Volume 9 Number 23 1981 Nucleic Acids Research

The 5′ termini of RNAs encoded by the transposable element copia

Andrew J.Flavell[1], Robert Levis[2], Meredith A.Simon[3] and Gerald M.Rubin[2]

Sidney Farber Cancer Center, 44 Binney Street, Boston, MA 02115, USA

Received 16 October 1981

ABSTRACT

The 5' termini of copia-specific RNAs in *Drosophila melanogaster* tissue culture cells were determined by S1 nuclease mapping and cap analysis. Both major *copia* RNAs share an identical set of heterogeneous 5' ends. Three major cap 1 structures M^7GpppCmpUp, M^7GpppCmpCp and M^7GpppGmpUp together with several other minor caps were found. Almost all the 5' termini, as judged by S-1 nuclease mapping, were located either in a pyrimidine-rich part of the terminal direct repeat or apparently outside of the *copia* element, suggesting that a proportion of *copia* transcripts derive from promoters external to the genetic element.

INTRODUCTION

Copia is a DNA sequence of approximately 5000 base pairs (5kb) present in about 50 closely conserved copies in the *Drosophila melanogaster* haploid genome. Although originally isolated as a class of moderately repetitious DNA encoding abundantly expressed RNA (1), *copia* and several analogous sequence families have generated wide interest owing to their ability to transpose to a large number of chromosomal sites in the *Drosophila* genome (2-4). Many additional moderately repeated sequence families have been described in *Drosophila* which share at least some of the properties of the *copia* sequence family (4-8) and together these sequences probably constitute about 5% of the *Drosophila* genome.

The general structural features of *copia*, including the presence of terminal direct repeats, are shared by the yeast transposable element *Ty1* (9-10) by certain prokaryotic transposons (11) and, surprisingly, by the integrated proviruses of some retroviruses (12,13). It is therefore likely that this class of genetic element is ubiquitous in living systems.

The effects of transposition in eukaryotes are varied. Eukaryotic transposable elements are associated with a wide variety of genetic effects such as unstable mutations, deletion formation and other chromosomal

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

rearrangements (14-20). Another way in which movable genetic elements can disturb cellular processes is by acting as a mobile promoter sequence. In tumours induced by avian leukosis virus a defective portion of the virus (always containing the long terminal repeat which possesses a promoter) is usually found integrated upstream of the *c-myc* cellular "oncogene" (21).

The structure of *copia* RNAs has been studied by Northern blot analysis (22,23). Two major *copia*-specific RNAs, approximately 5kb and 2kb long together constitute about 3% of the polyadenylated RNA in *D.melanogaster* cultured cells. Additionally, other species of approximately 1.3kb and 0.8kb (23) together with small amounts of RNA larger than the complete *copia* element (22) have been reported. The direct repeat sequences of *copia* (23; G. Rubin, unpublished) and the yeast transposable element *Ty1* (24) are at least in part transcribed and since the 5kb *copia* RNA is approximately the same size as the element it seemed likely that transcriptional initiation and termination occurred in these sequences. The long terminal repeat sequences of retrovirus proviruses contain the major promoters of these elements and it was therefore of interest to see whether the *copia* long terminal repeats were similar in this respect.

We have therefore studied the 5' ends of *copia* RNAs. We report here that the major 5' ends of both the 5kb and 2kb *copia* RNAs lie in the 276 nucleotide long terminal repeat sequence in a pyrimidine-rich region (with respect to the RNA coding strand of the DNA). We also show that both the 5kb and 2kb RNAs share a set of heterogeneous capped termini predominantly composed of pyridimine nucleotides suggesting that transcriptional initiation in this pyrimidine-rich region is staggered. Lastly we present DNA sequence and S1 nuclease mapping data which together suggest that a portion of *copia* transcripts are defective as mRNAs.

MATERIALS AND METHODS

Isolation of ^{32}P *copia* RNA and cap analysis. A subline of Eschaliers K_c. *Drosophila melanogaster* cells adapted for growth with no added serum (25, 6×10^8) was incubated with 10mCi (α-^{32}P) orthophosphate (Radiochemical Centre) for 6 hours and the polyadenylated RNA extracted by the method of Darnell (26). *Copia* RNA was isolated from cell polyadenylated RNA by hybridization to *copia* DNA cellulose (26). RNAase digestion and cap analysis was as described by Barrell (27) and Cory and Adams (28) respectively, with modifications (26).

