Volume 6 Number 7 1979 Nucleic Acids Research

Cloning of an almost full-length chicken conalbumin double-stranded cDNA

M.Cochet, F.Perrin, F.Gannon, A.Krust and P.Chambon
Lab. Génétique Moléculaire des Eucaryotes du CNRS, U 184 Biologie Moléculaire et Génie Génétique, INSERM, Inst. Chimie Biologique, Faculté de Médecine, Strasbourg, France

G.S.McKnight, D.C.Lee, K.E.Mayo and R.Palmiter
Department of Biochemistry, University of Washington, Seattle, WA 98195, USA

Received 9 March 1979

ABSTRACT

Chicken conalbumin double-stranded cDNA (con-dscDNA) was synthesized from a laying hen oviduct mRNA preparation enriched for conalbumin mRNA (con-mRNA). The dscDNA was inserted by blunt-end ligation into the SalI site of plasmid pBR322 which had been repaired with DNA polymerase I to create TaqI sites on each side of the inserted fragment. After bacterial transformation, one hybrid recombinant, pBR322-con1, which contains the largest inserted dscDNA (about 2350 bp) was shown to hybridize specifically to the RNA which is translated into conalbumin. Electron microscopic examination of hybrid molecules between con-mRNA and pBR322-con1 DNA indicate that the inserted con-dscDNA is an almost full-length double-stranded transcript of conalbumin mRNA.

INTRODUCTION

In the chicken oviduct the synthesis of the egg white proteins ovalbumin, conalbumin, ovomucoïd and lysozyme, is controlled by steroid hormones at the level of transcription (1-7). A comparison of the structural organization of the DNA within and around these genes should be helpful in understanding the molecular mechanisms involved in this hormonal control. The analysis of the structure of ovalbumin gene is well advanced (8-12). Comparison of the structure of the ovalbumin and the conalbumin genes should be very informative because, although both genes are induced by the presence of the steroid hormones oestradiol and progesterone, there are subtle differences in their responses. For example, the level of conalbumin messenger detectable in chicken oviduct mRNA increases very rapidly after administration of oestradiol whereas there is an apparent lag in the response of the ovalbumin gene (7,13). In addition, the ovalbumin gene is almost unexpressed in the liver (14) whereas the conal-

bumin gene is expressed constitutively (7) [the protein transferrin is the name given to the product of the "conalbumin" gene in the liver (7, 15)]. It is obvious that the complete analysis of the structure of the conalbumin gene can only be performed after its purification by molecular cloning. As a first step towards this goal we report here the cloning, identification and characterization of a plasmid recombinant which contains an almost full-length double-stranded cDNA copy of conalbumin mRNA.

MATERIALS AND METHODS

1. Enzymes.

Highly purified reverse transcriptase from avian myeloblastosis virus (AMV) was obtained from Dr. Beard (Life Sciences Inc., St. Petersbourg, USA). E.coli DNA polymerase I [15,000 poly dAT units/mg] was generously supplied by Dr. A. Kornberg and Dr. D. Brutlag. Endonuclease-free nuclease S1 from *Aspergillus oryzae* was purified according to Vogt (16). Bacterial alkaline phosphatase (BAPF) bought from Worthington (USA) was dialyzed against Tris-HCl 10mM, pH 7.5, NaCl 10 mM, and heated for 30 min at 80°C before use. T4 DNA ligase (3,300 units/ml using blunt-end DNA) was obtained from New England Biolabs (Beverly, USA). E.coli exonuclease III was purchased from Miles (U.K.). EcoRI, Bam H1 and HindIII restrictions enzymes were prepared according to the procedure of Sümegi et al. (17), Wilson and Young (18) and Humphries et al. (19), respectively. Other restriction enzymes were obtained from New England Biolabs or Bethesda Research Laboratories (Rockville, USA). Restriction endonuclease digestions were carried out as described in the New England Biolabs booklet. Labelled deoxyribonucleoside triphosphates were purchased from Amersham (U.K.).

2. Synthesis of complementary DNA (cDNA) and double-stranded cDNA (dscDNA) from partially purified conalbumin mRNA.

Nucleic acids were extracted from laying hen oviduct by proteinase K (Merck) digestion in the presence of SDS, phenol-chloroform extraction, and ethanol precipitation (7). After dissolution in water, high molecular weight RNA was precipitated by addition of 2 volumes of 3 M Na acetate, 5 mM EDTA pH 7.0

The precipitated RNA was then washed with 3 M Na acetate, dissolved in 0.5 M NaCl, 0.1% SDS, 4 mM EDTA, and 10 mM Tris-HCl pH 8.0 and applied to an oligo(dT)-cellulose column (Collaborative Research) as described elsewhere (7). After elution, the polyA-containing RNA was heated at 65°C in 1% SDS, 5 mM EDTA, 10 mM Tris-HCl pH 7.5 and sedimented on 5-20% sucrose gradients in the same buffer for 8 h at 280,000 g in a Beckman SW41 rotor. Fractions were assayed for both ovalbumin and conalbumin mRNA using cDNAs specific for each mRNA as previously described (7). The peak fractions of conalbumin mRNA (con-mRNA) were pooled and the mRNA further purified by two additional sucrose gradient centrifugations.

