





Review

# Overview of the Nucleic-Acid Binding Properties of the HIV-1 Nucleocapsid Protein in Its Different Maturation States

Assia Mouhand <sup>1,†</sup>, Marco Pasi <sup>2,†</sup>, Marjorie Catala <sup>1</sup>, Loussiné Zargarian <sup>2</sup>, Anissa Belfetmi <sup>2</sup>, Pierre Barraud <sup>1</sup>, Olivier Mauffret <sup>2,\*</sup> and Carine Tisné <sup>1,\*</sup>

<sup>1</sup> Expression Génétique Microbienne, UMR 8261, CNRS, Institut de Biologie Physico-Chimique (IBPC), Université de Paris, 75005 Paris, France; assia.mouhand@gmail.com (A.M.); marjorie.catala@ibpc.fr (M.C.); pierre.barraud@cnrs.fr (P.B.)

<sup>2</sup> Laboratoire de Biologie et de Pharmacologie Appliquée (LBPA), UMR 8113 CNRS, Institut D'Alembert, École Normale Supérieure Paris-Saclay, Université Paris-Saclay, 4, Avenue des Sciences, 91190 Gif sur Yvette, France; marco.pasi@ens-paris-saclay.fr (M.P.); loussine.zargarian@ens-paris-saclay.fr (L.Z.); anissa.belfetmi@hotmail.fr (A.B.)

\* Correspondence: olivier.mauffret@ens-paris-saclay.fr (O.M.); carine.tisne@cnrs.fr (C.T.)

† These authors contributed equally to this work.

Received: 31 August 2020; Accepted: 26 September 2020; Published: 29 September 2020



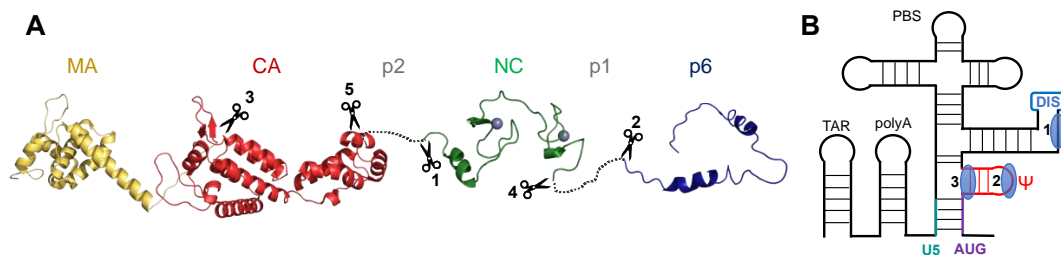
**Abstract:** HIV-1 Gag polyprotein orchestrates the assembly of viral particles. Its C-terminus consists of the nucleocapsid (NC) domain that interacts with nucleic acids, and p1 and p6, two unstructured regions, p6 containing the motifs to bind ALIX, the cellular ESCRT factor TSG101 and the viral protein Vpr. The processing of Gag by the viral protease subsequently liberates NCp15 (NC-p1-p6), NCp9 (NC-p1) and NCp7, NCp7 displaying the optimal chaperone activity of nucleic acids. This review focuses on the nucleic acid binding properties of the NC domain in the different maturation states during the HIV-1 viral cycle.

**Keywords:** HIV-1; nucleocapsid; nucleic acids; RNA; RNA-binding protein; NMR

## 1. Introduction

The HIV-1 Gag (group-specific antigen) protein coordinates all major steps in virion assembly. Gag comprises three ordered domains, namely the matrix (MA), the capsid (CA) and the nucleocapsid (NC) domains, which are connected by two linkers, p2 (also named SP1) and p1 (also named SP2) (Figure 1A). p6 was shown to be structured under membranous solution conditions [1]; however, p1–p6 was shown to be mostly disordered in the context of the C-terminal (Cter) part of Gag and to present regions that transiently adopt  $\alpha$ -helical structures [2]. HIV-1 Gag molecules exist as monomers or low-order multimers in the cell cytosol and only form higher-order multimers after binding to the plasma membrane [3], nucleating the assembly of thousands of Gag molecules into an immature virion [4]. Small numbers of Gag molecules selectively recruit a dimeric, unspliced viral RNA genome (gRNA) selected from a pool of excess cellular RNAs and spliced viral mRNAs [4–9] through the binding of the NC domain to the 5'-leader of gRNA (Figure 1B) [10–12]. The formation of infectious HIV-1 particles is triggered by the sequential proteolysis of Gag polyprotein by the viral protease [13,14]. During or shortly after the budding of the particle, the HIV-1 protease cleaves Gag to trigger HIV-1 maturation [15]. Three different forms of NC appear subsequently (Figure 1A). The first cleavage by the HIV-1 protease occurs between MA-CA-p2 and NC-p1–p6, thereby liberating NCp15. The second cleavage frees NCp9 (NC-p1), whereas the last one releases the NC domain itself, which constitutes the final maturation form (NCp7). The optimal RNA chaperone activity of the NC domain, essential notably

during reverse transcription (RT) [16–24], is observed for the mature form of NC. NCp15-containing virions are noninfectious [25–28], and the cleavage between p1 and p6 is therefore critical. The ability of the protease to achieve ordered Gag processing at specific cleavage sites, which display little to no sequence identity, primarily originates from the conformational dynamics of the protease flaps that cover its active site and that are involved in substrate recognition [29]. The modulation of the flap opening through protease/Gag interactions finely tunes the lifetime of the productive complex and hence the likelihood of Gag proteolysis.



**Figure 1.** Organization of the Gag protein and of the 5'-leader gRNA: (A) Juxtaposition of structures solved for the different domains of Gag (MA: PDB 1UPH, CA: PDB 3GV2, NC: PDB 1A1T, p6: PDB 2C55), the scissors indicating the cleavage sites by the viral protease and the numbers indicating the order of cleavage; (B) Secondary structure of the 5'-leader gRNA in the dimer promoting conformation (AUG: *gag* translation start codon, DIS: Dimerization Initiation Site that promotes genome dimerization, PBS: Primer Binding Site for initiation of RT, polyA: polyAdenylation signal, Ψ: packaging signal hairpin, TAR: *trans*-Activator Response region that stimulates transcription). The main binding sites of NCp7 or the NC domain of Gag described in the text are indicated by a blue ellipse.

Both MA and NC domains within Gag can bind to RNA in the cytoplasm [9,30]. However, the NC domain within Gag drives gRNA packaging by binding to the packaging signal Ψ. In contrast, MA binds a subset of tRNAs in the cytosol, which regulates MA and Gag-membrane binding [31,32]. Global changes in the RNA-binding specificity of Gag regulate virion synthesis. Before and after virion assembly and maturation, the NC within Gag preferentially binds to the Ψ region [6,33–35] and Rev Response Elements (RRE) of the gRNA [31]. The RNA-binding properties of NC in its mature forms hence resemble that of unassembled Gag in the cytosol. During assembly and budding, the NC within Gag nonspecifically binds to many sites on the gRNA [31], facilitating genome packaging. After assembly, the release step is mediated by the cellular ESCRT machinery, notably TSG101 and ALIX proteins [36–41]. For both TSG101 and ALIX, p6 bears the strongest interaction sites within Gag, but the NC cooperates with p6 for their binding and for promoting virus budding [42–44]. The coordination between protease activation, virion assembly and budding is critical and ensures the infectivity of progeny virions [15,42].

NC is a nucleic acid (NA)-binding protein that plays a role in almost all stages of the HIV-1 replication cycle: dimerization of gRNA, gRNA packaging [9], assembly of virions, hybridization of the primer tRNA<sup>Lys</sup><sub>3</sub> to the PBS [18,45–47], strand-transfer steps during RT [48–50] and integration of proviral DNA (for a review, see [51]). Many of these functions involve a structural reorganization of the NA structure. NC chaperones these rearrangements, leading to more thermodynamically stable structures [23]. The mature form of NC, NCp7, is a small protein of 55 residues, very basic (pI = 9.9), which contains two domains, each structured around a Zn<sup>2+</sup> ion (Figure 1A), named zinc-knuckle (ZK) domains, with the consensus sequence: CX<sub>2</sub>CX<sub>4</sub>HX<sub>4</sub>C. The electrostatic interactions represent an essential component of NC interactions with NAs, and NC binding is indeed dramatically dependent on the ionic strength [52–56]. Nevertheless, some NC interactions with NAs are more specific, with G-rich single-stranded sequences flanked by stable helices being the preferred sequence binding sites for NCp7 [20,31,57–63].

During the viral cycle, the nucleocapsid domain exists first within the Gag protein, then within NCp15 and NCp9 during the maturation of Gag by the viral protease, and as NCp7 in the mature form. This review focuses on the NA-binding properties of NC in its different forms of maturation.

## 2. NC Domain in Its Immature States (Gag, NCp15, NCp9)

### 2.1. Structure of the Packaging Signal and Binding of NC

During the HIV-1 viral cycle, viral RNA transcripts function as mRNAs: spliced transcripts encode the viral envelope and accessory proteins, whereas unspliced transcripts produce the Gag and Gag-Pol polyproteins by regulated frameshifting during translation. Some unspliced transcripts do not function as mRNAs but are instead selected for packaging into assembling virions as dimers of gRNA [5,6,64,65], a dimer being necessary for strand transfer-mediated recombination during RT [66]. Since the discovery of HIV-1, numerous studies have focused on the selection of a dimeric gRNA among cellular and viral RNAs, which still generates multiple debates in the literature. The two main questions regarding this process are: (1) the formation/structure of the gRNA dimer through the DIS/DIS interaction in the 5'-leader of gRNA, and (2) the sequence and structure determinants of the packaging signal in the 5'-leader. Both questions involve NC-mediated binding of Gag to the 5'-leader of gRNA. The binding of Gag at the level of the DIS sequence, notably its internal loop, is responsible for its slightly higher capacity, when compared to NCp7, to recruit gRNA versus spliced HIV-1 RNAs (Figure 1B, site 1) [35,67,68]. In cellulo, Gag recognizes particular motifs such as Rev Response Element (RRE) and the packaging signal (including SL1 and SL3) [31,69]. Recent studies involving a series of impressive NMR studies [6,7,62,64,70–72] have shown outstanding results using smart labeling strategies in order to answer these questions [73–75]. In these studies, the binding of NCp7, used as a substitute for Gag in the context of gRNA dimerization and its selective packaging, was extensively studied. A detailed description of these studies is beyond the scope of this review, but we chose to present the major results related to the binding of NC.

For gRNA dimerization, the dimer formation is promoted by the base-pairing of the AUG region (gag start codon, previously known as SL4 hairpin loop) with the U5 sequence (Figure 1B) [7,71]. The dimerization process appears to expose numerous NC binding sites in the 5'-leader region. The gRNA and mRNA diversification of functions was initially thought to be modulated by a riboswitch-like mechanism, in which the role of a single RNA transcript was modulated by changes in its structure [7,76,77]. It now seems that transcript dimerization and fate are controlled at the level of transcription by heterogeneous start site usage. Indeed, transcripts that begin with two or three sequential 5'-guanosines form monomers *in vitro* and are retained in cells for splicing and/or translation. In contrast, those starting with a single 5'-guanosine form dimers and are selectively packaged into virions [64,78,79].

