doi:10.1016/j.jmb.2006.08.065 J. Mol. Biol. (2006) 364, 496–511

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Transmission Electron Microscopy Reveals an Optimal HIV-1 Nucleocapsid Aggregation with Single-stranded Nucleic Acids and the Mature HIV-1 Nucleocapsid Protein

Gilles Mirambeau[1,2]*†, Sébastien Lyonnais[1]†, Dominique Coulaud[1] Laurence Hameau[1], Sophie Lafosse[1], Josette Jeusset[1] Anthony Justome[1], Etienne Delain[1], Robert J. Gorelick[3] and Eric Le Cam[1]*

[1]*Laboratoire de Microscopie Moléculaire et Cellulaire CNRS UMR 8126 Institut Gustave Roussy 94805 Villejuif, France*

[2]*Division de Biochimie UFR des Sciences de la Vie Université Pierre et Marie Curie-Paris6, Place Jussieu 75005 Paris, France*

[3]*AIDS Vaccine Program Basic Research Program SAIC-Frederick, Inc. NCI-Frederick, Frederick MD 21702-1201, USA*

**Corresponding authors*

HIV-1 nucleocapsid protein (NCp7) condenses the viral RNA within the mature capsid. In a capsid-free system, NCp7 promotes an efficient mechanism of aggregation with both RNA and DNA. Here, we show an analysis of these macromolecular complexes by dark-field imaging using transmission electron microscopy. Thousands of mature NCp7 proteins co-aggregate with hundreds of single-stranded circular DNA molecules (ssDNA) within minutes, as observed with poly(rA). These co-aggregates are highly stable but dynamic structures, as they dissociate under harsh conditions, and after addition of potent ssDNA or NCp7 competitive ligands. The N-terminal domain and zinc fingers of NCp7 are both required for efficient association. Addition of magnesium slightly increases the avidity of NCp7 for ssDNA, while it strongly inhibits co-aggregation with relaxed circular double-stranded DNA (dsDNA). This DNA selectivity is restricted to mature NCp7, compared to its precursors NCp15 and NCp9. Moreover, for NCp15, the linkage of NCp7 with the Gag C-terminal p6-peptide provokes a deficiency in ssDNA aggregation, but results in DNA spreading similar to prototypical SSB proteins. Finally, this co-aggregation is discussed in a dynamic architectural context with regard to the mature HIV-1 nucleocapsid. On the basis of the present data, we propose that condensation of encapsidated RNA requires the C-terminal processing of NCp. Subsequently, disassembly of the nucleocapsid should be favoured once dsDNA is produced by HIV-1 reverse transcriptase.

Published by Elsevier Ltd.

Keywords: TEM; HIV NCp; nucleic acids; nucleoprotein complex; micro-sedimentation

† G.M. and S.L. contributed equally to this work.

Present address: S. Lyonnais, Laboratoire d'Electronique Moléculaire, DRECAM/SPEC, CEA-Saclay, 91191 Gif sur Yvette, France.

Abbreviations used: HIV, human immunodeficiency virus; NCp, nucleocapsid protein; ss, single-stranded; ds, double-stranded; EM, electron microscopy; TEM, transmission electron microscopy; gp32, T4-encoded gene *32* protein; SSB, single-stranded binding; G4 DNA, parallel DNA quadruplex; RT, reverse transcriptase; IN, integrase; Vpr, viral protein R; PR, protease; RTC, reverse transcription complex; PIC, preintegration complex; Mo-MuLV, Moloney murine leukaemia virus.

E-mail addresses of the corresponding authors: gilles.mirambeau@free.fr; elecam@igr.fr

0022-2836/$ - see front matter. Published by Elsevier Ltd.

Introduction

In infectious human immunodeficiency virus type 1 (HIV-1) virions, the nucleocapsid architecture confined within the cone-shaped capsid contains two copies of the genomic RNA coated by the mature nucleocapsid protein (NCp7). This highly condensed ribonucleoprotein complex provides both the structures and functionality required for infectivity.[1,2] NCp7, characterized as an RNA/DNA-binding and nucleic-acids chaperone protein, contributes to both the stabilization and structural rearrangements of the viral genome within virions as well as during reverse transcription within infected cells.[3,4]

