Proc. Natl. Acad. Sci. USA
Vol. 77, No. 8, pp. 4683–4687, August 1980
Biochemistry

Electron microscope study of duck globin mRNA precursor crosslinked *in situ*

(splicing/pre-mRNA secondary structure/processing)

CHRISTOPHE REYMOND, ANNAMARIA AGLIANÓ, AND GEORGES SPOHR

Département de Biologie Animale de l'Université de Genève, Embryologie Moléculaire, 154, Route de Malagnou, 1224 Chene-Bougeries, Switzerland

Communicated by V. Prelog, May 27, 1980

ABSTRACT **Double-stranded RNA segments present in duck globin pre-mRNA were crosslinked *in situ* with aminomethyltrioxalen and UV light. The secondary structure of the crosslinked pre-mRNA was then studied by electron-microscopic analysis of pre-mRNA·cDNA hybrids. The data suggest that duck globin pre-mRNAs contain intervening sequences that are excised stepwise. Excision and subsequent ligation appears to occur on precursor molecules that are stabilized by base-paired regions.**

The presence of double-stranded RNA segments in heterogeneous nuclear RNA (hnRNA) of a variety of eukaryotic cells has been well established (1, 2). It is thought that these regions of secondary structure play a functional role in posttranscriptional modification, processing, or both of the primary RNA transcript to mature mRNA (3, 4). The biochemical properties of these double-stranded RNA structures are well known (5). In addition, they have also been visualized by electron microscopy (6). However, there is no direct evidence that base-paired regions observed in isolated hnRNA reflect the structural arrangement of this RNA *in vivo*.

In a previous paper (7) we showed that hydrogen-bonded molecules of 900,000 daltons hybridized to duck globin cDNA, suggesting that they represent precursors—or processing intermediates—to globin mRNA. Because we were not able to observe these precursor molecules when uncrosslinked RNA was isolated, they very likely represent, for the most part, partially processed molecules that are stabilized by hydrogen bonds and dissociate under denaturing conditions.

In the present report we present data on the investigation by electron microscopy of hybrids between crosslinked duck globin pre-mRNA and globin cDNA. Examination of the hybrid structures reveals that duck globin genes are split, similar to the globin genes of other species (8–11). The data furthermore suggest that the removal of intervening sequences in these pre-mRNAs occurs in a stepwise fashion.

MATERIALS AND METHODS

Cell Fractionation and RNA Extraction. Preparation of duck erythroblast nuclei and RNA extraction have been described (7, 12, 13). To purify globin pre-mRNA, crosslinked nuclear RNA was fractionated by chromatography on Bio-Gel A5 (Bio-Rad), denatured in boiling water, and immediately chromatographed on globin cDNA-cellulose (14). Denaturation and chromatography on cDNA cellulose were repeated a second time.

Treatment of Nuclei with 4-Aminomethyltrioxalen and UV Irradiation. The nuclei pelleted from the cell lysate were resuspended in 1 vol (referred to the volume of the cells used) of psoralen buffer (0.25 M sucrose/10 mM sodium acetate/1 mM magnesium acetate, pH 7) containing 4-aminomethyltrioxalen at 2 mg/ml and left for 10 min at 0–4°C, to allow the drug to penetrate into the nuclei. The nuclei were then diluted with 100 vol of psoralen buffer in order to reduce the absorbance at 360 nm to 1–2 and exposed to a high-pressure Hg vapor UV lamp covered with a cooling mantle that was filled with a circulating cobalt acetate solution (15).

The sample was irradiated from the bottom at 0–4°C for 20 min while being gently mixed. At the end of the irradiation period the nuclei were collected by centrifugation and extracted with phenol as described by Scherrer (12).

4-Aminomethyltrioxalen was synthesized according to the procedure of Isaacs *et al.* (15). Trioxalen was purchased from Paul Elder Co. (Bryan, OH).