Mapping of the 5' ends of *copia* RNA by S1 nuclease protection. Preparation of the DNA probe - the 3.5kb *PvuII* fragment of cDM 5002 (29) was purified and kinase end-labelled with ^{32}P to a specific activity of 7×10^5 cpm/µg as described by Levis *et al.* (30). The 5' end-labelled fragment was digested with *E.coli* exonuclease III to render it single stranded in the region complementary to *copia* RNA. The average size of the single stranded DNA after this treatment was 2.3kb. 1.7µg of the *PvuII* fragment was dissolved in 100µl 66mM Tris-Cl pH 8.0, 0.66mM MgCl2, 1mM DTT to which was added 1 U exonuclease III (New England Biolabs). After incubating for 30' at 20°C, EDTA was added to 12mM, the mixture was extracted once with an equal volume of phenol, 3x with ether, and nucleic acids recovered by ethanol precipitation.

Preparation of *Drosophila* cell cytoplasmic RNA. Total cytoplasmic RNA was prepared from Kc cells as described by Levis and Penman (31) with the modification that 5mM $MgCl_2$ was substituted for $CaCl_2$ in the cell lysis buffer.

S1 nuclease protection assay. The *Drosophila* cytoplasmic RNA (20µg) was combined with 15µg *E.coli* ribosomal RNA carrier and 20-40ng of the radioactive *PvuII* probe. The nucleic acids were precipitated with ethanol, dried under vacuum, then dissolved in 10µl of hybridization buffer (40mM PIPES pH 6.4, 400mM NaCl, 1mM EDTA, 80% (v/v) formamide). The solution was heated to 70°C for 10 min, then incubated at 42°C for 8-11 h. 0.3ml of ice-cold digestion buffer (30mM Na acetate pH 4.5, 300mM NaCl, 3mM $ZnCl_2$, 10µg/ml denatured salmon sperm DNA).S1 nuclease (Sigma Type III) was next added and digestion allowed to proceed for 30 min. In preliminary experiments it was found that digestion with 50 units S1 nuclease at 10°, 24° or 37°C gave equivalent results as did digestion at 37°C with 10-50 U S1. Routinely digestions were carried out at 37°C with 20 or 50 U S1 nuclease. The protected DNA fragments were recovered and dissolved in 15µl 80% formamide, 10mM NaOH, 1mM EDTA and heated to 95° for 1 min before electrophoresis on 3.5% acrylamide-urea slab gels (32). Gels were soaked in water to remove urea and then dried before autoradiography. All of the experiments described were performed under conditions of large DNA excess as judged by the proportionality of nuclease resistant hybrid to RNA input.

DNA sequence determination. DNA sequence determination was by the method of Maxam and Gilbert (32) with modifications (30)

RESULTS

Isolation of ^{32}P *copia* RNA and 5' terminal cap analysis

Copia-specific RNA was purified from ^{32}P-labelled tissue culture cells by hybridization of cytoplasmic RNA to a restriction fragment of the plasmid cDm 2056 (29) containing the entire *copia* element covalently bound to cellulose (26). The purity of the *copia* RNA was assayed by electrophoresis on an agarose gel containing 5M urea. *copia*-selected polyadenylated RNA was mainly composed of two size classes of RNA approximately 5kb and 2kb long together with small amounts of larger RNA up to 9kb long (Figure 1). These results were in agreement with data obtained from Northern blots by other workers (22,23). *Copia*-specific non-polyadenylated RNA, which was present in four-fold smaller amounts contained an additional minor species approximately 1.3kb long which presumably corresponds to the 1.3kb RNA described by Young and Schwartz (23). We did not detect the 0.8kb *copia*-specific RNA reported by these authors.

Copia-specific cytoplasmic polyadenylated RNA was analysed for the presence of 5' terminal cap structures by complete digestion with a mixture

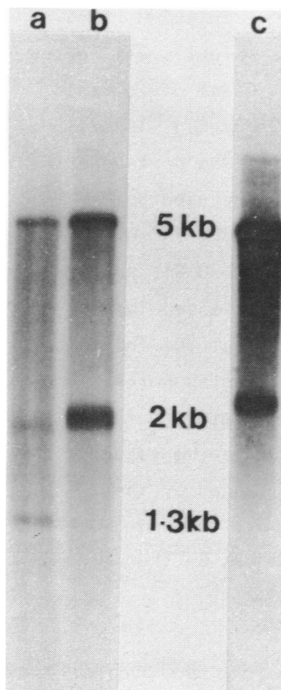

Figure 1. Agarose-urea electrophoresis of *in vivo*-labelled *copia* RNA. ^{32}P *copia*-specific non polyadenylated (lane a) and polyadenylated RNAs (lane b) were electrophoresed on a 0.8% agarose gel containing 5M urea. The sizes of the *copia* RNAs are deduced from parallel marker ribosomal RNAs. Lane c shows polyadenylated *copia* RNAs larger than 5kb which are visible upon prolonged autoradiography.