HIV-1 genomic RNA is incorporated into nascent virions through interactions with the nucleocapsid (NC) domain of Gag polyproteins. During viral maturation, the viral protease (PR) catalyzes a sequential and ordered processing of Gag, which drives nucleocapsid assembly and its confinement within the viral core.[5] Physical separation of the capsid and nucleocapsid domains of Gag (MA-CA-p2-NC-p1-p6) occurs *via* a rapid proteolytic cleavage at the p2-NC site, which produces a transient NC species, NCp15 (NC-p1-p6).[6] Thereafter, p6 removal gives way to the transient accumulation of a 71 amino acid residue protein, NCp9 (NC-p1), while the ultimate p1 removal liberates the mature NCp7 (55 amino acid residues).[5] Complete processing of NCp, along with the presence of the p1-sequence, reverse transcriptase (RT), and integrase (IN), are critical for both nucleocapsid condensation and stabilization of dimeric viral RNA.[7–10] However, the mechanism by which nucleocapsid condensation occurs is not clearly understood. After maturation, the HIV-1 conical capsid delineates an internal compartment that contains the nucleocapsid composed of NCp7 (between 1500 and 5000 copies) bound to a $tRNA^{Lys,3}$-primed dimer of genomic RNA (9250 nt per RNA monomer). Other viral components are incorporated: RT (~100 dimeric copies), IN (~200 monomeric copies), viral protein R (Vpr) (~700 copies), Nef, PR,[11–13] and some cellular components, notably small RNA species and a subset of cellular proteins.[14–16] The Gag-associated NC domain as well as mature NCp7 have been shown to participate in protein–protein interactions with RT, Vpr and cellular proteins, such as actin, topoisomerase I and APOBEC3G.[17–20] During reverse transcription and migration to the host nucleus, stability of the HIV-1 genome and subsequent reverse transcription products requires the presence of functional NCp7,[3] as HIV-1 mutants with an altered N-terminal domain[1] or histidine to cysteine replacements in either of the NCp zinc fingers resulted in unstable viral DNA.[21,22] In contrast, the majority of NCp7 appears to be dissociated from the full-length, linear double-stranded viral DNA extracted from infected cells within the reverse transcription/preintegration complexes (RTCs/PICs).[23–25] The loss of NCp7 *versus* an efficient retention of IN and Vpr within RTCs and PICs suggests that nucleocapsid disassembly and PIC assembly are directly interconnected during the reverse transcription process but the underlying mechanism is poorly understood.

The mature nucleocapsid protein (NCp7) is a small, basic and flexible polypeptide containing two highly structured zinc-finger domains separated by a proline-rich linker.[26–28] The rigid zinc finger structures are each stabilized *via* coordination of a zinc ion by the two highly conserved C-aa_2-C-aa_4-H-aa_4-C (CCHC) motifs.[29] They have been implicated in specific recognition of single-stranded (ss) nucleic acid regions and in nucleic acid chaperone activity.[4,30,31] The positive charges, especially those in the N-terminal sequence and the linker, are involved in the non-specific binding to nucleic acids and in the nucleic acid chaperone activity of NCp.[32] The nucleic acid chaperone activity of NCp7[4] has been shown to be necessary for viral RNA dimerization,[33–35] $tRNA^{Lys,3}$ annealing,[36,37] and throughout reverse transcription.[3,4] NCp7 and NCp9 (the ultimate precursor that behaves like NCp7 *in vitro*) non-specifically bind to nucleic acids *in vitro* with decreasing affinity from RNA to single-stranded DNA (ssDNA) and finally to double-stranded DNA (dsDNA).[38] They cover RNA or ssDNA lattices with occluded site sizes of 6–7 nt and 8 nt, respectively.[39] High-affinity binding of NCp7 is observed with HIV-1 RNA hairpins, and RNA hairpin aptamers,[27,40,41] and with oligodeoxyribonucleotides containing either TG repeats[42,43] or central flap-derived G-quartet structures.[44] Even though there is no detectable cooperativity, multiple binding of NCp7 occurs under saturating conditions with HIV-1 Psi-associated RNA hairpins[45] and parallel G-quadruplexes.[44]

A peculiar property of NCp7 and NCp9* (a one residue extended form of NCp9 at the C terminus) is the formation of nucleoprotein complexes of high molecular mass: with both their flexibility and their polycationic properties, these NCp capture nucleic acids lattices *in vitro* under saturating protein conditions. The corresponding aggregates have been observed with NCp and $tRNA^{Lys3}$,[46] RNA or rRNA,[47,48] and DNA.[49] With NCp9* and a model RNA (poly [rA]),[47] transmission electron microscopy (TEM) and quasi elastic light-scattering similarly show *in vitro* spherical aggregates that fall within a narrow distribution of sizes, after a salt-dependant ordered growth due to fusions of small aggregates.[47] They form at an NCp9*:occluded-site stoichiometric ratio of 1:1 and occur for a concentration of NCp9* in the micromolar range.[39] Optimum formation requires either 150 mM sodium or a combination of 50 mM sodium and 5 mM Mg^{2+}, which is similar to the salt concentration for optimum HIV-1 RT activity, *in vitro*.[47] Such an avidity between NCp9* and nucleic acids constitutes a distinctive feature over the prototypical single-stranded nucleic acid binding proteins like the *Escherichia coli* single-stranded binding (SSB) or phage T4 gene *32* proteins, which cover, spread out and maintain their DNA lattices individually.[50,51] Therefore, in order to approach both the architecture of the HIV-1 nucleocapsid and