Hybrids of Globin cDNA and pre-mRNA. Globin cDNA in 10- to 50-fold molar excess was hybridized to purified globin pre-mRNA by incubating for 1 hr at 60°C in 50 mM Tris·HCl, pH 8.0/5 mM EDTA/0.25 M NaCl.

To minimize formation of RNA aggregates, the hybrids were brought to a final nucleic acid concentration of 0.5 μg/ml in 80% (vol/vol) formamide/4 M urea/50 mM Tris·HCl, pH 8.5/5 mM EDTA/cytochrome *c* at 30 μg per ml and spread on distilled water (16). The nucleic acid/cytochrome *c* film was picked up on Parlodion-coated grids between 10 and 30 sec after spreading, stained for 30 sec in 10 μM uranyl acetate in 90% (vol/vol) ethyl alcohol, and rinsed for 5 sec in 90% alcohol. The grids were rotatory shadowed with Pt/Pd (80:20) at a 7° angle and observed under an AEI 6M electron microscope.

To establish whether DNA·RNA hybrids and double-stranded DNA have different lengths, R-loops of 18S rRNA from *Xenopus laevis* with a plasmid containing the 18S gene were analyzed (17). No difference in the average length per base pair was observed. The replicative form of phage φX174 was used as a length standard for DNA, and single-stranded 18S rRNA from duck was used as a standard for single-stranded RNA.

RESULTS

Experimental Approach. Duck erythroblast nuclei were treated with aminomethyltrioxalen and UV light to crosslink double-stranded RNA *in situ*. Nuclear RNA, which contains approximately 0.5% globin mRNA sequences, was extracted (7), denatured, and enriched for globin RNA sequences by affinity chromatography on cellulose covalently linked to globin cDNA (14). After two rounds of hybridization to cDNA-cellulose, 97% of the bulk hnRNA was eliminated, while approximately 70% of the globin sequences were recovered. We interpret these molecules as globin mRNA precursors (pre-mRNA). The purified, crosslinked pre-mRNA was then hy-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: hnRNA, heterogeneous nuclear RNA; bp, base pair(s).

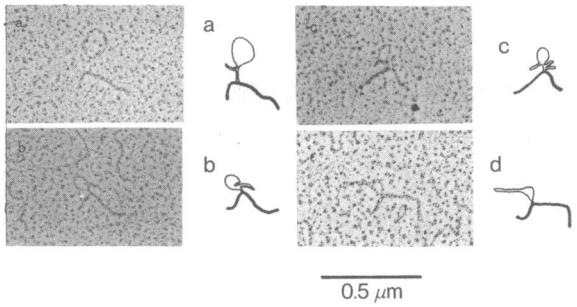

FIG. 1. Electron micrographs of type B hybrid structures with loops. Double-stranded hybrid segments can be distinguished from the single-stranded intervening sequences, which appear as a loop. The average length of these loops is approximately 1100 bases. Micrographs of four molecules are shown, with interpretative drawings.

bridized to excess globin cDNA. This single-stranded cDNA was prepared by reverse transcription of purified total globin mRNA (7). The resulting transcripts, which contained copies of β, α-major, and α-minor globin mRNAs, were homogeneous in size, had a length close to that of globin mRNA, and were at least 90% pure as determined by the kinetics of back hybridization to globin mRNA (7). Pre-mRNA·cDNA hybrids were subsequently spread for electron microscopy under conditions that allow one to distinguish between single- and double-stranded segments (16).

It should be noted that the majority or all of those sequences of the pre-mRNA that are preserved in mature mRNA are still available for hybridization after crosslinking. This has previously been shown by comparison of hybridization kinetics between crosslinked and uncrosslinked RNA to cDNA (7).