For gRNA packaging, the HIV-1 packaging signal is located within the 5'-leader of the gRNA [6,11,65], which is among the most conserved regions of the viral RNA [80,81] (Figure 1B). Early efforts to identify HIV-1 RNA packaging determinants within the 5'-leader identified stretches of nucleotides surrounding and including the  $\Psi$  hairpin (Figure 1B, site 2) that, if deleted, resulted in severe HIV-1 packaging defects [10,82,83]. The  $\Psi$  hairpin was thus proposed to be the major RNA packaging determinant. However, mutations in the GGAG loop of  $\Psi$  abolished NC binding but did not significantly affect RNA packaging [6,84], and it soon became clear that regions other than the apical loop of the  $\Psi$  hairpin were important for RNA packaging [65,70,84–87]. A region of the leader sufficient for packaging heterologous vectors into assembling virions was recently identified [70], and was shown to adopt a tandem three-way junction structure [72]. Besides, the conserved [UUUU]:[GGAG] element, located in the stem of the  $\Psi$  hairpin (Figure 1B) and adjacent to this tandem three-way junction structure, was shown to be critical both for high-affinity NC binding and competitive RNA packaging [62]. The HIV-1 leader binds approximately two dozen copies of the NC protein with affinities ranging from ~40 nM to 1.4  $\mu$ M. Each of the conserved [UUUU]:[GGAG] elements in the dimeric gRNA binds two NC

molecules with affinities (~40 nM) approximately one order of magnitude greater than those of other NC binding sites within the encapsidation-promoting region of the viral leader, and more than 30 times tighter than NC binding sites in other regions of the leader [62]. Binding to the four highest-affinity sites occurs with endothermic energetics attributed to NC-induced localized RNA melting, possibly helix melting at the level of the [UUUU]:[GGAG] element that presents a labile helical structure in the absence of NC (Figure 1B, site 3). Mutations that stabilize these sites inhibit NC binding in vitro and RNA packaging in transfected cells [62]. The absence of this [UUUU]:[GGAG] element in spliced viral transcripts is proposed to contribute to the selective packaging of the full-length gRNA. These findings appear consistent with chemical probing studies showing that residues of [UUUU]:[GGAG] were resistant to RNase V1 cleavage in the absence of NC but exhibited an enhanced chemical accessibility after Gag/NC binding [35,88].

Altogether, these studies underline the crucial role of unpaired guanines in the dimerization and packaging processes involved in the high-affinity binding to gRNA of NC within Gag.

## 2.2. Investigations of Gag, NCp15, NCp9 Structures and Dynamics, and NA-Binding Properties

The structural investigation of the full-length HIV-1 Gag protein has so far been prevented by its low solubility and its extreme sensitivity to proteolysis [89]. To date, only a Gag construct lacking p1 and p6 was amenable to NMR, and the structured domains within Gag (MA, CA and NC) were shown to retain the same fold as the isolated domains [90]. The structured domains of Gag reorient semi-independently from each other through p2 and p1 that are unstructured and that connect the structural domains. In this study, Gag exists as a dynamic equilibrium between monomers and dimers and shows a strong interaction with DNA through the NC domain and a weak secondary interaction through MA. They also confirmed that the proteolytic cleavage at site 1 (Figure 1A) was accelerated in the presence of single-stranded DNA, with no evidence of interaction between p2 and DNA, suggesting that substrate accessibility rather than substrate conformation constituted the rate limiting step for Gag proteolysis.

The affinities for NAs of NCp7 (NC), NCp9 (NC-p1) and NCp15 (NC-p1-p6), the Cter part of HIV-1 Gag, were shown to be similar [2,91]. However, distinct NA-binding properties were described concerning the kinetics of association/dissociation with NAs and the thermodynamics of interaction with NA. Slower NCp15 dissociation kinetics were observed when compared to NCp9 and NCp7, in line with the poorer NA-chaperone activity and NA-aggregation capacity of NCp15 [91]. More recently, the structure and dynamics of the Cter part of HIV-1 Gag were investigated by exploring the conformational landscape of the different stages of maturation of NC (NCp7, NCp9 and NCp15) using <sup>15</sup>N relaxation measurements [2]. For the three proteins, the two ZKs define two well-structured regions separated by a more flexible linker, showing that the global folding and dynamics of the NC domain is conserved during maturation. However, relative to NCp7 and NCp9, the N-terminal (Nter) part of NCp15 is more rigid, while residues of the ZKs and the linker are globally more flexible [2]. p1–p6 within NCp15 is disordered, but short-lived  $\alpha$ -helices are present in p1–p6 and are promoted by long-range and transient contacts with the NC domain. Notably, p1 transiently interacts at one site comprising the ZKs through hydrophobic contacts, whereas p6 has multiple sites of interaction in the Nter part of NC, which is unfolded and positively charged, through electrostatic interactions [2]. The three proteins can bind the SL3/ $\Psi$  hairpin using the same specific interactions, and p1–p6 do not participate directly in SL3 binding. This explains why the affinities for NAs are the same for NCp7, NCp9 and NCp15. However, the precursors NCp9 and NCp15 exhibit a higher entropic penalty for the binding of SL3 than NCp7. These changes in the entropic penalty for SL3 binding are correlated with changes in the fast dynamics observed for the different states of maturation of the NC domain. Indeed, the increase in global flexibility of the ZKs and of the linker does not change the global affinity of the NC domain for SL3, but instead results in a conformational entropic penalty to SL3 binding by the precursors. This study thus provides a new picture of the structural and dynamic properties of the NC domain during the assembly, budding and maturation of virions [2]. During the assembly

and budding of virions, concomitant with Gag oligomerization, the transient interactions between NC and p1–p6 become salient and responsible for (i) a higher level of p6 secondary structure favoring the recruitment of its budding partners and (ii) a higher entropic penalty to NC RNA-binding at specific sites, leading to the nonspecific RNA-binding of NC at a multitude of sites on gRNA that are necessary during assembly. The maturation by the protease cleaving off p6 and p1 from NC reverts the RNA-binding specificity of the mature NC. NCp7 no longer binds to multiple sites but binds to discrete specific sites on gRNA containing unpaired guanine. This model explains the apparent changes of the RNA-binding specificity of the NC domain during assembly and agrees with an ITC study showing that oligomerization-competent forms of Gag displayed a ~3-times stronger binding affinity for nonspecific RNA motifs over the cytosolic specific target of NC [92]. This was not observed for oligomerization-impaired forms of Gag, and for Gag lacking p6 and capable of high-order Gag oligomerization [92]. In conclusion, p1–p6 has a role in facilitating the binding of nonspecific RNA motifs to oligomeric forms of Gag present at the moment of the assembly of the virions.

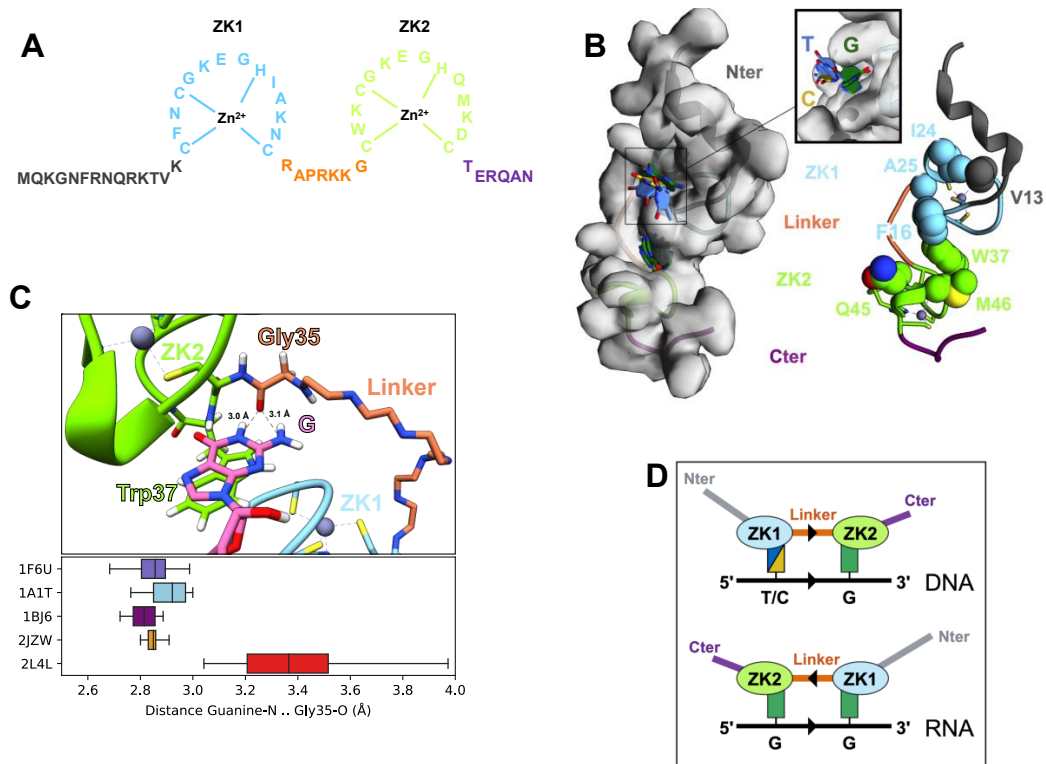
### 3. NCp7 and NA Interactions

#### 3.1. NMR Structures of NCp7-NA Complexes

To date, five structures of NCp7-NA complexes are available in the PDB and have been determined using NMR spectroscopy [58–61,93,94]. All NC-RNA complexes were obtained using the NC sequence from the NL4-3 strain (Figure 2A) and with RNA fragments that were either hairpins or a single-stranded RNA. The structures of NC-DNA complexes were solved with a truncated form of NC, lacking the first ten Nter residues, with the NC sequence from the MAL strain and with DNA hairpins (TAR, PBS) or a single-stranded DNA. Overall, the folding of the individual ZKs was the same in all complexes. A recent study has extensively analyzed these complexes, and using bioinformatic methods has quantified the interface between NC and NA in all the available 84 conformers [95]. The ZK2-NA interface was the same in all structures, while a broad diversity of ZK1-NA interfaces was observed. The reason for this diversity is not apparent and may originate from the differences in NA and/or NC sequences. The quantification of the contact areas shows that Phe16 and Arg/Lys26 in ZK1 and Trp37 in ZK2 are the most contacted residues in NC-NA complexes. Each ZK has a hydrophobic pocket for recognizing a base in a single-stranded part of the NA. The hydrophobic pockets are defined by Val13, Phe16, Ile24, Ala25 in ZK1 and Trp37, Gln45 and Met46 in ZK2 (Figure 2B). In all NCp7-NA complexes, a guanine is deeply inserted inside the hydrophobic pocket of ZK2 and interacts through  $\pi$ - $\pi$  stacking with Trp37. ZK1 binding is more versatile, and ZK1 can make contacts with either a guanine, a thymine or a cytosine (Figure 2B). However, for the complexes with a thymine or a cytosine bound in ZK1, the base is located at the periphery of the hydrophobic pocket. In contrast, a guanine is deeply inserted into the ZK1 hydrophobic pocket (Figure 2B, inset) through  $\pi$ - $\pi$  stacking with Phe16. In addition to  $\pi$ - $\pi$  stacking with aromatic residues, the binding of guanines in each ZK is reinforced by hydrogen bonds between the guanine base and the backbone of several residues, including residues of the linker. For instance, in all NCp7-NA complexes, Gly35 is involved in a network of hydrogen bonds with the guanine in ZK2 (Figure 2C). These structural studies agree with fluorescence spectroscopic studies [19,52,56,96–98], insofar as they show that ZKs contain hydrophobic residues at adequate positions to form an ideal surface for the recognition of guanine residues located in a single-stranded section of the NA.

