Volume 285, number 1, 31–34 FEBS 09864 July 1991

ADONIS 001457939100598C

Autoprocessing of *Drosophila copia gag* precursor to generate a unique laminate structure in *Escherichia coli*

Katsuji Yoshioka[1], Hiroma Kanda[1], Shunzo Kondo[2], Shin Togashi[3], Tadashi Miyake[3] and Tadayoshi Shiba[1]

[1]*Laboratory of Molecular Biology, School of Hygienic Sciences, Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 228,* [2]*Technical Section and* [3]*Laboratory of Cell Biology, Mitsubishi Kasei Institute of Life Sciences, 11 Minamioya, Machida, Tokyo 194, Japan*

Received 18 April 1991; revised version received 27 April 1991

Drosophila copia protease is likely to be encoded in the *gag* gene. We have expressed *copia gag* polyprotein precursor in *E. coli*. The *gag* precursor was correctly processed to generate a unique laminate structure in *E. coli*. The processing was almost completely blocked by a mutation at the putative active site of *copia* protease, and resulted in accumulation of the precursor. Furthermore, the laminate structure was not found in *E. coli* expressing the mutant precursor. These results indicate that the protease is involved in cleaving the *gag* precursor itself. Also, the assembly of *copia gag* protein should correlate to the autoprocessing of *copia gag* polyprotein precursor.

Autoprocessing; *Gag* polyprotein; Protease; Retrotransposon *copia*

1. INTRODUCTION

Drosophila transposable element *copia* is structurally related to retroviral proviruses [1–3]. It is 5 kb in length with long terminal repeats (LTRs) of 276 bp [4–6]. Major transcripts of *copia* are 5 kb and 2 kb in length in *Drosophila* cultured cells [7]. The larger one is a full-length RNA, and the 2 kb RNA generates by splicing of the 5-kb RNA [8–12]. Nucleotide sequence analyses show open reading frames (ORFs) in both the 5 kb [10,13] and 2 kb [11,12] RNAs. The ORF of the 5 kb RNA (termed ORF1) and that of the 2 kb RNA (termed ORF2) consist of 1049 amino acids and 426 amino acids, respectively. Translation products of ORF1 and ORF2 seem to be similar to retroviral *gag-pol* and *gag* polyprotein precursors, respectively [11–13].

The processing of *gag* and *gag-pol* precursors is an essential step in retroviral replication, and is directed by virus-encoded protease [14]. *Drosophila* retrotransposon *copia* seems to position the protease in the *gag* gene as in the case for avian retroviruses [11,12,14]. Our previous study [12] demonstrated that the 2 kb *copia* RNA contains sufficient information to make *copia* virus-like particle (VLP), probably through autoprocessing of *copia gag* precursor, in *Drosophila* cultured cells. To date, however, detailed characterization of *copia gag* precursor has not been done. Here we have expressed the *gag* precursor in *E. coli*, and have found that the precursor autocatalytically processes to generate a unique laminate structure.

Correspondence address: K. Yoshioka, Research Laboratory for Genetic Information, Kyushu University, 3-1-1 Maidashi, Fukuoka 812, Japan. Fax: (81) (92) 632 2375.

2. MATERIALS AND METHODS

2.1. *Materials*

Restriction enzymes, T4 polynucleotide kinase and T4 DNA ligase were purchased from TOYOBO. Oligonucleotides COP136 (5′-TTGAGTGAACCATGGACAAG-3′) and COP56 (5′-TGTC-CTTGCTTCTGGTG-3′), which were used for the *Nco*I site creation and introducing the mutation responsible for the Asp→Ala mutation at the putative active site (Asp-Ser-Gly) of *copia* protease, respectively, were synthesized by an automatic DNA synthesizer (Applied Biosystems, Model 381A). *Nco*I linker (5′-CAGCCATGGCTG-3′) and the expression vector pKK233-2 were obtained from Pharmacia.

2.2. *DNA manipulations*

The general DNA manipulations were carried out as described by Maniatis et al. [15].

2.3. *Bacterial cells*

E. coli JM109 [16] was used as a host for all expression plasmids.

2.4. *Bacterial cell growth for protein expression*

E. coli containing expression plasmid was grown in L-broth at 37°C until mid-log phase ($OD_{600} = 0.5–0.7$), after which induction of gene expression was initiated by addition of isopropyl-β-D-thiogalactoside to a final concentration of 50 μg/ml. The bacteria were further grown at 37°C for an appropriate period.