of RNAases which hydrolyse every RNA phosphodiester bond unprotected by 2'-O-methylation. RNAase-resistant cap structures ($M^7G^{5'}ppp^{5'}N^1mpNp^2$) were then resolved by two dimensional electrophoresis as described by Cory and Adams (28). The cap structure fingerprint of cytoplasmic polyadenylated *copia*-specific RNA is compared with that of total tissue culture RNA in Figure 2. Three major and four minor cap structures were observed in *copia*-specific RNAs whereas a completely different proportion of radioactivity was present in total *Drosophila* cap structures. The radioactivity in all *copia*-specific cap structures corresponded to one cap per 3.4kb, in good agreement with the expected value for an equimolar mixture of fully capped 5kb and 2kb RNAs. Uncapped 5' termini were not detected in *copia* RNA. The compositions of the numbered oligonucleotide spots from both *copia*-specific RNA and unselected *D.melanogaster* RNA in Figure 2 were analysed by further

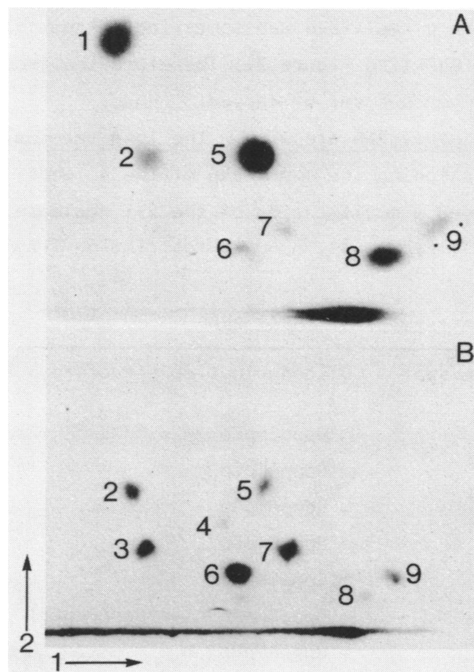

Figure 2. Comparison between *copia*-specific and total *Drosophila* cap structures. Fig. 2A; *copia*-specific cap structures. Fig. 2B; total Kc cultured cell capped structures. Digests were fractionated by electrophoresis on cellulose acetate at pH 3.5 (1) followed by electrophoresis on DEAE paper at pH 3.5 (2). Oligonucleotide spots 1-9 in both fingerprints (where present) were eluted and characterized (see Table 1).

enzymatic and separation techniques (26). The results of such analyses for all numbered oligonucleotide spots in Figure 2 are shown in Table 1. The assignments reached in Table 1 are in complete agreement with the data of Levis and Penman for total *Drosophila* capped RNAs (31). We conclude that *copia* RNAs possess a heterogeneous mixture of 5' terminal cap structures with M^7GpppCmpUp and M^7GpppCmpCp predominating.

The multiple cap structures are the 5' termini of both the 5kb and 2kb major *copia* RNAs.

To analyse the capping of the two major *copia* RNAs ^{32}P *copia* RNA was sedimented through a sucrose gradient containing 50% formamide (Figure 3). The 5kb and 2kb *copia* RNAs were well resolved from each other. Fractions containing each size class were pooled and assayed for cross contamination by electrophoresis on an agarose gel containing 5M urea (Figure 3). No cross contamination was observed between the two species. The pooled RNAs were then assayed for cap content as before. The 5kb and 2kb RNA preparations each yielded a pattern of multiple cap spots indistinguishable from unfractionated *copia* RNA (see Figure 2). Therefore the two major *copia* RNAs share a common heterogeneous set of capped 5' ends.

