Biophysical Chemistry 127 (2007) 64–68

Biophysical Chemistry

http://www.elsevier.com/locate/biophyschem

The HIV-1 central DNA flap region contains a "flapping" third strand

Besik I. Kankia *, Karin Musier-Forsyth

Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, United States

Received 22 November 2006; accepted 12 December 2006
Available online 19 December 2006

Abstract

Due to the discontinuous nature of HIV-1 plus-strand DNA synthesis, a 99-nt plus-strand overhang termed the "central DNA flap" is present near the center of the proviral DNA prior to integration. The flap appears to have stabilizing and/or protective effects on viral DNA, which has been hypothesized to be due to a specific conformation adopted by the three-stranded region. The 5′ end of the flap sequence is very purine rich and has the potential to adopt different secondary structures (e.g., duplex, triplex or quadruplex). In the present work, circular dichroism spectroscopy and thermal unfolding techniques were used to characterize an 89-nt long DNA sequence designed to mimic the three-stranded region at the 5′ end of HIV-1 proviral DNA. The effect of addition of the HIV-1 nucleocapsid protein (NC) on the nucleic acid structure was also examined. Although, guanine-rich short oligonucleotides derived from the DNA flap demonstrated CD spectra characteristic to parallel quadruplexes, this analysis reveals that the extended 89-nt construct folds into a canonical duplex with a "flapping" third strand both in the absence and presence of NC.

Keywords: HIV-1; Central DNA flap; Circular dichroism; Optical melting; Quadruplexes; Nucleocapsid protein

1. Introduction

Human immunodeficiency virus type 1 (HIV-1) carries its genetic information on a single-stranded RNA that is reverse transcribed into double-stranded DNA. Second or plus-strand DNA synthesis terminates with a strand displacement synthesis at the center of the DNA duplex, which generates a so-called "central DNA flap" of 99 nucleotides (nt) [1,2]. This three-stranded region of the proviral DNA contains two identical plus strands and a complementary minus-strand segment. Although the precise role of this structural feature is still an open question, the presence of the flap has been shown to enhance HIV-1 infection by facilitating nuclear import of proviral DNA [3,4]. The flap also appears to have a stabilizing and/or protective effect on viral DNA [5]. The protective effect has been hypothesized to be due to a specific conformation adopted by the three-stranded region, or alternatively, the flap may be involved in the recruitment of proteins [4–6]. Interestingly, the 5′ end of the flap contains two adjacent guanine tracts consisting of four and six guanines in a row, which could be involved in G-quartet formation [6]. In addition, a 27-nt long segment at the 5′ end of the flap contains 21 purines. Thus, the structure has the potential to form a pyrimidine–purine–purine triplex structure with the two purine strands in parallel orientation. Therefore, at least three possible structures are possible for the 5′ end of the flap region: (i) a canonical duplex with a "flapping" third strand; (ii) a quadruplex structure formed by guanine tracts; and (iii) a parallel pyrimidine–purine–purine triplex.

Previous studies to characterize the central DNA flap structure used native gel electrophoresis to analyze flap-derived oligonucleotides [5,6]. The appearance of slower migrating species under certain conditions was taken as evidence for a quadruplex formation. It was suggested that the G-quartets may form transiently and play a functional role in HIV-1 pre-integration complexes. These previous studies used tri-molecular DNA complexes, which gave individual strands conformational freedom to adopt preferred strand orientation.

In this work, the structure of the three-stranded segment at the 5′ end of HIV-1 proviral DNA is investigated using circular dichroism (CD) and ultraviolet (UV) spectroscopy. In the 89-nt long construct examined here, the three DNA strands are connected to each other by two 4-nt dT linkers (Fig. 1A), which

* Corresponding author. Tel.: +1 612 624 0510; fax: +1 612 626 7541.
E-mail address: bkankia@umn.edu (B.I. Kankia).