In all complexes, arginine and lysine residues, fifteen of which are present in full-length NCp7, are observed to participate in numerous nonspecific electrostatic interactions with the phosphodiester backbone of NA. Notably, the Nter part of NCp7, present in all NCp7-RNA complexes, is very basic and folds as a  $3_{10}$  helical structure upon binding to the RNA. Three salt bridges, not observed in free NCp7 structures [25,99], are observed within NCp7 in several complexes: Lys33–Glu42 (between linker and ZF2), Lys38–Glu51 (within ZF2) and Lys14–Glu21 (within ZF1). These structural data are in

agreement with a large number of physico-chemical studies pointing out the significant role of ionic forces in stabilizing the NC and Gag-NA complexes [33,55,96,100–102].



**Figure 2.** Structural determinants of NCp7-NA complexes: (A) Sequence of the HIV-1 NCp7 for the NL4.3 strain; (B) 3D structure representations of the interaction interface of NCp7 with NA. (Left) Surface of NCp7 in complex with the  $\Psi$  hairpin, also named SL3 (PDB: 1A1T), showing the location of the two NA-binding pockets in each ZK. The backbone trace is colored to highlight the different regions of NC. The bases bound in the pockets (shown as filled rings and colored according to the base type, see inset) are taken from all available structures of NC-NA complexes (each pocket is superposed using the  $C\alpha$  atoms of the corresponding ZK). (Right) The side chains lining the two NA-binding pockets are shown as Van der Waals spheres;  $Zn^{2+}$  ions are shown as gray spheres, while coordinating residues are shown as sticks. (C) Hydrogen bonds between the guanine bound in the ZK2 pocket and the backbone of Gly35 in NC-NA complexes. (Top) Interaction between the G base (pink) and the backbone of Gly35 (orange) in the NC-SL3 complex. G is stacked with the side chain of Trp37 (green). The linker is shown as the minimal backbone trace. (Bottom) Box plot of the distance between the backbone oxygen of Gly35 and the closest nitrogen of the Guanine bound in the ZK2 pocket, showing that the distances are compatible with a hydrogen bond in all conformations of all available NC-NA complexes. (D) Schematic representation of the respective polarity of NC and NA in NC-NA complexes. In NC-DNA complexes (top), the Nter→Cter direction of NC is parallel to the 5'→3' direction of the DNA, whereas in NC-RNA complexes (bottom), the two directions are antiparallel.

Last, the analysis of NCp7-NA structures shows that NCp7 can interact differently with the backbone riboses according to the nature of the NA [61]. Indeed, in NC-DNA complexes, ZK1 interacts with a base upstream (5'-side) of the G bound by ZK2, whereas in NC-RNA complexes, ZK1 interacts with a residue located downstream (3'-side) of this G (Figure 2D). Thus, NC can adapt to different sugar conformations to trap the G in the hydrophobic pocket of ZK2, suggesting that the binding of NCp7 to NA is driven by this interaction. This ability to recognize NA polarity is in agreement with the large differences in binding affinities measured according to the NA sequence polarity [56], as described for other single-stranded DNA-binding proteins [103].

Altogether, these structures explain the preference of NC for single-stranded NAs and for unpaired guanine and how NC can recognize either RNA or DNA.

### 3.2. Dynamics of NCp7 and NCp7-NA Complexes

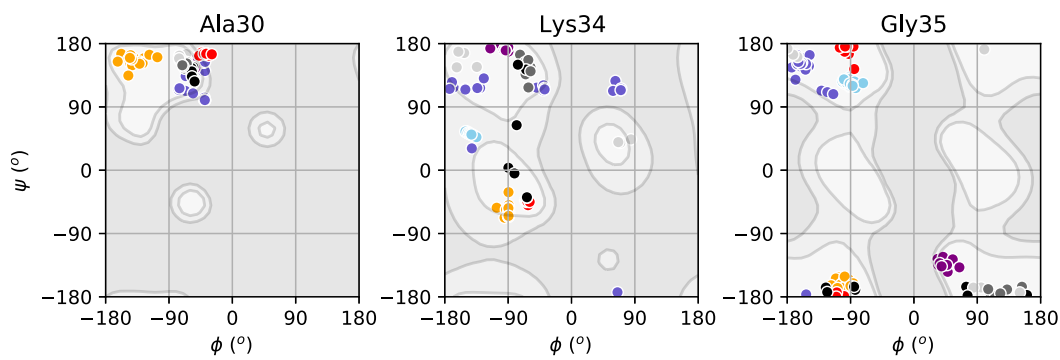
Early study of the dynamics of NCp7 using  $^{15}\text{N}$  NMR relaxation demonstrated that NCp7 is in a rapid equilibrium between several conformations from weakly interacting to noninteracting ZKs [99]. The linker, the Nter and Cter parts of NCp7 are disordered and particularly flexible. In contrast, both ZKs adopt a folded and more rigid conformation and reorient independently of each other due to the linker's flexibility. This suggests that the two ZKs could function independently from each other in the different steps of the NA recognition process.

NCp7 dynamics were further analyzed using  $^{15}\text{N}$  relaxation measurements performed at several magnetic fields [2,104], combining residual dipolar coupling (RDC), SAXS and ensemble-simulated annealing [105], using relaxation dispersion and chemical exchange saturation transfer measurements [106], as well as molecular simulations [107]. Globally, these data demonstrated that the two ZKs were not equivalent. While the overall conformational space sampled by NCp7 was large, the conformational distribution of ZK1 was distinct from that of ZK2 as a result of differential flexibility along the linker region. The largest conformational freedom was observed for Lys34 and Gly35, comparable to a hinge, located at Gly35 just before the beginning of ZK2. This feature was also observed in NCp9 and NCp15 [2]. In contrast, linker residues from Cys28 to Lys33 displayed restricted motions and numerous contacts with residues of ZK1. This differential flexibility within the linker resulted in an asymmetric motion of the ZKs, which may explain their distinct role in NA-binding despite their high sequence identity (see the following paragraphs).

NCp7-NA complexes globally exhibit a loss of flexibility upon binding for both the NA and NCp7 [2,97,98,108,109]. To compare the flexibility within the linker in NCp7 and NCp7-NA complexes, we analyzed the backbone dihedral angles ( $\varphi$  and  $\psi$ ) of linker residues (Ala30 to Gly35) in all the available structures (Figure 3). In the unbound NCp7 [105], Lys34 and Gly35 exhibited by far the largest backbone dihedral fluctuations among the residues in the linker (Figure 3, grey points), confirming the role of these two residues as a hinge for ZK motions. In particular, this backbone flexibility seems to involve the  $\psi$  angle of Lys34 and the  $\varphi$  angle of Gly35. Binding to NAs significantly reduces the conformational freedom of Lys34 and Gly35, narrowing the range of accessible  $\varphi$  and  $\psi$  values, particularly for Gly35, which is involved in the guanine binding (Figure 2C). It is tempting to speculate that directing the backbone flexibility at the level of a few degrees of freedom involving the backbone of Gly35 could create a molecular switch that would entail a more rigid structure of NC upon NA binding.

### 3.3. Conclusions

Altogether, the structural and dynamic data discussed here highlight several aspects of NC-NAs interactions: 1) ZK2 interacts with unpaired guanines, the preferred substrate of NC, and is more prone to do so than ZK1; 2) NC orients on NA chains with a definite polarity, depending on the nature of the NA (DNA or RNA); 3) the dynamics of the two ZKs are remarkably different, due to their different localizations in the sequence rather than to their slightly different compositions in amino acids. These biophysical properties explain a number of observations reported in the literature, and the next paragraph is devoted to this point.



**Figure 3.** Ramachandran plot for selected linker residues. The  $\phi$  and  $\psi$  values for Ala30 (left), Lys34 (middle) and Gly35 (right) are superposed to the probability distribution of the same residue in globular proteins taken from high-resolution structures in the PDB (<https://www.rcsb.org/>), where lighter shades of gray indicate a higher probability of occurrence. Ala30 is shown as a reference for poor variation in free NC. Backbone dihedral angles were calculated on the following PDB structures: 511R [105] (NCp7, grey, 21 conformers); 1A1T [59] (NCp7-SL3, sky blue, 25 conformers); 1F6U [60] (SL2-NC, slate blue, 20 conformers); 2L4L [61] (cTAR-NC, red, 10 conformers); 2JZW [58] (NCp7-PBS, orange, 19 conformers); 1BJ6 [93] (NCp7-d(ACGCC), purple, 10 conformers). The 21 conformers of 511R are further subdivided according to the definitions given in the original paper: cluster 1 (grey, ~40%), cluster 2 (dark grey, ~50%), cluster 3 (light grey, ~10%).

#### 4. The Two ZKs Are Not Equivalent

##### 4.1. Biochemical and Biological Evidence

Numerous studies focused on the distinct roles of ZKs in the various steps of the replication cycle. In Table 1, we have summarized the modifications or mutations incorporated in NC and the phenotype observed for in vivo studies, or the defects observed in one step of the viral cycle for in vitro studies.

**Table 1.** Implication of NC protein modifications or mutations on the phenotype and/or defects of the viral cycle.

Modifications/Mutations	Phenotype/Defects	Conclusions	References
Cx->Sx or Hy->Ay	Defect in RNA packaging, assembly not correct, defects in NA annealing	More deleterious effects for mutations in ZK1	[12,110,111]
Cx->Hx and Hy->Cy	Normal RNA packaging but virus noninfectious	More deleterious effects for mutations in ZK1	[112–114]
ZK2-ZK1 ZK1-ZK1 ZK2-ZK2	Defect in strand-transfer and self-priming reaction during RT	ZK1 must be in the first position, a major role of ZK1 for destabilization activity of NC	[100,115]
ZK2-ZK1 ZK1-ZK1 ZK2-ZK2	Defect in NA annealing	ZK1 must be in the first position, a major role of ZK1 for destabilization activity of NC	[116]
ZK2-ZK1 ZK1-ZK1 ZK2-ZK2	Defects in DNA stretching ability as a test of chaperone activity	ZK1 must be in the first position, a major role of ZK1 for destabilization activity of NC	[117]
NC(1–35)	Sufficient for in vitro chaperone activity	Critical need of ZK1 for chaperone activity	[118]
NC(29–55)	Not able to recognize packaging signal, low affinity for NA	Critical need of ZK1 for interaction with NA	[118,119]



Table 1. Cont.