The major 5' ends of *copia* RNA are within the long terminal repeat sequence

To study the location on the *copia* DNA of the 5' ends of the major *copia* RNAs, we employed a modification of the S-1 nuclease mapping technique of Berk and Sharp (33; Figure 4). A *PvuII* restriction fragment of *copia*

Table 1. Capped 5' termini of *copia* RNA

Oligonucleotide spot	Oligonucleotide structure	Relative Abundance (percent)
1	m^7GpppCmpCp	19
2	m^7GpppAmpCp	7
3	m^7GpppAmpAp	-
4	m^7GpppGmpCp	-
5	m^7GpppCmpUp	38
6	m^7GpppGmpAp (+m^7GpppAmpGp)	4
7	m^7GpppAmpUp	4
8	m^7GpppGmpUp	14
9	m^7GpppUmpUp	4

Oligonucleotide spots 1-9 in Figure 2A and 2B were eluted from cap fingerprints and digested with P 1 nuclease. The digestion products were subjected to electrophoresis on DEAE paper at pH 3.5 in parallel with marker cap cores and nucleoside 5' monophosphates. The structure deduced from the digestion products of each spot is shown.

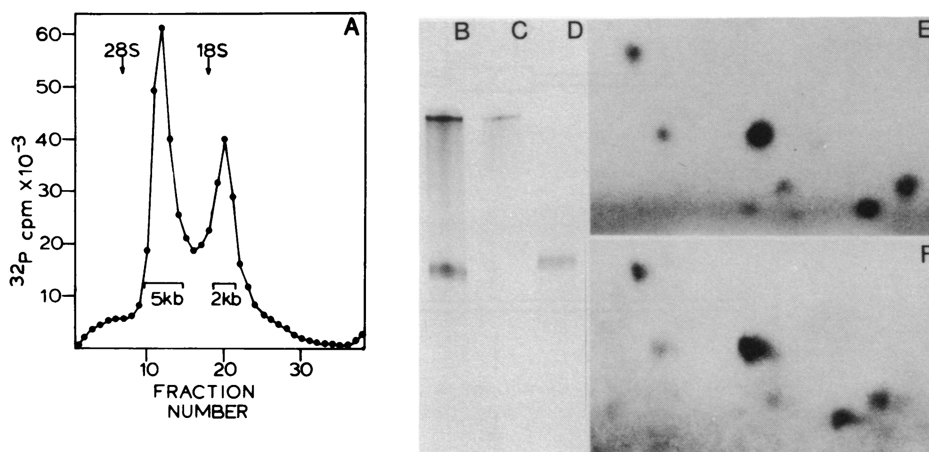

Figure 3. The two major *copia*-specific RNAs share a common set of capped 5' termini.
Figure 3A: *In vivo* labelled *copia* RNA was sedimented through a sucrose gradient containing 50% formamide. Indicated fractions were pooled and assayed for cross contamination by electrophoresis on an agarose gel containing 5M urea.
Figure 3B: Agarose-urea electrophoresis of unsedimented *copia* RNA.
Figure 3C: Electrophoresis of gradient-purified *copia* 5kb RNA.
Figure 3D: Electrophoresis of gradient-purified *copia* 2kb RNA.
Figure 3E and F: Cap analysis of gradient-purified 5kb and 2kb *copia* RNA respectively.

containing approximately the terminal 820 nucleotides of the element together with 2.7kb of flanking genomic DNA was labelled at it's 5' ends with ^{32}P. The fragment was digested with exonuclease III to destroy the DNA strand with the same sense as *copia* RNA and this treated fragment was annealed with total *D.melanogaster* polyadenylated cytoplasmic RNA. Unhybridized nucleic acids were destroyed with S1 nuclease and the radioactive products were fractionated on a urea-polyacrylamide gel (Figure 5). One major band was observed together with one minor and several very faint bands.

The major S-1-nuclease-resistant species was approximately 680 nucleotides in length. To determine more exactly the location of the 5' end corresponding to this fragment, the following method was used. The labelled *PvuII* fragment was digested with *BalI* or *HaeIII* endonucleases (which cleave *copia* between the end of the element and the *PvuII* site) to yield fragments with ends spanning the 5' terminus of the 680 nucleotide protected

