; Esté, J. A.; Botta, M. A dynamic target-based pharmacophoric model mapping the CD4 binding site on HIV-1 gp120 to identify new inhibitors of gp120-CD4 protein-protein interactions. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6087–6091.
- (10) Lazzarin, A. Enfuvirtide: the first HIV fusion inhibitor. *Expert Opin. Pharmacother.* **2005**, *6*, 453–464.
- (11) Johnson, V. A.; Brun-Vezinet, F.; Clotet, B.; Gunthard, H. F.; Kuritzkes, D. R.; Pillay, D.; Schapiro, J. M.; Richman, D. D. Update of the Drug Resistance Mutations in HIV-1. *Top. HIV Med.* **2008**, *16*, 138–145.
- (12) Ikuta, K.; Kameoka, M.; Luftig, R. B. AIDS pathogenesis: the role of accessory gene mutations, leading to formation of long-lived persistently infected cells and/or apoptosis-inducing HIV-1 particles. *Virus Res.* **1997**, *52*, 145–156.
- (13) Druilennec, S.; Roques, B. P. HIV-1 NCp7 as a target for the design of novel antiviral agents. *Drug News Perspect.* **2000**, *13*, 337–349.
- (14) Roques, B. P.; Morellet, N.; de, R. H.; Demene, H.; Schueler, W.; Julian, N. Structure, biological functions and inhibition of the HIV-1 proteins Vpr and NCp7. *Biochimie* **1997**, *79*, 673–680.
- (15) Dannull, J.; Surovoy, A.; Jung, G.; Moelling, K. Specific binding of HIV-1 nucleocapsid protein to Psi RNA in vitro requires N-terminal zinc finger and flanking basic amino acid residues. *EMBO J.* **1994**, *13*, 1525–1533.
- (16) Barat, C.; Schatz, O.; Le, G. S.; Darlix, J. L. Analysis of the interactions of HIV-1 replication primer tRNA(Lys,3) with nucleocapsid protein and reverse transcriptase. *J. Mol. Biol.* **1993**, *231*, 185–190.
- (17) Azoulay, J.; Clamme, J. P.; Darlix, J. L.; Roques, B. P.; Mely, Y. Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J. Mol. Biol.* **2003**, *326*, 691–700.
- (18) Bampi, C.; Jacquinet, S.; Lener, D.; Decimo, D.; Darlix, J. L. The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 1668–1686.
- (19) Gabus, C.; Auxilien, S.; Pechoux, C.; Dormont, D.; Swietnicki, W.; Morillas, M.; Surewicz, W.; Nandi, P.; Darlix, J. L. The prion protein has DNA strand transfer properties similar to retroviral nucleocapsid protein. *J. Mol. Biol.* **2001**, *307*, 1011–1021.
- (20) Gabus, C.; Derrington, E.; Leblanc, P.; Chnaiderman, J.; Dormont, D.; Swietnicki, W.; Morillas, M.; Surewicz, W. K.; Marc, D.; Nandi, P.; Darlix, J. L. The prion protein has RNA binding and chaperoning properties characteristic of nucleocapsid protein NCp7 of HIV-1. *J. Biol. Chem.* **2001**, *276*, 19301–19309.
- (21) Bombarda, E.; Cherradi, H.; Morellet, N.; Roques, B. P.; Mely, Y. Zn(2+) binding properties of single-point mutants of the C-terminal zinc finger of the HIV-1 nucleocapsid protein: evidence of a critical role of cysteine 49 in Zn(2+) dissociation. *Biochemistry* **2002**, *41*, 4312–4320.
- (22) Bombarda, E.; Roques, B. P.; Mely, Y.; Grell, E. Mechanism of zinc coordination by point-mutated structures of the distal CCHC binding motif of the HIV-1 NCp7 protein. *Biochemistry* **2005**, *44*, 7315–7325.
