JOURNAL OF VIROLOGY, June 2002, p. 5452–5462 Vol. 76, No. 11
0022-538X/02/$04.00+0 DOI: 10.1128/JVI.76.11.5452–5462.2002

Rous Sarcoma Virus Gag Protein-Oligonucleotide Interaction Suggests a Critical Role for Protein Dimer Formation in Assembly

Yu May Ma and Volker M. Vogt*

Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853

Received 12 December 2001/Accepted 1 March 2002

The structural protein Gag is the only viral product required for retrovirus assembly. Purified Gag proteins or fragments of Gag are able in vitro to spontaneously form particles resembling immature virions, but this process requires nucleic acid, as well as the nucleocapsid domain of Gag. To examine the role of nucleic acid in the assembly in vitro, we used a purified, slightly truncated version of the Rous sarcoma virus Gag protein, ΔMBDΔPR, and DNA oligonucleotides composed of the simple repeating sequence GT. Apparent binding constants were determined for oligonucleotides of different lengths, and from these values the binding site size of the protein on the DNA was calculated. The ability of the oligonucleotides to promote assembly in vitro was assessed with a quantitative assay based on electron microscopy. We found that excess zinc or magnesium ion inhibited the formation of virus-like particles without interfering with protein-DNA binding, implying that interaction with nucleic acid is necessary but not sufficient for assembly in vitro. The binding site size of the ΔMBDΔPR protein, purified in the presence of EDTA to remove zinc ions at the two cysteine-histidine motifs, was estimated to be 11 nucleotides (nt). This value decreased to 8 nt when the protein was purified in the presence of low concentrations of zinc ions. The minimum length of DNA oligonucleotide that promoted efficient assembly in vitro was 22 nt for the zinc-free form of the protein and 16 nt for the zinc-bound form. To account for this striking 1:2 ratio between binding site size and oligonucleotide length requirement, we propose a model in which the role of nucleic acid in assembly is to promote formation of a species of Gag dimer, which itself is a critical intermediate in the polymerizaton of Gag to form the protein shell of the immature virion.

In retroviruses, the structural protein Gag is necessary and sufficient to direct the assembly, budding, and release of virus-like particles (VLPs) from the plasma membrane. This assembly process involves multiple steps, including Gag-RNA interaction, Gag-Gag interaction, and Gag-membrane interaction. Late in the budding process, the Gag polyprotein is cleaved by viral protease (PR) to generate the mature proteins matrix (MA), capsid (CA), and nucleocapsid (NC), which are ubiquitous in all retroviruses, as well as several small peptides or proteins specific to the genus. MA lines the inner side of the viral membrane. CA is the major structural protein that forms the shell of viral core. In human immunodeficiency virus type 1 (HIV-1), CA dimerizes in vitro, a process that is mediated through its C-terminal domain (CTD) (21). In contrast, Rous sarcoma virus (RSV) CA (10, 36), as well as human T-cell leukemia virus type 1 (35) and equine infectious anemia virus CA (4), remains as a monomer even at very high protein concentrations. NC is a highly basic protein that coats the two strands of genomic RNA, forming a nucleocapsid complex within the viral core. Retroviral NC proteins have one or two Cys-His motifs (CCHC) that tightly bind zinc, and virions contain stoichiometric amounts of this metal ion. In solution structures of NC proteins of HIV-1 (42), Mason-Pfizer monkey virus (23), mouse mammary tumor virus (37), and murine leukemia virus (MLV) (17), the CCHC motifs form a "finger" or "knuckle" surrounding the ion, but the rest of the NC protein is unstructured. Although the CCHC motifs are dispensable for assembly both in vivo and in vitro, they are critical for binding specifically to the RNA packaging sequence that allows the incorporating of the virus genome into the virion (3, 15). From deletion analyses, three short sequences within Gag have been found to serve important functions during budding. The M (membrane-binding) sequence is responsible for plasma membrane targeting and binding. This sequence includes the basic amino acid residues within the first half of MA domain, as well as the N-terminal myristate group present in most of retroviral Gag proteins but not in RSV Gag (40, 48, 50, 58, 61, 70). The L (late) sequence, which is located either between MA and CA or at the C terminus of Gag, facilitates the pinching off and release of the budding virions. Deletion of this sequence results in the accumulation of virus particles on the plasma membrane, where they appear to remain tethered by membrane stalks (2, 13, 24, 25, 51, 52, 62, 65, 66). The I (interaction) sequence, which is poorly defined but appears to be present in one or two copies in the NC domain, is composed mainly of basic amino acid residues. Deletion of this sequence is reported to lead to the formation of particles with abnormally low density, presumably by abolishing RNA binding and thereby somehow leading to loose packing of the Gag polyprotein (5, 12, 16). However, the correlation between basic residues and particle density was recently challenged, and the mechanism by which I domains exert their function in assembly remains controversial (11).

In vitro systems have been developed to study assembly for several retroviruses, including Mason-Pfizer monkey virus (38), RSV (8, 9, 31, 68), and HIV-1 (6, 22, 26–29, 45, 60). Spherical or tubular VLPs form spontaneously in vitro from purified Gag protein or fragments of Gag. In the RSV system, morphology is dictated in part by the presence or absence of

* Corresponding author. Mailing address: Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853. Phone: (607) 255-2443. Fax: (607) 255-2428. E-mail: vmv1@cornell.edu.

the 25 C-terminal amino acid residues of the p10 domain just N-terminal to CA (8, 9, 31). In the HIV-1 system, CA itself, or CA-NC in the presence of RNA, forms conical particles and tubular particles of different widths (18, 22, 26, 27, 60). However, in some contexts N-terminal extensions of CA, the presence of the spacer region (SP1) between CA and NC, or a basic pH all promote the formation of spherical particles (28, 29, 60). Although Gag proteins missing part of MA form particles closely resembling immature virions, full-length HIV-1 Gag assembles into particles of improper size unless an inositol polyphosphate-containing compound is present (7). Although under some circumstances the purified mature RSV CA or HIV-1 CA can form tubular particles by itself, usually the NC domain, as well as RNA, is required for the efficient assembly of spherical particles resembling wild-type immature virions (18, 22, 26, 27, 36, 60, 68). Previous studies showed that the sequence of the RNA is irrelevant; various species of RNA, as well as DNA oligonucleotides and a polyanion, are able to promote efficient assembly in vitro, suggesting that the electrostatic interaction between Gag protein and the nucleic acid is the driving force for assembly (68). In a model proposed previously, the role of nucleic acid is to act as a scaffold. Binding of Gag to RNA via the NC domain would bring the upstream domains of Gag close together, thus facilitating protein-protein interactions (8).

We used two independent assays for protein-nucleic acid binding, as well as quantitative electron microscopic (EM) analysis, to study the relationship between protein-nucleic acid binding and in vitro assembly of an RSV Gag fragment missing the N-terminal membrane-binding domain and the C-terminal PR domain. When prepared in the presence of zinc, the Gag protein has a smaller binding site on DNA oligonucleotides than the protein prepared in the presence of EDTA. However, in both cases, the shortest DNA oligonucleotide that supports efficient assembly is twice the size of the binding site of protein. This ratio is striking, suggesting a model in which the role of nucleic acid in assembly is to promote formation of Gag dimers; these, in turn, are the critical building blocks for polymerization of the Gag shell in the immature virion.

MATERIALS AND METHODS

DNA constructs and protein purification. The plasmid pET3xc-ΔMBDΔPR was constructed by using common subcloning techniques and propagated in the DH5α strain of *Escherichia coli* as described elsewhere (9). *E. coli* BL21(DE3)/pLysS cells were grown and induced for protein expression as previously described (9).

Protein ΔMBDΔPR was purified by using a method modified from previous reports (9, 68). The frozen cell pellets were resuspended on ice in buffer A (20 mM Tris [pH 7.5], 0.5 M NaCl, 0.1% Nonidet P-40, 10% glycerol, 1 mM EDTA, 10 mM dithiothreitol [DTT], 1 mM phenylmethylsulfonyl fluoride [PMSF]) at 25 ml/liter of cell culture, and the cells were broken by sonication. Insoluble debris, ribosomes, and nucleic acids were removed by the addition of 0.3% polyethyleneimine, followed by ultracentrifugation at 45,000 rpm for 3 h by using a 50.2 Ti rotor. Soluble protein was first precipitated with 25% saturated ammonium sulfate and then resuspended in buffer B (20 mM Tris [pH 7.5], 1 mM EDTA, 10 mM DTT, 1 mM PMSF) plus 0.1 M NaCl at 5 ml/liter of cell culture. Insoluble material was removed by centrifugation, and the supernatant was applied to a DEAE-cellulose (DE-52) column. The resin was washed with buffer B plus 0.1 M NaCl. The flowthrough and wash fractions were pooled and loaded onto a phosphocellulose (Whatman P11) column. The resin with bound protein was washed with buffer B plus 0.1 M NaCl and then with buffer B plus 0.3 M NaCl. Protein was eluted with buffer B plus 0.5 M NaCl and dialyzed against buffer B plus 0.1 M NaCl overnight at 4°C.