Modifications/Mutations	Phenotype/Defects	Conclusions	References
Mutation of Trp37 in ZK2 or Phe16 in ZK1	Defect in NC chaperone activity, NA-binding strongly affected	Critical role of Trp37 in ZK2, less critical for Phe16 in ZK1	[120]
Basic residues	Inhibition of gRNA dimerization, reduction of gRNA packaging, strong effect on virus assembly, reduction of infectivity	Critical for RNA packaging, optimal chaperone activity and infectivity	[101,121–123]
Deletion of the linker	Decrease of NA-binding affinity	Linker contributes to NA-binding and chaperone activity of NC	[118]
Deletion of ZK1 and/or ZK2 within Gag	Production of viral particles containing DNA	ZK2 more important than ZK1 to block premature RT	[42,124]
Deletion of ZK1 and/or ZK2 within Gag	Loss of affinity in Gag-gRNA complexes	ZK2 more important to ensure Gag-gRNA complexes of high affinities	[102]
Deletion of ZK1 and/or ZK2 within Gag	Recruitment of Gag-gRNA complexes at the plasma membrane (PM)	ZK2 more important for the accumulation of ribonucleoprotein complexes at the PM	[125]

Initial studies on NC modifications covered: (i) alterations in the coordination array of each  $\text{Zn}^{2+}$  ion (mutation of C to S or H and H to C or A) [10,12,111–115,126]; (ii) swapped mutants that exchanged the position between ZK1 and ZK2 (ZK2-ZK1) or duplication of one ZK (ZK1-ZK1 or ZK2-ZK2) [115]; (iii) modifications of the basic residues located inside and/or outside of the two ZKs [110,122,123,127]. In most of these studies, plasmids containing the NC-mutated HIV genome were transfected in cells, and viral particles were then extracted and analyzed after several days of post-transfection cultures [10,12,110,115]. The major result from these studies was that both ZKs were required for an optimal RNA packaging and efficient viral infectivity. However, alterations in ZK1 provoked more deleterious defects than in ZK2 [12]. Interestingly, these studies were crucial to identifying that NC was involved in an early step of the replication cycle occurring before the RNA packaging: subsequently, it was found that this step was the reverse transcription (RT) and that an NC mutant able to partially perform RNA packaging could be totally unable to perform RT, leading to noninfectious viruses [12,112].

The discovery of the involvement of NCp7 or NCp9 in gRNA dimerization and in different steps of RT, such as the initiation of RT through the hybridization of  $\text{tRNA}^{\text{Lys}_3}$  to the PBS in the 5'-leader of gRNA and the stimulation of the two strand-transfers during RT [16,23,24,98,128,129], fueled further investigations of the NA chaperone capability of NC, as well as of its NA-binding properties. Indeed, effective binding is a necessary, but not sufficient, condition for NA-chaperone activity. Modifications of the two ZKs were found to affect the RNA-packaging and NA-chaperone activities of NC (strand-transfer, self-priming, NA annealing and NA-helix destabilization) differently. Pioneering studies have shown that ZK1 is absolutely required for RNA annealing, such as HIV-1 gRNA dimerization or annealing of  $\text{tRNA}^{\text{Lys}_3}$  to the PBS [45,130]. ZK2-ZK2 and ZK2-ZK1 constructs of NCp7 did not exhibit any chaperone activity, whereas ZK1-ZK1 showed a satisfying (but nonoptimal) chaperone activity [100]. ZK1 is thus more critical for chaperone activity than ZK2, and must be located in the Nter part of NC for a correct chaperone activity [100,116,117]. ZK2 is, however, necessary to increase the NC annealing rate [116]. In addition, the ZK1-ZK1 construct is a weaker binder than NCp7 but performs better than the ZK2-ZK2 and ZK2-ZK1 mutants [131–133]. More recently, a peptide comprising the first 35 Nter residues of NC, NC(1–35), was shown to be necessary and sufficient for NA-chaperone activity in vitro [118], highlighting the critical role of ZK1 in this context. NC(1–35) displays a binding affinity for NAs similar to that of NC, whereas NC(1–28) (further lacking the linker) and NC(29–55) (containing just the linker and ZK2) have decreased affinities [118]. NC(1–35) can also recognize the

packaging signal  $\Psi$ , while NC(29–55) cannot [119]. Similar ZK constructs produced in the context of the Gag $\Delta$ p6 precursor gave totally different results in assays of  $\Psi$  binding in a 1M concentration of NaCl, thus favoring the nonelectrostatic component of the interaction [102]. While the construct with ZK2 has almost the same affinity for  $\Psi$  as NC, those with ZK1 (i.e.,  $\Delta$ ZK2) exhibit a dramatic decrease of three orders of magnitude [102], in agreement with previous data [55]. The contradictory results of these studies underline the importance of the conditions used to analyze the contribution of the two ZKs to NA binding affinities. A critical point could also lie in the ability of the linker residues to interact with ZK1 residues. Indeed, such interactions have been proposed as exerting a significant influence on NA binding [104,105,134]. A comprehensive study of NC using polyetheno-adenylic acid showed that NA binding stems more from ZK1 than from ZK2, notably with two lysines from ZK1 involved in electrostatic contacts with NA instead of only one for ZK2 [55]. These results are rather well corroborated by the structure of the NC-SL3  $\Psi$  complex [59]. Additionally, when working at a high salt concentration (1M NaCl) to focus on the nonelectrostatic contribution of NA-binding, results are similar to those obtained in the context of Gag [102].

The molecular basis of the ability of ZK1 to better destabilize the NA secondary structure, i.e., helix unwinding ability, was investigated in several studies [55,97,116,131,135]. Point mutations in ZK1 to reproduce the sequence of ZK2 show that the I24Q mutation impairs helix unwinding, while N27D alters both RNA binding and helix unwinding [135]. The more hydrophobic nature of ZK1, relative to ZK2, was generally proposed to explain the higher helix-unwinding ability of ZK1 [55,97,116,131,135]. In the case of the PBS (-), whose stem's destabilization by NC has been established [58], the complex structure shows that the stem is in contact with hydrophobic residues of ZK1, in particular Val13, Phe16, Thr24 and Ala25 [58,104]. In all structures of NC in a complex with hairpins, whether RNA or DNA [58–60], the stem of the hairpin interacts with the hydrophobic residues of ZK1, but not with those of ZK2 (Figure 2B), corroborating the implication of ZK1 in the helix-unwinding activity.

Biophysical methods, like optical tweezers and FRET, were also extensively used to probe the NA chaperone activity in response to different alterations of the ZKs [117,120,131,136]. These studies confirmed the major role of ZK1 in the NA-destabilization activity of NC, its necessity to be located at the Nter part of NC and the critical role of Trp37 in ZK2 [131,135,137]. This role of Trp37 was further described in several publications [97,109,120,121,131], showing the interdependence of the two ZKs in this activity. The mutation of Phe16, the aromatic residue in ZK1 that occupies the same position as Trp37 in ZK2, to a nonaromatic residue (Ala or Leu) also appears to reduce the chaperone activity of NC. Still, these effects are less severe than for Trp37 [120,131]. These aromatic residues, Phe16 and Trp37, also play a crucial role in NA binding: mutating either of these to a nonaromatic residue strongly affects the NA-binding affinities of NC [56,97,120,131,132,138], in agreement with NA-NC structures; again, mutating Trp37 was more deleterious than mutating Phe16 [97,120].

Altogether, this large corpus of data suggests that ZK1 plus the linker, NC(1-35), possess all the needed elements to achieve chaperone activity and NA binding, including the selective binding of  $\Psi$  [118,119]. However, the mutation of Trp37 in ZK2 to a nonaromatic residue dramatically alters both the NA-binding and chaperone activities [55,120,131]. This apparent contradiction may be reconciled in light of the recent analysis of NC-RNA complexes [95], where both ZKs interact with a guanine. The largest contact area with RNA is observed for ZK1, whereas Trp37 in ZK2 presents the highest individual contribution. The following picture thus emerges: ZK1, due to its larger contact area both in terms of electrostatic and hydrophobic interactions with NA, is necessary in order to provide NA-NC complexes with a sufficient affinity to interact, and the proper interaction of Trp37 with a guanine is required to obtain a complex with an optimal affinity and specificity. Additionally, our analysis of the dynamic properties of NC-NA complexes (see Section 3.2) suggests that guanine binding in ZK2 leads to a conformational switch in the linker, reducing its flexibility at the level of the Gly35 hinge and thus potentially modifying the ZK1 interaction properties.

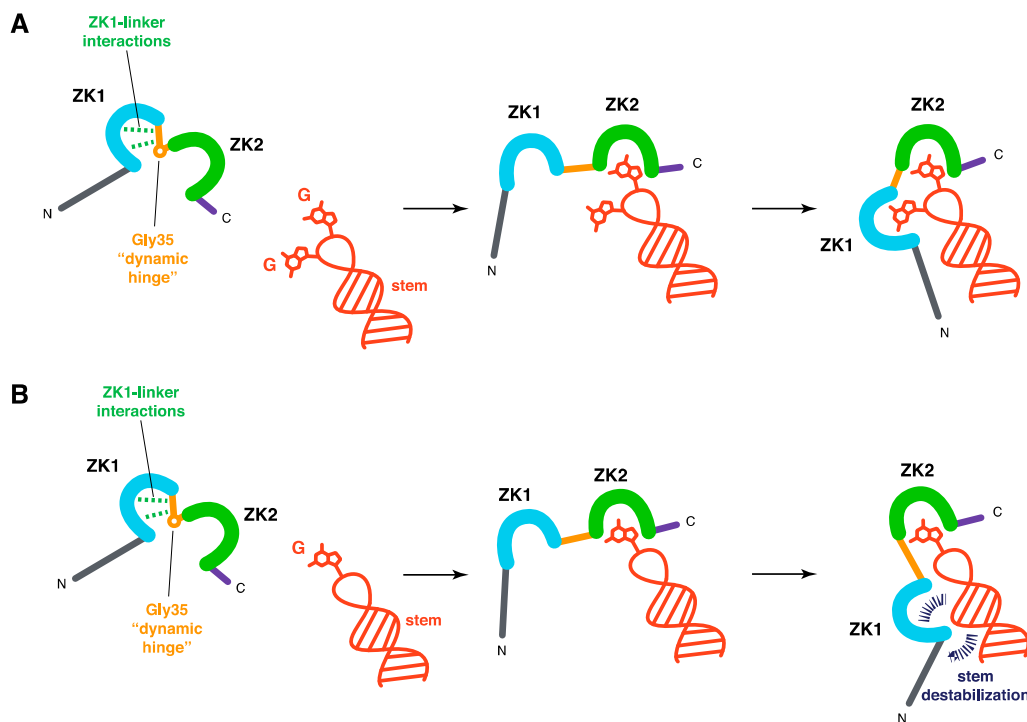
A last example of the difference between the two ZKs can be found in the studies that uncovered, upon deletion of one or two ZKs, an unexpected premature viral DNA synthesis starting in virus-producing cells and leading to the production of noninfectious particles containing a high level of viral DNA [124,139]. A substantial difference is observed in the quantity of viral DNA measured in virions depending on which ZK was deleted, with the most considerable effect marked for the deletion of ZK2 [124]. The fine analysis of these mechanisms points to the importance of NC and of its adjacent domains p1 and p6 in the precise timing of the budding events; the perturbation of this schedule leads to premature RT in infected cells [42,124,140]. These latter events involve interactions of NC and p6 with cellular factors (TSG101, ALIX) [14,15,42–44]. Interestingly, the deletion of essential parts of the packaging signal  $\Psi$  (SL1, SL3) also results in premature RT, like the deletion of ZK2, suggesting that direct contact between SL1-SL3 and ZK2 could be needed to restrict the premature RT process and trigger the correct succession of events leading to the budding of virions [15]. These data thus suggest a specific recognition between NC and gRNA in the control of the budding process, complementing the role of cellular proteins of the ESCRT family, with the optimal recognition requiring the presence of ZK2 [14,42].

Interestingly, a recent study examining the contribution of each ZK to the accumulation of Gag-gRNA complexes at the plasma membrane identified ZK2 as the major contributor in this task, underlining one more time the importance of ZK2 in the recruitment of the ribonucleoprotein complexes involved in the assembly process [125].