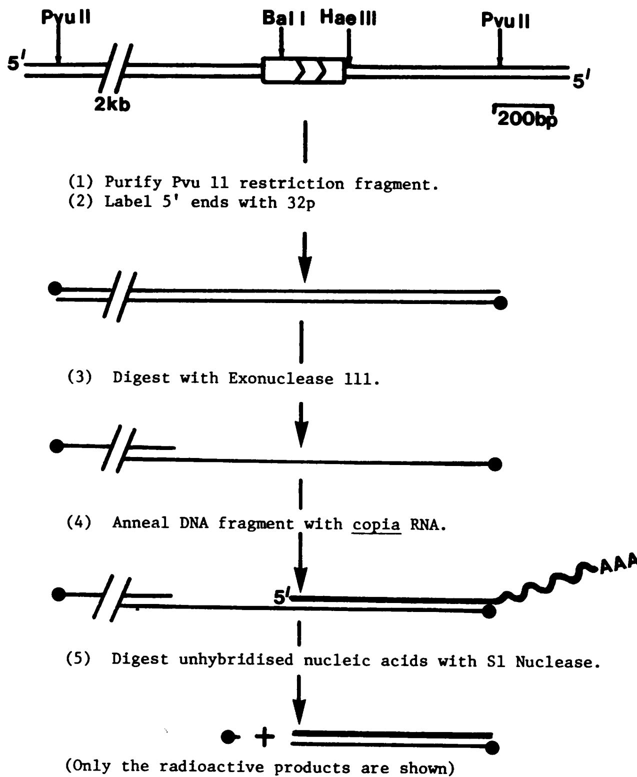

Figure 4. S1-nuclease mapping protocol. The ^{32}P labelled products from the above procedure were sized on a polyacrylamide-urea gel. The location of the long terminal repeat on the *copia* element in cDm 5002 is shown ( ). The *copia* element extends to the right of the long terminal repeat.

DNA fragment. These restriction fragments were run in parallel with the S-1-nuclease-resistant hybrids (Figure 5). The exact size of these fragments were determined from the DNA sequence of this region (Figure 6). In this way the end of the 680 nucleotide fragment was placed at nucleotide 137 of the *copia* DNA with an estimated accuracy of ±10 nucleotides. While it is possible that this end marks the position of a splice acceptor site we think it more likely that it represents the location of the heterogeneous major capped 5' ends of *copia*. The concurrence between the DNA sequence of this region (CTTTCCTTCTTGTACGTTTTT) and the major capped dinucleotides in *copia* RNA (CU, CC and GU) supports this

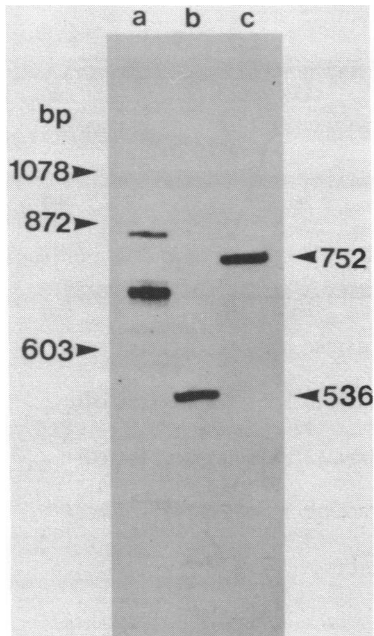

Figure 5. Sizing the S1 nuclease resistant hybrids between *copia* RNA and the 3.5kb *PvuII* fragment of cDM 5002. The products from S1 nuclease mapping of *copia* using the method described in Figure 4 were electrophoresed on a 3.5% poly-acrylamide-urea gel (lane a), together with DNA size markers derived from *HaeIII* digested ØX174 DNA (marked on the left of the Figure), the 536bp *PvuII/HaeIII* (lane b) and the 752bp *PvuII/BalI* (lane c) DNA fragments from cDM 2056 (see Figures 4 and 6).

proposition.

The size of the other prominent band in *copia* S-1-nuclease-resistant hybrids was estimated in a similar manner as 820±10 nucleotides. This places the end of this hybrid at, or very near to, the extreme end of the element. We suggest that this fragment corresponds to RNAs initiated outside of *copia* which read through into the element. As there are approximately 150 copies of *copia* at different sites in the Kc cell genome it is very unlikely that the flanking DNA of cDm 5002 could protect more than a very small proportion of such RNAs from nuclease digestion. In support of this suggestion we have evidence that the 820 nucleotide band is enriched, relative to the 680 nucleotide band, in RNAs larger than 5kb (data not shown).