To obtain ΔMBDΔPR protein with bound zinc (ΔMBDΔPR-Zn), a modified scheme was used to include the divalent metal ion in the purification steps. We added 50 μM $ZnCl_2$ to each buffer instead of 1 mM EDTA. The final purified protein was first dialyzed against buffer B plus 0.1 M NaCl and 200 μM $ZnCl_2$ overnight and then dialyzed against buffer B plus 0.1 M NaCl and 20 μM $ZnCl_2$ for 4 h to remove the bulk of the free zinc. Protein aliquots were stored in liquid nitrogen for no more than 2 months.

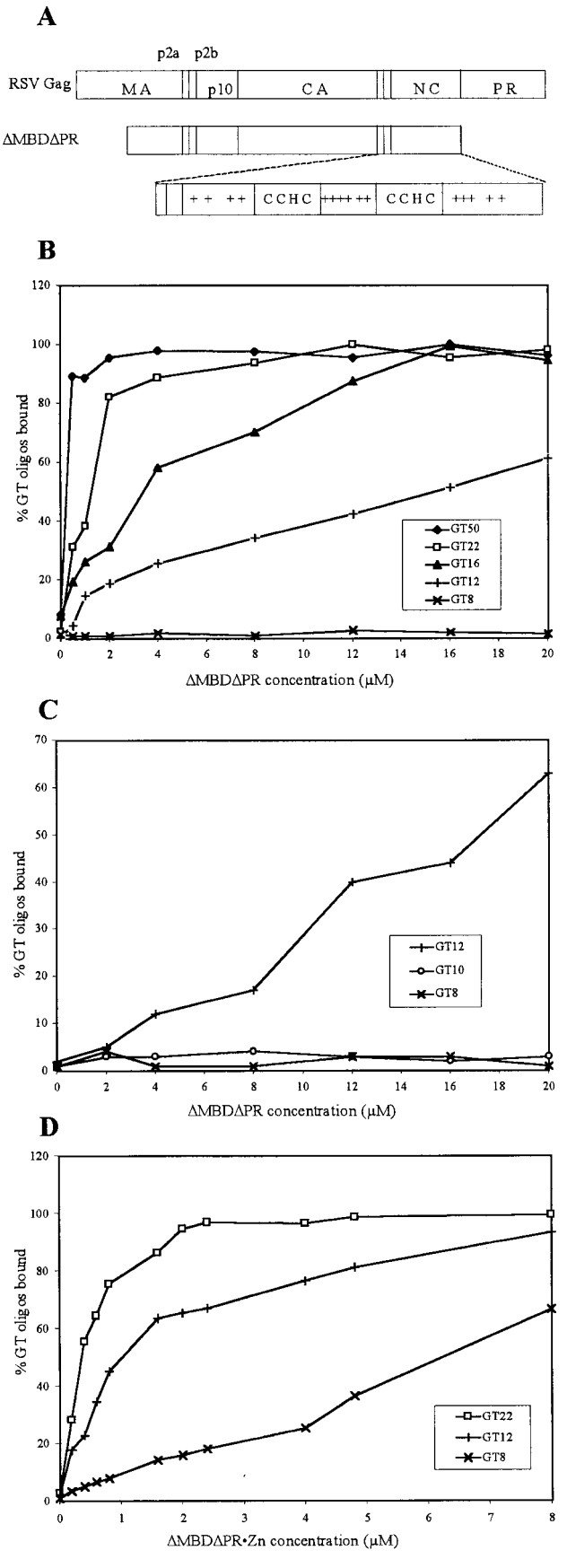
Protein concentration was determined by spectrophotometry. The A_{260}/A_{280} ratio of purified protein was determined to be ca. 0.7, indicating a lack of nucleic acid contamination. The final protein was ca. 90% pure as judged by Coomassie blue staining after sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The protein used in the following binding and assembly experiments was frozen and thawed only once.

Binding assays. Nitrocellulose filter-binding assays were performed as described before (53). In brief, standard reaction mixtures (50 μl) contained buffer C (20 mM Tris [pH 7.5], 100 mM NaCl, 1 mM DTT), 0.2 to 1 ng of ^{32}P-labeled DNA oligonucleotide, and various concentrations of purified protein. After 30 min of incubation at room temperature, the reaction mixtures were filtered through 25-mm presoaked nitrocellulose filters (BA85; Schleicher & Schuell, Inc.) on a large manifold with gentle suction (ca. 0.5 ml/min per filter). The filters were then washed three times with 0.5 ml of buffer C and counted for radioactivity. The apparent binding constant (K_a) of the protein is defined as the inverse of the protein concentration at which 50% of maximal binding occurs. The binding site size of ΔMBDΔPR and of ΔMBDΔPR-Zn was calculated by pairwise comparison of the K_a of each protein with DNA oligonucleotides of various lengths and composed of the simple repeat sequence GT, as described in Results. This analysis is based on the following assumptions. (i) One protein molecule bound to the oligonucleotide will cause retention of the complex with 100% efficiency. (ii) None of the protein dissociates from the complex on the filter during the wash step. (iii) The intrinsic binding constant (K_i) is the same for all GTmers that are at least as long as the binding site size. (iv) At the protein concentration at which half-maximal binding occurs, not more than one protein molecule is bound to any oligonucleotide. (v) There is only one binding mode, governed by one K_i, for the protein on any GTmer, and thus the binding is "all or nothing."

Assumption i above is standard in this type of assay and is likely to hold for large proteins, although it may not hold for very small proteins such as NC itself (53). Assumption ii also is standard. Assumption iii might be in error by a factor of two, since possibly the NC domain could bind differently to the sequence GTGT. . . than to the sequence TGTG. . . . For both assumptions ii and iii, the conclusions that we draw are unlikely to be affected because they are based on comparisons. Assumption iv should hold for oligonucleotides that are less than twice the binding site size, since no matter how the first protein binds there will not be enough room for a second protein. However, for larger DNAs, e.g., GT50, some of the DNA molecules retained on the filter will have more than one bound protein. (The complications in an analysis of the protein-DNA binding for nonspecific interactions, in which there are overlapping binding sites, are discussed in reference 19.) In our experiments, the K_a for GT50 estimated by the protein concentration at half-maximal binding underestimates the real K_a by ca. 23%, which would cause the binding site size to be underestimated by less than 1 nt. Assumption v is a simplification for two reasons. First, published evidence suggests that under some conditions the NC domain can become more compact and thus have a smaller binding site size (34). Second, it seems likely that the NC domain also could bind to a shorter stretch of nucleotides than the calculated binding site size, e.g., 7 nt instead of 8 nt, but with a lower affinity. Unless the affinity for such a putative smaller binding site is of a magnitude similar to that of a full binding site, our treatment of NC-oligonucleotide binding is likely to remain a good approximation.

For gel mobility shift assays, in a typical reaction (10 μl) various amounts of purified protein were mixed with 0.2 to 1 ng of ^{32}P-labeled DNA oligonucleotide and buffer C. The mixtures were incubated at room temperature for 30 min, after which 2 μl of 50% glycerol was added to each reaction. The mixtures were then separated on a 10% polyacrylamide gel in TA buffer (40 mM Tris-acetate, 20 μM $ZnCl_2$). The gel was subsequently dried and exposed to film. The RSV and HIV-1 NC proteins used in this assay were generously provided by Robert Gorelick (64).

Transmission EM. A direct dilution method was used for in vitro assembly as described previously (68). In brief, protein (typically 5 mg/ml) in buffer B plus 0.1 M NaCl was first diluted with 4 volumes of buffer D (50 mM morpholineethanesulfonic acid [pH 6.0], 0.1 M NaCl), followed by the addition of DNA oligonucleotide (10% [wt/wt] ratio of nucleic acid to protein). Assembly reactions were incubated at room temperature for 30 min or at 4°C for 24 h. Particles formed under these conditions were negatively stained with 2% uranyl acetate (pH 5) on



Formvar-carbon coated grids. To estimate the assembly efficiency under defined conditions, the total number of particles was counted from five EM micrographs, each from an independent assembly reaction.

RESULTS

Analysis of protein-oligonucleotide interaction. To be tractable for in vitro assembly, Gag proteins must be soluble under the buffer conditions used. The largest soluble fragment of RSV Gag as purified from *E. coli* is a protein missing the N-terminal 84-amino-acid membrane-binding domain, as well as the 123-amino-acid protease domain. We have been using this protein, ΔMBDΔPR, as a model for full-length Gag (Fig. 1A) (9, 68). ΔMBDΔPR assembles into spherical VLPs in the presence of a nucleic acid. The VLPs closely resemble immature virions in size and morphology. Diverse RNAs, double-stranded DNA, and single-stranded DNA oligonucleotides as short as 22 nt all promote assembly efficiently. While at least a portion of the NC domain is required for assembly both in vitro and in vivo, the CCHC motifs that bind Zn^{2+} ions are not required. In previous work, ΔMBDΔPR was purified in the presence of 1 mM EDTA, which results in the ejection of zinc. We have modified the purification scheme to include 50 μM $ZnCl_2$, as well as fresh DTT, at every purification step to prevent the loss of the divalent cation. The resulting protein, called ΔMBDΔPR-Zn, was used in the experiments described below in parallel with protein prepared with EDTA (ΔMBDΔPR).