0301-4622/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2006.12.005

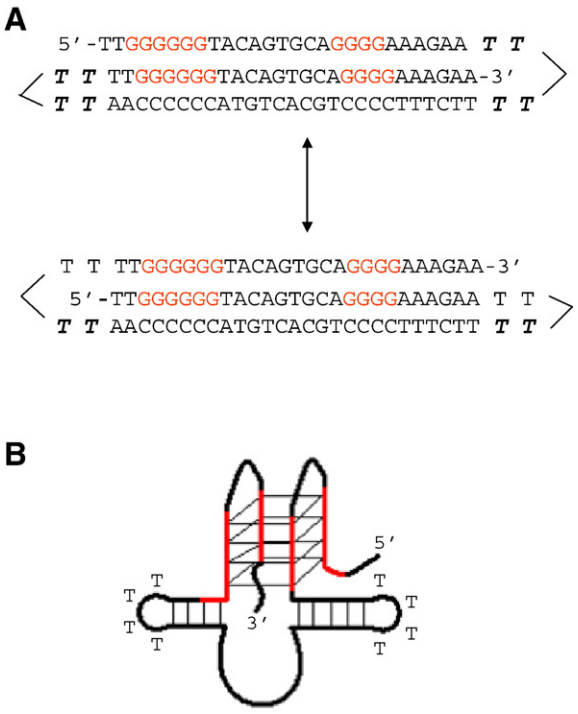

Fig. 1. (A) Sequence of the 89-nt DNA flap construct studied in this work. (B) Scheme of a possible quadruplex structure formed by the sequence shown in panel A.

restricts the orientation of the strands and is designed to more closely mimic the *in vivo* conformation of the DNA flap region. A schematic diagram of a possible quadruplex structure formed by this molecule is shown in Fig. 1B. To establish whether such a structure is formed, optical and thermodynamic studies were performed under various solution conditions both in the absence of protein and in the presence of HIV-1 nucleocapsid protein (NC). The latter is a nucleic acid chaperone known to destabilize nucleic acids to facilitate formation of thermodynamically stable structures [7–11].

2. Experimental section

DNA oligonucleotides were purchased from Integrated DNA Technologies, purified on 12% denaturing polyacrylamide gels, and dialyzed against 10 mM Cs-HEPES, pH 7.5, at 4 °C using Spectrum dialysis tubing with a molecular weight cutoff of 500 Da. The concentration of oligonucleotides was determined at 260 nm and 80 °C, using a molar extinction coefficient obtained from the tabulated values of the dimers and monomer bases using procedures reported earlier [12]. NC was prepared by solid-phase synthesis, cleaved, purified, and activated with Zn^{2+}, as has been described elsewhere [13].

CD spectra were obtained with a JASCO J710 spectropolarimeter equipped with a water-jacketed cell holder using 1-mm path-length cells. The measurements were performed either a few minutes after addition of the cations in solutions kept at 25 °C, after heating to 90 °C and slowly cooling to 25 °C, or after incubation for 48 h at 4 °C.

UV absorption readings were taken as a function of temperature, using a GBC 918 spectrophotometer equipped with thermoelectrically controlled cell holders. Melting curves recorded at 260 and 300 nm allowed an estimation of melting temperature, T_m, the midpoint temperature of the unfolding process. Van't Hoff enthalpies, ΔH_{vH}, were calculated using the following equation [14]:

$$\Delta H_{vH} = 6RT_m^2 \delta\alpha/\delta T$$

where, R is the gas constant and $\delta\alpha/\delta T$ is the slope of the normalized optical absorbance versus temperature curve at T_m.

3. Results and discussion

CD spectroscopy is a sensitive technique to study the secondary structures of DNA molecules. In general, canonical B-form DNA duplexes are characterized by a positive band around 260–270 nm and a negative band with the same intensity around 240 nm. DNA quadruplexes have different CD profiles depending on strand orientation. Parallel quadruplexes, where all strands have the same orientation, are characterized by a spectrum with a positive band around 260 nm, and a much less intense negative band around 240 nm [15]. In the case of antiparallel quadruplexes, both bands are shifted to higher wavelengths; the positive band is between 290–310 nm, and the negative band is between 260–280 nm [16]. CD bands of DNA triplexes are usually shifted to higher wavelengths relative to the bands of their parent duplexes, and are characterized by increased intensity in the positive band [17].