2.5. *Electron microscopy*

After the wild-type or mutant *gag* gene was induced for 22 h, the bacterial cells were collected by centrifugation at 2000 × *g* for 10 min, fixed with 2.5% glutaraldehyde for 2 h at 4°C, washed with 0.1 M sodium phosphate buffer (pH 7.4), post-fixed in 1% osmium tetroxide for 2 h at 4°C, dehydrated, and embedded in epoxy resin (Epok 812). Specimens were stained with uranyl acetate and lead compounds. The same procedures were adopted also in the case of *E. coli* containing only the expression vector pKK233-2.

2.6. *Other methods*

Site-directed mutagenesis and Western blot analysis were described previously [12]. *Copia* VLP was purified essentially by the methods previously described [17,18].

3. RESULTS AND DISCUSSION

To express entire ORF2 in *E. coli*, we have used plasmid pZY2 *copia* (Fig. 1; see also [12]), in which the 2 kb *copia* RNA's intron has been removed at the DNA level, for the starting material. Two nucleotides (ApA) neighboring with the first methionine codon of ORF2 were converted to CpC using oligonucleotide COP136 in order to create a *Nco*I site (5'-CCATGG-3'). Furthermore, the *Hpa*I site locating downstream of ORF2 was changed using *Nco*I linker. The 1.4 kb *Nco*I fragment, which covers the entire region of ORF2, was inserted into the *Nco*I site of the expression vector pKK233-2. The resultant plasmid was designated pEC1. Fig. 1 shows construction scheme of the expression plasmid pEC1.

Immunological analysis of *E. coli* containing pEC1 (termed EC1) was carried out and the result was shown in Fig. 2. By using anti-VLP serum, two major polypeptides were detected. One is a 48 kDa polypeptide, the size of which is correctly corresponding to the predicted size of the *gag* precursor. The other one is a 33 kDa polypeptide which co-migrates with the major *copia* VLP protein. The apparent bands of the 33 kDa and 48 kDa polypeptides appeared 1 h and 3 h, respectively,

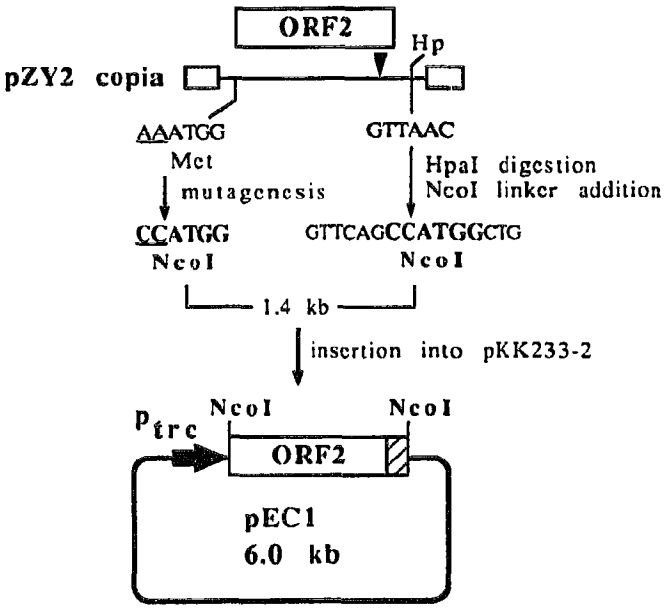

Fig. 1. Construction scheme of the expression plasmid pEC1. Plasmid pZY2 *copia* [12], in which the 2 kb RNA's intron has been removed at the DNA level, was used for the starting material. The dotted box indicates the *copia* LTR. The arrowhead represents the spliced junction. The ORF of the 2 kb *copia* RNA, ORF2, is shown. Two nucleotides (ApA) neighboring with the first methionine codon of ORF2, and the *Hpa*I site were changed to create *Nco*I sites, respectively. The 1.4 kb *Nco*I fragment was inserted into the expression vector pKK233-2, and the resultant plasmid was termed pEC1. The bold horizontal arrow indicates the *trc* promoter (a *trp-lac* fusion promoter with the consensus 17 bp spacing between the *trp* −35 region and the *lac* UV5 −10 region). The dashed box indicates the *copia* sequence other than ORF2.