In order to probe the role of nucleic acid binding in particle assembly, we used a set of defined DNA oligonucleotides composed of the dinucleotide repeat GT (GTmers). Both binding constant and assembly efficiency were measured. It had been shown previously that HIV-1 NC binds to runs of GT repeats much more tightly than to other sequences in single-stranded DNA, but it is unknown whether RSV NC has a similar preference (20). We first determined the binding constants of the two forms of ΔMBDΔPR by means of the classic nitrocellulose filter assay. In this assay protein is retained quantitatively on the nitrocellulose, while naked DNA flows through. If as little as one protein molecule is bound to a DNA molecule, the complex remains on the filter.

If the protein in the assay is in large molar excess over the radioactively labeled DNA, the apparent binding constant (K_a) is equal to the inverse of protein concentration required for retention of half of the total radioactive oligonucleotide on the filter. This condition is met in our experiments, since the molar ratio of protein to DNA was more than 1,000. Longer oligo-

FIG. 1. Filter-binding analysis. (A) Diagrammatic representation of proteins. The rectangle at the top represents the RSV Gag protein, with vertical lines marking the sites of proteolytic cleavages during maturation. The ΔMBDΔPR protein is missing the N-terminal membrane-binding domain, as well as the C-terminal protease domain. +, lysine or arginine residue; CCHC, the conserved zinc-binding motif. (B to D) Representative nitrocellulose filter-binding data for ΔMBDΔPR (B and C) and ΔMBDΔPR-Zn (D) with a series of GT oligonucleotides. The total length of each DNA is given by the number following GT. The apparent binding constant (K_a) is defined as the inverse of protein concentration that allows 50% of the radioactively labeled oligonucleotide to be retained on the nitrocellulose filter. At least four independent binding assays were performed, some with lower concentrations of protein to expand the curves for the larger DNAs.

TABLE 1. Compilation of apparent binding constants[a]

GTmer	ΔMBDΔPR		ΔMBDΔPR-Zn	
	K_a (M^{-1})	n	K_a (M^{-1})	n
GT50	$(7.4–9.1) \times 10^6$	5	ND	
GT22	$(1.1–1.2) \times 10^6$	5	$(2.6–2.9) \times 10^6$	4
GT16	$(2.9–3.2) \times 10^5$	4	ND	
GT12	$(6.6–6.9) \times 10^4$	4	$(8.2–9.8) \times 10^5$	4
GT10	—		ND	
GT8	—		$(1.5–1.7) \times 10^5$	4

[a] The range of values from independent experiments is shown. n is the number of experiments. ND, not determined; —, no binding observed.

nucleotides have a higher K_a than shorter oligonucleotides because fewer binding events are required for one-half of the counts to be retained. For a synthetic DNA with a simple repeating sequence, in this case the dinucleotide GT, the intrinsic binding constant (K_i) governing interaction of the protein with the DNA does not depend on DNA length (as long as the DNA is larger than the binding site size), since the protein always sees the same sequence. As a consequence, in a set of GT oligonucleotides, differences in K_a are due entirely to differences in length. Systematic analysis of these differences can yield estimates both of K_i and of the binding site size of the protein. In simplified form, $K_a = N \cdot K_i$, where N is the number of binding sites on the DNA. N is a function of length of the oligonucleotide, L, and the size of the binding site size, S, as given by the formula: $N = L - S + 1$. For example, for $S = 10$ and $L = 20$, there would be 11 different overlapping binding sites. Since K_i is constant, K_a can be measured, L is known, and S is assumed to be a constant property of the protein, any pairwise comparison between the K_a measured for two oligonucleotides will yield an estimate for S. Thus, $K_i = K_a'/N' = K_a''/N''$. Several of the assumptions and approximations on which this treatment is based are presented in Materials and Methods (19).

Role of zinc in binding site size of ΔMBDΔPR. Five GT oligonucleotides were tested in the nitrocellulose filter-binding assay at 100 mM NaCl. Four of these were able to be retained by ΔMBDΔPR on the filter, while one was not (Fig. 1B). GT50, GT22, GT16, and GT12 were bound by the protein, and as expected on a theoretical basis, K_a decreased with decreasing DNA size (Table 1). No GT8 binding was detectable under the standard condition for in vitro assembly (1 mg of protein/ml). By using pairwise comparisons of the K_a measured for each of the four oligonucleotides, we deduced that the binding site size of ΔMBDΔPR is ca. 9 to 13 nt (Table 2). This value is consistent with the observation that GT8 failed to bind to ΔMBDΔPR even when the protein concentration was as high as 20 μM (~1 mg/ml) (Fig. 1B). To refine the measurements of binding site size, we carried out a second set of binding experiments with three small GTmers that differed in length by only 2 nt: GT12, GT10, and GT8 (Fig. 1C). While ΔMBDΔPR protein bound to GT12 as shown previously, no binding was detected for GT8 and GT10 with a protein concentration as high as 1 mg/ml. Based on this assay the binding site size for ΔMBDΔPR on GTmers thus appears to be approximately 11 to 12 nt, a finding consistent with the value obtained by pairwise comparisons of K_a. We also performed less-rigorous filter-binding assays to estimate the K_a for three DNA oligonucleotides made up of the repeat sequence AC: AC50, AC22, and AC12 (data not shown). Pairwise comparison of the resulting binding constants implies a binding site size of ca. 11 to 16 nt, suggesting that the binding site size is an intrinsic property of the NC domain, independent of the DNA sequence.

TABLE 2. Binding site size of ΔMBDΔPR protein[a]

GTmer	Binding site size (nt) with GTmer			
	GT50	GT22	GT16	GT12
GT50				
GT22	9			
GT16	12	13		
GT12	11	11	10	

[a] The binding site size of the zinc-free form of ΔMBDΔPR was calculated by pairwise comparisons of the apparent binding constant (K_a) of protein and four GTmers of the indicated lengths.

The binding site size on GTmers determined as described above for ΔMBDΔPR without Zn^{2+} is consistent with a previous report, in which the molar ratio of ΔMBDΔPR to nucleic acid was determined for particles assembled in vitro, from the A_{260}/A_{280} ratio after dissolution of the particles in sodium dodecyl sulfate (68). The conclusion that this ratio reflects the binding site size of the protein is indirect, being based on the assumption that in the particles all protein molecules contact nucleic acid in the same way and cover the entire nucleic acid used in assembly. However, all published calculations of the binding site size of retroviral NC proteins are in the range of 5 to 8 nt, smaller than the value we estimate here for ΔMBDΔPR (20, 32, 56, 57, 67). We envisioned two possible explanations for this discrepancy. First, the larger binding site size could be an intrinsic property of the Gag precursor, reflecting a fundamental difference between a Gag protein containing an NC domain and a mature NC protein. Second, the binding of the CCHC motifs to Zn^{2+} might play a role in binding site size.

In order to differentiate between these two possible explanations, we performed gel mobility shift assays to compare the ability of three proteins prepared in the presence of $ZnCl_2$, ΔMBDΔPR-Zn, RSV NC, and HIV-1 NC, to bind to a GT8 oligonucleotide in the presence or absence of 5 mM EDTA (Fig. 2). RSV and HIV-1 NC had been purified by using high-pressure liquid chromatography and were resuspended in buffer containing $ZnCl_2$, conditions known to allow stoichiometric binding of the metal ion to the protein (64). In the absence of protein, the ^{32}P-labeled GT8 probe migrated as a single band in a 10% polyacrylamide gel (Fig. 2, lanes 1). Consistent with findings of previous reports, both RSV and HIV-1 NC were able to bind to this labeled DNA and form stable complexes, which were retained in the well (Fig. 2A, lanes 6 and 7; Fig. 2B, lanes 2 and 3 and lanes 6 and 7). The lack of migration out of the well may be accounted for by the small size of the oligonucleotide, since most or all of its negative charges would be expected to be neutralized by the charged residues on NC. In contrast to the filter-binding results described above for ΔMBDΔPR in absence of Zn, the ΔMBDΔPR-Zn protein was able to immobilize GT8 in the wells in the absence of EDTA (Fig. 2A, lanes 2 and 3). However, in the presence of EDTA, presumably resulting in ejec-