Fig. 1 shows the DNA construct designed to mimic the central DNA flap region of the HIV-1 proviral DNA. Initial characterization of the flap construct was performed under conditions expected to favor formation of DNA duplexes. The effect of adding specific cations known to favor triplex or quadruplex formation was also tested. While DNA duplexes can be folded in the presence of any alkaline or alkaline-earth cation, formation of triplexes and quadruplexes occurs only in the presence of specific cations. Due to the high linear charge density of phosphates in triplexes, effective screening by site-specific binding of divalent cations, such as Mg^{2+}, is required for triplex formation. This is especially true for purine–purine–pyrimidine triplexes [18]. Metal ion requirements for quadruplex formation are even more strict since the cations must fit into the inner core of G-quartets and form inner sphere complexes with the O6 carbonyl groups of the guanine bases [16,19]. As a result, quadruplex formation strongly depends on cationic size, with the most stable quadruplexes formed with cations of 1.3–1.4 Å ionic radius, such as K^+ and Sr^{2+} [16].

Due to the large ionic radius (1.7 Å) and low charge density, Cs^+ is an ideal cation to avoid formation of triplexes or quadruplexes. As expected, the CD profile of the 89-nt DNA flap mimic in 10 mM Cs-HEPES, pH 7.5, resembles the CD spectrum of B-DNA (Fig. 2A). The positive band is somewhat more intense than the negative band, which could be induced by the presence in our construct of an unstructured or "flapping"

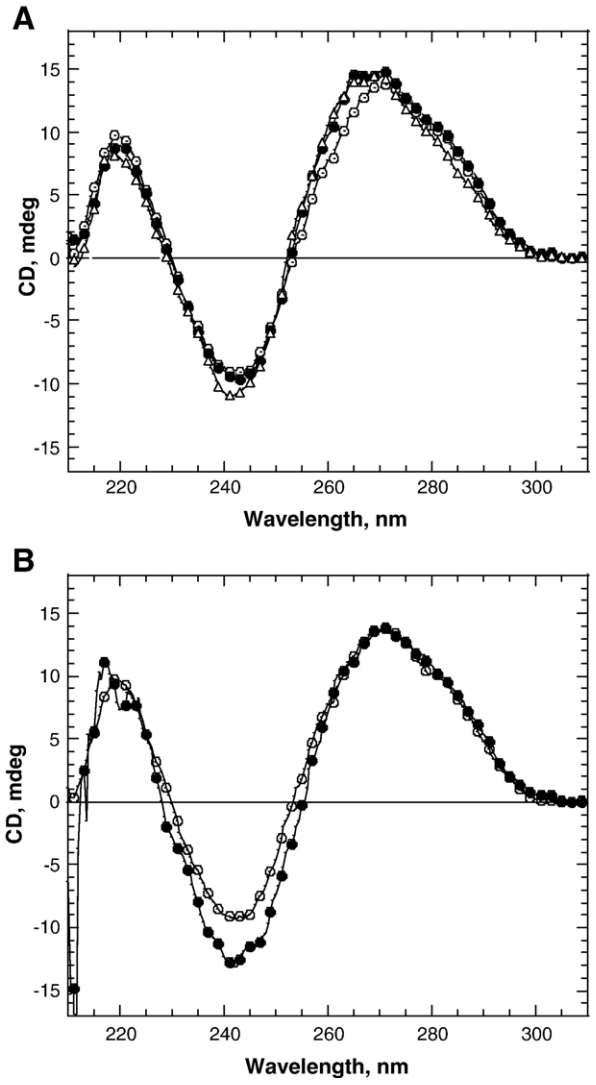

Fig. 2. CD spectra demonstrating the effect of different cations and NC on the secondary structure of the DNA flap construct shown in Fig. 1A. (A) Spectra of the construct in 10 mM Cs-HEPES, pH 7.5 (○), and after addition of 120 mM KCl (●) or 10 mM $SrCl_2$ (Δ). (B) Spectrum in 10 mM Cs-HEPES, pH 7.5, in the absence (○) and presence of 1 μM NC (●). All spectra were recorded in a 1 mm cell at 20 °C using 9 μM oligonucleotide.

third strand. Addition of 120 mM KCl did not result in any changes in the spectrum, while 10 mM $SrCl_2$ slightly increased the intensity of the negative band at ~245 nm (Fig. 2A). Addition of the same amount of $MgCl_2$ showed a similar effect (data not shown). These results suggest that both cations bind to the DNA construct in a nonspecific manner, stabilizing the B-conformation. Thus, all CD measurements reveal that the flap construct forms a canonical DNA duplex with an unassociated third strand even in the presence of buffer conditions favorable for triplexes or quadruplexes.