#### 4.2. Conclusions about the Functions of the Two ZKs

In conclusion, the two ZKs have different functions in the processes involving NC-NA interactions, namely the selective packaging of the gRNA, the NC chaperone activity and the control of the late RT process in the final steps of the cell infection. ZK1 appears necessary for high-affinity binding and is more critical in the chaperone activity than ZK2. Trp37 in ZK2 is crucial for the specific recognition of unpaired guanine. Notably, NC (1–35) appears necessary and sufficient to achieve most of the essential functions of NC; however, in NCp7, the mutation of Trp37 leads to a loss of these functions. The dynamics of the two ZKs are also different, since the interactions between the linker and ZK1 and the variations in flexibility along the linker render the motions of the two ZKs asymmetric. The literature is thus entirely consistent with a two-step model, in which Trp37 in ZK2 is first required to bind one unpaired guanine, followed by the binding of ZK1 to NA to produce a more stable complex and to allow NA stem destabilization if needed (Figure 4).

This model assumes that, in NCp7, ZK1 is not as accessible as ZK2 for RNA binding. The ZK1-linker contacts would prevent the ZK1 domain from interacting with unpaired guanines, whereas the ZK2 domain would be more accessible and competent to interact with unpaired guanines. In contrast, ZK1, with its sizeable hydrophobic plateau, would be able to destabilize the double-stranded regions adjacent to the guanines bound by ZK2. NMR data support the notion that the two ZKs are not equivalent and present different hydrodynamic properties. Moreover, small molecules able to eject zinc ions preferentially target ZK2 rather than ZK1 [106,141–144]. Similarly, alkylating agents, such as N-ethylmaleimide, react preferentially with the thiolates of the cysteines in ZK2. Interestingly, for N-ethylmaleimide, it was shown that the reactivity is different for the isolated ZKs, with a higher reactivity for isolated ZK1 compared to ZK2, in contrast to the results observed within the native NCp7 [145]. The linker residues and the internal dynamics of NC therefore regulate the different functions of the two ZKs that are required for an optimal chaperone activity.



**Figure 4.** Schematic representation of the task specialization of each ZK upon binding to NA. The NA fragment is in red, NC ZKs are in cyan (ZK1) and green (ZK2), NC Nter and Cter tails are in grey and purple, respectively, and the inter-ZK linker is in orange. In the NA-free form of NC, interactions between the more rigid part of the linker and ZK1 are represented with dashed lines in green. In addition, within the linker, the Gly35 dynamic hinge is shown (see paragraph 3.2 and Figure 3). This dynamic hinge is not retained in NA-bound forms of NC, Gly35 being directly involved in the guanine recognition within ZK2 (see Figure 2C). **(A)** Situation 1, with two accessible guanines. The ZK1-linker interactions render ZK1 less accessible than ZK2. Thus, ZK2 binds first to an accessible guanine, and ZK1 binds the remaining guanine. The stem is not destabilized. **(B)** Situation 2, with only one accessible guanine. Similarly, ZK2 binds first to the unique accessible guanine, and ZK1 remains free to contact the stem via its large hydrophobic platform. The stem is destabilized (in dark blue).

## 5. Conclusions

NCp7, the final maturation state of the NC domain, is the most studied form of this domain since the discovery of HIV-1. It is now clear that the two ZKs are not equivalent, due to their position in the NC sequence, the interactions between the linker and ZK1, and the variations in flexibility along the linker. Recent studies demonstrate that the NC domain of Gag, NCp15 and NCp9 behaves exactly the same as in the mature form. However, the interactions between p1–p6 and the ZKs modulate the dynamics of the ZKs and finely tune the interaction of the NC domain with NAs. These interactions can explain how the NC domain can change its NA-binding preferences during the assembly and budding of virions.

**Funding:** This research was funded by ANRS and SIDACTION, A.M. was the recipient of a SIDACTION fellowship to finance her last year of Ph.D.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fossen, T.; Wray, V.; Bruns, K.; Rachmat, J.; Henklein, P.; Tessmer, U.; Maczurek, A.; Klinger, P.; Schubert, U. Solution structure of the human immunodeficiency virus type 1 p6 protein. *J. Biol. Chem.* **2005**, *280*, 42515–42527. [[CrossRef](#)] [[PubMed](#)]

2. Mouhand, A.; Belfetmi, A.; Catala, M.; Larue, V.; Zargarian, L.; Brachet, F.; Gorelick, R.J.; Van Heijenoort, C.; Mirambeau, G.; Barraud, P.; et al. Modulation of the HIV nucleocapsid dynamics finely tunes its RNA-binding properties during virion genesis. *Nucleic Acids Res.* **2018**, *46*, 9699–9710. [[CrossRef](#)] [[PubMed](#)]
3. Kutluay, S.B.; Bieniasz, P.D. Analysis of the initiating events in HIV-1 particle assembly and genome packaging. *PLoS Pathog.* **2010**, *6*. [[CrossRef](#)]
4. Jouvenet, N.; Simon, S.M.; Bieniasz, P.D. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 19114–19119. [[CrossRef](#)]
5. Kuzembayeva, M.; Dilley, K.; Sardo, L.; Hu, W.S. Life of psi: How full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology* **2014**, *454*, 362–370. [[CrossRef](#)]
6. Lu, K.; Heng, X.; Summers, M.F. Structural determinants and mechanism of HIV-1 genome packaging. *J. Mol. Biol.* **2011**, *410*, 609–633. [[CrossRef](#)] [[PubMed](#)]
7. Lu, K.; Heng, X.; Garyu, L.; Monti, S.; Garcia, E.L.; Kharytonchyk, S.; Dorjsuren, B.; Kulandaivel, G.; Jones, S.; Hiremath, A.; et al. NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. *Science* **2011**, *334*, 242–245. [[CrossRef](#)]
8. Ferrer, M.; Clerté, C.; Chamontin, C.; Basyuk, E.; Lainé, S.; Hottin, J.; Bertrand, E.; Margeat, E.; Mougél, M. Imaging HIV-1 RNA dimerization in cells by multicolor super-resolution and fluctuation microscopies. *Nucleic Acids Res.* **2016**, *44*, 7922–7934. [[CrossRef](#)]
9. Rein, A. RNA Packaging in HIV. *Trends Microbiol.* **2019**, *27*, 715–723. [[CrossRef](#)]
10. Aldovini, A.; Young, R.A. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J. Virol.* **1990**, *64*, 1920–1926. [[CrossRef](#)]
11. Berkowitz, R.D.; Luban, J.; Goff, S.P. Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. *J. Virol.* **1993**, *67*, 7190–7200. [[CrossRef](#)] [[PubMed](#)]
12. Gorelick, R.J.; Nigida, S.M.; Bess, J.W.; Arthur, L.O.; Henderson, L.E.; Rein, A. Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* **1990**, *64*, 3207–3211. [[CrossRef](#)]
13. Konvalinka, J.; Kräusslich, H.G.; Müller, B. Retroviral proteases and their roles in virion maturation. *Virology* **2015**, *479*, 403–417. [[CrossRef](#)] [[PubMed](#)]
14. Freed, E.O. HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* **2015**, *13*, 484–496. [[CrossRef](#)] [[PubMed](#)]
15. Bendjennat, M.; Saffarian, S. The Race against Protease Activation Defines the Role of ESCRTs in HIV Budding. *PLoS Pathog.* **2016**, *12*, 1–25. [[CrossRef](#)]
16. Tisne, C. Structural Bases of the Annealing of Primer Lys tRNA to the HIV-1 Viral RNA. *Curr. HIV Res.* **2005**, *3*, 147–156. [[CrossRef](#)]
17. Tisné, C.; Roques, B.P.; Dardel, F. Heteronuclear NMR studies of the interaction of tRNA<sup>Lys</sup> with HIV-1 nucleocapsid protein. *J. Mol. Biol.* **2001**, *306*, 443–454. [[CrossRef](#)] [[PubMed](#)]
18. Tisné, C.; Roques, B.P.; Dardel, F. The Annealing Mechanism of HIV-1 Reverse Transcription Primer onto the Viral Genome. *J. Biol. Chem.* **2004**, *279*, 3588–3595. [[CrossRef](#)] [[PubMed](#)]
19. Barraud, P.; Gaudin, C.; Dardel, F.; Tisné, C. New insights into the formation of HIV-1 reverse transcription initiation complex. *Biochimie* **2007**, *89*, 1204–1210. [[CrossRef](#)] [[PubMed](#)]
20. Belfetmi, A.; Zargarian, L.; Tisné, C.; Sleiman, D.; Morellet, N.; Lescop, E.; Maskri, O.; René, B.; Mély, Y.; Fossé, P.; et al. Insights into the mechanisms of RNA secondary structure destabilization by the HIV-1 nucleocapsid protein. *RNA* **2016**, *22*, 506–517. [[CrossRef](#)] [[PubMed](#)]
21. Sleiman, D.; Bernacchi, S.; Guerrero, S.X.; Brachet, F.; Larue, V.; Paillart, J.C.; Tisné, C. Characterization of RNA binding and chaperoning activities of HIV-1 Vif protein Importance of the C-terminal unstructured tail. *RNA Biol.* **2014**, *11*, 906–920. [[CrossRef](#)] [[PubMed](#)]
22. Sleiman, D.; Barraud, P.; Brachet, F.; Tisne, C. The Interaction between tRNA<sup>Lys</sup>3 and the Primer Activation Signal Deciphered by NMR Spectroscopy. *PLoS ONE* **2013**, *8*. [[CrossRef](#)] [[PubMed](#)]
23. Levin, J.G.; Guo, J.; Rouzina, I.; Musier-Forsyth, K. Nucleic Acid Chaperone Activity of HIV-1 Nucleocapsid Protein: Critical Role in Reverse Transcription and Molecular Mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* **2005**, *80*, 217–286. [[CrossRef](#)] [[PubMed](#)]
24. Levin, J.G.; Mitra, M.; Mascarenhas, A.; Musier-Forsyth, K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol.* **2010**, *7*, 754–774. [[CrossRef](#)]