A proportion of *copia* RNA does not appear to be functional messenger RNA

Two lines of evidence suggest that only a small proportion of *copia* RNAs in cultured cells can act as messenger RNA. First the majority of steady state cytoplasmic *copia* RNA is not found associated with polysomes, suggesting that it is not being efficiently translated (23,34). Second, *copia*-specific RNA, isolated from Kc cells, is only translated *in vitro* with low efficiency (for 2kb RNA) or apparently not at all (for 5kb, RNA; 34,35). The DNA sequence of cDm 2056 corresponding to the first 1000 nucleotides of *copia* RNA (Figure 6) and the S-1 nuclease mapping data of

```
TGTTGGAATA TACTATTCAA CCTACAAAAA TAACGTTAAA CAACACTACT TTATATTTGA TATGAATGGC
1                   20                  40                  60            B

CACACCTTTT ATGCCATAAA ACATATTGTA AGAGAATACC ACTCTTTTTA TTCCTTCTTT CCTTCTTGTA
         80                  100                 120                  140

CGTTTTTTGC TGTGAGTAGG TCGTGGTGCT GGTGTTGCAG TTGAAATAAC TTAAAATATA AATCATAAAA
                    160                  180                  200

CTCAAACATA AACTTGACTA TTTATTTATT TATTAAGAAA GGAAATATAA ATTATAAATT ACAACAGGTT
         220                  240                  260                  280
    H
ATGGGCCCAG TGCATGCCTA ATAAACAATT AAATTGTGAA TTAAAGATTG TGAAAATAAA TTGTGAAATA
                    300                  320                  340

GCATTTTTTC ACATTCTTGT GAAATAGCTT TTTTTTTCAC ATTCTTGTGA AATTATTTCC TTCTCAGAAT
         360                  380                  400                  420

TTGAGTGAAA AATGGACAAG GCTAAACGTA ATATTAAGCC GTTTGATGGC GAGAAGTACG CGATTTGGAA
                    440                  460                  480

ATTTAGAATT AGGGCTCTTT TAGCCGAGCA AGATGTGCTT AAAGTAGTTG ATGGTTTAAT GCCTAACGAG
         500                  520                  540                  560

GTAGATGACT CCTGAAAAAG GCAGAGCGTT GTGCAAAAAG TACAATAATA GAGTACCTAA GCGACTCGTT
                    580                  600                  620

TTTAAATTTC GCAACAAGCG ACATTACGGC GCGTCAGATT CTTGAGAATT TGGACGCCGT TTATGAACGA
         640                  660                  680                  700

AAAAGTTTGG CGTCGCAACT GGCGCTGCGA AAACGTTTGC TTTCTCTGAA GCTATCGAGT GAGATGTCAC
                    720                  740            P     760

TATTAAGCCA TTTTCATATT TTTGACGAAC TTATAAGTGA ATTGTTGGCA GCTGGTGCAA AAATAGAAGA
         780                  800                  820                  840

GATGGATAAA ATTTCTCATC TACTGATCAC ATTGCCTTCG TGTTACGATG GAATTATTAC AGCGATAGAG
                    860                  880                  900

ACATTATCTG AAGAAAATTT GACATTGGCG TTTGTGAAAA ATAGATTGCT GGATCAAGAA ATTAAAATTA
         920                  940                  960                  980

AAAATGACCA CAACGATACA AGCAAGAAAG TTATGAACGC GATCGTGCAC AACAATAATA ACACTTATAA
                    1000                 1020                 1040

AAATAATTTG TTTAAAAATC GGGTAACTAA ACCAAAGAAA ATATTCAAGG GAAATTCAAA GTATAAAGTC
         1060                 1080                 1100                 1120

AAGTGTCACC ACTGTGGCGG CC
                    1140
```

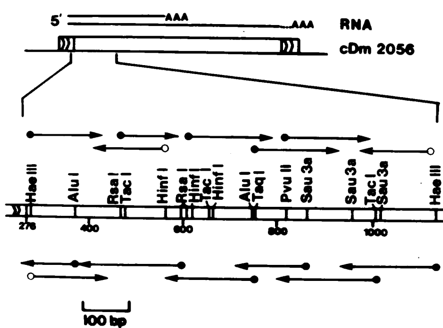

Figure 6. Sequence of a part of the *copia* element in cDm 2056 encoding the first 1000bp of the major *copia* RNAs. The DNA strand corresponding to the *copia* RNAs is shown. The long terminal repeat lies between nucleotide 1-276 (30). The positions of the *Bal1* (B), *Hae111* (H), and *Pvu11* (P) restriction sites are marked. The major 5' ends of *copia* RNAs lie at nucleotides 137±10 (overlined) or 1±10. An open reading frame extends from nucleotide 569 past nucleotide 1142. Stop codons in all three reading frames betweennucleotides 562 and 575 are underlined.The sequencing strategy for nucleotides 276-1142 is shown.● 5'end-labeled.O 3'labeled.

this RNA further supports this hypothesis. There is no continuous open reading frame beginning with an ATG codon and extending into the interior of *copia* before nucleotide 764 on cDM 2056 despite evidence that sequences essential for the translation of a 51,000 dalton *copia*-specific protein lie to the 5' side of nucleotide 567 on this DNA strand (35). Therefore the transcripts from the cDm 2056 *copia* are either defective, or the RNA from this element is spliced in this region. However, the S-1 mapping using cDM 5002 shows that there is no discontinuity in the large majority of *copia* RNAs between the long terminal repeat and the *PvuII* site. It is unlikely that cDm 2056 and cDm 5002 have differential splicing patterns since these elements have apparently identical restriction maps in this region (data not shown). We therefore suggest that either cDm 2056 encodes defective mRNA or only a small minority of *copia* transcripts are spliced to produce functional mRNA.