- (23) Anzellotti, A. I.; Liu, Q.; Bloemink, M. J.; Scarsdale, J. N.; Farrell, N. Targeting retroviral Zn finger-DNA interactions: a small-molecule approach using the electrophilic nature of trans-platinum-nucleobase compounds. *Chem. Biol.* **2006**, *13*, 539–548.
- (24) Loo, J. A.; Holler, T. P.; Sanchez, J.; Gogliotti, R.; Maloney, L.; Reily, M. D. Biophysical characterization of zinc ejection from HIV nucleocapsid protein by anti-HIV 2,2'-dithiobis[benzamides] and benzisothiazolones. *J. Med. Chem.* **1996**, *39*, 4313–4320.
- (25) Mayasundari, A.; Rice, W. G.; Diminnie, J. B.; Baker, D. C. Synthesis, resolution, and determination of the absolute configuration of the enantiomers of cis-4,5-dihydroxy-1,2-dithiane 1,1-dioxide, an HIV-1NCp7 inhibitor. *Bioorg. Med. Chem.* **2003**, *11*, 3215–3219.
- (26) Tisne, C.; Roques, B. P.; Dardel, F. Specific recognition of primer tRNA Lys 3 by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension. *Biochimie* **2003**, *85*, 557–561.
- (27) Zhang, Z.; Xi, X.; Scholes, C. P.; Karim, C. B. Rotational dynamics of HIV-1 nucleocapsid protein NCp7 as probed by a spin label attached by peptide synthesis. *Biopolymers* **2008**, *89*, 1125–1135.
- (28) Vuilleumier, C.; Bombarda, E.; Morellet, N.; Gerard, D.; Roques, B. P.; Mely, Y. Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. *Biochemistry* **1999**, *38*, 16816–16825.
- (29) Raja, C.; Ferner, J.; Dietrich, U.; Avilov, S.; Ficheux, D.; Darlix, J. L.; de, R. H.; Schwalbe, H.; Mely, Y. A tryptophan-rich hexapeptide inhibits nucleic acid destabilization chaperoned by the HIV-1 nucleocapsid protein. *Biochemistry* **2006**, *45*, 9254–9265.
- (30) Shvadchak, V.; Sanglier, S.; Rocle, S.; Villa, P.; Haiech, J.; Hibert, M.; Van, D. A.; Mely, Y.; de, R. H. Identification by high throughput screening of small compounds inhibiting the nucleic acid destabilization activity of the HIV-1 nucleocapsid protein. *Biochimie* **2009**, *91*, 916–923.
- (31) Ramboarina, S.; Srividya, N.; Atkinson, R. A.; Morellet, N.; Roques, B. P.; Lefevre, J. F.; Mely, Y.; Kieffer, B. Effects of temperature on the dynamic behaviour of the HIV-1 nucleocapsid NCp7 and its DNA complex. *J. Mol. Biol.* **2002**, *316*, 611–627.
- (32) Xi, X.; Sun, Y.; Karim, C. B.; Grigoryants, V. M.; Scholes, C. P. HIV-1 nucleocapsid protein NCp7 and its RNA stem loop 3 partner: rotational dynamics of spin-labeled RNA stem loop 3. *Biochemistry* **2008**, *47*, 10099–10110.
- (33) Protein Data Base; Brookhaven National Laboratory: Upton, NY; <http://www.rcsb.org/pdb>. Accessed August 2, 2009.
- (34) De Guzman, R. N.; Wu, Z. R.; Stalling, C. C.; Pappalardo, L.; Borer, P. N.; Summers, M. F. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **1998**, *279*, 384–388.
- (35) Bourbigot, S.; Ramalanjaona, N.; Boudier, C.; Salgado, G. F.; Roques, B. P.; Mely, Y.; Bouaziz, S.; Morellet, N. How the HIV-1 nucleocapsid protein binds and destabilises the (−)primer binding site during reverse transcription. *J. Mol. Biol.* **2008**, *383*, 1112–1128.