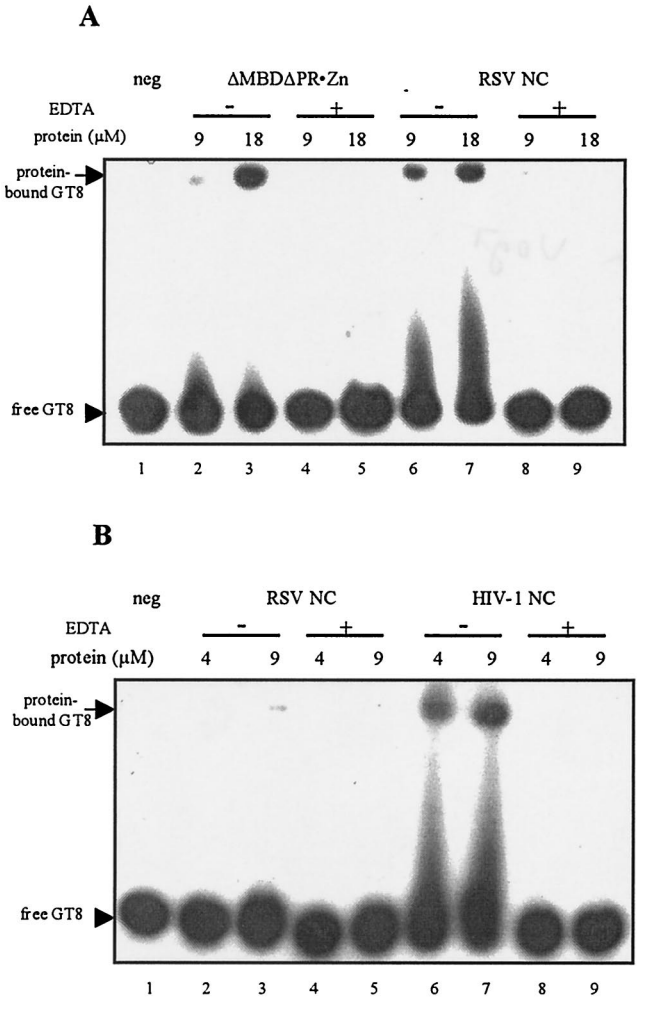


FIG. 2. Gel mobility shift analysis. Three zinc-bound proteins—ΔMBDΔPR-Zn, RSV NC, and HIV-1 NC--were mixed with a ^{32}P-labeled GT8 probe in the absence or presence of 5 mM EDTA. The binding reactions were incubated at room temperature for 30 min before electrophoresis was performed on a polyacrylamide gel. The concentrations of protein and the location of free and protein-bound GT8 are indicated. neg, no protein added.

tion of the zinc ions, the binding between GT8 and all three proteins was abolished (Fig. 2, lanes 4 and 5 and lanes 8 and 9).

We draw the following conclusions from these results. First, the binding site sizes of both the Gag precursor and the mature NC protein that contain functional zinc fingers are 8 nt or smaller. However, after the removal of zinc, the proteins adopt a more extended conformation, resulting in a larger binding site size. For RSV ΔMBDΔPR, this binding site size is ca. 11 to 12 nt. Second, little difference in binding site size or binding affinity can be detected between the Gag precursor and the mature NC protein by this assay. Thus, the NC domain as part of the much larger Gag precursor acts in a way that is similar to that of the mature NC. Third, as inferred from the mobility shift assays in Fig. 2B (lanes 2 and 3 and lanes 6 and 7), HIV-1 NC binds to GT8 much more tightly than RSV NC does. Since previous reports suggested a similar binding constant for both proteins interacting with poly(rA), the higher affinity of HIV-1 NC for GT8 could be due to a difference in sequence preference (32, 56, 67).

TABLE 3. Binding site size of ΔMBDΔPR-Zn protein[a]

GTmer	Binding site size (nt) with:		
	GT22	GT12	GT8
GT22			
GT12	7.4		
GT8	8.0	8.2	

[a] The binding site size of the zinc-bound form of ΔMBDΔPR was calculated by pairwise comparisons of the apparent binding constant (K_a) of protein and three GTmers of the indicated lengths.

To determine the binding constant of ΔMBDΔPR-Zn and GT oligonucleotides, we measured K_a values for three small GTmers by means of nitrocellulose filter binding (Table 1). The zinc-bound form of the protein retained GT8 on the filter, confirming the results from the gel mobility shift assays (Fig. 1D). From pairwise comparisons of the K_a of the protein with three GT oligonucleotides, the binding site size of ΔMBDΔPR-Zn was estimated to be ca. 8 nt (Table 3), a finding consistent with the results previously reported for retroviral NC proteins (32, 56, 57, 67). If we assume that the binding site sizes of ΔMBDΔPR and ΔMBDΔPR-Zn are 12 and 8 nt, respectively, the intrinsic binding constants for the two forms of the protein would be similar, approximately 0.7×10^5 to 1.7×10^5 M^{-1}, a value that also is consistent with the binding constant previously reported for various retroviral NC proteins at 100 mM NaCl (32, 67). More recent data obtained by using plasmon resonance (Biacore) technology suggest a much higher intrinsic binding constant for HIV to GT sequences, with a value of as high as 10^7 to 10^8 M^{-1} (20). One possible explanation for this apparent discrepancy is that RSV NC might not have the same specificity on GT sequence as HIV-1 NC does, as also suggested by the gel mobility shift assays described above (Fig. 2B).

Dependence of assembly on oligonucleotide size. We analyzed the role of oligonucleotide length in assembly by testing the ability of several GTmers to support the formation of VLPs by ΔMBDΔPR or by ΔMBDΔPR-Zn. Particles assembled under standard conditions were counted under EM after negative staining. At least five independent experiments were carried out, each with a parallel assembly reaction containing GT50 (for ΔMBDΔPR) or GT22 (for ΔMBDΔPR-Zn) as an internal control, and the data were normalized to this control. The results showed that for both proteins no VLPs whatsoever formed with GT8 (Fig. 3). Since ΔMBDΔPR-Zn is able to bind to GT8, the absence of particles implies that binding itself is not sufficient for assembly in vitro. For larger oligonucleotides, both proteins showed a cutoff in DNA length, below which few assembled particles were detected. For ΔMBDΔPR this cutoff was ca. 20 to 22 nt (Fig. 3A). GT22 was almost as efficient at promoting assembly as the much larger GT50, which was itself similar to total *E. coli* RNA (data not shown). However, GT20 and GT18 were much less efficient, showing only 23 and 15% as many VLPs as GT22 after 30 min, respectively. The zinc-bound form of the protein, ΔMBDΔPR-Zn, showed a slightly different profile of assembly efficiency with GTmers, with a sharper cutoff in size. Maximal efficiency was achieved with

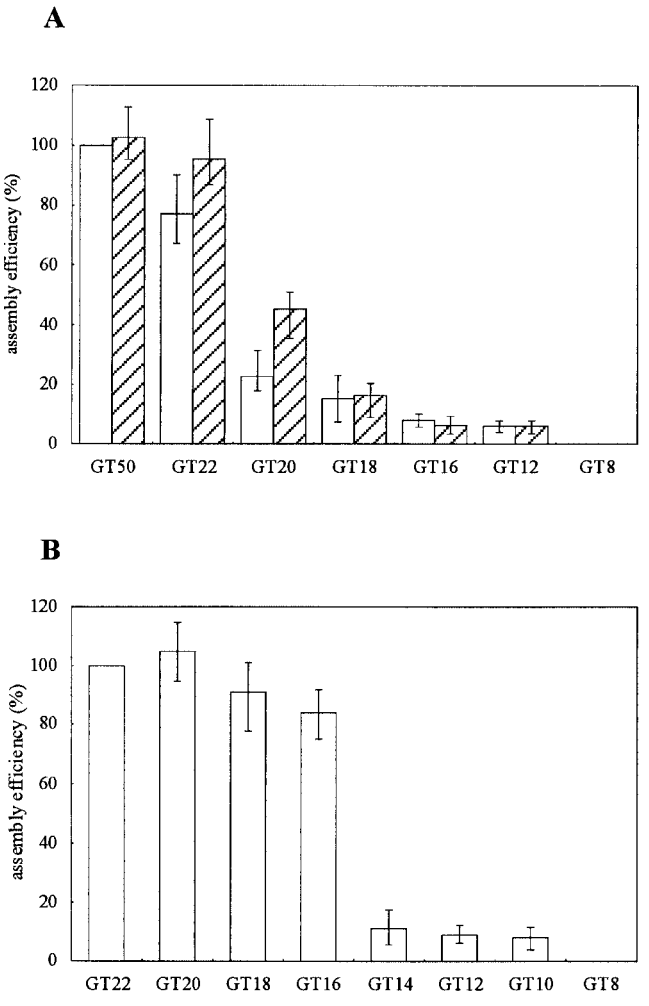


FIG. 3. EM analysis of assembly efficiency. Different GT oligonucleotides were used to support assembly in vitro with ΔMBDΔPR protein (A) or ΔMBDΔPR-Zn (B). VLPs were counted on the EM grids after incubation of the assembly reaction at room temperature for 30 min (white bars) or at 4°C for 24 h (hatched bars). The number of VLPs was normalized to the number of particles when either GT50 or GT22 was used to promote assembly. The averages from at least five independent assembly reactions are shown. Error bars indicate the range of VLP number observed.

oligonucleotides at least 16 nt in length, while shorter oligonucleotides supported assembly poorly (Fig. 3B).

The dramatic effect of size on assembly efficiency is not due to a binding defect for shorter DNAs, is not sequence specific, and is not kinetic in nature. The K_a of shorter oligonucleotides decreased gradually, as predicted from the assumption that K_i is constant (Table 1). When AC22 and AC20 were used to promote assembly with ΔMBDΔPR, a similar difference was observed as for GT22 compared with GT20 (data now shown). To determine whether there is interplay between oligonucleotide length and assembly kinetics, the assembly reactions were incubated for 24 h at 4°C (instead of the usual 30 min at 22°C) with ΔMBDΔPR and the same set of GTmers (Fig. 3A). Again, GT22 promoted assembly as efficiently as GT50. The relative assembly efficiency of GT20 increased compared with the 30-min reactions but only to ca. 40% of that for GT22. The longer incubation did not change the results for DNAs shorter than 20 nt.