25. Morellet, N.; Jullian, N.; De Rocquigny, H.; Maignet, 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. [[CrossRef](#)] [[PubMed](#)]
26. Coren, L.V.; Thomas, J.A.; Chertova, E.; Sowder, R.C.; Gagliardi, T.D.; Gorelick, R.J.; Ott, D.E. Mutational Analysis of the C-Terminal Gag Cleavage Sites in Human Immunodeficiency Virus Type 1. *J. Virol.* **2007**, *81*, 10047–10054. [[CrossRef](#)]
27. Kafaie, J.; Dolatshahi, M.; Ajamian, L.; Song, R.; Mouland, A.J.; Rouiller, I.; Laughrea, M. Role of capsid sequence and immature nucleocapsid proteins p9 and p15 in Human Immunodeficiency Virus type 1 genomic RNA dimerization. *Virology* **2009**, *385*, 233–244. [[CrossRef](#)]
28. Lee, S.-K.; Harris, J.; Swanstrom, R. A Strongly Transdominant Mutation in the Human Immunodeficiency Virus Type 1 gag Gene Defines an Achilles Heel in the Virus Life Cycle. *J. Virol.* **2009**, *83*, 8536–8543. [[CrossRef](#)]
29. Deshmukh, L.; Tugarinov, V.; Louis, J.M.; Clore, G.M. Binding kinetics and substrate selectivity in HIV-1 protease–Gag interactions probed at atomic resolution by chemical exchange NMR. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E9855–E9862. [[CrossRef](#)]
30. Alfadhli, A.; Still, A.; Barklis, E. Analysis of Human Immunodeficiency Virus Type 1 Matrix Binding to Membranes and Nucleic Acids. *J. Virol.* **2009**, *83*, 12196–12203. [[CrossRef](#)]
31. Kutluay, S.B.; Zang, T.; Blanco-Melo, D.; Powell, C.; Jannain, D.; Errando, M.; Bieniasz, P.D. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell* **2014**, *159*, 1096–1109. [[CrossRef](#)] [[PubMed](#)]
32. Gaines, C.R.; Tkacik, E.; Rivera-Oven, A.; Somani, P.; Achimovich, A.; Alabi, T.; Zhu, A.; Getachew, N.; Yang, A.L.; McDonough, M.; et al. HIV-1 Matrix Protein Interactions with tRNA: Implications for Membrane Targeting. *J. Mol. Biol.* **2018**, *430*, 2113–2127. [[CrossRef](#)] [[PubMed](#)]
33. Bernacchi, S.; Abd El-Wahab, E.W.; Dubois, N.; Hijnen, M.; Smyth, R.P.; Mak, J.; Marquet, R.; Paillart, J.C. HIV-1 Pr55Gag binds genomic and spliced RNAs with different affinity and stoichiometry. *RNA Biol.* **2017**, *14*, 90–103. [[CrossRef](#)] [[PubMed](#)]
34. Paillart, J.C.; Shehu-Xhilaga, M.; Marquet, R.; Mak, J. Dimerization of retroviral RNA genomes: An inseparable pair. *Nat. Rev. Microbiol.* **2004**, *2*, 461–472. [[CrossRef](#)]
35. El-Wahab, E.W.A.; Smyth, R.P.; Mailler, E.; Bernacchi, S.; Vivet-Boudou, V.; Hijnen, M.; Jossinet, F.; Mak, J.; Paillart, J.C.; Marquet, R. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat. Commun.* **2014**, *5*. [[CrossRef](#)]
36. VerPlank, L.; Bouamr, F.; LaGrassa, T.J.; Agresta, B.; Kikonyogo, A.; Leis, J.; Carter, C.A. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55Gag. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7724–7729. [[CrossRef](#)]
37. Garrus, J.E.; Von Schwedler, U.K.; Pornillos, O.W.; Morham, S.G.; Zavitz, K.H.; Wang, H.E.; Wettstein, D.A.; Stray, K.M.; Côté, M.; Rich, R.L.; et al. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **2001**, *107*, 55–65. [[CrossRef](#)]
38. Votteler, J.; Sundquist, W.I. Virus budding and the ESCRT pathway. *Cell Host Microbe* **2013**, *14*, 232–241. [[CrossRef](#)]
39. Van Engelenburg, S.B.; Shtengel, G.; Sengupta, P.; Waki, K.; Jarnik, M.; Ablan, S.D.; Freed, E.O.; Hess, H.F.; Lippincott-Schwartz, J. Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. *Science* **2014**, *343*, 653–656. [[CrossRef](#)]
40. Strack, B.; Calistri, A.; Craig, S.; Popova, E.; Göttlinger, H.G. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell* **2003**, *114*, 689–699. [[CrossRef](#)]
41. Martin-Serrano, J.; Bieniasz, P.D. A Bipartite Late-Budding Domain in Human Immunodeficiency Virus Type 1. *J. Virol.* **2003**, *77*, 12373–12377. [[CrossRef](#)] [[PubMed](#)]
42. Chamontin, C.; Rassam, P.; Ferrer, M.; Racine, P.J.; Neyret, A.; Laine, S.; Milhiet, P.E.; Mougél, M. HIV-1 nucleocapsid and ESCRT-component Tsg101 interplay prevents HIV from turning into a DNA-containing virus. *Nucleic Acids Res.* **2015**, *43*, 336–347. [[CrossRef](#)] [[PubMed](#)]
43. Sette, P.; Dussupt, V.; Bouamr, F. Identification of the HIV-1 NC Binding Interface in Alix Bro1 Reveals a Role for RNA. *J. Virol.* **2012**, *86*, 11608–11615. [[CrossRef](#)] [[PubMed](#)]
44. El Meshri, S.E.; Boutant, E.; Mouhand, A.; Thomas, A.; Larue, V.; Richert, L.; Vivet-Boudou, V.; Mély, Y.; Tisné, C.; Muriaux, D.; et al. The NC domain of HIV-1 Gag contributes to the interaction of Gag with TSG101. *Biochim. Biophys. Acta Gen. Subj.* **2018**, *1862*, 1421–1431. [[CrossRef](#)]

45. De Rocquigny, H.; Gabus, C.; Vincent, A.; Fournie-Zaluski, M.C.; Roques, B.; Darlix, J.L. Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6472–6476. [[CrossRef](#)]
46. 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. [[CrossRef](#)]
47. Hargittai, M.R.S.; Mangla, A.T.; Gorelick, R.J.; Musier-Forsyth, K. HIV-1 nucleocapsid protein zinc finger structures induce tRNA<sup>Lys</sup> structural changes but are not critical for primer/template annealing. *J. Mol. Biol.* **2001**, *312*, 985–997. [[CrossRef](#)]
48. Ben-Artzi, H.; Shemesh, J.; Zeelon, E.; Amit, B.; Kleiman, L.; Gorecki, M.; Panet, A. Molecular analysis of the second template switch during reverse transcription of the HIV RNA template. *Biochemistry* **1996**, *35*, 10549–10557. [[CrossRef](#)]
49. Johnson, P.E.; Turner, R.B.; Wu, Z.R.; Hairston, L.; Guo, J.; Levin, J.G.; Summers, M.F. A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry* **2000**, *39*, 9084–9091. [[CrossRef](#)]
50. Peliska, J.A.; Balasubramanian, S.; Giedroc, D.P.; Benkovic, S.J. Recombinant HIV-1 Nucleocapsid Protein Accelerates HIV-1 Reverse Transcriptase Catalyzed DNA Strand Transfer Reactions and Modulates RNase H Activity. *Biochemistry* **1994**, *33*, 13817–13823. [[CrossRef](#)]
51. Mailler, E.; Bernacchi, S.; Marquet, R.; Paillart, J.C.; Vivet-Boudou, V.; Smyth, R.P. The life-cycle of the HIV-1 gag–RNA complex. *Viruses* **2016**, *8*, 248. [[CrossRef](#)] [[PubMed](#)]
52. Fisher, R.J.; Rein, A.; Fivash, M.; Urbaneja, M.A.; Casas-Finet, J.R.; Medaglia, M.; Henderson, L.E. Sequence-Specific Binding of Human Immunodeficiency Virus Type 1 Nucleocapsid Protein to Short Oligonucleotides. *J. Virol.* **1998**, *72*, 1902–1909. [[CrossRef](#)] [[PubMed](#)]
53. Mely, Y.; De Rocquigny, H.; Sorinas-Jimeno, M.; Keith, G.; Roques, B.P.; Marquet, R.; Gerard, D. Binding of the HIV-1 nucleocapsid protein to the primer tRNA<sup>3</sup>(Lys), in vitro, is essentially not specific. *J. Biol. Chem.* **1995**, *270*, 1650–1656. [[CrossRef](#)] [[PubMed](#)]
54. 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. [[CrossRef](#)] [[PubMed](#)]
55. 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. [[CrossRef](#)]
56. Vuilleumier, C.; Bombarda, E.; Morellet, N.; Gérard, D.; Roques, B.P.; Mély, Y. Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: A fluorescence study. *Biochemistry* **1999**, *38*, 16816–16825. [[CrossRef](#)]
57. Wilkinson, K.A.; Gorelick, R.J.; Vasa, S.M.; Guex, N.; Rein, A.; Mathews, D.H.; Giddings, M.C.; Weeks, K.M. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* **2008**, *6*, 883–899. [[CrossRef](#)]
58. Bourbigot, S.; Ramalanjaona, N.; Boudier, C.; Salgado, G.F.J.; Roques, B.P.; Mély, 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. [[CrossRef](#)]
59. 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. [[CrossRef](#)]
60. 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. [[CrossRef](#)]
61. Bazzi, A.; Zargarian, L.; Chaminade, F.; Boudier, C.; De Rocquigny, H.; René, B.; Mély, Y.; Fossé, P.; Mauffret, O. Structural insights into the cTAR DNA recognition by the HIV-1 nucleocapsid protein: Role of sugar deoxyriboses in the binding polarity of NC. *Nucleic Acids Res.* **2011**, *39*, 3903–3916. [[CrossRef](#)] [[PubMed](#)]
62. Ding, P.; Kharytonchyk, S.; Waller, A.; Mbaekwe, U.; Basappa, S.; Kuo, N.; Frank, H.M.; Quasney, C.; Kidane, A.; Swanson, C.; et al. Identification of the initial nucleocapsid recognition element in the HIV-1 RNA packaging signal. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 17737–17746. [[CrossRef](#)] [[PubMed](#)]