- (36) Urbaneja, M. A.; Kane, B. P.; Johnson, D. G.; Gorelick, R. J.; Henderson, L. E.; Casas-Finet, J. R. Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: elucidation of the structural determinants for function. *J. Mol. Biol.* **1999**, *287*, 59–75.
- (37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian03*, revision C.02, Gaussian, Inc.: Wallingford, CT, 2004.
- (38) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (39) Cieplak, P.; Cornell, W. D.; Bayly, C. I.; Kollman, P. A. Application of the multimolecule and multiconformational RESP methodology to

- biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.* **1995**, *16*, 1357–1377.
- (40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.
- (41) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.
- (42) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabre, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, CA, 2008.
- (43) Merz, K. M. Carbon dioxide binding to human carbonic anhydrase II. *J. Am. Chem. Soc.* **1991**, *113*, 406–411.
- (44) Kottalam, J.; Case, D. A. Langevin modes of macromolecules: applications to crambin and DNA hexamers. *Biopolymers* **1990**, *29*, 1409–1421.
- (45) Paoletti, A. C.; Shubsda, M. F.; Hudson, B. S.; Borer, P. N. Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. *Biochemistry* **2002**, *41*, 15423–15428.
- (46) Amarasinghe, G. K.; De Guzman, R. N.; Turner, R. B.; Chancellor, K. J.; Wu, Z. R.; Summers, M. F. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* **2000**, *301*, 491–511.
- (47) Morellet, N.; Demene, H.; Teilleux, V.; Huynh-Dinh, T.; de, R. H.; Fournie-Zaluski, M. C.; Roques, B. P. Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *J. Mol. Biol.* **1998**, *283*, 419–434.
- (48) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948.
- (49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.
- (50) Mahoney, M. W.; Jorgensen, W. L. A five-site model liquid water and the reproduction of the density anomaly by rigid, non-polarizable model. *J. Chem. Phys.* **2000**, *112*, 8910–8922.
- (51) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
- (52) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.
- (53) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.* **1977**, *23*, 327–341.
- (54) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide. *Biopolymers* **1992**, *32*, 523–535.
- (55) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.
- (56) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401–9409.
- (57) Morellet, N.; Julian, N.; de, R. H.; Maigret, B.; Darlix, J. L.; Roques, B. P. Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by ¹H NMR. *EMBO J.* **1992**, *11*, 3059–3065.
- (58) Giessner-Prettre, C.; Jacob, O. A theoretical study of Zn⁺⁺ interacting with models of ligands present at the thermolysin active site. *J. Comput.-Aided Mol. Des.* **1989**, *3*, 23–37.
- (59) Stote, R. H.; Kellenberger, E.; Muller, H.; Bombarda, E.; Roques, B. P.; Kieffer, B.; Mely, Y. Structure of the His44 → Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein: a combined NMR, molecular dynamics simulation, and fluorescence study. *Biochemistry* **2004**, *43*, 7687–7697.
- (60) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037–10041.
- (61) *Pymol Molecular Graphics System*; DeLano Scientific: Palo Alto, CA, 2002.
- (62) Lu, Q.; Tan, Y. H.; Luo, R. Molecular dynamics simulations of p53 DNA-binding domain. *J. Phys. Chem. B* **2007**, *111*, 11538–11545.
- (63) South, T. L.; Summers, M. F. Zinc- and sequence-dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the Psi-site analog, dACGCC. *Protein Sci.* **1993**, *2*, 3–19.
- (64) Surovoy, A.; Dannull, J.; Moelling, K.; Jung, G. Conformational and nucleic acid binding studies on the synthetic nucleocapsid protein of HIV-1. *J. Mol. Biol.* **1993**, *229*, 94–104.
- (65) Fogolari, F.; Brigo, A.; Molinari, H. The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *J. Mol. Recognit.* **2002**, *15*, 377–392.