Conalbumin cDNA (con-cDNA) was synthesized from the partially purified con-mRNA using reverse transcriptase as previously reported (7). In the preparation which was used for con-dscDNA synthesis, actinomycin D was not added and [^{3}H]-dCTP (280 cpm/pmole) was used as the radioactive label. The con-cDNA preparation which was used as hybridization probe, was synthesized in the presence of Actinomycin D and [^{3}H]-TTP (100 Ci/mmole) as radioactive label.

Con-dscDNA was prepared as follows. cDNA molecules larger than about 1500 nucleotides were selected by centrifugation through alkaline sucrose gradient (20). 0.16 to 0.20 µg of these cDNA molecules was incubated for 1 h at 37°C with 18 units of E.coli DNA polymerase I in 30 mM Tris-HCl pH 8.0, 4 mM $MgCl_2$, 10 mM β-mercaptoethanol, 200 µg/ml DNase-free bovin serum albumin (21), 0.2 mM [^{32}P]-dCTP (200 cpm/pmole) and 0.2 mM of dATP, dGTP and TTP. The dscDNA was phenol-extracted, ethanol-precipitated and treated for 90 min at 37°C in 500 µl of 30 mM sodium acetate pH 4.5, 3 mM $ZnCl_2$ and 0.3 M NaCl with 100 units of S1 nuclease (16) to cut the hairpins. After phenol extraction and ethanol precipitation the dscDNA was centrifuged in a neutral 5-25% sucrose gradient (SW56, 31,000 rpm, 15 h at 20°C). Fractions containing molecules which on the average were longer than 1500 bp (about 120 ng) were pooled and ethanol-precipitated using E.coli tRNA (10 µg/ml) as a carrier. Finally, the extremities of the dscDNA were made perfectly flush (blunt) by incubation with E.coli DNA polymerase I as described by Seeberg et al. (22).

3. Construction of recombinant DNA plasmids.

a) Preparation of a pBR322 vector with TaqI sites at its extremities. The extremities of 8.5 µg of pBR322 linearized with SalI restriction enzyme were repaired to create TaqI sites by incubation for 20 min at 14°C with 15 units of E.coli DNA polymerase I in 250 µl of the same reaction mixture as above (22). The 5' terminal phosphates of the vector DNA were removed by treatment (15 min at 65°C) with 1.5 units of bacterial alkaline phosphatase (23). After phenol extraction, ethanol precipitation and ligation, 1 µg of such a treated vector yielded about 1000 transformant colonies, whereas 1 µg of untreated pBR322 DNA yielded 2.10^6 transformant colonies.

b) Ligation. 50 ng of pBR322 DNA treated as described above and 20 ng of dscDNA were "blunt-end" ligated by incubation (in 10 µl of 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol and 1 mM ATP) for 18 h at 12°C in the presence of 1 µl of T4 ligase (24).

4. Transformation and identification of recombinant clones.

E.coli C600 $r_k^- m_k^-$ rec $B^- C^-$ was transformed with half of the hybrid plasmid DNA as described in Humphries et al. (25). Transformed ampicillin-resistant and tetracycline sensitive colonies were selected (see Result section) and replated. The plasmid DNA of each colony was then analyzed for its size essentially as described by Telford et al. (26). A part of each selected colony was suspended in 20 µl of 50 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.5% SDS and incubated for 2 h at 37°C with 50 µg of proteinase K. The lyzed material was brought to 1% sucrose, 0.01% bromophenol blue, mixed with 20 µl of 0.8% agarose and run on 0.8% agarose slab gel at 50 volts for 6 h in 40 mM Tris-acetate pH 7.8, 1 mM EDTA buffer. The colonies containing the longest hybrid plasmid DNA molecules were grown in 10 ml of L broth, and their plasmid DNA was amplified by chloramphenicol (see below). The DNA was extracted and hybridized to conalbumin [^{3}H]-TMP-labelled cDNA as previously described by Humphries et al. (25) for ovalbumin.

5. Preparation of plasmid DNA.

Cultures were grown to the logarithmic phase before amplification by addition of 200 µg/ml of chloramphenicol (27). Extraction and purification of plasmid DNA was achieved essentially by the cleared lysate technique of Clewell et al. (28) followed by two successive ethidium bromide-ClCs equilibrium gradients.

6. Orientation of the inserted conalbumin dscDNA.

The method of Cohn et al. (29) was followed. 2.5 µg of purified pBR322-con1 DNA (see Result section) was linearized with HindIII restriction enzyme and treated (22°C, 100 µl) with 20 units of exonuclease III in 66 mM Tris-HCl pH 8.0, 90 mM NaCl, 5 mM $MgCl_2$ and 5 mM dithiothreitol. Aliquots containing 0.625 µg of plasmid DNA were taken over a period of 2 h, diluted five times with 1 mM EDTA and 50 mM Tris-HCl pH 8.0, hybridized with [^{3}H]-TMP-labelled cDNA either before or after heat-denaturation, and the percentage of the [^{3}H] counts resistant to nuclease S1 digestion was determined.