It is provocative that the minimum size of DNA oligonucleotides that allows efficient assembly (22 or 16 nt) corresponds to two binding sites (11 or 8 nt) for both ΔMBDΔPR and ΔMBDΔPR-Zn proteins. From these observations we hypothesize that the essential role of nucleic acid in assembly is to promote formation of Gag dimers. According to this model, the binding of two Gag molecules to a nucleic acid brings the upstream CA domains in close proximity, thus facilitating the formation of dimers via CA-CA contacts (Fig. 5). Oligonucleotides not long enough to allow binding of at least two protein molecules cannot promote dimer formation. Gag dimers thus are envisioned as the fundamental building blocks for assembly, and dimer-dimer interaction results in the polymerization of Gag into the spherical shell of the VLP, and by extension in vivo, the shell of the immature budding virion.

Inhibitory effect of excess magnesium and zinc ions on assembly. For most Gag proteins it is known that high ionic strength prevents in vitro assembly, presumably by blocking the stable binding of the NC domain to nucleic acid. In the process of further optimizing the RSV in vitro assembly system, we found that pH and the presence of some divalent cations can compromise particle formation without interfering with nucleic acid binding. At between pH 6 and 8 there was no difference in protein-oligonucleotide binding, but the assembly efficiency, as well as the integrity of the VLPs, was greatly reduced at the higher pH values (68; also data not shown). Surprisingly, Zn^{2+} and Mg^{2+} also had profound effects on assembly without affecting binding. $ZnCl_2$ in the several hundred micromolar concentration range inhibited the in vitro formation of VLPs by ΔMBDΔPR-Zn in the presence of GT22, as evidenced by lower numbers of VLPs counted on EM grids (data not shown). Similar inhibition has been observed for in vitro assembly in the HIV-1 system (I. Gross, unpublished data). However, the binding between ΔMBDΔPR-Zn and GT22 was not affected; essentially the same binding constants were determined without extra zinc or with 200 μM $ZnCl_2$ in the reaction (data not shown). In contrast to in vitro assembly of ΔMBDΔPR-Zn protein, assembly of ΔMBDΔPR was not inhibited by zinc (data not shown), implying that the inhibitory effect is mediated by functional zinc finger motifs. Presumably, once the Zn^{2+} has been removed from the protein by EDTA, the binding motif undergoes irreversible oxidative changes preventing Zn^{2+} binding.

Magnesium ions also inhibited assembly in vitro (Fig. 4). In the In the filter-binding assay, we detected no difference in the intrinsic binding constant K_i of ΔMBDΔPR or ΔMBDΔPR-Zn to GT22, in the presence or absence of several millimolar $MgCl_2$, even though the apparent binding constant K_a differed somewhat due to the difference in binding site size (Fig. 4A and B). On the other hand, $MgCl_2$ had a distinct effect on formation of VLPs. In the case of ΔMBDΔPR, assembly efficiency decreased dramatically with increasing magnesium concentrations, with very few particles being observed at 5 mM $MgCl_2$, a concentration that is within the physiological range (Fig. 4C). In contrast, Mg^{2+} did not affect the assembly of ΔMBDΔPR-Zn, with similar numbers of VLPs being observed. We speculate that inhibition by Mg^{2+} leads to incorrect folding

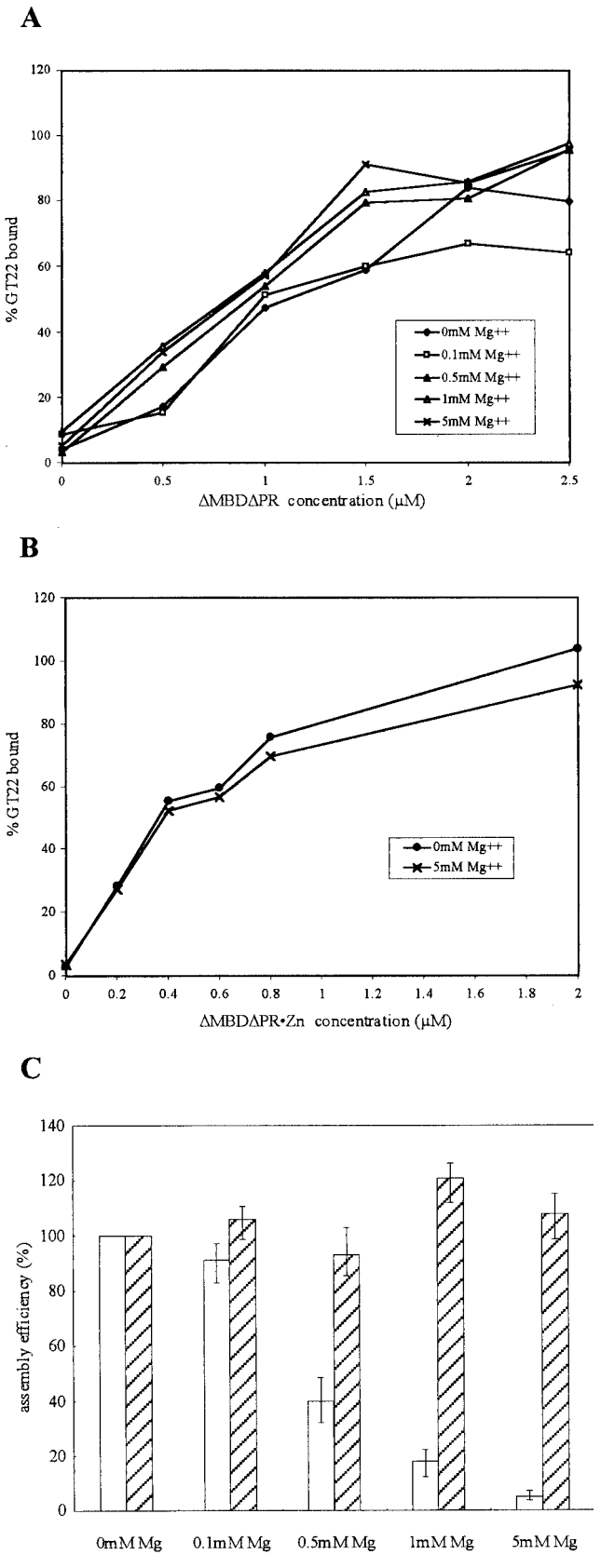


FIG. 4. Effect of magnesium on binding and assembly. (A and B) Nitrocellulose filter-binding analysis with ΔMBDΔPR (A) or ΔMBDΔPR-Zn (B) with GT22 at various $MgCl_2$ concentrations. (C) EM analysis of assembly efficiency with ΔMBDΔPR (open bars) or ΔMBDΔPR-Zn (hatched bars) at various $MgCl_2$ concentrations.

of a part of the polypeptide, presumably encompassing the CCHC motifs in NC or nearby sequences. Correctly prebound Zn^{2+} could mask this effect.

DISCUSSION

The results presented in this work show a correlation between the in vitro assembly efficiency of an RSV Gag protein on DNA oligonucleotides of different lengths and the binding site size of the protein on the DNA. The oligonucleotides all were made of the repeating sequence $(GT)_n$, so that the sequence recognized by the protein is identical. In two independent assays ΔMBDΔPR, a slightly truncated and soluble version of RSV Gag, had a binding site size of 8 or 11 nt in the presence or absence of Zn^{2+} bound to the Cys-His motifs of the NC domain, respectively. In a quantitative EM assay, the smallest oligonucleotide that could support efficient assembly was twice the size of the binding site, i.e., 16 or 22 nt, respectively. Under certain conditions the presence of either excess Zn^{2+} or Mg^{2+} inhibited assembly without altering the binding constant, implying that interaction with nucleic acid is necessary but not sufficient for the formation of VLPs in vitro. Based on the 1:2 ratio of binding site size to oligonucleotide size needed for efficient assembly, we propose a model in which the role of nucleic acid in assembly is to promote formation of Gag dimers, which themselves are crucial intermediates in assembly.

Our estimates of the binding site size of RSV Gag and NC proteins on GTmers are congruent with published values for the binding site sizes of other retroviral NC proteins on homopolymeric RNAs. Fluorescence-quenching measurements for MLV NC and RSV NC originally indicated that that protein covers approximately 6 to 8 nt (32). Later, HIV-1 and SIV NC were shown to have similar binding site sizes (56, 57, 67). In the presence of EDTA, the binding site size of HIV-1 NC was reported to increase from 6 to 9 nt, paralleling our observation with ΔMBDΔPR (56). Such an increase might be anticipated from unraveling of the zinc fingers, which would cause the protein to become more extended. Not all reports on binding site sizes of retroviral NC proteins or NC domains are consistent, however. Two studies on larger versions of NC reported larger binding site sizes. In one case the authors inferred the existence of two types of protein-nucleic acid complexes formed under low-salt conditions with an HIV-1 NC extended by 16 amino acids at the C terminus (NC71) (33, 34). The binding site sizes were 8 and 14 nt, respectively. Proteolytic truncation of the C-terminal 14 amino acids from NC71 removed the binding site size heterogeneity, with the smaller binding site being the only one observed. In the only report of the binding site size of a full-length Gag precursor, MLV Gag appeared to cover up to 20 nt in a fluorescence-quenching assay (32). However, that protein had been purified under denaturing conditions, and thus the data may not be biologically relevant.