63. McCauley, M.J.; Rouzina, I.; Manthei, K.A.; Gorelick, R.J.; Musier-Forsyth, K.; Williams, M.C. Targeted binding of nucleocapsid protein transforms the folding landscape of HIV-1 TAR RNA. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13555–13560. [[CrossRef](#)] [[PubMed](#)]
64. Brown, J.D.; Kharytonchyk, S.; Chaudry, I.; Iyer, A.S.; Carter, H.; Becker, G.; Desai, Y.; Glang, L.; Choi, S.H.; Singh, K.; et al. Structural basis for transcriptional start site control of HIV-1 RNA fate. *Science* **2020**, *368*, 413–417. [[CrossRef](#)] [[PubMed](#)]
65. D'Souza, V.; Summers, M.F. How retroviruses select their genomes. *Nat. Rev. Microbiol.* **2005**, *3*, 643–655. [[CrossRef](#)] [[PubMed](#)]
66. Onafuwa-Nuga, A.; Telesnitsky, A. The Remarkable Frequency of Human Immunodeficiency Virus Type 1 Genetic Recombination. *Microbiol. Mol. Biol. Rev.* **2009**, *73*, 451–480. [[CrossRef](#)]
67. Dubois, N.; Khoo, K.K.; Ghossein, S.; Seissler, T.; Wolff, P.; McKinstry, W.J.; Mak, J.; Paillart, J.C.; Marquet, R.; Bernacchi, S. The C-terminal p6 domain of the HIV-1 Pr55Gag precursor is required for specific binding to the genomic RNA. *RNA Biol.* **2018**, *15*, 923–936. [[CrossRef](#)]
68. Smyth, R.P.; Despons, L.; Huili, G.; Bernacchi, S.; Hijnen, M.; Mak, J.; Jossinet, F.; Weixi, L.; Paillart, J.C.; Von Kleist, M.; et al. Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat. Methods* **2015**, *12*, 866–872. [[CrossRef](#)]
69. Bieniasz, P.; Telesnitsky, A. Multiple, switchable protein: RNA interactions regulate human immunodeficiency virus type 1 assembly. *Annu. Rev. Virol.* **2018**, *5*, 165–183. [[CrossRef](#)]
70. Heng, X.; Kharytonchyk, S.; Garcia, E.L.; Lu, K.; Divakaruni, S.S.; Lacotti, C.; Edme, K.; Telesnitsky, A.; Summers, M.F. Identification of a minimal region of the HIV-1 5'-leader required for RNA dimerization, NC binding, and packaging. *J. Mol. Biol.* **2012**, *417*, 224–239. [[CrossRef](#)]
71. Keane, S.C.; Summers, M.F. NMR studies of the structure and function of the HIV-1 5'-leader. *Viruses* **2016**, *8*, 338. [[CrossRef](#)] [[PubMed](#)]
72. Keane, S.C.; Heng, X.; Lu, K.; Kharytonchyk, S.; Ramakrishnan, V.; Carter, G.; Barton, S.; Hosic, A.; Florwick, A.; Santos, J.; et al. Structure of the HIV-1 RNA packaging signal. *Science* **2015**, *348*, 917–921. [[CrossRef](#)] [[PubMed](#)]
73. Kotar, A.; Foley, H.N.; Baughman, K.M.; Keane, S.C. Advanced approaches for elucidating structures of large RNAs using NMR spectroscopy and complementary methods. *Methods* **2020**, in press. [[CrossRef](#)] [[PubMed](#)]
74. Alvarado, L.J.; Leblanc, R.M.; Longhini, A.P.; Keane, S.C.; Jain, N.; Yildiz, Z.F.; Tolbert, B.S.; D'Souza, V.M.; Summers, M.F.; Kreutz, C.; et al. Regio-selective chemical-enzymatic synthesis of pyrimidine nucleotides facilitates RNA structure and dynamics studies. *ChemBioChem* **2014**, *15*, 1573–1577. [[CrossRef](#)]
75. Asadi-Atoi, P.; Barraud, P.; Tisne, C.; Kellner, S. Benefits of stable isotope labeling in RNA analysis. *Biol. Chem.* **2019**, *400*, 847–865. [[CrossRef](#)]
76. Darlix, J.L.; Gabus, C.; Nugeyre, M.T.; Clavel, F.; Barré-Sinoussi, F. Cis elements and Trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Mol. Biol.* **1990**, *216*, 689–699. [[CrossRef](#)]
77. Ooms, M.; Huthoff, H.; Russell, R.; Liang, C.; Berkhout, B. A Riboswitch Regulates RNA Dimerization and Packaging in Human Immunodeficiency Virus Type 1 Virions. *J. Virol.* **2004**, *78*, 10814–10819. [[CrossRef](#)]
78. Kharytonchyk, S.; Monti, S.; Smaldino, P.J.; Van, V.; Bolden, N.C.; Brown, J.D.; Russo, E.; Swanson, C.; Shuey, A.; Telesnitsky, A.; et al. Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13378–13383. [[CrossRef](#)]
79. Masuda, T.; Sato, Y.; Huang, Y.L.; Koi, S.; Takahata, T.; Hasegawa, A.; Kawai, G.; Kannagi, M. Fate of HIV-1 cDNA intermediates during reverse transcription is dictated by transcription initiation site of virus genomic RNA. *Sci. Rep.* **2015**, *5*, 1–15. [[CrossRef](#)]
80. Coffin, J.M.; Hughes, S.H.; Varmus, H. *Retroviruses*; Cold Spring Harbor Laboratory Press: Plainview, NY, USA, 1997; ISBN 978-0-87969-571-2.
81. Lever, A.M.L. HIV-1 RNA Packaging. In *Advances in Pharmacology*; Academic Press: Cambridge, MA, USA, 2007; Volume 55, pp. 1–32. ISBN 0123736102.
82. Lever, A.; Gottlinger, H.; Haseltine, W.; Sodroski, J. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol.* **1989**, *63*, 4085–4087. [[CrossRef](#)]
83. Clavel, F.; Orenstein, J.M. A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal particle morphology. *J. Virol.* **1990**, *64*, 5230–5234. [[CrossRef](#)] [[PubMed](#)]



84. Russell, R.S.; Hu, J.; Bériault, V.; Mouland, A.J.; Laughrea, M.; Kleiman, L.; Wainberg, M.A.; Liang, C. Sequences Downstream of the 5' Splice Donor Site Are Required for both Packaging and Dimerization of Human Immunodeficiency Virus Type 1 RNA. *J. Virol.* **2003**, *77*, 3891. [[CrossRef](#)]
85. McBride, M.S.; Panganiban, A.T. The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J. Virol.* **1996**, *70*, 2963–2973. [[CrossRef](#)] [[PubMed](#)]
86. Richardson, J.H.; Child, L.A.; Lever, A.M. Packaging of human immunodeficiency virus type 1 RNA requires cis-acting sequences outside the 5' leader region. *J. Virol.* **1993**, *67*, 3997–4005. [[CrossRef](#)]
87. Singh, G.; Rife, B.D.; Seufzer, B.; Salemi, M.; Rendahl, A.; Boris-Lawrie, K. Identification of conserved, primary sequence motifs that direct retrovirus RNA fate. *Nucleic Acids Res.* **2018**, *46*, 7366–7378. [[CrossRef](#)]
88. Zeffman, A.; Hassard, S.; Varani, G.; Lever, A. The major HIV-1 packaging signal is an extended bulged stem loop whose structure is altered on interaction with the Gag polyprotein. *J. Mol. Biol.* **2000**, *297*, 877–893. [[CrossRef](#)]
89. Campbell, S.; Rein, A. In Vitro Assembly Properties of Human Immunodeficiency Virus Type 1 Gag Protein Lacking the p6 Domain. *J. Virol.* **1999**, *73*, 2270–2279. [[CrossRef](#)]
90. Deshmukh, L.; Ghirlando, R.; Clorea, G.M. Conformation and dynamics of the Gag polyprotein of the human immunodeficiency virus 1 studied by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3374–3379. [[CrossRef](#)]
91. Wang, W.; Naiyer, N.; Mitra, M.; Li, J.; Williams, M.C.; Rouzina, I.; Gorelick, R.J.; Wu, Z.; Musier-Forsyth, K. Distinct nucleic acid interaction properties of HIV-1 nucleocapsid protein precursor NCp15 explain reduced viral infectivity. *Nucleic Acids Res.* **2014**, *42*, 7145–7159. [[CrossRef](#)]
92. Tanwar, H.S.; Khoo, K.K.; Garvey, M.; Waddington, L.; Leis, A.; Hijnen, M.; Velkov, T.; Dumsday, G.J.; McKinsty, W.J.; Mak, J. The thermodynamics of Pr55Gag-RNA interaction regulate the assembly of HIV. *PLoS Pathog.* **2017**, *13*. [[CrossRef](#)]
93. Morellet, N.; Déméné, H.; Teilleux, V.; Huynh-Dinh, T.; De Rocquigny, H.; Fournié-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. [[CrossRef](#)] [[PubMed](#)]
94. Spriggs, S.; Garyu, L.; Connor, R.; Summers, M.F. Potential intra- and intermolecular interactions involving the unique-5' region of the HIV-1 5'-UTR. *Biochemistry* **2008**, *47*, 13064–13073. [[CrossRef](#)] [[PubMed](#)]
95. Retureau, R.; Oguey, C.; Mauffret, O.; Hartmann, B. Structural Explorations of NCp7–Nucleic Acid Complexes Give Keys to Decipher the Binding Process. *J. Mol. Biol.* **2019**, *431*, 1966–1980. [[CrossRef](#)]
96. Athavale, S.S.; Ouyang, W.; McPike, M.P.; Hudson, B.S.; Borer, P.N. Effects of the nature and concentration of salt on the interaction of the HIV-1 nucleocapsid protein with SL3 RNA. *Biochemistry* **2010**, *49*, 3525–3533. [[CrossRef](#)] [[PubMed](#)]
97. Avilov, S.V.; Piemont, E.; Shvadchak, V.; de Rocquigny, H.; Mély, Y. Probing dynamics of HIV-1 nucleocapsid protein/target hexanucleotide complexes by 2-aminopurine. *Nucleic Acids Res.* **2008**, *36*, 885–896. [[CrossRef](#)]
98. Godet, J.; Mély, Y. Biophysical studies of the nucleic acid chaperone properties of the HIV-1 nucleocapsid protein. *RNA Biol.* **2010**, *7*, 687–699. [[CrossRef](#)]
99. Lee, B.M.; De Guzman, R.N.; Turner, B.G.; Tjandra, N.; Summers, M.F. Dynamical behavior of the HIV-1 nucleocapsid protein. *J. Mol. Biol.* **1998**, *279*, 633–649. [[CrossRef](#)]
100. Guo, J.; Wu, T.; Kane, B.F.; Johnson, D.G.; Henderson, L.E.; Gorelick, R.J.; Levin, J.G. Subtle Alterations of the Native Zinc Finger Structures Have Dramatic Effects on the Nucleic Acid Chaperone Activity of Human Immunodeficiency Virus Type 1 Nucleocapsid Protein. *J. Virol.* **2002**, *76*, 4370–4378. [[CrossRef](#)]
101. Wu, H.; Mitra, M.; Nauffer, M.N.; McCauley, M.J.; Gorelick, R.J.; Rouzina, I.; Musier-Forsyth, K.; Williams, M.C. Differential contribution of basic residues to HIV-1 nucleocapsid protein's nucleic acid chaperone function and retroviral replication. *Nucleic Acids Res.* **2014**, *42*, 2525–2537. [[CrossRef](#)]
102. Webb, J.A.; Jones, C.P.; Parent, L.J.; Rouzina, I.; Musier-Forsyth, K. Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: Implications for viral genomic RNA packaging. *RNA* **2013**, *19*, 1078–1088. [[CrossRef](#)]
103. Gamsjaeger, R.; Kariawasam, R.; Gimenez, A.X.; Touma, C.; McIlwain, E.; Bernardo, R.E.; Shepherd, N.E.; Ataide, S.F.; Dong, Q.; Richard, D.J.; et al. The structural basis of DNA binding by the single-stranded DNA-binding protein from *Sulfolobus solfataricus*. *Biochem. J.* **2015**, *465*, 337–346. [[CrossRef](#)] [[PubMed](#)]
104. Zargarian, L.; Tisné, C.; Barraud, P.; Xu, X.; Morellet, N.; René, B.; Mély, Y.; Fossé, P.; Mauffret, O. Dynamics of linker residues modulate the nucleic acid binding properties of the HIV-1 nucleocapsid protein zinc fingers. *PLoS ONE* **2014**, *9*, 27–30. [[CrossRef](#)]