7. Purification of conalbumin mRNA by hybridization with pBR322-con1 DNA immobilized on filters.

Wild type pBR322 DNA, pBR322-con1 DNA (see below) and pCR1-Ov2.1 DNA (25) were immobilized on nitrocellulose filters as follows. DNA in 2 M NaCl, 0.2 M NaOH was heated at 100° for 1 min, cooled, and spotted on 7 mm diameter Sartorius filters at a concentration of 1-2 µg per filter. The filters were baked for 2 h at 80°C, washed thoroughly in 0.3 M NaCl, 2 mM EDTA, 0.1% SDS, 10 mM Tris-HCl pH 7.5 and allowed to dry at room temperature.

The hybridization reactions (60 µl) were prepared in 5 ml polypropylene tubes (Falcon) and contained 50 µg/ml of partially purified mRNA up to the first sucrose gradient (see above) in 30% formamide, 0.5 M NaCl, 0.4% SDS, 2 mM EDTA, 50 mM Pipes pH 7.5, and one of each filter type. The reaction was overlaid with light paraffin oil and incubated at 47°C for 2 h with gentle shaking. The reaction was stopped by dilution with 10 volumes of 0.1 M NaCl. After hybridization, the filters were washed several times in 15 ml of 0.3 M NaCl, 0.1% SDS, 2 mM EDTA, 10 mM Tris-

HCl pH 7.5, by shaking in a 45°C water bath, and subsequently they were washed several times in the same solution without SDS at 4°C. The filters were then placed individually in polypropylene tubes containing 30 µl of 5 mM KCl, and heated for 1 min in a boiling water bath to elute the hybridized RNA; the tubes were then immersed in ice water and the filters removed as quickly as possible.

8. Translation of purified conalbumin mRNA.

Aliquots of the solution into which the mRNA had been eluted were translated in a cell-free, mRNA-dependent, lysate derived from rabbit reticulocytes. Translation was essentially as described (30) and included [^{35}S]-Cysteine and ribonuclease-free antibody to conalbumin at 46 µg/ml. At the end of the incubation period, aliquots of the translation mixture (2-10 µl) were diluted to 30 µl with distilled water and added to an equal volume of gel sample buffer, 40% glycerol, 4% SDS, 1% β-mercapthoethanol, 0.1 M Tris-HCl pH 6.8. The samples were boiled for 1 min, cooled, and electrophoresed on a 10% polyacrylamide/SDS slab gel by the method of Studier (31). Electrophoresis was for 5 h at 25 mA. The gel was stained with Coomassie brilliant blue, destained, and prepared for fluorography by the method of Bonner and Laskey (32).

Aliquots of the translation mixture were also precipitated with 5 µg of unlabelled conalbumin and excess anti-conalbumin serum (30); the immunoprecipitates were collected, resuspended in sample buffer, and electrophoresed as described above.

9. Electron microscopy.

Heteroduplex molecules between pBR322 and pBR322-con1 were formed by hybridizing, at 25°C for 30 min in 70% deionized formamide, 300 mM NaCl, 10 mM Tris-HCl pH 8.5, and 1 mM EDTA, 2 µg/ml of each DNA linearized with Hind III and heat-denatured. RNA-DNA hybrid molecules were obtained by incubating purified con-mRNA (3 µg/ml) and heat-denatured pBR322-con1 DNA (3 µg/ml) linearized with HindIII at 59°C for 4 h in 70% deionized formamide, 300 mM NaCl, 10 mM Tris-HCl pH 8.5 and 1 mM EDTA. Samples were mounted for

electron microscopy at a concentration of about 0.3 µg/ml of DNA as previously described (33).

10. Biohazards.

Biohazards associated with the experiments described in this publication were examined previously by the French National Control Committee, and carried out according to the rules followed by this Committee (Le Progrès Scientifique, N° 191, Nov/Dec 1977), under L3-B1 conditions.

RESULTS

I. Construction and selection of hybrid plasmids containing conalbumin double-stranded cDNA.

Conalbumin mRNA was purified from total nucleic acids of laying hen oviduct by oligo(dT)-cellulose chromatography and three centrifugations through sucrose gradients (Materials and Methods). This purification yields an mRNA fraction highly enriched in con-mRNA which contains less than 1% of ovalbumin mRNA when assayed by hybridization with ovalbumin cDNA (7). The partially purified con-mRNA preparation was then used as a template for AMV reverse transcriptase in the absence of actinomycin D to increase the proportion of molecules with hairpins (Materials and Methods). As shown in Fig. 1, such a cDNA preparation contains about 50% of sequences complementary to con-mRNA and 50% of sequences synthesized from contaminating minor mRNA species.

The longest con-cDNA molecules were selected by centrifugation through a sucrose gradient and converted to double-stranded cDNA (con-dscDNA) with E.coli DNA polymerase I (Materials and Methods). After treatment with nuclease S1 to cleave the hairpin, the longest molecules were again selected by sucrose gradient centrifugation and incubated in the presence of E.coli DNA polymerase I to obtain perfect blunt ends.