One might expect that the binding of ΔMBDΔPR to long

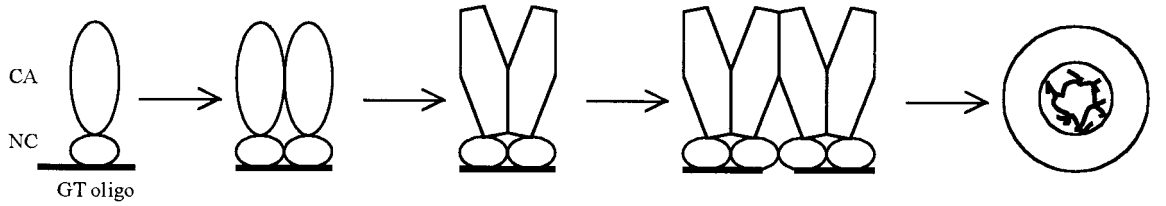


FIG. 5. Dimerization model. Only the CA and NC domains of Gag are shown. Upon binding to a nucleic acid, two adjacent Gag protein molecules undergo a conformational change. The resulting dimer has a newly exposed surface that mediates dimer-dimer interaction. The Gag dimers are used as building blocks for assembly of the virus particle.

oligonucleotides should exhibit cooperativity. That is, if there are protein-protein interactions in a complex, then binding of a free protein molecule to a complex of one protein with a DNA should be favored over binding of the free protein to a free DNA molecule. Our filter-binding assays gave no evidence for such cooperativity. First, binding to large oligonucleotides able to hold more than one protein molecule was not tighter than the binding to oligonucleotides able to hold only one protein, once the number of binding sites was taken into account (Tables 2 and 3). Second, the rise in the percent oligonucleotide retained on the filter as a function of protein concentration was approximately linear and not sigmoidal as expected for a cooperative reaction. The limited accuracy of this type of measurement cannot exclude a low level of cooperativity, but overall the data suggest that protein-protein interactions must be weak.

We previously suggested a scaffold model to explain the role of nucleic acid in assembly in vitro. Nucleic acid was envisioned to serve as a binding platform, concentrating Gag molecules in space and thereby promoting intrinsically weak protein-protein interactions between the CA domains or other parts of Gag (8). The idea that nucleic acid plays an essential role in assembly is consistent with recent observations in vivo, in which a variety of RNA species was found to be incorporated into MLV virions in the absence of a genomic packaging sequence (47). However, the notion that nucleic acid serves only to concentrate Gag molecules does not adequately explain how DNA oligonucleotides and small RNAs can support assembly in vitro as efficiently as much longer RNAs such as retroviral genomes. A variation of the scaffold model (called "bridging") that might provide such an explanation is based on the observation that under some conditions HIV-1 NC can form ternary complexes (A. Rein and R. Fisher, unpublished data). If one NC domain could bind simultaneously to the ends of two oligonucleotides, it might be able to link many short nucleic acid molecules together to form a chain, which would be functionally like a single long strand. Were bridging to play a paramount role in assembly in vitro, the smallest assembly-competent oligonucleotide should be ca. 8 nt in length, long enough to accommodate two half sites for the NC domain of Gag. Our results with ΔMBDΔPR in the presence of Zn^{2+} clearly are not consistent with this expectation, since no assembly occurred in the presence of GT8, even though binding was readily detectable at protein concentrations typically used in assembly reactions. These findings and observations on the inhibition of assembly by some divalent cations underscore the conclusion that protein-nucleic acid interaction, while necessary, is not sufficient for assembly.

We propose a new model to explain how the interaction of Gag with nucleic acid in general, and with short DNA oligonucleotides in particular, leads to assembly in vitro (Fig. 5). In this model the function of nucleic acid is to promote the formation of Gag dimers, which themselves are the immediate building blocks for assembly. For the dimerization step to occur, the oligonucleotide must be long enough to accommodate two Gag molecules simultaneously. For ΔMBDΔPR in presence of Zn^{2+}, this size is 16 nt. We envision that Gag dimerization is accompanied by conformational changes that facilitate dimer-dimer interactions, ultimately leading to polymerization of the dimers into a spherical shell. To explain the inefficient assembly observed with oligonucleotides smaller than two binding sites but larger than one (10 to 14 nt for ΔMBDΔPR in the presence of Zn^{2+}), we speculate that the NC domain can adopt a conformation that allows it to bind to a smaller site, presumably with lower affinity. This speculation is based on the known flexibility of the NC polypeptide, which might allow it become more compact, as suggested previously for a C-terminally extended version of HIV-1 NC (34). Such an alternative conformation and binding mode would give fewer contacts with the nucleic acid backbone and be less energetically favorable.

Support for Gag dimerization as a critical step in assembly comes from the properties of chimeric Gag molecules in which the NC domain is replaced by a leucine zipper (LZ), a protein dimerization motif also called a coiled-coil domain. Two groups have shown in the HIV-1 system that in transfected cells expressing Gag-LZ proteins lacking the NC domain, budding occurs almost as efficiently as for wild-type Gag (2, 69). While these reports did not address the morphology of the resulting particles, we have shown recently that similar RSV Gag-LZ chimeric proteins lacking NC, when expressed in a baculovirus-insect cell system, form regular budding structures and VLPs (M. Johnson and V. M. Vogt, unpublished results). The particles are either spherical, resembling immature virions, or tubular, resembling some Gag mutants, with the morphology depending on how the LZ is juxtaposed to the CA domain. These results imply that dimer formation mediated by an LZ can drive morphologically similar assembly in vivo.

It remains unknown what Gag-Gag interactions underlie the formation of Gag dimers or the further polymerization of dimers into the shell of ~1,500 Gag molecules that comprise an immature virion (59). Mature HIV-1 CA, which like all CA

proteins is composed of two domains, readily forms dimers in solution, with a K_d of 10 to 30 μM (21). The dimer interface is in the CTD and includes part of the highly conserved major homology region. By extrapolation it is sometimes assumed that all retroviral CA proteins form dimers in a fashion analogous to that in HIV-1 CA, but this assumption may not be warranted. RSV CA remains monomeric in solution even at concentrations approaching 1 mM (10, 36). Similarly, the CA proteins of equine infectious anemia virus and of human T-cell leukemia virus type 1 do not show obvious dimerization in solution (4, 35). Protein-protein contacts in CA crystals may provide clues to the nature of Gag multimerization, but the clues are difficult to decipher since these contacts are different for different species of CA. The most enlightening paradigm for Gag multimerization is derived from the cryo-EM reconstruction of tubes assembled in vitro with HIV-1 CA (43). In this structure, which represents the mature core of the virus, planar hexagonal rings of CA are held together by contacts between the N-terminal domains (NTDs) at the surface of the tube. At the same time each CA molecule is linked to a molecule in an adjoining hexagonal ring by interactions between the CTDs, which project inward from the surface of the tube. What may be similar hexagonal rings also have been observed in MLV as well as RSV CA proteins polymerized as two-dimensional arrays on lipid layers (44a, 73, 74). It is unclear to what extent these structures apply to the immature Gag shell, but at the least they provide a conceptual framework for picturing Gag multimerization. In this framework both the NTD and the CTD of CA provide contacts that help hold the spherical particle together.

For RSV assembly in vitro, we speculate that NC binding to nucleic acid first promotes an interaction between CTDs of CA and/or the 12-amino-acid spacer (SP) separating CA from NC and that polymerization of Gag dimers then occurs by interaction of the NTDs. We envision the dimer formation to be accompanied by a conformational change that helps lock the dimer together. The SP region may be the locus of such a conformational change, since both in RSV Gag and in the analgous portion of HIV-1 Gag (called SP1 or p2), the amino acid sequence is critically important for assembly. Deletions or mutations in SP lead to grossly aberrant budding structures in vivo, and deletion of HIV-1 SP1 changes the morphology of in vitro assembled particles from spheres to tubes (1, 14, 29, 39, 41). The SP1 sequence in HIV-1 is predicted to form an alpha helix, but as a C-terminal extension of the RSV or HIV-1 CTD it has been found to be unstructured (36, 63). Other portions of Gag also are likely to provide protein-protein interactions in vivo. For example, HIV-1 MA can form trimers, both as crystals and in solution (30, 46). However, in RSV ΔMBDΔPR the N-terminal membrane-binding domain of MA is missing, and most of the remaining amino acid residues in MA and the adjoining p2 and p10 domains are dispensable for spherical particle assembly in vitro. Hence, at least in RSV, interprotein contacts in vitro may be restricted to the CA, SP, and NC domains.