Visualization of Pre-mRNAs. Several kinds of hybrid structures could be observed: (*i*) double-stranded molecules of approximately 600 base pairs (bp), which represent hybrids of full-length cDNA and mature mRNA; (*ii*) two hybrid regions that present at their junction a single-stranded RNA loop (Fig. 1); (*iii*) crosslike structures consisting of two hybrid segments and two single-stranded tails (Fig. 2); (*iv*) two hybrid regions that present at their junction either two (or more) RNA loops or RNA loops and single-stranded tails (Fig. 3).

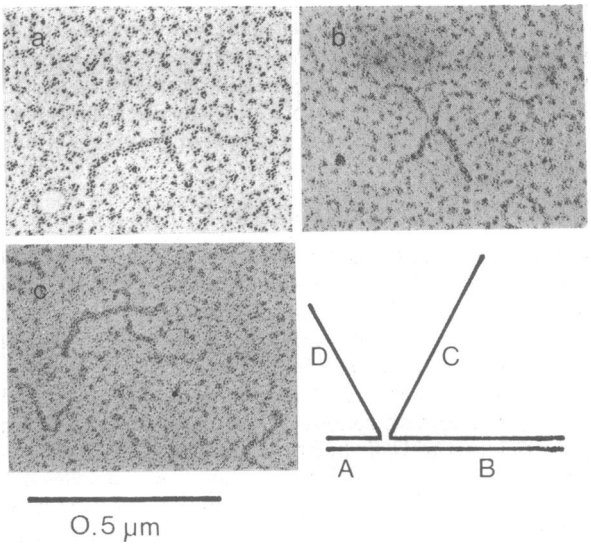

FIG. 2. Electron micrographs and diagrammatic representation of globin pre-mRNA·cDNA hybrids, showing crosslike structures. Double-stranded segments correspond to globin mRNA sequences.

In order to classify the various hybrid structures, we first determined the length of the double-stranded segments in the different molecules. Fig. 4 shows the length distribution of the hybridized segments, and it is obvious that preferential length occurs. Analysis of individual molecules reveals, furthermore, that each hybrid contains a long and a short double-stranded arm (Fig. 5). All molecules exhibit a double-stranded segment of approximately 400 bp; the shorter arms are 140, 230, or 340 bp long (Table 1). In the following we will classify the molecules according to the length of the short double-stranded regions.

Hybrids displaying double-stranded segments of 140, 210, and 340 bp are called type A, B, and C respectively. They represent 47%, 44%, and 9%, respectively, of the examined pre-mRNA structures. The relative frequency of these three types of hybrids is in accordance with the different amount of the three globin mRNAs found in duck reticulocytes (18). They might therefore result from the hybridization of three pre-mRNA species to the corresponding cDNAs.

The existence of single-stranded segments located at the junction of the segments that hybridize with globin cDNA reveals the presence of transcribed intervening sequences in the pre-mRNAs. Hybrid structures of types A and B were therefore further characterized by length measurement of the single-stranded RNA loops and tails. Type C hybrids were found only at a low frequency and could therefore not be characterized conclusively.

Analysis of Type B Hybrids. Type B hybrids contain either one single-stranded loop or two single-stranded tails flanked by double-stranded structures of 400 bp and 210 bp (Figs. 1 and 2). Six out of 59 type B hybrids presented intact loops. Whereas all of the observed loops have a total length of about 1100 bases, the single-stranded tails differ in size. The histogram in Fig. 6*b* shows that the majority of the molecules contain single strands that add up to a total length of either 1050 or 1400 bases. Less frequently a total length of 2100 and 2800 bases is found. The length measurement of about 1050 bases is in good agreement with the length of the observed intact loop. The two single-stranded arms might therefore result from a first cleavage event in the intronic RNA sequence. The appearance of a second peak and two minor ones representing a length of more than 1000 nucleotides might be explained by the fact that in one hybrid structure two partially processed pre-mRNAs are present.