The secondary structure of the DNA flap construct was further characterized in the presence of NC, which is an abundant component of the retrovirus. NC is a small, highly basic protein containing 15 positively charged amino acids out of a total of 55 residues [20–23]. Due to its basic nature, NC binds tightly and nonspecifically to most nucleic acids [7]. It has been reported that in addition to facilitating reverse transcription by accelerating tRNA primer annealing [9,13,24] and minus-and plus-strand transfer reactions [25–27], NC facilitates strand displacement synthesis that produces the central DNA flap [2]. It was also shown that NC preferentially recognizes the intermolecular G-quadruplex structures formed by the sequences derived from the DNA flap [5]. Fig. 2B reveals that at a nt:NC ratio of 10, NC has a similar effect on the CD cpectrum of the DNA construct as the addition of 10 mM divalent cation, slightly increasing the intensity of the negative band. This result is in good agreement with a previous study using a perfect 15 nt DNA duplex, which also revealed insignificant changes in the CD profiles upon addition of similar ratios of NC [28]. As expected, further increase in the NC concentration induced nucleic acid aggregation, prohibiting accurate recording of CD spectra.

The CD spectra of short oligonucleotides derived from the DNA flap, 5′-TTGGGGGGTACAGTGCA-3′ (ODN1) and 5′-TTGGGGGGTACAGTGCAGGGGAAA-3′ (ODN2), in the presence of K^+ ions were also recorded. In agreement with previous studies [6], CD profiles characteristic of parallel quadruplexes were observed (Fig. 3). ODN1 is characterized by peaks at 245 nm and 265 nm that are ~1.5 times more intense than those of ODN2. This could be explained by formation of a 4-stranded structure with 6 G-quartets in the case of ODN1. In contrast, ODN2 may form a bimolecular structure with 4 G-quartets. In the latter case, to keep strands parallel, the dimer should contain crossed loops (for instance [29]).

UV melting experiments monitored at 260 nm and performed with 0.3 μM DNA in 10 mM Cs-HEPES, show that the flap construct undergoes a monophasic melting

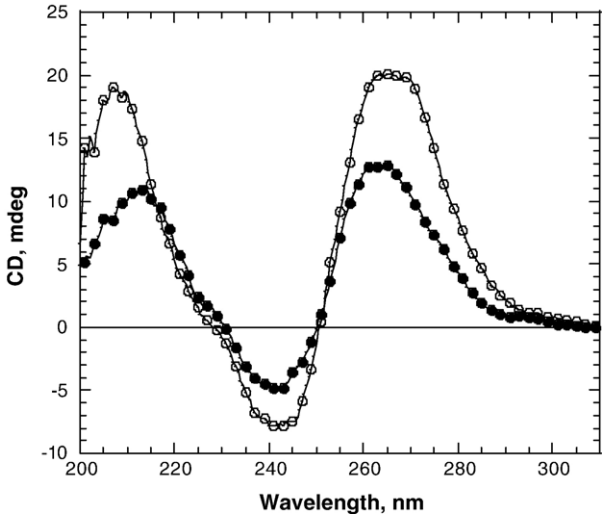

Fig. 3. CD spectra of short oligonucleotides derived from the DNA flap region, 5′-TTGGGGGGTACAGTGCA-3′ (50 μM (○)) and 5′-TTGGGGGGTACAGTGCAGGGGAAA-3′ (25 μM (●)) in 50 mM KCl, 10 mM Cs-HEPES, pH 7.5. The spectra are recorded in a 1 mm cell at 20 °C.

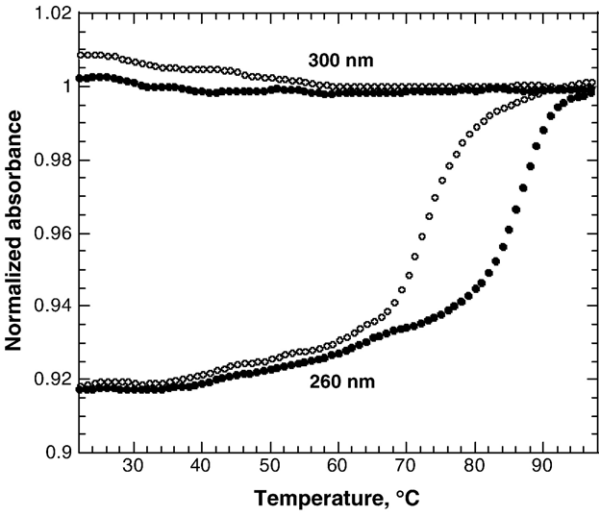

Fig. 4. Thermal unfolding of the DNA flap construct shown in Fig. 1 measured at 260 nm (lower curves) and 300 nm (upper curves). Measurements were obtained using 0.3 μM oligonucleotide in 10 mM Na-HEPES, pH 7.5 (○), or in 100 mM KCl, 10 mM HEPES, pH 7.5 (●).