The hypothesis that Gag dimer formation is critical for assembly needs to be supported by direct evidence for the existence of dimeric species of Gag. In vitro assembly carried out under published condition shows no evidence of such species (data not shown), which we interpret to mean that dimers are transient intermediates that are rapidly incorporated into the growing particle. More dilute conditions might allow the visualization of dimers. Formation of HIV-1 particles in a translationally programmed crude extract is reported to be accompanied by assembly intermediates, but the nature of these intermediates remains to be clarified (44). Dynamic light scattering has been used successfully to identify many assembly features of hepatitis B virus and cowpea chlorotic virus and thus may prove a useful tool for studying retrovirus assembly in vitro (71, 72). A paradigm for the role of protein dimers as dynamic assembly intermediates is provided by alphaviruses. The NC protein of Sindbis virus is similar to retroviral Gag protein in that it assembles spontaneously into viral cores in vitro in the presence of nucleic acid. Sindbis virus NC protein is monomeric in solution but dimeric in the virus crystal structure. During the in vitro assembly reaction dimers can be trapped by chemical cross-linking, and purified cross-linked dimers are readily incorporated into the assembled core in the presence of wild-type protein (54, 55). Similar techniques may help elucidate the role of dimerization in retrovirus assembly.

ACKNOWLEDGMENTS

We thank Marc Johnson, Gisela Schatz, Ingolf Gross, and Amanda Dalton for critical reading of the manuscript and Alan Rein, John Wills, Rebecca Craven, and Stephen Campbell for extensive discussions. We thank Robert Gorelick for generously providing the purified HIV-1 and RSV NC proteins used in this study.

The work was supported by USPHS grant CA20081.

REFERENCES

1. **Accola, M. A., S. Höglund, and H. G. Göttlinger.** 1998. A putative α-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. J. Virol. **72:**2072–2078.
2. **Accola, M. A., B. Strack, and H. G. Göttlinger.** 2000. Efficient particle production by minimal Gag constructs which retain the carboxy-terminal domain of human immunodeficiency virus type 1 capsid-p2 and a late assembly domain. J. Virol. **74:**5395–5402.
3. **Berkowitz, R., J. Fisher, and S. P. Goff.** 1996. RNA packaging. Curr. Top. Microbiol. Immunol. **214:**177–218.
4. **Birkett, A. J., B. Yelamos, I. Rodriguez-Crespo, F. Gavilanes, and D. L. Peterson.** 1997. Cloning, expression, purification and characterization of the major core protein (p26) from equine infectious anemia virus. Biochim. Biophys. Acta **1339:**62–72.
5. **Bowzard, J. B., R. P. Bennett, N. K. Krishna, S. M. Ernst, A. Rein, and J. W. Wills.** 1998. Importance of basic residues in the nucleocapsid sequence for retrovirus Gag assembly and complementation rescue. J. Virol. **72:**9034–9044.
6. **Campbell, S., and A. Rein.** 1999. In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J. Virol. **73:**2270–2279.
7. **Campbell, S., R. J. Fisher, E. M. Towler, S. Fox, H. J. Issaq, T. Wolfe, L. R. Phllips, and A. Rein.** 2001. Modulation of HIV-like particle assembly in vitro by inositol phosphates. Proc. Natl. Acad. Sci. USA **98:**10875–10879.
8. **Campbell, S., and V. M. Vogt.** 1995. Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. **69:**6487–6497.
9. **Campbell, S., and V. M. Vogt.** 1997. In vitro assembly of virus-like particles with Rous sarcoma virus Gag deletion mutants: identification of the p10 domain as a morphological determinant in the formation of spherical particles. J. Virol. **71:**4425–4435.
10. **Campos-Olivas, R., J. L. Newman, and M. F. Summers.** 2000. Solution structure of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. J. Mol. Biol. **296:**633–649.
11. **Cimarelli, A., and J. Luban.** 2000. Human immunodeficiency virus type 1 virion density is not determined by nucleocapsid basic residues. J. Virol. **74:**6734–6740.
12. **Cimarelli, A., S. Sandin, S. Höglund, and J. Luban.** 2000. Basic residues in human immunodeficiency virus type-1 nucleocapsid promote virion assembly via interaction with RNA. J. Virol. **74:**3046–3057.
13. **Craven, R. C., R. N. Harty, J. Paragas, P. Palese, and J. W. Wills.** 1999. Late domain function identified in the vesicular stomatitis virus M protein by use of rhabdovirus-retrovirus chimeras. J. Virol. **73:**3359–3365.

14. **Craven, R. C., A. E. Leure-duPree, C. R. Erdie, C. B. Wilson, and J. W. Wills.** 1993. Necessity of the spacer peptide between CA and NC in the Rous sarcoma virus Gag protein. J. Virol. **67:**6246–6252.
15. **Dannull, J., A. Surovoy, G. Jung, and K. Moelling.** 1994. 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. **13:**1525–1533.
16. **Dawson, L., and X.-F. Yu.** 1998. The role of nucleocapsid of HIV-1 in virus assembly. Virology **251:**141–157.
17. **Déméné, H., N. Julian, H. deRocquigny, F. Cornille, B. Maigret, and B. P. Roques.** 1994. Three-dimensional ^{1}H NMR structure of the nucleocapsid protein NCp10 of Moloney murine leukemia virus. J. Biomol. NMR **4:**153–170.
18. **Ehrlich, L. S., B. E. Agresta, and C. A. Carter.** 1992. Assembly of recombinant human immunodeficiency virus type 1 capsid protein in vitro. J. Virol. **66:**4874–4883.
19. **Epstein, I. R.** 1978. Cooperative and non-cooperative binding of large ligands to a finite one-dimensional lattice: a model for ligand-oligonucleotide interactions. Biophys. Chem. **8:**327–339.
20. **Fisher, R. J., A. Rein, M. Fivash, M. A. Urbenja, J. R. Casas-Finet, M. Medaglia, and L. E. Henderson.** 1998. Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. J. Virol. **72:**1902–1909.
21. **Gamble, T. L., S. Yoo, F. F. Vajdos, U. K. von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist, and C. P. Hill.** 1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. Science **278:**849–853.
22. **Ganser, B. K., S. Li, V. Y. Klishko, J. T. Finch, and W. I. Sundquist.** 1999. Assembly and analysis of conical models for the HIV-1 core. Science **283:**80–83.
23. **Gao, Y., K. Kaluarachchi, and D. P. Giedroc.** 1998. Solution structure and backbone dynamics of Mason-Pfizer monkey virus nucleocapsid protein. Protein Sci. **7:**2265–2280.
24. **Garnier, L., J. W. Wills, M. F. Verderame, and M. Sudol.** 1996. WW domains and retrovirus budding. Nature **381:**744–745.
25. **Göttlinger, H. G., T. Dorfman, J. G. Sodroski, and W. A. Haseltine.** 1991. Effect of mutations affecting the p6 Gag protein on human immunodeficiency virus particle release. Proc. Natl. Acad. Sci. USA **88:**3195–3199.
26. **Grättinger, M., H. Hohenberg, D. Thomas, T. Wilk, B. Müller, and H.-G. Kräusslich.** 1999. In vitro assembly properties of wild-type and cyclophilin-binding defective human immunodeficiency virus capsid proteins in the presence and absence of cyclophilin A. Virology **257:**247–260.
27. **Gross, I., H. Hohenberg, and H.-G. Kräusslich.** 1997. In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. Eur. J. Biochem. **249:**592–600.
28. **Gross, I., H. Hohenberg, C. Huckhagel, and H.-G. Kräusslich.** 1998. N-terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. J. Virol. **72:**4798–4810.
29. **Gross, I., H. Hohenberg, T. Wilk, K. Wiegers, M. Grättinger, B. Müller, S. Fuller, and H.-G. Kräusslich.** 2000. A conformational switch controlling HIV-1 morphogenesis. EMBO J. **19:**103–113.
30. **Hill, C. P., D. Worthylake, D. P. Bancroft, A. M. Christensen, and W. I. Sundquist.** 1996. Crystal structures of the trimeric HIV-1 matrix protein: implications for membrane association and assembly. Proc. Natl. Acad. Sci. USA **93:**3099–3104.
31. **Joshi, S. M., and V. M. Vogt.** 2000. Role of the Rous sarcoma virus p10 domain in shape-determination of Gag virus-like particles assembled in vitro and within *Escherichia coli*. J. Virol. **74:**10260–10268.
32. **Karpel, R. L., L. E. Henderson, and S. Oroszlan.** 1987. Interaction of retroviral structural proteins with single-stranded nucleic acids. J. Biol. Chem. **262:**4961–4967.
33. **Khan, R., and D. P. Giedroc.** 1992. Recombinant human immunodeficiency virus type 1 nucleocapsid (NC^{p7}) protein unwinds tRNA. J. Biol. Chem. **267:**6689–6695.
34. **Khan, R., and D. P. Giedroc.** 1994. Nucleic acid binding properties of recombinant Zn_2 HIV-1 nucleocapsid protein are modulated by COOH-terminal processing. J. Biol. Chem. **269:**22538–22546.
35. **Khorasanizadeh, S., R. Campos-Olivas, and M. F. Summers.** 1999. Solution structure of the capsid protein from the human T-cell leukemia virus type-1. J. Mol. Biol. **291:**491–505.
36. **Kingston, R., E. Z. Eisenmesser, T. Fitzon-Ostendorp, G. W. Schatz, V. M. Vogt, C. B. Post, and M. G. Rossmann.** 2000. Structure and self-association of the Rous sarcoma virus capsid protein. Structure **8:**617–628.
37. **Klein, D., P. E. Johnson, E. S. Zollars, R. N. De Guzman, and M. F. Summers.** 2000. The NMR structure of the nucleocapsid protein from the mouse mammary tumor virus reveals unusual folding of the C-terminal zinc knuckle. Biochemistry **39:**1604–1612.
38. **Klikova, M., S. S. Rhee, E. Hunter, and T. Ruml.** 1995. Efficient in vivo and in vitro assembly of retroviral capsids from Gag precursor proteins expressed in bacteria. J. Virol. **69:**1093–1098.
39. **Kräusslich, H.-G., M. Fäcke, A.-M. Heuser, J. Konvalinka, and H. Zentgraf.** 1995. The spacer peptide between human immunodeficiency virus capsid and nucleocapsid proteins is essential for ordered assembly and viral infectivity. J. Virol. **69:**3407–3419.
40. **Kräusslich, H.-G., and R. Welker.** 1996. Intracellular transport of retroviral capsid components. Curr. Top. Microbiol. Immunol. **214:**25–63.
41. **Krishna, N. K., S. Campbell, V. M. Vogt, and J. W. Wills.** 1998. Genetic determinants of Rous sarcoma virus particle size. J. Virol. **72:**564–577.
42. **Lee, B. M., R. N. De Guzman, B. G. Turner, N. Tjandra, and M. F. Summers.** 1998. Dynamical behavior of the HIV-1 nucleocapsid protein. J. Mol. Biol. **279:**633–649.
43. **Li, S., C. P. Hill, W. I. Sundquist, and J. T. Finch.** 2000. Image reconstructions of helical assemblies of the HIV-1 CA protein. Nature **407:**409–413.
44. **Lingappa, J. R., R. L. Hill, M. L. Wong, and R. S. Hegde.** 1997. A multistep ATP-dependent pathway for assembly of human immunodeficiency virus capsids in a cell-free system. J. Cell Biol. **136:**567–581.