In Fig. 6*a* the length measurements obtained for the single-stranded tails are plotted individually. Prominent peaks at 350, 600, and 1000 bases are found. Because the longest single-stranded segments are of the same size as the intact loops (1000 bases) they might result from a cleavage that occurred at or near the intron/exon junction. The shorter single-stranded tails of 350 and 600 bases might result from a cleavage that occurred inside the intron. The size distribution of the tails suggests that a specific cleavage event has occurred.

As already mentioned, two cleaved pre-mRNAs may be

Table 1. Length of double-stranded segments present in globin pre-mRNA·cDNA hybrids

Type	mol %	Mean length ± SD, bp: Short segment	Long segment	Total	Mean M_r
A	47	140 ± 20	400 ± 70	540 ± 72	177,500
B	44	230 ± 30	390 ± 60	600 ± 47	199,000
C	9	340 ± 50	420 ± 100	760 ± 111	251,000

Double-stranded segments as shown in Figs. 1, 2, and 3 were measured with a Numonics digitizer. Each molecule contains one short and one long double-stranded segment. The classes were defined according to the length of the short segment. In total, 134 hybrid molecules were analyzed.

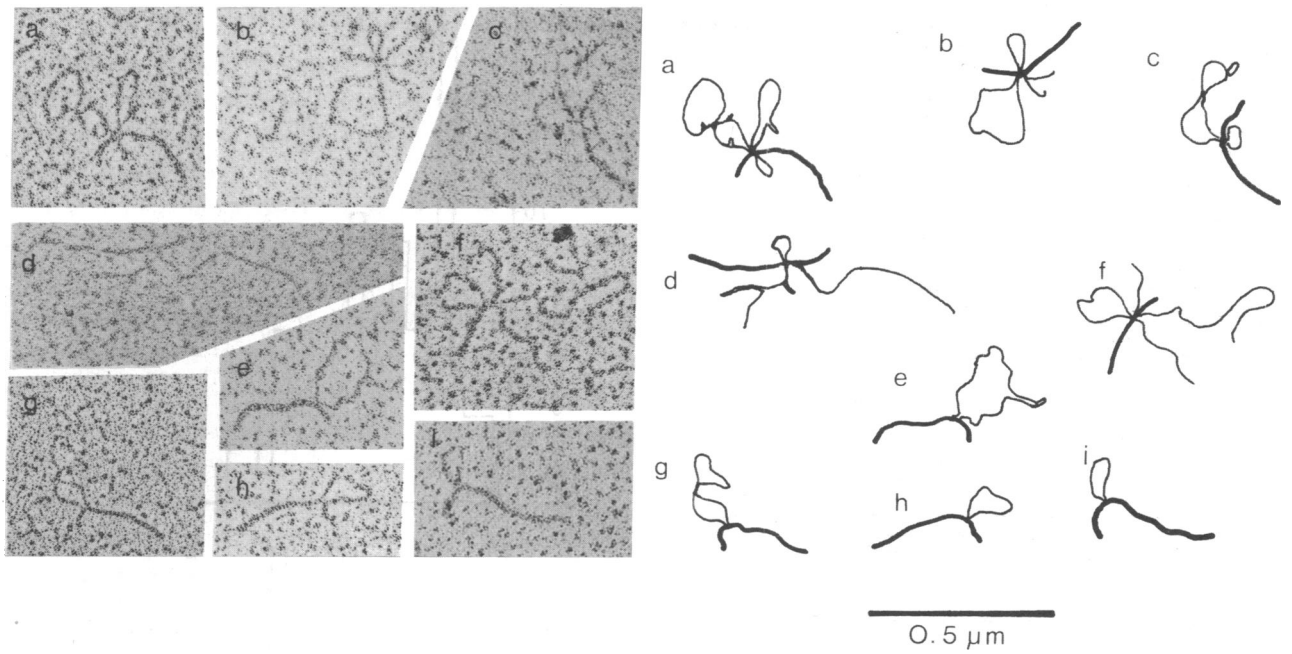

FIG. 3. Electron micrographs of type A hybrid molecules with loops. We observed hybrids with three loops (*a*), two loops (*b, c*), and one loop (*d–i*). In *d* and *f* single-stranded segments that may represent cleaved loops are also present. The average lengths of the loops are approximately 350, 800, and 1300 bases.

found in one hybrid structure. This would result in single-stranded tails that sum up to 2100 bases (two complete introns), to 1600 bases (one complete intron and a 600-base fragment), or to 1400 bases (one intron and a 300-base fragment) (Fig. 6*b*) (see *Discussion*).