transition with a T_m of 72 °C (Fig. 4). Addition of 100 mM KCl into the buffer shifted the T_m value by 15 °C, as expected due to the stabilizing effect of increased ionic strength on the DNA duplex. Melting experiments conducted with 10-fold higher oligonucleotide concentrations (3 μM) revealed a similar value for T_m (data not shown), supporting the monomolecular nature of the construct. This result also excludes association of strands into multimolecular quadruplexes. The van't Hoff enthalpy estimated from the melting curves is ~200 kcal/mol, which is in very good agreement with enthalpies for melting of a 27 base-pair DNA duplex predicted using DINAMelt (207 kcal/mol) [30]. A very small increase in absorbance at 260 nm also occurs under both conditions over the range 35–60 °C. This increase is consistent with unstacking of the bases in the flap strand. Monitoring the optical absorbance of the DNA construct at 300 nm as a function of temperature (Fig. 4), did not reveal any melting transition. Since this wavelength is sensitive to formation of antiparallel quadruplexes, these experiments rule out this type of structure.

In conclusion, the optical and thermodynamic studies reported here reveal that an 89-nt oligonucleotide construct (Fig. 1) designed to mimic the central DNA flap region of HIV-1 proviral DNA, forms a canonical DNA duplex with an unassociated or "flapping" third strand. No evidence was obtained for formation of either parallel or antiparallel quadruplexes, despite the fact that short oligonucleotides derived from the DNA flap sequence can fold into parallel quadruplex structures. It is unlikely that a longer construct designed to mimic the DNA flap region even more closely would favor quadruplex formation either, since this would simply result in elongation of the hairpin regions and single-stranded tails shown in Fig. 1B. Even in the absence of an unusual structure, the single stranded DNA flap may be used to attract nucleic acid binding proteins such as NC, which help to stabilize the proviral DNA prior to integration [5].

Acknowledgments

We would like to thank Prof. George Barany, Dr Daniel G. Mullen and Ms Brandie Kovaleski for chemical synthesis of NC. This work was supported by NIH grant GM065056.