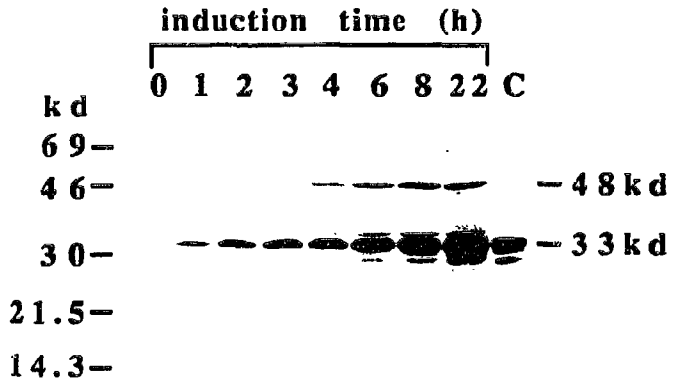

Fig. 2. Western blot analysis of the wild-type *copia gag* precursor expressed in *E. coli* EC1. At the times indicated, cells (2×10^7) were suspended in 10 μl of the lysis buffer, consisting of 50 mM Tris-HCl pH 6.8, 2% SDS, 5% glycerol, 5% 2-mercaptoethanol, 0.01% bromophenol blue. The proteins were separated on a 12.5% SDS-polyacrylamide gel, transferred to nitrocellulose, and subjected to immunoblot analysis using anti-VLP serum. As a control, *copia* VLP prepared from *Drosophila melanogaster* Kc cells was also electrophoresed (lane C). The positions and mol. wts of the *gag* precursor and the major VLP protein are indicated on the right. Mol. wt. markers are given on the left.

after the *gag* expression was induced. Both the polypeptides increased with time. These results suggest that the *gag* precursor expressed in *E. coli* leads to correctly processed *copia* VLP protein, probably through autoprocessing of the precursor.

Next, we constructed plasmid pEC2 which harbors the GAT → GCT mutation responsible for the Asp → Ala mutation at the putative active site (Asp-Ser-Gly) of *copia* protease. The result of immunological

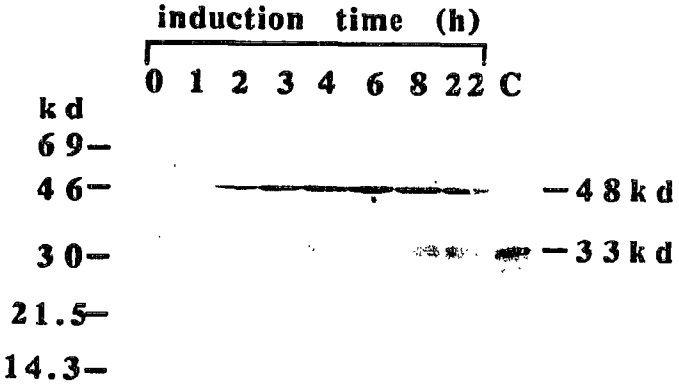

Fig. 3. Western blot analysis of the protease mutant *gag* precursor expressed in *E. coli* EC2. The GAT → GCT substitution corresponding to the Asp → Ala mutation at the putative active site of *copia* protease was made using oligonucleotide COP56. Lane C: *copia* VLP prepared from *Drosophila melanogaster* Kc cells. Equal amount of Kc *copia* VLP was used as controls in both Figs. 2 and 3. Procedures were as described in the legends to Fig. 2. The positions and mol. wts of the *gag* precursor and the major VLP protein are indicated on the right. Mol. wt. markers are given on the left.