105. Deshmukh, L.; Schwieters, C.D.; Grishaev, A.; Clore, G.M. Quantitative Characterization of Configurational Space Sampled by HIV-1 Nucleocapsid Using Solution NMR, X-ray Scattering and Protein Engineering. *ChemPhysChem* **2016**, *17*, 1548–1552. [[CrossRef](#)] [[PubMed](#)]
106. Deshmukh, L.; Tugarinov, V.; Appella, D.H.; Clore, G.M. Targeting a Dark Excited State of HIV-1 Nucleocapsid by Antiretroviral Thioesters Revealed by NMR Spectroscopy. *Angew. Chemie Int. Ed.* **2018**, *57*, 2687–2691. [[CrossRef](#)] [[PubMed](#)]
107. El Khoury, L.; Célerse, F.; Lagardère, L.; Jolly, L.H.; Derat, E.; Hobaika, Z.; Maroun, R.G.; Ren, P.; Bouaziz, S.; Gresh, N.; et al. Reconciling NMR Structures of the HIV-1 Nucleocapsid Protein NCp7 Using Extensive Polarizable Force Field Free-Energy Simulations. *J. Chem. Theory Comput.* **2020**, *16*, 2013–2020. [[CrossRef](#)]
108. Ramboarina, S.; Srividya, N.; Atkinson, R.A.; Morellet, N.; Roques, B.P.; Lefèvre, J.F.; Mély, 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. [[CrossRef](#)]
109. Godet, J.; Ramalanjaona, N.; Sharma, K.K.; Richert, L.; De Rocquigny, H.; Darlix, J.L.; Duportail, G.; Mély, Y. Specific implications of the HIV-1 nucleocapsid zinc fingers in the annealing of the primer binding site complementary sequences during the obligatory plus strand transfer. *Nucleic Acids Res.* **2011**, *39*, 6633–6645. [[CrossRef](#)]
110. Dorfman, T.; Luban, J.; Goff, S.P.; Haseltine, W.A.; Göttlinger, H.G. Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* **1993**, *67*, 6159–6169. [[CrossRef](#)]
111. Guo, J.; Wu, T.; Anderson, J.; Kane, B.F.; Johnson, D.G.; Gorelick, R.J.; Henderson, L.E.; Levin, J.G. Zinc Finger Structures in the Human Immunodeficiency Virus Type 1 Nucleocapsid Protein Facilitate Efficient Minus- and Plus-Strand Transfer. *J. Virol.* **2000**, *74*, 8980–8988. [[CrossRef](#)]
112. Gorelick, R.J.; Chabot, D.J.; Ott, D.E.; Gagliardi, T.D.; Rein, A.; Henderson, L.E.; Arthur, L.O. Genetic analysis of the zinc finger in the Moloney murine leukemia virus nucleocapsid domain: Replacement of zinc-coordinating residues with other zinc-coordinating residues yields noninfectious particles containing genomic RNA. *J. Virol.* **1996**, *70*, 2593–2597. [[CrossRef](#)]
113. Gorelick, R.J.; Gagliardi, T.D.; Bosche, W.J.; Wiltrout, T.A.; Coren, L.V.; Chabot, D.J.; Lifson, J.D.; Henderson, L.E.; Arthur, L.O. Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: Characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. *Virology* **1999**, *256*, 92–104. [[CrossRef](#)]
114. Tanchou, V.; Decimo, D.; Péchoux, C.; Lener, D.; Rogemond, V.; Berthoux, L.; Ottmann, M.; Darlix, J.-L. Role of the N-Terminal Zinc Finger of Human Immunodeficiency Virus Type 1 Nucleocapsid Protein in Virus Structure and Replication. *J. Virol.* **1998**, *72*, 4442–4447. [[CrossRef](#)] [[PubMed](#)]
115. Gorelick, R.J.; Chabot, D.J.; Rein, A.; Henderson, L.E.; Arthur, L.O. The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J. Virol.* **1993**, *67*, 4027–4036. [[CrossRef](#)] [[PubMed](#)]
116. Heath, M.J.; Derebail, S.S.; Gorelick, R.J.; DeStefano, J.J. Differing roles of the N- and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-enhanced nucleic acid annealing. *J. Biol. Chem.* **2003**, *278*, 30755–30763. [[CrossRef](#)] [[PubMed](#)]
117. Williams, M.C.; Gorelick, R.J.; Musier-Forsyth, K. Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8614–8619. [[CrossRef](#)]
118. Mitra, M.; Wang, W.; Vo, M.N.; Rouzina, I.; Barany, G.; Musier-Forsyth, K. The N-terminal zinc finger and flanking basic domains represent the minimal region of the human immunodeficiency virus type-1 nucleocapsid protein for targeting chaperone function. *Biochemistry* **2013**, *52*, 8226–8236. [[CrossRef](#)]
119. 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. [[CrossRef](#)]
120. Wu, H.; Mitra, M.; McCauley, M.J.; Thomas, J.A.; Rouzina, I.; Musier-Forsyth, K.; Williams, M.C.; Gorelick, R.J. Aromatic residue mutations reveal direct correlation between HIV-1 nucleocapsid protein's nucleic acid chaperone activity and retroviral replication. *Virus Res.* **2013**, *171*, 263–277. [[CrossRef](#)]
121. Kafaie, J.; Song, R.; Abrahamyan, L.; Moulant, A.J.; Laughrea, M. Mapping of nucleocapsid residues important for HIV-1 genomic RNA dimerization and packaging. *Virology* **2008**, *375*, 592–610. [[CrossRef](#)]

122. Poon, D.T.; Wu, J.; Aldovini, A. Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity. *J. Virol.* **1996**, *70*, 6607–6616. [[CrossRef](#)]
123. 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. [[CrossRef](#)] [[PubMed](#)]
124. Houzet, L.; Morichaud, Z.; Didierlaurent, L.; Muriaux, D.; Darlix, J.L.; Mougél, M. Nucleocapsid mutations turn HIV-1 into a DNA-containing virus. *Nucleic Acids Res.* **2008**, *36*, 2311–2319. [[CrossRef](#)] [[PubMed](#)]
125. Boutant, E.; Bonzi, J.; Anton, H.; Nasim, M.B.; Cathagne, R.; Réal, E.; Dujardin, D.; Carl, P.; Didier, P.; Paillart, J.C.; et al. Zinc Fingers in HIV-1 Gag Precursor Are Not Equivalent for gRNA Recruitment at the Plasma Membrane. *Biophys. J.* **2020**, *119*, 419–433. [[CrossRef](#)] [[PubMed](#)]
126. Zhang, Y.; Barklis, E. Nucleocapsid protein effects on the specificity of retrovirus RNA encapsidation. *J. Virol.* **1995**, *69*, 5716–5722. [[CrossRef](#)]
127. Wang, S.-W.; Noonan, K.; Aldovini, A. Nucleocapsid-RNA Interactions Are Essential to Structural Stability but Not to Assembly of Retroviruses. *J. Virol.* **2004**, *78*, 716–723. [[CrossRef](#)]
128. Darlix, J.L.; Lapadat-Tapolsky, M.; De Rocquigny, H.; Roques, B.P. First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J. Mol. Biol.* **1995**, *254*, 523–537. [[CrossRef](#)]
129. Sleiman, D.; Goldschmidt, V.; Barraud, P.; Marquet, R.; Paillart, J.C.; Tisné, C. Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. *Virus Res.* **2012**, *169*, 324–339. [[CrossRef](#)]
130. Lapadat-tapolsky, M.; De Rocquigny, H.; Van Gent, D.; Roques, B.; Plasterk, R.; Darlix, J.-L. Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. *Nucleic Acids Res.* **1993**, *21*, 831–839. [[CrossRef](#)]
131. Beltz, H.; Clauss, C.; Piémont, E.; Ficheux, D.; Gorelick, R.J.; Roques, B.; Gabus, C.; Darlix, J.L.; De Rocquigny, H.; Mély, 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. [[CrossRef](#)]
132. Fisher, R.J.; Fivash, M.J.; Stephen, A.G.; Hagan, N.A.; Shenoy, S.R.; Medaglia, M.V.; Smith, L.R.; Worthy, K.M.; Simpson, J.T.; Shoemaker, R.; et al. Complex interactions of HIV-1 nucleocapsid protein with oligonucleotides. *Nucleic Acids Res.* **2006**, *34*, 472–484. [[CrossRef](#)]
133. Cruceanu, M.; Gorelick, R.J.; Musier-Forsyth, K.; Rouzina, I.; Williams, M.C. Rapid Kinetics of Protein-Nucleic Acid Interaction is a Major Component of HIV-1 Nucleocapsid Protein's Nucleic Acid Chaperone Function. *J. Mol. Biol.* **2006**, *363*, 867–877. [[CrossRef](#)]
134. Ren, W.; Ji, D.; Xu, X. Metal cofactor modulated folding and target recognition of HIV-1 NCp7. *PLoS ONE* **2018**, *13*, 1–21. [[CrossRef](#)]
135. Narayanan, N.; Gorelick, R.J.; DeStefano, J.J. Structure/function mapping of amino acids in the N-terminal zinc finger of the human immunodeficiency virus type 1 nucleocapsid protein: Residues responsible for nucleic acid helix destabilizing activity. *Biochemistry* **2006**, *45*, 12617–12628. [[CrossRef](#)] [[PubMed](#)]
136. Williams, M.C.; Rouzina, I.; Wenner, J.R.; Gorelick, R.J.; Musier-Forsyth, K.; Bloomfield, V.A. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6121–6126. [[CrossRef](#)] [[PubMed](#)]
137. Remy, E.; De Rocquigny, H.; Petitjean, P.; Muriaux, D.; Theilleux, V.; Paoletti, J.; Roques, B.P. The annealing of tRNA<sup>3</sup>(Lys) to human immunodeficiency virus type 1 primer binding site is critically dependent on the NCp7 zinc fingers structure. *J. Biol. Chem.* **1998**, *273*, 4819–4822. [[CrossRef](#)] [[PubMed](#)]
138. Maki, A.H.; Ozarowski, A.; Misra, A.; Urbaneja, M.A.; Casas-Finet, J.R. Phosphorescence and optically detected magnetic resonance of HIV-1 nucleocapsid protein complexes with stem-loop sequences of the genomic  $\psi$ -recognition element. *Biochemistry* **2001**, *40*, 1403–1412. [[CrossRef](#)] [[PubMed](#)]
139. Didierlaurent, L.; Houzet, L.; Morichaud, Z.; Darlix, J.L.; Mougél, M. The conserved N-terminal basic residues and zinc-finger motifs of HIV-1 nucleocapsid restrict the viral cDNA synthesis during virus formation and maturation. *Nucleic Acids Res.* **2008**, *36*, 4745–4753. [[CrossRef](#)]
140. Racine, P.J.; Chamontin, C.; De Rocquigny, H.; Bernacchi, S.; Paillart, J.C.; Mougél, M. Requirements for nucleocapsid-mediated regulation of reverse transcription during the late steps of HIV-1 assembly. *Sci. Rep.* **2016**, *6*, 1–11. [[CrossRef](#)]

141. Sancineto, L.; Iraci, N.; Tabarrini, O.; Santi, C. NCp7: Targeting a multitasking protein for next-generation anti-HIV drug development part 1: Covalent inhibitors. *Drug Discov. Today* **2018**, *23*, 260–271. [[CrossRef](#)]
142. Basrur, V.; Song, Y.; Mazur, S.J.; Higashimoto, Y.; Turpin, J.A.; Rice, W.G.; Inman, J.K.; Appella, E. Inactivation of HIV-1 nucleocapsid protein p7 by pyridinioalkanoyl thioesters. Characterization of reaction products and proposed mechanism of action. *J. Biol. Chem.* **2000**, *275*, 14890–14897. [[CrossRef](#)]
143. Maynard, A.T.; Covell, D.G. Reactivity of zinc finger cores: Analysis of protein packing and electrostatic screening. *J. Am. Chem. Soc.* **2001**, *123*, 1047–1058. [[CrossRef](#)] [[PubMed](#)]
144. Miller Jenkins, L.M.; Hara, T.; Durell, S.R.; Hayashi, R.; Inman, J.K.; Piquemal, J.P.; Gresh, N.; Appella, E. Specificity of acyl transfer from 2-mercaptobenzamide thioesters to the HIV-1 nucleocapsid protein. *J. Am. Chem. Soc.* **2007**, *129*, 11067–11078. [[CrossRef](#)] [[PubMed](#)]
145. Chertova, E.N.; Kane, B.P.; McGrath, C.; Johnson, D.G.; Sowder, R.C.; Arthur, L.O.; Henderson, L.E. Probing the topography of HIV-1 nucleocapsid protein with the alkylating agent N-ethylmaleimide. *Biochemistry* **1998**, *37*, 17890–17897. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).