References

[1] P. Charneau, G. Mirambeau, P. Roux, S. Paulous, H. Buc, F. Clavel, HIV-1 reverse transcription. A termination step at the center of the genome, J. Mol. Biol. 241 (1994) 651–662.
[2] L. Hameau, J. Jeusset, S. Lafosse, D. Coulaud, E. Delain, T. Unge, T. Restle, E. Le Cam, G. Mirambeau, Human immunodeficiency virus type 1 central DNA flap: dynamic terminal product of plus-strand displacement DNA synthesis catalyzed by reverse transcriptase assisted by nucleocapsid protein, J. Virol. 75 (2001) 3301–3313.
[3] Z. Ao, X. Yao, E.A. Cohen, Assessment of the role of the central DNA flap in human immunodeficiency virus type 1 replication by using a single-cycle replication system, J. Virol. 78 (2004) 3170–3177.
[4] V. Zennou, C. Petit, D. Guetard, U. Nerhbass, L. Montagnier, P. Charneau, HIV-1 genome nuclear import is mediated by a central DNA flap, Cell 101 (2000) 173–185.
[5] S. Lyonnais, R.J. Gorelick, J.L. Mergny, E. Le Cam, G. Mirambeau, G-quartets direct assembly of HIV-1 nucleocapsid protein along single-stranded DNA, Nucleic Acids Res. 31 (2003) 5754–5763.
[6] S. Lyonnais, C. Hounsou, M.P. Teulade-Fichou, J. Jeusset, E. Le Cam, G. Mirambeau, G-quartets assembly within a G-rich DNA flap. A possible event at the center of the HIV-1 genome, Nucleic Acids Res. 30 (2002) 5276–5283.
[7] J.G. Levin, J. Guo, I. Rouzina, K. Musier-Forsyth, Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism, Prog. Nucleic Acid Res. Mol. Biol. 80 (2005) 217–286.
[8] A. Rein, L.E. Henderson, J.G. Levin, Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication, Trends Biochem. Sci. 23 (1998) 297–301.
[9] J.-L. Darlix, M. Lapadat-Tapolsky, H. de Rocquigny, B.P. Roques, First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses, J. Mol. Biol. 254 (1995) 523–537.
[10] Z. Tsuchihashi, P.O. Brown, DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein, J. Virol. 68 (1994) 5863–5870.
[11] C. Bampi, S. Jacquenet, D. Lener, D. Decimo, J.L. Darlix, The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance, Curr. HIV Educ. Res. 2 (2004) 79–92.
[12] B.I. Kankia, Optical absorption assay for strand-exchange reactions in unlabeled nucleic acids, Nucleic Acids Res. 32 (2004) e154.
[13] H.W. Liu, G. Cosa, C.F. Landes, Y. Zeng, B.J. Kovaleski, D.G. Mullen, G. Barany, K. Musier-Forsyth, P.F. Barbara, Single-molecule FRET studies of important intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in HIV-1 reverse transcription, Biophys. J. 89 (2005) 3470–3479.
[14] L.A. Marky, K.J. Breslauer, Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves, Biopolymers 26 (1987) 1601–1620.
[15] M. Lu, Q. Guo, N.R. Kallenbach, Thermodynamics of G-tetraplex formation by telomeric DNAs, Biochemistry 32 (1993) 598–601.
[16] B.I. Kankia, L.A. Marky, Folding of the thrombin aptamer into a G-quadruplex with Sr(2+): stability, heat, and hydration, J. Am. Chem. Soc. 123 (2001) 10799–10804.
[17] A.M. Soto, J. Loo, L.A. Marky, Energetic contributions for the formation of TAT/TAT, TAT/CGC(+), and CGC(+)/CGC(+) base triplet stacks, J. Am. Chem. Soc. 124 (2002) 14355–14363.
[18] R.H. Shafer, Stability and structure of model DNA triplexes and quadruplexes and their interactions with small ligands, Prog. Nucleic Acid. Res. Mol. Biol. 59 (1998) 55–94.

[19] C.C. Hardin, A.G. Perry, K. White, Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids, Biopolymers 56 (2000) 147–194.
[20] L.M. Green, J.M. Berg, Retroviral nucleocapsid protein-metal ion interactions: folding and sequence variants, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 6403–6407.
[21] L.E. Henderson, T.D. Copeland, R.C. Sowder, G.W. Smythers, S. Oroszlan, Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses, J. Biol. Chem. 256 (1981) 8400–8406.
[22] J.M. Berg, Potential metal-binding domains in nucleic acid binding proteins, Science 232 (1986) 485–487.
[23] S.N. Covey, Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus, Nucleic Acids Res. 14 (1986) 623–633.
[24] H. De Rocquigny, C. Gabus, A. Vincent, M.C. Fournié-Zaluski, B. Roques, J.L. Darlix, Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 6472–6476.
[25] J. Guo, L.E. Henderson, J. Bess, B. Kane, J.G. Levin, Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA, J. Virol. 71 (1997) 5178–5188.
[26] S. Auxilien, G. Keith, S.F.J. Le Grice, J.-L. Darlix, Role of post-transcriptional modifications of primer $tRNA^{Lys,3}$ in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription, J. Biol. Chem. 274 (1999) 4412–4420.
[27] T. Wu, J. Guo, J. Bess, L.E. Henderson, J.G. Levin, Molecular requirements for human immunodeficiency virus type 1 plus-strand transfer: analysis in reconstituted and endogenous reverse transcription systems, J. Virol. 73 (1999) 4794–4805.
[28] B.I. Kankia, G. Barany, K. Musier-Forsyth, Unfolding of DNA quadruplexes induced by HIV-1 nucleocapsid protein, Nucleic Acids Res. 33 (2005) 4395–4403.
[29] A. Risitano, K.R. Fox, Inosine substitutions demonstrate that intramolecular DNA quadruplexes adopt different conformations in the presence of sodium and potassium, Bioorg. Med. Chem. Lett. 15 (2005) 2047–2050.
[30] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. 31 (2003) 3406–3415.