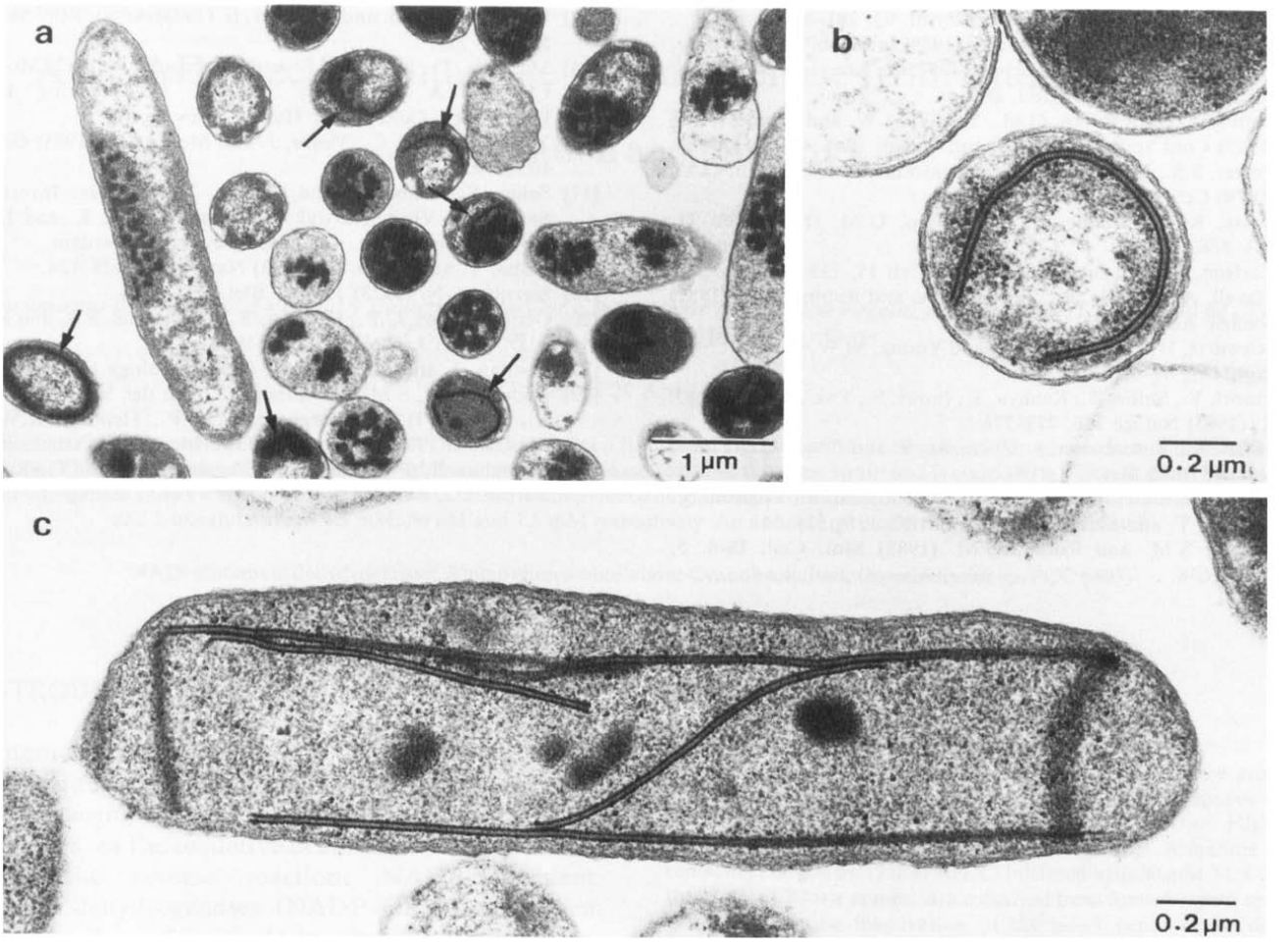

Fig. 4. Electron-micrographs of *E. coli* EC1 expressing the wild-type *gag* precursor. Specimens were prepared for electron microscopy as described in Materials and Methods. (a) The laminate structure is indicated by the arrow. Scale bar = 1 μm. (b) A cross-section. Scale bar = 0.2 μm. (c) A vertical section. Scale bar = 0.2 μm.

analysis of *E. coli* containing pEC2 (termed EC2) was shown in Fig. 3. When the mutation was introduced, the processing was almost completely blocked, and resulted in accumulation of the mutant precursor. Taking into account the observation of Fig. 2, this result indicates that the wild-type *gag* precursor expressed in *E. coli* autocatalytically processes to produce the major *copia* VLP protein, and the Asp → Ala mutation drastically reduces the efficiency of the autoprocessing.