Analysis of Type A Hybrids. Common to all type A hybrids (Figs. 3 and 5) are double-stranded segments of 400 bp and 140 bp. The analysis of unhybridized loops and single-stranded tails, however, is more difficult than that of the type B structures. Twelve out of 63 type A hybrids presented loops. One can distinguish by size three types of loops, which are 350, 800, or 1300 bases long. Molecules containing one, two, or rarely three of these loops—which are located close to each other—can be found. Furthermore, one observes hybrids (5 out of 12) that contain not only loops but in addition single-stranded tails (Fig. 3). The presence of more than one loop indicates that this type of pre-mRNA contains more than one intervening sequence. However, the data do not allow us to establish whether two or three introns are present; one intervening sequence may, for example, be folded in a specific manner and show up as two loops.

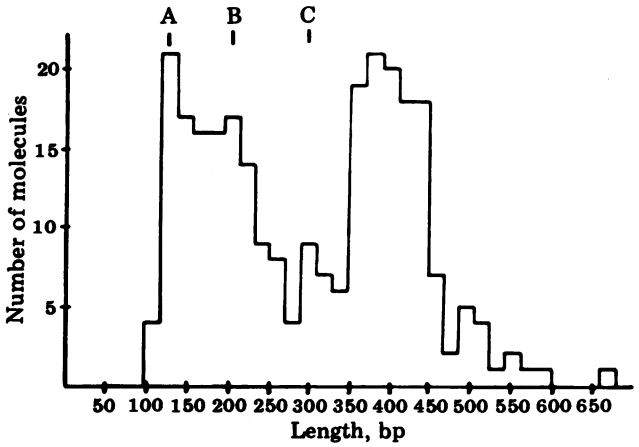

FIG. 4. Length distribution of the double-stranded segments present in hybrid molecules. Double-stranded segments as shown in Figs. 1, 2, and 3 were measured with a Numonics digitizer. Each molecule bears two double-stranded segments, which were measured independently.

When a detailed analysis of the length of the single stranded tails is carried out the resulting histogram shows a complex pattern (data not shown). When one considers that this pre-mRNA contains several introns that might be cleaved internally as in the case of the type B pre-mRNAs, one expects a large number of cleaved intermediates. This number is too large to allow discrimination between individual species in the histogram.

However, the existence of hybrids containing only one loop that is 350, 800, or 1300 bases long suggests that the intervening sequences in the primary transcript are not removed in a given order.

DISCUSSION

Globin mRNA is initially transcribed in the form of a precursor that has to undergo a maturation process. We recently showed (7) that high molecular weight precursors for globin mRNA are found in duck erythroblasts. We were able to observe these pre-mRNAs under denaturing conditions after crosslinking *in situ*. The biochemical properties of these crosslinked pre-mRNAs have been described (7). The principal objective of the present investigation was to study in more detail their structural features by electron microscopy.