- (66) De Rocquigny, H.; Petitjean, P.; Tanchou, V.; Decimo, D.; Drouot, L.; Delaunay, T.; Darlix, J. L.; Roques, B. P. The zinc fingers of HIV nucleocapsid protein NCp7 direct interactions with the viral regulatory protein Vpr. *J. Biol. Chem.* **1997**, *272*, 30753–30759.
- (67) Ramboarina, S.; Morellet, N.; Fournie-Zaluski, M. C.; Roques, B. P. Structural investigation on the requirement of CCHH zinc finger type in nucleocapsid protein of human immunodeficiency virus 1. *Biochemistry* **1999**, *38*, 9600–9607.
- (68) Ramboarina, S.; Druillennec, S.; Morellet, N.; Bouaziz, S.; Roques, B. P. Target specificity of human immunodeficiency virus type 1 NCp7 requires an intact conformation of its CCHC N-terminal zinc finger. *J. Virol.* **2004**, *78*, 6682–6687.
- (69) McGaughey, G. B.; Gagne, M.; Rappe, A. K. pi-Stacking interactions. Alive and well in proteins. *J. Biol. Chem.* **1998**, *273*, 15458–15463.
- (70) Hagan, N.; Fabris, D. Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. *Biochemistry* **2003**, *42*, 10736–10745.
- (71) Shubsda, M. F.; Kirk, C. A.; Goodisman, J.; Dabrowski, J. C. Binding of human immunodeficiency virus type 1 nucleocapsid protein to psi-RNA-SL3. *Biophys. Chem.* **2000**, *87*, 149–165.
- (72) Shubsda, M. F.; Paoletti, A. C.; Hudson, B. S.; Borer, P. N. Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry* **2002**, *41*, 5276–5282.
- (73) Darlix, J. L.; Vincent, A.; Gabus, C.; De Rocquigny, H.; Roques, B. Trans-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA. *C. R. Acad. Sci., III* **1993**, *316*, 763–771.
- (74) Druillennec, S.; Meudal, H.; Roques, B. P.; Fournie-Zaluski, M. C. Nucleomimetic strategy for the inhibition of HIV-1 nucleocapsid protein NCp7 activities. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 627–632.
- (75) Feng, Y. X.; Campbell, S.; Harvin, D.; Ehresmann, B.; Ehresmann, C.; Rein, A. The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol.* **1999**, *73*, 4251–4256.
- (76) Schmalzbauer, E.; Strack, B.; Dannull, J.; Guehmann, S.; Moelling, K. Mutations of basic amino acids of NCp7 of human immunodeficiency virus type 1 affect RNA binding in vitro. *J. Virol.* **1996**, *70*, 771–777.
- (77) Dietz, J.; Koch, J.; Kaur, A.; Raja, C.; Stein, S.; Grez, M.; Pustowka, A.; Mensch, S.; Ferner, J.; Moller, L.; Bannert, N.; Tampe, R.; Divita, G.; Mely, Y.; Schwalbe, H.; Dietrich, U. Inhibition of HIV-1 by a peptide ligand of the genomic RNA packaging signal Psi. *ChemMed-Chem.* **2008**, *3*, 749–755.
- (78) Takahashi, T.; Ueno, A.; Mihara, H. Nucleobase amino acids incorporated into the HIV-1 nucleocapsid protein increased the binding affinity and specificity for a hairpin RNA. *ChemBioChem* **2002**, *3*, 543–549.
- (79) Laughrea, M.; Shen, N.; Jette, L.; Darlix, J. L.; Kleiman, L.; Wainberg, M. A. Role of distal zinc finger of nucleocapsid protein in genomic RNA dimerization of human immunodeficiency virus type 1; no role for the palindrome crowning the R-U5 hairpin. *Virology* **2001**, *281*, 109–116.
- (80) Beltz, H.; Clauss, C.; Piemont, E.; Ficheux, D.; Gorelick, R. J.; Roques, B.; Gabus, C.; Darlix, J. L.; de, R. H.; Mely, Y. Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis. *J. Mol. Biol.* **2005**, *348*, 1113–1126.