TY3 GAG3 protein forms ordered particles in Escherichia coli

Permalink
https://escholarship.org/uc/item/81t898mq

Journal
Virology, 370(2)

ISSN
0042-6822

Authors
Larsen, LSZ
Kuznetsov, Y
McPherson, A
et al.

Publication Date
2008-01-20

DOI
10.1016/j.virol.2007.09.017

License
CC BY 4.0

Peer reviewed
Ty3 Gag3 Protein Forms Ordered Particles in Escherichia Coli

Liza S.Z. Larsen¹,²,³, Yurii Kuznetsov⁴, Alex McPherson⁴, G.Wesley Hatfield²,³,⁵, and Suzanne Sandmeyer¹,²,³,*

¹Department of Biological Chemistry, University of California, Irvine, CA 92697
²School of Information and Computer Sciences and Institute for Genomics and Bioinformatics, University of California, Irvine, CA 92697
³Department of Microbiology and Molecular Genetics, University of California, Irvine, CA 92697
⁴Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697
⁵CODA Genomics, Inc. Laguna Hills, CA 92653

Abstract

The yeast retrovirus-like element Ty3 GAG3 gene encodes a Gag3 polyprotein analogous to retroviral Gag. Gag3 lacks matrix, but contains capsid, spacer, and nucleocapsid domains. Expression of a Ty3 Gag3 or capsid domain optimized for expression in Escherichia coli, was sufficient for Ty3 particle assembly. Virus-like ordered particles assembled from Gag3 were similar in size to immature particles from yeast and contained nucleic acid. However, particles assembled from the CA domain were variable in size and displayed much less organization than native particles. These results indicate that assembly can be driven through interactions among capsid subunits in the particle, but that the nucleocapsid domain, likely in association with RNA, confers order upon this process.

Keywords

retrotransposon; retrovirus; virus-like particles; capsid; Escherichia coli; yeast Ty3

Introduction

Gag, the major structural component of retrovirus cores, typically contains matrix (MA), capsid (CA), nucleocapsid (NC) and one or more spacer (SP) domains as well as other virus-specific domains. Within the Gag context, MA functions in membrane targeting; CA mediates Gag interaction; and NC packages RNA and promotes Gag interaction (Adamson & Jones, 2004; Scarlata & Carter, 2003). NC is a small domain rich in basic residues that typically contains one or two copies of a CX₂CX₄HX₄C zinc-binding motif. After particle assembly and budding, these domains are separated by proteolytic cleavage and characteristic cores are formed. Expression of Gag in heterologous systems, such as baculovirus and Escherichia coli has shown formation of virus-like particles (VLPs). In the case of HIV, depending on salt and pH conditions CA can form tubes, spheres, and cones in vitro. Inclusion of the NC domain and RNA enhances formation of sheet-like structures or cylindrical forms (Ehrlich, Agresta et

*Corresponding author: Phone: 949-824-7571, Email: sbsandme@uci.edu.

Publisher’s Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Formation of spheres approximating immature particle forms occurs when regions amino-terminal to CA are included. This effect is attributed to constraint of the amino terminus of CA which rearranges during maturation and is engaged in an internal salt bridge in the mature form (von Schwedler, Stemmler et al., 1998; Gross, Hohenberg et al., 1998; Joshi & Vogt, 2000; Rumlova-Klikova, Hunter et al., 2000; Campbell & Vogt, 1997).

The budding yeast retrotransposon Ty3 has a Gag3 domain comprised of CA, SP3 and NC domains (Fig. 1) (Sandmeyer, Aye et al., 2002). Initial assembly of Ty3 results in a roughly spherical VLP which remains in a similar form even after proteolytic maturation (Kuznetsov, Zhang et al., 2005). The Ty3 precursor differs from the retroviral precursor in that it lacks the amino-terminal extension of CA into MA, as well as the conserved amino-terminal proline of CA. In this work we determined that expression of Ty3 Gag3 was sufficient to form spherical particles in a bacterial expression system. The organization of particles was greatly enhanced by inclusion of the NC domain.