Because aminomethyltrioxalen specifically crosslinks base-paired regions in RNA (15), the treated molecules retain the secondary structure that is normally present in the cell and can therefore be isolated under stringent denaturing conditions. Denaturation steps during isolation of hnRNA are necessary to eliminate unspecific aggregates that are frequently found in hnRNA. As already mentioned, we did not find high molecular weight globin pre-mRNAs in significant amounts when the trioxalen treatment was omitted. This observation is consistent with the electron microscopic data presented here. In fact only a low number of intact pre-mRNA molecules was observed. This amount is too low to be measured by the biochemical analysis described earlier (7, 13). This suggests that

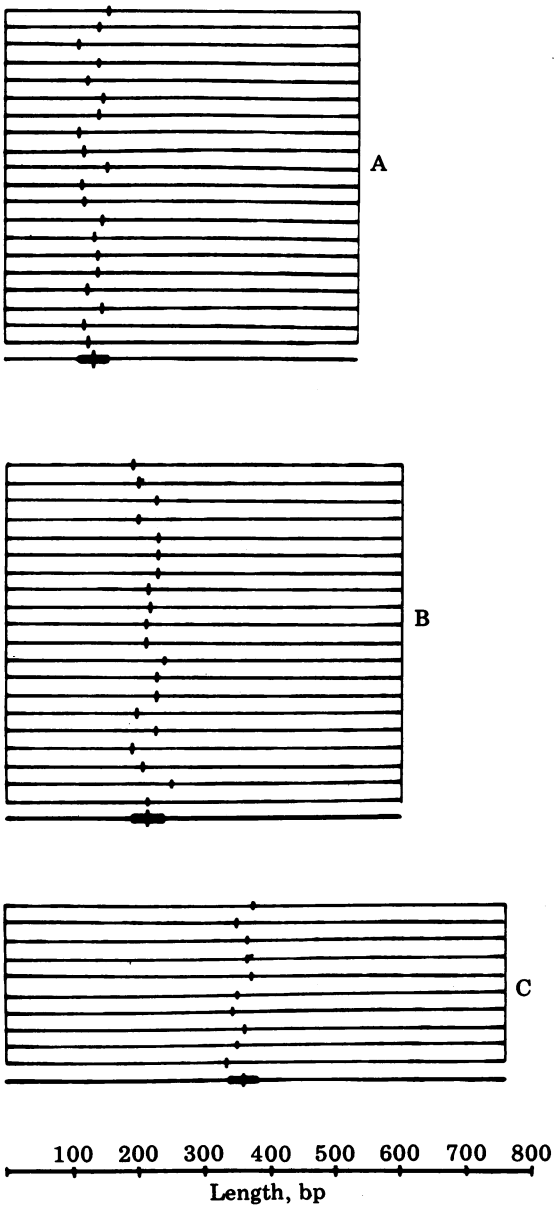

FIG. 5. Gallery of hybrids. Randomly chosen pre-mRNA hybrids were classified according to the length of the short double-stranded segments. Total length was normalized to the average length of each class. The bottom bar of each set of molecules represents the average position and the standard deviation of the junction of the double-stranded segments.

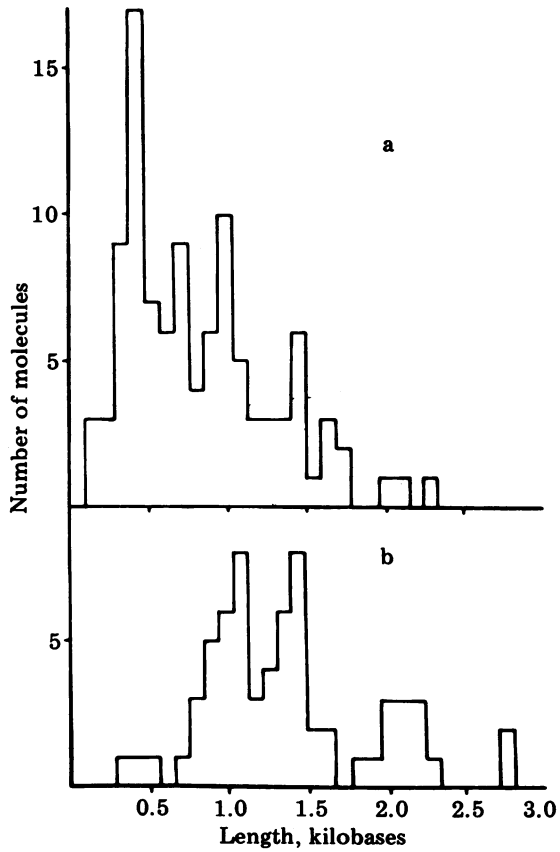

FIG. 6. Length distribution of the single-stranded tails of the type B molecules. The lengths C and D depicted in Fig. 2 were measured either independently (*a*) or added together (*b*). In *b* two predominant lengths can be seen at 1100 and 1400 bases. In *a* at least three predominant lengths of 390, 1000, and 1400 bases are observed.