Furthermore, we analyzed *E. coli* EC1 and EC2 using electron microscopy. A laminate structure was found in EC1 (Fig. 4). The structure is quite different from *copia* VLP produced in *Drosophila* cultured cells. However, since we could not detect the laminate structure in EC2 nor *E. coli* containing only the expression vector pKK233-2 (data not shown), the structure is specific for EC1. These results strongly suggest that *copia gag* protein assembles also in *E. coli*, and the assembly should correlate to the autoprocessing of *copia gag* precursor. An open question is why the laminate structure but not *copia* VLP is produced in EC1. In some cases, such as the head protein of bacteriophage lambda [19-21] and ribulose biphosphate carboxylase [22,23], the assemblies of proteins require molecular chaperons. If *copia* VLP formation needs a molecular chaperon, the structural difference in between *Drosophila* and *E. coli* may reflect the difference of the molecular chaperon. Further study will clarify the mechanism of *copia* VLP formation.

Finally, the expression system of *copia gag* polyprotein in *E. coli* should be the aid for identification and more, biochemical study of *copia* protease. Experiments of this nature are presently under way.

Acknowledgements: We thank Dr. Y. Sakaki for critically reviewing the manuscript, and Dr. N. Takamatsu for helpful discussion. This work was supported in part by a grant from the Kitasato Research Foundation to K.Y., and was also supported in part by a grant from the Ministry of Education, Sciences and Culture, Japan to T.M. and T.S.

REFERENCES

[1] Rubin, G.M. (1983) in: Mobile Genetic Elements (Shapiro, J.A., ed.) pp. 329-361, Academic Press, New York.

[2] Finnegan, D.J. (1985) Int. Rev. Cytol. 93, 281-326.
[3] Bingham, P.M. and Zachar, Z. (1989) in: Mobile DNA (Berg, D.E. and Howe, M.M. eds.) pp. 485-502, American Society for Microbiology, Washington, DC.
[4] Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S (1973) Cold Spring Harbor Symp. Quant. Biol. 42, 1053-1063.
[5] Potter, S.S., Brorein, W.J.,Jr., Dunsmuir, P. and Rubin, G.M. (1979) Cell 17, 415-427.
[6] Levis, R., Dunsmuir, P. and Rubin, G.M. (1980) Cell 21, 581-588.
[7] Carlson, M. and Brutlag, D. (1978) Cell 15, 733-742.
[8] Flavell, A.J., Levis, R., Simon, M.A. and Rubin, G.M. (1981) Nucleic Acids Res. 9, 6279-6291.
[9] Schwartz, H.E., Lockett, T.J. and Young, M.W. (1982) J. Mol. Biol. 157, 49-68.
[10] Emori, Y., Shiba, T., Kanaya, S., Inoue, S., Yuki, S. and Saigo, K. (1985) Nature 316, 773-776.
[11] Miller, K., Rosenbaum, J., Zbrzezna, V. and Pogo, A.O. (1989) Nucleic Acids Res. 17, 2134.
[12] Yoshioka, K., Honma, H., Zushi, M., Kondo, S., Togashi, S., Miyake, T. and Shiba, T. (1990) EMBO J. 9, 535-541.
[13] Mount, S.M. and Rubin, G.M. (1985) Mol. Cell. Biol. 5, 1630-1638.
[14] Kraüsslich, H.-G. and Wimmer, E. (1988) Annu. Rev. Biochem. 57, 701-754.
[15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
[16] Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene 33, 103-119.
[17] Saigo, K., Shiba, T. and Miyake, T. (1980) in: Invertebrate Systems In Vitro (Kurstak, E., Maramorosch, K. and Dübendorfer, A. eds.), pp. 411-424, Elsevier, Amsterdam.
[18] Shiba, T. and Saigo, K. (1983) Nature 302, 119-124.
[19] Sternberg, N. (1973) J. Mol. Biol. 76, 25-44.
[20] Georgopoulos, C.P., Hendrix, R.W., Casjens, S.R. and Kaiser, A.D. (1973) J. Mol. Biol. 76, 45-60.
[21] Kochan, J. and Murialdo, H. (1983) Virology 131, 100-115.
[22] Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W. and Ellis, R.J. (1988) Nature 333, 330-334.
[23] Goloubinoff, P., Gatenby, A.A. and Lorimer, G.H. (1989) Nature 337, 44-47.