The con-dscDNA molecules were inserted by blunt-end ligation at the SalI site of plasmid pBR322. Before ligation, the

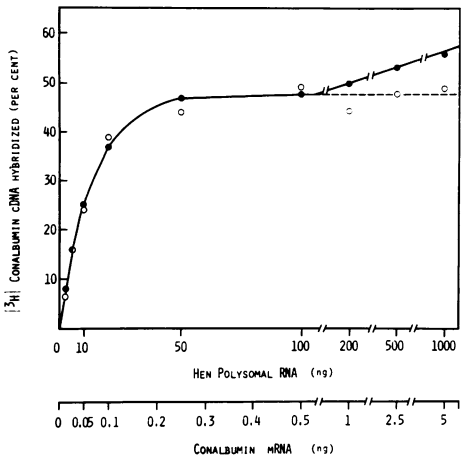

Fig. 1 : [^{3}H]-TMP labelled cDNA was hybridized with laying hen polysomal RNA (closed circles) or with pure con-mRNA (open circles) as described by Lee et al. (7). After hybridization the samples were treated with nuclease S1 and the acid-insoluble radioactivity was counted.

SalI linearized plasmid was first treated with E.coli DNA polymerase I to generate one TaqI site on each side of the inserted sequence, and then with alkaline phosphatase to prevent self-ligation (Materials and Methods). In addition, since the SalI site is located within the tetracycline-resistance gene (34), a double screening, first on ampicillin and then on tetracycline plates, allows one to eliminate wild type pBR322-containing colonies which are both ampicillin and tetracycline-resistant.

120 ampicillin-resistant and tetracycline-sensitive colonies were obtained as described in Materials and Methods. These colonies were analyzed for the size of their plasmid DNA by agarose gel electrophoresis. The 10 colonies which contain the largest recombinant plasmids were grown in 10 ml cultures, "amplified" with chloramphenicol, and their plasmid DNA was extracted and hybridized to [^{3}H]-labelled con-cDNA as described in Materials and Methods. One clone, pBR322-con1, was found after hybridization to protect 45% of the sequences present in the con-cDNA preparation against nuclease S1 digestion. Since not more than 50% of the con-cDNA sequences were actually conalbumin-specific (see above and Fig. 1), we concluded that clone pBR322-con1 was very likely to contain extensive conalbumin dscDNA sequences.

II. Plasmid pBR322-con1 contains an almost "full-length" conalbumin double-stranded cDNA copy.

1. Translation of the mRNA which hybridizes to the plasmid pBR322-con1 demonstrates that the inserted DNA is conalbumin-specific.

A partially purified mRNA preparation containing both conalbumin and ovalbumin (ov) mRNAs was hybridized to pBR322-con1 DNA, pCR1-ov2.1 DNA (25) and pBR322 DNA immobilized onto nitrocellulose filters (Materials and Methods). The hybridized RNA was then eluted from the different filters and translated in a mRNA-dependent cell-free lysate. As shown in Fig. 2, the RNA eluted from the pBR322-con1 filter is specifically translated into a single band (lane 4) which comigrates with a [^{14}C]-labelled conalbumin standard (lane 6), whereas the RNA eluted from the pCR1-ov2.1 filter is specifically translated into ovalbumin (lane 3). The band migrating slightly faster than ovalbumin in lanes 3 and 5 is frequently observed and is immunoprecipitable with anti-

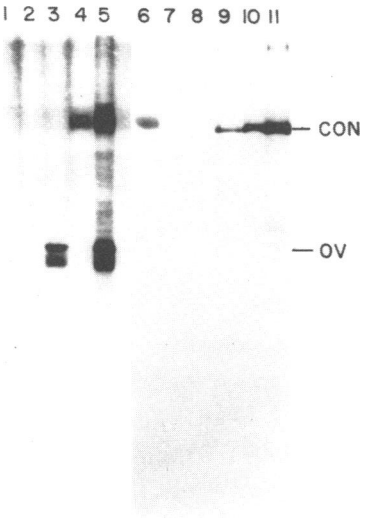

Fig. 2 : Fluorograph from 10% polyacrylamide/SDS slab gels of proteins translated *in vitro* from mRNA specifically bound to filters containing plasmid DNA. Samples 1 through 5 are products applied directly to the gel following translation. Samples 7 through 11 are the precipitates which result from immunoprecipitation of samples 1 through 5 with antibody specific for conalbumin. Lane 1: Control translation with no added mRNA. Lanes 2, 3, 4: Translation products of mRNA specifically bound to filters containing DNA from the plasmids pBR322 (2), pCR1-ov2.1 (3) and pBR322-con1 (4). Lane 5: Translation products from partially purified mRNA used in the hybridization. This represents the total RNA population available for binding to the filters. Lane 6 : [^{14}C]-conalbumin marker synthesized in oviduct explants. Lanes 7, 8, 9, 10: Conalbumin specific immunoprecipitates of samples 2,3,4,5, respectively. Lane 11 : same as 10, four times as much sample.

ovalbumin; it presumably represents premature termination products. As expected, no translatable RNA was eluted from the control pBR322 filter (lane 2). That the protein band seen in lane 4 is conalbumin was confirmed by precipitation with anti-conalbumin serum and electrophoresis of the precipitated proteins. All of the bands attributed to conalbumin were still present after immunoprecipitation (lanes 9 to 11), whereas those corresponding to ovalbumin were lost after precipitation with the anti-conalbumin serum (lanes 8 and 10). We conclude from these results that the dscDNA inserted in clone pBR322-con1 is conalbumin specific.