most of the precursor molecules might be nicked or partly processed at specific sites. Regions of secondary structure that *in vivo* might stabilize such processing intermediates would dissociate under denaturing extraction conditions and could not be observed unless stabilized by aminomethyltrioxalen. In order to test this hypothesis and to localize those sequences of the globin pre-mRNA that are preserved in mature mRNA, crosslinked globin pre-mRNA was hybridized to cDNA complementary to duck globin mRNA.

Examination by electron microscopy of the resulting hybrids revealed the presence of two main populations of hybrids (type A, type B) and a minor species (type C), which were characterized by the length of their double-stranded segments. We would like to point out that the relative proportion of the type A, B, and C structures is very similar to the relative amount of the three globin mRNA species of duck reticulocytes (18). The globin cDNA preparation that was used in these experiments was 90% pure as judged by the kinetics of back hybridization to globin mRNA. Though it cannot be excluded that minor nonglobin sequences are present in the cDNA preparation, they would not have obscured our analysis because only the most frequently occurring hybrid structures were further examined. The number of type C hybrids was already too low to allow interpretable length measurements. Less frequent globin pre-mRNA species would therefore also have escaped our analysis. For this reason it is not surprising that the duck globin precursor RNA of 2×10^6 daltons that has been described by Imaizumi *et al.* (19) was not detectable in our hybrid population.

Examination of type A and B hybrids reveals that both kinds of pre-mRNAs contain intervening sequences, an observation that is not surprising, because all genomic globin genes investigated so far have been shown to be interrupted by introns (8–11). These intervening sequences are transcribed and removed during pre-mRNA processing (20–22). Whereas type B pre-mRNAs contain only one large intron of 1000 bases, type A contain at least two intervening sequences. Furthermore, it is necessary to point out that smaller intervening sequences (100 bases) might have escaped detection. Fig. 2 shows that type B hybrids frequently display intronic RNA sequences that do not form a loop but rather consist of two single-stranded tails of defined length. If these single-stranded tails result from degradation during sample preparation, one would not expect them to be of defined size. Furthermore, the sum of the length of the two RNA tails is frequently equal to the length of the uncleaved intron. These results might therefore reflect specific processing steps that are necessary for the removal of the intervening sequence.

As suggested by the presence of single-stranded tails that add up to more than one intervening sequence, two pre-mRNAs may be present in one hybrid structure. This could be due to the fact that, in spite of the cDNA excess used during hybridization, two pre-mRNAs have hybridized to one cDNA molecule. Alternatively, it is theoretically possible that the two pre-mRNAs were associated already in the nucleus.

The data furthermore suggest that the intervening sequences are removed in a stepwise manner in which a first cleavage can occur either at the intron/exon junction or within the intron itself. A stepwise excision of intronic sequences has already been found by Kinniburgh and Ross (20) for mouse globin pre-mRNA. These authors showed that the removal of the larger intervening sequence involves at least two cleavage–ligation reactions.

As already pointed out, the interpretation of type A pre-mRNA is complicated. The intervening sequences in this pre-mRNA are located so close to each other that it is impossible to correlate the single-stranded tails to the loops from which they have derived. However, the existence of several loops of defined size in one hybrid molecule and the presence of molecules that contain only one loop of either size is suggestive of a stepwise removal of the intervening sequences that occurs not in a given order.