**Results**

**Assembly of recoded GAG3**

Because Ty3 Gag3 protein is comprised only of CA, SP, and NC domains, it was of interest to determine whether its expression was sufficient to support spherical particle formation in a heterologous system. A synthetic Ty3 GAG3 gene optimized for bacterial protein expression produced significant levels of Gag3 (Fig. 2, lanes 1-3). Little if any protein expression was observed from the native yeast Ty3 GAG3 gene (Fig. 2, lanes 4-6). About 40% of the recoded Gag3 was soluble (Fig. 3A). The nature of the insoluble protein fraction was not explored, but it could have contained Gag3 inclusion bodies or aggregates.

**Atomic force microscopy analysis of E. coli cell lysates expressing Gag3**

Approximately 500 μg of cleared lysate from cells expressing recoded Gag3 was fractionated over a 20/30/70% sucrose step gradient as previously described (Kuznetsov, Zhang, Menees, McPherson, & Sandmeyer, 2005). Gradient fractions were analyzed by denaturing polyacrylamide gel electrophoresis. A protein of the same size as Gag3 that reacted with anti-CA antibody was the major component of fractions 7 and 8 (Fig. 3B); this is the position of migration of native Ty3 VLPs produced in yeast [(Kuznetsov, Zhang, Menees, McPherson, & Sandmeyer, 2005) and data not shown]. A 5 μl aliquot of the 70/30% interface fraction was used for AFM imaging as described previously. Imaging of the interface fraction showed particles, ranging from 30 to 50 nm in diameter (Fig. 4 A-C). These were of comparable size to protease mutant Ty3 VLPs produced in budding yeast, which had a mean of 49 nm +/- 2.3 nm [(Kuznetsov, Zhang, Menees, McPherson, & Sandmeyer, 2005) and data not shown]. A small proportion of VLPs appeared to be joined, mostly in pairs (e.g. Fig. 4C). Similar associations were observed in AFM imaging of wt Ty3 VLPs from yeast (Kuznetsov, Zhang, Menees, McPherson, & Sandmeyer, 2005). Overall, VLPs displayed ordered capsomeres and were generally similar to those previously observed. Tapping disruption of the particles resulted in release of fibrous material from some which resembled RNA (Fig. 4 G-H), suggesting that similar to retroviruses, Ty3 Gag3 can package heterologous RNA in the absence of its native genome.

**Expression of CA and AFM analysis of E. coli expression lysates**

In order to determine the contribution of SP and NC domains to VLP structure, the CA-coding domain was amplified using recoded GAG3 sequence as the template. Expression of this domain from the IPTG-inducible T7 promoter was efficient, but failed to produce sufficient
levels of soluble protein for analysis (data not shown). The construct was recloned into an arabinose-inducible expression vector (pBAD-CODA-CA) where expression was controlled to optimize solubility (data not shown). Under these optimal conditions, much less CA protein was produced than Gag3, but the proportion of soluble protein was comparable to that in the Gag3 expression system (Fig. 3A and C). Approximately 500 μg of cleared lysate from cells expressing pBAD-CODA-CA was fractionated over a 20/30/70% sucrose step gradient and 5 μl of the 30/70% interface fraction was used for AFM imaging as described above. Analysis showed partially structured particles, as well as unstructured amorphous protein aggregates (Fig. 4 D-F and data not shown). Some of these aggregates resembled double particles (Fig. 4F). A random aliquot of particulate material was measured. Particles in this fraction ranged in size from 40 to 157 nm with an overall mean of 94.2 nm.