DISCUSSION

The observation that transcription of *copia* initiates primarily in the long terminal repeat sequences further emphasises the close similarity between this element and retrovirus proviruses. Furthermore, since *copia* is flanked by a pair of such sequences it is possible that both contain promoters. In this case *copia* may promote the expression of flanking sequences downstream from it in the *Drosophila* genome, thus functioning as a mobile promoter.

The heterogeneity of the 5' ends of *copia* RNAs may derive from a corresponding heterogeneity in the approximately 150 elements present in each *D.melanogaster* tissue culture cell. However, we consider this possibility to be unlikely for two reasons. First, sequence analysis of the two complete pairs of long terminal repeats in cDm 2056 and cDm 5002 yielded only two base pair substitutions between the pairs, neither of which was located near to the major 5' ends (30). Therefore there is probably little sequence divergence between *copia* elements in the long terminal repeats. Second, heterogeneity in the 5' ends of RNAs encoded by homogeneous isolates of DNA tumour viruses is well established (26, 36-38). We therefore believe that the 5' ends of the major *copia* RNAs are staggered on the DNA of the long terminal repeat in a manner similar to that observed for Polyoma (26,36), SV40 (37) and Adenovirus (38). The major capped 5' termini in *copia*-specific RNAs contain the dinucleotides CC and CU. These presumably correspond to several such dinucleotides between nucleotides 127

and 147 in the direct repeat sequence (Figure 6) where the major 5' ends of the RNAs are located. To our knowledge this sequence is unlike any other encoding RNA 5' ends studied to date, in that the RNA coding strand of the DNA contains pyrimidines almost exclusively (19 out of 21). Furthermore, pyrimidine-terminal caps consitute greater than 60% of the capped RNA species whereas virtually all RNAs studied to date initiate with purine nucleotides (although total unselected RNA does contain pyrimidine caps (28,31). The reasons for such singular behaviour in the case of *copia* remain to be found.

The majority of transcripts not initiated in the above region appear to derive from promoters located upstream of some or all of the *copia* elements. In such cases it appears that *copia* has transposed, without killing the host cell, into a site in the *Drosophila* chromosome which is being actively transcribed.

We could detect no difference between the 5' ends of the 2kb and 5kb major *copia* RNAs, either as judged by their cap pattern (Figure 3) or by S-1 nuclease protection experiments (data not shown). However, the 2kb RNA contains messenger activity while the 5kb species apparently does not (35). We favour the hypothesis that this apparently paradoxical behaviour is due to the majority of both RNAs being non functional messengers. In such a case, the minor functional 2kb species may not be discernible against the background of inactive RNA species. Such a hypothesis is supported by the known low messenger efficiency of *copia* RNAs (34,35). Furthermore, the DNA sequence analysis presented above shows that cDm 2056 cannot encode an active mRNA between nucleotides 1 and 821 without RNA splicing, yet S-1 nuclease mapping of *copia* RNA detected no major splices.

ACKNOWLEDGEMENTS

We thank David Ish-Horowicz and Julian Burke for critical reading of this manuscript. Footnotes: [1]present address, ICRF Laboratories, Burtonhole Lane, London, NW7 1AD; [2]present address, Carnegie Institution, 115 W University Parkway, Baltimore, Md 21210, USA; [3]present address, Tufts Veterinary School, Boston, Ma 02111, USA.