2. Size, restriction enzyme mapping and orientation with respect to transcription of the inserted conalbumin dscDNA.

The size and the restriction enzyme map of the inserted con-dscDNA were established by digestion with various enzymes and comparison with wild type pBR322 (the detailed results are not shown, but are available upon request). For instance, digestion of pBR322-con1 DNA with HaeII yielded only one additional fragment of 2575 bp not present in the pBR322 digest. Since the SalI site of pBR322 is located within a 227 bp HaeII fragment (35), we conclude that the inserted con-dscDNA is about 2350 bp long. Fig. 3 shows the map of pBR322-con1 within (a) and around (b) the inserted con-dscDNA. The TaqI sites ($TaqI_1$ and $TaqI_4$) have been effectively created at the plasmid insertion sites by repairing the pBR322 SalI site with E.coli DNA polymerase I (see above). Among the enzymes which cut rather unfrequently the chicken DNA, the con-dscDNA contains one EcoRI, one BglI, one BglII, two BamH1 and at least two HhaI sites, whereas there are no sites for HpaI, XhoI, SstI, XbaI, Kpn I, HindIII and PstI enzymes. In addition to the restriction enzyme sites shown in Fig. 3, we have found two or more sites for MboII, HpaII, HgaI, HphI, TacI, HaeIII and AluI.

The inserted con-dscDNA sequence was oriented with respect to transcription by digestion with E.coli exonuclease III and hybridization with con-cDNA according to Cohn et al. (29). pBR322-con1 linearized by digestion with HindIII (Fig. 3b) was

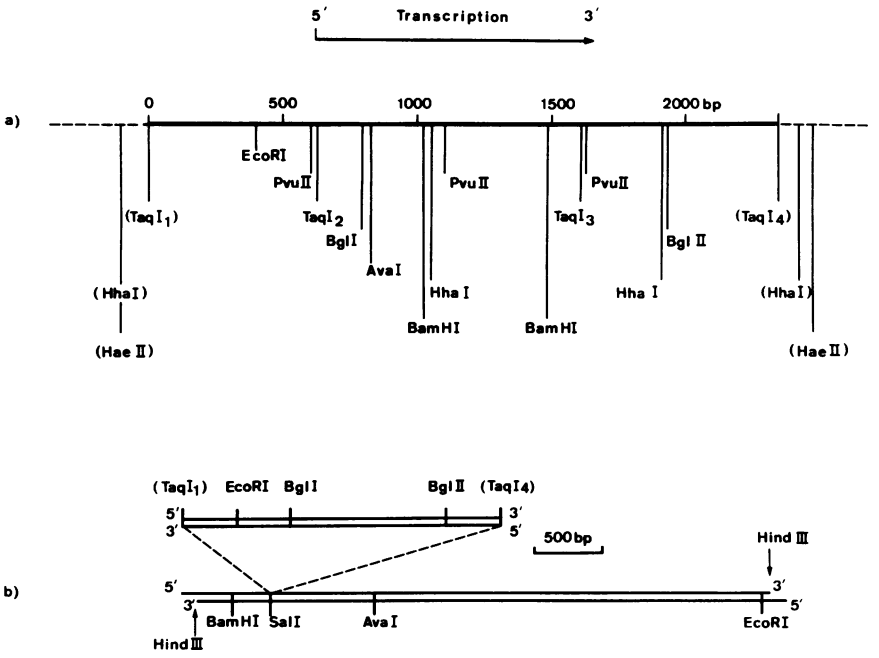

Fig. 3 : Restriction enzyme mapping of pBR322-con1 DNA. (a) Restriction enzyme map of the con-dscDNA segment inserted in pBR322-con1 DNA. The heavy and the dashed lines represent the con-dscDNA and pBR322 vector DNA sequences, respectively. The restriction enzyme sites found in the pBR322 DNA are between brackets. There might be an additional HhaI site between BglII and $TaqI_4$ sites. (b) Orientation of the con-dscDNA segment within the pBR322 vector DNA. The lower lines represent pBR322 DNA linearized with HindIII endonuclease, whereas the upper lines represent the con-dscDNA sequences inserted at the SalI site of the vector DNA.

digested with E.coli exonuclease III. Aliquots were taken at increasing lengths of time and hybridized with [^{3}H]-labelled con-cDNA before and after heat-denaturation. We found (Table 1) that with increasing time of digestion, an increasing proportion of the cDNA was hybridized by the non-denatured samples, up to the value obtained with the denatured ones. From these results we concluded that the strand complementary to the cDNA (and therefore not coding for the mRNA) is not digested by exonuclease III and that the orientation of transcription is as indicated in