**Discussion**

Expression of retroviral gag in heterologous systems is sufficient for particle formation. In the case of retroviruses in vitro assembly using CA with amino-terminal extensions into upstream domains results in spherical particles thought to mimic the immature capsid form, whereas in vitro assembly of CA-NC results in tubes and sheets. This difference is thought to result from constraint in the former situation of the conserved amino-terminal proline of CA, which when released in processing, forms a salt bridge required for the mature particle structure (von Schwedler, Stemmler, Klishko, Li, Albertine, Davis, & Sundquist, 1998; Gross, Hohenberg, Huckhagel, & Krausslich, 1998; Johansen & Vogt, 2000; Rumlova-Klikova, Hunter, Nermut, Pichova, & Ruml, 2000; Campbell & Vogt, 1997). In vitro assembly using MPMV mutant CA-NC lacking the amino-terminal proline residue and unable to form this bridge also resulted in spherical particles (Ulbrich, Haubova et al., 2006). Modeling of the Ty3 CA domain suggests that, similar to retroviral CA, Ty3 CA is composed of amino-terminal and carboxyl-terminal clusters of alpha helices (Larsen, Zhang et al., 2007). However, Ty3 lacks a CA amino-terminal extension and the conserved proline of CA found in retroviruses and is roughly spherical in both immature and mature states. Roughly spherical particles were formed in *E. coli* upon Ty3 Gag3 expression. This may be similar in some respects to the formation of spherical particles in vitro from MPMV CA-NC lacking the amino-terminal proline. In any case, Ty3 CA, SP and NC domains, likely together with RNA, are sufficient to specify this structure. The nature of interactions which stabilize the mature Ty3 VLPs is not known.

Studies of retroviruses have shown that NC is unnecessary for *in vitro* retroviral particle formation (Muriaux, Mirro et al., 2001). However, inclusion of the NC domain and RNA in assembly reactions promotes interaction of CA domains and multimer formation (e.g. RSV (Johnson, Scobie, Ma, & Vogt, 2002) reviewed in (Adamson & Jones, 2004)).

Expression of Ty3 CA in *E. coli* resulted in formation of highly irregular particles, although it is possible that sheets also formed and remained in the extract pellet. These results demonstrated that the SP-NC coupled to an RNA scaffold is not necessary for Ty3 Gag3 multimerization *per se*. However, the difference in Gag3 and CA particle morphology indicates that the SP-NC domains present in Gag3 and lacking in CA contribute to Ty3 particle organization and symmetry, potentially via an RNA scaffold. Although a role is implied for RNA by the contribution of SP-NC, this role is not likely to be sequence specific. Recoded GAG3 mRNA included none of the 5′ untranslated region previously shown to be important in *cis* for transposition (Kirchner, Sandmeyer et al., 1992) and by virtue of recoding, was 25% different in sequence within the translated region from the natural GAG3 sequence.

Expression of the Ty1 TYA structural protein from native RNA also resulted in particles in *E. coli* (Luschnig, Hess et al., 1995). Ty1 is more distantly related to Ty3 than are retroviruses and Ty1 does not contain a zinc-binding motif. However, in this case also, deletion of the
nucleic acid binding domain disrupted particle formation. Thus, these yeast retrotransposons both of which assemble spherical, cytoplasmic particles that undergo maturation in the absence of a membrane envelope, showed dependence upon a nucleic acid binding domain for assembly in *E. coli*.

In summary, Gag3 in the absence of Ty3 native RNA, *POL3*, and other yeast proteins is sufficient for VLP assembly into structures resembling, but not identical to, native VLPs. CA is sufficient for multimer formation. However, this work shows that the SP-NC domain, possibly together with RNA, contributes to ordering of the particle structure.

**Materials and Methods**

**Strains and culture conditions for *E. coli***

*E. coli* strain DH5α (Invitrogen, Carlsbad, CA) was used for DNA cloning and plasmid preparation of CODA-designed genes. *E. coli* strains BL21 (DE3) ((Novagen, Madison, WI), and Top10 (Invitrogen Inc., Carlsbad, CA) were used for protein expression.