REFERENCES

1. Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S. (1978 Cold Spring Harbor Symp. Quant. Biol. 42, 1053-1063
2. Potter, S.S., Brorein, W.J., Dunsmuir, P. and Rubin, G.M. (1979) Cell 17, 415-427
3. Strobel, E., Dunsmuir, P. and Rubin, G.M. (1979) Cell 17, 429-439

4. Ilyin, Y.V., Tchurikov, N.A., Ananiev, E.V., Ryskow, A.P., Yenikolopov, G.N., Limborska, S.A., Maleeva, N.E., Gvozdev, V.A. and Georgiev, G.P. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 959-970
5. Ilyin, Y.V., Chmeliauskaite, V.G., Ananiev, E.V., Kulguskin, V.V., Bayev, Jr, A.A., Lyubomirskaya, N.V. and Georgiev, G.P. (1980) Nucleic Acids Res. 8, 5333-5346
6. Ilyin, Y.V., Chmeliauskaite, V.G., Ananiev, E.V. and Georgiev, G.P. (1980) Chromosoma, 81, 27-53.
7. Rubin, G.M., Brorein, W.J., Dunsmuir, P., Flavell, A.J., Levis, R., Strobel, E., Toole, J.J. and Young, E. (1980) Cold Spring Harbor Symp. Quant. Biol. 45, 619-628
8. Spradling, A.C. and Rubin, G.M.R. Ann. Rev. Genetics (1981) 15, in press.
9. Farabaugh, P.J. and Fink, G.R. (1980) Nature 286, 352-356
10. Gafner, J. and Philippsen, P. (1980) Nature, 286, 414-418
11. Kleckner, N. (1977) Cell 11, 11-23
12. Shimotohno, K., Mizutani, S. and Temin, H.M. (1980) Nature 285, 550-554
13. Dhar, R., McClements, W.L., Enquist, L.W. and Vande Woude, G.F. (1980) Proc. Natl. Acad. Sci. USA 77, 3937-3941
14. Bingham, P.M., Levis, R. and Rubin, G.M. (1981) Genetics, in press
15. Rasmuson, B., Westerberg, B.M., Rasmuson, A., Gvozdev, V.A., Belyaeva, E.S. and Ilyin, Y.V. (1980) Cold Spring Harbor Symp. Quant. Biol. 45, 545-551
16. Fink, G., Farabaugh, P., Roeder, G. and Chaleff, D. (1980) Ibid 45, 575-580
17. Greer, H., Igo, M. and de Bruijn, F. (1980) Ibid 45, 567-574
18. Gehring, W.J. and Paro, R. (1980) Cell 19, 897-904
19. Fincham, J.R.S. and Sastry, G.R.K. (1974) Ann. Rev. Genetics 8, 15-50
20. Green, M.M. (1976) The Genetics and Biology of *Drosophila* (Ashburner and Novitski, Eds), London Academic Press pp929-946
21. Hayward, W.S., Neel, B.G. and Astrin, S.M. (1981) Nature 290, 475-480
22. Carlson, M. and Brutlag, D. (1978) Cell 15, 733-742
23. Young, M.W. and Schwartz, H.E. (1980) Cold Spring Harbor Symp. Quant. Biol. 45, 629-640
24. Elder, R.T., St. John, T.P., Stinchcomb, D.T. and Davis, R.W. (1980) Ibid 581-591
25. Rubin, G.M. and Hogness, D.S. (1975) Cell 6, 207-213
26. Flavell, A.J., Cowie, A., Legon, S. and Kamen, R. (1979) Cell 16, 357-371
27. Barrell, B. (1974) Proc. Nucleic Acid Res. 2, 751-779
28. Cory, S. and Adams, J.M. (1975) J. Mol. Biol. 99, 519-547
29. Dunsmuir, P., Brorein, W.J., Simon, M.A. and Rubin, G.M. (1980) Cell 21, 575-579
30. Levis, R., Dunsmuir, P. and Rubin, G.M. (1980) Cell 21, 581-588
31. Levis, R. and Penman, S. (1978) J. Mol. Biol. 120, 487-515
32. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology 65, 499-560 Academic Press.
33. Berk, A.J. and Sharp, P.A. (1978) Proc. Natl. Acad. Sci. 75, 1274-1278
34. Falkenthal, S. and Lengyel, J.A. (1980) Biochemistry 19, 5842-5840
35. Flavell, A.J., Ruby, S.W., Toole, J.J., Roberts, B.E. and Rubin, G.M. (1980) Proc. Natl. Acad. Sci. USA 77, 7107-7111
36. Flavell, A.J., Cowie, A., Arrand, J.R. and Kamen, R. (1980) J. Virol. 33, 902-908
37. Ghosh, P.V., Reddy, V.B., Swinscoe, J., Lebowitz, P. and Weissman, S.M. (1978) J. Mol. Biol. 126, 813-846
38. Baker, C.C., Herisse, J., Courtois, G., Galibert, F. and Ziff, E. (1979) Cell 18, 569-580.