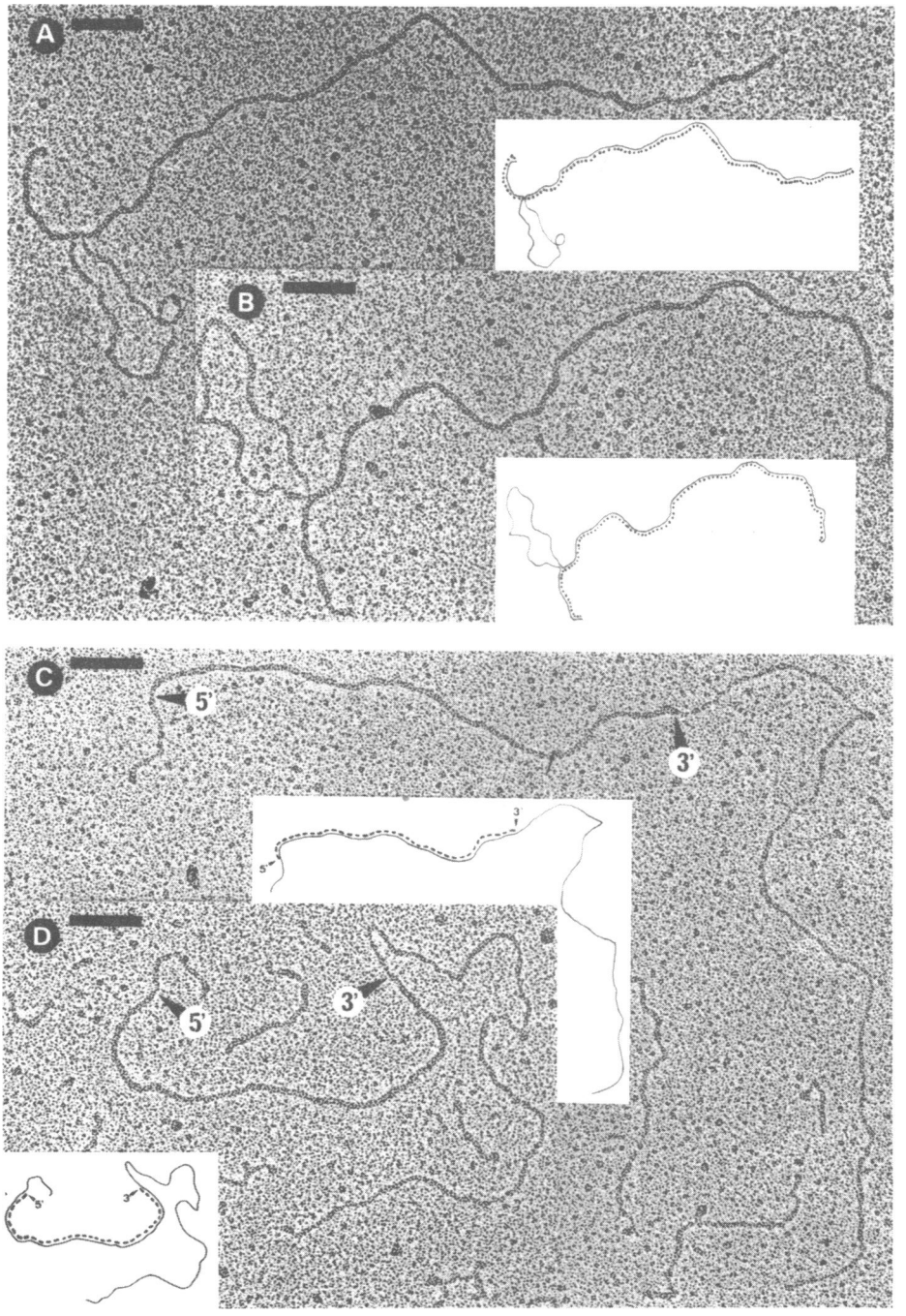
A
B
C
5'
3'
D
5'
3'

Fig. 3, with the $TaqI_1$-EcoRI region of the transcribed strand of DNA coding for the 5' end of con-mRNA.

3. Electron microscopic studies reveal that the inserted con-dscDNA is almost "full-length".

Electron microscopic examination of heteroduplex molecules between pBR322-con1 and pBR322 DNAs linearized with Hind III confirmed that the length of the inserted con-dscDNA is actually about 2350 bp. The length of the con-dscDNA loop (Fig. 4, A and B) is 2290 ± 182 nucleotides, in excellent agreement with the length determined by agarose gel electrophoresis. That this length is very close to that of con-mRNA was shown by hybridizing pBR322-con1 DNA with con-mRNA under conditions which favor the formation of DNA-RNA hybrids (Fig. 4, C and D). No RNA tails were seen at the 5' and 3' ends of the hybrid region (2200 ± 264 bp). Since a 100 nucleotide-long tail could be detected under these electron microscopic conditions (F. Perrin, unpublished results), we conclude 1) that the 3' poly A tail of con-mRNA is on the average shorter than that of ovalbumin mRNA which can be visualized by electron microscopy (33), and 2) that the con-dscDNA present in pBR322-con1 is an almost full-length copy of con-mRNA which, at most, could lack the sequences corresponding to the first 5'-terminal 100 nucleotides of con-mRNA.

Fig. 4 : A) and B) : Electron micrographs of heteroduplex molecules between pBR322 and pBR322-con1 DNAs linearized by HindIII restriction enzyme. A single-stranded DNA loop, 2290 ± 182 nucleotides long is located at 655 ± 74 bp from one extremity and at 3870 ± 298 bp from the other one. C) and D) : Electron micrographs of hybrid molecules between purified conalbumin mRNA and heat-denatured pBR322-con1 DNA linearized with Hind III endonuclease. A double-stranded, thicker RNA-DNA segment of 2200 ± 264 bp lies at 740 ± 144 nucleotides from one DNA end and 3820 ± 514 nucleotides from the other one. 5' and 3' arrow-heads refer to the corresponding ends of con-mRNA. The bar represents 0.1 µm.
In the line drawings, the solid line represents the linearized single-stranded pBR322-con1 DNA, the dotted line represents the linearized pBR322 single-stranded DNA, and the dashed line represents the mRNA.