44a. **Mayo, K., M. L. Vana, J. McDermott, D. Huseby, J. Leis, and E. Barklis.** 2002. Analysis of Rous sarcoma virus capsid protein variants assembled on lipid bilayers. J. Mol. Biol. **316:**667–678.

45. **Morikawa, Y., T. Toshiyuki, and K. Sano.** 1999. In vitro assembly of human immunodeficiency virus type 1 Gag protein. J. Biol. Chem. **274:**27997–28002.
46. **Morikawa, Y., W.-H. Zhang, D. J. Hockley, M. V. Nermut, and I. M. Jones.** 1998. Detection of a trimeric human immunodeficiency virus type 1 gag intermediate is dependent on sequences in the matrix protein p17. J. Virol. **72:**7659–7663.
47. **Muriaux, D., J. Mirro, D. Harvin, and A. Rein.** 2001. RNA is a structural element in retrovirus particles. Proc. Natl. Acad. Sci. USA **98:**5246–5251.
48. **Ono, A., and E. O. Freed.** 1999. Binding of human immunodeficiency virus type 1 Gag to membrane: role of the matrix amino terminus. J. Virol. **73:**4136–4144.
49. **Parent, L. J., R. P. Bennett, R. C. Craven, T. D. Nelle, N. K. Krishna, J. B. Bowzard, C. B. Wilson, B. A. Puffer, R. C. Montelaro, and J. W. Wills.** 1995. Positionally independent and exchangeable late budding functions of the RSV and HIV Gag proteins. J. Virol. **69:**5455–5460.
50. **Parent, L. J., C. B. Wilson, M. D. Resh, and J. W. Wills.** 1996. Evidence for a second function of the MA sequence in the Rous sarcoma virus Gag protein. J. Virol. **70:**1016–1026.
51. **Puffer, B. A., L. J. Parent, J. W. Wills, and R. C. Montelaro.** 1997. Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein. J. Virol. **71:**6541–6546.
52. **Puffer, B. A., S. C. Watkins, and R. C. Montelaro.** 1998. EIAV Gag polyprotein late domain recruits the cellular AP-2 adaptor protein complex to facilitate viral budding. J. Virol. **72:**10218–10221.
53. **Steeg, C. M., and V. M. Vogt.** 1990. RNA-binding properties of the matrix protein ($p19^{Gag}$) of avian sarcoma and leukemia viruses. J. Virol. **64:**847–855.
54. **Tellinghuisen, T. L., and R. J. Kuhn.** 2000. Nucleic acid-dependent cross-linking of the nucleocapsid protein of Sindbis virus. J. Virol. **74:**4302–4309.
55. **Tellinghuisen, T. L., R. Perera, and R. J. Kuhn.** 2001. In vitro assembly of Sindbis virus core-like particles from cross-linked dimers of truncated and mutant capsid proteins. J. Virol. **75:**2810–2817.
56. **Urbaneja, M. A., B. P. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson, and J. R. Casas-Finet.** 1999. 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. **287:**59–75.
57. **Urbaneja, M. A., C. F. McGrath, B. P. Kane, L. E. Henderson, and J. R. Casas-Finet.** 2000. Nucleic acid binding properties of the simian immunodeficiency virus nucleocapsid protein NCp8. J. Biol. Chem. **275:**10394–10404.
58. **Verderame, M. F., T. D. Nelle, and J. W. Wills.** 1996. The membrane-binding domain of the Rous sarcoma virus Gag protein. J. Virol. **70:**2664–2668.
59. **Vogt, V. M., and M. N. Simon.** 1999. Mass determination of Rous sarcoma virus virions by scanning transmission electron microscopy (STEM). J. Virol. **73:**7050–7055.
60. **von Schwedler, U. K., T. L. Stemmler, V. Y. Klishko, S. Li, K. H. Albertine, D. R. Davis, and W. I. Sundquist.** 1998. Proteolytic refolding of the HIV-1 capsid protein amino terminus facilitates viral core assembly. EMBO J. **17:**1555–1568.
61. **Wills, J. W., R. C. Craven, R. A. Weldon, Jr., T. D. Nelle, and C. R. Erdie.** 1991. Suppression of retroviral MA deletions by the amino-terminal membrane-binding domain of $p60^{src}$. J. Virol. **65:**3804–3812.
62. **Wills, J. W., C. E. Cameron, C. B. Wilson, Y. Xiang, R. P. Bennett, and J. Leis.** 1994. An assembly domain of the Rous sarcoma virus Gag protein required late in budding. J. Virol. **68:**6605–6618.
63. **Worthylake, D. K., H. Wang, S. Yoo, W. I. Sundquist, and C. P. Hill.** 1999. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. Acta Crystallogr. D **55:**85–92.
64. **Wu, W., L. E. Henderson, T. D. Copeland, R. J. Gorelick, W. J. Bosche, A. Rein, and J. G. Levin.** 1996. Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract. J. Virol. **70:**7132–7142.
65. **Xiang, Y., C. E. Cameron, J. W. Wills, and J. Leis.** 1996. Fine mapping and characterization of Rous sarcoma virus $Pr76^{gag}$ late assembly domain. J. Virol. **70:**5695–5700.
66. **Yasuda, J., and E. Hunter.** 1998. A proline-rich motif (PPPY) in the Gag

polyprotein of Mason-Pfizer monkey virus plays a maturation-independent role in virion release. J. Virol. **72:**4095–4103.

67. **You, J. C., and C. S. McHenry.** 1993. HIV nucleocapsid protein: expression in *Escherichia coli*, purification, and characterization. J. Biol. Chem. **268:** 16519–16527.
68. **Yu, F., S. M. Joshi, Y. M. Ma, R. L. Kingston, M. N. Simon, and V. M. Vogt.** 2001. Characterization of Rous saroma virus Gag particles assembled in vitro. J. Virol. **75:**2753–2764.
69. **Zhang, Y., H. Qian, Z. Love, and E. Barklis.** 1998. Analysis of the assembly function of the human immunodeficiency virus type 1 Gag protein nucleocapsid domain. J. Virol. **72:**1782–1789.
70. **Zhou, W., L. J. Parent, J. W. Wills, and M. D. Resh.** 1994. Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. J. Virol. **68:**2556–2569.
71. **Zlotnick, A., R. Aldrich, J. M. Johnson, P. Ceres, and M. J. Young.** 2000. Mechanism of capsid assembly for an icosahedral plant virus. Virology **277:** 450–456.
72. **Zlotnick, A., J. M. Johnson, P. W. Wingfield, S. J. Stahl, and D. Endres.** 1999. A theoretical model successfully identifies features of hepatitis B virus capsid assembly. Biochemistry **38:**14644–14652.
73. **Zuber, G., J. McDermott, S.Karanjia, W. Zhao, M. F. Schmid, and E. Barklis.** 2000. Assembly of retrovirus capsid-nucleocapsid proteins in the presence of membranes or RNA. J. Virol. **74:**7431–7441.
74. **Zuber, G., and E. Barklis.** 2000. Atomic force microscopy and electron microscopy analysis of retrovirus Gag proteins assembled in vitro on lipid bilayers. Biophys. J. **78:**373–384.