**Gene cloning**

The Ty3 *GAG3* open reading frame was amplified from the pDLC201 plasmid using the primers, 5′-CTATATCTAGCATATGAGCTTTATGGATCA-3′ and 5′-GTATTGGATCCTTATTAAGATCGGTAGAACTCGCC-3′. The amplified product was digested with *NdeI* and *BamHI* and ligated into the complementary sites of the pET-3a expression vector (Novagen, Madison, WI) to give plasmid pLZL2423 (pET3a-yeast-native-GAG3). In order to optimize protein production, the full-length, 876-bp Ty3 *GAG3* ORF was recoded to reflect *E. coli* codon bias and assembled *in vitro* according to manufacturer's instructions (CODA Genomics Inc., Laguna Hills, CA) (Hatfield GW & Gutman GA, 1992) (Larsen et al., in preparation) (For assembly strategy and intermediate and full-length products, see Supplemental Data Fig. 1; for sequence of oligonucleotides, see Supplemental Data, Table 1.) The yeast CODA Ty3 *GAG3* gene was cloned into pET-3a, under an IPTG responsive promoter, to yield clone pLZL2422 (pET3a-CODA-GAG3). The portion of this *GAG3* encoding CA (residues 1-207 of Gag3) was amplified using the polymerase chain reaction and primers which annealed to the *GAG3* sequence in pET3a-CODA-GAG3 (5′-TACTCCTCGAGATGTCATTCATGGACCAGATTCCGGGCGG-3′ and 5′-GGCATGCAAGCTTTTATTAGCCTATGATGGTATCTCCGTCCGGCCTCC-3′). The amplified fragment was restricted and cloned into the *XhoI* and *HindIII* sites of the araBAD-based promoter pBAD vector (Invitrogen Inc.) to create clone pLZL2424 (pBAD-CODA-CA). Clones were verified by DNA sequence analysis (Cogenics, Morrisville, NC).

**Protein expression**

The pET3a-yeast-native-GAG3 and pET3a-CODA-GAG3 were each transformed into *E. coli* strain BL21(DE3). For optimal protein solubility, a fresh overnight culture was diluted 1:100 into 200 ml LB/amp medium and was grown at 37°C to OD<sub>600</sub>=1.0. The T7 RNA polymerase gene in BL21(DE3) under the control of the lac promoter exhibits leaky expression sufficient to achieve soluble Gag3 by growth at 4°C for 72 h without the use of IPTG. Pelleted cells were resuspended in 20 ml lysis Buffer A (20 mM TrisCl pH 7.5, 20 mM NaCl, 10% glycerol, 5mM DTT, 1mM PMSF, 1mM EDTA) and disrupted by three serial passages through a French pressure cell at 15,000 lb/in<sup>2</sup> at 4°C. These extracts were clarified by centrifugation at 17K × g for 15 min at 4°C. Pellets were resuspended in 20 ml of buffer A. Laemmli sample buffer (LSB) was added to 10 μl of either resuspended pellet or supernatant to a concentration of 1 × and proteins were analyzed by SDS-10% polyacrylamide gel electrophoresis.
The pBAD-CODA-CA was transformed into *E. coli* strain Top10 (Invitrogen, Carlsbad, CA). A single *E. coli* Top10- pBAD-CODA-CA colony was inoculated into 200 ml of LB medium plus 100 μg/ml ampicillin at 30 °C to OD₆₀₀= 0.5. Protein expression was induced with the addition of 0.002% arabinose, and growth was continued for 3 h. Processing of cell pellets was as described above.

**VLP preparation from *E. coli* lysates**

Approximately, 500 μg of protein in clarified whole cell lysates produced as described above from *E. coli* cells expressing Gag3 or CA, was fractionated over a 20/30/70% (542 μl, 253 μl, and 181 μl respectively) sucrose step gradient by centrifugation in a TLS55 rotor in the Optima TL-100 ultracentrifuge (Beckman-Coulter, Fullerton, CA) at 86K × g for 68 min at 4°C. Eight 175 μl fractions were successively removed from the top of the gradient and the interface fraction which contained the majority of Ty3 protein as determined by immunoblotting using anti-Ty3 CA and Coo massie Brilliant Blue R-250 staining (Fig. 3 and data not shown) was used for AFM.