TABLE 1

Orientation of the inserted conalbumin ds-cDNA in pBR322-con1.

Experimental conditions*		S_1 - resistant $[^3H]$-cDNA* (in percent)
Time of incubation (min)	Heat denaturation	
0	-	2.5
	+	100
30	-	41
	+	98.5
60	-	80
	+	101
120	-	98
	+	99

*See Materials and Methods. 1150 cpm of $[^3H]$-TMP labelled cDNA were used in each assay. 100% correspond to 550 cpm (see Fig. 1).

DISCUSSION

Starting from partially purified con-mRNA we have prepared a dscDNA conalbumin clone. The approach which we used, and which is generally applicable, was to blunt-end ligate the dscDNA with the plasmid pBR322 in a manner which retained TaqI restriction sites at the extremities of the integrated DNA. By hybridization studies we showed that the plasmid which we called pBR322-con1 corresponded to the major component present in the partially purified con-mRNA. An unequivocal demonstration that the dscDNA cloned in this recombinant plasmid is specific for con-mRNA was obtained by the *in vitro* translation of the RNA purified by hybridization of laying hen oviduct RNA with the plasmid pBR322-con1. The product obtained was the same size as conalbumin and was immunoprecipitated by antibodies to conalbumin. Furthermore, preliminary DNA sequence analysis (R. Hen, personal communica-

tion) of the cloned dscDNA also shows that the 5' terminal region of the dscDNA contains the sequence coding for the NH_2-terminal amino-acid sequence of conalbumin which has been recently established by Thibodeau et al. (30).

The electron microscopic results suggest that the cloned dscDNA present in pBR322-con1 is an almost full-length transcript of con-mRNA, since no RNA tails were visible at either the 5' or the 3' ends of the DNA:RNA hybrid region (Fig. 4). There are indications (R. Hen, personal communication) that indeed only the first 20 nucleotides coding for con-mRNA are lacking at the 5' end of the dscDNA. We do not know at present whether any mRNA-coding sequence is missing at the 3' end of the inserted dscDNA. Nevertheless the length of such a putative missing sequence cannot be longer than 50 to 80 bp in view of the absence of a visible tail at the 3' end of the hybrid region, in spite of the presence of the 3' terminal poly(A) in this region.

Since the inserted dscDNA is about 2300-2350 bp long as determined by agarose gel electrophoresis or electron microscopy of heteroduplex molecules, the actual length of con-mRNA should be about the same size. This size is significantly lower than the previous estimation of Lee et al. (7), who reported that conalbumin mRNA is about 3200 nucleotides long, as determined from sedimentation and polyacrylamide gel electrophoresis under non-denaturing conditions. This discrepancy is most likely related to the secondary structure of con-mRNA, since determination of the size of con-mRNA by agarose gel electrophoresis under fully denaturating conditions in the presence of methyl mercuric hydroxyde (36), yields a value of about 2400 nucleotides (M. LeMeur and A. Krust, unpublished results) in keeping with the value found for the inserted dscDNA.

It appears therefore that the 5' and 3' terminal non-coding regions of con-mRNA must be short, since conalbumin, which has a molecular weight of approximately 77,000 daltons (37) requires a coding region of about 2300 nucleotides. Sequencing studies are in progress to determine the exact length of those untranslated regions. In any case, it is already very likely from our present results that the con-dscDNA present in pBR322-con1 contains most, if not all, of the sequence coding

for the conalbumin amino acids.

It is clear that the pBR322-con1 plasmid will be very useful. It is obviously an invaluable tool for studying the organization of the conalbumin gene and the regulation of its expression. It could also be used to determine the complete amino acid sequence of conalbumin and to study the synthesis of a large eukaryotic polypeptide in heterogeneous prokaryotic or lower eukaryotic environments.

ACKNOWLEGEMENTS

We thank Drs. J.P. LePennec and P. Gerlinger for gifts of EcoRI, BamHI and HindIII restriction enzymes, and Dr. M. Bellard for a gift of nuclease S1. We thank the Viral Cancer Program, National Cancer Institute (Dr. Beard) for gifts of avian myeloblastosis virus reverse transcriptase and Dr. A. Kornberg and D. Brutlag for gifts of purified E.coli DNA polymerase I. The excellent technical assistance of Mrs. C. Wasylyk, E. Sittler and of Mr. J.M. Garnier, B. Boulay and E. Taubert is greatly acknowledged. Electron microscopy facilities were made available to F. Perrin in Strasbourg by the CNRS and the INSERM. This work was supported in Strasbourg by grants to P. Chambon from the INSERM (CRT 76.5.462 and 76.5.468), the CNRS (ATP 2117) and the Fondation pour la Recherche Médicale Française, and in Seattle by grants to R. Palmiter from the National Institutes of Health (GM-15731 and HD 09172). Dr. C. Lee is supported by the Research Service Award GM-07270 from the National Institutes of Health. G.S. McKnight and R.D. Palmiter are Associate Investigator and Investigator of the Howard Hughes Medical Institute, respectively.