Though the duck globin genes have not been isolated and described in detail so far, the presented data do nevertheless allow to describe some aspects of pre-mRNA maturation. It is evident that base-paired structures in pre-mRNAs play an important role during the various steps that a precursor molecule has to undergo during processing. We do not know whether the observed processing intermediates are hydrogen bonded by inverted repeated sequences present in the pre-mRNA itself or whether small RNAs are involved (23).

Stabilization of base-paired regions in pre-mRNA by cross-linking with aminomethyltrioxalen may be a useful tool for the more detailed analysis of the complex steps involved in the processing of the primary transcript.

We thank Dr. Joseph Beard, The Office of Program Resources and Logistics, Viral Cancer Program, National Cancer Institute, Bethesda, MD, for the gift of reverse transcriptase, Bernard Dumont for help in electron microscopy, Noëlle Charpentier and Christine Gonet for excellent technical assistance, and Freddy Bourquin and Hélène Felder for help during preparation of the manuscript. In addition we are grateful to Irmi Sures for help and advice and to Brian Jackson for critical reading of the manuscript. This investigation was supported by the Swiss National Science Foundation (Grant 3.170.0.77).

1. Jelinek, W., Molloy, G., Fernandez-Munoz, R., Salditt, M. & Darnell, J. E. (1974) *J. Mol. Biol.* **82,** 361–370.
2. Ryskov, A. P., Saunders, G. F., Farashyan, V. R. & Georgiev, G. P. (1973) *Biochim. Biophys. Acta* **312,** 152–163.
3. Calvet, J. P. & Pederson, T. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3705–3709.
4. Jelinek, W. R. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2679–2683.
5. Robertson, H. D., Dickson, E. & Jelinek, W. (1977) *J. Mol. Biol.* **115,** 571–589.
6. Fedoroff, N., Wellauer, P. K. & Wall, R. (1977) *Cell* **10,** 597–610.
7. Aglianó, A. M., Nacci, A., Reymond, C., Appleby, D. & Spohr, G. (1979) *Eur. J. Biochem.* **95,** 203–213.
8. Jeffreys, A. J. & Flavell, R. A. (1977) *Cell* **12,** 1097–1108.
9. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15,** 1157–1174.
10. Tiemeier, D. C., Tilghman, S. M., Polsky, F. I., Seidman, J. G., Leder, A., Edgell, M. H. & Leder, P. (1978) *Cell* **14,** 237–245.
11. Dodgson, J. B., Strommer, J. & Engel, J. D. (1979) *Cell* **17,** 879–887.
12. Scherrer, K. (1969) in *Fundamental Techniques in Virology,* eds. Habel, K. & Salzman, N. P. (Academic, New York), pp. 413–432.
13. Spohr, G., Dettori, G. & Manzari, V. (1976) *Cell* **8,** 505–512.
14. Wood, T. G. & Lingrel, J. B. (1977) *J. Biol. Chem.* **252,** 457–463.
15. Isaacs, S. T., Shen, C. J., Hearst, J. E. & Rapoport, H. (1977) *Biochemistry* **16,** 1058–1064.
16. Wellauer, P. K. & Dawid, I. B. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 2827–2831.
17. White, R. L. & Hogness, D. S. (1977) *Cell* **10,** 177–192.
18. Spohr, G., Imaizumi, T., Stewart, A. & Scherrer, K. (1972) *FEBS Lett.* **28,** 165–168.
19. Imaizumi, T., Diggelmann, H. & Scherrer, K. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 1122–1126.
20. Kinniburgh, A. J. & Ross, J. (1979) *Cell* **17,** 915–921.
21. Kinniburgh, A. J., Mertz, J. E. & Ross, J. (1978) *Cell* **14,** 681–693.
22. Tilghman, S. M., Curtis, P. J., Tiemeier, D. A., Leder, P. & Weissman, C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1309–1313.
23. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. (1980) *Nature (London)* **283,** 220–224.