**Atomic force microscopy**

A 5 μl aliquot of the 70/30% interface fraction collected as described above was used for AFM imaging in fluid using tapping mode (Kuznetsov, Zhang, Menees, McPherson, & Sandmeyer, 2005). The atomic force microscope was a Nanoscope IIIa (VEECO, Santa Barbara, CA). VLPs were scanned in tapping mode with an oscillation frequency of 9.2 KHz and a scan frequency of 1 Hz using oxide sharpened silicon nitride tips in a 75 μl fluid cell (Hansma, Browne et al., 1994; Hansma & Hoh, 1994). Scanning was performed at room temperature. Nucleic acid was extruded from particles which were broken in 1.5M KCl at pH 12 with addition of sodium hydroxide.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

This research was supported by Public Health Services grants GM33281 to S.B.S., GM68903 to G. W. H., and GM58868 to AM. We thank M. Zhang for helpful discussions and M. Oakes for assistance with fluorescence microscopy.

**References**


Fig. 1. Ty3 organization
Ty3 ORFs GAG3 and POL3 are shown with protein domains capsid (CA), spacer (SP), nucleocapsid (NC), protease (PR), reverse transcriptase (RT), and integrase (IN). PR and RT domains are separated by a domain of unknown function of approximately 10 kDa. Gag is a 290 aa polyprotein which is processed into 207 aa CA, 26 aa spacer, and 57 aa NC.
Fig. 2. Recoded Gag3 is expressed at a higher level in *E. coli* than native yeast Gag3
DH5α cells were induced to express Gag3 from pET3a-CODA-GAG (lanes 1-3) and pET3a-
yeast-native-GAG3 (lanes 4-6) for 1, 2, 3 h, respectively. Protein extracts of 50 μl of cells
expressing Ty3 Gag3 were fractionated on SDS 10% polyacrylamide gels and stained with
Coomassie Brilliant Blue R-250 as described in Materials and Methods.
Fig. 3. Expression of Gag3 and CA in *E. coli*

DH5α cells were induced to express pET3a-CODA-GAG3 (A) and pBAD-CODA-CA (C) as described in Materials and Methods. Whole cell extracts were centrifuged and pellets were resuspended in a volume equal to the supernatant. Equal aliquots of pellets (P) and supernatant (S) were fractionated on SDS 10% polyacrylamide gels. Gag3 (A) was visualized by staining with Coomassie Brilliant Blue R-250. (B) Clarified whole cell extract prepared from cells expressing pET3a-CODA-GAG3 as described in Materials and Methods was fractionated on a 20%/30%/70% sucrose step gradient and analyzed by SDS 10% gel electrophoresis. Fractions 7 and 8, corresponding to the 30%/70% interface and below, contained Ty3 Gag3 protein. (C) CA, which was expressed at a lower level than Gag3, was visualized by immunoblot analysis. Size markers were included in all gels, and relevant markers are indicated with dots in A and the same markers are indicated with dots in B.
Fig. 4. AFM analysis of particles resulting from expression of Ty3 Gag3 and CA in *E. coli*

Extracts of *E. coli* expressing pET3a-CODA-GAG3 (panels A-C) and pBAD-CODA-CA (panels D-F) were fractionated as described in Materials and Methods and in the Fig. 3B legend, and the 30/70% interface fractions were imaged by AFM. Dimension of the scan is indicated at lower right of each panel. Size determinations were based on height and cannot be as accurately judged from horizontal distances while scanning in tapping mode. Tapping mode extrusion of nucleic acid from Ty3 particles (panels G-H).