REFERENCES

1. Palmiter, R.D. (1972) J. Biol. Chem. 247, 6450-6461.
2. Schimke, R.T., McKnight, G.S. and Shapiro, D.J. (1975) in "Biochemical Actions of Hormones", Vol. III (Ed. G.Litwack), Academic Press, New York, pp. 246-270.
3. Rosen, J.M. and O'Malley, B.W. (1975) in "Biochemical Actions of Hormones", Vol. III (Ed. G. Litwack), Academic

Press, New York, pp. 271-318.
4. Palmiter, R.D., Moore, P.B. and Mulvihill, E.R. (1976) Cell 8, 557-572.
5. Nguyen-Huu, M.C., Sippel, A.E., Hynes, N.E., Groner, B. and Schütz, G. (1978) Proc. Natl. Acad. Sci. USA 75, 686-690.
6. Bellard, M., Gannon, F. and Chambon, P. (1977) Cold Spring Harbor Symp. Quant. Biol. V. 42, 779-790.
7. Lee, D.C., McKnight, G.S. and Palmiter, R.D. (1978) J. Biol. Chem. 253, 3494-3503.
8. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) Proc. Natl. Acad. Sci. USA 75, 4853-4857.
9. Dugaiczyk, A., Woo, S.L.C., Lai, E.C., Mace, Jr., M.L., McReynolds, L. and O'Malley, B.W. (1978) Nature 274, 328-333.
10. Kourilsky, P. and Chambon, P. (1978) Trends in Biochemical Sciences 3, 244-247.
11. Chambon, P., Benoist, C., Breathnach, R., Cochet, M., Gannon, F., Gerlinger, P., Krust, A., LeMeur, M., LePennec, J.P., Mandel, O'Hare, K. and Perrin, F. (1979) in "Proceedings of the 11th Miami Winter Symp. "From Gene to Protein : Transfer in normal and abnormal cells". Academic Press, New York, in press.
12. Gannon, F., O'Hare, K., Perrin, F., LePennec, J.P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B. and Chambon, P. (1979) Nature, in press.
13. McKnight, S.G. (1978) Cell 14, 403-413.
14. LePennec, J.P. (1976) Diploma work, University of Strasbourg.
15. Williams, J. (1962) Biochem. J. 83, 355-364.
16. Vogt, V.M. (1973) Eur. J. Biochem. 33, 192-200.
17. Sümegi, J., Breedveld, D., Hossenlopp, P. and Chambon, P. (1977) Biochem. Biophys. Res. Commun. 76, 78-85.
18. Wilson, G.A. and Young, F.E. (1975) J. Mol. Biol. 97, 123-125.
19. Humphries, P., Gordon, R.L., McConnel, D.J. and Connolly, P. (1974) Virology 58, 25-31.
20. Diggelman, H., Faust, C.H. and Mach, B. (1973) Proc. Natl. Acad. Sci. USA 70, 693-696.
21. Hossenlopp, P., Sümegi, J. and Chambon, P. (1978) Eur. J. Biochem. 90, 615-631.
22. Seeburg, P.H., Shine, J., Martial, J.A., Baxter, J.D. and Goodman, H.M. (1977) Nature 270, 486-494.
23. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M. (1977) Science 196, 1313-1319.
24. Sugino, A., Goodman, H.M., Heyneker, H.L., Shine, J., Boyer, H.W. and Cozzarelli, N.R. (1977) J. Biol. Chem. 252, 3987-3994.
25. Humphries, P., Cochet, M., Krust, A., Gerlinger, P., Kourilsky, P. and Chambon, P. (1977) Nucl. Acids Res. 4, 2389-2407.
26. Telford, J., Bosele, P., Schaffner, W. and Birnstiel, M. (1977) Science 195, 391-392.
27. Clewell, D.B. (1972) J. Bacteriol. 110, 667-676.
28. Clewell, D.B. and Helinski, D.R. (1969) Proc. Natl. Acad. Sci. USA 62, 1159-1166.

29. Cohn, R.H., Lowry, J.C. and Kedes, L.H. (1976) Cell 9, 147-161.
30. Thibodeau, S.N., Lee, D.C. and Palmiter, R.D. (1978) J. Biol. Chem. 253, 3771-3774.
31. Studier, F.W. (1973) J. Mol. Biol. 79, 237-248.
32. Bonner, W.M. and Laskey, R.A. (1974) Eur. J. Biochem. 46, 83-88.
33. Garapin, A.C., LePennec, J.P., Roskam, W., Perrin, F., Cami, B., Krust, A., Breathnach, R., Chambon P. and Kourilsky, P. (1978) Nature 273 349-354.
34. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. and Boyer, H.W. (1977) Gene 2, 95-113.
35. Sutcliffe, J.G. (1978) Nucl. Acids Res. 5, 2721-2728.
36. Bailey, J.M. and Davidson, N. (1976) Analytical Biochem. 70, 75-85.
37. Greene, F.C. and Feeney, R.E. (1968) Biochemistry 7, 1366-1371.